



HAL
open science

Etude de la stabilité et de la précision des modèles utilisés dans la correction des effets de proximité optique en photolithographie

Mazen Saied

► **To cite this version:**

Mazen Saied. Etude de la stabilité et de la précision des modèles utilisés dans la correction des effets de proximité optique en photolithographie. Autre. Université de Grenoble, 2011. Français. NNT : 2011GRENT079 . tel-00682907

HAL Id: tel-00682907

<https://theses.hal.science/tel-00682907>

Submitted on 27 Mar 2012

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 L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Micro et Nano Électronique**

Arrêté ministériel : 7 août 2006

Présentée par

Mazen / SAIED

Thèse dirigée par **Isabelle / SCHANEN** et
codirigée par **Emek / YESILADA**

préparée au sein de **l'Institut de Micoélectronique, Électromagnétisme
et Photonique (IMEP)**
dans **l'École Doctorale d'Électronique, Électrotechnique, Automatique
et Traitement du Signal (EEATS)**

Étude de la stabilité et de la précision des modèles utilisés dans la correction des effets de proximité optique en photolithographie

Thèse soutenue publiquement le **30 septembre 2011** ,
devant le jury composé de :

M. François FLORY

PR, École Centrale de Marseille, Président

M. Olivier PARRIAUX

PR, Université Jean Monnet St Etienne, Rapporteur

M. Régis OROBTCHOUK

MCF, INSA Lyon, Rapporteur

M. Patrick SCHIAVONE

PhD, Aselta Nanographics, Examineur

Mme. Isabelle SCHANEN

PR, Grenoble INP, Directrice de thèse

M. Emek YESILADA

PhD, ST Microelectronics, Encadrant

M. Olivier TOUBLAN

PhD, Mentor Graphics, Invité

M. Laurent DEPPE

PhD, ASML Brion Technologies, Invité



À mes chers parents ...

Remerciements

Je tiens tout d'abord à remercier les membres du jury qui ont évalué ce travail de thèse : Monsieur François Flory pour m'avoir fait l'honneur de présider ce jury, Messieurs Olivier Parriaux et Régis Orobtcouk pour avoir accepté d'évaluer ce mémoire en qualité de rapporteurs, Monsieur Patrick Schiavone pour avoir accepté de faire partie de mon jury en qualité d'examineur, et enfin Messieurs Olivier Toublan et Laurent Depré pour avoir accepté mon invitation pour participer à ce jury.

Je tiens à remercier sincèrement Madame Isabelle Schanen, ma directrice de thèse, qui s'est toujours montrée à l'écoute et très disponible tout au long de la réalisation de ce travail et sans qui ce mémoire n'aurait jamais vu le jour. Je la remercie également pour la grande patience dont elle a su faire preuve ainsi que pour l'aide et le temps précieux qu'elle a bien voulu me consacrer, malgré ses nombreuses charges professionnelles.

Cette thèse a été effectuée dans le cadre d'une convention CIFRE conclue entre le laboratoire IMEP/LAHC de Grenoble et Freescale Semiconductors à Crolles durant les deux premières années, puis reconduite avec la société STMicroelectronics en troisième année. Je tiens donc à remercier sincèrement Monsieur Yorick Trouiller de m'avoir accueilli dans son équipe sur le site de STMicroelectronics. Mes sincères remerciements vont plus particulièrement à mon encadrant industriel, Monsieur Franck Foussadier, pour la confiance et la grande liberté qu'il m'avait accordées et pour les nombreuses remarques et suggestions qu'il m'avait fait part afin d'améliorer la qualité de ce mémoire.

Je ne sais comment exprimer ma gratitude à Monsieur Vincent Farys, ingénieur à STMicroelectronics, pour m'avoir encadré de façon bénévole, pour sa gentillesse, son soutien, ses conseils très avisés et sa confiance envers moi. Je le remercie également pour ses nombreuses contributions à améliorer ce travail de thèse et pour m'avoir aidé dans la préparation de ma soutenance. Je remercie également Monsieur Gérard Ghibaudo, directeur du laboratoire IMEP/LAHC, pour sa générosité, pour son soutien aussi bien matériel que moral, pour tous les conseils que j'ai eu l'honneur de recevoir de sa part mais aussi pour m'avoir aidé dans toutes les démarches administratives précédant la soutenance. Je lui en suis très reconnaissant.

Je remercie tous ceux sans qui cette thèse ne serait pas ce qu'elle est : aussi bien par les discussions que j'ai eu la chance d'avoir avec eux, leurs suggestions ou contributions. Je pense ici en particulier à Monsieur Alexandre Lagrange pour les longues discussions sur la modélisation des effets topographiques du masque, Monsieur Jérôme Belledent qui m'a fait part de son expertise en photolithographie, Monsieur Frédéric Robert pour ses suggestions en matière de programmation et pour ses conseils en simulation et modélisation, Monsieur Gurwan Kerrien pour son aide précieuse dans les manips, Monsieur Alexandre Villaret pour nos longues discussions sur les modèles empiriques et Monsieur Emek Yesilada qui sans lui je n'aurais jamais pensé à déposer un brevet pour ma méthode de modélisation. Je remercie aussi tous les membres de l'équipe RET à STMicroelectronics avec qui il a été très agréable de travailler.

Je souhaite remercier tous mes amis et toutes les personnes que j'ai eu le plaisir de côtoyer au CEA/Léti durant ces quelques années. Merci à Mickael Martin pour sa générosité et son soutien, à Jonathan Pradelles qui a eu la gentillesse de lire et de corriger ce travail et pour les nombreuses sorties d'amis, à Sébastien Soulan pour nos nombreuses discussions et pour sa gentillesse et ses encouragements, à David Rio pour nos parties de tennis et pour nos multiples conversations, à Jessy Bustos pour les moments de délire et pour les pauses café, à Béatrice Icard qui a su me reconforter dans des moments difficiles, à Claire Sourd pour sa bonne humeur, à Johann Foucher

pour sa gentillesse et son soutien et à Laurent Pain qui a su des fois me motiver et me redonner confiance en moi. Merci également à tous les autres membres du laboratoire de lithographie au CEA/Léti pour leur soutien.

Je passe ensuite une dédicace spéciale à tous mes amis pour leurs soutiens et leurs encouragements : Kévin, Meshkan, Sébastien, Erwine, Abdi, Jean-Fred, Manu, Rached, Mouna, Nader, Walid, Oussama et j'en oublie certainement d'autres.

Enfin, je remercie mes parents, mes frères et mes soeurs pour leurs encouragements et leurs soutiens qui m'ont permis de faire cette thèse dans de bonnes conditions.

Table des matières

Remerciements	v
Table des matières	i
Glossaire	viii
Introduction	ix
La photolithographie et ses enjeux	1
Chapitre 1. Les enjeux de la photolithographie	3
1.1 Introduction	5
1.2 Éléments d'un système de photolithographie	6
1.2.1 L'illumination	6
1.2.2 Le masque	6
1.2.3 L'optique de projection	7
1.2.4 L'empilement multicouche de la résine	8
1.2.4.1 La résine photosensible	8
1.2.4.2 Les couches antireflet	8
1.3 Techniques d'amélioration de la résolution	9
1.3.1 La résolution	9
1.3.2 Choix de l'illumination	10
1.3.2.1 La longueur d'onde	10
1.3.2.2 La géométrie de la source	10
1.3.2.3 La polarisation de la source	12
1.3.3 Choix du masque	14
1.3.4 La correction des effets de proximité optique	17
1.3.5 La lithographie par immersion	19
1.3.6 La lithographie par double exposition	20
1.4 Modélisation du procédé photolithographique	22
1.4.1 Modélisation de la partie optique	22
1.4.1.1 Optique de Fourier et formule d'Abbé	22
1.4.1.2 Formule d'Hopkins pour une lumière partiellement cohérente	23
1.4.2 Modélisation de la partie résine	24
1.4.2.1 Modèles simples	26
1.4.2.1.1 Modèle à seuil constant	26
1.4.2.1.2 Modèle à diffusion gaussienne	26
1.4.2.2 Modèles avancés	27
1.4.2.2.1 Modèle à seuil variable	27
1.4.2.2.2 Modèle à convolution étendue	28
1.4.2.2.3 Modèle à opérateur différentiel	28
1.4.3 Discussions	29
1.5 Conclusion	30

Étude de la stabilité des modèles empiriques	33
---	-----------

Chapitre 2. Causes et effets de l'instabilité des modèles empiriques	35
---	-----------

2.1	Introduction	37
2.2	Position du problème	37
2.2.1	Expérimentation	38
2.2.2	Résultats et discussions	39
2.3	État de l'art sur la stabilité des modèles empiriques	40
2.3.1	Méthode de la réponse en surface	41
2.3.2	Méthode des différences entre modèle et imagerie	41
2.3.3	Comparatif des différentes méthodes	42
2.4	Origines de l'instabilité des modèles empiriques	43
2.4.1	Étude qualitative	43
2.4.2	Comparatif des modèles	44
2.4.3	Origines de l'instabilité des modèles	44
2.4.3.1	Définition de la stabilité	44
2.4.3.2	Cas des paramètres du procédé	45
2.4.3.2.1	Nombre de motifs de calibration	45
2.4.3.2.2	Type des motifs de calibration	46
2.4.3.2.3	Impact du bruit de mesures	48
2.4.3.3	Cas des paramètres de modélisation	50
2.4.3.3.1	Dépendance élevée aux paramètres empiriques	50
2.4.3.3.2	Différence dans la nature physique des termes	51
2.4.3.3.3	Existence de termes à effets antagonistes	53
2.5	Conclusion	54

Chapitre 3. Étude de la stabilité de modèles empiriques par la méthode perturbative	55
--	-----------

3.1	Introduction	57
3.2	Principe de la méthode perturbative	57
3.3	Stabilité d'un modèle empirique du nœud 65 nm	58
3.3.1	Stabilité par rapport à la position du plan focal	58
3.3.2	Stabilité par rapport à l'ouverture numérique et aux facteurs de cohérence partielle	60
3.3.3	Stabilité par rapport à l'épaisseur de la résine	61
3.3.4	Application à un circuit quelconque	62
3.3.5	Vérification de la couverture du modèle	64
3.3.5.1	Définition de l'espace de paramètres du modèle	64
3.3.5.2	Choix de la distance d'un site par rapport à l'espace des paramètres	65
3.3.5.3	Détermination des motifs sensibles	65
3.4	Stabilité de modèles empiriques du nœud 45 nm	66
3.4.1	Stabilité par rapport à la position du plan focal	67
3.4.2	Stabilité par rapport à l'ouverture numérique	68
3.4.3	Stabilité par rapport aux facteurs de cohérence partielle	69
3.4.4	Comparatif des différents modèles	70
3.4.5	Vérification expérimentale	71
3.5	Conclusion	72

Modélisation des effets topographiques du masque	75
---	-----------

Chapitre 4. Méthodes de modélisation des effets topographiques du masque	77
---	-----------

4.1	Introduction	79
-----	------------------------	----

4.2	Position du problème	79
4.3	Méthodes rigoureuses	80
4.3.1	La méthode FDTD	80
4.3.2	La méthode RCWA	81
4.4	L'approximation de Kirchhoff	82
4.4.1	Principe	82
4.4.2	Limitations	83
4.4.3	Analyse du champ proche	84
4.4.3.1	Masque étudié	84
4.4.3.2	Cas d'un bord diffractant	84
4.4.3.3	Cas d'une ligne isolée	88
4.4.3.4	Cas d'un réseau de lignes denses	89
4.4.4	Impact sur l'intensité lumineuse	90
4.4.4.1	Contraste	92
4.4.4.2	Position du meilleur focus	93
4.4.4.3	Profondeur de champ	93
4.4.4.4	Effets de proximité	94
4.4.4.5	Effets sur le seuil d'impression des SRAF	95
4.4.5	Discussion des résultats	98
4.5	Méthodes semi-rigoureuses	98
4.5.1	Méthode de décomposition de domaine DDM	98
4.5.2	Méthode de compensation des effets de diffraction M3D	99
4.5.3	Évaluation par rapport au calcul rigoureux	100
4.5.3.1	Définition des conditions de simulation	100
4.5.3.2	Fenêtres de procédé et courbes de Bossung	101
4.5.3.3	Position du meilleur focus	102
4.5.3.4	Profondeur de champ	103
4.5.3.5	Effets de proximité	104
4.5.3.6	Prédiction de l'impression des SRAF	104
4.5.4	Discussion des résultats	105
4.6	Conclusion	107
Chapitre 5. Correction des effets topographiques du masque		109
5.1	Introduction	111
5.2	Méthodes heuristiques	111
5.2.1	Correction par adaptation de la dose	112
5.2.2	Correction par un modèle empirique	113
5.2.2.1	Introduction d'un facteur de dose	114
5.2.2.2	Régression quadratique	115
5.2.2.2.1	Récapitulatif des résultats	117
5.2.2.3	Optimisation des modèles	117
5.2.2.4	Précision et stabilité	118
5.2.3	Amélioration de la prédiction de l'impression des SRAF	119
5.2.3.1	Correction par réduction de la taille des SRAF	119
5.2.3.2	Détermination des spécifications des dimensions des SRAF	120
5.2.3.2.1	Cas d'une diffusion nulle dans la résine	120
5.2.3.2.2	Cas d'une diffusion non nulle dans la résine	121
5.2.3.3	Correction par modification de la transmission et de la phase des SRAF	122
5.3	Méthode de modélisation par décomposition multi-niveaux du masque fin	123
5.3.1	Principe	124
5.3.1.1	Expression du champ scalaire effectif	124
5.3.1.2	Décomposition multi-niveaux du champ électrique	125

5.3.2	Évaluation par rapport au calcul rigoureux	128
5.3.2.1	Définition des conditions de simulation	128
5.3.2.2	Fenêtres de procédé et courbes de Bossung	131
5.3.2.3	Position du meilleur focus	132
5.3.2.4	Profondeur de champ	132
5.3.2.5	Effets de proximité	132
5.3.2.6	Prédiction de l'impression des SRAF	134
5.3.2.7	Ressources nécessaires et temps d'exécution	134
5.3.2.8	Intégration en traitement OPC	135
5.4	Conclusion	136
Conclusion générale		139
Perspectives		141
Annexes		142
Annexe A. Éléments de la photolithographie		143
A.1	Éléments d'un système de photolithographie	143
A.1.1	L'illumination	144
A.1.2	Le masque	144
A.1.3	L'optique de projection	145
A.1.4	L'empilement multicouche de la résine	146
A.1.4.1	La résine photosensible	146
A.1.4.2	Les couches antireflet	148
A.2	Formation de l'image	149
A.2.1	Cas d'une source cohérente	149
A.2.2	Cas d'une source partiellement cohérente	151
A.3	Critères de performances en photolithographie	154
A.3.1	La résolution	154
A.3.2	La profondeur de champ	155
A.3.3	La latitude en énergie	156
A.3.4	La fenêtre de procédé	156
A.3.5	Le facteur d'erreur du masque	157
Annexe B. État de l'art sur la stabilité des modèles empiriques		159
B.1	Méthode d'un seul facteur à la fois	159
B.2	Méthode de la sensibilité aux paramètres	159
B.3	Concept des modèles MSM	160
B.4	La méthode des motifs sensibles à l'imagerie	162
Annexe C. Présentation de modèles empiriques avancés		164
C.1	Modèle à opérateur différentiel	164
C.2	Modèle à convolution étendue	165
Annexe D. Théorie scalaire de la diffraction		167
D.1	Historique	167
D.2	Principe de Huygens-Fresnel	167
D.3	Zones de Fresnel	168
D.4	Théorème intégral de Helmholtz-Kirchhoff	170

D.5	Formule de Fresnel-Kirchhoff	172
Annexe E. La méthode FDTD		174
E.1	Historique	174
E.2	Principe de la méthode FDTD	175
E.2.1	Approximation des différences finies	175
E.2.2	Application à l'onde électromagnétique	176
E.3	Dispersion numérique de la méthode FDTD	178
E.4	Critère de stabilité de la méthode FDTD	179
Liste des publications		180
1	Brevets	180
2	Publications premier auteur	180
3	Publications co-auteur	180
Bibliographie		182
Résumé		192

Glossaire

- ARC** “*Anti Reflective Coating*”, couche antireflet qui sert à atténuer les ondes stationnaires dans la résine. 8, 149
- BARC** “*Bottom Anti Reflective Coating*”, couche antireflet enterrée. 91, 93, 149
- BL** “*Boundary Layer*”, méthode de modélisation des effets topographiques du masque qui consiste à rajouter aux motifs du circuit des bords d’une certaine largeur et des valeurs préajustées de la transmission et de la phase. 100, 111, 124
- CD** “*Critical Dimension*”, désigne la taille minimale d’un motif pouvant être résolu par le système optique. Ce terme est communément utilisé pour désigner également la largeur d’une ligne ou d’un espace que ce soit dans le circuit avant sa fabrication ou sur le masque ou sur la plaquette. 9, 14, 17, 19, 29, 41, 44, 45, 49, 51–53, 57, 59, 60, 62, 66, 67, 71, 72, 80, 88–95, 97, 98, 101, 102, 104, 105, 114, 117, 119–121, 129–134, 136, 137, 148, 155–157, 159, 160
- CM1** “*Compact Model 1*”, modèle à opérateur différentiel. 26, 28–30, 42, 44, 66, 96, 164, 165
- CPL** “*Chromeless Phase-shift Lithography*”), type de masque à décalage de phase. 16
- CTR** “*Constant Threshold Resist*”, modèle à seuil constant. 26, 29, 38, 39, 43–45, 51, 52, 59, 61–63, 67
- DDM** “*Domain Decomposition Method*”, méthode de décomposition de domaine. 98–101, 105–107, 123, 124, 128–134, 137, 140
- DOF** “*Depth Of Focus*”, profondeur de champ. 80, 93, 94, 98, 100, 103, 132, 155–157
- DPL** “*Double Patterning Lithography*”, lithographie par double exposition. 20
- DRAM** “*Dynamic Random Access Memory*”, mémoire vive dynamique. 22
- DTD** “*Dual Tone Development*”, technique de double exposition par deux étapes de lithographie où l’on utilise deux solutions de natures différentes pour le développement. 21
- EL** “*Exposure Latitude*”, latitude en énergie. 93, 94, 100, 103, 132, 156, 157
- EPE** “*Edge Placement Error*”, erreur de placement de bord. 57, 59, 61–63, 66, 67, 70, 91, 115–118
- EUV** Extrême Ultraviolet. 10, 139
- FDTD** “*Finite Difference Time Domain*”, méthode des différences finies dans le domaine temporel, sert la modélisation du champ électromagnétique dans une topographie quelconque. xi, 79–81, 83–98, 100–106, 111–116, 119, 120, 122, 127, 128, 130–134, 137, 140, 141, 174–176, 178, 179
- ITRS** “*International Technology Roadmap for Semiconductors*”, feuille de route de l’industrie du semi-conducteur. 10, 14, 22, 30, 54, 79, 111, 139, 148
- λ Longueur d’onde de la source d’illumination. 6, 9, 10, 29, 86, 87, 90, 144, 149, 154
- LELE** “*Litho-Etch, Litho-Etch*”, technique de double exposition par deux étapes successives de lithographie et de gravure. 20, 21
- LER** “*Line Edge Roughness*”, rugosité de bord de ligne. 21, 22, 148, 149
- LFLE** “*Litho-Freeze, Litho-Etch*”, technique de double exposition par une première étape de lithographie suivie d’un gel de la résine puis une seconde étape de lithographie et une gravure. 20, 21
- LWR** “*Line Width Roughness*”, rugosité de la largeur de ligne. 30
- M3D** “*Mask 3D*”, méthode de modélisation des effets topographiques du masque. 100, 101, 105–107, 123, 124, 128, 130–134, 137, 140
- MEEF** “*Mask Error Enhancement Factor*”, facteur d’erreur du masque. 17, 80, 98, 157–159
- MMFE** “*Modal Method by Fourier Expansion*”, méthode modale par l’expansion de Fourier, également appelée RCWA, sert à la modélisation du champ électromagnétique dans une topographie quelconque. 81
- MSM** “*Manufacturing Sensitivity Models*”, modèles de la sensibilité du procédé de fabrication. 40, 42, 160, 161

- NGL** “*Next Generation Lithography*”, lithographie de nouvelle génération. 139
- OFAT** “*One-Factor-At-a-Time*”, méthode d'un facteur à la fois. 40, 42, 57, 159
- ON** Ouverture Numérique. 9, 11, 14, 22, 23, 38, 58, 60, 68–72, 90, 97, 100, 120, 131, 151–155
- OPC** “*Optical Proximity Correction*”, correction des effets de proximité optique. x, xi, 3, 5, 17–19, 22, 24–30, 35, 37, 38, 40–43, 54, 62, 63, 68, 72, 77, 79, 83, 96, 98–100, 106, 107, 111, 115, 123, 135, 136, 139–141, 159, 161, 164
- OSP** “*Optically Sensitive Patterns*”, motifs optiquement sensibles. 42, 162
- PAG** “*Photo Acid Generator*”, photogénérateur d'acide. 8, 147
- PEB** “*Post Exposure Bake*”, recuit après exposition. 8, 24, 25, 39, 41, 147, 148
- PS** “*Parameter Sensitivity*”, méthode de la sensibilité des paramètres. 42, 57, 159
- PSM** “*Phase Shift Mask*”, masque à décalage de phase. 14–16, 30, 38, 42, 84, 97, 100, 120, 131
- RCWA** “*Rigorous Coupled Wave Analysis*”, analyse rigoureuse des ondes couplées, également appelée MMFE, sert la modélisation du champ électromagnétique dans une topographie quelconque. 81, 82
- RET** “*Resolution Enhancement Techniques*”, techniques d'amélioration de la résolution. x, 5, 17, 30, 37, 139
- RMS** “*Root Mean Square*”, moyenne quadratique. 49, 50, 54, 66, 67
- RSM** “*Response Surface Method*”, méthode de la réponse de surface. 40–42, 50
- SADP** “*Self-Aligned Double Patterning*”, technique de double exposition, également appelée SIT, qui consiste à une première étape de lithographie suivie d'un dépôt d'une fine couche de matériau gardant la forme initiale de la résine, qui, par la suite, sera gravée pour faire apparaître des espaceurs de part et d'autre des bords de la résine qui serviront comme couche protectrice lors de la gravure. 21
- SCAA** “*Sidewall Chrome Alternating Aperture*”, type de masque à décalage de phase. 15
- SEMCD** “*Scanning Electron Microscopy for Critical Dimensions*”, microscope électronique à balayage dédié à la mesure des dimensions critiques. 38, 96, 122
- SIT** “*Sidewall Image Transfer*”, technique de double exposition, voir l'entrée SADP du glossaire. 21
- SMO** “*Source-Mask Optimization*”, méthode d'optimisation simultanée de la source d'illumination et du dessin à graver sur le réticule du masque. 12
- SRAF** “*Sub Resolution Assist Feature*”, motif diffractant non résolu. 15, 18, 19, 30, 62, 71, 95–98, 104–107, 112, 119–123, 129–132, 134, 136, 137, 140
- SRAM** “*Static Random Access Memory*”, mémoire vive statique. 112
- TARC** “*Top Anti Reflective Coating*”, couche antireflet superficielle. 149
- TCC** “*Transfer Cross Coefficients*”, matrice de dimension 4 qui sert dans le calcul des ordres diffractés. 24, 100–102, 111, 123, 128, 129, 140
- TE** Transverse Électrique. 12–14, 84–90, 99, 106, 124, 125, 177
- TM** Transverse Magnétique. 12–14, 83–90, 99, 124, 125, 177
- TMA** “*Thin Mask Approximation*”, approximation du masque fin. 83
- VT5** “*Variable Threshold 5*”, modèle à seuil variable. 26–28, 38, 39, 43, 44, 46–53, 59, 61–63, 65, 66, 96, 113, 114, 116, 119, 165

Introduction

Depuis sa naissance en 1959, l'industrie de la micro-électronique a connu un véritable essor technologique et une croissance incessante qui ont littéralement transformé nos modes de vie. En effet, nous assistons, depuis une vingtaine d'années, à de nombreuses avancées technologiques dans beaucoup de secteurs de l'ingénierie moderne, à l'instar des techniques de l'information et de la communication. Année après année, les circuits intégrés sont devenus la technologie principale et continuent à être le moteur de l'innovation. Les progrès technologiques dans ce secteur n'ont qu'un seul objectif : intégrer toujours plus de transistors sur une même surface. La réduction incessante des dimensions des circuits a fortement contribué à augmenter considérablement la densité d'intégration et a permis de produire des circuits réalisant des tâches de plus en plus complexes plus rapidement et à faible consommation.

Au delà des performances des circuits intégrés, le facteur économique est d'une importance capitale pour le devenir de l'industrie des semi-conducteurs. L'augmentation de la densité d'intégration et la maîtrise des coûts de fabrication n'ont été envisageables qu'avec le développement de techniques de miniaturisation collectives à fort volume de production. C'est parmi ces techniques que s'inscrit la lithographie, étape déterminante du procédé de fabrication d'un circuit intégré.

La photolithographie, ou lithographie optique, est la technique la plus répandue pour une utilisation industrielle en raison de sa capacité de fabrication à grand volume. C'est l'étape clef de la définition des motifs des circuits et son évolution permet de suivre le rythme de la miniaturisation imposé par la loi de Moore¹.

Depuis son adoption par les industriels, la photolithographie a parcouru plusieurs périodes traduisant l'évolution des règles de dessin utilisées pour la fabrication des circuits intégrés. Chacune de ces périodes traduit une génération des outils de lithographie et se caractérise par l'utilisation d'une longueur d'onde particulière. La figure 1 décrit l'évolution des règles de dessin et de la résolution des outils de lithographie depuis les années 1980. Cette évolution se caractérise par la réduction, qui s'est effectuée par sauts, de la longueur d'onde d'exposition utilisée dans les outils de lithographie. Selon cette figure, l'évolution de la résolution était plus rapide que celle de la longueur d'onde d'insolation. En effet, la résolution a diminué d'un facteur 60 ($3\mu\text{m}$ dans les années 1980 contre 45 nm aujourd'hui) alors que la longueur d'onde n'a diminué que d'un facteur deux (436 nm dans les années 1980 contre 193 nm aujourd'hui). Bien que l'on ait commencé avec une résolution six à sept fois supérieure à la longueur d'onde, nous arrivons aujourd'hui à un régime "*sub-wavelength*" [1–3] où les dimensions des motifs sont bien en dessous de la longueur d'onde d'insolation. Cette contrainte rend une génération lithographique difficile à maintenir.

Au début des années 2000, les chercheurs prévoient des sauts vers les longueurs d'onde 157 nm et 126 nm pour viser les nœuds technologiques 65 et 45 nm respectivement [4] (voir tableau 1). Ces technologies ont été vite abandonnées pour des raisons techniques concernant la difficulté de mettre en place des sources d'illumination fiables ou de fabriquer des lentilles transparentes aux longueurs d'onde utilisées. L'unique solution était donc de maintenir la lithographie optique 193 nm. Cependant, avec la dégradation de la résolution, ce maintien n'a été possible qu'avec le

1. La dernière loi de Moore, émise par Gordon E. Moore en 1975, est une interpolation empirique qui stipule que le nombre de transistors par unité de surface double tous les deux ans dans les circuits intégrés. Cette loi s'est jusqu'ici révélée exacte et elle est encore un repère pour les fabricants du semi-conducteur.

développement de nouvelles solutions permettant d'améliorer le procédé lithographique. En effet, la mise au point de nouvelles résines photosensibles plus fines et peu absorbantes aux longueurs d'onde utilisées a permis de réduire la dose d'exposition limitant ainsi les effets de la diffusion chimique. L'amélioration des optiques de projection a permis de réduire les effets parasites liés à la diffraction. L'augmentation de l'ouverture numérique des systèmes optiques (par des techniques telles que l'utilisation d'un liquide d'immersion) et le développement de nouveaux types de masque ont permis d'améliorer le contraste de l'intensité lumineuse. Finalement, la correction des effets de proximité optique (de l'anglais "*Optical Proximity Correction*" ou OPC) [3] et l'utilisation des techniques de double exposition ont rendu possible de résoudre des pas de réseau plus petits que la longueur d'onde d'insolation. Aujourd'hui, grâce aux techniques d'amélioration de la résolution (de l'anglais "*Resolution Enhancement Techniques*" ou RET) [5], la lithographie optique à 193 nm, qui était réservée aux nœuds technologiques 90 et 65 nm dans un premier temps, a été maintenue comme solution industrielle pour la génération 45 nm et au delà [4, 6].

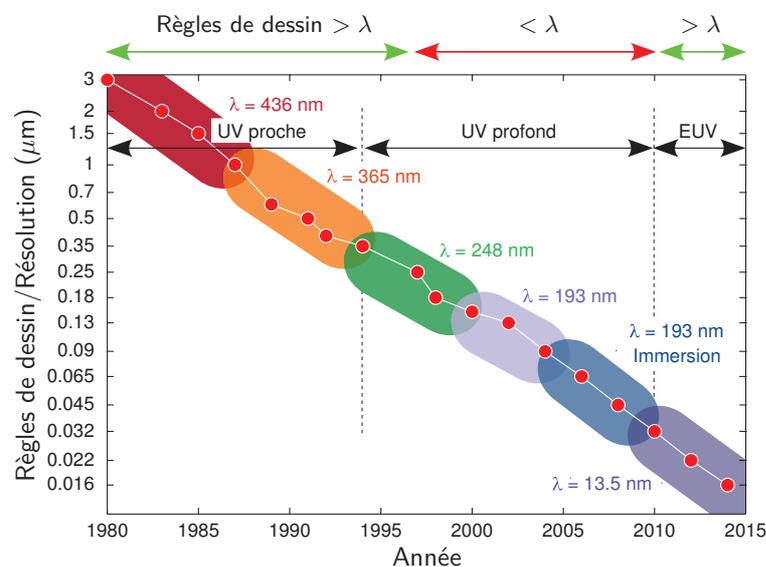


Figure 1 – Évolution des règles de dessin et de la résolution au cours du temps. Cette courbe montre également la succession des générations des outils de photolithographie. Chaque génération correspond à une longueur d'onde unique. Source : [4]

Source	λ (nm)	Résolution	Année d'introduction
Raie g	436	1 μm	
Raie i	365	0.5 μm	1984
KrF	248	0.25 μm	1989
ArF	193	100 nm	2001
F ₂	157	65 nm	(abandonnée)
Ar ₂	126	45 nm	(abandonnée)

Tableau 1 – Récapitulatif des outils de lithographie optique passés et prévisions des chercheurs au début des années 2000 pour viser les nœuds 65 nm et 45 nm. Source : [7].

Le maintien de la lithographie optique à 193 nm pour les générations de circuits intégrés actuelles implique l'amélioration des techniques existantes et en particulier la correction des effets de proximité optique (OPC). Les exigences technologiques en terme de résolution deviennent aujourd'hui si fortes que la lithographie optique est poussée dans ses limites ultimes. Pour répondre à ces contraintes, la correction OPC doit évoluer dans ce sens par la mise en place de nouvelles techniques plus agressives tout en respectant les contraintes de temps et de coûts qu'impliquent la préparation et la fabrication des masques. C'est pour cela qu'un certain nombre de modèles OPC empiriques plus avancés et de techniques de modélisation optique sont apparus

dans les logiciels de traitement OPC. Ces nouvelles techniques permettent de corriger les effets optiques et chimiques de la photolithographie, mais aussi de la gravure. Les modèles empiriques deviennent alors de plus en plus complexes pour permettre de capturer le maximum d'effets dans un temps raisonnable. À l'heure actuelle, les modèles récemment développés incluent plus de 50 paramètres à ajuster afin de reproduire aussi proche que possible les contours dans la résine. La question de la stabilité de tels modèles purement empiriques devient alors d'actualité. De plus, à grande ouverture numérique, des effets optiques complexes peuvent se produire (la polarisation, les effets topographiques du masque, les aberrations optiques. . .). Ces effets, auparavant négligés, modifient l'intensité lumineuse simulée et faussent l'extraction des contours dans la résine.

L'objectif de ce travail de thèse est d'étudier la stabilité des modèles empiriques actuels et de dégager les principales causes d'instabilité afin de mettre une stratégie préventive pour les éventuelles dérives des modèles étudiés. Parallèlement à cette étude, afin de simplifier les modèles utilisés tout en améliorant la précision du calcul et la stabilité de ces modèles, il est nécessaire d'examiner les approximations utilisées dans la modélisation de l'intensité lumineuse et d'étudier leurs limites de validité.

Avant de détailler les travaux réalisés durant cette thèse, nous allons présenter, dans le premier chapitre, les différentes étapes du procédé de photolithographie et les paramètres de l'outil photolithographique. Ensuite, nous détaillerons les principales techniques permettant d'améliorer la résolution d'un outil de lithographie optique. Puis, nous présenterons les principaux axes dans la modélisation du procédé photolithographique, étape importante dans la correction des effets de proximité optique (OPC). Enfin, nous présenterons les différents modèles empiriques que nous avons rencontrés dans les logiciels de traitement OPC.

Dans le chapitre 2, nous présenterons un état de l'art des différentes méthodes permettant d'étudier la stabilité d'un modèle empirique. Ensuite, en se basant sur la nature physique des modèles, nous présenterons une définition de la stabilité. Enfin, à la lumière de cette définition, nous décrirons qualitativement les origines d'instabilité des modèles empiriques actuels.

La plupart des méthodes proposées dans la littérature pour l'étude de la stabilité ont pour principal objectif l'évaluation de la sensibilité du procédé photolithographique plutôt que l'étude de la stabilité numérique des modèles. La difficulté de trouver une méthode compatible avec notre définition de la stabilité nous a permis de développer une nouvelle méthode perturbative que nous présenterons dans le chapitre 3 de ce mémoire. Ainsi, nous présenterons, dans un premier lieu, le principe de cette nouvelle méthode. Ensuite, nous évaluerons cette méthode pour l'étude de la stabilité de différents modèles empiriques utilisés pour les nœuds technologiques 65 et 45 nm. Finalement, nous discuterons les résultats théoriques avec ceux de l'expérience.

Dans l'objectif de simplifier les modèles empiriques et d'améliorer ainsi leur stabilité, nous allons nous intéresser aux approximations effectuées dans la modélisation de l'intensité lumineuse à travers le système optique. Parmi lesquelles, l'approximation de Kirchhoff est la méthode la plus communément utilisée dans la partie optique d'un modèle OPC afin de décrire le champ électrique en sortie du masque. Dans cette méthode, le masque est considéré infiniment fin et les effets topographiques sont négligés. Dans le chapitre 4, nous étudierons les limites de cette méthode. Puis, nous présenterons et évaluerons les différentes méthodes approchées semi-rigoureuses disponibles dans les logiciels de traitement OPC et comparerons les résultats de simulation avec ceux obtenus avec la méthode rigoureuse FDTD.

À la lumière des résultats obtenus, nous montrerons que contrairement aux algorithmes de calcul de l'approximation de Kirchhoff, les méthodes semi-rigoureuses sont moins rapides et nécessitent des ressources informatiques importantes. Dans le chapitre 5, nous proposerons différentes façons de corriger les effets topographiques du masque, tout en gardant l'approximation de Kirchhoff dans la modélisation de la partie optique. Parmi les méthodes proposées, nous exploiterons celle permettant de réduire les erreurs liées aux effets topographiques du masque par l'intermédiaire d'un second modèle empirique. Nous montrerons que pour garantir une précision adéquate, il est nécessaire d'augmenter la complexité du modèle en rajoutant des termes additionnels, pouvant ainsi dégrader sa stabilité. Enfin, pour s'affranchir du problème de la

stabilité, nous introduirons une nouvelle méthode approchée, permettant d'inclure les effets topographiques par décomposition multi-niveaux du masque fin et discuterons ses avantages et ses limites.



Première partie

La photolithographie et ses enjeux

Chapitre 1

Les enjeux de la photolithographie

La lithographie optique (ou photolithographie) est, depuis des décennies, une étape déterminante dans la fabrication des circuits intégrés. Elle est l'étape clef de la définition des motifs des circuits et son évolution permet de suivre le rythme de la miniaturisation imposé par la loi de Moore. Ses principales évolutions passent, tout d'abord, par la diminution, qui s'est effectuée par sauts, de la longueur d'onde d'exposition. Les lampes à mercure émettant des radiations dans l'ultraviolet proche et moyen (raies g,h,i du spectre d'émission du mercure) ont laissé la place aux lasers à excimère. À l'heure actuelle, l'ultraviolet profond (248 et 193 nm) semble encore être le monopole des outils de lithographie optique. Cette génération qui, préalablement visait les nœuds technologiques 90 et 65 nm, a vu son domaine s'étendre aux nœuds technologiques 45 nm, 32 nm et probablement le 22 nm et ce grâce à l'amélioration du procédé lithographique et au développement de nouvelles techniques dites "d'amélioration de la résolution".

Parmi ces techniques, la correction des effets de proximité optique ou OPC (de l'anglais "*Optical Proximity Correction*") est devenue aujourd'hui indispensable pour résoudre des dimensions plus petites que la longueur d'onde d'exposition. C'est à travers une modélisation fine du procédé lithographique qu'elle permet d'anticiper les éventuelles dérives et de les corriger. Les modèles employés en OPC doivent ainsi suivre le rythme de la miniaturisation et relever les défis imposés par la loi de Moore en matière de résolution, de la feuille de route de l'ITRS en terme de précision, mais aussi les contraintes industrielles en terme de coût et de temps.

Sommaire

1.1	Introduction	5
1.2	Éléments d'un système de photolithographie	6
1.2.1	L'illumination	6
1.2.2	Le masque	6
1.2.3	L'optique de projection	7
1.2.4	L'empilement multicouche de la résine	8
1.2.4.1	La résine photosensible	8
1.2.4.2	Les couches antireflet	8
1.3	Techniques d'amélioration de la résolution	9
1.3.1	La résolution	9
1.3.2	Choix de l'illumination	10
1.3.2.1	La longueur d'onde	10
1.3.2.2	La géométrie de la source	10
1.3.2.3	La polarisation de la source	12
1.3.3	Choix du masque	14
1.3.4	La correction des effets de proximité optique	17
1.3.5	La lithographie par immersion	19
1.3.6	La lithographie par double exposition	20
1.4	Modélisation du procédé photolithographique	22
1.4.1	Modélisation de la partie optique	22
1.4.1.1	Optique de Fourier et formule d'Abbé	22
1.4.1.2	Formule d'Hopkins pour une lumière partiellement cohérente	23
1.4.2	Modélisation de la partie résine	24
1.4.2.1	Modèles simples	26
1.4.2.2	Modèles avancés	27
1.4.3	Discussions	29
1.5	Conclusion	30

1.1 Introduction

La lithographie, du grec “*lithos*” (pierre) et “*graphein*” (écrire), est une technique ancienne permettant la reproduction à de multiples exemplaires d’un tracé à l’encre ou au crayon sur une pierre. Dans le domaine de la micro-électronique, le procédé lithographique est l’étape permettant de reproduire le tracé d’un circuit dans une couche de résine étalée sur une plaquette de silicium.

La photolithographie, ou lithographie optique, est aujourd’hui la technique la plus répandue pour une utilisation industrielle en raison de sa capacité de fabrication à grand volume. L’opération consiste en l’exposition via une optique réductrice d’une couche de résine photosensible au travers d’un masque représentant le dessin d’un circuit. L’objectif est de transférer son image dans l’empilement de la résine qui servira de couche sélective pour une étape ultérieure (gravure, dépôt, dopage, implantation ionique. . .). Dans les régions exposées, la lumière réagit avec les chaînes polymères de la résine pour modifier leurs structures moléculaires. Ceci permet de modifier les propriétés de solubilité de la résine dans les endroits exposés. Les motifs sont ensuite révélés lors de l’étape du développement qui consiste à plonger la résine dans une solution aqueuse, le développeur. Selon la tonalité de la résine employée, les zones exposées ou non de celle-ci seront retirées après développement comme le montre le schéma de la figure 1.1.

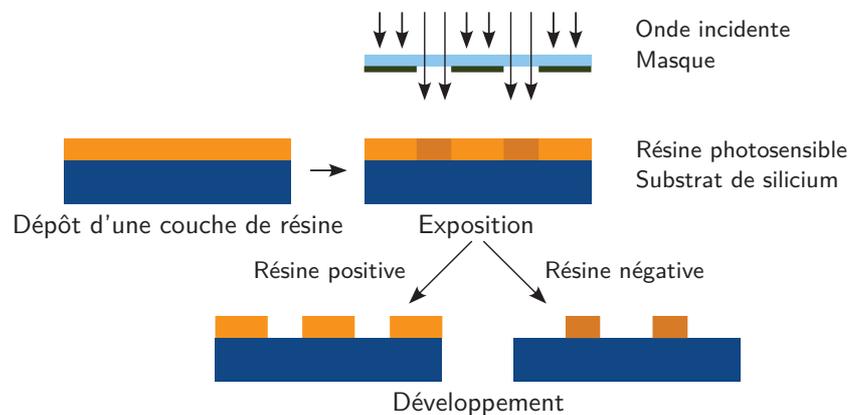


Figure 1.1 – Représentation des différentes étapes du procédé photolithographique. Pour une résine à tonalité positive, la partie exposée est retirée lors du développement, alors que pour une résine à tonalité négative, les zones exposées demeurent intactes.

La lithographie est une étape importante dans une chaîne de production en micro-électronique. C’est l’étape clef de la miniaturisation des circuits intégrés et son évolution permet de suivre la loi de Moore. Cependant, au risque de développer de nouveaux outils lithographiques très coûteux en temps et en argent, le maintien d’une solution industrielle existante pour viser plusieurs générations technologiques successives n’a été envisageable qu’avec la mise au point de nouvelles résines moins absorbantes et plus performantes et le développement des techniques dites d’amélioration de la résolution (de l’anglais “*Resolution Enhancement Techniques*” ou RET) [5]. L’une d’entre-elles, la correction des effets de proximité optique (de l’anglais “*Optical Proximity Correction*” ou OPC) est basée sur la modification du dessin du masque afin de compenser les effets indésirables du procédé photolithographique.

Dans ce chapitre, nous allons, dans un premier temps, présenter brièvement les éléments d’un système de photolithographie par projection. Le lecteur trouvera, en annexe A de ce manuscrit, un complément d’information sur les éléments d’un outil de photolithographie et les critères de performance du procédé photolithographique. Ensuite, nous présenterons les solutions permettant d’améliorer la résolution de l’outil de photolithographie, notamment la correction des effets de proximité optique (OPC). Finalement, nous décrirons la modélisation du procédé photolithographique qui constitue le cadre du travail effectué dans cette thèse.

1.2 Éléments d'un système de photolithographie

Un système de photolithographie par projection est, de façon générale, constitué de quatre parties principales : la source d'illumination, le masque, l'optique de projection et la plaquette de silicium sur laquelle une résine photosensible est déposée. L'éclairage utilisé dans cette chaîne est de type Köhler. Un schéma simplifié de celui-ci est représenté sur la figure 1.2.

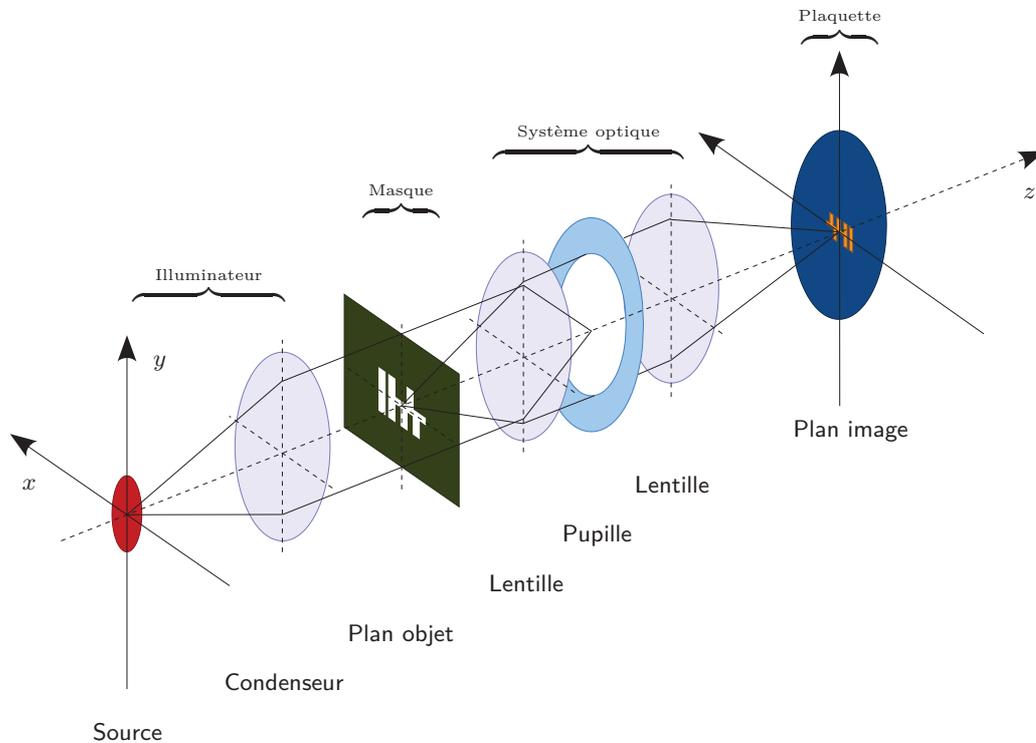


Figure 1.2 – Schéma d'un système d'illumination de type Köhler.

Dans ce type d'éclairage, l'image de la source d'illumination est projetée dans la pupille de l'optique de projection afin de la séparer du masque. L'uniformité de l'éclairage sur toute la surface du masque est assurée grâce à un système de lentilles, appelé le condenseur, permettant de projeter l'image de la source à l'infini au niveau du plan du masque. Les différents ordres de diffraction captés par l'optique de projection contribuent à la formation de l'image finale des éléments du masque dans la résine.

1.2.1 L'illumination

L'illumination constitue l'ensemble des dispositifs permettant d'éclairer le masque. Elle est composée d'une source lumineuse et d'un système de diaphragmes et de lentilles, appelé l'illuminateur, permettant de filtrer la lumière et de la projeter uniformément sur la surface du masque. La source lumineuse est constituée d'une lampe à décharge ou un laser émettant un rayonnement ultraviolet permettant d'insoler une résine photosensible et de modifier ses propriétés de solubilité. L'illumination est définie par une longueur d'onde d'exposition λ , une polarisation et une forme géométrique.

1.2.2 Le masque

Le masque est un élément important en photolithographie car il comporte le dessin du circuit à reproduire sur la plaquette. Il est composé d'un substrat transparent épais (de l'ordre de quelques millimètres) sur lequel est déposée une fine couche (de l'ordre de quelques dizaines

à quelques centaines de nanomètres) d'un matériau absorbant à la longueur d'onde. Suivant le facteur de réduction du système optique, les éléments du circuit sont gravés dans cette couche absorbante avec des dimensions 4 ou 5 fois supérieures à celles que l'on veut obtenir sur la plaquette. L'une des propriétés d'un masque de photolithographie est sa polarité. On parle de masque en champ clair lorsque les motifs sont définis dans la couche absorbante et de masque en champ obscur lorsqu'ils y sont ajourés (voir figure 1.3).

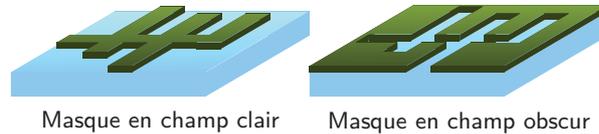


Figure 1.3 – Polarité d'un masque utilisé en lithographie optique : a) Masque en champ clair, b) Masque en champ obscur.

1.2.3 L'optique de projection

L'optique de projection est composée d'un système de lentilles qui permet de collecter une partie des ordres diffractés au niveau du masque et de les projeter dans le plan de la résine. La reconstitution de l'image des éléments du masque s'effectue alors par l'interférence des ondes collectées dans la résine. L'image aérienne désigne l'intensité lumineuse formée au niveau de la plaque dans l'air (ou le liquide d'immersion), c'est-à-dire en l'absence de la résine.

L'optique de projection se caractérise par son facteur de réduction et son ouverture numérique. Le facteur de réduction définit le rapport entre les dimensions des motifs sur plaquette et leurs dimensions sur le masque. L'ouverture numérique définit l'angle maximal que fait un rayon diffracté avec l'axe optique pouvant être capté par l'optique de projection. Du côté image, l'ouverture numérique est définie par :

$$ON_i = n_i \sin \alpha_i, \quad (1.1)$$

n_i étant l'indice de réfraction du milieu dans lequel les faisceaux se propagent en sortie de l'optique de projection (air ou liquide d'immersion) et α_i l'angle maximal d'obliquité que peut faire un rayon projeté avec l'axe optique. De la même façon, on exprime l'ouverture numérique, côté objet par :

$$ON_o = n_o \sin \alpha_o, \quad (1.2)$$

où n_o désigne l'indice de réfraction du milieu dans lequel baigne l'objet¹ et α_o le demi-angle maximal du cône de lumière créé par tous les faisceaux émanant de la source et atteignant un point du masque.

L'ouverture numérique maximale d'un outil lithographique est limitée par les facteurs géométriques des lentilles constituant l'optique de projection, en l'occurrence, leurs diamètres et leurs courbures. Nous verrons, dans le cas d'une lithographie à immersion, que l'on peut augmenter l'ouverture numérique en insérant un liquide en sortie du système optique permettant de régler le problème de la réflexion totale de la lumière entre les deux milieux.

Un dernier aspect de l'optique de projection concerne les aberrations optiques et la qualité de polissage des lentilles. Les techniques de projections actuelles, par répétition et balayage (de l'anglais "*step and scan*"), permettent de réduire les aberrations des lentilles en les moyennant sur toute la surface de l'optique de projection. Les derniers outils lithographiques (X1900i de la société ASML) sont équipés de systèmes de correction d'aberrations pour un maximum de performance.

1. L'objet ici est le masque. Une valeur typique de n_o serait donc égale à 1 qui est l'indice de réfraction de l'air ambiant.

1.2.4 L'empilement multicouche de la résine

1.2.4.1 La résine photosensible

La résine photosensible est un élément essentiel en photolithographie. C'est elle qui permet de transférer les motifs du masque sur la plaquette. La résine photosensible est un matériau organique photosensible habituellement utilisé à des longueurs d'onde dans le spectre ultraviolet. L'interaction de la lumière ultraviolette avec la résine photosensible permet de modifier ses propriétés physico-chimiques pour la rendre soluble dans un développeur aqueux. Outre la photosensibilité, les résines utilisées en photolithographie doivent avoir des propriétés de viscosité et d'adhérence pour permettre un étalement uniforme sur toute la surface du substrat. Elles doivent aussi avoir des propriétés mécaniques et thermiques leur permettant de résister aux traitements des étapes ultérieures tels que la gravure ou l'implantation ionique.

À l'heure actuelle, les résines les plus communément utilisées dans l'industrie, en raison de leur grande photosensibilité à la longueur d'onde 193 nm et de leurs performances à la fois mécaniques et thermiques, sont à amplification chimique. Elles sont composées :

- d'un polymère, pratiquement insoluble dans un développeur basique aqueux par le greffage partiel d'un groupement protecteur (inhibiteur de dissolution) sur sa chaîne principale,
- d'un composé photo-générateur d'acide (PAG) qui se transforme en un acide sous un rayonnement ultraviolet de longueur d'onde 284 ou 193 nm,
- et d'une base, qui lors de l'insolation, consomme en partie l'acide formé afin de contrôler sa concentration et limiter sa diffusion

1.2.4.2 Les couches antireflet

Les nombreuses réflexions de la lumière par les surfaces séparant la résine des autres couches de l'empilement conduisent à la formation d'ondes stationnaires dans la résine. Cela entraîne l'apparition de maxima et de minima d'énergie qui traduisent des concentrations non homogènes de l'acide photo-généré. Le profil final de la résine après développement présente des ondulations latérales (voir figure 1.4) pouvant représenter une fraction de la dimension du motif. Les ondes stationnaires entraînent une dégradation du contraste ainsi que le contrôle des dimensions.



Figure 1.4 – Profil d'une ligne de résine après développement : a) sans ondes stationnaires, b) avec ondes stationnaires.

Bien que les ondes stationnaires puissent être réduites par un recuit après exposition (de l'anglais “*Post-Exposure Bake*” ou PEB) [8], une couche antireflet ARC (de l'anglais “*Anti Reflective Coating*”) peut être insérée afin d'accomplir le même objectif [9,10]. Cette couche est un revêtement de quelques dizaines à quelques centaines de nanomètres formé d'un matériau organique ou inorganique. Elle assure une différence de phase de 180 degrés entre l'onde réfléchie et celle retransmise dans la résine, permettant ainsi de créer des interférences destructives au sein de la résine et d'atténuer les ondes stationnaires. De plus, elle permet d'absorber une partie de l'onde afin d'atténuer son amplitude quand elle sera retransmise dans la résine.

1.3 Techniques d'amélioration de la résolution

1.3.1 La résolution

La limite de résolution est un critère permettant d'affirmer si l'on arrive à distinguer deux objets éloignés très rapprochés. Elle est donnée par le critère de Rayleigh [11] qui permet d'exprimer l'ouverture angulaire limite permettant d'observer distinctement l'image de deux points à travers une ouverture (voir figure 1.5). Dans le cas d'une ouverture circulaire, la figure de diffraction obtenue, connue sous le nom de tâche d'Airy, est formée d'un disque lumineux central entouré de plusieurs anneaux lumineux. Selon Rayleigh, deux images sont discernables lorsque le maximum central de la première coïncide avec le premier minimum de la seconde. La séparation angulaire critique entre deux sources, correspondant au critère de Rayleigh s'écrit ainsi :

$$\sin \theta_c = 1.22 \frac{\lambda}{D}, \quad (1.3)$$

λ étant la longueur d'onde et D le diamètre de l'ouverture. D'après l'équation 1.3, on remarque que plus le diamètre D est petit plus l'angle θ_c est grand.

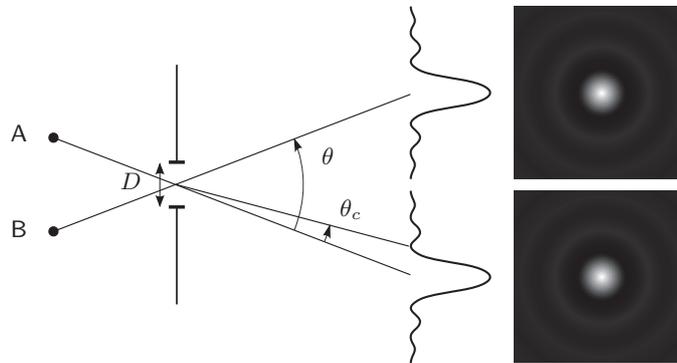


Figure 1.5 – Montage simplifié pour la démonstration du critère de Rayleigh.

Dans un instrument d'optique, l'angle θ_c est limité par l'ouverture numérique ($\text{ON} = \sin \theta$). Suivant le critère de Rayleigh, le diamètre minimal de la fente pouvant être résolu par le système optique correspond à un angle θ_c correspondant à deux fois l'ouverture numérique ON :

$$D_{\min} = 1.22 \frac{\lambda}{2 \times \text{ON}} = 0.61 \frac{\lambda}{\text{ON}} \quad (1.4)$$

En photolithographie, on introduit un facteur k_1 permettant de prendre en compte les paramètres du procédé lithographique (type d'illumination, cohérence de la source, type du masque, type de résine...). La résolution de l'outil lithographique est définie par rapport au plus petit pas de réseau :

$$p_{\lim} = k_1 \frac{\lambda}{\text{ON}}, \quad (1.5)$$

La limite inférieure du facteur k_1 est égale à 0.5 (pour une ouverture rectangulaire), ce qui correspond au fait que seulement deux ordres diffractés (0,1) ou (-1,0) puissent être captés par la pupille ($2 \times \text{ON}$). Dans la littérature, nous pouvons rencontrer des valeurs de k_1 de l'ordre de 0.3; la résolution étant exprimée cette fois-ci par rapport à la dimension minimale résolvable d'un motif (CD) et non pas par rapport au pas du réseau. Dans ce cas, la limite inférieure du facteur k_1 est égale à 0.25.

D'après l'équation 1.5, la résolution du système est fonction de la longueur d'onde λ , de l'ouverture numérique et du facteur k_1 , qui dépend des conditions de l'exposition et des paramètres du procédé lithographique [12]. Nous verrons dans la suite que l'on peut jouer sur ces trois paramètres afin d'améliorer la résolution de l'outil de lithographie optique.

1.3.2 Choix de l'illumination

1.3.2.1 La longueur d'onde

L'amélioration de la résolution passe, entre autres, par une diminution de la longueur d'onde λ . Cette technique était bien exploitée au cours du temps. En effet, au début des années 1980, les lampes à vapeurs de mercure : raies g, et i du spectre d'émission du mercure permettaient d'obtenir des longueurs d'ondes égales à 436 et 365 nm respectivement. Puis, au milieu des années 1990, l'apparition des lasers à excimère a permis de diminuer encore la longueur d'onde d'exposition. Les lasers krypton-fluor (KrF) ont permis d'atteindre une longueur d'onde égale à 248 nm. À l'heure actuelle, les outils de lithographie optique utilisent des lasers à base d'un mélange d'argon-fluor (ArF) émettant un rayonnement de longueur d'onde égale à 193 nm. L'étape suivante devait être le laser F₂ de longueur d'onde égale à 157 nm, mais cette solution a été quasiment abandonnée en raison de la difficulté de fabriquer des lentilles en CaF₂, matériau transparent à cette longueur d'onde. Aujourd'hui, la feuille de route de l'ITRS (*"International Technology Roadmap for Semiconductors"*) souligne la fin de l'ultraviolet profond en lithographie optique et prévoit un saut vers l'extrême ultraviolet (EUV) avec une longueur d'onde égale à 13,5 nm.

1.3.2.2 La géométrie de la source

La longueur d'onde est un paramètre important de la source d'illumination car elle est directement liée à la résolution du système optique. La diminution de la longueur d'onde a donc pour effet d'améliorer la résolution. Néanmoins, arrivant à des tailles de motifs avoisinant la longueur d'onde, l'optimisation du procédé lithographique devient nécessaire. L'une des techniques permettant d'améliorer la résolution est d'utiliser un éclairage permettant d'améliorer le contraste de l'intensité lumineuse au niveau de la résine. Cela consiste à utiliser des sources d'illumination de formes géométriques plus sophistiquées permettant un éclairage dit "hors axe". La figure 1.6 présente une liste non exhaustive de formes pouvant prendre une source d'illumination.

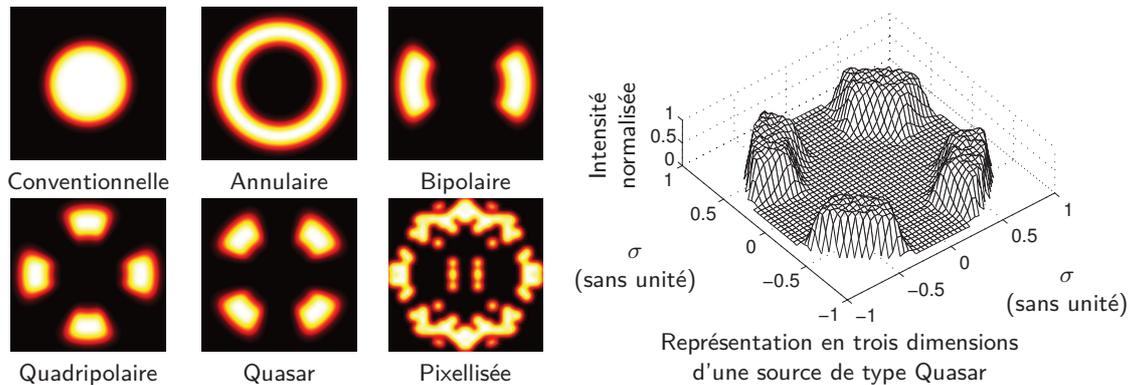


Figure 1.6 – Exemples de formes pour une source d'illumination (à gauche) et représentation en trois dimensions de l'intensité lumineuse normalisée d'une source de type Quasar (les facteurs de cohérence partielle de cette source sont $\sigma_{in}=0.6$ et $\sigma_{out}=0.9$ et l'angle d'ouverture est égal à 45 degrés) (à droite). Ces formes sont classées par ordre de complexité. En passant par les formes géométriques les plus simples comme les disques d'intensité, les sources d'illumination prennent des formes de plus en plus complexes. L'étendue spatiale de chacune des sources schématisées sur ce graphique est liée au facteur de cohérence partielle σ qui est compris entre 0 et 1.

Les sources conventionnelles sont les plus simples et se présentent dans l'espace sous forme d'un disque d'intensité. L'unique paramètre réglable de ces sources est leur diamètre qui est relié au facteur de cohérence partielle (voir paragraphe A.1.1, page 144). En plus complexe, les sources annulaires possèdent deux paramètres définissant les deux diamètres extérieur et intérieur de l'anneau. Ces sources permettent un éclairage "hors axe" car l'onde incidente émise

en tout point de l'anneau arrive avec un angle oblique sur le masque. Cette technique permet de faire passer d'avantage d'ordres diffractés à travers la pupille et améliore ainsi l'image obtenue. Le schéma de la figure 1.7 montre que l'on puisse arriver à faire passer les ordres ± 2 en utilisant un éclairage "hors axe".

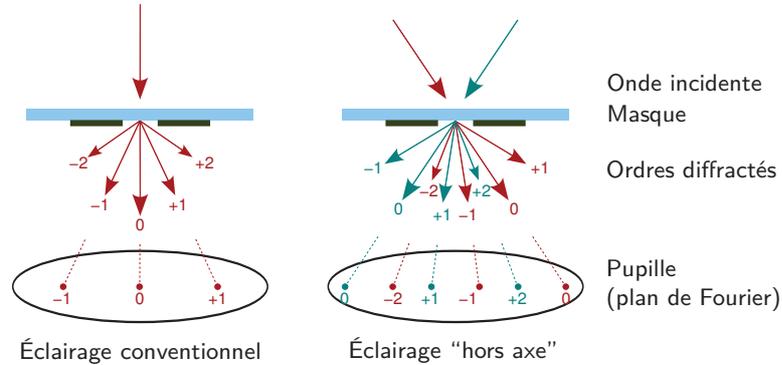


Figure 1.7 – Comparaison entre éclairages conventionnel et "hors axe". Ce dernier permet de faire passer plus d'ordres diffractés à travers la pupille du système optique de projection. L'interférence des ordres captés au niveau de la plaque permettent de reconstruire l'image du circuit du masque.

Dans le cas d'une lumière partiellement cohérente, l'amélioration du contraste de l'intensité pour les motifs de petites dimensions est le résultat de la diminution de la surface de non-recouvrement dans la pupille, c'est-à-dire la surface de l'ordre zéro qui ne participe pas à la formation de l'image, tout en augmentant celles des autres ordres ± 1 et ± 2 . Nous avons tracé sur la figure 1.8 les taux de surface de recouvrement et de non-recouvrement² en fonction du pas de réseau pour une source conventionnelle ($\sigma = 0.75$) et une source annulaire ($\sigma_{in} = 0.6$, $\sigma_{out} = 0.9$) avec une ouverture numérique égale à 1.35. D'après cette figure, nous pouvons remarquer que, pour les réseaux denses, c'est-à-dire, avec un pas de réseau inférieur à 140 nm, l'illumination annulaire permet de réduire la surface de non-recouvrement.

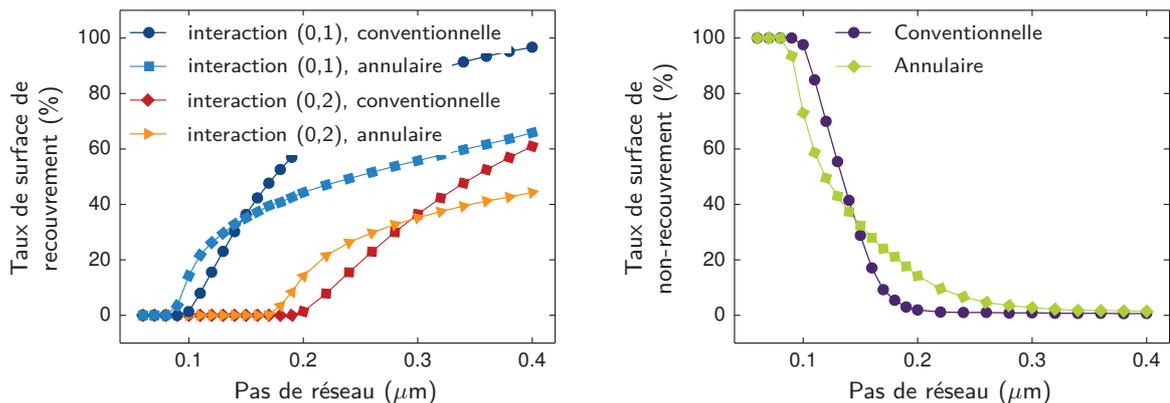


Figure 1.8 – Taux de surface de recouvrement (à gauche) et de non-recouvrement (à droite) de la source en fonction du pas de réseau pour deux sources d'illumination de formes conventionnelle ($\sigma = 0.75$) et annulaire ($\sigma_{in} = 0.6$, $\sigma_{out} = 0.9$). L'ouverture numérique ON utilisée ici est égale à 1.35.

Les sources bipolaires sont constituées de deux pôles orientés selon un certain angle par rapport à l'axe des abscisses. Elles permettent d'améliorer la qualité de l'image selon une seule direction. Les sources quadripolaires sont constituées de quatre pôles (deux pôles sont situés sur l'axe x et deux sur l'axe y) et permettent d'améliorer la qualité de l'image selon deux directions.

2. Pour le calcul des taux de surface de recouvrement et de non-recouvrement, le lecteur trouvera deux exemples illustrant la superposition des différents ordres diffractés pour deux sources conventionnelle (figure A.13) et annulaire (figure A.14) en annexe A de ce mémoire.

Il existe aujourd'hui des techniques, basées sur des simulations, permettant d'optimiser la source d'illumination pour un ensemble donné de structures. La technique SMO (de l'anglais "*Source-Mask Optimization*"), par exemple, permet d'optimiser simultanément la source d'illumination et le dessin à graver dans le réticule du masque [13]. Le résultat obtenu est une source de forme quelconque pouvant être décomposée en une multitude de points lumineux de différentes intensités ; on parle ici de sources pixellisées. À l'heure actuelle, ces sources sont encore au stade du développement et leur optimisation est coûteuse en temps et en argent. Le choix de ces sources dépend donc du gain que l'on peut obtenir par rapport à d'autres techniques concurrentes.

1.3.2.3 La polarisation de la source

La polarisation traduit la trajectoire du champ électrique \vec{E} (ou magnétique \vec{H}) de l'onde dans un plan perpendiculaire à l'axe de propagation mené par le vecteur d'onde \vec{k} . Lorsqu'une lumière est polarisée, le mouvement du vecteur \vec{E} n'est plus arbitraire mais peut être décrit par une trajectoire bien définie. Comme le montre le schéma de la figure 1.9, la trajectoire du vecteur \vec{E} d'une onde polarisée est un segment de droite (cas de la polarisation rectiligne), une ellipse (cas de la polarisation elliptique) ou un cercle (cas de la polarisation circulaire).

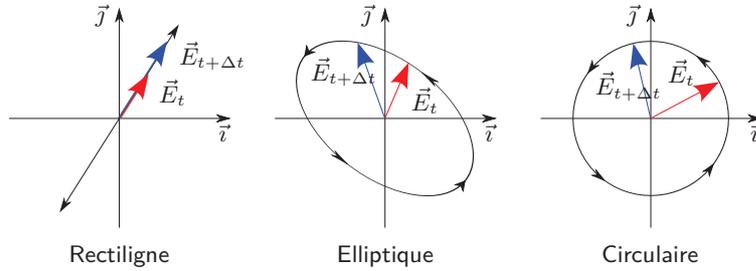


Figure 1.9 – Trajectoire du champ électrique \vec{E} d'une onde polarisée dans le plan perpendiculaire à l'axe de propagation. Cette trajectoire traduit l'état de polarisation de l'onde : rectiligne, elliptique ou circulaire.

Pour une lumière de polarisation quelconque, le vecteur \vec{E} du champ électrique peut s'exprimer en la somme de deux composantes perpendiculaires. Nous utilisons les notations TM (Transverse Magnétique) pour la composante radiale, c'est-à-dire, située dans le plan de propagation formé par le vecteur d'onde \vec{k} et l'axe optique, et TE (Transverse Électrique) pour la composante perpendiculaire à ce plan. Ainsi, une lumière polarisée en mode TE (ou TM) uniquement admet une polarisation rectiligne car la seule composante du champ électrique est confinée dans un unique plan.

Les résultats de l'interférence de deux ondes polarisées issues de deux points lumineux de la source sont différents selon que la polarisation est en TE ou en TM. En effet, considérons deux rayons lumineux portés par les vecteurs d'onde \vec{k}_1 et \vec{k}_2 et notons \vec{E}_1 et \vec{E}_2 leurs vecteurs champs électriques respectifs. Le champ résultant de l'interférence de ces deux rayons s'écrit ainsi :

$$\vec{E} = \vec{E}_1 + \vec{E}_2. \quad (1.6)$$

L'intensité obtenue est proportionnelle à la moyenne temporelle du carré du champ \vec{E} qui s'écrit (en simplifiant le terme de proportionnalité) [14] :

$$\begin{aligned} I &= \langle \vec{E}^2 \rangle_T \\ &= \langle (\vec{E}_1 + \vec{E}_2)^2 \rangle_T \\ &= \langle \vec{E}_1^2 + \vec{E}_2^2 + 2\vec{E}_1 \cdot \vec{E}_2 \rangle_T \\ &= \langle \vec{E}_1^2 \rangle_T + \langle \vec{E}_2^2 \rangle_T + 2 \langle \vec{E}_1 \cdot \vec{E}_2 \rangle_T \\ &= I_1 + I_2 + I_{12} \end{aligned} \quad (1.7)$$

avec $I_1 = \langle \vec{E}_1^2 \rangle_T$, $I_2 = \langle \vec{E}_2^2 \rangle_T$ et $I_{12} = 2 \langle \vec{E}_1 \cdot \vec{E}_2 \rangle_T$.

Si l'on considère deux ondes polarisées linéairement de la forme $\vec{E}_1(\vec{r}, t) = \vec{E}_{01} \cos(\vec{k}_1 \cdot \vec{r} - \omega t + \varepsilon_1)$ et $\vec{E}_2(\vec{r}, t) = \vec{E}_{02} \cos(\vec{k}_2 \cdot \vec{r} - \omega t + \varepsilon_2)$, le terme d'interférence I_{12} vaut :

$$I_{12} = 2 \langle \vec{E}_1 \cdot \vec{E}_2 \rangle_T = \vec{E}_{01} \cdot \vec{E}_{02} \cos(\delta)$$

avec δ , égal à $(\vec{k}_1 \cdot \vec{r} - \vec{k}_2 \cdot \vec{r} + \varepsilon_1 - \varepsilon_2)$, est la différence de phase résultant à la fois de la différence des chemins optiques et des angles de phase initiaux. Notons ici que si \vec{E}_{01} et \vec{E}_{02} (et donc \vec{E}_1 et \vec{E}_2) sont perpendiculaires, alors le terme l'intensité I_{12} vaut zéro. L'intensité d'une onde polarisée en mode TE ou en mode TM uniquement s'écrit donc ainsi :

$$I_{\text{TE}} = I_1 + I_2 \pm \sqrt{I_1 \cdot I_2} \cos(\delta) \quad (1.8a)$$

$$I_{\text{TM}} = I_1 + I_2 \pm \sqrt{I_1 \cdot I_2} \cos(\delta) |\cos(\alpha)| \quad (1.8b)$$

avec α l'angle que forme \vec{E}_1 avec \vec{E}_2 , comme le montre le schéma de la figure 1.10.

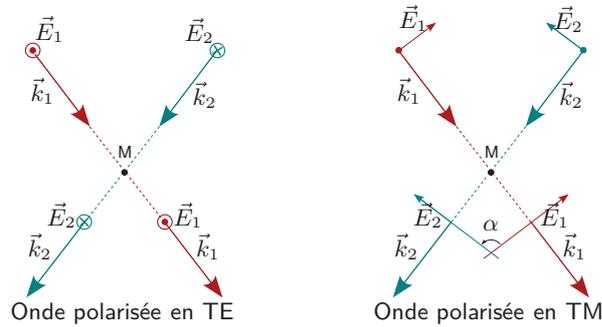


Figure 1.10 – Interférence de deux ondes obliques polarisées en mode TE (à gauche) et TM (à droite) dans la résine. Contrairement au cas d'une onde polarisée en mode TE où les vecteurs du champ électrique restent parallèles en tout point du milieu, ceux du mode TM font un angle quelconque dans l'empilement multicouche de la résine.

Le contraste (ou visibilité) v de l'intensité lumineuse d'une onde polarisée en TE ou en TM uniquement est alors donné par :

$$v_{\text{TE}} = 2 \frac{\sqrt{I_1 \cdot I_2}}{I_1 + I_2} \cos(\delta) \quad (1.9a)$$

$$v_{\text{TM}} = 2 \frac{\sqrt{I_1 \cdot I_2}}{I_1 + I_2} \cos(\delta) |\cos(\alpha)| \quad (1.9b)$$

$$= v_{\text{TE}} |\cos(\alpha)| \quad (1.9c)$$

En photolithographie, la polarisation de la source s'effectue par l'intermédiaire d'un système de polariseurs situés en amont du masque permettant de filtrer certaines composantes du champ électrique. Le principal intérêt d'utiliser une lumière polarisée en TE réside dans le fait qu'elle permet d'améliorer le contraste de l'intensité lumineuse dans la résine, en éliminant la composante TM dont l'intensité lumineuse obtenue s'atténue quand l'angle des rayons projetés s'approche de 90 degrés. En effet, en mode TE, les vecteurs du champ électrique issus de deux points lumineux diamétralement opposés arrivent parallèles dans l'empilement multicouche de la résine. Le contraste de l'intensité obtenue ne dépend donc pas des angles des rayons qui interfèrent. Au contraire, comme le montre le schéma de la figure 1.10, les vecteurs du mode TM arrivent avec un certain angle dans la résine, ce qui diminue le contraste de l'image obtenue.

Nous avons tracé, sur la figure 1.11, le cosinus de l'angle formé par deux rayons lumineux de polarisation TM interférant dans la résine en fonction de celui formé en sortie de l'optique de projection. Les deux courbes de cette figure représentent les deux cas d'une lithographie à 193 nm conventionnelle (dans un environnement sec) et à immersion. D'après cette figure, nous pouvons clairement remarquer que le cosinus de l'angle dans la résine décroît plus rapidement pour la lithographie 193 nm à immersion. Il en est de même pour le contraste de l'intensité

lumineuse formée par l'interférence de deux ondes polarisée en TM. De plus, la courbe montre que l'on arrive à inverser le contraste (le cosinus devient négatif) lorsque les angles dans la résine deviennent supérieurs à 90 degrés.

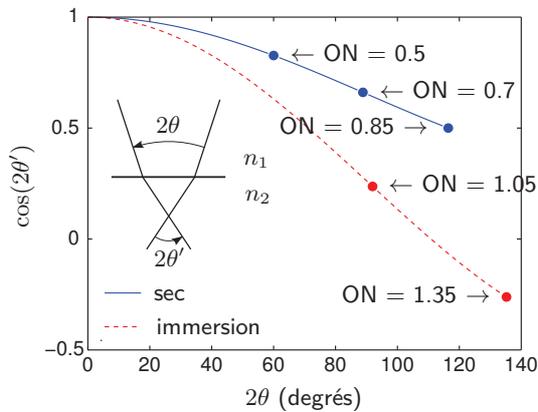


Figure 1.11 – Tracé du cosinus de l'angle formé par deux rayons interférant dans la résine (l'indice de réfraction de la résine pris ici est égal à 1.7) en fonction de celui formé en sortie de l'optique de projection. Les deux courbes désignent les deux cas d'une lithographie 193 nm conventionnelle (dans un environnement sec) et à immersion. Le liquide d'immersion utilisé ici est l'eau (l'indice de réfraction pris est égal à 1.46 à la longueur d'exposition).

Les résultats de simulation, présentés sur la figure 1.12, de l'intensité lumineuse à travers un masque formé par un réseau de lignes ($CD = 45$ nm, pas = 45 nm) illustrent cette inversion de contraste dans le cas d'une lumière polarisée en TM et d'une ouverture numérique égale à 1.3. Nous pouvons également remarquer une nette amélioration du contraste dans le cas d'une lumière polarisée en TE.

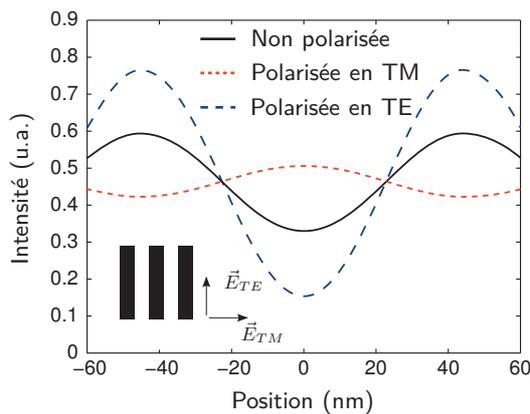


Figure 1.12 – Intensités lumineuses simulées pour un réseau de lignes ($CD=45$ nm, pas= 45 nm) avec $ON = 1.3$ et une illumination dipolaire $(\sigma_c, \sigma_r) = (0.9, 0.2)$ dans le cas d'une source non polarisée, une source polarisée en mode TE et source polarisée en mode TM. Source : [15]

Ainsi, nous avons vu que la polarisation de la source peut jouer un rôle important dans l'amélioration du contraste de l'intensité lumineuse; ce qui permet d'améliorer la résolution du procédé photolithographique. L'utilisation des sources polarisées a vu son domaine s'élargir lorsque l'on a commencé à utiliser des outils lithographiques à grande ouverture numérique [15–20], en particulier les systèmes à immersion. En 2005, la feuille de route de l'ITRS prévoyait l'utilisation d'illuminations polarisées pour les systèmes optiques opérant avec des ouvertures numériques supérieures à 0.85, et ce pour les nœuds technologiques 65 nm et au delà.

1.3.3 Choix du masque

Une des techniques permettant d'améliorer la résolution est d'utiliser des masques dits à décalage de phase ou PSM (de l'anglais "Phase Shift Mask"). Le concept de ces masques est assez ancien. Il a été proposé par Levenson [21] en 1982. L'idée de base était de supprimer une partie de la lumière diffractée du masque en introduisant un déphasage égal à 180 degrés entre les différentes régions du masque. Le schéma de la figure 1.13 montre que l'on peut obtenir un meilleur contraste de l'intensité lumineuse en utilisant ce type de masque par rapport aux masques binaires traditionnels.

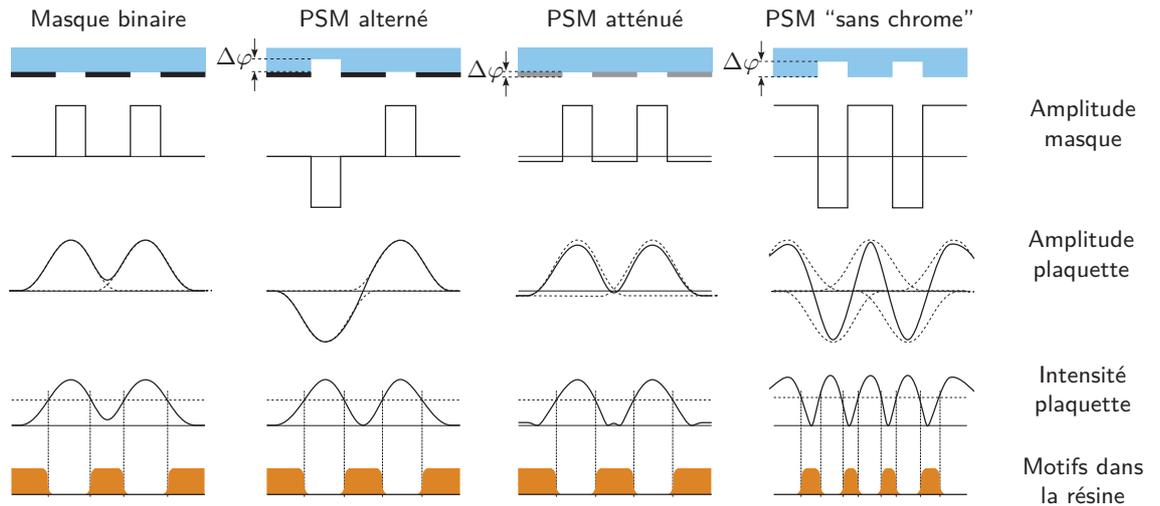


Figure 1.13 – Schéma simplifié illustrant une comparaison entre les intensités lumineuses obtenues avec un masque binaire et un masque à décalage de phase (PSM). Cette figure montre également que le contraste de l'intensité lumineuse est amélioré en utilisant un masque à décalage de phase. Selon le type de masque utilisé, le profil et les dimensions d'un motif dans la résine peuvent varier.

Il existe différentes méthodes pour réaliser un déphasage de 180 degrés. Tout dépend de la manière où et comment placer le déphasage entre les éléments du masque. La figure 1.14 présente une liste non exhaustive de configurations possibles, comparée au masque binaire conventionnel. Ces méthodes sont basées sur la gravure du substrat³, cas des masques "sans chrome" ou PSM alternés, ou l'utilisation d'un matériau déphaseur (par exemple, MoSi), cas d'un masque PSM atténué. On parle souvent de masques "forts" lorsqu'il s'agit d'une gravure dans le substrat et de masques "faibles" dans le cas où un matériau déphaseur est utilisé.

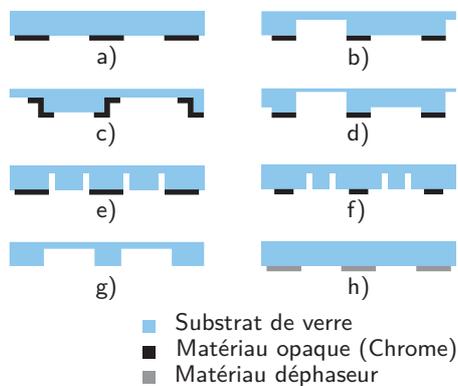


Figure 1.14 – Schéma illustrant différentes manières de réaliser un déphasage pour un masque de lithographie optique : a) masque conventionnel binaire (sans déphasage) b) masque PSM alterné c) masque PSM alterné avec des bords opaques, appelé aussi SCAA (de l'anglais "Sidewall Chrome Alternating Aperture") d) masque PSM alterné à double tranchées e) masque binaire avec un déphasage au niveau des bords f) masque binaire avec déphasage des motifs diffractants non résolus (SRAF) g) masque sans chrome présentant un déphasage entre les zones pleines et les zones creuses h) masque PSM atténué.

L'amélioration de la résolution est nettement supérieure dans le cas des masques forts par rapport à d'autres types de masques. En effet, les deux graphiques de la figure 1.15 montre un gain considérable en contraste de l'intensité lumineuse avec un masque PSM alterné ou sans chrome par rapport aux masques binaires ou PSM atténués.

3. La profondeur gravée permet d'avoir une différence de chemins optiques qui se traduit par une différence de phase :

$$\Delta\varphi = (n - 1) \frac{2\pi e}{\lambda}$$

avec e l'épaisseur gravée du substrat, n l'indice du matériau à graver et λ la longueur d'onde.

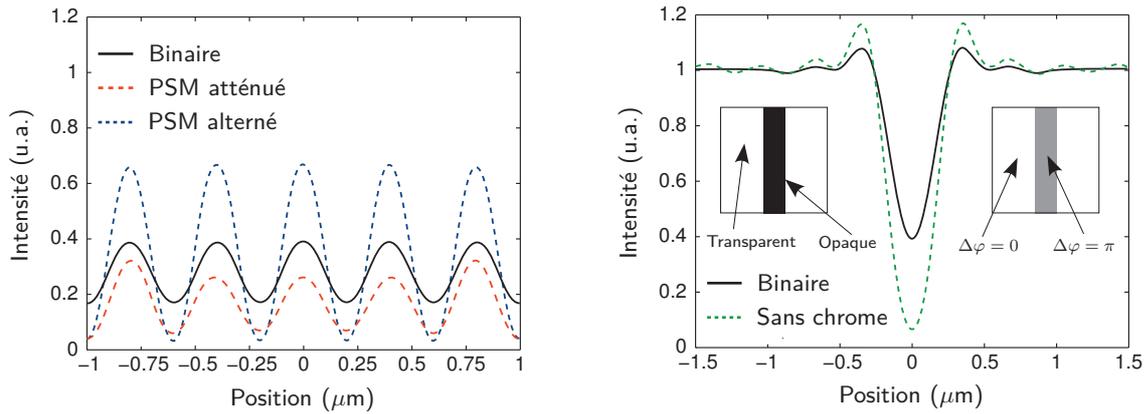


Figure 1.15 – Comparaison entre les intensités lumineuses simulées avec différents types de masques pour un réseau de lignes (à gauche) et une ligne isolée (à droite). Une amélioration du contraste est observée pour les masques PSM alterné et sans chrome par rapport au masque binaire conventionnel. Source : [22]

Bien que les masques PSM alternés permettent d’atteindre d’excellentes performances lithographiques, ils présentent des inconvénients faisant obstacle à une utilisation dans un environnement de production à fort volume. En effet, ces masques sont convenables pour les motifs de forte périodicité (cas des motifs 1D comme les réseaux de lignes par exemple) mais il est difficile d’effectuer un déphasage lorsqu’il s’agit d’un dessin arbitraire (cas des motifs 2D). En outre, dans des configurations particulières, des conflits de phase peuvent apparaître entre les régions à 0 et 180 degrés. Ces zones à conflit causent (par interférences destructives) l’impression de lignes indésirables dans la résine. Le problème peut être résolu de deux façons. La première est d’effectuer une deuxième étape d’exposition avec un second masque permettant d’enlever les lignes superflues. Cette méthode est coûteuse car elle nécessite la fabrication d’un second masque et augmente le temps de l’étape de lithographie. La deuxième méthode est d’insérer des régions à déphasage intermédiaire (60 ou 120 degrés) entre les zones à conflit. Cependant, cette fois, il faut prévoir de l’espace pour l’insertion de ces zones dans le dessin, ce qui n’est pas une tâche aisée lorsqu’il s’agit de circuits à forte densité. Le dernier inconvénient des masques PSM alternés vient de la complexité de la fabrication. La réalisation du déphasage nécessite plusieurs étapes d’exposition, de dépôt et de gravure. De plus, les plages de spécifications concernant l’alignement des motifs entre les différentes étapes et l’erreur sur la phase sont assez restreintes. Tout ceci justifie le prix dix fois plus élevé d’un masque PSM alterné par rapport à celui d’un masque binaire conventionnel.

Contrairement aux PSM alternés, les masques atténués sont moins performants au niveau de l’amélioration de la résolution mais ils sont moins coûteux et plus accessibles. Le déphasage dans les PSM atténués est réalisé par différents types de matériaux. La transmission de tels matériaux doit être bien contrôlée (typiquement entre 0 à 15%) afin d’éviter l’apparition des maxima locaux d’intensité provoquant l’impression de motifs indésirables dans la résine. De plus, à 248 et 193 nm, ces matériaux doivent avoir des propriétés diélectriques tout en gardant un caractère métallique pour faciliter la dissipation des charges électriques lors de la gravure du réticule. Dans la littérature, différents types de matériaux ont été proposés. L’approche commune est d’utiliser des matériaux non stœchiométriques [23–28] (exemples : $MoSi_xO_yN_z$, SiN_x , CrO_xF_y ou $(AlN)_{(1-x)}(CrN)_x$) dont les compositions peuvent être ajustées pour atteindre les propriétés optiques voulues. L’avantage de cette méthode est qu’elle permet de déposer une couche unique du matériau. Cependant, la composition chimique des éléments est instable et peut rapidement changer suite à une oxydation chimique par exemple. Cette approche a donc été abandonnée laissant le champ libre aux développements de déphaseurs formés par un empilement de plusieurs couches de matériaux différents [28, 29] (exemples : SiN/TiN , Si_3N_4/TiN , AlN/CrN).

Les masques sans chrome, appelés également CPL (de l’anglais “*Chrome-less Phase-shift Lithography*”), sont similaires aux PSM atténués, mais possèdent un taux de transmission égal à 100%. Pour ce type de masques, les motifs sont directement gravés dans le substrat de verre

de telle façon qu'un déphasage égal à 180 degrés est créé. Au niveau de la résine, l'intensité lumineuse s'annule aux bords des motifs, par interférence destructive des ondes à 0 et à 180 degrés. Après développement, deux lignes de résine apparaissent au niveau de chaque bord d'un motif isolé. Si l'on rapproche suffisamment les bords d'un motif, nous obtiendrons donc une seule ligne de résine. L'exposition des lignes larges nécessite donc leurs découpes en plusieurs lignes fines très rapprochées.

1.3.4 La correction des effets de proximité optique

La correction des effets de proximité optique ou OPC (de l'anglais "*Optical Proximity Correction*") est l'une des techniques d'amélioration de la résolution (RET). Elle est communément utilisée afin de compenser les erreurs dues à la diffraction et aux effets liés au procédé de fabrication. En effet, les structures de petites tailles perdent une partie importante de leurs spectres de fréquences spatiales, coupée par la pupille de l'optique de projection. Leurs images apparaissent ainsi avec des distorsions telles que des coins arrondis et des lignes moins larges que celles du dessin du circuit. Les conséquences peuvent être plus ou moins pénalisantes, allant d'une diminution des performances ou une perte de rendement, jusqu'à une défaillance totale du circuit. Le traitement du dessin par OPC permet d'anticiper les distorsions de forme et de taille en appliquant des corrections sur les masques permettant d'obtenir des motifs plus proches des motifs visés (voir figure 1.16).

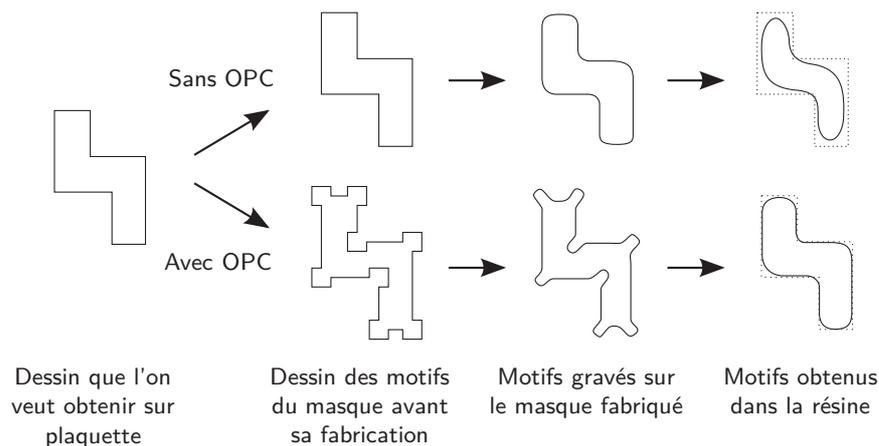


Figure 1.16 – Principe de la correction des effets de proximité optique (OPC). Les motifs présentés ici sont dessinés à l'échelle de la plaquette.

Les effets de proximité optique apparaissent lorsque l'on s'approche de la limite de résolution. L'image d'un motif présentera des irrégularités de forme et de taille en fonction de la proximité des différents motifs sur le masque. Les principaux effets de proximité optique sont de plusieurs types :

- La densité environnante : cet effet représente les variations des dimensions critiques CD en fonction de la densité des motifs voisins. Une ligne du dessin du masque s'imprime différemment dans la résine selon qu'elle est isolée ou qu'elle appartient à un réseau. Il peut y avoir également d'autres effets d'origine optique : la lumière parasite, appelée "flare" [30–32] ou chimique : le "flare" chimique [33, 34].
- La non-linéarité : cet effet s'exprime en calculant les rapports de tailles des motifs sur le masque et dans la résine. Une ligne isolée sur le masque de dimension égale à deux fois celle d'une autre ligne isolée ne traduit pas forcément ce même rapport de dimensions dans la résine. La non-linéarité est strictement liée au facteur d'erreur du masque MEEF.
- Les retraits des bouts de ligne et l'arrondissement des coins : cet effet est causé par la perte des hautes fréquences spatiales des motifs. Il est de plus en plus visible pour les motifs de faibles dimensions.

- Les effets liés à une anomalie ou un dysfonctionnement de l'outil de lithographie : ces effets s'expriment en terme d'instabilité de la dose d'exposition par exemple. Nous pouvons avoir une variation importante de la dose dans une même puce (intra-champs) ou entre différentes puces (inter-champs).

Il existe un dernier type d'effets de proximité non lithographique : celui de la gravure. En effet, la gravure n'est pas isotrope suivant que le motif est situé dans un environnement dense ou isolé. Le biais (qui signifie la différence de tailles avant et après gravure) dépend de la surface gravée avoisinant le motif. Les effets de la gravure sont donc des effets de densité à courte distance.

Il existe deux types de traitement OPC : par règles géométriques ou par modèle (de l'anglais "*Rule Based OPC*" et "*Model Based OPC*"). La correction par règles consiste à rajouter de la matière lorsqu'il y a un retrait de bouts de ligne ou des coins convexes arrondis, et d'en retirer dans le cas de coins concaves arrondis. La correction de la non-linéarité (et de la gravure) se fait par une table formée des valeurs de biais pour plusieurs couples ligne/espace. Ces biais seront additionnés aux dimensions des motifs sur le masque. La correction des effets de la densité se fait en insérant des motifs diffractants non résolus, dits SRAF (de l'anglais "*Sub Resolution Assist Feature*"), à proximité des bords isolés permettant de les densifier.

La deuxième méthode de traitement OPC se fait par l'utilisation de modèles permettant de simuler les contours des motifs dans la résine ou après gravure. Le traitement se fait de façon itérative où l'on fait bouger à chaque itération la position des bords des motifs sur le masque auquel on applique les modèles afin de prédire les contours dans la résine. Le processus s'arrête lorsque l'erreur quadratique moyenne entre les bords des motifs visés et les contours est minimisée. En pratique, un nombre entre 4 et 8 itérations est suffisant pour avoir un résultat convenable.

Ces deux méthodes de traitement OPC peuvent être utilisées séparément ou simultanément durant les différentes étapes du processus. Un exemple décrivant les étapes d'un traitement OPC est décrit sur la figure 1.17. Si l'on considère un dessin de départ (le design) avec des dimensions visées après gravure, la première étape consiste à redimensionner par règles géométriques les motifs du dessin afin de viser les dimensions souhaitées dans la résine. Ensuite, les motifs diffractants non résolus (SRAF) sont insérés à proximité des bords isolés. Les bords du dessin sont ensuite divisés en fragments pouvant bouger séparément. Finalement, un processus itératif permet de simuler les contours dans la résine par l'intermédiaire d'un modèle empirique et de déplacer les fragments afin de minimiser l'écart entre les contours et les bords du dessin souhaité. Le résultat obtenu est celui qui va être gravé sur un masque.

Les SRAF sont des motifs de tailles réduites qui servent à améliorer la résolution des structures isolées sans pour autant être imprimés dans la résine. Ils sont non résolubles car dessinés avec des dimensions bien inférieures à la résolution du scanner. Les SRAF existent dans les deux tonalités possibles du masque et peuvent être insérés dans le dessin du circuit indépendamment du type de masque choisi. On parle de barres diffractantes (de l'anglais "*Scattering Bars*") lorsqu'il s'agit de rajouter de la matière sur le réticule, c'est-à-dire, d'insérer des lignes de petites tailles autour des motifs du circuit. Ces barres diffractantes servent à améliorer la résolution des lignes isolées. Inversement, les tranchées diffractantes (de l'anglais "*Scattering Trenches*") consistent à retirer de la matière à l'intérieur des motifs de grandes tailles et permettent d'améliorer la résolution des espaces isolés. Finalement, les SRAF peuvent être insérés lorsqu'il y a suffisamment d'espace entre les motifs du masque et les algorithmes utilisés dans ce placement opèrent soit par règles géométriques ou alors en utilisant un modèle OPC.

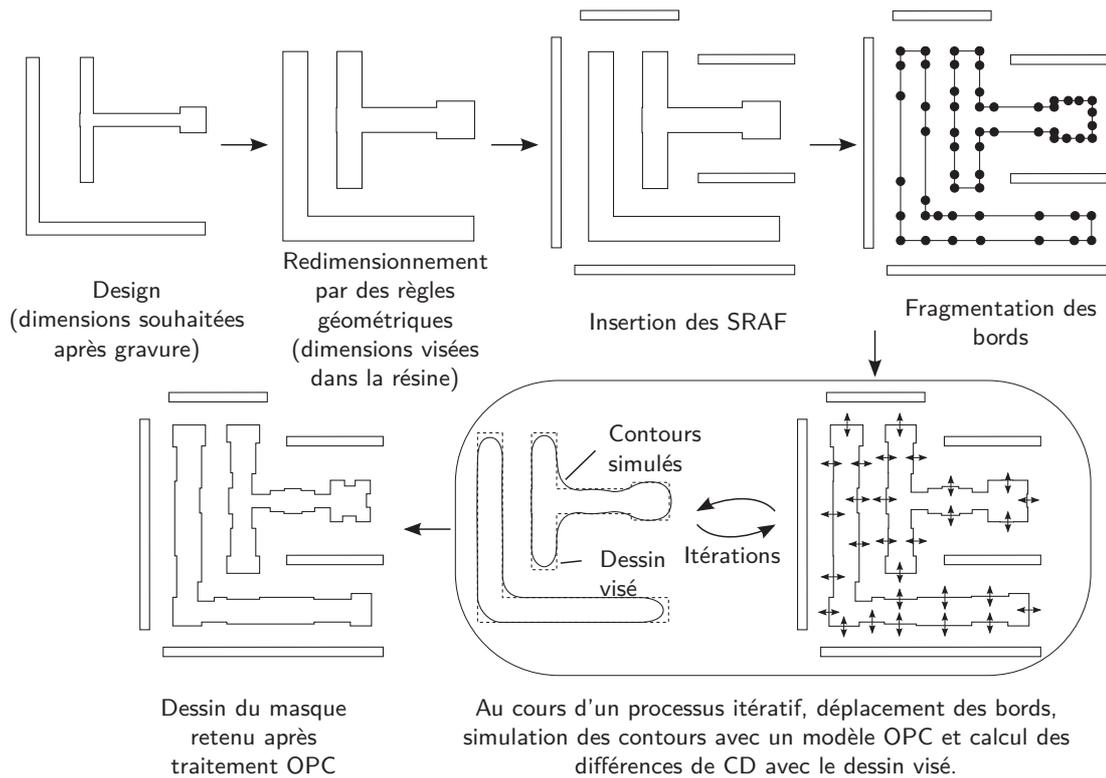


Figure 1.17 – Représentation simplifiée des différentes étapes de traitement OPC du dessin d'un circuit intégré.

1.3.5 La lithographie par immersion

La lithographie par immersion est la technique qui est actuellement utilisée afin de pousser les limites du procédé lithographique à 193 nm en terme de résolution. Le principe est simple. En insérant un liquide d'indice de réfraction supérieur à 1 entre la dernière lentille de projection et la résine (qui est l'air dans un environnement sec), on arrive à mieux adapter les impédances entre le dernier élément optique et le liquide d'immersion permettant d'éliminer les phénomènes de réflexion optique à cette interface. Un schéma simplifié de ce principe est illustré sur la figure 1.18. Le fait de passer plus d'ordres diffractés permet de résoudre des réseaux de lignes de pas plus faibles. L'immersion a donc pour effet d'atteindre des ouvertures numériques supérieures à 1, qui est la limite physique dans l'air. À l'heure actuelle, les outils de lithographie par immersion à base d'eau ($n = 1.44$ à la longueur d'onde de 193 nm) permettent d'atteindre une ouverture numérique de l'ordre de 1.2.

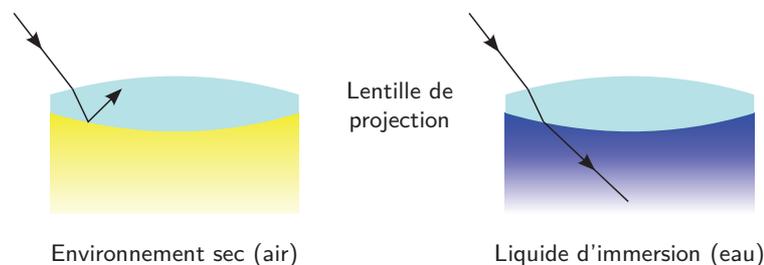


Figure 1.18 – Principe de la lithographie à immersion. L'immersion permet d'éviter que le rayon soit complètement réfléchi à l'interface séparant l'élément optique et le liquide d'immersion.

1.3.6 La lithographie par double exposition

Les techniques dites d'expositions multiples sont des méthodes permettant d'augmenter la densité des motifs par unité de surface. Elles permettent à travers de multiples expositions alignées d'obtenir des dimensions bien inférieures à la limite de résolution de l'outil lithographique. La lithographie par double exposition DPL (de l'anglais "*Double Patterning Lithography*") est un cas particulier où l'on effectue seulement deux expositions successives. Le principe de cette technique est illustré sur la figure 1.19.

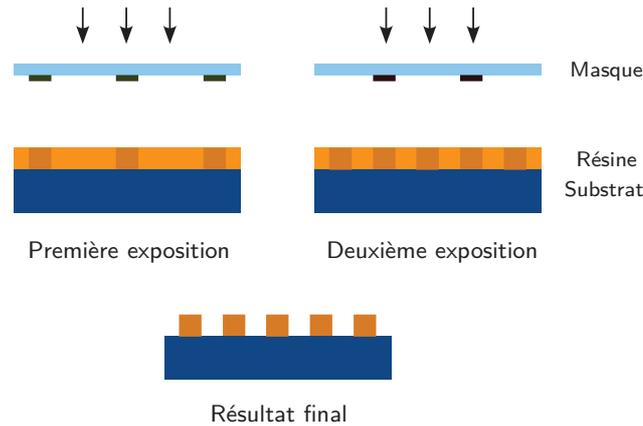


Figure 1.19 – Principe de la lithographie par double exposition

Le procédé de la lithographie DPL peut se faire par plusieurs façons. L'une des méthodes utilisées, appelée LELE (de l'anglais "*Litho-Etch, Litho-Etch*"), s'effectue par deux premières étapes d'exposition et de gravure, suivies de deux secondes étapes similaires [35]. Le principe de cette méthode est illustré sur la figure 1.20. Cette approche nécessite le dépôt à la surface du substrat d'une couche de sélectivité à la gravure qui est différente de celle du substrat. L'avantage de cette technique est qu'elle est stable et qu'elle est tout à fait appropriée pour les nœuds technologiques supérieurs à 22 nm. Cependant, le dédoublement des étapes d'exposition et de la gravure la rend coûteuse. De plus, l'alignement est d'autant plus problématique que les dimensions des motifs sont faibles.

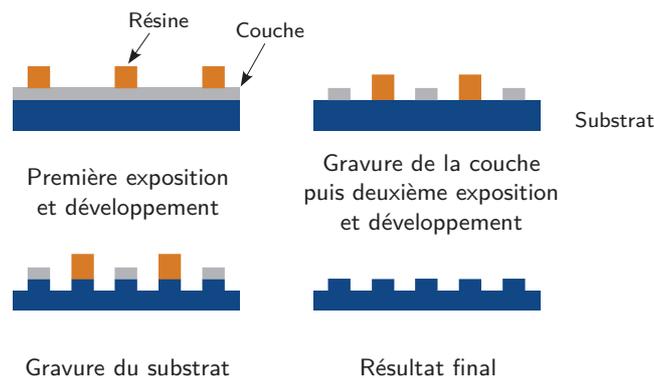


Figure 1.20 – L'une des variétés de la lithographie par double exposition est la technique LELE. Elle permet une division par deux du pas de réseau par l'intermédiaire de deux étapes d'exposition et de gravure.

Une méthode alternative permettant de réduire le nombre des étapes du procédé est la technique LFLE (de l'anglais "*Litho-Freeze, Litho-Etch*"). Elle s'effectue par un durcissement thermique ou chimique de la première résine développée, suivi d'une deuxième exposition et d'une gravure [36]. Cette approche permet donc de réduire les coûts de fabrication. En revanche, elle nécessite un contrôle supplémentaire des dimensions de la première résine durcie qui peut se dégrader aux étapes de durcissement et du second développement. À l'heure actuelle, le problème

de la dégradation des résines en durcissement a fait obstacle pour une utilisation industrielle de la technique LFLE, en faveur de l'approche LELE qui est plus souvent utilisée.

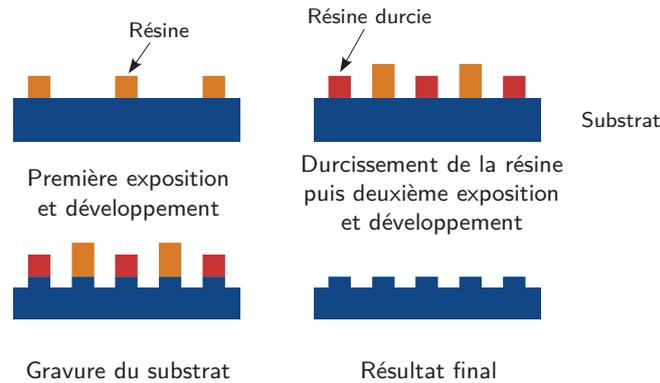


Figure 1.21 – Représentation des différentes étapes d'une lithographie de double exposition par durcissement de la résine, LFLE

Une autre approche prometteuse de la double exposition est la technique DTD (de l'anglais "*Dual Tone Development*"), au cours de laquelle, une exposition classique est suivie de deux étapes de développement. Les résines employées sont des résines à double tonalité, c'est-à-dire qu'elles réagissent différemment à la lumière lorsqu'elles sont exposées à différentes doses. Un premier seuil de dose permet, après développement, d'enlever la partie exposée, tandis qu'un deuxième seuil permet d'enlever la partie non-exposée. L'approche DTD se présente comme un moyen de réduire les étapes du procédé, et par conséquent, le coût de la fabrication. Cependant, à l'heure actuelle, les résines à double tonalité ne sont pas encore assez matures (au niveau des propriétés thermiques et chimiques) pour être utilisées dans un environnement de production [37]. De plus, la rugosité des lignes LER reste un problème majeur pour ce type de résines qui, après deux étapes de développement successives, peuvent présenter des LER en dehors des spécifications.

Contrairement aux trois approches citées ci-dessus, où l'alignement est problématique pour les nœuds technologiques avancés, l'approche dérivée des techniques de double exposition SADP (de l'anglais "*Self-Aligned Double Patterning*"), également appelée SIT (de l'anglais "*Sidewall Image Transfer*") permet un auto-alignement des motifs. Dans cette technique, la résine est développée puis une fine couche d'un matériau est déposée de manière à garder la forme initiale de la topographie. La gravure de cette couche fait apparaître des espaceurs de part et d'autre des bords de la résine. Celle-ci est ensuite enlevée et les espaceurs restant sont utilisés comme couche protectrice pour graver le substrat. La figure 1.22 illustre les différentes étapes de cette approche.

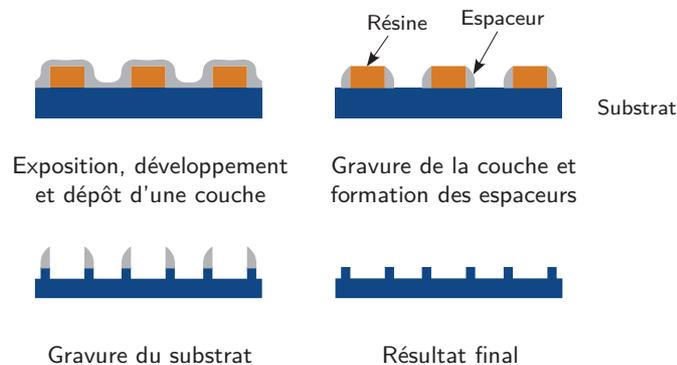


Figure 1.22 – Principe de l'approche SADP permettant de diviser par deux le pas d'un réseau de lignes.

L'avantage d'une telle approche vient du fait qu'une seule exposition est nécessaire et que le problème de l'alignement est résolu. De plus, l'uniformité des dimensions critiques et la ru-

gosité des lignes LER sont améliorées par rapport à d'autres techniques conventionnelles [38] et correspondent aux exigences de la feuille de route de l'ITRS pour le nœud 22 nm. Cependant, le problème majeur de cette technique est qu'elle n'est pas tout à fait adaptée à un circuit arbitraire, notamment les structures 2D où les configurations sont difficiles à réaliser ou impossibles. Même si le recours à une deuxième exposition permet de nettoyer des régions indésirables, cette technique reste restreinte aux cas où la disposition des motifs est régulière (DRAM, contacts. . .) [39, 40].

1.4 Modélisation du procédé photolithographique

L'une des méthodes de correction des effets de proximité optique OPC est d'utiliser des modèles permettant de simuler le procédé de photolithographie afin d'anticiper les éventuelles distorsions géométriques des éléments imprimés dans la résine afin de les corriger sur le masque. Le processus se fait de manière itérative où, après chaque itération, le masque est réajusté (par déplacement des bords des motifs) puis les contours de la résine sont simulés. Le but est de parvenir au dessin du masque qui minimise l'écart entre les contours simulés et le dessin que l'on souhaite obtenir sur la plaquette.

À la manière des outils lithographiques modernes, les modèles OPC sont divisés en deux parties. Une première partie purement optique qui permet de simuler l'intensité lumineuse arrivant dans la résine. Cette partie, communément appelée "modèle optique", permet de décrire toutes les étapes régissant la diffraction et la propagation de la lumière depuis la source d'illumination jusqu'à la plaquette en passant par le masque et l'optique de projection. La seconde partie est celle qui décrit les phénomènes physico-chimiques qui surgissent lors de l'interaction entre la lumière et la résine, mais aussi les effets de la diffusion chimique lors des étapes de recuit et de développement. Commençons tout d'abord par présenter les modèles optiques.

1.4.1 Modélisation de la partie optique

Un modèle optique peut être considéré comme une fonction de transfert $H(x, y)$ permettant de calculer l'intensité lumineuse dans la résine à partir du dessin du masque. Cette intensité est donnée par la figure d'interférence des ondes diffractées au niveau du masque. Les modèles optiques actuels sont basés sur l'optique de Fourier afin de calculer l'image aérienne du masque. Nous allons brièvement présenter le principe de cette méthode.

1.4.1.1 Optique de Fourier et formule d'Abbé

L'optique de Fourier considère que l'onde diffractée se propage selon des angles correspondant aux fréquences spatiales du masque. Chaque angle définit un ordre de diffraction bien déterminé. En traversant le système optique, les ondes se propageant suivant des angles θ_i tels que $\sin \theta_i \geq \text{ON}$ sont coupées par la pupille. En conséquence, seulement une partie de l'onde de départ peut atteindre la plaque. Ainsi, l'image du masque est approximativement reconstruite dans la résine par interférence des ondes résultantes.

Le masque est présenté par deux fonctions scalaires discrètes de transmission $t(x, y)$ et de phase $\varphi(x, y)$ telles que le champ électrique s'écrit :

$$e_m(x, y) = \sqrt{t(x, y)} e^{j\varphi(x, y)}, \quad (1.10)$$

où

$$t(x, y) = \begin{cases} t_0 & \text{si } (x, y) \in \text{circuit du masque} \\ 1 & \text{sinon} \end{cases}$$

et

$$\varphi(x, y) = \begin{cases} \varphi_0 & \text{si } (x, y) \in \text{circuit du masque} \\ 0 & \text{sinon} \end{cases}$$

La fonction de transfert du système optique (représenté par les lentilles de projection et la pupille) agit comme un filtre passe-bas. Pour une ouverture circulaire de la pupille, la fonction de transfert peut s'écrire, en l'absence d'un décalage en focus, de la façon suivante :

$$K(u, v) = \begin{cases} 1 & \text{si } \sqrt{u^2 + v^2} \leq ON \\ 0 & \text{sinon} \end{cases}, \quad (1.11)$$

ON étant l'ouverture numérique de la pupille et u et v les fréquences spatiales sur les axes horizontal et vertical respectivement du masque. Les fréquences spatiales au niveau de la plaque sont calculées par produit de la transformée de Fourier du masque par la fonction de transfert K selon l'équation suivante :

$$E_p(\nu_x, \nu_y) = K(\lambda\nu_x, \lambda\nu_y) \cdot E_m(\nu_x, \nu_y) \quad (1.12)$$

où ν_x et ν_y désignent les fréquences spatiales, $E_m(\nu_x, \nu_y)$ la transformée de Fourier du champ électrique $e_m(x, y)$, et λ la longueur d'onde d'exposition.

$$\begin{aligned} E_m(\nu_x, \nu_y) &= TF(e_m(x, y))(\nu_x, \nu_y) \\ &= \iint_{x,y} e_m(x, y) e^{-j2\pi(x\nu_x + y\nu_y)} dx dy \end{aligned} \quad (1.13)$$

Ainsi, le champ électrique peut être calculé au niveau de la plaque par l'application de la transformée de Fourier inverse :

$$\begin{aligned} e_p(x, y) &= TF^{-1}(E_p(\nu_x, \nu_y))(x, y) \\ &= \iint_{\nu_x, \nu_y} K(\lambda\nu_x, \lambda\nu_y) \cdot E_m(\nu_x, \nu_y) \cdot e^{j2\pi(x\nu_x + y\nu_y)} d\nu_x d\nu_y \end{aligned} \quad (1.14)$$

L'intensité sur la plaque est alors calculée par :

$$I(x, y) = |e_p(x, y)|^2 \quad (1.15)$$

Cette expression de l'intensité n'est valable que pour une source ponctuelle. Pour une lumière partiellement cohérente, l'intensité totale est calculée par la somme de toutes les intensités provenant de chaque point individuel de la source. Ainsi, pour une source de forme conventionnelle et de facteur de cohérence partielle σ , on obtient :

$$I(x, y) = \iint_{\alpha^2 + \beta^2 \leq \sigma \cdot ON} S(\alpha, \beta) \left| \iint_{\nu_x, \nu_y} K(\lambda\nu_x + \alpha, \lambda\nu_y + \beta) \cdot E_m(\nu_x, \nu_y) \cdot e^{j2\pi(x\nu_x + y\nu_y)} d\nu_x d\nu_y \right|^2 d\alpha d\beta, \quad (1.16)$$

où $S(\alpha, \beta)$ est la fonction de transmission de la source d'illumination dans le plan de la pupille.

Cette dernière expression de l'intensité est souvent appelée la formule d'Abbé pour une lumière partiellement cohérente. Il est intéressant de dire que l'intégration sur toute la source se fait en dernier lieu et que par conséquent, la formule d'Abbé permet dans certaines applications d'optimiser la source. En effet, il suffit de calculer la seconde intégrale de l'équation 1.16 et de faire varier l'intensité de chaque point de la source en vue de l'optimiser.

1.4.1.2 Formule d'Hopkins pour une lumière partiellement cohérente

La double intégrale dans l'équation 1.14 est approchée par deux sommes discrètes sur les fréquences spatiales. On obtient alors l'expression suivante :

$$\begin{aligned}
I(x, y) &= \iint_{\alpha^2 + \beta^2 \leq \sigma \cdot ON} S(\alpha, \beta) \left| \sum_{\nu_x} \sum_{\nu_y} K(\lambda\nu_x + \alpha, \lambda\nu_y + \beta) \cdot E_m(\nu_x, \nu_y) \cdot e^{j2\pi(x\nu_x + y\nu_y)} \right|^2 d\alpha d\beta \\
&= \iint_{\alpha^2 + \beta^2 \leq \sigma \cdot ON} S(\alpha, \beta) \sum_{\nu_x} \sum_{\nu_y} \sum_{\nu'_x} \sum_{\nu'_y} E_m(\nu_x, \nu_y) E_m^*(-\nu'_x, -\nu'_y) \cdot e^{j2\pi(x(\nu_x + \nu'_x) + y(\nu_y + \nu'_y))} \\
&\quad \cdot K(\lambda\nu_x + \alpha, \lambda\nu_y + \beta) \cdot K^*(-\lambda\nu'_x + \alpha, -\lambda\nu'_y + \beta) d\alpha d\beta \tag{1.17}
\end{aligned}$$

À cause de la linéarité de l'intégrale, l'équation 1.17 peut être réécrite en intégrant sur la source en premier puis en sommant sur les fréquences spatiales du masque, ce qui nous donne l'expression suivante :

$$\begin{aligned}
I(x, y) &= \sum_{\nu_x} \sum_{\nu_y} \sum_{\nu'_x} \sum_{\nu'_y} E_m(\nu_x, \nu_y) E_m^*(-\nu'_x, -\nu'_y) \cdot e^{j2\pi(x(\nu_x + \nu'_x) + y(\nu_y + \nu'_y))} \\
&\quad \cdot \iint_{\alpha^2 + \beta^2 \leq \sigma \cdot ON} S(\alpha, \beta) K(\lambda\nu_x + \alpha, \lambda\nu_y + \beta) \cdot K^*(-\lambda\nu'_x + \alpha, -\lambda\nu'_y + \beta) d\alpha d\beta \tag{1.18}
\end{aligned}$$

Cette dernière expression est la formule d'Hopkins [41] pour une lumière partiellement cohérente (dans le cas d'une source conventionnelle de facteur de cohérence partielle σ). Le dernier terme, la double intégrale, est généralement appelé TCC (de l'anglais "*Transfer Cross Coefficients*").

$$TCC(\nu_x, \nu_y, \nu'_x, \nu'_y) = \iint_{\alpha^2 + \beta^2 \leq \sigma \cdot ON} S(\alpha, \beta) K(\lambda\nu_x + \alpha, \lambda\nu_y + \beta) \cdot K^*(-\lambda\nu'_x + \alpha, -\lambda\nu'_y + \beta) d\alpha d\beta \tag{1.19}$$

$$I(x, y) = \sum_{\nu_x} \sum_{\nu_y} \sum_{\nu'_x} \sum_{\nu'_y} E_m(\nu_x, \nu_y) E_m^*(-\nu'_x, -\nu'_y) \cdot e^{j2\pi(x(\nu_x + \nu'_x) + y(\nu_y + \nu'_y))} TCC(\nu_x, \nu_y, \nu'_x, \nu'_y) \tag{1.20}$$

Le TCC est une fonction des fréquences spatiales qui ne dépend que de la source d'illumination et de la pupille. Ainsi, étant indépendant du masque, le TCC est très utilisé dans les logiciels commerciaux de traitement OPC pour l'optimisation des structures d'un masque. En Effet, l'OPC change la géométrie des structures du masque afin de reproduire les motifs visés sur la plaquette de silicium. Ceci explique le fait que plusieurs logiciels de traitement OPC utilisent cette méthode.

Malgré l'avantage d'être indépendant du masque, le TCC ne peut être utilisé que dans le cas où les ordres de diffraction sont connus et sont indépendants de la source d'illumination utilisée. Pour une source "hors-axe" émettant des rayons avec une incidence oblique, cette hypothèse n'est plus valide du moment que les angles d'incidence deviennent suffisamment grands pour que les amplitudes des ordres ± 1 soient modifiées [42]. Il existe néanmoins des solutions à ce problème, notamment la modification des tables TCC par l'ajout de termes résiduels dits "non-hermitiens" [43] permettant de prendre en compte les effets de l'incidence oblique dans le calcul de l'intensité lumineuse.

1.4.2 Modélisation de la partie résine

Pendant l'exposition de la plaque, la composition chimique d'une résine photosensible est modifiée. Le recuit après exposition PEB (de l'anglais "*Post-Exposure Bake*") [8] permet l'accélération de la réaction chimique, c'est-à-dire, la diffusion de l'acide photo-généré afin d'activer la solubilité des parties exposées. Il sert aussi à homogénéiser les concentrations locales des espèces acides et basiques et de lisser leurs profils afin d'éviter la formation des ondes planes stationnaires dans la résine [8]. La diffusion des espèces chimiques dépend principalement de la quantité de lumière localement absorbée et de la durée du recuit. Cependant, d'autres effets optiques ou chimiques peuvent se rajouter afin de modifier la concentration locale de l'acide photo-généré

(flare optique [30–32], flare chimique [33,34], fluctuation de la température [44], etc) et modifier ainsi la position de l’interface entre la partie exposée ou non. Finalement, pendant l’étape du développement, la position de l’interface de transition peut aussi être modifiée par diffusion des espèces chimiques entre le liquide de développement et la résine.

Dans la littérature, de nombreux modèles ont été détaillés pour la modélisation des différentes étapes de la formation des contours. Il existe des modèles plus ou moins complexes basés uniquement sur des lois physiques. Les modèles physiques utilisés dans les logiciels actuels de simulation se servent des techniques proposées par Dill et ses collaborateurs [45–48] en 1975. Les modèles de Dill permettent de prendre en compte les effets des étapes d’exposition et du développement. Durant l’exposition, la distribution de l’acide photo-généré dans le film de la résine est calculée à partir du profil de l’image aérienne. Cette distribution d’acide forme ce que l’on appelle l’image latente. Pendant le développement, le taux de dissolution de la résine est relié à la concentration instantanée de l’acide à l’interface du liquide de développement. Des méthodes itératives permettent le calcul du profil de la résine à partir de l’image latente obtenue. Avec l’apparition des résines à amplification chimique, une étape supplémentaire, le recuit PEB, a été introduite dans le procédé. La diffusion de l’acide photo-généré pendant le recuit a été modélisée par l’application de la seconde loi de Fick à l’image initiale [49]. Les concentrations de l’acide photo-généré et de celui utilisé dans la réaction de déprotection sont calculées simultanément par voie itérative. L’image latente est alors obtenue par le calcul de la concentration du polymère non-protégé. Le taux de dissolution est calculé à partir du taux de déprotection de la résine.

Plusieurs équations ont été proposées dans la littérature pour le calcul du taux de dissolution [46, 48, 50–52]. Parmi les modèles publiés, Mack a proposé un modèle cinétique pour le développement de résine de type Novolak/DNQ [50]. Ce premier modèle présente des problèmes pour les faibles et forts taux de dissolution. Un second modèle amélioré [51, 52] permettant de compenser les erreurs a été ensuite proposé par Mack. Puis, ses travaux avec Arthur ont permis de créer un autre modèle plus précis appelé le modèle “Notch” [53–55]. La figure 1.23 présente les courbes du taux de dissolution en fonction de la concentration de l’inhibiteur pour les trois modèles de Mack.

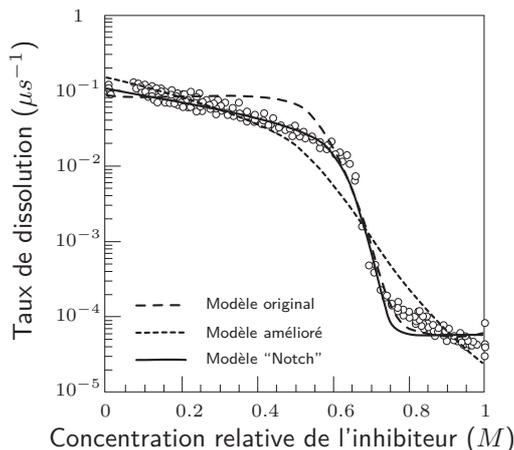


Figure 1.23 – Tracés du taux de dissolution de la résine en fonction de la concentration relative de l’acide inhibiteur (mesures expérimentales) et courbes de régression de trois modèles de développement basés sur les équations de Mack : modèle original, modèle amélioré et modèle “Notch” [55].

En traitement OPC, le problème majeur de ces modèles réside dans le fait qu’ils comportent un nombre élevé de paramètres à ajuster. De plus, étant donné que ces modèles décrivent successivement les étapes d’exposition, de recuit et du développement, le temps de calcul est élevé. C’est pourquoi, les outils de traitement OPC utilisent des modèles simplifiés, qui ne décrivent pas toutes les étapes du procédé mais sont capables de modéliser le procédé dans sa globalité. Ces modèles, ne comportant qu’un petit nombre de paramètres, peuvent être ajustés par des mesures expérimentales.

L’équation générale des modèles empiriques peut être décrite comme suit :

$$f(x, y) = \sum_k c_k T_k(x, y). \quad (1.21)$$

où $f(x, y)$ est la fonction caractéristique⁴ du modèle, c_k les coefficients à ajuster et T_k les termes du modèle. Ces derniers sont formés par des opérations mathématiques simples (dérivés partiels, produits simples ou de convolution, maximum et minimum d'intensité. . .) appliquées à l'intensité lumineuse initiale ou à l'image du masque et constituent les paramètres du modèle. Nous allons maintenant présenter, par ordre de complexité, les modèles empiriques que nous allons utiliser pour l'étude de la stabilité.

1.4.2.1 Modèles simples

1.4.2.1.1 Modèle à seuil constant

Le modèle à seuil constant CTR (de l'anglais "*Constant Threshold Resist*") [56] est le modèle empirique le plus simple. Comme son nom l'indique, il permet de calculer les contours des motifs dans la résine en appliquant un seuil constant à l'intensité lumineuse :

$$I(x, y) = I_s = C^{te}, \quad (1.22)$$

I_s étant l'intensité seuil d'extraction des contours.

La figure 1.24 ci-après illustre les principales étapes suivies pour la modélisation des contours d'un circuit dans la résine. L'image du circuit dans la résine a été simulée à l'aide de la partie optique. L'application d'un seuil constant à l'intensité lumineuse obtenue a permis l'extraction des contours des motifs. L'avantage d'utiliser ce modèle, à la seule variable définie par le seuil d'extraction I_s , est qu'il ne dépend pas du type de motifs exposé, et qu'il est simple et facile à implémenter. Toutefois, ce modèle ne tient pas compte de la diffusion dans la résine. Pour les nœuds technologiques avancés, où les effets dans la résine sont non-négligeables, le modèle à seuil constant peut présenter certaines limitations pour la prédiction des contours de la résine.

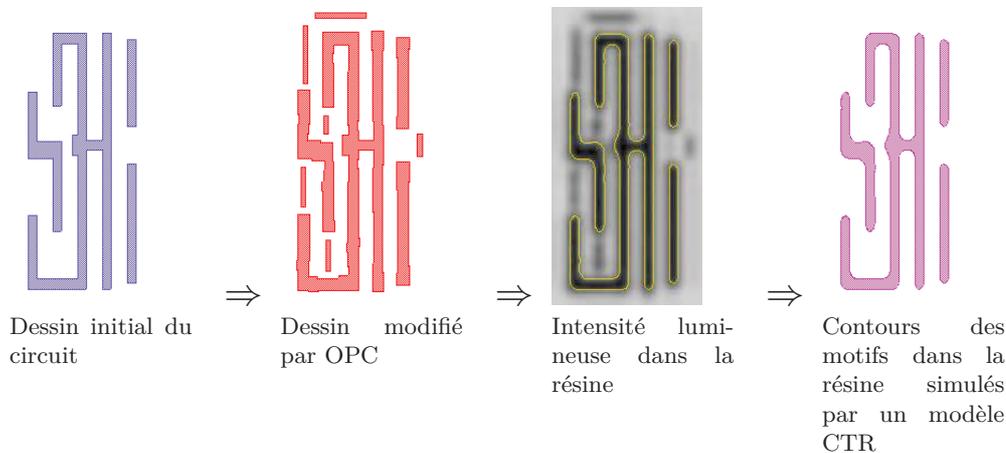


Figure 1.24 – Schéma illustrant les principales étapes suivies pour la modélisation des contours d'un circuit dans la résine dans le cas d'un modèle CTR.

1.4.2.1.2 Modèle à diffusion gaussienne

La convolution de l'intensité lumineuse dans la résine par une fonction gaussienne est une manière simple pour la modélisation de la diffusion chimique dans la résine. La somme pondérée de plusieurs produits de convolution de l'intensité nous permet de modéliser des effets multiples de la diffusion chimique. Comme l'indique l'équation 1.23, le modèle à diffusion gaussienne n'est autre qu'un modèle à seuil constant appliqué à cette somme pondérée :

$$c_1 I(x, y) \otimes G_1(x, y) + \dots + c_N I(x, y) \otimes G_N(x, y) = I_s, \quad (1.23)$$

4. $f(x, y)$ peut être un seuil variable (cas du modèle VT5) ou une surface résine (cas du modèle CM1).

où N est le nombre de gaussiennes utilisées et c_k sont tels que $\sum_{k=1}^N c_k = 1$.

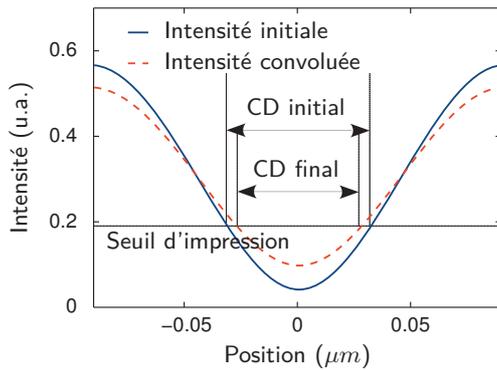


Figure 1.25 – Intensités lumineuses d'une coupe transversale d'une ligne isolée avant et après convolution par une fonction gaussienne d'écart type σ égal à 20 nm.

La figure 1.25 montre les profils des intensités lumineuses d'une ligne isolée avant et après convolution par une fonction gaussienne d'écart type $\sigma = 20\text{nm}$. La dimension de la ligne isolée obtenue en appliquant un seuil constant à l'intensité lumineuse est modifiée et la convolution diminue le contraste de l'intensité. Ce qui permet d'approcher les effets physiques réels dans la résine. Cependant, cette approche suppose que la diffusion se fait d'une manière isotrope et uniforme selon les directions x et y et qu'elle est négligée selon z . Ce qui n'est pas tout à fait juste pour modéliser les contours en chaque point de la résine. Elle permet néanmoins de prédire les allures des profils de concentration des acides photo-générés dans un unique plan.

1.4.2.2 Modèles avancés

1.4.2.2.1 Modèle à seuil variable

Le dernier modèle à seuil variable, connu sous le nom VT5 (de l'anglais "Variable Threshold 5"), a été introduit par Y. Granik et N. Cobb [57] et implémenté dans l'outil de traitement OPC, Calibre™ de la société Mentor Graphics. Ce modèle définit le seuil d'impression en fonction des propriétés de l'intensité lumineuse dans la résine. Ces propriétés sont les valeurs maximale et minimale de l'intensité I_{max} et I_{min} , sa pente Slope et sa courbure Factor⁵.

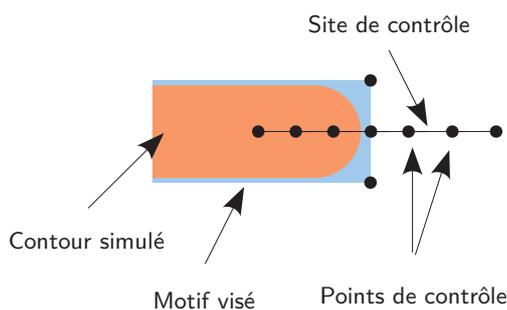


Figure 1.26 – Placement d'un site de simulation dans le cas d'un modèle VT5.

Ces paramètres sont calculés le long d'un motif sur ce que l'on appelle site de contrôle. Un site de contrôle est un concept introduit par Mentor Graphics définissant une ligne sur laquelle un fragment du bord du motif se déplace durant un traitement OPC, comme le montre l'exemple de la figure 1.26. Un site de contrôle est constitué de plusieurs points uniformément espacés définissant les lieux où l'intensité lumineuse dans la résine est calculée [57, 58]. Les paramètres sont extraits dans un intervalle, appelé fenêtre d'exploration, autour du seuil de référence I_r (voir figure 1.27).

5. La courbure est calculée au seuil de référence I_r

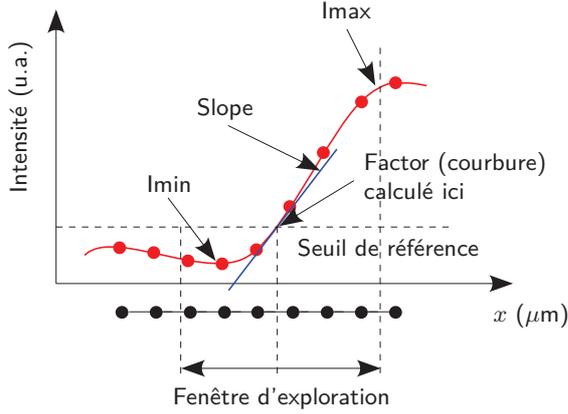


Figure 1.27 – Paramètres de l'intensité lumineuse dans la résine.

L'intensité maximale I_{max} est évaluée par le maximum d'intensité dans la fenêtre d'exploration. I_{min} est le minimum d'intensité. La valeur de la pente $Slope$ est approchée au seuil de référence I_r par une régression linéaire de la droite passant par les deux points de contrôle les plus proches. La courbure $Factor$ est calculée par une approximation des gradients d'intensité dans les directions parallèle et perpendiculaire au site de contrôle considéré.

Le seuil d'impression calculé par un modèle VT5 est une fonction polynomiale de second ordre des paramètres I_r , I_{max} , I_{min} , la pente $Slope$ et la courbure $Factor$ que l'on peut exprimer par l'équation suivante :

$$I_s = I_r + \sum_{m=0}^2 \sum_{i \neq j} c_k P_i^m P_j^{2-m}, \quad (1.24)$$

où P_i est l'un des paramètres I_{max} , I_{min} , $Slope$ et $Factor$, et c_k les coefficients du polynôme. Il est évident que le modèle à seuil constant est une forme particulière du modèle VT5 (en fixant $c_k = 0, \forall k$).

1.4.2.2.2 Modèle à convolution étendue

Le modèle à convolution étendue est celui qui a été implémenté dans le logiciel de traitement OPC Tachyon™ et est la propriété de la société Brion. Il permet de calculer sur une grille de simulation la surface $R(x, y)$ de l'image latente dans la résine. Cette surface est la somme de plusieurs termes permettant de modéliser différents effets physiques et chimiques du procédé lithographique. Le calcul des contours dans la résine s'effectue par la résolution de l'équation $R(x, y) = 0$ où R s'exprime selon cette équation :

$$\begin{aligned} R = & c_{AA}A + c_{M_{av}}M_{av} + c_{A_p}A_p \otimes G_{A_p} + c_{B_p}B_p \otimes G_{B_p} + c_{A_m}A_m \otimes G_{A_m} \\ & + c_{B_n}B_n \otimes G_{B_n} + c_{AG_1}A \otimes G_{AG_1} + c_{AG_2}A \otimes G_{AG_2} + c_{MG_1}M \otimes G_{MG_1} \\ & + c_{MG_2}M \otimes G_{MG_2} + c_{Slope}Slope + theta - thresh \end{aligned} \quad (1.25)$$

où c_A , $c_{M_{av}} \dots c_{AG}$ et $theta$ sont des coefficients à optimiser, A_p et A_m sont les parties supérieures de l'intensité lumineuse tronquées au seuil b_0 et b_m respectivement, B_p et B_n sont les parties inférieures de l'intensité lumineuse tronquées au seuil b_0 et b_n respectivement, M_{av} est la moyenne locale de l'image du masque M dans un rayon autour du point (x, y) et $Slope$ est la pente locale de l'intensité lumineuse au point (x, y) . Pour comprendre la signification des termes et leur contribution dans la modélisation des effets physiques, le lecteur trouvera un complément d'information en annexe C de ce mémoire. Le modèle à convolution étendue est une extension du modèle à diffusion gaussienne avec la possibilité de modéliser d'autres phénomènes optiques ou chimiques tels que les effets à courte, moyenne et longue distances (par convolution de l'image du masque).

1.4.2.2.3 Modèle à opérateur différentiel

C'est un modèle qui a été introduit par Y. Granik sous le nom CM1 (de l'anglais "Compact Model 1") pour viser les nœuds technologiques avancés (45 nm et au delà) [59,60]. Il est disponible

dans l'outil de traitement OPC, CalibreTM, et est la propriété de Mentor Graphics. Dans ce modèle, le seuil d'impression est fixé comme dans le cas du modèle à seuil constant CTR. En revanche, il n'est plus appliqué à l'intensité lumineuse de départ $I(x, y)$ mais à une surface $R(x, y)$ définissant l'intensité de l'image latente du circuit. Cette dernière est calculée par une combinaison linéaire de termes M_i selon l'équation :

$$R(x, y) = \sum_i c_i M_i(x, y). \quad (1.26)$$

Les termes $M_i(x, y)$ sont des opérateurs mathématiques qui sont appliqués à l'intensité lumineuse initiale dans la résine $I(x, y)$:

$$M_i(x, y) = \left(\left(\nabla^k I_{\pm b}(x, y) \right)^n \otimes G_{s,p}(x, y) \right)^{\frac{1}{n}}. \quad (1.27)$$

où

$G_{s,p}(x, y)$ est la famille de fonctions noyaux de Gauss-Laguerre et $I_{\pm b}(x, y)$ est la partie supérieure ou inférieure de l'intensité lumineuse initiale tronquée au seuil b .

Le lecteur trouvera un complément d'information en annexe C de ce mémoire sur la signification de ces termes ainsi que sur leurs principales fonctions. Le modèle CM1 représente un cas généralisé du modèle à convolution étendue (sans convolution de l'image du masque).

1.4.3 Discussions

À l'heure actuelle, la formulation de Hopkins [41], basée sur la théorie de la diffraction de Fourier, reste la méthode la plus utilisée pour calculer l'image d'un masque entier à travers un système optique. Une application potentielle de cette méthode est le traitement OPC d'un circuit qui permet de corriger les distorsions des motifs liées aux procédés de la lithographie et de la gravure. L'avantage majeur de cette méthode est qu'elle est rapide et permet d'obtenir une approximation adéquate de l'intensité lumineuse dans la résine, ce qui séduit les industriels trouvant un compromis entre la précision de calcul et le temps de cycle.

Seulement, l'adoption de la théorie de Fourier nécessite de considérer trois approximations. La première stipule que les champs électrique et magnétique ne sont pas couplés et qu'il n'existe aucune interaction entre ces deux champs. Cela signifie que le champ électrique seul peut servir à calculer l'intensité lumineuse sur la plaque. La seconde approximation considère que les effets vectoriels provenant du masque sont négligeables. Ceci vient du fait qu'on ne tient pas compte de la polarisation de la lumière à travers le masque. La troisième approximation considère que le masque utilisé est infiniment fin et que la topographie du masque n'a aucune influence sur le calcul des ordres de diffraction. Dans ces conditions, l'adoption d'une fonction scalaire pour décrire le masque est légitime.

Toutefois, ces trois approximations ne sont valables que dans le cas où les dimensions critiques (CD) des structures dessinées sur le masque sont grandes par rapport à la longueur d'onde d'exposition λ . Pour les nœuds technologiques avancés où les dimensions critiques sont de l'ordre de la longueur d'onde, voire plus petites, ces conditions ne sont plus valides. De plus, l'utilisation des masques à décalage de phase annule l'approximation de non-couplage des champs électrique et magnétique. L'introduction de systèmes optiques à grande ouverture numérique modifie l'état de polarisation de l'onde après avoir parcouru le masque. Dans ces cas, les effets vectoriels ne peuvent plus être négligés et le calcul de l'image aérienne par la théorie de Fourier ne sera plus aussi précis. D'autres méthodes approchées basées sur un calcul rigoureux du champ électrique peuvent être utilisées (voir chapitre 4).

Les approximations employées dans la partie optique des modèles OPC actuels rendent la tâche plus difficile aux modèles empiriques pour une modélisation des contours dans la résine avec des contraintes de plus en plus strictes pour répondre aux exigences de la feuille de route de l'ITRS. Là, où auparavant, il suffisait d'appliquer un seuil constant à l'intensité lumineuse pour

extraire les contours dans la résine, les modèles empiriques actuels deviennent de plus en plus complexes afin de capturer le maximum d'effets. Les différences de nature des termes empiriques utilisés dans ces modèles risquent de créer des phénomènes incontrôlables pouvant affecter à leur stabilité numérique.

1.5 Conclusion

La résolution de l'outil lithographique à 193 nm devient problématique pour les techniques de lithographie optique classiques. L'extension, dans un premier lieu, de la lithographie 193 nm par immersion, puis, l'utilisation massive des techniques d'amélioration de la résolution (RET) ont rendu possible le développement des technologies 45 nm et 32 nm. Néanmoins, le coût et la complexité des masques, avec l'utilisation d'algorithmes de traitement OPC de plus en plus agressifs, risquent de rendre cette méthode trop coûteuse pour la réalisation de circuits en petite série.

Au delà des coûts ascendants des masques, les recommandations de la feuille de route de l'ITRS en terme de dimensions critiques, par exemple, sont d'une importance capitale afin de continuer à suivre le rythme de la miniaturisation prédit par la loi de Moore. Cependant, les exigences spécifiées par l'ITRS deviennent de plus en plus difficiles à respecter tant au niveau de la plaquette qu'au niveau du masque. Les contraintes imposées à la rugosité des lignes (LWR), par exemple, sont calculées comme étant une fraction des dimensions critiques et sont par conséquent plus fortes lorsque les éléments du masque sont de plus en plus petits. À cela se rajoutent les règles de dessin (de l'anglais "*Design Rule Manual*" ou DRM) regroupant les contraintes de fabrication et de performance du circuit intégré comme la longueur et la largeur des grilles des transistors, les arrondissements des coins (de l'anglais "*Corner Rounding*") et les retraits au niveau des bouts de lignes (de l'anglais "*end cap*") , mais aussi les contraintes de fabrication du masque comme les espaces limites autorisés entre deux lignes ou deux coins et les dimensions minimales autorisées pour les SRAF.

Outre les spécifications de la feuille de route de l'ITRS et les contraintes de fabrication, la correction OPC doit tenir compte des effets indésirables du procédé de fabrication qui s'accroissent avec l'utilisation des systèmes à grandes ouvertures numériques et avec la réduction des dimensions critiques. Pour toutes ces raisons, l'OPC devient une tâche ardue qui nécessite une bonne compréhension de la physique afin de mieux modéliser les effets de la photolithographie et de la gravure. Ainsi, les modèles OPC doivent garantir une simulation fine des contours dans la résine et prédire les éventuelles dérives du procédé de la photolithographie dans plusieurs conditions du focus et de la dose. En plus d'être précis et prédictifs, ils doivent être rapides afin de minimiser les coûts en temps et en argent requis pour la préparation et la fabrication des masques. Tout cela justifie les nombreuses approximations utilisées dans la modélisation optique de l'intensité lumineuse et la mise au point de modèles purement empiriques pour la modélisation des contours dans la résine.

À l'heure actuelle, les dimensions critiques des motifs (à l'échelle du masque, c'est-à-dire, en 4X) avoisinent la longueur d'onde d'exposition ($\lambda = 193$ nm). L'utilisation des masques atténués à décalage de phase (de l'anglais "*Phase Shift Mask*" ou PSM) annule l'approximation de non-couplage entre les champs électrique et magnétique et limite ainsi la validité des approximations effectuées dans la partie optique. De plus, à haute ouverture numérique, les effets topographiques du masque ne peuvent plus être négligés. Le maintien des approximations dans la partie optique du modèle rend la tâche plus difficile aux modèles empiriques qui deviennent de plus en plus complexes. Ainsi, nous avons vu, dans ce chapitre, différents modèles empiriques avec différents niveaux de complexité. À l'instar des modèles à opérateur différentiel (CM1) ou à convolution étendue (Brion), les modèles avancés incluent de nombreux termes empiriques agissant sur la reconstitution de l'image latente et permettant de modéliser différents effets. Toutefois, les paramètres des modèles ne sont pas tous de nature physique. Ils peuvent présenter ainsi des effets

incontrôlables et/ou imprévisibles dans la modélisation des contours. La question de la stabilité de tels modèles devrait donc être posée.



Deuxième partie

Étude de la stabilité des modèles
empiriques

Chapitre 2

Causes et effets de l'instabilité des modèles empiriques

En photolithographie, la fabrication des masques constitue une étape importante de la technologie. Durant cette étape, le dessin du circuit est traité par OPC avant d'être transféré dans le réticule. Dans le cas de la photolithographie à 193 nm, le traitement OPC du masque semble être une étape essentielle pour les technologies actuelles (45 nm et au delà). Au fil des années, le traitement OPC n'a cessé d'évoluer. Les techniques utilisées permettent de corriger les effets optiques et de la résine, mais aussi les effets de la gravure. Avec la réduction des dimensions critiques et l'utilisation de systèmes à grandes ouvertures numériques, les modèles utilisés en traitement OPC deviennent de plus en plus complexes pour permettre de capturer le maximum d'effets dans un temps raisonnable. La question de la stabilité de tels modèles purement empiriques devient alors d'actualité.

Sommaire

2.1	Introduction	37
2.2	Position du problème	37
2.2.1	Expérimentation	38
2.2.2	Résultats et discussions	39
2.3	État de l'art sur la stabilité des modèles empiriques	40
2.3.1	Méthode de la réponse en surface	41
2.3.2	Méthode des différences entre modèle et imagerie	41
2.3.3	Comparatif des différentes méthodes	42
2.4	Origines de l'instabilité des modèles empiriques	43
2.4.1	Étude qualitative	43
2.4.2	Comparatif des modèles	44
2.4.3	Origines de l'instabilité des modèles	44
2.4.3.1	Définition de la stabilité	44
2.4.3.2	Cas des paramètres du procédé	45
2.4.3.3	Cas des paramètres de modélisation	50
2.5	Conclusion	54

2.1 Introduction

La correction des effets de proximité optique (OPC) constitue une des techniques d'amélioration de la résolution (RET). L'OPC est un traitement du dessin du masque qui a pour objectif de corriger les dimensions des motifs dans la résine et de compenser les erreurs liées aux divers effets optiques et chimiques du procédé. Parmi les différentes façons de traitement OPC, l'une est basée sur la modélisation du procédé de photolithographie et permet de corriger le dessin du masque afin de minimiser les écarts entre les contours dans la résine et les dimensions visées. La simulation des contours dans la résine s'effectue grâce à un modèle OPC qui, comme nous l'avons vu dans le premier chapitre de ce mémoire, est composé de deux parties différentes mais complémentaires. La première, purement optique, est basée sur des postulats d'optique et permet de décrire les étapes de formation de l'image du circuit dans la résine. La seconde, appelée modèle empirique, permet la reconstitution des contours du circuit dans la résine à partir de l'intensité lumineuse de départ.

Contrairement à la partie optique du modèle qui est basée sur des lois physiques connues (Kirchhoff, Hopkins, Abbé, Fourier...), la modélisation de l'interaction de la lumière avec la résine est assez compliquée car elle nécessite d'appréhender les effets physico-chimiques (i.e. diffusion des espèces chimiques) des étapes d'exposition, de recuit et de développement. La difficulté de modéliser avec précision ces effets complexes en un temps de calcul raisonnable a fait obstacle à une utilisation dans un environnement de production à grand volume. C'est la raison pour laquelle des modèles empiriques ont été progressivement développés et sont disponibles dans les logiciels de traitement OPC. Les modèles empiriques sont basés sur des fonctions mathématiques à plusieurs paramètres pouvant être facilement ajustés à partir de mesures expérimentales de dimensions critiques.

Au cours du temps, la complexité des modèles empiriques n'a cessé d'augmenter. À l'heure actuelle, ces modèles comprennent de nombreux termes permettant de modéliser des effets divers tels que la densité environnante, la diffusion des espèces chimiques ou bien le filtrage du bruit de mesures. Toutefois, ils sont assez limités en précision car ils ne sont pas capables de prendre en compte la majorité des effets qui se produisent dans la résine. En outre, les approximations utilisées¹ dans la partie optique du modèle entraînent des erreurs résiduelles, difficiles à corriger par un modèle empirique. La réduction des dimensions des structures du masque et l'utilisation de systèmes à grandes ouvertures numériques nous ramènent à reconsidérer le domaine de validité des approximations de la partie optique. L'augmentation du nombre de termes dans un modèle empirique nous permet souvent de gagner en précision ; mais cela peut dégrader la stabilité du modèle obtenu car il devient dépendant du procédé et des conditions de calibration.

Dans le premier chapitre de ce mémoire, nous avons présenté le critère de la précision comme l'unique métrique permettant de définir la qualité d'un modèle. Nous verrons dans celui-ci que ce critère est insuffisant pour les nœuds technologiques avancés et qu'il est nécessaire de définir un nouveau critère, celui de la stabilité. Pour ce faire, nous présenterons un état de l'art des méthodes existantes pour l'analyse de la stabilité. Nous justifierons cette étude par une discussion des principales origines d'instabilité des modèles empiriques par rapport aux paramètres du procédé.

2.2 Position du problème

Depuis quelques années, les modèles empiriques utilisés pour prédire les contours des structures dans la résine étaient simples et ne comprenaient qu'un nombre limité de paramètres. Nous pouvons citer, par exemple, le modèle à seuil constant (voir définition page 26), et le modèle à convolution gaussienne (voir définition page 26). Malgré leur simplicité, ces modèles nous permettent d'obtenir une précision satisfaisante pour les nœuds technologiques passés. En revanche,

1. Les aberrations du système optique et les effets topographiques du masque sont négligées. L'arrondissement des coins des motifs n'est pas pris en compte...

ils s'avèrent limités pour modéliser correctement les contours des motifs dans la résine pour les technologies actuelles où divers effets tels que la diffusion chimique et la topographie du masque ne peuvent plus être négligés.

2.2.1 Expérimentation

Nous allons comparer les mesures expérimentales et les résultats de simulation obtenus avec deux modèles empiriques de complexités différentes. Pour le besoin de cette étude, nous avons préparé un masque de test comprenant des structures rondes de différents diamètres (voir figure 2.1) avec les deux tonalités possibles. Le masque utilisé est un PSM à 6% de transmission et à décalage de phase égal à 180 degrés. Nous avons ensuite exposé les différentes structures du masque dans la résine. La source d'illumination est de type Quasar (σ_{in} : 0.6, σ_{out} : 0.9, angle d'ouverture : 30 degrés) et l'ouverture numérique ON est égale à 1.05. Après exposition, nous avons mesuré les diamètres de ces structures dans la résine (voir figure 2.2).

Le choix des structures rondes peut être justifié par les critères suivants :

- La symétrie cylindrique des structures rondes favorise une diffusion isotrope des espèces chimiques dans les direction x et y, ce qui nous permet de négliger les effets à courte distance (causés par une densité locale non-uniforme à chaque bord de la structure) et ceux engendrés par les interactions entre les bords (causés par l'interférence des ondes diffractées par deux bords très proches).
- Le calcul mathématique de l'intensité de lumière est simple, ce qui nous permet de facilement vérifier les résultats de simulation par la partie optique du modèle OPC utilisé et de quantifier les erreurs liées aux approximations adoptées.

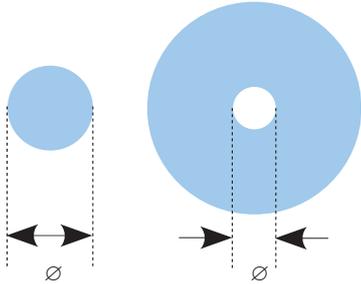


Figure 2.1 – Structures rondes suivant les deux tonalités possibles du masque : disque plein (à gauche) et disque creux (à droite). Les diamètres des structures varient de 100 nm à 1 μm (à l'échelle de la plaquette).

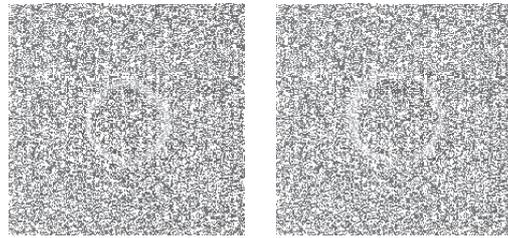


Figure 2.2 – Images d'un plot (à gauche) et d'un trou (à droite) de résine prises par un microscope électronique à balayage (SEMCD). Les diamètres moyens respectifs sont 197.6 nm et 206.6 nm (à l'échelle de la plaquette).

Afin de réduire le bruit des mesures, nous avons effectué la moyenne de trois mesures prises dans des champs différents de la plaquette. Ensuite, nous avons simulé les contours de ces structures dans la résine en utilisant deux modèles empiriques différents : le modèle à seuil constant (CTR) et celui à seuil variable (VT5). Le modèle CTR est un modèle simple permettant de reconstituer les contours des structures dans la résine à partir de l'intensité lumineuse initiale uniquement, c'est-à-dire, sans inclure les effets de diffusion des étapes de recuit et de développement. La comparaison directe des diamètres simulés par ce modèle et mesurés sur plaquette nous permet de quantifier ces effets. Le modèle VT5 est un modèle de complexité supérieure. Il permet d'adapter le seuil d'impression pour chaque bord du masque à partir de certaines propriétés du profil de l'intensité lumineuse (voir définition page 27). Bien qu'il ne gère pas directement le problème de la diffusion dans la résine, le modèle VT5 nous permet souvent d'améliorer la précision.

2.2.2 Résultats et discussions

Nous avons tracé sur la figure 2.3 les courbes des différences de diamètres simulés par ces deux modèles et mesurés sur plaquette en fonction des dimensions sur le masque (dessin). D'après cette figure, nous remarquons que plus le diamètre des structures sur le masque diminue, plus l'écart entre les diamètres simulés et mesurés augmente. Nous pouvons expliquer ces différences de diamètres par :

- L'apparition d'effets physico-chimiques complexes dans la résine au niveau des structures de petites dimensions. Ces effets sont de deux types :
 - o Des effets chimiques tels que la diffusion des espèces acides et basiques dans la résine ainsi que le flare chimique [33,34].
 - o Des effets optiques tels que les effets à moyenne et à grande distance [30–32] causés par une distribution non-uniforme de la lumière sur la plaquette.
- Les erreurs d'écriture du masque [61], c'est-à-dire, les différences de dimensions critiques des motifs du masque avant et après sa fabrication² (voir figure 1.16).
- Les erreurs de simulation dues aux approximations effectuées pour réduire le temps de calcul :
 - o Utilisation de l'incidence normale pour les sources hors-axes.
 - o Effets de la topographie au niveau du masque (voir chapitre 4, page 77) négligés et non pris en compte.
 - o Lentilles parfaites, pas d'aberrations optiques, etc.

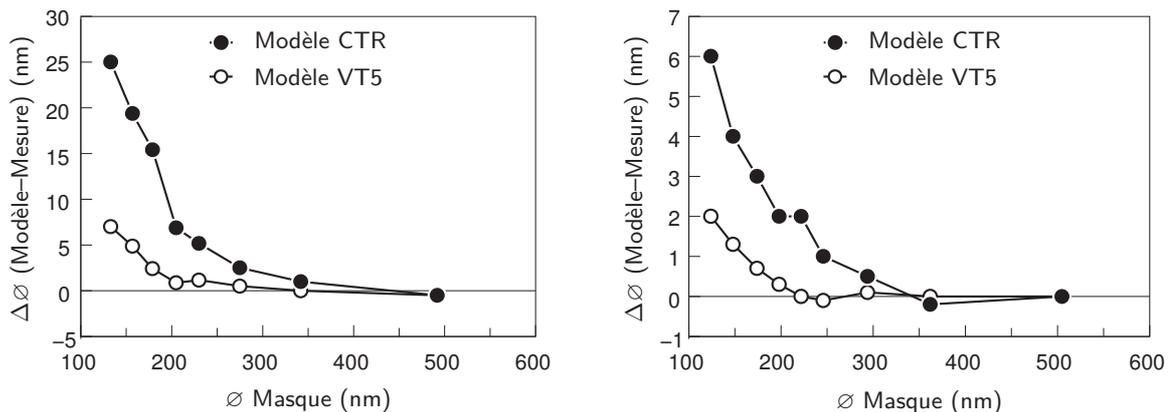


Figure 2.3 – Différences entre les diamètres simulés par un modèle à seuil constant et un second à seuil variable et ceux mesurés sur plaquette pour les plots (à gauche) et trous (à droite) de résine en fonction des dimensions dessinées sur le masque (à l'échelle de la plaquette). Les différences de diamètres suivent une loi exponentielle décroissante du type $y(x) = Ae^{-\frac{x}{D}}$. Ceci montre que plus le diamètre de la structure augmente, plus l'écart entre la mesure et la simulation diminue. De plus, ces écarts sont plus élevés pour le modèle basique à seuil constant que pour celui à seuil variable.

De plus, on constate que, pour un même diamètre sur le masque, l'écart de diamètres mesurés et simulés pour les plots de résine est plus important que celui obtenu pour les trous. Ceci s'explique par des effets de diffusion chimique plus importants dans le cas des plots que pour les trous. Si l'on compare les profils de l'intensité lumineuse initiale dans la résine pour un plot et un trou de même diamètre, nous pouvons noter une baisse significative de contraste pour les trous de faibles dimensions contrairement aux plots de résine (voir figure 2.4). Notons ici qu'un faible contraste traduit des gradients chimiques plus faibles lors du recuit après exposition (PEB), c'est-à-dire, une diffusion chimique moins importante. La baisse de contraste pour les trous de petits diamètres s'explique par un faible rapport entre l'intensité de l'ordre 0 et celles des ordres diffractés ± 1 .

2. Ce sont les différences entre les dimensions au niveau du dessin du masque que nous utilisons dans la modélisation et les dimensions réelles des motifs gravés sur le masque (à l'échelle de la plaquette).

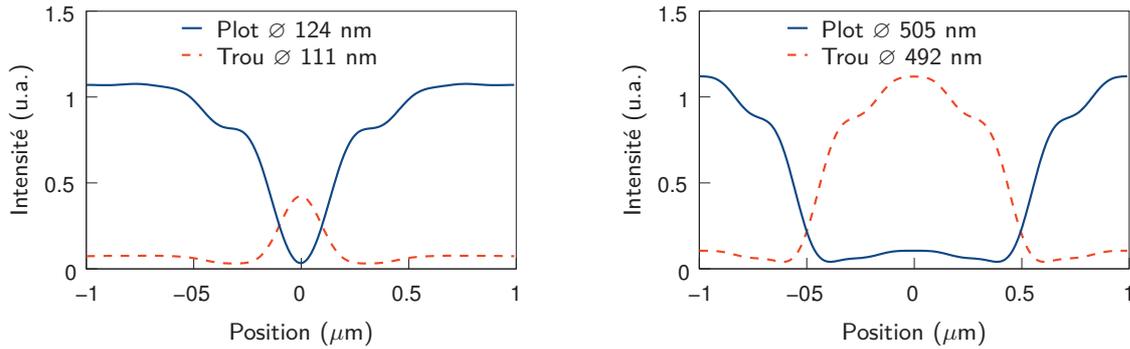


Figure 2.4 – Profils simulés de l'intensité lumineuse initiale dans la résine pour des structures rondes de diamètres avoisinant 100 (à gauche) et 500 nm (à droite). Ces profils ont été calculés pour des plots et trous de résine respectivement avec des diamètres légèrement modifiés par l'outil de traitement OPC. Contrairement aux profils calculés pour les structures de grands diamètres (à droite), ceux des structures de faibles diamètres (à gauche) présentent des contrastes très différents. La baisse de contraste est nettement visible pour les trous de résine de diamètre égal à 111 nm.

Les résultats obtenus sont valables pour les deux modèles utilisés. Par contre, la figure 2.3 nous montre que l'utilisation d'un modèle de complexité supérieure (par exemple le modèle à seuil variable) nous permet de réduire les erreurs de la simulation par rapport aux mesures expérimentales, en particulier pour les structures de petites dimensions. Par exemple, l'écart entre les diamètres simulés et mesurés obtenu pour un plot d'environ 100 nm de diamètre atteint 25 nm dans le cas du modèle à seuil constant, alors qu'il est égal à 8 nm dans le cas du modèle à seuil variable. Ceci nous amène à dire qu'un modèle empirique simplifié n'est pas assez prédictif pour les structures de petites dimensions. Il est donc limité pour les derniers nœuds technologiques (45 nm et au delà). Le besoin d'utiliser des modèles avancés pour la modélisation des structures de petites dimensions nous semble évident. Toutefois, on a besoin de s'assurer que ces modèles avancés sont assez stables vis-à-vis des paramètres utilisés afin d'appréhender les éventuelles dérives du procédé photolithographique.

La stabilité d'un modèle empirique est son aptitude à maintenir un comportement similaire lorsque la valeur d'un paramètre est modifiée. Pour mieux comprendre ce concept, nous allons présenter un état de l'art des méthodes utilisées dans ce contexte.

2.3 État de l'art sur la stabilité des modèles empiriques

L'étude de la stabilité des modèles empiriques est une idée nouvelle en traitement OPC. Cette technique est devenue nécessaire ces dernières années, lorsque les modèles empiriques utilisés sont devenus complexes. Dans la littérature, les premiers objectifs des méthodes proposées concernent la compréhension de l'influence des paramètres du modèle utilisé sur sa réponse finale (dimensions critiques) dans le but de le simplifier en négligeant les termes les moins significatifs. Ensuite, d'autres méthodes ont été développées dans le but de vérifier, par l'intermédiaire de simulations par un modèle empirique, la sensibilité des motifs à imprimer par rapport aux variations des paramètres de procédé, notamment le décalage en focus et la dose d'exposition. Outre la vérification, le besoin de détecter les structures sensibles aux variations des paramètres de procédé a permis de développer d'autres méthodes permettant d'améliorer la couverture du modèle et sa régression.

Dans cette partie, nous décrivons les méthodes développées pour qualifier la stabilité : la méthodologie de la réponse en surface RSM et la méthode des différences entre modèle et imagerie. Nous décrivons également l'approche d'un facteur à la fois OFAT, la méthode de la sensibilité aux paramètres, le concept des modèles MSM et la méthode des motifs sensibles à l'imagerie, en annexe B de ce mémoire.

2.3.1 Méthode de la réponse en surface

La méthode de la réponse en surface RSM (de l'anglais "*Response Surface Methodology*") est une approche méthodologique utilisée pour analyser l'influence d'un ensemble de facteurs indépendants en entrée sur une réponse en sortie. Elle permet, par des analyses statistiques sur des observations, de trouver une relation de corrélation entre la réponse du modèle et les paramètres utilisés comme décrit par l'équation suivante :

$$y = f(x_1, x_2, \dots, x_k), \quad (2.1)$$

où y est la réponse à exprimer, $\{x_i\}_{1 \leq i \leq k}$ est l'ensemble des facteurs indépendants et f est la fonction de corrélation.

Cette méthode a été employée par D.S. Shon et al. [62] afin de corrélérer les dimensions critiques CD avec les paramètres d'exposition de Dill : A, B et C³. La fonction de corrélation obtenue par Shon est de forme polynomiale de second ordre et s'exprime de la manière suivante :

$$CD = \beta_0 + \beta_1 A + \beta_2 B + \beta_3 C + \beta_{11} A^2 + \beta_{22} B^2 + \beta_{33} C^2 + \beta_{12} AB + \beta_{23} BC + \beta_{13} AC$$

L'étude de la stabilité est basée sur la quantification de la sensibilité de la réponse CD par rapport aux paramètres A, B et C. Cette quantification est simplement basée sur l'identification des termes ayant les coefficients les plus élevés dans la fonction polynomiale de corrélation. D'une manière similaire, la technique RSM a été utilisée par S.K. Kim [63] et al. pour étudier l'impact des paramètres de modélisation et de procédé sur les dimensions critiques. Dans ses travaux de recherche, Kim a construit un modèle de second ordre reliant le CD à onze paramètres de modélisation et de procédé. L'un des résultats importants obtenus par le biais de cette technique est que le temps et la température de recuit après exposition (PEB) peuvent avoir une forte influence sur les dimensions critiques.

Bien que cette technique permette d'appréhender le problème de la stabilité d'une manière simple, son domaine d'application se limite à des paramètres indépendants (elle nous permet d'identifier, d'une manière directe, les paramètres ayant une forte influence sur la réponse du modèle). Cette technique ne peut donc pas être utilisée dans le cas de la plupart des paramètres du procédé lithographique qui peuvent interagir entre-eux⁴.

2.3.2 Méthode des différences entre modèle et imagerie

Une autre méthode intéressante pour l'étude de la stabilité des modèles empiriques est celle qui est actuellement implémentée dans le logiciel de traitement OPC, Calibre™, de Mentor Graphics. Cette technique consiste à évaluer l'écart entre l'image latente calculée par le modèle empirique et l'intensité lumineuse de départ calculée par la partie optique. La formule suivante permet d'exprimer cet écart :

$$S = \int_{site} (I_{\text{image aérienne}} - I_{\text{image latente}})^2 \quad (2.2)$$

La figure 2.5 présente les courbes de l'intensité lumineuse dans la résine et de l'image latente calculées le long d'un site de simulation. L'aire de la zone située entre les deux courbes correspond à la valeur de S pour ce site. Sur l'exemple de la figure 2.5, nous pouvons clairement remarquer qu'un maximum local apparaît au niveau de l'image latente contrairement à ce que l'on observe sur l'intensité lumineuse dans la résine. Ce maximum local devient critique lorsque l'image latente à ce point se trouve au dessus du seuil d'impression, provoquant ainsi l'apparition d'un trou de résine. Il peut donc correspondre à un échec du modèle empirique à modéliser l'image

3. voir [46] pour la signification de ces paramètres.

4. Nous pouvons citer par exemple le cas du focus qui dépend de l'ouverture numérique, des coefficients de cohérence partielle ou alors des caractéristiques géométriques (épaisseur) et optiques (indices d'extinction et de réfraction) de la résine

latente de la structure. Par conséquent, le modèle peut être considéré instable pour ce type de structure. Dans un prochain paragraphe de ce chapitre, nous présenterons une étude qualitative afin d'expliquer l'apparition de ce maximum local au niveau de l'image latente.

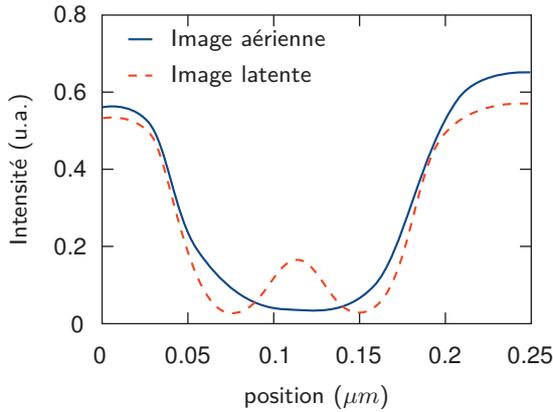


Figure 2.5 – Schéma illustrant la méthode implémentée dans le logiciel de traitement OPC, Calibre™, de Mentor Graphics pour l'évaluation de la stabilité d'un modèle empirique. Le cas ici est celui d'une ligne appartenant à un circuit arbitraire du noeud technologique 45 nm. L'intensité de l'image aérienne est calculée avec une source de type Quasar^a ($\sigma_{in} : 0.6$, $\sigma_{out} : 0.9$, angle d'ouverture : 30 degrés), une ouverture numérique ON égale à 1.05 et un masque PSM avec 6% de transmission et un décalage de phase de 180 degrés. L'image latente est obtenue avec un modèle empirique de type CM1.

a. voir définition page 10

Contrairement aux autres méthodes présentées précédemment où l'on parle de sensibilité d'un modèle empirique aux variations d'un paramètre donné, celle-ci nous permet d'aborder ce problème indépendamment du paramètre choisi. Elle permet essentiellement de détecter les endroits où le modèle diverge de façon inattendue par rapport à l'intensité lumineuse de départ. Les différences sont ensuite examinées et une tolérance est appliquée pour extraire les cas les plus critiques.

2.3.3 Comparatif des différentes méthodes

Ainsi, nous avons présenté différentes méthodes de la littérature pour l'étude de la stabilité d'un modèle empirique. Le tableau 2.1 présente un récapitulatif des propriétés de ces méthodes et les possibilités de les appliquer dans notre étude. D'après ce tableau, les trois premières méthodes sont basées sur les différences des dimensions critiques des motifs alors que les autres sont plutôt basées sur les différences d'intensité. L'inconvénient de ces dernières est que la sensibilité est calculée en moyennant les différences d'intensité au voisinage du site (soit sur une ligne ou dans un cercle) ; ce qui peut fausser l'évaluation de la sensibilité. Ceci nous permet d'éliminer les trois dernières méthodes.

	Méthode basée sur les variations	Permet d'évaluer la sensibilité	Sensibilité évaluée par type de paramètre	Stabilité du modèle évaluée directement
RSM	des dimensions	du procédé	Oui	Non
OFAT	des dimensions	du procédé	Oui	Non
PS	des dimensions	du procédé	Oui	Non
MSM	de l'intensité	du procédé	Oui	Non
OSP	de l'intensité	du procédé	Oui	Non
Mentor	de l'intensité	du modèle	Non	Oui

Tableau 2.1 – Comparatif des méthodes présentées dans la littérature pour l'étude de la stabilité d'un modèle empirique.

Nous avons également éliminé la méthode de réponse en surface RSM car elle ne s'applique qu'à des paramètres indépendants. Nous avons donc gardé les méthodes OFAT et PS dont le principe est similaire. Cependant, d'après le tableau 2.1, la seule méthode qui nous permet d'évaluer la stabilité d'un modèle empirique est celle de Mentor Graphics. Les autres méthodes nous permettent seulement d'étudier la sensibilité du procédé. Nous avons donc eu l'idée de réunir son principe avec celui des méthodes OFAT et PS afin de les adapter à notre étude. Sur cette base, nous avons développé une méthode perturbative que nous présenterons dans le chapitre 3 de ce mémoire.

2.4 Origines de l'instabilité des modèles empiriques

Dans le premier chapitre de ce mémoire, nous avons brièvement présenté un état de l'art des modèles empiriques basés sur des lois physiques. Ces derniers, bien qu'ils soient précis pour la modélisation des effets de la résine, sont très coûteux en temps de calcul. C'est pour cette raison que d'autres types de modèles empiriques et simplifiés ont été mis en place. Dans la suite de ce chapitre, nous effectuerons brièvement une analyse qualitative des modèles empiriques qui existent actuellement dans les logiciels de traitement OPC. Nous donnerons ensuite une définition de la stabilité. Finalement, nous présenterons les principales causes d'instabilité des modèles étudiés.

2.4.1 Étude qualitative

Dans le cas du modèle à seuil constant, l'extraction des contours s'effectue à partir de l'intensité lumineuse initiale dans la résine. Ce modèle n'apporte donc aucune modification à l'intensité de départ dont le calcul est basé sur des lois physiques connues et vérifiées. Le modèle CTR est donc numériquement stable. C'est pour cette raison que nous l'avons choisi comme modèle de référence pour l'étude de la stabilité.

En ce qui concerne le modèle à diffusion gaussienne, la stabilité de tels modèles dépend des écarts types des fonctions gaussiennes et des coefficients de pondération. En effet, il existe des cas où la somme pondérée de deux ou plusieurs gaussiennes présente des minima d'intensité pouvant fausser l'extraction des contours lorsque l'on applique un seuil constant au résultat de la convolution. La figure 2.6 montre que la somme pondérée de deux fonctions gaussiennes peut donner lieu à des minima locaux dans le cas où les coefficients de pondération sont de signes opposés. Comme le montre le schéma de la figure 2.7, la convolution de l'intensité lumineuse initiale avec la somme des gaussiennes fait apparaître un maximum local d'intensité. Dans ce cas particulier, la forme des contours dépend de la position du seuil par rapport au maximum local formé.

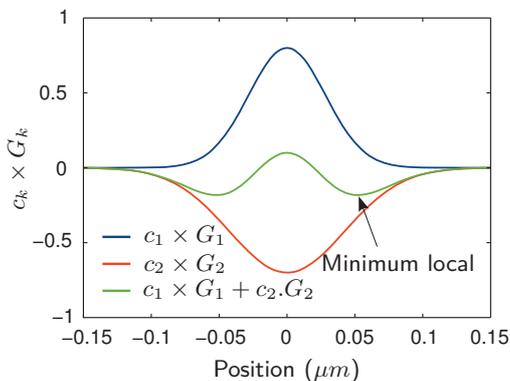


Figure 2.6 – Somme pondérée de deux fonctions gaussiennes d'écarts types différents. Les coefficients de pondération sont de signes opposés.

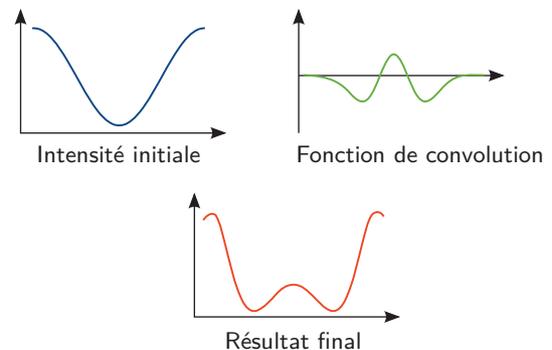


Figure 2.7 – Convolution de l'intensité lumineuse initiale par une fonction présentant deux minima locaux. Le résultat de convolution présente un maximum local d'intensité.

Quant au modèle à seuil variable (VT5), le seuil est une fonction polynomiale des paramètres I_r , I_{max} , I_{min} , la pente Slope et la courbure Factor. Ce modèle est une généralisation de celui à seuil constant avec l'existence d'autres termes pondérés avec des coefficients c_k . La stabilité du modèle VT5 dépend donc des valeurs des coefficients utilisés ainsi que de la nature physique des termes.

Finalement, la stabilité des modèles à convolution étendue et à opérateur différentiel n'est pas évidente puisqu'elle dépend des écarts types des fonctions gaussiennes utilisées, des signes des coefficients de pondération et des positions des seuils b_0 , b_n et b_m dans le cas du modèle à convolution étendue et $\pm b$ dans le cas du modèle à opérateur différentiel.

2.4.2 Comparatif des modèles

Dans l'objectif de comparer les modèles de cette étude⁵, nous avons récapitulé dans le tableau 2.2 les propriétés de chacun d'entre-eux : le nombre des paramètres utilisés, les nombres minimal et maximal de termes, la nature du modèle, le type de la simulation et finalement une description de la stabilité.

	CTR	VT5	Brion	CM1
Nombre de paramètres ^a	1	5	14	6
Nombre minimal de termes ^b	1	1	3	1
Nombre maximal de termes	1	15	13	∞
Nature	optique pur	optique et empirique	optique et empirique	optique et empirique
Simulation sur	sites ou grille	sites	grille	grille
Stabilité numérique	stable	inconnue	inconnue	inconnue

a. Les coefficients de pondération ne sont pas comptés parmi les paramètres du modèle.

b. L'équation du modèle se traduit par une somme pondérée de termes empiriques qui s'écrivent à leur tour en fonction des paramètres.

Tableau 2.2 – Récapitulatif des modèles empiriques utilisés dans l'étude de la stabilité.

Dans la suite de ce chapitre, nous allons examiner les causes d'instabilité d'un modèle empirique vis-à-vis des paramètres du procédé lithographique. Nous nous basons sur le modèle à seuil constant pour établir une définition de la stabilité. Ensuite, nous nous intéressons aux modèles VT5, Brion et CM1 qui peuvent être instables.

2.4.3 Origines de l'instabilité des modèles

2.4.3.1 Définition de la stabilité

Un modèle est stable si sa sensibilité aux variations d'un paramètre du procédé lithographique est équivalente à celle du modèle à seuil constant. Cette définition de la stabilité est basée sur le fait que le modèle à seuil constant est naturellement stable. Nous exprimons la sensibilité aux variations d'un paramètre P de la façon suivante :

$$S = \frac{\partial CD}{\partial P}, \quad (2.3)$$

∂CD étant les variations de CD et ∂P celles du paramètre P . Nous comparons ainsi analytiquement la sensibilité du modèle empirique étudié à celle de l'optique pur qui est basé sur des postulats physiques vérifiés et sont par conséquent numériquement stables. Bien que la partie optique du modèle inclue différentes approximations comme celle de Kirchhoff, la modélisation optique se fait de la même manière si l'on fait varier un paramètre donné; ce qui lui procure une certaine stabilité numérique. Nous ne cherchons donc pas à étudier la stabilité du procédé par une comparaison directe de la réponse du modèle avec les mesures expérimentales mais la stabilité numérique du modèle, c'est-à-dire, sa capacité à extrapoler au delà de son domaine d'apprentissage.

Un modèle empirique peut avoir différents comportements à estimer les dimensions des motifs lorsque l'on fait varier un paramètre donné. Nous avons reporté sur la figure 2.8 quatre cas de modèles traduisant les combinaisons possibles des états de précision et de stabilité d'un modèle. Sur cette figure, on peut clairement noter que le modèle du cas a) est stable selon notre définition puisque la sensibilité du modèle étudié est équivalente à celle du modèle à seuil constant CTR et précis car les résultats de simulation (CD en fonction des variations du paramètre ΔP) coïncident avec les mesures expérimentales. Ce cas particulier ne peut exister que lorsque la sensibilité du modèle CTR est équivalente à celle des mesures; ce qui n'est pas toujours vrai car les effets

5. À l'exception du modèle à convolution gaussienne

chimiques dans la résine peuvent varier en fonction du paramètre. Dans le cas où les sensibilités du modèle CTR et des mesures expérimentales sont différentes, le comportement du modèle empirique peut se présenter sous trois formes : le cas b) est celui d'un modèle précis et instable, c) est celui d'un modèle stable et imprécis et finalement d) est le cas d'un modèle instable et imprécis.

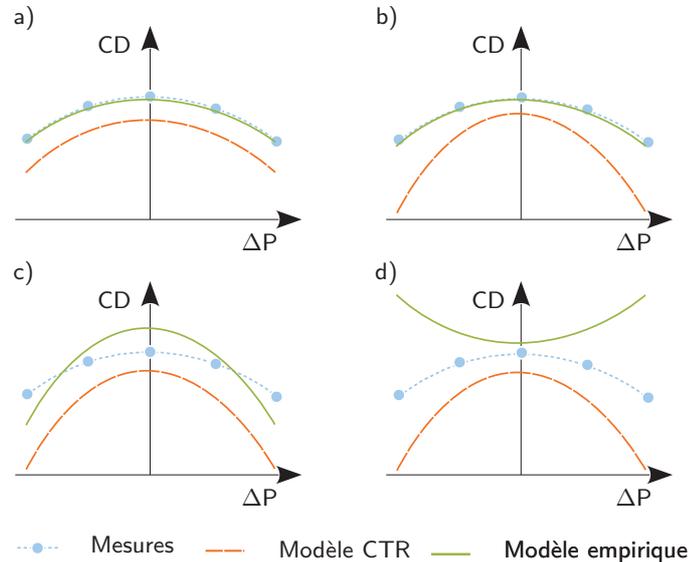


Figure 2.8 – Courbes de CD en fonction des variations d'un paramètre donné illustrant différents comportements d'un modèle empirique pour une structure donnée : a) modèle stable et précis, b) modèle instable mais précis, c) modèle imprécis mais stable et finalement d) modèle imprécis et instable.

2.4.3.2 Cas des paramètres du procédé

Les paramètres du procédé sont des grandeurs physiques définissant les variables du procédé lithographique. Ils sont utilisés dans la modélisation de l'image aérienne (partie optique) et des contours de la résine (modèle empirique) sous forme de constantes telles que les épaisseurs de l'empilement de la plaque ou alors des variables telles que la position du plan focal. Les paramètres du procédé sont mesurables. Ils sont donc des paramètres physiques. Notons également que les paramètres du procédé interviennent dans la partie optique du modèle qui permet de calculer l'intensité lumineuse de départ, c'est-à-dire, sans tenir compte des effets dans la résine. Du fait que la partie résine du modèle dépend de l'intensité lumineuse calculée par la partie optique, il en résulte qu'une variation dans l'un des paramètres de la partie optique peut altérer la stabilité du modèle empirique.

La stabilité d'un modèle empirique par rapport à un paramètre physique dépend des mesures de calibration aussi bien en terme de quantité qu'en terme de qualité. En effet, les mesures collectées pour différentes structures de calibration dans une ou plusieurs conditions de procédé représentent des contraintes qui orientent la régression du modèle et modifient son comportement. Un nombre plus élevé des mesures de calibration permet donc de mieux répartir les effets des termes empiriques du modèle. En terme de qualité, les informations apportées par les nombreuses mesures collectées doivent être non-redondantes, c'est-à-dire, qu'elles doivent être uniformément réparties pour que la régression du modèle ne soit pas statistiquement biaisée. En revanche, le bruit des mesures ne représente pas vraiment un facteur d'instabilité tant qu'il est uniformément réparti. Finalement, les mesures doivent être également effectuées dans des conditions décalées du paramètre.

2.4.3.2.1 Nombre de motifs de calibration

Il est intéressant d'étudier la stabilité des modèles empiriques par rapport à la taille de l'ensemble formé par des structures de calibration. Pour ce faire, nous avons ajusté plusieurs

modèles empiriques en faisant varier le nombre de mesures utilisées pendant la calibration. Une liste non-exhaustive de structures utilisées dans la calibration des modèles est donnée par la figure 2.9.

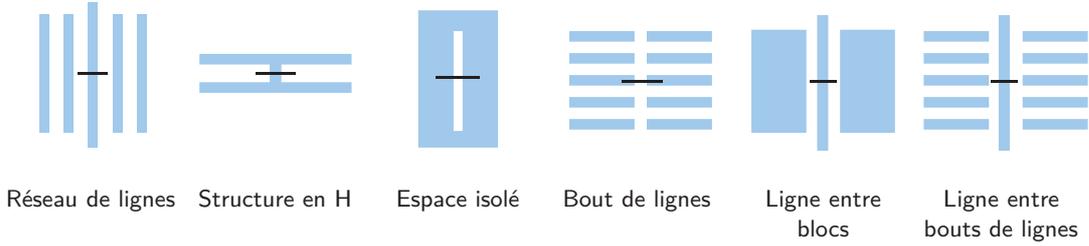


Figure 2.9 – Quelques structures utilisées pour la calibration d'un modèle empirique

Les modèles de cette étude sont à seuil variable dont la fonction caractéristique est donnée par l'équation :

$$\begin{aligned}
 I_s = & c_0 + c_1 I_{\max} + c_2 I_{\max}^2 + c_3 I_{\min} + c_4 I_{\min}^2 + c_5 \text{Slope} + c_6 \text{Slope}^2 + c_7 \text{Factor} \\
 & + c_8 \text{Factor}^2 + c_9 I_{\max} \cdot I_{\min} + c_{10} I_{\max} \cdot \text{Slope} + c_{11} I_{\max} \cdot \text{Factor} + c_{12} I_{\min} \cdot \text{Slope} \\
 & + c_{13} I_{\min} \cdot \text{Factor} + c_{14} \text{Slope} \cdot \text{Factor}
 \end{aligned} \quad (2.4)$$

où I_{\max} , I_{\min} , Slope et Factor sont les paramètres du modèle, et c_k les coefficients à ajuster. Notons ici que c_0 désigne le seuil d'extraction de l'intensité lumineuse initiale dans la résine (modèle optique pur), alors que les autres coefficients traduisent les poids attribués aux termes du modèle et permettent d'équilibrer leurs effets. La figure 2.10 décrit les valeurs des coefficients c_k de chaque modèle en fonction du nombre de motifs de calibration. D'après cette figure, on remarque que les coefficients c_k , $k \neq 0$, sont assez importants par rapport au coefficient c_0 (coefficient du modèle optique pur) lorsque le nombre des mesures de calibration est petit (c'est-à-dire inférieur à 100). En augmentant le nombre des mesures de calibration, ces coefficients tendent vers zéro et le modèle empirique se rapproche de plus en plus à un modèle à seuil constant qui est stable. Nous pouvons donc conclure qu'un modèle à seuil variable est d'autant plus stable que le nombre des mesures de calibration est élevé.

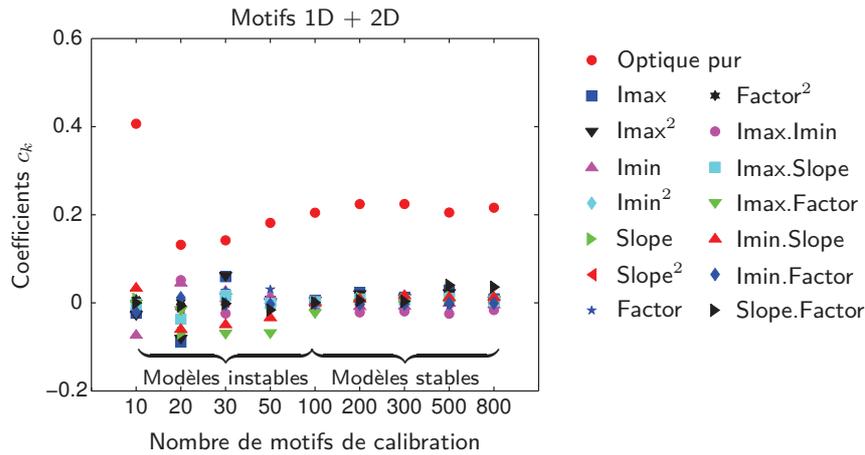


Figure 2.10 – Coefficients c_k de modèles empiriques à seuil variable VT5 calibrés en utilisant différents nombres n de mesures réalisées pour différents types de structures. En augmentant le nombre de mesures expérimentales de calibration, le modèle VT5 tend vers un modèle à seuil constant CTR.

2.4.3.2.2 Type des motifs de calibration

Étudions maintenant la stabilité du modèle à seuil variable par rapport au type des motifs utilisés en calibration. Pour ce faire, nous avons créé plusieurs modèles à seuil variable de forme identique en n'utilisant que des structures 1D ou 2D uniquement. Nous avons également fait varier le nombre des mesures utilisées pour chacun des deux cas. La figure 2.11 présente les

coefficients c_k des modèles obtenus en fonction du nombre n des motifs de calibration, calculés pour les structures 1D et 2D respectivement. D'après cette figure, l'augmentation du nombre des motifs 1D utilisés en calibration entraîne une meilleure stabilité du modèle obtenu. De plus, il est important d'avoir un nombre minimal de motifs 1D pour qu'un modèle à seuil variable soit stable. Dans notre cas, le modèle se stabilise à partir d'un nombre de motifs 1D égal à 100. En ce qui concerne les motifs 2D, la figure montre qu'il est difficile d'obtenir un modèle entièrement stable en utilisant uniquement ce type de structures. Dans ce cas particulier, un nombre de motifs de calibration inférieur à 200 est insuffisant pour avoir un modèle stable. Si nous comparons ces résultats avec ceux trouvés dans le cas des motifs 1D, nous pouvons en déduire que les structures 2D apportent moins d'informations utiles pour la calibration d'un modèle à seuil variable. Nous expliquons ce résultat par :

- Une meilleure couverture de l'espace de calibration pour les structures 1D contrairement à celles 2D.
- Les informations (Imax, Imin, Slope, Factor) extraites des profils d'intensité des structures 1D sont plus homogènes et possèdent un comportement continu permettant d'aligner la réponse du modèle.
- Le bruit de mesures des structures 2D est plus élevé que celui des motifs 1D.

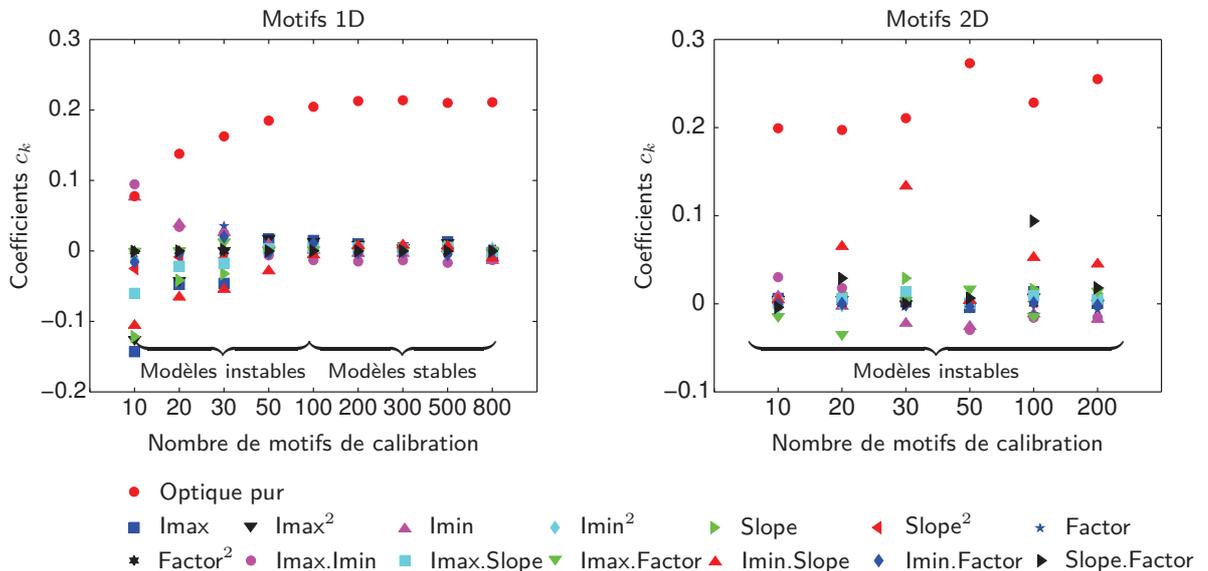


Figure 2.11 – Évolution des coefficients c_k de modèles empiriques à seuil variable calibrés en fonction du nombre de mesures effectuées a) pour des structures 1D uniquement b) pour des structures 2D uniquement.

Le résultat de la figure 2.11 pour les motifs 1D s'explique par le fait que la régression du modèle est étroitement liée aux informations apportées par les profils d'intensité calculés sur les sites mesurés. Nous avons tracé sur la figure 2.12 les seuils simulés pour 125 motifs⁶ 1D par un modèle à seuil variable "stable" en fonction de Imax. Ensuite, nous avons extrait un sous-ensemble comportant 20 motifs. Le but ici est de montrer que les seuils empiriques⁷ recueillis pour les 20 motifs ne sont pas suffisants pour garantir la stabilité numérique du modèle. Pour ce faire, nous avons appliqué une régression polynômiale de second ordre aux seuils empiriques en utilisant la totalité des 125 mesures et le sous-ensemble formé par les 20 mesures. La figure 2.12 nous montre que la courbe de régression s_2 obtenue pour seulement 20 mesures diverge par rapport à s_3 qui correspond à la régression du seuil simulé par le modèle VT5. En revanche, celle qui est obtenue pour 125 mesures, s_1 , reste similaire au seuil simulé du modèle VT5. Certes, l'utilisation d'une régression linéaire de premier ordre avec les 20 mesures nous permet de nous rapprocher de la réponse finale du modèle ; mais cette manoeuvre modifie la forme initiale du

6. Les 125 mesures ont été choisies parmi le nombre total des mesures (environ 1000) ayant servi pour calibrer le modèle et sont telles que les paramètres Slope, Imin et Factor restent à peu près constants.

7. Les seuils empiriques sont les seuils qui correspondent exactement aux mesures expérimentales.

modèle en éliminant le terme I_{\max}^2 . Finalement, en observant les forts coefficients c_k des termes I_{\max} et I_{\max}^2 dans les équations de s_2 , comparés à ceux de s_1 , nous pouvons affirmer que le modèle de régression s_2 est instable.

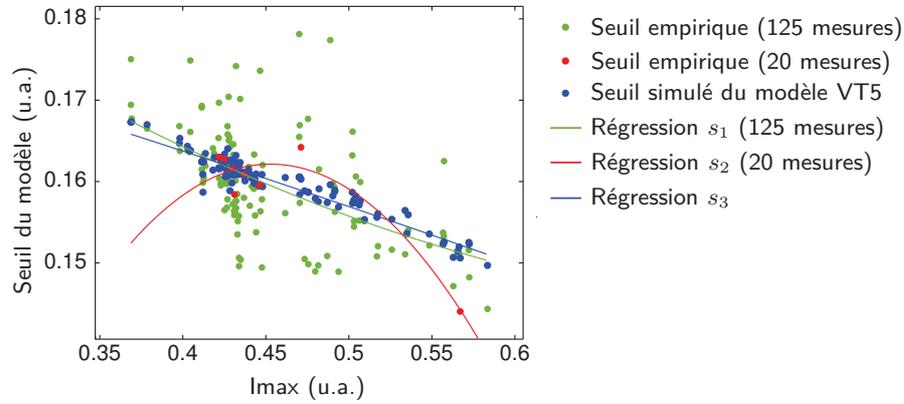


Figure 2.12 – Seuils d'extraction des contours en fonction de I_{\max} (Slope, I_{\min} et Factor étant fixés) pour des structures 1D. Les seuils présentés sont ceux qui sont calculés par rapport aux mesures (seuils empiriques) et ceux qui sont simulés par un modèle à seuil variable. Les fonctions polynomiales de régression sont $s_1 = 0.1096I_{\max}^2 - 0.1840I_{\max} + 0.2204$ et $s_2 = -1.3726I_{\max}^2 + 1.2431I_{\max} - 0.1193$ pour les seuils empiriques avec 125 et 20 mesures respectivement et $s_3 = -0.0130I_{\max}^2 - 0.0565I_{\max} + 0.1884$ pour celui simulé par le modèle VT5.

2.4.3.2.3 Impact du bruit de mesures

Nous allons maintenant étudier l'impact du bruit de mesures sur la stabilité d'un modèle à seuil variable. Pour ce faire, nous avons créé plusieurs modèles VT5 en utilisant des mesures obtenues en additionnant un bruit gaussien⁸ (voir figure 2.13) aux valeurs initiales. Dans le but d'analyser l'impact d'un bruit gaussien sur la stabilité du modèle à seuil variable, nous avons fait varier l'écart type σ et l'espérance μ du bruit de 0.5 à 3 nm et de 1 à 3 nm respectivement. Nous avons ensuite ajusté les coefficients c_k du modèle.

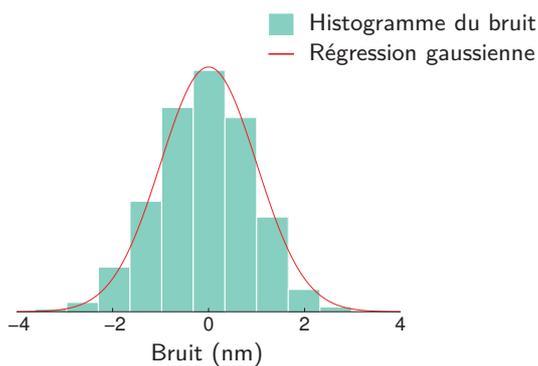


Figure 2.13 – Histogramme d'un bruit gaussien de moyenne nulle et d'écart type égal à 1 nm

La figure 2.14 montre l'évolution des coefficients c_k des modèles obtenus en fonction de l'écart type du bruit injecté aux mesures. Nous pouvons remarquer que les modèles obtenus sont pratiquement similaires et que les coefficients c_k ne varient que très peu lorsque l'écart type du bruit est inférieur à 5 nm. Les coefficients c_k obtenus avec les fortes valeurs d'écart-type (10 et 20 nm) sont légèrement modifiés et les modèles correspondants deviennent peu à peu instables. Néanmoins, les perturbations que présentent les coefficients c_k sont suffisamment

8. La densité de probabilité du bruit créé suit une loi normale. La loi normale (également appelée la loi normale gaussienne ou encore la loi de Laplace-Gauss) d'espérance μ et d'écart type σ est une loi de probabilité telle que la densité de probabilité f s'écrit : $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

légères pour que nous puissions les négliger. Par conséquent, nous avons considérés ces modèles comme stables. Nous montrons ainsi que l'effet d'un bruit de mesures de moyenne nulle sur la stabilité d'un modèle à seuil variable est faible.

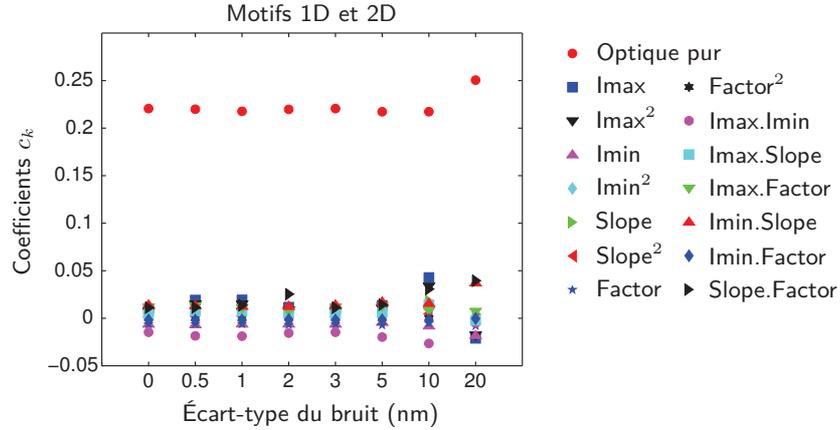


Figure 2.14 – Tracés des coefficients c_k d'un modèle à seuil variable VT5 en fonction de l'écart type du bruit que nous avons rajouté aux mesures expérimentales de calibration, la moyenne du bruit étant nulle.

D'une manière similaire, nous avons tracé sur la figure 2.15 les coefficients des modèles à seuil variable en fonction de la moyenne du bruit de mesures d'écart-type constant (égal à 0.5nm). Nous pouvons remarquer que les coefficients c_k varient très peu. La modification de la moyenne du bruit à quelques nanomètres près ne modifie donc pas la stabilité du modèle. En effet, cette modification oblige le modèle à adapter le coefficient c_0 (coefficient du modèle optique pur) afin de compenser l'erreur en CD causée par une moyenne non nulle du bruit. Il en résulte que l'effet du bruit de mesures à moyenne non nulle sur la stabilité d'un modèle empirique à seuil variable est faible.

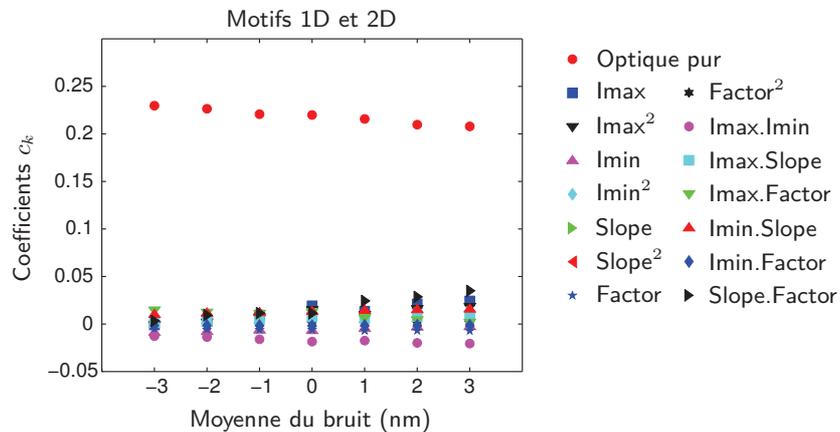


Figure 2.15 – Tracés des coefficients c_k d'un modèle à seuil variable VT5 en fonction de la moyenne du bruit que nous avons rajouté aux mesures expérimentales de calibration, l'écart-type du bruit étant égal à 0.5 nm.

Nous pouvons expliquer les résultats obtenus dans le cas d'un modèle à seuil variable par l'existence d'un paramètre de simulation appelé "critère de régression". Ce dernier permet au modèle de filtrer le bruit de mesures. Il permet d'augmenter ou diminuer la régression du modèle aux mesures expérimentales. Une meilleure régression du modèle est obtenue lorsque l'on fait tendre la valeur de ce paramètre vers les faibles valeurs ($<10^{-4}$). Dans ce cas, l'erreur quadratique moyenne (RMS) obtenue est très faible et la précision du modèle est améliorée. La figure 2.16 présente les courbes de l'erreur quadratique moyenne du modèle et la somme des coefficient c_k , tels que $k \neq 0$, en fonction du critère de régression. D'après cette figure, nous pouvons différencier trois cas possibles :

- Cas où le critère de régression est inférieur à 10^{-4} : le modèle obtenu est précis mais pas assez stable.
- Cas où le critère de régression est supérieur à 10^{-2} : le modèle obtenu est stable mais pas assez précis.
- Cas où le critère de régression est proche de 10^{-3} : le modèle obtenu est à la fois stable et précis.

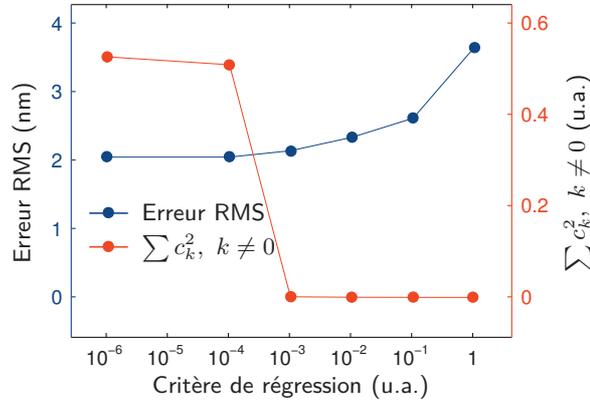


Figure 2.16 – Évolution de la stabilité et la précision d'un modèle à seuil variable en fonction du critère de régression.

En conclusion, l'erreur quadratique moyenne (RMS) ne constitue plus l'unique métrique pour accepter ou rejeter un modèle donné. L'analyse de stabilité du modèle s'avère importante si l'on veut obtenir un modèle stable et précis.

2.4.3.3 Cas des paramètres de modélisation

Nous allons nous intéresser aux différentes origines d'instabilité d'un modèle empirique vis-à-vis des paramètres de modélisation. Ces derniers peuvent être classés en deux catégories. La première regroupe les termes utilisés dans la formule mathématique du modèle empirique. Ces paramètres servent à modéliser différents effets tels que la diffusion des espèces chimiques ou la diffraction optique et ils sont ajustés par des mesures expérimentales de dimensions critiques. Ils sont donc des paramètres empiriques. La deuxième catégorie regroupe les paramètres qui sont utilisés par l'algorithme de calcul durant la simulation des contours ou l'optimisation du modèle. Ils servent à remplir des fonctions telles que la définition de la zone de calcul et le filtrage du bruit de mesures. L'instabilité d'un modèle vis-à-vis des paramètres de modélisation peut être envisagée dans trois cas.

2.4.3.3.1 Dépendance élevée aux paramètres empiriques

Dans ce cas, le risque d'instabilité vient d'une dépendance forte du modèle aux paramètres empiriques utilisés. Une technique simple permettant de quantifier cette dépendance est la méthode de la réponse en surface RSM (voir définition page 41). Cette dernière nous permet, par le biais d'un plan d'expérience, d'exprimer la réponse du modèle en fonction des paramètres empiriques utilisés en considérant les dépendances linéaires (ou au plus quadratiques) et croisées. Par l'intermédiaire de cette approche, la réponse du modèle $f(x, y)$ en fonction des combinaisons des termes s'écrit :

$$f(x, y) = c_0 + c_1 P_1 + \dots + c_n P_n + c_{11} P_1^2 + \dots + c_{ij} P_i P_j + \dots + c_{nn} P_n^2, \quad (2.5)$$

où $\{c_k\}$ est l'ensemble des coefficients, P_k les paramètres empiriques du modèle et n leur nombre.

Nous prenons comme exemple d'étude le modèle à seuil variable VT5 car sa fonction caractéristique (voir équation 2.4) est similaire à la forme donnée par l'équation 2.5. La figure 2.17 présente une comparaison des coefficients c_k de deux modèles VT5 de forme identique.

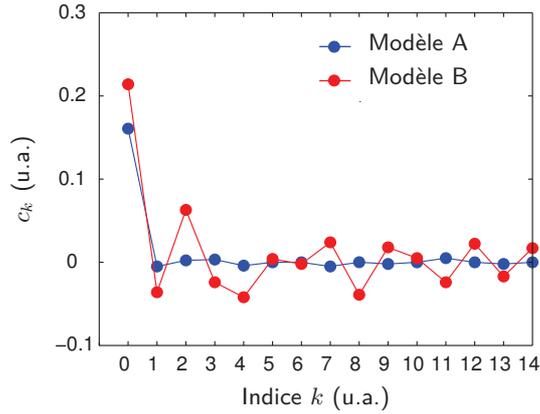


Figure 2.17 – Comparaison des coefficients de deux modèles empiriques à seuil variable (VT5) présentant différents degrés d'instabilité liée aux paramètres empiriques. Notons ici que c_0 est le coefficient attribué au modèle optique pur.

D'après la figure 2.17, le modèle B présente une forte dépendance aux paramètres empiriques utilisés, contrairement au modèle A qui reste plus proche du modèle optique pur (les coefficients c_k , tels que $k \neq 0$, sont faibles devant c_0). Dans l'objectif d'étudier la sensibilité des deux modèles A et B aux variations du focus, nous avons comparé sur la figure 2.18 les courbes de Bossung simulées avec ces deux modèles à celles obtenues en utilisant un modèle à seuil constant. Cette figure nous montre que les courbes de CD en fonction du focus (à différentes doses d'exposition) simulées avec le modèle A présentent un comportement similaire à celles obtenues avec le modèle CTR. Par contre, celles simulées avec le modèle B présentent des courbures de signe opposé. Le modèle A, contrairement au modèle B, garde une sensibilité similaire par rapport à celui à seuil constant. Il est donc plus stable que le modèle B.

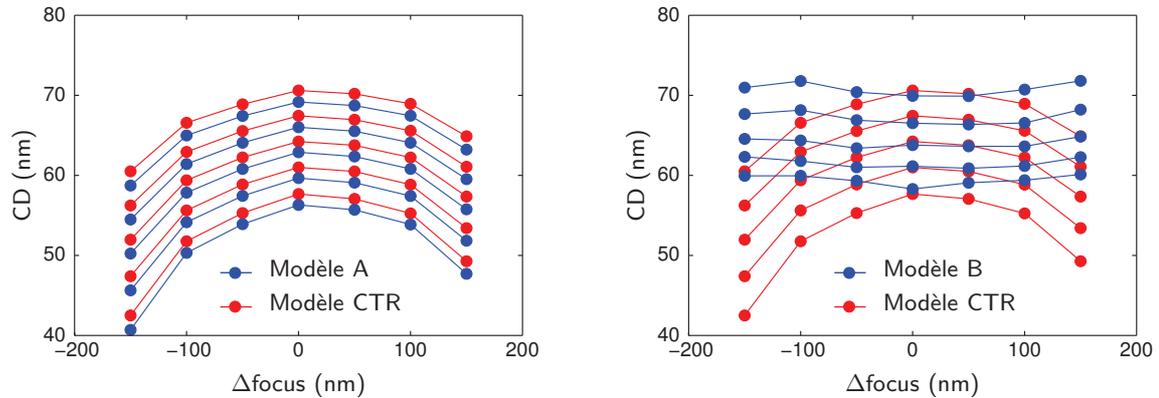


Figure 2.18 – Comparaison des courbes de Bossung simulées pour une ligne isolée avec un modèle à seuil constant et deux modèles à seuil variable (voir figure 2.17) : modèle A (à gauche) et B (à droite). Les niveaux de dose sont calculés pour un réseau de lignes de pas égal à 180 nm tels que les variations de CD sont 0, ± 5 et $\pm 10\%$.

2.4.3.3.2 Différence dans la nature physique des termes

Le risque d'instabilité peut surgir lorsque les paramètres empiriques du modèle sont de natures physiques différentes. Ceci vient du fait que la sensibilité du modèle est différente suivant la nature du paramètre empirique. C'est la raison pour laquelle le modèle peut être plus ou moins instable par rapport à un terme donné. Pour qu'un modèle soit stable, le critère des faibles dépendances n'est pas suffisant. Car même si le poids attribué à un terme est faible, le modèle peut être instable ; et ceci dépend de la nature physique du terme étudié.

Pour justifier ceci, nous avons pris un modèle à seuil variable dont la fonction caractéristique est donnée par l'équation 2.4 telle que $c_k = 0, \forall k \neq 0$. Nous avons ensuite fait varier le coefficient c_6 relatif au terme $Slope^2$ de $-2 \cdot 10^{-3}$ à $6 \cdot 10^{-3}$. Le seuil du modèle obtenu est donc donné par :

$$I_s = c_0 + c_6 Slope^2.$$

Nous avons tracé sur la figure 2.20 les courbes des dimensions critiques d'une ligne isolée en fonction du décalage en focus avec la dose nominale. D'après cette figure, bien que les valeurs de c_6 soient faibles par rapport à c_0 (le rapport c_6/c_0 est de l'ordre de 10^{-2}), le modèle passe d'un état stable par rapport au focus (sa sensibilité aux variations du focus est comparable à celle d'un modèle CTR) à un état quasi instable. Le modèle est donc sensible au terme Slope^2 .

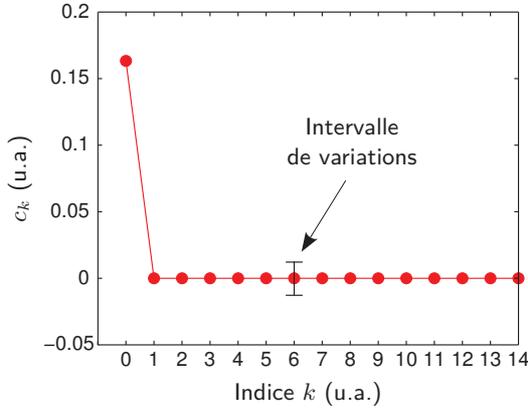


Figure 2.19 – Coefficients attribués aux termes d'un modèle à seuil variable VT5 de fonction caractéristique décrite par l'équation 2.4 telle que $c_k = 0$, $\forall k \neq 0$. Le coefficient c_6 relatif au terme Slope^2 varie de $-2 \cdot 10^{-3}$ à $6 \cdot 10^{-3}$.

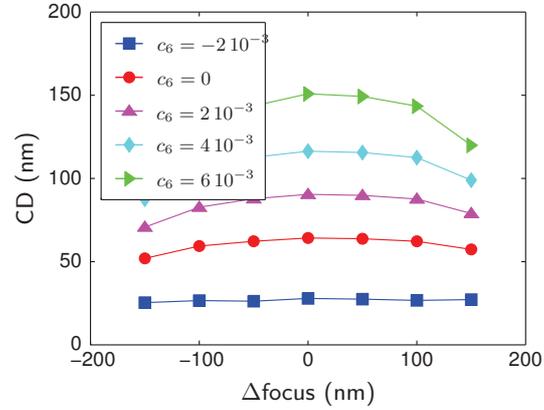


Figure 2.20 – Valeurs de CD en fonction du décalage en focus à dose nominale simulées par un modèle VT5 (voir figure 2.19) pour différentes valeurs du coefficient c_6 relatif au terme Slope^2 . La structure simulée est une ligne isolée.

Pour compléter cette étude, nous avons également fait varier séparément les coefficients c_1 , c_2 , c_3 , c_4 et c_5 relatifs aux termes Imax , Imax^2 , Imin , Imin^2 et Slope respectivement. Pour chaque cas, nous avons tracé les courbes des dimensions critiques en fonction du focus à dose nominale et appliqué une régression polynômiale de second ordre. Ensuite, nous avons calculé le rayon de courbure⁹ des polynômes de régression au point d'abscisse $\Delta\text{focus} = 0$. La figure 2.21 montre l'évolution des rayons de courbure, normalisés par rapport à celui du modèle à seuil constant, en fonction du coefficient c_k . Les rayons de courbure normalisés nous permettent de quantifier les sensibilités des modèles obtenus par rapport à celle du modèle à seuil constant. D'après cette figure, les valeurs des rayons de courbure normalisés sont proches de 1. Les différences entre les courbures normalisées des modèles obtenus et celle du modèle à seuil constant CTR sont inférieures à 4% en valeurs absolues. Nous en déduisons donc que la stabilité du modèle VT5 n'est pas perturbée par des petites variations au niveau des poids attribués aux termes Imax , Imax^2 , Imin , Imin^2 et Slope . En revanche, si l'on regarde la courbe de la figure 2.22, les valeurs des rayons de courbure calculés en faisant varier le coefficients c_6 relatif au terme Slope^2 peuvent être 6 fois plus importantes que celles calculées pour un modèle à seuil constant. La stabilité du modèle VT5 est donc fortement dépendante du poids attribué au terme Slope^2 .

9. Le rayon de courbure au point $(x, y(x))$ d'une fonction réelle $y(x)$ continue et deux fois dérivable est donné par l'équation suivante :

$$R = \frac{y''(x)}{(1 + y'(x)^2)^{3/2}} \quad (2.6)$$

où $y'(x)$ et $y''(x)$ désignent les dérivées première et seconde de $y(x)$ respectivement.

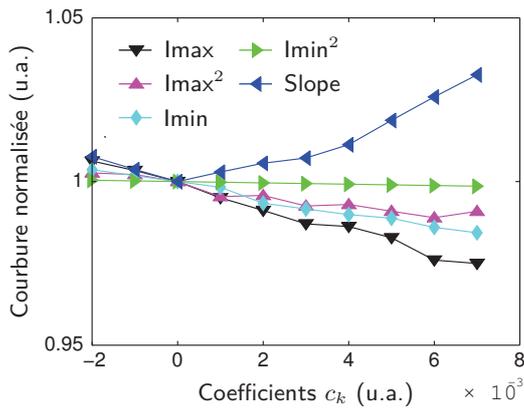


Figure 2.21 – Rayons de courbure normalisés des profils des courbes de Bossung simulées par le modèle VT5 (voir figure 2.19) en fonction de la valeur du coefficient c_k pour les termes l_{max} , l_{max}^2 , l_{min} , l_{min}^2 et $Slope$. Les rayons de courbure ont été calculés en utilisant une régression polynomiale des valeurs de CD simulées par le modèle au point tel que $\Delta focus = 0$.

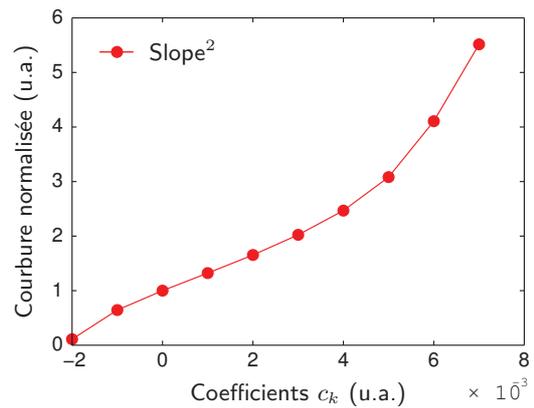


Figure 2.22 – Courbure normalisée des profils des courbes de Bossung simulées par le même modèle en fonction de la valeur du coefficient c_6 relatif au terme $Slope^2$. Cette figure montre que contrairement aux autres termes, une petite variation du coefficient c_6 peut causer l'instabilité du modèle étudié.

D'après les résultats obtenus dans cette partie, nous pouvons conclure que dans le cas où les termes ne sont pas normalisés (cas du modèle VT5 par exemple), les faibles coefficients attribués aux termes ne peuvent confirmer la stabilité du modèle étudié.

2.4.3.3 Existence de termes à effets antagonistes

Le risque d'instabilité peut être important lorsque deux termes du modèle sont à effets antagonistes, c'est-à-dire, pouvant s'annuler mutuellement. Ces termes peuvent être par exemple des produits de convolution de l'intensité lumineuse par des fonctions gaussiennes d'écart types similaires (voir page 26). Pour un modèle à convolution gaussienne, il existe des cas où l'on obtient des maxima et minima d'intensité susceptibles de fausser l'extraction des contours (voir figure 2.23). Dans ces cas, le modèle peut être instable.

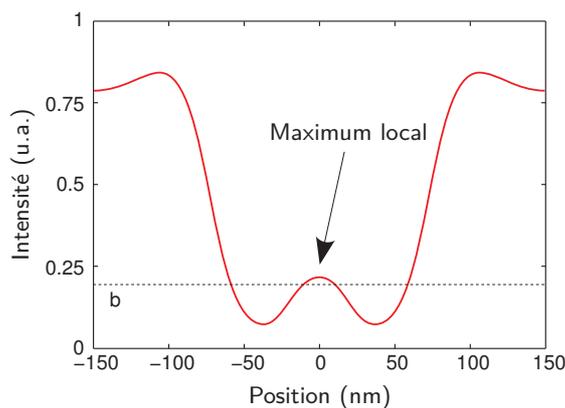


Figure 2.23 – Profil de l'image latente d'une ligne isolée simulée par un modèle empirique à convolution gaussienne. Ce profil présente un maximum local d'intensité qui peut fausser l'extraction des contours lorsque l'on applique un seuil constant.

Il est possible de comparer les profils de l'intensité calculés par le modèle empirique étudié et celui du modèle à seuil constant. Nous pouvons ainsi définir une fonction erreur de la forme :

$$\varepsilon(x, y) = \frac{f(x, y) - I(x, y)}{I(x, y)}, \quad (2.7)$$

où $f(x, y)$ est l'intensité calculée par le modèle empirique et $I(x, y)$ est celle du modèle à seuil constant.

Dans l'équation 2.7, la valeur moyenne de la fonction erreur $\varepsilon(x, y)$ calculée sur un site de simulation peut être utilisée afin de quantifier la sensibilité du modèle empirique par rapport à l'optique.

2.5 Conclusion

Depuis des années, la seule métrique utilisée afin de qualifier un modèle OPC donné était basée sur son erreur quadratique moyenne (RMS error) calculée à la calibration [64]. Le besoin d'étudier la stabilité des modèles OPC est apparu avec l'augmentation de la complexité de ces derniers. En effet, la course à capturer la majorité des effets physiques du procédé lithographique a favorisé l'apparition de modèles empiriques complexes à grand nombre de termes. Étant empiriques, ces derniers peuvent être de natures différentes et sans rapport avec la physique. En outre, leurs comportements peuvent être incompréhensibles et/ou imprévisibles lorsque l'on fait varier des paramètres du procédé (tels que la position du plan focal et la dose d'exposition). Devant un tel problème, il est devenu nécessaire d'étudier la stabilité de ces modèles par rapport aux variations des paramètres du procédé.

Dans ce chapitre, nous avons montré que la modélisation du procédé photolithographique pour les nœuds technologiques avancés nécessite des modèles empiriques de complexité supérieure capables de capturer une bonne partie des effets indésirables du procédé dans un temps raisonnable et de garantir ainsi une précision adéquate qui respecte la feuille de route de l'ITRS. Pour l'étude de la stabilité, nous avons dressé un état de l'art des méthodes existantes et étudié leurs avantages et inconvénients. Nous avons également décrit les principales causes d'instabilité des modèles empiriques avancés, causes qui sont principalement liées aux paramètres physiques et empiriques d'une part, et à la qualité et au nombre des mesures expérimentales de calibration d'autre part. Dans ce contexte, nous avons démontré que la stabilité d'un modèle à seuil variable dépend du type et du nombre de mesures utilisées pendant la calibration. Nous avons également étudié l'effet d'un bruit de mesures gaussien sur la stabilité d'un modèle à seuil variable. À cette étape, nous avons pu démontrer que la stabilité de ce type de modèle ne dépend pas du taux de bruit injecté, mais que, au contraire, la métrique de l'erreur quadratique moyenne est fortement impactée par celui-ci et devient donc compromise.



Chapitre 3

Étude de la stabilité de modèles empiriques par la méthode perturbative

La majorité des méthodes proposées dans la littérature pour l'étude de la stabilité ont pour principal objectif l'évaluation de la sensibilité du procédé photolithographique plutôt que l'étude de la stabilité elle-même. La difficulté de trouver une méthode compatible avec notre définition de la stabilité nous a permis de développer une nouvelle méthode perturbative. Cette dernière est basée sur la réponse du modèle à des petites variations dans l'un des paramètres du procédé. Son principe consiste à comparer la sensibilité du modèle lui-même pour une structure quelconque par rapport à celle de l'optique pur lorsque l'on fait varier un paramètre du procédé. Dans ce chapitre, nous allons, d'abord, détailler le principe de cette nouvelle méthode. Ensuite, nous allons l'utiliser pour l'étude de la stabilité de plusieurs modèles empiriques avancées que nous avons préalablement ajustés avec des mesures de dimensions critiques des nœuds 65 nm et 45 nm. Nous comparerons et discuterons les résultats de la simulation avec ceux de l'expérience.

Sommaire

3.1	Introduction	57
3.2	Principe de la méthode perturbative	57
3.3	Stabilité d'un modèle empirique du nœud 65 nm	58
3.3.1	Stabilité par rapport à la position du plan focal	58
3.3.2	Stabilité par rapport à l'ouverture numérique et aux facteurs de cohérence partielle	60
3.3.3	Stabilité par rapport à l'épaisseur de la résine	61
3.3.4	Application à un circuit quelconque	62
3.3.5	Vérification de la couverture du modèle	64
3.3.5.1	Définition de l'espace de paramètres du modèle	64
3.3.5.2	Choix de la distance d'un site par rapport à l'espace des paramètres	65
3.3.5.3	Détermination des motifs sensibles	65
3.4	Stabilité de modèles empiriques du nœud 45 nm	66
3.4.1	Stabilité par rapport à la position du plan focal	67
3.4.2	Stabilité par rapport à l'ouverture numérique	68
3.4.3	Stabilité par rapport aux facteurs de cohérence partielle	69
3.4.4	Comparatif des différents modèles	70
3.4.5	Vérification expérimentale	71
3.5	Conclusion	72

3.1 Introduction

Dès lors que la nécessité d'utiliser des modèles empiriques complexes pour la modélisation des contours dans la résine commence à apparaître, quelques approches ont été développées afin d'étudier la stabilité des modèles utilisés [62, 65–74] ou, d'une façon plus générale, la sensibilité du procédé photolithographique. Nous avons présenté dans le chapitre précédent quelques unes de ces méthodes et discuté leurs avantages et inconvénients. Nous devons donc choisir parmi toute une série de méthodes existantes celle qui convient le plus avec notre définition de la stabilité. Néanmoins, la seule méthode qui nous permet éventuellement d'évaluer la stabilité numérique d'un modèle est celle qui est proposée dans le logiciel Calibre™ de la société Mentor Graphics. Les autres semblent plutôt s'intéresser à l'étude de la sensibilité du procédé. Contrairement à notre définition de la stabilité qui est plutôt basée sur les variations des dimensions critiques (CD), l'approche de Mentor Graphics est basée sur les différences entre l'intensité lumineuse de départ et du profil de l'image latente simulée par le modèle. Nous avons donc réuni son principe avec celui des méthodes OFAT et PS afin de les adapter à notre étude. Dans ce contexte, nous avons développé une nouvelle méthode que nous appelons la méthode perturbative.

3.2 Principe de la méthode perturbative

Le principe de cette méthode est basé sur la quantification de la variation de dimensions critiques CD engendrée par une petite variation d'un paramètre donné autour de sa valeur nominale [75]. Les différences de CD correspondent aux différences de contours dans la résine (ΔEPE) comme représentées sur la figure 3.1.

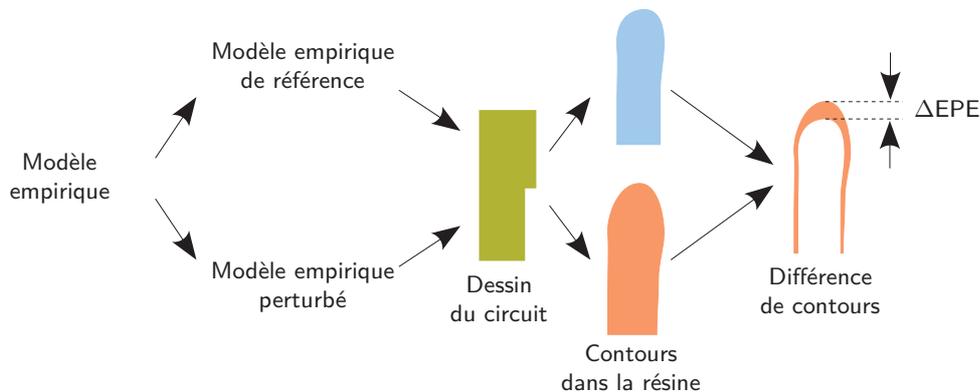


Figure 3.1 – Schéma illustrant les étapes suivies pour le calcul de ΔEPE pour l'étude de la stabilité.

La sensibilité d'un modèle empirique par rapport aux variations d'un paramètre P est donnée par :

$$S_{\text{modèle}} = \frac{\partial EPE_{\text{modèle}}}{\partial P} \quad (3.1)$$

En l'occurrence, la sensibilité d'un modèle à seuil constant est donnée par :

$$S_{\text{optique}} = \frac{\partial EPE_{\text{optique}}}{\partial P} \quad (3.2)$$

D'après notre définition de la stabilité (voir définition page 44), un modèle empirique est considéré stable si sa sensibilité est égale à celle du modèle à seuil constant. Nous pouvons alors évaluer la stabilité du modèle en calculant le rapport des sensibilités du modèle étudié et de celui à seuil constant.

$$\frac{S_{\text{modèle}}}{S_{\text{optique}}} \approx \frac{\Delta EPE_{\text{modèle}}}{\Delta EPE_{\text{optique}}} \quad (3.3)$$

Les variations du paramètre P étant faibles, l'approximation dans l'équation 3.3 est donc justifiée. Nous introduisons ici un facteur de stabilité, noté β , que l'on exprime par l'équation suivante :

$$\beta = \begin{cases} \log\left(\frac{\Delta EPE_{\text{modèle}}}{\Delta EPE_{\text{optique}}}\right) & \text{si } \Delta EPE_{\text{modèle}} \text{ et } \Delta EPE_{\text{optique}} \text{ sont du même signe,} \\ \infty & \text{sinon} \end{cases} \quad (3.4)$$

De la même façon, nous pouvons évaluer la contribution de la partie résine seule dans l'instabilité du modèle par un facteur α que l'on exprime ainsi :

$$\alpha = \frac{|S_{\text{modèle}} - S_{\text{optique}}|}{|S_{\text{optique}}|} \approx \frac{|\Delta EPE_{\text{modèle}} - \Delta EPE_{\text{optique}}|}{|\Delta EPE_{\text{optique}}|} \quad (3.5)$$

D'après l'équation 3.4, un modèle est considéré stable si β ou α sont nuls.

3.3 Stabilité d'un modèle empirique du nœud 65 nm

Nous allons utiliser la méthode perturbative pour étudier la stabilité d'un modèle à seuil variable utilisé pour le niveau grille du nœud 65 nanomètres. Le modèle a été préalablement ajusté avec environ 1000 mesures effectuées sur des motifs 1D et 2D. Le tableau 3.1 résume les paramètres de procédé que nous avons utilisés dans cette étude. Ces paramètres regroupent l'ouverture numérique ON du système optique, les facteurs de cohérence partielle de la source, la position du plan focal et l'épaisseur de la résine. Les variations en entrée n'ont que peu d'importance tant qu'elles restent suffisamment petites tout en ayant une réponse suffisante pour être mesurée.

Paramètre	Variations
focus	± 40 nm
ON	$\pm 3\%$ ON
σ_{in}	$\pm 3\%$ σ_{in}
σ_{out}	$\pm 3\%$ σ_{out}
Épaisseur résine	± 5 nm

Tableau 3.1 – Intervalles de variations des paramètres de procédé pour l'étude de la stabilité d'un modèle à seuil variable par la méthode perturbative

3.3.1 Stabilité par rapport à la position du plan focal

Nous nous intéressons à étudier la stabilité du modèle à seuil variable par rapport aux variations du focus, c'est-à-dire la position du plan focal. Sur la figure 3.2, nous avons tracé l'histogramme de β pour les structures de calibration. La distribution de β est pratiquement gaussienne avec une moyenne égale à 0.3 et un écart type égal à 0.84.

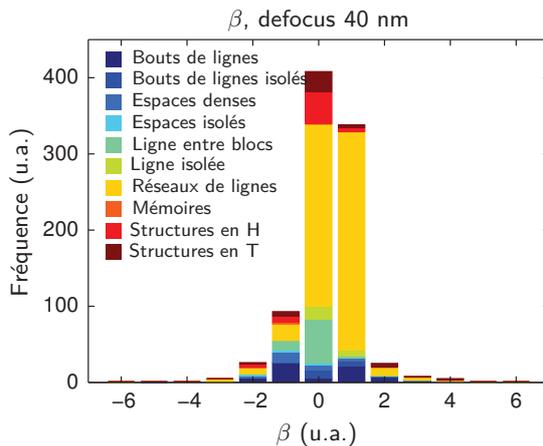


Figure 3.2 – Histogramme de β calculé pour un décalage en focus égal à 40 nm.

En essayant de comprendre la sensibilité du modèle au focus pour les réseaux de lignes, nous avons tracé sur la figure 3.3 la courbe de β en fonction du pas pour des réseaux de lignes de dimension 100 nm ($\pm 10\%$). Cette courbe montre un comportement hyperbolique avec deux asymptotes perpendiculaires menées par les droites $x = 0.03$ et $y = 272.45$. L'asymptote horizontale nous indique que plus le pas du réseau est grand, plus la valeur de β est proche de zéro et plus le modèle est stable pour ces structures. L'asymptote verticale nous indique que le modèle risque d'être instable pour les valeurs du pas de réseau avoisinant 273 nm.

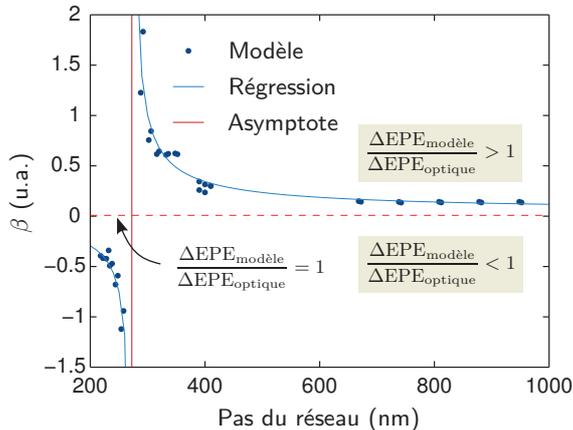


Figure 3.3 – Courbe de β , calculées pour un defocus égal à 40 nm, en fonction du pas pour des réseaux de lignes de dimension 100 nm (± 10 nm). La régression de cette courbe est une hyperbole d'équation $f(t) = 0.03 + 11.38(t - 272.45)^{-0.74}$.

Au voisinage de l'asymptote verticale, la valeur de β tend vers plus ou moins l'infini. Une valeur fortement négative de β signifie que le modèle est insensible aux variations du focus ($\Delta EPE_{\text{modèle}} = 0$) tandis qu'une valeur positive élevée représente une forte sensibilité du modèle par rapport à celle du modèle à seuil constant. Afin de comprendre ce qui se passe, nous avons tracé sur la figure 3.4 les courbes de ΔEPE en fonction du décalage en focus calculées par les deux modèles CTR et VT5 en prenant deux pas de réseau situés de part et d'autre de l'asymptote verticale (positionnée à 273 nm). Ces courbes nous donnent une vision des profils des courbes de Bossung avec la dose nominale. D'après cette figure, les courbes de ΔEPE calculées par le modèle VT5 sont décentrées et le meilleur focus est décalé par rapport à celui prédit par le modèle CTR. Ce décalage en focus est assez fréquent lorsque le seuil d'extraction calculé par le modèle coïncide avec l'isofocale pour ces types de réseaux. Ceci est confirmé par les valeurs très faibles de ΔEPE de la figure 3.4; ce qui traduit des courbes de Bossung qui restent quasiment plates dans cette gamme de focus.

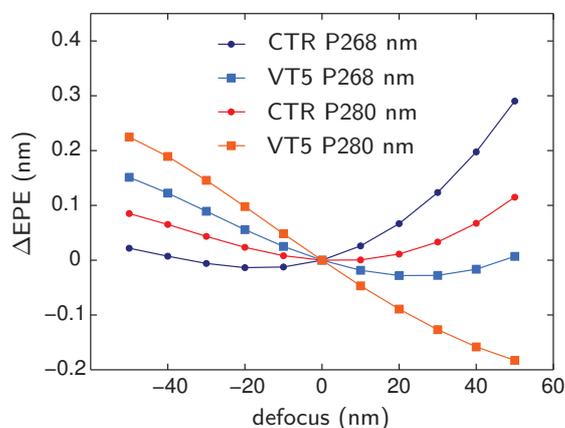


Figure 3.4 – Courbes de ΔEPE en fonction du décalage en focus pour deux réseaux de lignes ($CD = 102$ nm) de pas 268 et 280 nm respectivement calculées par les modèles CTR et VT5.

Nous avons également tracé sur la figure 3.5 les courbes de β en fonction du pas de réseau pour des lignes de dimensions 70, 100, 140 et 200 nm ($\pm 10\%$) où l'on observe le même comportement

asymptotique que pour le cas de la figure 3.3. Toutefois, la position de l'asymptote verticale est différente à chaque fois que l'on prend un CD différent. Pour des lignes de CD égal à 70 nm, celle-ci apparaît pour un pas de réseau avoisinant 200 nm puis vers 550 nm. Dans les 2 autres cas, elle est positionnée à 360 nm (pour les lignes de CD égal à 140 nm) et à 480 nm (pour les lignes de CD égal à 200 nm). Ces discontinuités correspondent aux réseaux de lignes avec des isofocales égales aux seuils d'extraction calculés par le modèle.

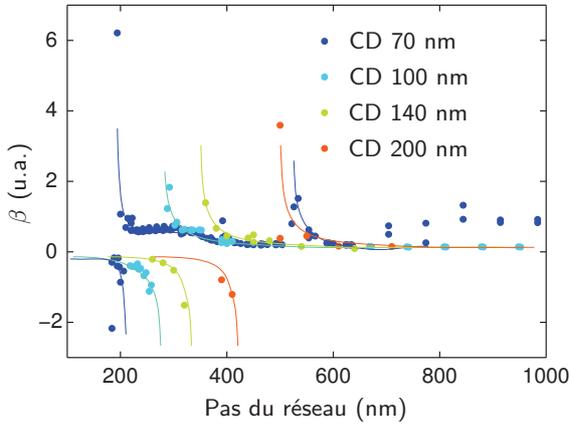


Figure 3.5 – Courbes de β , calculées pour un defocus égal à 40 nm, en fonction du pas pour des réseaux de lignes de dimensions 70, 100, 140 et 200 nm ($\pm 10\%$). Les courbes sont de forme hyperbolique avec différentes positions de l'asymptote verticale qui présentent les réseaux de lignes auxquels le modèle risque d'être instable.

3.3.2 Stabilité par rapport à l'ouverture numérique et aux facteurs de cohérence partielle

Nous avons étudié la stabilité du modèle à seuil variable par rapport aux variations de l'ouverture numérique ON et des facteurs de cohérence partielle de la source. Pour cette étude, les variations de l'ouverture numérique ON sont choisies égales à $\pm 3\%ON$ tandis que celles des facteurs de cohérence partielle σ_{in} et σ_{out} sont égales à $\pm 3\%\sigma_{out}$. Nous avons également étudié la stabilité du modèle pour toutes les combinaisons possibles des paramètres σ_{in} et σ_{out} en prenant le centre σ_c ($(\sigma_{in} + \sigma_{out})/2$) et la largeur σ_{rin} ($\sigma_{out} - \sigma_{in}$) de la source annulaire (voir figure 3.6).

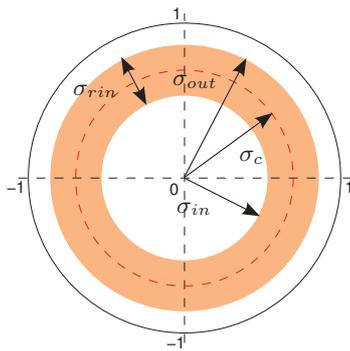


Figure 3.6 – Schéma illustrant les coefficients de cohérence partielle de la source, σ_{in} et σ_{out} ainsi que les paramètres $\sigma_c = \frac{\sigma_{in} + \sigma_{out}}{2}$ et $\sigma_{rin} = \sigma_{out} - \sigma_{in}$.

Nous avons tracé sur les figures 3.7 et 3.8 les histogrammes de β pour des variations de 3% de l'ouverture numérique ON et du facteur de cohérence partielle σ_{out} (pire cas). D'après ces deux figures, il existe peu de structures présentant un risque d'instabilité pour le modèle. Pour ces deux cas, les motifs sensibles sont des réseaux de lignes de pas petits (inférieurs à 260 nm) ou des structures de faibles dimensions (CD inférieur à 80 nm).

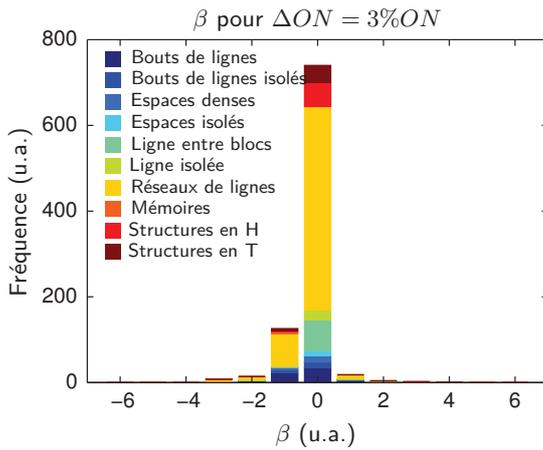


Figure 3.7 – Histogramme de β calculé pour $\Delta ON = 3\% ON$.

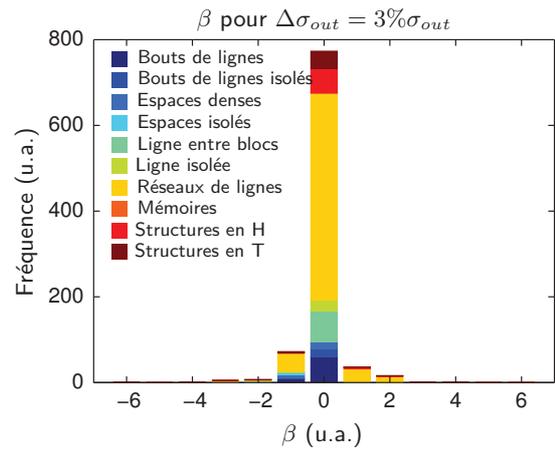


Figure 3.8 – Histogramme de β calculé pour $\Delta\sigma_{out} = 3\%\sigma_{out}$.

3.3.3 Stabilité par rapport à l'épaisseur de la résine

La dispersion du facteur β calculé pour une variation de l'épaisseur de résine de ± 5 nm (voir figure 3.9 nous indique que le modèle présente un risque élevé d'instabilité pour différents types de motifs.

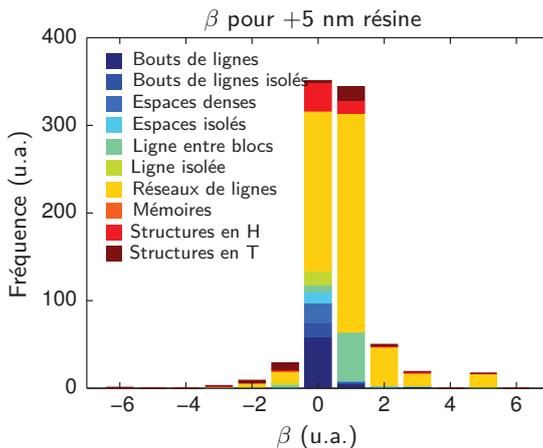


Figure 3.9 – Histogramme de β calculé pour une variation de +5 nm de l'épaisseur de résine.

Comme dans le cas du focus, nous avons tracé sur la figure 3.10 la courbe de β en fonction du pas pour des réseaux de lignes de dimensions 100 nm ($\pm 10\%$). Cette courbe présente un comportement hyperbolique avec une asymptote verticale située à 251 nm, puis un petit pic situé à 385 nm. En observant les courbes de ΔEPE simulées pour deux pas de réseau situés de part et d'autre de l'asymptote (voir figure 3.11), il est clair que pour le pas situé à gauche de l'asymptote (pas = 248 nm), le modèle VT5 est peu sensible aux variations de l'épaisseur de résine contrairement au modèle CTR. Ce qui explique une valeur de β négative et une asymptote qui tend vers moins l'infini. D'autre part, les sensibilités des deux modèles sont inversées pour le pas de réseau situé à droite de l'asymptote (pas = 254 nm), la valeur de β tend vers l'infini. L'instabilité du modèle pour les réseaux de lignes de pas proches de 251 nm vient du fait qu'une variation de l'épaisseur de résine se traduit par un léger décalage en focus. Or, nous avons vu dans le cas du focus le même comportement pour des réseaux de lignes de dimensions 100 nm ($\pm 10\%$) et de pas égal à 273 nm. Le pas 254 nm peut donc bien correspondre à une valeur de l'isofocale coïncidant parfaitement avec le seuil d'extraction calculé par le modèle. Cette remarque est justifiée par les faibles valeurs de ΔEPE observées sur la figure 3.11.

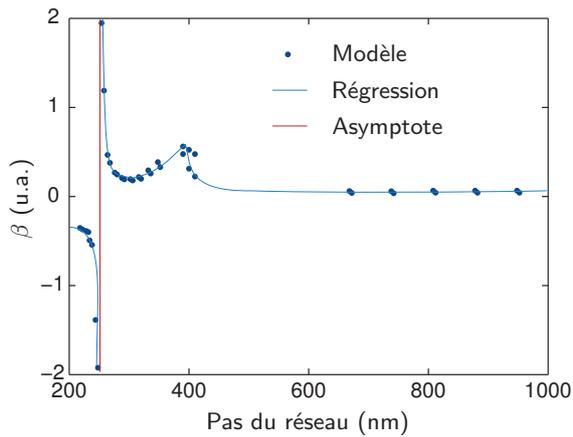


Figure 3.10 – Courbe de β , calculées pour une variation de l'épaisseur de la résine égale à $+5$ nm, en fonction du pas pour des réseaux de lignes de dimension 100 nm (± 10 nm). Cette courbe présente un comportement hyperbolique au voisinage de 251 nm puis un petit pic à 385 nm.

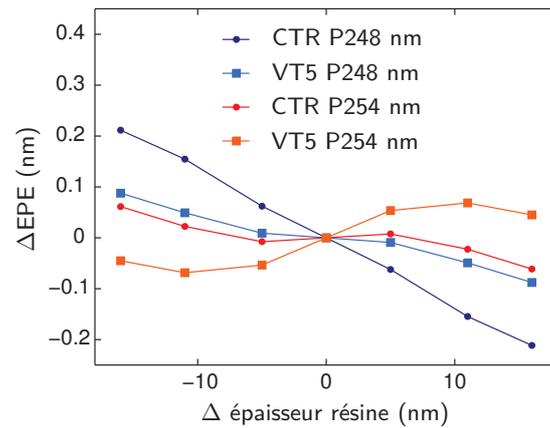


Figure 3.11 – Courbes ΔEPE en fonction des variations de l'épaisseur de la résine pour deux réseaux de lignes (CD = 102 nm) de pas 248 et 254 nm respectivement calculées par les modèles CTR et VT5.

3.3.4 Application à un circuit quelconque

Nous avons vu précédemment que le modèle à seuil variable de cette étude présente un risque d'instabilité vis-à-vis du focus pour plusieurs structures de calibration, comme les réseaux de lignes avec des pas particuliers. Il est donc intéressant de vérifier la stabilité du modèle pour un circuit quelconque. Nous avons calculé le facteur β pour l'ensemble des sites de simulation du circuit. Les sites sont positionnés de sorte qu'ils soient perpendiculaires aux contours de la résine simulés par le modèle optique de référence (voir figure 3.12).

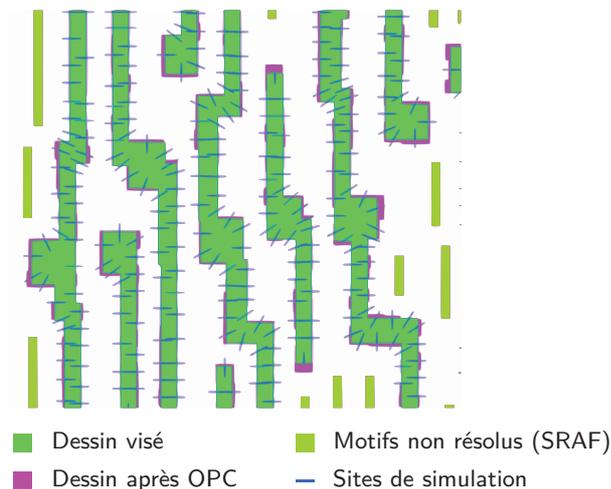


Figure 3.12 – Exemple d'une région du circuit de cette étude munie d'un ensemble de sites de simulation. Ces derniers sont positionnés sur les bords du circuit perpendiculairement aux contours du modèle optique.

Nous avons tracé sur la figure 3.13 l'historgramme de β pour une variation de focus égale à 40 nm. D'après cette figure, il est clair que la majorité des sites simulés présentent des valeurs de β proches de zéro. Ce résultat nous indique que la sensibilité du modèle VT5 est égale à celle du modèle optique pour ces sites. Néanmoins, on peut remarquer qu'il existe des motifs pour lesquels le modèle VT5 risque d'être instable, notamment les bouts de lignes, les interconnexions, les coins convexes et les grilles de transistors. Nous avons récapitulé sur le tableau 3.2 les valeurs

maximales de β en fonction du type de motifs dans le circuit. Puis, nous avons représenté sur la figure 3.14 quelques cas de motifs présentant des valeurs de β élevées.

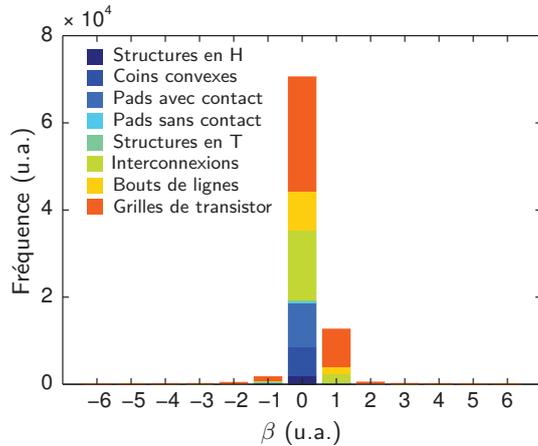


Figure 3.13 – Histogramme de β pour l'ensemble des sites du circuit pour une variation du focus égale à 40 nm

Type de motif	ΔEPE (nm)	β
Structures en H	-1.39 -1.91	0.31
Coins convexes	0.01 -0.67	∞
Pads avec contact	-0.2 -0.98	1.59
Pads sans contact	0.16 0.28	0.56
Structures en T	-0.07 -0.62	2.18
Interconnexions	-0.01 -0.56	4.02
Bouts de lignes	-0.01 -1.79	5.18
Grilles de transistor	-0.01 -0.72	4.27
CTR	VT5	

Tableau 3.2 – Valeurs maximales de β en fonction du type de motif pour le circuit de l'étude. Les valeurs de ΔEPE sont également données pour les deux modèles VT5 et CTR.

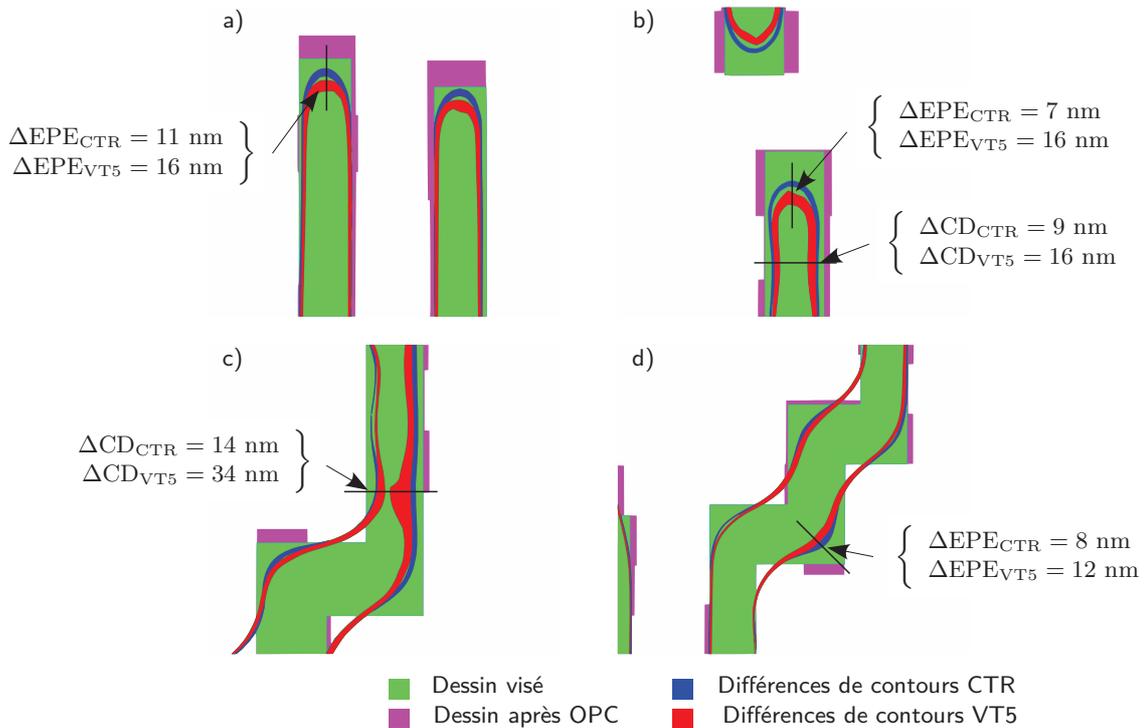


Figure 3.14 – Quelques cas de motifs présentant des valeurs élevées de $|\beta|$ calculées par le modèle VT5 pour des variations de focus égales à 40 nm. Les cas a) et b) présentent des bouts de lignes sensibles. Les cas b) et c) présentent un amincissement de lignes qui peuvent constituer des risques d'instabilité du modèle. Le cas d) est celui d'un coin convexe sensible.

Dans les cas a) et b), nous avons une différence de contours plus importante pour le modèle VT5 contrairement au cas du modèle optique. Dans le cas b), nous remarquons un début d'amincissement à proximité du bout de ligne qui est visiblement plus marqué avec le modèle VT5. Le cas c) montre un amincissement beaucoup plus important avec une dimension finale égale à 35 nm avec le modèle VT5 et 55 nm avec le modèle optique. Finalement, le cas d) représente une différence de contours plus importante au niveau des coins convexes des motifs pour le modèle VT5. La sensibilité élevée du modèle VT5 qui prédit ou pas l'apparition de lieux d'amincissement (ou de coupure) des lignes dans la résine suggère une vérification expérimentale

des données avant d'utiliser le modèle. Les conséquences d'une prédiction erronée des contours de la résine peuvent être plus ou moins pénalisantes notamment sur la précision de l'algorithme de correction et le temps additionnel requis pour la vérification expérimentale sur plaquette.

3.3.5 Vérification de la couverture du modèle

3.3.5.1 Définition de l'espace de paramètres du modèle

Lors de la calibration d'un modèle empirique, des mesures expérimentales de dimensions critiques sont collectées sur un ensemble de structures types choisies dans le but de représenter celles d'un circuit quelconque de production. Vu le grand nombre de structures que peut inclure un circuit de production, il est pratiquement impossible de les prendre toutes en compte durant la calibration. Une des méthodes utilisées pour réduire considérablement leur nombre est celle publiée par K. Patterson et al. [76]. Elle consiste à filtrer les informations redondantes à partir des quantités I_{\max} (maximum d'intensité), I_{\min} (minimum d'intensité), Slope (pente du profil d'intensité) et Factor (courbure du motif).

L'ensemble des vecteurs formés par les quadruples $\{I_{\max}, I_{\min}, \text{Slope}, \text{Factor}\}$ simulés pour les structures de calibration définit l'espace de paramètres du modèle et constitue son domaine d'apprentissage. La superposition des deux espaces de paramètres calculés pour les motifs de calibration et les sites d'un circuit quelconque de production permet d'analyser la couverture du modèle [77].

Nous avons vu précédemment que la sensibilité du modèle est différente suivant le type de motif. Les bouts de lignes, par exemple, sont les motifs les plus sensibles pour une variation de focus. Existe-t-il un lien donc entre la couverture et la sensibilité du modèle? Pour répondre à cette question, nous allons analyser la couverture du modèle en utilisant cette notion d'espace de paramètres. La figure 3.15 montre la superposition (en deux dimensions) des espaces de paramètres d'un circuit quelconque et des motifs utilisés en calibration. D'après cette figure, on peut remarquer qu'il existe des sites du circuit qui ne sont pas bien représentés par l'espace des paramètres de calibration. Cela constitue probablement un risque d'instabilité pour ce modèle notamment pour les motifs des zones non-couvertes. La prédiction du modèle peut donc être fautive lorsqu'un site est très distant de l'espace de calibration.

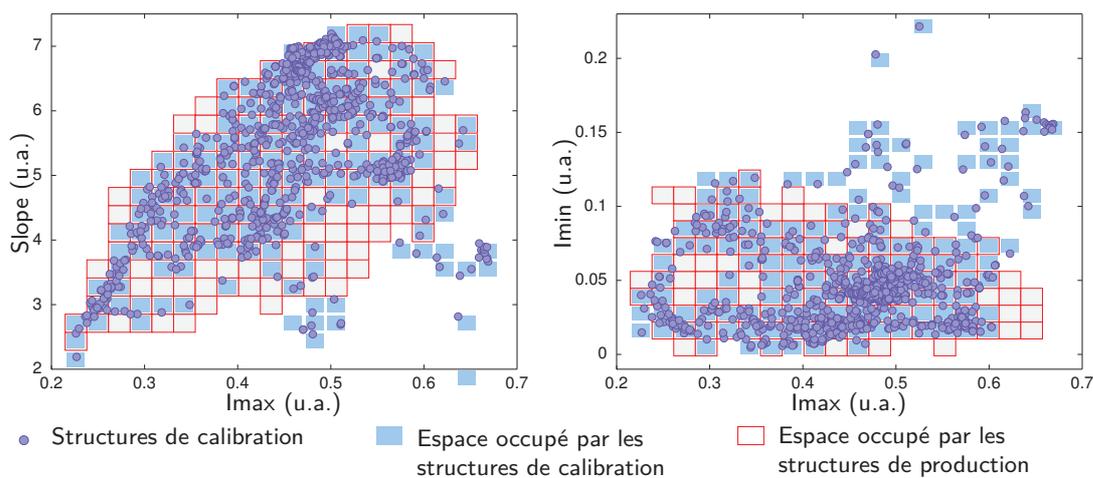


Figure 3.15 – Superposition de deux espaces de paramètres : celui des motifs de calibration (domaine d'apprentissage) et celui des sites d'un circuit de production.

Nous allons maintenant essayer de démontrer que la sensibilité du modèle pour un motif donné augmente quand ce dernier n'est pas bien représenté dans l'espace des paramètres de calibration. Pour ce faire, nous allons évaluer la distance de chaque site du circuit de production par rapport à l'espace des paramètres.

3.3.5.2 Choix de la distance d'un site par rapport à l'espace des paramètres

Il existe de multiples façons pour le calcul de la distance d'un point par rapport à un ensemble donné. La manière la plus simple est d'utiliser la forme euclidienne qui consiste à calculer la plus courte distance séparant le point à l'ensemble. Néanmoins, cette méthode ne nous permet pas de distinguer les sites qui sont complètement à l'extérieur de l'ensemble de ceux qui sont à l'intérieur de celui-ci mais très distants de leurs voisins. Pour résoudre ce problème d'environnement local, nous avons choisi une forme gaussienne pour le calcul de la distance que nous décrivons par l'équation suivante :

$$d_i = \frac{1}{N} \sum_{j=1}^N \omega_j e^{-\frac{d_{ij}^2}{2\sigma^2}}, \quad \text{avec } d_{ij} = \left(\sum_{k=1}^4 (P_i^k - P_j^k)^2 \right)^{1/2}. \quad (3.6)$$

Dans cette équation, d_{ij} est égale à la distance euclidienne entre les sites i (circuit de production) et j (circuit de calibration), ω_j est le poids attribué au site j lors de la calibration du modèle, σ est l'écart type de la fonction gaussienne donné par la distance moyenne entre les sites de calibration et P_i^k est la valeur normalisée des paramètres I_{max}, I_{min}, Slope et Factor, calculée pour le site i .

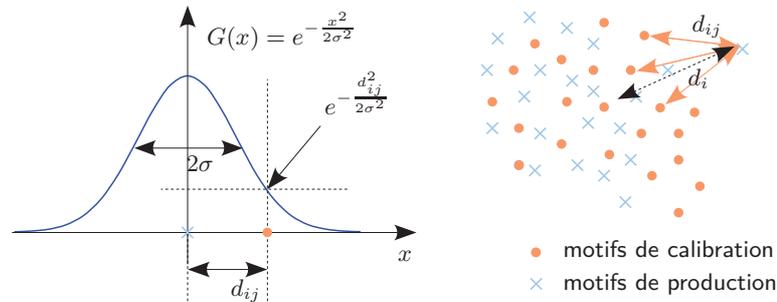


Figure 3.16 – Calcul de la distance d'un site du circuit de production par rapport à l'espace des paramètres de calibration d'un modèle empirique (c'est-à-dire, le domaine d'apprentissage du modèle).

3.3.5.3 Détermination des motifs sensibles

Nous avons calculé le facteur α de tous les sites d'un circuit de production en utilisant le modèle VT5 pour une variation du focus égale à 40 nm. Ensuite, nous avons utilisé la formule de l'équation 3.6 afin de calculer la distance d_i de chaque site du circuit de production à l'espace de paramètres de calibration (domaine d'apprentissage du modèle). La figure 3.17 présente les valeurs de $\alpha \times |\Delta EPE_{CTR}|$ en fonction des distances normalisées d_i .

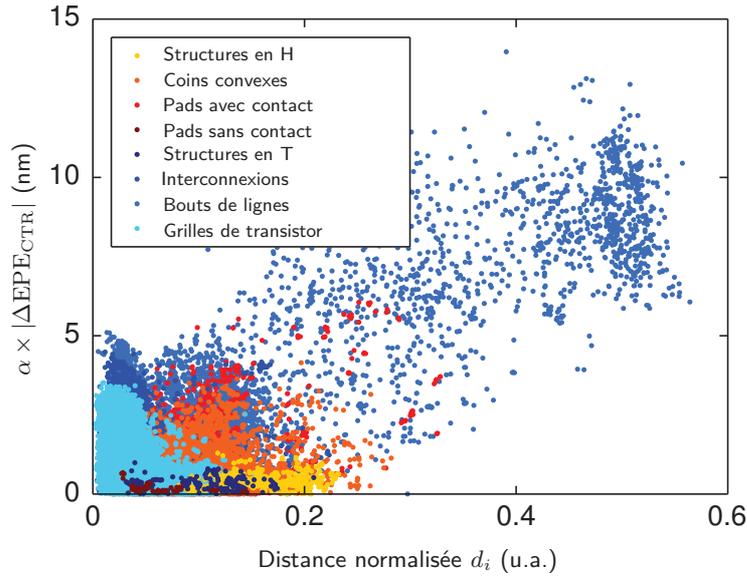
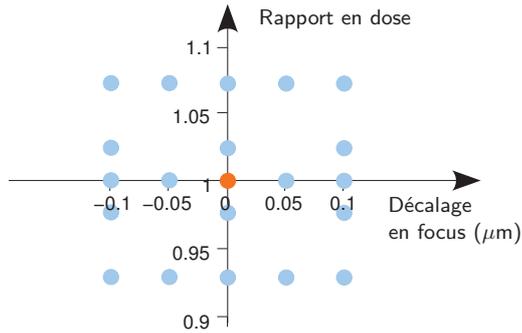


Figure 3.17 – Tracé de $\alpha \times |\Delta EPE_{CTR}| = |\Delta EPE_{VT5} - \Delta EPE_{CTR}|$ en fonction de la distance normalisée d_i calculée en utilisant la forme gaussienne de la distance. Les valeurs sont obtenues pour des variations du focus égales à 40 nm.

D’après la figure 3.17, la sensibilité du modèle est plus importante lorsque la distance à l’espace de paramètres de calibration (domaine d’apprentissage) augmente. En outre, cette figure montre des problèmes de couverture du modèle pour quelques structures 2D, en particulier les bouts de lignes, les pads et les coins convexes. En ce qui concerne les motifs 1D, le modèle présente une bonne couverture de l’espace des paramètres, ce qui traduit les faibles différences de ΔEPE avec le modèle optique. En appliquant un seuil maximal pour ces différences, nous pouvons détecter les motifs du circuit qui ne sont pas bien représentés par l’ensemble des motifs de calibration. Les structures du circuit dont les différences de ΔEPE sont supérieures au seuil fixé sont plus sensibles à celles des motifs de calibration. Ainsi, cette manière de détecter les sites sensibles est une bonne approche pour compléter l’ensemble des motifs de calibration en y rajoutant les structures du circuit dont les différences de ΔEPE sont fortes. Finalement, une calibration du modèle avec des mesures effectuées sur un ensemble plus complet de motifs de calibration semble être une bonne approche pour obtenir une meilleure stabilité.

3.4 Stabilité de modèles empiriques du nœud 45 nm

Nous allons étudier, par la méthode perturbative, la stabilité de trois modèles empiriques avancés par rapport aux paramètres physiques présentés dans le tableau 3.1. Les modèles de l’étude ont été ajustés avec les mêmes mesures expérimentales effectuées sur des structures de test en faisant varier le focus et la dose (cf. figure 3.18). Les mesures effectuées s’élèvent à plus de 5500 mesures et sont réparties sur 21 conditions de focus et de dose. De plus, pour une même condition de focus et de dose, elles ont été moyennées sur plusieurs champs afin de diminuer le bruit. Le poids attribué à chaque motif de calibration est égal au nombre des mesures utilisées dans le calcul de la moyenne. Le tableau 3.3 présente les résultats des erreurs quadratiques moyennes (RMS error) en CD après l’étape de la calibration pour les trois modèles de cette étude. Pour des besoins de confidentialité, nous utiliserons les lettres A, B et C pour désigner les modèles étudiés qui sont les propriétés des sociétés Brion Technologies (pour le cas du modèle à convolution étendue) et Mentor Graphics (pour les modèles à seuil variable VT5 et opérateur différentiel CM1).



Modèles	RMS (nm)
CTR	6.78
A	4.04
B	3.36
C	4.81

Tableau 3.3 – Erreurs quadratiques moyennes en CD de chacun des modèles après calibration.

Figure 3.18 – Ensemble des conditions de procédé utilisées pour la calibration des modèles empiriques de l'étude.

Nous avons ensuite ajusté le seuil d'exposition (seuil de référence) du modèle optique qui sera notre modèle de référence. Nous disposons maintenant de quatre modèles calibrés avec les mêmes mesures expérimentales. Pour chacun de ces modèles, nous avons introduit une petite variation à l'un des paramètres de l'étude (dans la partie optique ou dans celle de la résine) afin de créer un modèle modifié. Nous avons ensuite généré les contours de la résine pour un circuit quelconque du niveau de la grille dessiné selon les règles du nœud technologique 45 nm (cf. figure 3.1). Finalement, les différences ΔEPE des contours prédits par ces modèles ont été extraites¹ puis analysées. Le facteur de stabilité β a été calculé pour tous les sites du circuit².

3.4.1 Stabilité par rapport à la position du plan focal

Nous avons tracé, sur la figure 3.19, les histogrammes de β calculés avec les trois modèles de cette étude pour des variations du focus égales à 40 nm. Les distributions observées sur ces histogrammes sont équivalentes avec un léger avantage au modèle C qui présente l'écart type le plus petit (0.60, contre 0.68 et 0.76 pour les modèles A et B respectivement). Bien que les écarts types des trois distributions de β sont faibles, il existe des sites présentant des valeurs absolues élevées de β . Ces sites peuvent éventuellement présenter un risque d'instabilité pour les trois modèles.

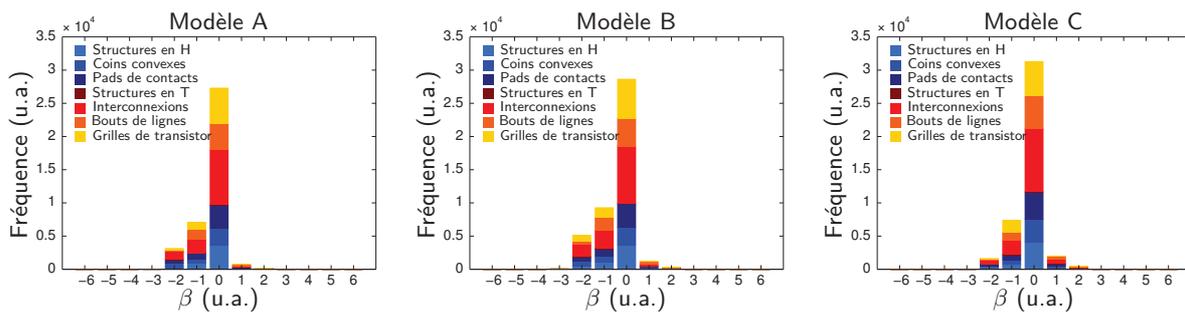


Figure 3.19 – Histogrammes de β calculés pour des variations du plan focal égales à 40 nm en utilisant les trois modèles A, B et C.

Nous avons ainsi représenté sur la figure 3.20 les distributions des sites sensibles tels que $|\beta| > 2$ pour chacun des modèles. La figure montre que les structures les plus sensibles sont principalement des bouts de lignes, des interconnexions et des structures en H. En ce qui concerne les grilles de transistors, la figure montre très peu de grilles sensibles qui, en observant leurs positions dans le circuit, correspondent aux interfaces avec les interconnexions. Le risque d'instabilité de ces modèles pour les grilles de transistor est donc faible. En revanche, les interconnexions

1. L'extraction s'est effectuée par intersection des polygones constituant les sites avec les différences de contours prédits par les modèles. Les signes de ΔEPE ont été attribués par rapport à la position du contour prédit par le modèle au paramètre modifié avec celui prédit avec le modèle de référence.

2. Pour le calcul de β , nous avons arrondi les valeurs de ΔEPE à 10^{-2} nm près.

peuvent constituer un risque élevé d’instabilité, entraînant des prédictions erronées des contours réels dans la résine qui peuvent impacter la correction OPC.

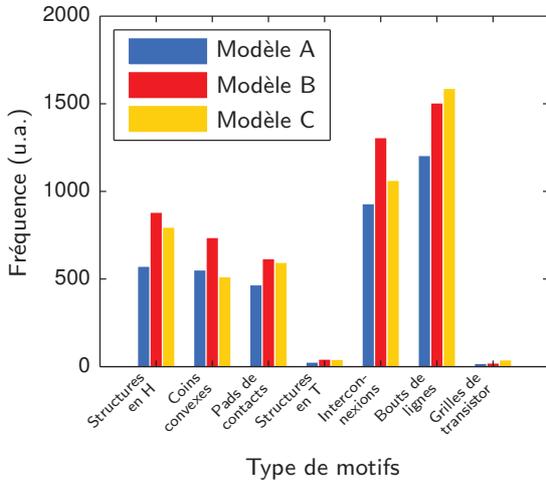


Figure 3.20 – Distributions des motif sensibles ($|\beta| > 2$) du circuit calculées par les trois modèles pour des variations du focus égales à 40 nm. Les motifs les plus sensibles sont principalement des bouts de lignes, des interconnexions et des structures en H.

3.4.2 Stabilité par rapport à l’ouverture numérique

Étudions maintenant la stabilité des trois modèles par rapport à des variations de l’ouverture numérique. Comme dans le cas précédent, nous avons tracé sur la figure 3.21 les histogrammes de β pour des variations de l’ouverture numérique ΔON égales à 3%ON. D’après cette figure, les distributions des modèles A et B sont similaires avec des écarts types respectifs égaux à 0.80 et 0.86. Le modèle C présente un avantage avec une distribution moins dispersée (écart type égal à 0.58) et un comportement symétrique.

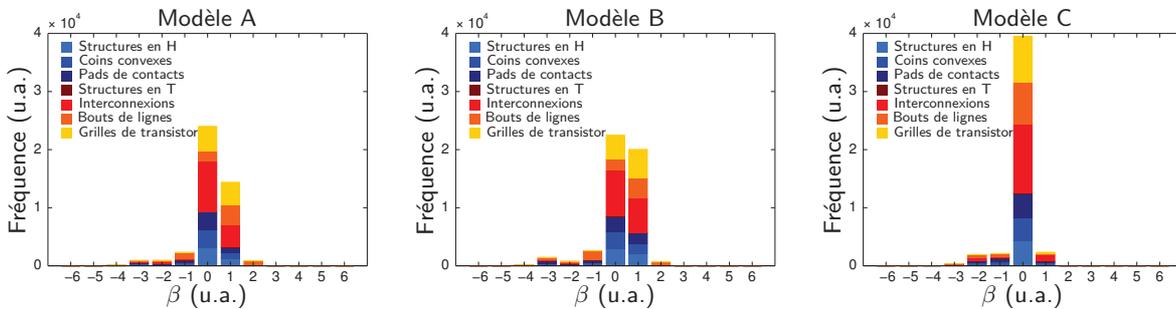


Figure 3.21 – Histogrammes de β calculés pour des variations de l’ouverture numérique ΔON égales à 3%ON en utilisant les trois modèles A, B et C.

D’après la figure 3.22, les motifs les plus sensibles tels que $|\beta| > 2$ sont principalement des bouts de lignes, des interconnexions, des pads de contacts et des coins convexes. Les trois modèles montrent une similitude au niveau des résultats avec un léger avantage au modèle A qui semble avoir un nombre moins élevé de motifs sensibles. Bien que le modèle C admette le plus petit écart type, il présente néanmoins un risque d’instabilité plus important pour les interconnexions et les grilles de transistor.

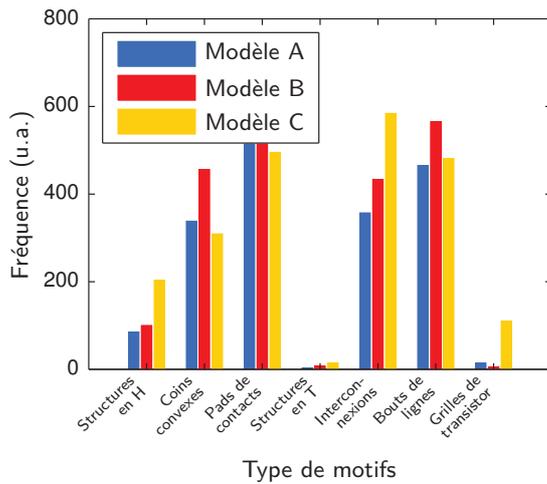


Figure 3.22 – Distributions des motif sensibles ($|\beta| > 2$) du circuit calculées par les trois modèles pour des variations de l'ouverture numérique ΔON égales à $3\%ON$. Les motifs les plus sensibles sont principalement des bouts de lignes, des interconnexions, des pads de contacts et des coins convexes.

3.4.3 Stabilité par rapport aux facteurs de cohérence partielle

Nous avons évalué la stabilité des trois modèles par rapport aux variations des facteurs de cohérence partielle de la source. La figure 3.23 montre que les distributions de β calculées pour des variations du facteur de cohérence partielle σ_{out} égales à $3\%\sigma_{out}$ (pire cas) sont similaires pour les trois modèles avec un avantage au modèle C qui présente l'écart type le plus faible (0.62 contre 0.79 et 0.84 pour les modèles A et B respectivement).

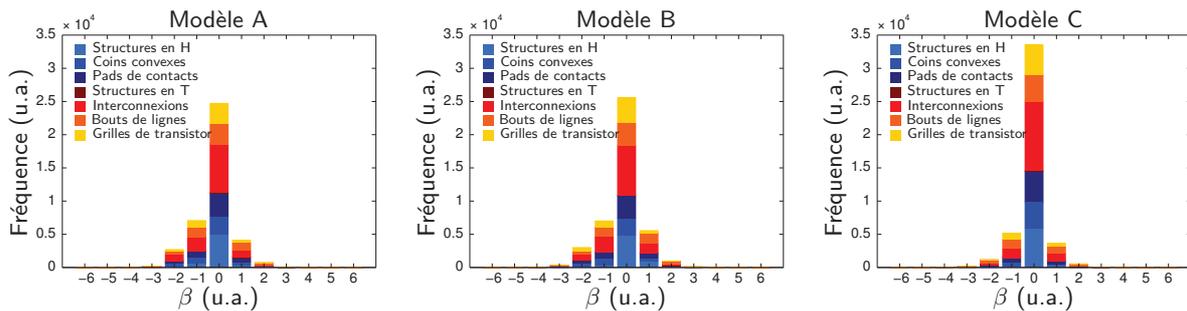


Figure 3.23 – Histogrammes de β calculés pour des variations du facteur de cohérence partielle σ_{out} égales à $3\%\sigma_{out}$ en utilisant les trois modèles A, B et C.

La figure 3.24 montre que les motifs les plus sensibles aux variations du facteur de cohérence partielle σ_{out} sont principalement des bouts de lignes, des interconnexions, des pads de contacts et des coins convexes. Néanmoins, le modèle C montre une moindre sensibilité pour tous ces motifs (à l'exception des bouts de lignes). En ce qui concerne les grilles de transistor, nous n'avons pas détecté de risque d'instabilité pour ces structures. La stabilité de ces trois modèles est donc confirmée pour les grilles de transistor.

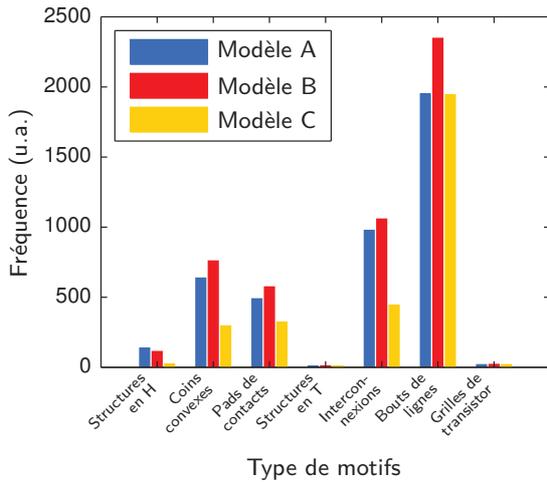


Figure 3.24 – Distributions des motif sensibles ($|\beta| > 2$) du circuit calculées par les trois modèles pour des variations du facteur de cohérence partielle σ_{out} égales à $3\% \sigma_{out}$. Les motifs les plus sensibles sont principalement des bouts de lignes, des interconnexions, des pads de contacts et des coins convexes.

3.4.4 Comparatif des différents modèles

Nous avons récapitulé sur le tableau 3.4 les valeurs maximales de $|\beta|$ calculées sur les sites du circuit pour les trois modèles de cette étude. Bien que les distributions de β soient plus dispersées pour le paramètre de l’ouverture numérique, les résultats du tableau montrent que les modèles A, B et C sont assez stables vis-à-vis de ce paramètre. Cependant, ils montrent des valeurs assez élevées pour pratiquement tous les motifs dans les cas du focus et du coefficient de cohérence partielle. Ces modèles sont donc instables vis-à-vis de ces deux paramètres.

Type de motifs	Modèle A			Modèle B			Modèle C			
	ΔEPE (nm)		$ \beta $	ΔEPE (nm)		$ \beta $	ΔEPE (nm)		$ \beta $	
$\Delta focus = 40 \text{ nm}$	Structures en H	0.06	1.7	3.34	-0.08	1.71	∞	0.06	2.56	3.75
	Coins convexes	-0.08	3.44	∞	0.06	2.97	3.90	0.04	1.66	3.73
	Pads de contacts	0.06	1.88	3.44	0.06	2.56	3.75	0.06	-2	∞
	Structures en T	0.04	0.47	2.46	0.04	1.84	3.83	0.06	-1.23	∞
	Interconnexions	0.04	2.5	4.14	0.04	1.72	3.76	0.04	4	4.61
	Bouts de lignes	0.06	3.06	3.93	0.06	3.06	3.93	0.06	11.54	5.26
	Grilles de transistor	0.06	-1.5	∞	0.06	1.56	3.26	0.3	1.59	1.67
$\Delta ON = 3\% ON$	Structures en H	2	17	2.14	1	8	2.08	2	7	1.25
	Coins convexes	1	7.21	1.98	1	8.25	2.11	1.41	5.83	1.42
	Pads de contacts	1	6.32	1.84	1	7.62	2.03	1	5.39	1.68
	Structures en T	2.24	9.22	1.41	2	9	1.50	2.24	5.39	0.88
	Interconnexions	1	8.25	2.11	3	23	2.04	1	3.16	1.15
	Bouts de lignes	1	18	2.89	1	13	2.56	2.24	15.13	1.91
	Grilles de transistor	2	10	1.61	2	10	1.61	2	5	0.92
$\Delta \sigma_{out} = 3\% \sigma_{out}$	Structures en H	-0.18	4.09	∞	0.05	3.34	4.20	0.06	1.24	3.03
	Coins convexes	0.04	3.66	4.52	0.06	4.14	4.23	0.06	2.6	3.77
	Pads de contacts	0.04	3.62	4.51	0.06	3.62	4.10	0.06	2.68	3.80
	Structures en T	-0.08	4.12	∞	1.4	6.82	1.58	0.47	1.4	1.09
	Interconnexions	0.06	4.14	4.23	0.01	0.81	4.39	0.06	2.18	3.59
	Bouts de lignes	0.06	5.17	4.46	0.06	6	4.61	0.06	13.5	5.42
	Grilles de transistor	0.06	2.18	3.59	0.06	3.16	3.96	0.06	1.12	2.93

■ $|\beta| > 4$

■ $3 < |\beta| \leq 4$

■ $2 < |\beta| \leq 3$

■ $1 < |\beta| \leq 2$

■ $|\beta| \leq 1$

Tableau 3.4 – Valeurs maximales de $|\beta|$ en fonction du type de motifs pour le circuit de l’étude. Les valeurs de ΔEPE sont également données pour chacun des modèles A, B et C.

Bien que le modèle C ait l’erreur quadratique moyenne la plus élevée (cf. tableau 3.3), ce modèle montre un moindre risque d’instabilité par rapport aux paramètres du procédé. Les résultats trouvés dans cette étude sont en accord avec le niveau de complexité des modèles. En effet, le modèle C utilise moins de paramètres de simulation et présente le plus petit écart avec le modèle optique. Ce fait lui procure un sens physique et une meilleure stabilité numérique.

Les modèles A et B, quant à eux, présentent de nombreux termes empiriques ajustés avec des mesures expérimentales effectuées dans des conditions données du procédé. L'extrapolation de ces modèles en dehors des conditions de calibration nécessite d'effectuer d'autres mesures dans les conditions décalées, notamment, dans le cas du facteur de cohérence partielle de la source.

Finalement, le fait que les trois modèles ont été ajustés avec des mesures expérimentales de plusieurs conditions de focus nous a permis d'obtenir une moindre dispersion du facteur β (par comparaison aux autres paramètres). Cependant, les résultats du tableau 3.3 montrent qu'il existe des cas de motifs présentant une forte sensibilité dans le cas du focus. Ceci peut être justifié par une mauvaise couverture du modèle par rapport à ces cas de motifs.

3.4.5 Vérification expérimentale

Des mesures de dimensions critiques ont été effectuées pour différentes structures 1D et 2D préalablement dessinées (puis gravées sur un masque) pour le niveau de grille du nœud technologique 45 nm. Durant l'exposition, nous avons fait varier le focus, l'ouverture numérique (ON) et le facteur de cohérence partielle σ_{out} de la source d'illumination. Ensuite, nous avons calculé les valeurs de ΔCD entre une condition particulière et la condition nominale. La validation expérimentale de la sensibilité du modèle par rapport à des petites variations d'un paramètre de procédé dépend principalement de la précision du modèle et des erreurs de mesures. Nous allons maintenant présenter quelques types de structures présentant un écart important entre les dimensions critiques simulées et mesurées. Pour ce faire, nous avons comparé les valeurs de ΔCD mesurées avec celles simulées en utilisant le modèle empirique et le modèle optique pur. Nous allons nous intéresser uniquement au cas de l'ouverture numérique qui présente une dispersion plus importante du facteur β et au modèle B.

La figure 3.25 présente les différences des dimensions critiques mesurées et simulées pour le cas d'une ligne isolée protégée par des SRAF. D'après cette figure, le modèle B n'est pas assez précis quant à l'estimation des variations des dimensions critiques suite à des variations de l'ouverture numérique. En effet, les variations des dimensions critiques sont surestimées par le modèle B. Ce phénomène est assez fréquent quand il s'agit de simuler des structures critiques comme les structures isolées ou les réseaux de lignes denses de pas réduits. La simulation par le modèle optique pur nous permet, bien évidemment, de quantifier la contribution du modèle optique seul.

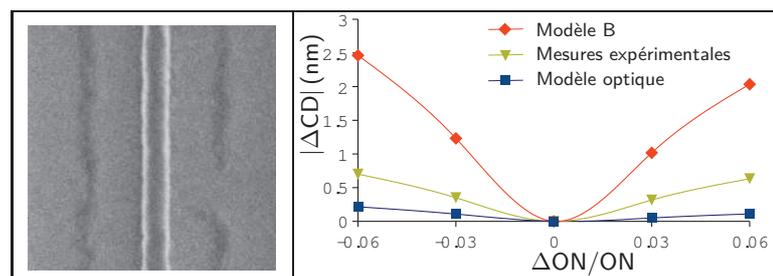


Figure 3.25 – Ligne isolée avec insertion de motifs diffractants non résolus (dimensions sur le dessin du masque : CD 86.5 nm)

La figure 3.26 présente une comparaison des variations de dimensions critiques pour une structure H isolée à bras très rapprochés. Les dimensions critiques simulées par le modèle B pour cette structure sont surestimées par rapport aux mesures expérimentales. Étant donné que les valeurs obtenues avec le modèle B et le modèle optique pur coïncident, cette structure n'est donc pas considérée sensible pour le modèle B ($\beta \approx 0$).

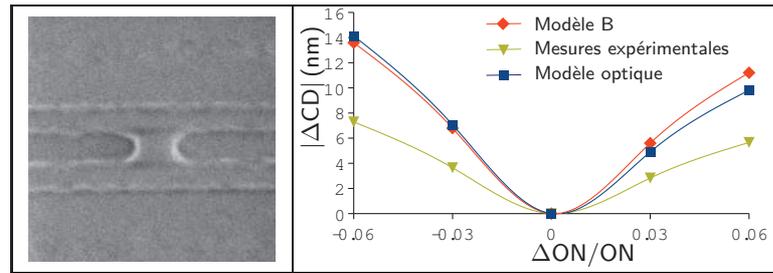


Figure 3.26 – Structure de forme H (dimensions sur le dessin du masque : CD 46 nm, espacement des bras 74 nm)

La figure 3.27 présente une ligne traversant des bouts de lignes denses. Ici, le modèle B est considéré stable ($\beta \approx 0$) pour les valeurs positives de ΔON mais moins stable pour les valeurs négatives ($\beta \approx 0.9$). Les variations de CD calculées par le modèle optique sont souvent plus fidèles à l'expérience pour les structures 1D, contrairement aux motifs 2D. Notre modèle de référence n'est donc pas toujours précis pour bien prédire les variations de CD. Nous verrons dans les chapitres suivants que l'on peut améliorer la précision du modèle optique afin de mieux modéliser divers effets optiques.

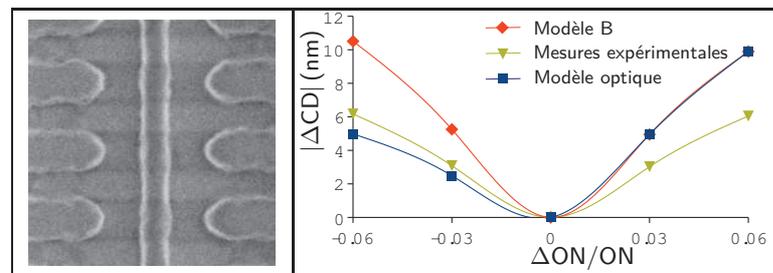


Figure 3.27 – Ligne traversant des bouts de lignes denses (dimensions sur le dessin du masque : CD de la ligne centrale 75 nm, CD des lignes du réseau 96.5 nm, Pas du réseau 258 nm, espacement des bouts de lignes 223 nm)

3.5 Conclusion

Le principal objectif de la calibration d'un modèle empirique est d'obtenir une précision adéquate pour un grand nombre de structures, mais aussi un comportement stable permettant au modèle d'extrapoler en dehors des données expérimentales de calibration. La stabilité des modèles empiriques est un critère difficile à quantifier car nécessitant un effort supplémentaire de mesures expérimentales dans plusieurs conditions décalées du procédé lithographique.

Dans l'objectif d'étudier la stabilité d'un modèle empirique quelconque, nous avons développé une méthode perturbative. Son principe consiste à quantifier la sensibilité d'une structure donnée (en terme de dimensions critiques) par rapport à des petites variations d'un paramètre du procédé. Cette méthode permet de détecter les structures sensibles aux variations d'un paramètre afin de pouvoir les corriger par un traitement OPC spécifique. Elle permet également de définir des intervalles de stabilité d'un modèle par rapport au paramètre concerné. Ces intervalles sont définis par les plages de variation du paramètre dans lesquelles le modèle reste stable. L'application directe de la méthode perturbative nous a permis d'étudier quatre cas de modèles empiriques avancés utilisés en 65 nm et 45 nm respectivement. Les résultats de simulation montrent que ces modèles présentent une forte sensibilité aux variations du plan focal contrairement à celles de l'ouverture numérique et du facteur de cohérence partielle σ_{out} . Cette étude nous a également montré que la précision d'un modèle empirique est toutefois souvent mise en balance avec la stabilité du modèle lui-même. En effet, l'utilisation d'un nombre élevé de termes dans un modèle empirique entraîne une meilleure précision mais dégrade sa stabilité. Au niveau

de la calibration, nous avons démontré que l'utilisation de mesures faites dans les conditions décalées permet d'améliorer la stabilité numérique d'un modèle empirique. Néanmoins, la méthode perturbative dépend fortement de la précision et de la qualité du modèle optique utilisé.

L'étude de la couverture du modèle à seuil variable du nœud technologique 65 nm par l'analyse de son espace de paramètres nous a permis de corréler sensibilité et couverture du modèle et de détecter les structures critiques. À travers cette étude, nous avons pu montrer qu'une structure sensible d'un circuit peut ne pas être bien représentée par l'espace des paramètres du modèle. Finalement, nous avons vérifié la validité de la méthode perturbative à travers des mesures expérimentales effectuées sur des structures de calibration.



Troisième partie

Modélisation des effets
topographiques du masque

Chapitre 4

Méthodes de modélisation des effets topographiques du masque

Afin de gagner en temps de cycle, des approximations sont souvent employées dans la partie optique du modèle OPC. L'approximation de Kirchhoff est la méthode la plus communément utilisée dans la partie optique d'un modèle OPC afin de décrire le champ électrique en sortie du masque. Dans cette méthode, le masque est considéré infiniment fin et est simplement décrit par une unique fonction scalaire de transmission et de phase. Les effets topographiques sont ainsi négligés.

Bien que ces méthodes approchées permettent de diminuer les ressources en mémoire et le temps de calcul, elles constituent une source d'erreurs dans la modélisation de l'image du masque en sortie du système optique. Les erreurs obtenues sont d'autant plus pénalisantes que lorsque les dimensions des motifs approchent la longueur d'onde d'exposition. La dégradation dans la précision du calcul de l'intensité de l'image aérienne implique l'utilisation de modèles empiriques de plus en plus complexes afin de capturer les erreurs optiques résiduelles. L'augmentation de la complexité des modèles empiriques employés dégrade leur stabilité. Ainsi, l'utilisation de modèles empiriques plus simples permet probablement de s'affranchir du problème de l'instabilité. Les approximations effectuées dans la partie optique doivent alors être examinées.

Sommaire

4.1	Introduction	79
4.2	Position du problème	79
4.3	Méthodes rigoureuses	80
4.3.1	La méthode FDTD	80
4.3.2	La méthode RCWA	81
4.4	L'approximation de Kirchhoff	82
4.4.1	Principe	82
4.4.2	Limitations	83
4.4.3	Analyse du champ proche	84
4.4.3.1	Masque étudié	84
4.4.3.2	Cas d'un bord diffractant	84
4.4.3.3	Cas d'une ligne isolée	88
4.4.3.4	Cas d'un réseau de lignes denses	89
4.4.4	Impact sur l'intensité lumineuse	90
4.4.4.1	Contraste	92
4.4.4.2	Position du meilleur focus	93
4.4.4.3	Profondeur de champ	93
4.4.4.4	Effets de proximité	94
4.4.4.5	Effets sur le seuil d'impression des SRAF	95
4.4.5	Discussion des résultats	98
4.5	Méthodes semi-rigoureuses	98
4.5.1	Méthode de décomposition de domaine DDM	98
4.5.2	Méthode de compensation des effets de diffraction M3D	99
4.5.3	Évaluation par rapport au calcul rigoureux	100
4.5.3.1	Définition des conditions de simulation	100
4.5.3.2	Fenêtres de procédé et courbes de Bossung	101
4.5.3.3	Position du meilleur focus	102
4.5.3.4	Profondeur de champ	103
4.5.3.5	Effets de proximité	104
4.5.3.6	Prédiction de l'impression des SRAF	104
4.5.4	Discussion des résultats	105
4.6	Conclusion	107

4.1 Introduction

Les principales théories classiques appliquées aux phénomènes optiques, tels que la diffraction et l'interférence, furent formulées à la fin du 19^{ème} siècle. Dès lors, ces théories subirent une forte opposition et de nombreuses critiques conjointement avec la découverte de l'électrodynamique quantique. Celle-ci, bien qu'elle ait profondément affecté les visions sur la nature de la lumière, a permis de rassembler les efforts pour ré-étudier les limitations des théories classiques et de définir leurs intervalles de validité. En effet, ces théories classiques traitent la propagation de la lumière comme un phénomène scalaire et, par conséquent, négligent sa nature électromagnétique. Car, comme formulé par les équations de Maxwell (1831–1879), le champ électromagnétique doit être caractérisé par ses deux composantes vectorielles couplées, le champ électrique et le champ magnétique. La théorie scalaire ne considère qu'une seule composante radiale, mais elle reste suffisamment correcte sous trois conditions :

- La taille de l'objet doit être grande par rapport à la longueur d'onde de la lumière.
- L'écran d'observation doit être situé à une distance suffisamment grande par rapport à l'objet.
- Les angles d'incidence et de diffraction doivent être suffisamment faibles pour pouvoir négliger la composante axiale du champ électrique.

Dans l'objectif de simplifier le calcul du champ électrique, les modèles de traitement OPC actuels utilisent encore la théorie scalaire sans pour autant remplir ces conditions. En effet, la taille des structures sur le masque pour les nœuds actuels (45 nm et au delà) est de l'ordre de la longueur d'onde, voire plus petite. Pour les systèmes optiques actuels opérant en immersion, l'ouverture numérique devient grande (supérieure à 1) et les angles de diffraction sont de moins en moins petits. Comme l'image est projetée à l'infini dans le système optique de Köhler (voir définition en annexe A), l'approximation de Fraunhofer (1787–1826) est utilisée pour calculer le champ scalaire en sortie de l'optique de projection. Cette approximation néglige les effets de l'incidence oblique (régime paraxial) puisqu'elle suppose que le front d'onde est plan. Ainsi, elle reste précise tant que l'objet est grand devant la longueur d'onde et que l'ouverture numérique du scanner est réduite.

À l'heure actuelle, les dimensions des motifs du masque (l'objet) deviennent équivalentes à la longueur d'onde d'exposition et l'ouverture numérique du scanner ne cesse d'augmenter après chaque génération lithographique. Dès lors, la modélisation de l'image du masque dans la résine par la théorie scalaire est insuffisante pour garantir une précision respectant les spécifications de la feuille de route de l'ITRS. Bien que les erreurs engendrées par l'utilisation de l'approche scalaire dans la modélisation du système optique puissent être réduites par la partie résine du modèle [78], cette manière ne fait que dégrader la stabilité du modèle et le rendre plus dépendant aux paramètres du procédé lithographique.

Afin d'améliorer la stabilité d'un modèle empirique, les approximations effectuées dans la partie optique doivent être examinées. Dans ce chapitre, nous allons étudier les limites de validité de l'approximation de Kirchhoff, méthode qui, jusqu'à présent, est la plus utilisée dans la modélisation du champ proche d'un masque. Cette étude est illustrée par la comparaison des résultats de simulation de différents cas de structures en utilisant cette approximation (Kirchhoff) et une méthode rigoureuse pour le calcul du champ proche (la méthode FDTD). Ensuite, nous présenterons un état de l'art des méthodes approchées semi-rigoureuses disponibles dans les logiciels de traitement OPC et étudierons leurs avantages et inconvénients.

4.2 Position du problème

Le problème qui surgit aujourd'hui dans la modélisation de l'image du masque à travers un système optique vient des limitations que présente l'approche scalaire et des conditions non respectées. Nous pouvons résumer ce problème en deux points :

- La définition de l'objet : la topographie du masque n'est pas prise en compte dans le calcul du champ électromagnétique.
- La définition de l'onde incidente : l'incidence oblique des rayons incidents n'est pas prise en compte dans le calcul de l'image du masque.

En effet, la théorie scalaire ne permet pas de calculer le champ électromagnétique proche du masque. Elle considère uniquement que ce dernier est une fonction scalaire définie par une transmission et un déphasage bien déterminés. Cependant, comme le montre le schéma de la figure 4.1, le masque possède une topographie qui peut être décrite par un ensemble de paramètres (les épaisseurs des couches de l'empilement, les indices optiques des matériaux, les arrondissements des coins etc). Le calcul du champ électromagnétique en sortie du masque nécessite la propagation de l'onde à l'intérieur de la topographie.

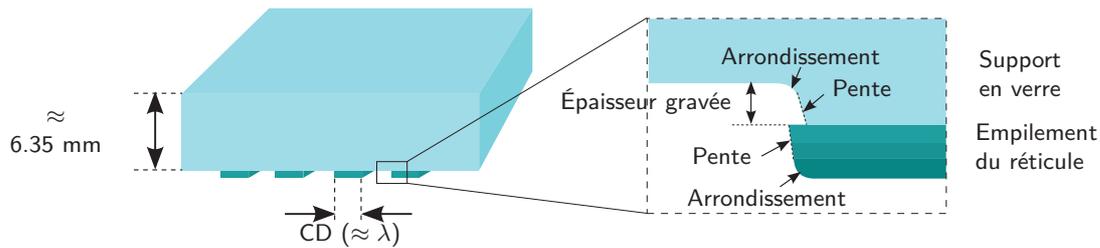


Figure 4.1 – Schéma d'un masque utilisé en lithographie optique par projection (à gauche) et représentation simplifiée de l'empilement multicouche du réticule permettant de définir les paramètres de la topographie (à droite).

En ce qui concerne l'incidence oblique, la théorie scalaire suppose que les angles des rayons incidents sont suffisamment faibles pour pouvoir négliger la composante axiale du champ. Ceci correspond à considérer que le front d'onde est plan et que l'incidence des rayons est normale. Comme conséquences à cette hypothèse, les amplitudes des ordres diffractés à travers le masque restent constantes quelque soit l'angle d'incidence. Il suffit donc de calculer les ordres de diffraction pour le point central de la source d'illumination et le décaler dans l'espace des fréquences pour obtenir le spectre de diffraction des autres points. Cependant, l'apparition de la composante axiale du champ électromagnétique entraîne la réduction des amplitudes des autres composantes. Plus l'angle d'incidence est grand, plus l'amplitude de la composante axiale est élevée et moins elle est pour les autres. Il en résulte que l'intensité lumineuse finale dans la résine est impactée, entraînant de ce fait une mauvaise estimation des critères de performances du procédé lithographique (CD, DOF, MEEF, etc).

Dans l'objectif de prendre en compte les effets de la topographie du masque, il existe des méthodes rigoureuses de calcul permettant de résoudre les équations de Maxwell dans un domaine donné. Nous verrons plus loin que le champ électrique que l'on obtient par des méthodes rigoureuses est bien différent de celui calculé avec la théorie scalaire. Nous allons maintenant présenter brièvement deux méthodes rigoureuses les plus répandues dans la modélisation électromagnétique.

4.3 Méthodes rigoureuses

4.3.1 La méthode FDTD

La méthode des différences finies dans le domaine temporel (de l'anglais "*Finite Difference Time Domain*" ou FDTD), introduite par Yee en 1966, est une méthode rigoureuse pour la modélisation des ondes électromagnétiques dans un domaine quelconque. Son principe consiste à résoudre les équations de Maxwell régissant la propagation des champs électrique et magnétique par l'intermédiaire d'une méthode astucieuse des différences finies. Cette méthode permet d'exprimer un champ électromagnétique à partir d'un développement limité de second ordre

permettant de relier ses première et seconde dérivées partielles. La première étape consiste à effectuer un maillage spatial de l'objet selon une grille tridimensionnelle et à définir un plan d'excitation. Puis, un processus itératif permet de calculer les champs électriques à un instant t à partir des champs magnétiques obtenus aux instants $t - \frac{1}{2}\Delta t$ et $t + \frac{1}{2}\Delta t$ et réciproquement. C'est pendant ce processus que les différences finies sont utilisées. Après chaque itération, une impulsion constante vient se rajouter aux champs calculés au niveau du plan d'excitation. Le calcul s'arrête lorsque les champs convergent vers une solution unique que l'on retiendra. Pour plus d'informations sur cette méthode, une petite présentation est donnée en annexe E de ce manuscrit. Le lecteur peut également consulter l'un de ses ouvrages [79–81].

L'avantage de la méthode FDTD est son caractère général. Elle permet de suivre l'évolution du champ électromagnétique en tout point de la grille de calcul et dans le temps. Ainsi, par exemple, nous pouvons facilement suivre, depuis le régime transitoire jusqu'au régime stationnaire, le champ électromagnétique ou la courbe d'absorption dans un empilement multicouche de plusieurs matériaux. Cependant, bien qu'elle soit exacte, la méthode FDTD nécessite un maillage très fin (le pas de grille doit être inférieur à $\lambda/20$) afin de converger vers une solution unique. De ce fait, elle requiert une capacité de calcul importante et entraîne des temps de calcul élevés. Ce qui est le cas d'une modélisation d'un masque de lithographie optique. En effet, la discrétisation d'un masque nécessite un maillage de l'ordre de quelques nanomètres afin de garantir une précision adéquate. Mais vu les dimensions d'un masque (de l'ordre d'une dizaine de centimètres), la méthode FDTD n'est pas utilisable compte tenu des contraintes de temps et de la taille gigantesque en mémoire requise. Il est donc judicieux de préférer des méthodes approchées.

Dans la suite de cette thèse, nous utiliserons la méthode FDTD, disponible dans le logiciel de simulation électromagnétique EM-Suite™ de la société Panoramic, comme méthode de référence afin de comparer différentes méthodes de modélisation.

4.3.2 La méthode RCWA

La RCWA (de l'anglais "*Rigorous Coupled Wave Analysis*"), également connue sous le nom de MMFE (de l'anglais "*Modal Method by Fourier Expansion*"), est une méthode rigoureuse de calcul électromagnétique qui a été mise au point par M.G. Moharam et T.K. Gaylord au début des années 1980 pour la modélisation de réseaux plans diélectriques ou métalliques modulés sinusoidalement [82,83]. Elle consiste en une décomposition en série de Fourier de la permittivité électrique de l'objet diffractant et de résoudre les équations de Maxwell que l'on projette dans une base de Floquet-Fourier. Elle fut généralisée aux réseaux présentant des reliefs de surface [84] puis au cas tridimensionnel d'incidence conique [85]. La méthode a également été étendue aux cas de réseaux anisotropes [86]. Les deux chercheurs ont tiré profit de la décomposition en série de Fourier de la permittivité initiée par Burckhardt [87], Kaspar [88] et Knop [89] (méthode BKK), en y ajoutant une description du champ dans la région modulée sous la forme d'ondes couplées se propageant selon l'axe z et correspondant aux ordres diffractés hors du réseau. De plus l'approximation, alors courante, sur la dérivée seconde du champ, n'était plus considérée.

L'algorithme de la méthode RCWA requiert que l'objet soit de forme périodique (par exemple : un réseau de lignes). Dans ce cas particulier, une décomposition en ondes planes et la résolution d'un système aux valeurs propres issues des équations de Maxwell permettent de calculer les champs électrique et magnétique. Dans le cas d'un objet de forme quelconque, l'objet est découpé en réseaux élémentaires et le calcul des champs électromagnétiques se fait de la même manière que pour un objet périodique. Puis, une écriture sous forme matricielle des conditions aux limites permet de relier les différents champs entre-eux.

L'avantage de la méthode RCWA par rapport à d'autres méthodes de modélisation est sa rapidité de calcul lorsque les structures étudiées sont périodiques. Par contre, la non-périodicité de l'objet représente son inconvénient majeur. De plus, la modélisation de la composante TM du champ électromagnétique peut être instable et nécessite des temps de calcul plus longs.

Dans les années 90, de nouvelles méthodes de calcul ont été proposées dans la littérature afin de résoudre le problème de la non-convergence de la composante TM [90,91]. Puis, plus récemment, le problème de la non-périodicité a été résolu par l'introduction d'une périodicité virtuelle et l'incorporation de milieux artificiels entre les bords des cellules élémentaires périodiques [92]. La méthode RCWA a été finalement généralisée aux milieux non-périodiques avec un remarquable succès pour les structures 2D d'épaisseurs finies [93–97]. D'autres techniques de factorisation rapide ont été également proposées afin d'accélérer les temps de calcul [98,99]. Aujourd'hui, la méthode RCWA reste très avantageuse par rapport à d'autres techniques en terme de temps et de ressources requises pour le calcul.

4.4 L'approximation de Kirchhoff

La difficulté de résoudre les équations de Maxwell a contribué au développement de nouvelles méthodes approchées pour la modélisation de phénomènes électromagnétiques, principalement dérivées de la théorie scalaire de la diffraction et de l'optique géométrique. Ces méthodes se sont rapidement développées au milieu du siècle écoulé, entraînées par la vague des télécommunications et la nécessité de mieux comprendre la transmission des signaux¹. Parmi ces méthodes, nous pouvons citer, celle qui va nous intéresser dans ce mémoire, l'approximation de Kirchhoff.

L'origine de l'approximation de Kirchhoff remonte aux hypothèses effectuées par le physicien allemand Gustav Robert Kirchhoff (1824–1887) pour calculer l'onde diffractée par une fente située dans un plan opaque supposé infini (cf. Annexe D). Kirchhoff supposa que la valeur de l'onde et de ses dérivées partielles dans le plan de la fente sont égales à celles calculées en l'absence de l'écran et qu'elles sont nulles en tout point de celui-ci. Ces hypothèses sont appelées les conditions aux limites de Kirchhoff et consistent à une idéalisation du champ scalaire de la fente.

Bien que cette méthode nous permette d'appréhender le problème de la diffraction d'une manière simple, son domaine d'application n'en reste pas moins très limité. En effet, l'approximation de Kirchhoff a attiré l'attention des chercheurs et a fait l'objet de plusieurs publications scientifiques. Dans la littérature, nous pouvons citer les travaux de Ripoll et al. [103] qui a étudié la diffraction de la lumière diffuse dans des milieux arbitraires en utilisant une forme modifiée de la fonction de Green. Soto-Crespo [104] comme Pinel [105] et autres l'ont utilisée pour résoudre des problèmes de diffraction par des surfaces arbitraires très rugueuses. D'autres publications ont proposé des techniques pour l'extension asymptotique [106] et l'utilisation des termes de seconds ordres dans l'approximation de Kirchhoff [107] pour l'étude de problèmes plus complexes tels que la diffraction par des objets transparents non-homogènes.

4.4.1 Principe

Le principe de cette méthode consiste à considérer que le masque est idéal et qu'il est entièrement défini par une fonction complexe constante définissant la transmission et la phase des motifs dessinés sur le masque. Si l'on considère que le champ électrique incident est une onde plane d'amplitude et de phase égales à A_0 et φ_0 respectivement :

$$E_i(x, y) = A_0 e^{j\varphi_0},$$

nous pouvons alors définir le champ électrique à la sortie du masque comme suit :

1. À cette époque, les sujets ayant attiré l'attention des chercheurs étaient variés. Par exemple, nous pouvons citer les problèmes liés à la propagation d'une onde dans un milieu non-homogène, un sujet qui a intéressé Bremmer [100] au début des années 1950 où il a utilisé l'approximation de W.K.B. (Wentzel-Kramers-Brillouin). Durant les années 1960, Jones s'est intéressé aux effets d'un obstacle sur la propagation d'une onde électromagnétique [101]. Il a étudié la diffraction par des objets sans bords en utilisant une méthode approchée asymptotique pour l'évaluation de l'intégrale de Kirchhoff (voir annexe D). Bouwkamp a présenté quelques méthodes approchées pour la résolution de problèmes de diffraction et discuté leurs domaines de validité [102].

$$E(x, y) = \begin{cases} A_0 e^{j\varphi_0} & \text{en champ clair} \\ Ae^{j\varphi} & \text{en champ obscur} \end{cases} \quad (4.1)$$

avec

$$A = \sqrt{T}A_0 \quad \text{et} \quad \varphi = \varphi_0 + \Delta\varphi$$

sont l'amplitude et la phase du champ sortant du masque, T et $\Delta\varphi$ étant le taux de transmission et le décalage de phase de la région sombre du masque.

L'application directe de cette approximation nous permet d'adopter l'approximation du masque fin TMA (de l'anglais "*Thin Mask Approximation*") qui considère que le masque est un objet infiniment fin puisqu'elle nous permet aussi d'appliquer les conditions aux limites de Kirchhoff.

4.4.2 Limitations

À l'heure actuelle, l'approximation de Kirchhoff est la méthode la plus communément employée dans les modèles optiques en OPC pour la modélisation de champ électrique au niveau du masque. Cette méthode définit le masque par une fonction scalaire constante et possède l'avantage d'être simple et facile à implémenter. Elle permet également de réduire le temps de calcul et les ressources en mémoire requises pour simuler un masque complet. Toutefois, cette méthode est scalaire et présente des limitations par rapport aux méthodes rigoureuses qui utilisent un calcul vectoriel. En effet, cette méthode nous permet de négliger les effets liés à la topographie du masque qui sont principalement :

- La diffraction de l'onde au niveau des bords des structures dessinées sur le masque. En effet, sur la figure 4.2 (a), nous pouvons remarquer que le champ électrique calculé en utilisant l'approximation de Kirchhoff est maintenu constant à une position z quelconque à l'intérieur du masque. En se propageant à l'intérieur de l'empilement du masque, l'onde électromagnétique ne subit aucune diffraction par les bords et elle est maintenue constante jusqu'à la sortie, où l'on peut remarquer une distorsion de l'amplitude de l'onde causée par la diffraction. Dans le cas d'un calcul effectuée avec la méthode rigoureuse FDTD, la figure 4.2 (b) nous montre clairement que l'onde à l'intérieur du masque subit une diffraction le long des bords et que son amplitude varie le long de l'axe z . Par conséquent, l'amplitude et la phase de l'onde à la sortie du masque sont différentes de celles que l'on a calculées en utilisant l'approximation de Kirchhoff.

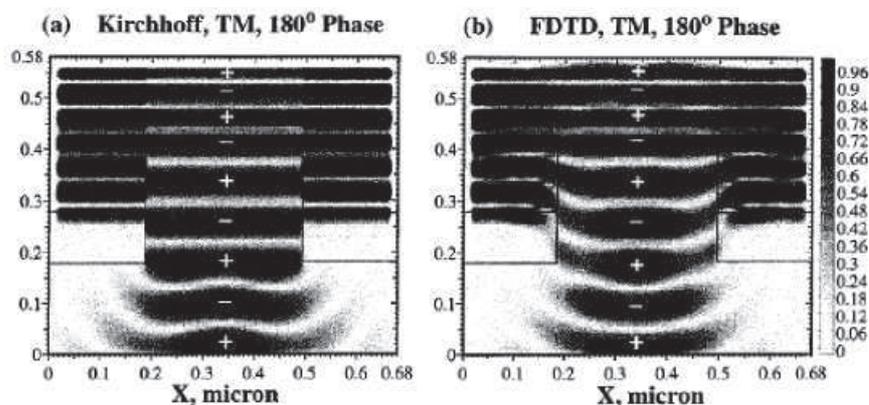


Figure 4.2 – Amplitude instantanée du champ électrique au voisinage d'une ouverture de 2λ par 12λ d'un masque déphasé de 180° , calculé en incidence normale avec $\lambda = 157\text{nm}$. Les symboles $+$ et $-$ représentent les signes du vecteur du champ électrique polarisé en TM. Ce dernier est calculé par a) l'approximation de Kirchhoff et b) la méthode rigoureuse FDTD. Le masque simulé ici est de type binaire (en Chrome). Source : [108].

- L'absence du couplage entre les deux champs magnétique et électrique. En effet, l'onde est scalaire et est entièrement définie par une fonction complexe unique. Elle peut caractériser

à la fois le champ électrique et le champ magnétique. Ceci veut dire que si l'on veut considérer les deux champs ensemble, nous aurons alors à définir deux fonctions scalaires qui se propagent indépendamment de l'autre. Ce qui est inutile puisqu'au final, l'intensité est calculée à partir du champ électrique uniquement.

- L'effet de l'incidence oblique sur la polarisation du champ électrique n'est pas pris en compte. En effet, la topographie du masque peut modifier la polarisation de l'onde qui le traverse. Dans d'autres études, nous avons montré qu'une onde polarisée en TE arrivant avec une incidence oblique sur un masque à décalage de phase aura une composante TM non nulle en sortie [109]. Ce qui n'est pas le cas d'une onde scalaire dans l'approximation de Kirchhoff.

4.4.3 Analyse du champ proche

4.4.3.1 Masque étudié

Le masque que nous allons utiliser dans la suite de cette partie est un PSM atténué avec une transmission et un déphasage valant 6% et 180 degrés respectivement. L'empilement du masque est constitué d'un support en verre d'épaisseur 6.35 mm avec un réticule constitué d'une couche de MoSi oxydée en surface. Une faible épaisseur du verre a été gravée afin de corriger le déphasage de 180 degrés du masque comme le montre la figure 4.3. Le tableau 4.1 récapitule les différents paramètres de la topographie du masque.

Matériau	Épaisseur	Indice de réfraction n (u.a.)	Indice d'extinction k (u.a.)
Support en verre	6.35 mm	1.563	0.0
Verre gravé	4 nm	1.563	0.0
MoSi	64 nm	2.440	0.6
Oxyde	4 nm	2.240	0.4

Tableau 4.1 – Épaisseurs et propriétés optiques des matériaux constituant l'empilement du masque

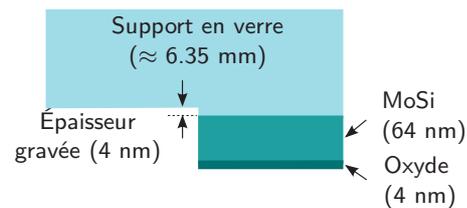


Figure 4.3 – Topographie utilisée pour un masque PSM à 6% de transmission et à décalage de phase égal à 180 degrés.

4.4.3.2 Cas d'un bord diffractant

La figure 4.4 présente les courbes de l'amplitude et de phase en champ proche d'un bord diffractant. Les valeurs de ces courbes ont été obtenues en utilisant l'approximation de Kirchhoff et la méthode rigoureuse FDTD où l'on a représenté les deux composantes TE et TM du champ électrique. D'après cette figure, nous pouvons clairement constater que les courbes calculées par la méthode rigoureuse présentent des oscillations qui s'amplifient à proximité du bord. Le comportement oscillatoire du champ électrique est le résultat de la diffraction de l'onde électromagnétique au niveau du bord du masque et l'interférence des rayons diffractés. Le but ici est de savoir à quelles fréquences spatiales ces oscillations correspondent-elles et dans quelles mesures pouvons-nous les négliger.

Nous avons tracé sur la figure 4.5 (a) le spectre de l'amplitude du champ électrique d'un bord diffractant dans l'approximation de Kirchhoff. Le nombre m , qui représente l'abscisse du spectre, correspond à l'indice de Bragg défini par

$$m = \frac{k_x \cdot L_x}{2\pi} = \frac{\sin \theta_x}{\lambda} \cdot L_x, \quad (4.2)$$

k_x et L_x étant la composante du vecteur d'onde selon x et la dimension du masque respectivement. De la même manière, nous avons calculé ce même spectre d'amplitude en utilisant la méthode rigoureuse FDTD et tracé la différence des deux spectres obtenus sur la figure 4.5-b). Cette dernière nous montre clairement que l'erreur sur l'amplitude causée par l'approximation

de Kirchhoff augmente lorsque le nombre m augmente. Ceci veut dire que les pseudo-oscillations observées sur la figure 4.4-a) se traduisent par des harmoniques à fortes fréquences spatiales dans le spectre de Fourier. Notons ici que ces hautes fréquences spatiales sont partiellement filtrées par la pupille du système optique. Mais en augmentant l'ouverture numérique du système, une partie des hautes fréquences peut passer à travers la pupille et les erreurs sur l'amplitude et la phase deviennent non négligeables.

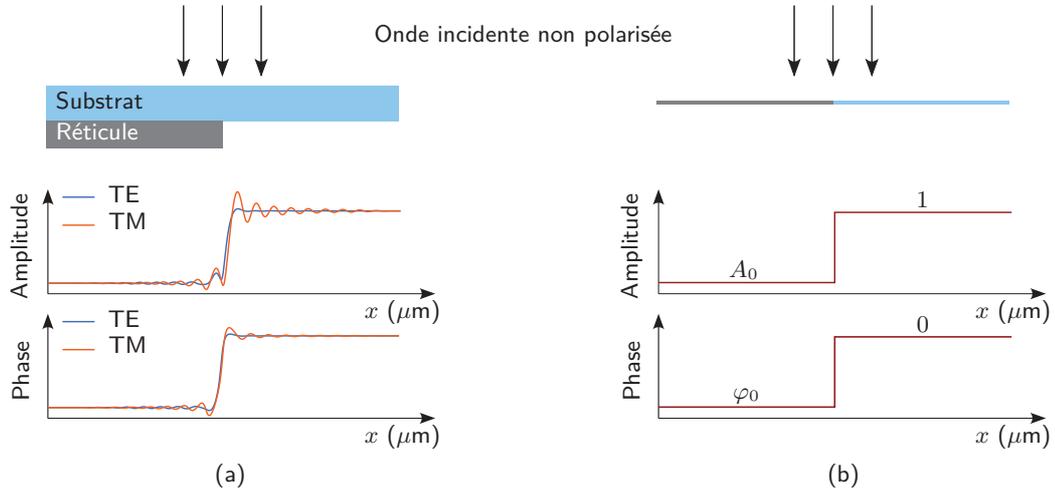


Figure 4.4 – Comparaison de l'amplitude et de la phase du champ électrique en sortie du masque, calculé en utilisant (a) la méthode rigoureuse FDTD et (b) l'approximation de Kirchhoff. Le masque utilisé est atténué à 6% de transmission et à décalage de phase égal à 180 degrés.

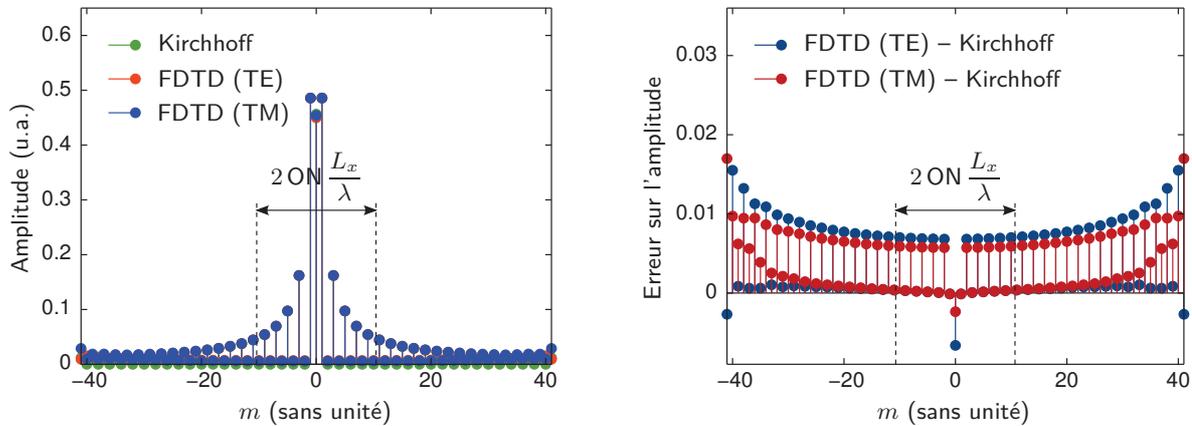


Figure 4.5 – Spectres de Fourier de l'amplitude du champ électrique d'un bord diffractant isolé calculés avec l'approximation de Kirchhoff et la méthode rigoureuse FDTD (à gauche) et différences entre les ordres calculés par FDTD et l'approximation de Kirchhoff pour les deux composantes TE et TM (à droite). m étant l'indice de Bragg défini par $\frac{k_x \cdot L_x}{2\pi}$, k_x et L_x étant la composante du vecteur d'onde et la dimension du masque selon l'axe x respectivement.

Ce que nous pouvons constater également sur la figure 4.5 est que contrairement aux ordres ± 1 où l'erreur en amplitude est pratiquement nulle, l'ordre zéro présente une sur-estimation de l'amplitude avec l'approximation de Kirchhoff par rapport à la méthode rigoureuse FDTD. Ceci entraîne une estimation erronée du contraste de l'image aérienne au niveau de la plaque pouvant introduire un décalage en seuil d'extraction des dimensions critiques.

Nous avons calculé, en incidence normale, le champ électrique en sortie du masque par la méthode rigoureuse FDTD. La figure 4.6 présente les profils de l'amplitude et de la phase des composantes TE et TM du champ obtenu. Les pseudo-périodes des oscillations (les distances entre différents maxima et minima consécutifs de l'amplitude et de la phase) du champ électrique

sont de l'ordre de la longueur d'onde ($\lambda \approx 193nm$). Ce qui correspond bien à la taille des motifs dessinés sur le masque (à l'échelle du masque, c'est-à-dire 4 fois la taille sur la plaquette) pour le nœud technologique 45 nm. L'approximation du bord idéal de Kirchhoff ne peut donc être utilisée puisque l'erreur sur l'amplitude de l'onde transmise devient importante (l'erreur maximale sur l'amplitude vaut 20%).

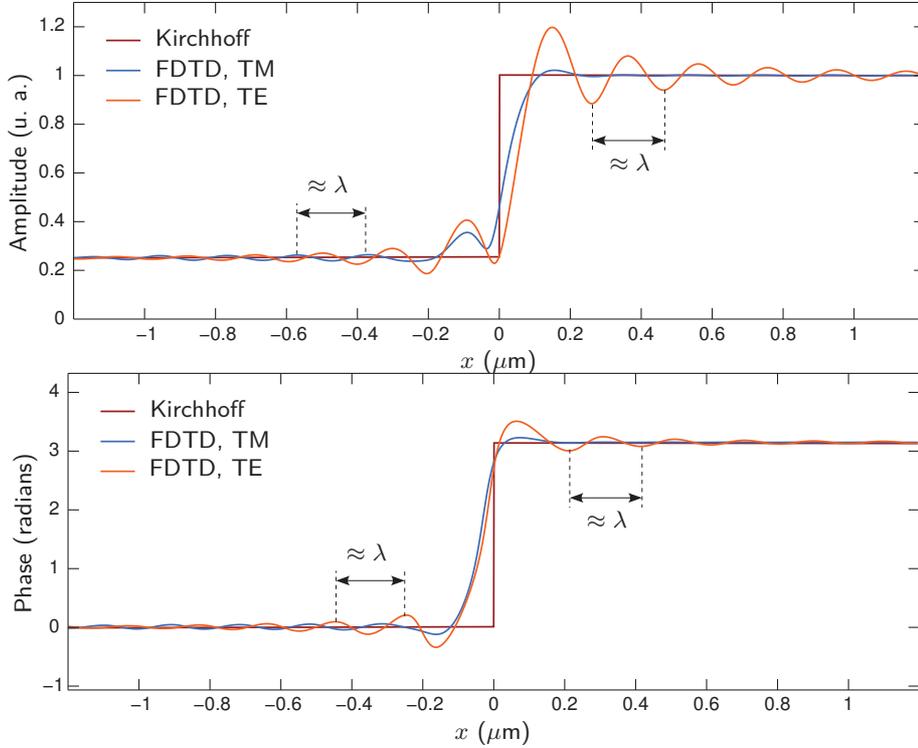


Figure 4.6 – Tracés de l'amplitude et de la phase des composantes TE et TM du champ électrique à proximité d'un bord diffractant calculées par la méthode rigoureuse FDTD avec une incidence normale. Le masque est atténué à 6% de transmission et à décalage de phase de 180 degrés. Le plan d'observation a été fixé à 50 nm du masque.

Les pseudo-oscillations observées sur la figure 4.6 sont le résultat de la diffraction de l'onde au niveau du masque ainsi que l'interférence entre les ondes transmises ayant effectué une ou plusieurs réflexions à travers les différentes interfaces de la topographie. Nous allons maintenant justifier les valeurs des pseudo-périodes des oscillations qui sont proches de la longueur d'onde λ . Si nous décrivons le champ proche du bord par la théorie scalaire de la diffraction, et plus particulièrement par le principe d'Huygens-Fresnel (voir annexe D), l'amplitude de l'onde diffractée est la somme des contributions de tout point élémentaire du front d'onde. En utilisant les zones de Fresnel, l'amplitude de l'onde reçue en un point quelconque P est égale à la moitié de celle émise par la première zone de Fresnel selon l'équation (voir démonstration en annexe D) :

$$U(P) = \frac{U_1(P)}{2} \quad (4.3)$$

avec

$$U_p(P) = 2j\lambda A \frac{e^{jk(r+b)}}{r+b} K_p (-1)^{p+1}$$

Considérons le schéma de la figure 4.7, l'amplitude de l'onde reçue au point P est égale à :

$$U(P) = \frac{U_1(P)}{2} = j\lambda A \frac{e^{jk(r+b)}}{r+b} K_1. \quad (4.4)$$

De la même façon, nous exprimons l'amplitude reçue au point Q par :

$$U(Q) = \frac{U_1(Q)}{2} = j\lambda A \frac{e^{jk(r+b(x))}}{r+b(x)} K_1, \quad (4.5)$$

Cette dernière expression s'écrit ainsi :

$$U(Q) = U(P) \frac{r+b}{r+b(x)} e^{jk(b(x)-b)} \quad (4.6)$$

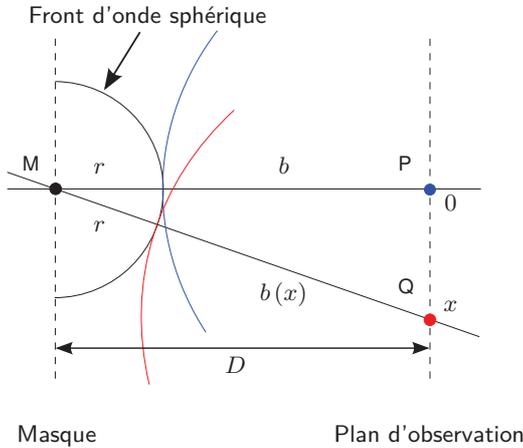


Figure 4.7 – Représentation de la diffraction au niveau du masque en utilisant les zones de Fresnel. L'amplitude de l'onde reçue en P est égale à la moitié de celle de l'onde émise par la première zone de Fresnel définie par l'intersection du front d'onde sphérique de rayon r avec la sphère de rayon $b + \frac{\lambda}{2}$. Pour le point Q, la première zone de Fresnel est définie par l'intersection du même front d'onde avec la sphère de rayon $b(x) + \frac{\lambda}{2}$.

Cherchons maintenant les valeurs de x telles que l'amplitude au point Q est en phase ou en opposition de phase par rapport au point P. D'après l'équation 4.6, cette condition doit satisfaire :

$$k(b(x) - b) = \frac{2\pi}{\lambda}(b(x) - b) = p\pi, \quad p \in \mathbb{Z}$$

qui s'écrit également :

$$b(x) = b + p\frac{\lambda}{2}, \quad p \in \mathbb{Z}. \quad (4.7)$$

En utilisant $(b(x) + r)^2 = ((b+r)^2 + x^2)$ et $D = r+b$, les solutions en x telles que le déphasage entre P et Q est égal à $p\pi$, $p \in \mathbb{Z}$ s'expriment ainsi :

$$x_p = \pm \sqrt{p\lambda D + p^2 \frac{\lambda^2}{4}}, \quad p \in \mathbb{Z}. \quad (4.8)$$

Les pseudo-périodes des oscillations sont donc données par :

$$T = x_{p+2} - x_p, \quad p \in \mathbb{Z} \quad (4.9)$$

Pour un plan d'observation placé à une distance D égale à 50 nm du masque, nous avons tracé sur la figure 4.8 les valeurs des pseudo-périodes des oscillations de l'onde diffractée. D'après cette figure, la pseudo-période décroît rapidement vers la longueur d'onde λ égale à 193 nm. Ce résultat théorique justifie les valeurs des pseudo-périodes observées sur la figure 4.6 et qui sont proches de la longueur d'onde λ d'exposition. L'équation 4.6 justifie également l'amortissement de l'amplitude des composantes TE et TM du champ électrique. Nous avons tracé sur la figure 4.9 l'allure de la fonction théorique d'amortissement $g(x)$ telle que

$$g(x) = \frac{D}{\sqrt{D^2 + x^2}}. \quad (4.10)$$

Cette fonction dépend de la position du plan d'observation D . Le champ électrique calculé par la méthode rigoureuse FDTD dépend donc de la position du plan d'observation. Elle montre également que l'amortissement est plus rapide lorsque D est petit et réciproquement. Le calcul de l'image aérienne doit donc tenir compte du décalage en focus qui résulte du choix de cette position. Dans la suite de ce chapitre, nous avons choisi une distance D égale à 50 nm à partir de la sortie du masque pour tous les résultats de simulation que nous présenterons. Ce choix n'influe guère sur le calcul de l'intensité lumineuse en sortie puisque nous avons pris en compte le décalage en focus résultant.

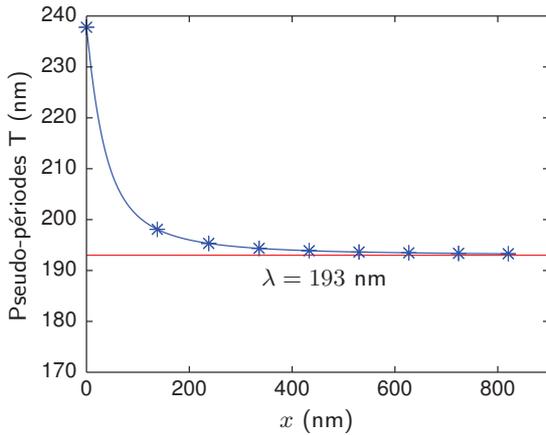


Figure 4.8 – Courbe des pseudo-périodes des oscillations de l'onde diffractée en fonction de la distance x (en nm, à l'échelle du masque), le plan d'observation étant placé à une distance D égale à 50 nm à partir du masque. Les pseudo-périodes ont été calculées en utilisant l'équation 4.9.

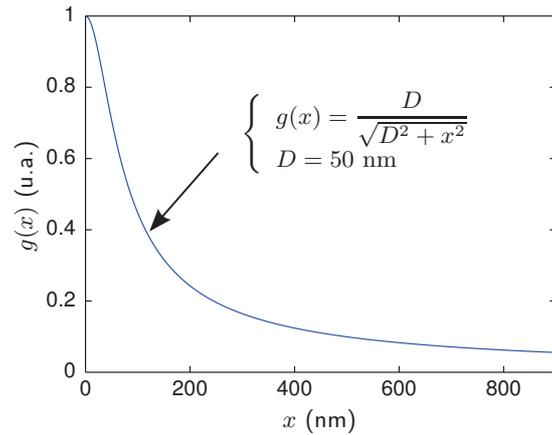


Figure 4.9 – Fonction théorique d'amortissement des oscillations de l'onde diffractée en fonction de la distance x (en nm, à l'échelle du masque), le plan d'observation étant placé à une distance D égale à 50 nm à partir du masque. La fonction $g(x)$ est donnée par l'équation 4.10.

Finalement, bien que l'on ait utilisé les zones de Fresnel qui nécessitent des valeurs de b et r grandes devant la longueur d'onde λ , les résultats obtenus sont valables pour des faibles distances D . En effet, les expressions de l'amplitude de l'onde aux points P et Q (voir equations 4.4 et 4.5) sont celles d'une onde sphérique qui émane du point M et que l'on calcule à des distances égales à $r + b$ et $r + b(x)$ respectivement. L'expression de l'amplitude au point Q en fonction de celle au point P donnée par l'équation 4.6 reste donc valable pour la suite du raisonnement.

4.4.3.3 Cas d'une ligne isolée

Nous avons simulé en incidence normale les champs électriques en sortie du masque d'une ligne isolée ($CD = 75$ nm) avec la méthode rigoureuse FDTD et l'approximation de Kirchhoff. Le masque que nous avons utilisé est atténué à 6% de transmission et à déphasage égal à 180 degrés. La figure 4.10 présente les profils obtenus en amplitude et en phase pour les deux composantes TE et TM ainsi que le profil idéal de Kirchhoff.

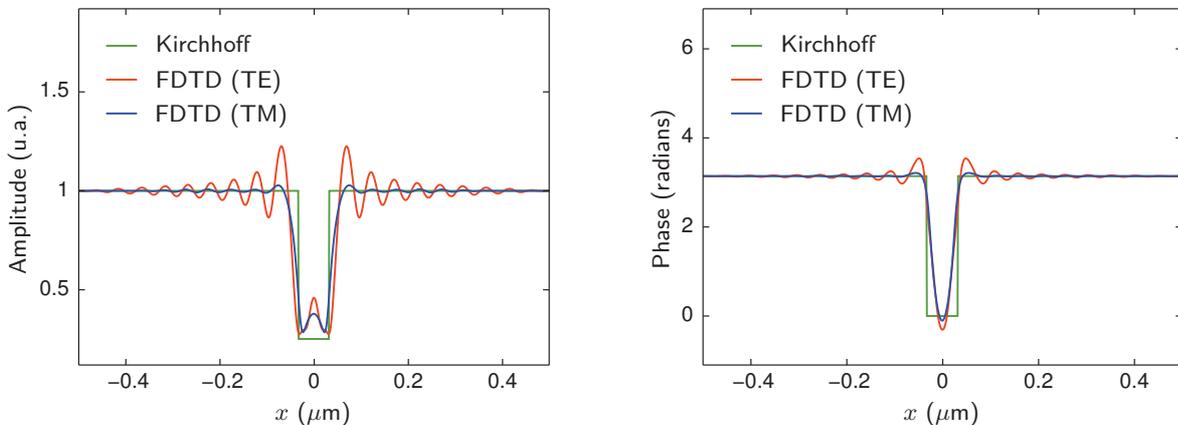


Figure 4.10 – Tracés des amplitudes (à gauche) et des phases (à droite) des champs électriques d'une ligne isolée ($CD = 75$ nm) calculés en sortie du masque avec l'approximation de Kirchhoff et la méthode rigoureuse FDTD (composantes TE et TM) en incidence normale. Le masque utilisé est un masque atténué à 6% de transmission et à décalage de phase égal à 180 degrés.

D'après la figure 4.10, nous retrouvons les mêmes pseudo-oscillations avec des amplitudes qui s'amortissent de la même manière que dans le cas du bord diffractant isolé. À l'intérieur de la ligne, les amplitudes des composantes TE et TM sont légèrement plus élevées que dans le cas de l'approximation de Kirchhoff.

Afin de mieux visualiser ces écarts, nous avons tracé sur la figure 4.11 les spectres de Fourier des deux composantes TE et TM et du profil idéal de Kirchhoff ainsi que les différences entre ces spectres. Cette figure nous montre que les écarts en amplitude sont du même ordre de grandeur que ceux obtenus dans le cas du bord diffractant isolé.

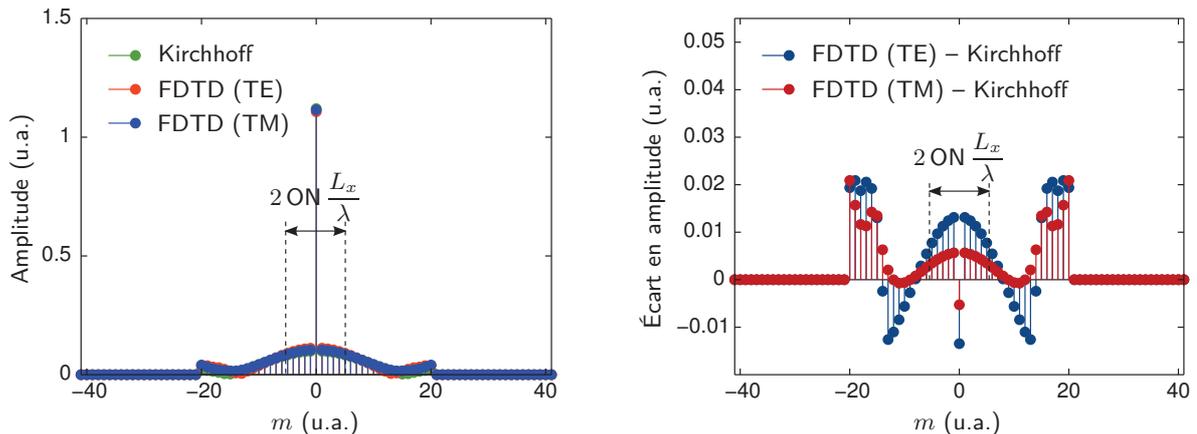


Figure 4.11 – Tracés des spectres de Fourier (en amplitude) des champs électriques d'une ligne isolée ($CD = 75$ nm) calculés en sortie du masque avec l'approximation de Kirchhoff et la méthode rigoureuse FDTD (composantes TE et TM) en incidence normale (à gauche) et de l'écart d'amplitude entre les spectres des deux composantes TE et TM par rapport à celui obtenu par l'approximation de Kirchhoff (à droite).

4.4.3.4 Cas d'un réseau de lignes denses

Nous nous intéressons maintenant aux champs électriques obtenus pour un réseau de lignes denses de pas égal à 180 nm et de dimension CD égale à 75 nm. La figure 4.12 illustre une comparaison des champs électriques obtenus par la méthode rigoureuse FDTD (composantes TE et TM) et l'approximation de Kirchhoff.

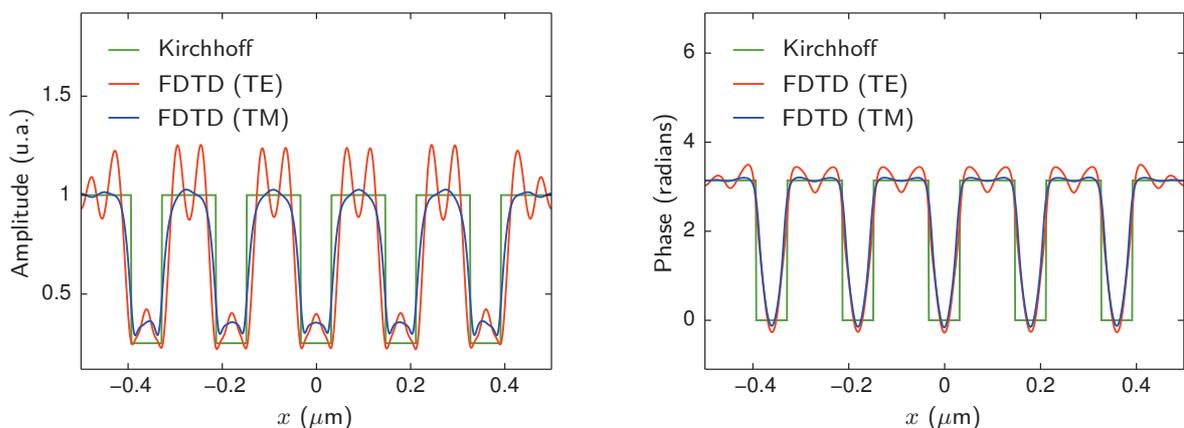


Figure 4.12 – Tracés des amplitudes (à gauche) et des phases (à droite) des champs électriques d'un réseau de lignes de pas égal à 180 nm ($CD = 75$ nm) calculés en sortie du masque avec l'approximation de Kirchhoff et la méthode rigoureuse FDTD (composantes TE et TM) en incidence normale. Le masque utilisé est un masque atténué à 6% de transmission et à décalage de phase égal à 180 degrés. La position x est donnée à l'échelle de la plaquette.

Au niveau de l'amplitude, comme pour la phase, nous pouvons clairement remarquer que les différences entre le calcul rigoureux et le profil idéal de Kirchhoff sont importantes. À l'échelle du masque, l'espacement entre les lignes du réseau ($4 \times 105 \text{ nm}$) correspond approximativement à deux fois la longueur d'onde λ . D'où, l'existence des deux pseudo-périodes sur les courbes de l'amplitude et de la phase des composantes TE et TM.

Pour ce même réseau de lignes, nous avons tracé sur la figure 4.13 les spectres de Fourier des amplitudes des composantes TE et TM et du profil idéal de Kirchhoff ainsi que l'écart entre ces spectres. Les parties des spectres pouvant traverser le système optique d'ouverture numérique ON égale à 1.05 sont comprises entre les valeurs de l'indice m égales à $\pm 5.44^2$. Les ordres ± 1 correspondent aux deux premiers pics ($m = \pm 5.55$) visibles sur les trois spectres. Les ordres ± 2 et ± 3 correspondent aux autres pics de moindres amplitudes. Les écarts entre les spectres des composantes TE et TM avec l'approximation de Kirchhoff sont plus visiblement importants à proximité des positions des ordres de diffraction. Ce résultat est valable pour les ordres qui traversent le système optique, en particulier l'ordre zéro et les ordres ± 1 (partiellement). De plus, l'écart en amplitude de l'ordre zéro est environ dix fois plus grand que celui observé dans le cas du bord diffractant isolé. Ceci se traduira par une image aérienne moins intense et un seuil d'exposition plus faible dans le cas d'un calcul rigoureux par la méthode FDTD.

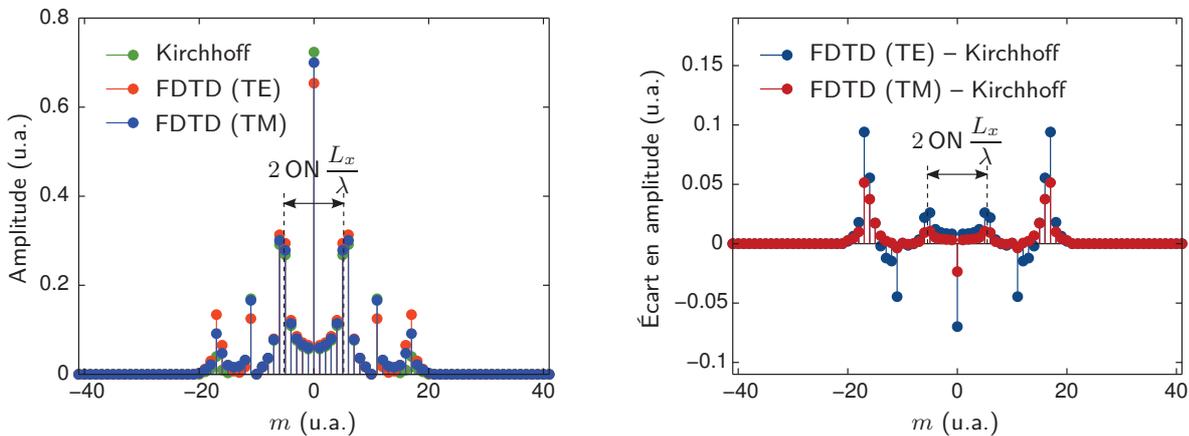


Figure 4.13 – Tracés des spectres de Fourier (en amplitude) des champs électriques d'un réseau de lignes de pas égal à 180 nm ($CD = 75 \text{ nm}$) calculés en sortie du masque avec l'approximation de Kirchhoff et la méthode rigoureuse FDTD (composantes TE et TM) en incidence normale (à gauche) et de l'écart d'amplitude entre les spectres des deux composantes TE et TM par rapport à celui obtenu par l'approximation de Kirchhoff (à droite).

Les écarts en amplitude observés sur les spectres de Fourier des composantes TE et TM par rapport à l'approximation de Kirchhoff se traduisent par une modification de l'intensité de l'image aérienne. Dans la suite, nous allons examiner les effets topographiques du masque sur l'intensité lumineuse dans la résine pour quelques cas de structures. Ensuite, nous analyserons l'impact sur le contraste de l'image aérienne, la position du meilleur focus, la profondeur de champs et le facteur d'erreur du masque. Finalement, nous étudierons l'impact sur le seuil d'impression des motifs diffractants non résolus.

4.4.4 Impact sur l'intensité lumineuse

Afin de simuler l'intensité lumineuse dans la résine, nous avons utilisé un masque atténué à 6% de transmission et à déphasage égal à 180 degrés, une ouverture numérique ON égale à 1.05 et une source d'illumination non polarisée de type Quasar³ avec des facteurs de cohérence partielle $\sigma_{in} = 0.6$ et $\sigma_{out} = 0.9$ et un angle d'ouverture égal à 30 degrés. Les profils d'intensité

2. $|m| \leq \text{ON} \frac{L_x}{\lambda}$ avec $\text{ON}=1.05$, $L_x=1\mu\text{m}$ et $\lambda = 0.193\mu\text{m}$

3. voir définition page 10

sont calculés dans un film de résine photosensible à amplification chimique d'épaisseur 130 nm comme le montre le schéma de la figure 4.14. La diffusion dans la résine est choisie nulle ; ce qui nous permet de réduire les effets chimiques et de ne considérer que les effets topographiques.

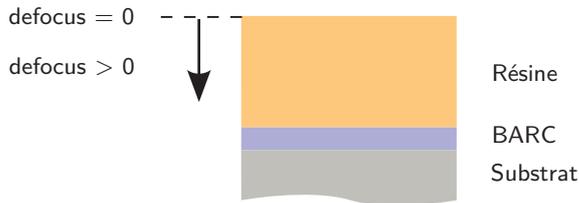


Figure 4.14 – Empilement de la résine servi pour réaliser les simulations pour l'étude des effets topographiques du masque.

Afin de calculer la position du meilleur focus, nous avons fait varier la position du plan focal ainsi que du plan d'observation dans la résine et calculé le contraste⁴ de l'intensité lumineuse pour un réseau de lignes denses de pas égal à 180 nm avec un CD égal à 75 nm. La figure 4.15 nous donne deux couples de plans focal et d'observation pour lesquels le contraste est maximal. Ces deux couples traduisent les positions selon z dans la résine où l'amplitude des ondes stationnaires est maximale, c'est-à-dire au quart et au trois quarts de la période des ondes stationnaires : $\frac{\lambda}{4n} = 28.4$ nm et $\frac{3\lambda}{4n} = 85.1$ nm (n étant l'indice de réfraction de la résine et vaut 1.7). Nous avons retenu cette dernière valeur comme position du meilleur focus correspondant à un defocus égal à 35 nm.

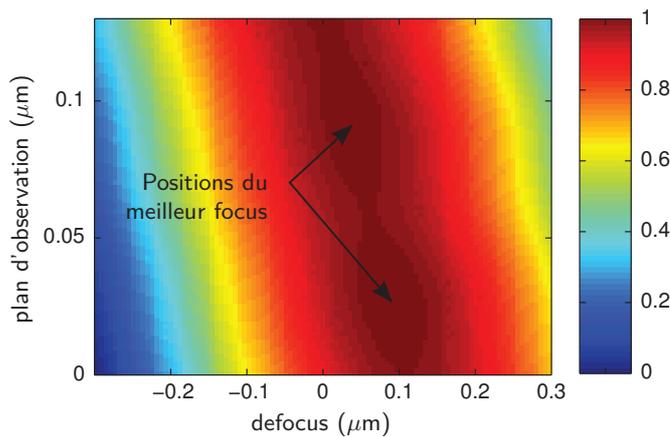


Figure 4.15 – Contraste de l'intensité lumineuse dans la résine en fonction du focus et de la position du plan d'observation, pour un réseau de lignes de pas égal à 180 nm et CD égal à 75 nm. Le maximum du contraste est obtenu pour deux valeurs correspondant au premier quart et au trois quarts de la période des ondes stationnaires : $\frac{\lambda}{4n} = 28.4$ nm et $\frac{3\lambda}{4n} = 85.1$ nm.

Nous avons simulé les intensités lumineuses dans la résine pour quelques structures et tracé leurs profils sur la figure 4.16. Dans les cas du bord diffractant et de la ligne isolée, les différences en intensité lumineuse entre le calcul rigoureux FDTD et l'approximation de Kirchhoff sont faibles. En effet, pour un seuil fixé à 0.2, les différences d'EPE⁵ par bord sont inférieures à 0.5 nm. Dans le cas des réseaux de lignes denses, nous pouvons clairement remarquer une baisse considérable du contraste de l'intensité lumineuse qui s'accroît pour les réseaux de faibles pas. Pour ce même seuil égal à 0.2, les différences d'EPE par bord sont de 1.75 nm et de 5.1 nm pour les pas 180 nm et 140 nm respectivement. Ce qui correspond à des différences de CD respectives égales à 3.5 nm et 10.2 nm qui sont au delà des spécifications pour le niveau de grille du nœud technologique 45 nm.

4. Le contraste est donné par

$$\frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{max}} + I_{\text{min}}}$$

où I_{max} et I_{min} désignent le maximum et le minimum de l'intensité lumineuse.

5. L'EPE (de l'anglais "Edge Placement Error") est l'écart entre le contour de la résine et le motif visé dans la résine.

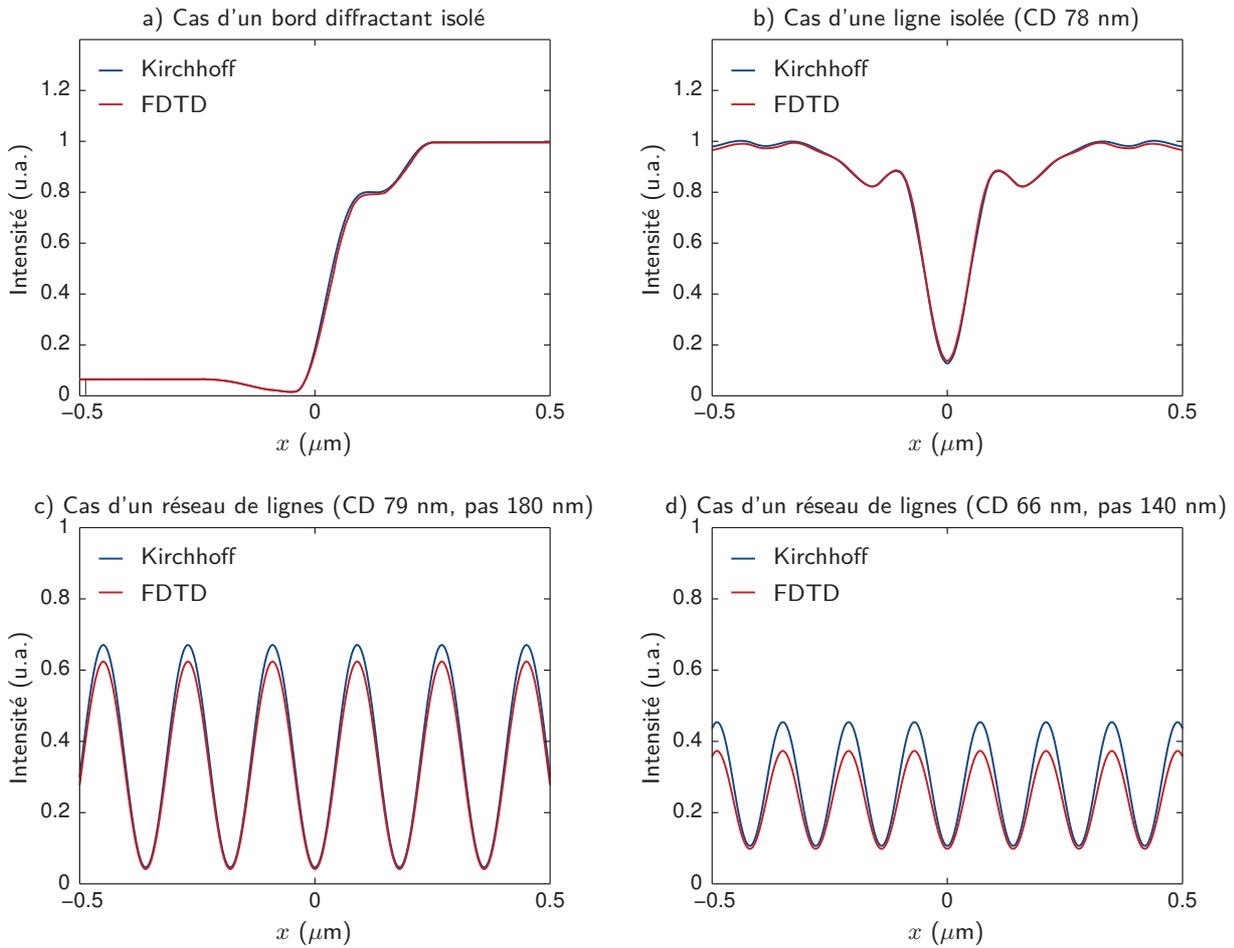


Figure 4.16 – Profils de l'intensité lumineuse dans la résine calculés avec l'approximation de Kirchhoff et la méthode rigoureuse FDTD pour a) un bord diffractant b) une ligne isolée (CD 78 nm) c) un réseau de lignes denses (CD 79 nm, pas 180 nm) d) un réseau de lignes denses (CD 66 nm, pas 140 nm).

4.4.4.1 Contraste

Nous avons simulé l'intensité lumineuse dans la résine pour des réseaux de lignes de même taille (CD 75 nm) avec des pas allant de 140 nm à 360 nm. La figure 4.17 montre les courbes du maximum (I_{\max}) et du minimum (I_{\min}) de l'intensité lumineuse obtenues par la méthode rigoureuse FDTD et l'approximation de Kirchhoff.

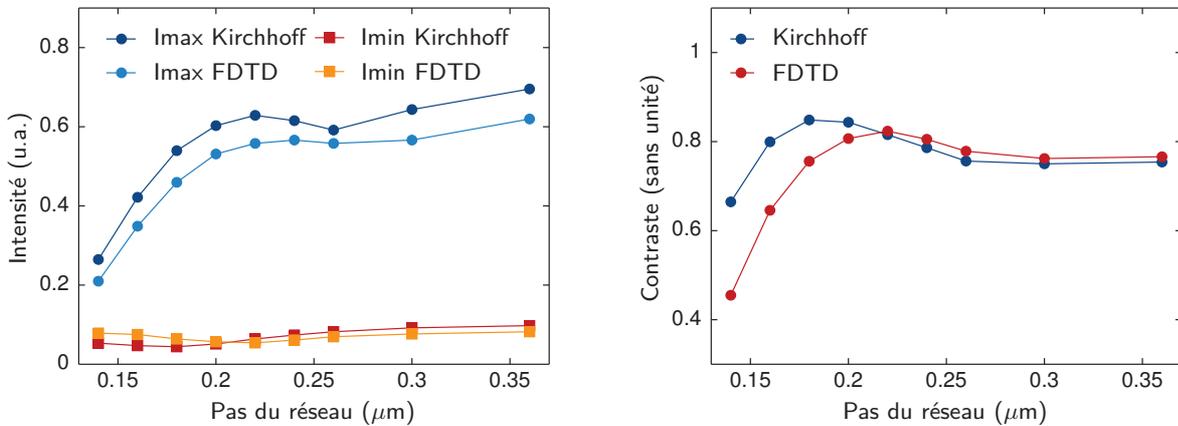


Figure 4.17 – Courbes de I_{\max} et I_{\min} pour des réseaux de lignes de pas allant de 140 nm à 360 nm (à gauche) et du contraste de l'intensité lumineuse dans la résine en fonction du pas de réseau (à droite)

Nous pouvons clairement remarquer que, contrairement au minimum d'intensité I_{\min} qui reste à peu près du même ordre de grandeur, le maximum d'intensité I_{\max} calculé par la méthode

rigoureuse FDTD est visiblement plus faible que celui de l'approximation de Kirchhoff. Les courbes du contraste de l'intensité lumineuse calculé à partir des valeurs de I_{\max} et I_{\min} montre une dégradation significative de celui-ci avec la méthode FDTD pour les pas de réseau inférieurs à 200 nm. Au contraire, le contraste est légèrement amélioré pour les autres pas de réseau.

4.4.4.2 Position du meilleur focus

Nous avons calculé la position du meilleur focus dans la résine pour des réseaux de lignes denses de même dimension ($CD = 75$ nm) et de pas allant de 140 nm à 360 nm. Les profils de l'intensité lumineuse sont obtenus dans le plan d'observation situé à la surface de la résine. La courbe 4.18 montre que, contrairement au cas de l'approximation de Kirchhoff, la position du plan focal est modifiée par les effets topographiques du masque et devient dépendante du pas de réseau et des dimensions des structures dessinées sur le masque. L'écart maximal en position du plan focal, égal à 170 nm, est visiblement supérieur aux erreurs intrinsèques du scanner qui sont de l'ordre de 40 nm. Ce résultat indique que l'on pourrait être à tout moment hors focus si l'on travaille avec une position unique du plan d'observation.

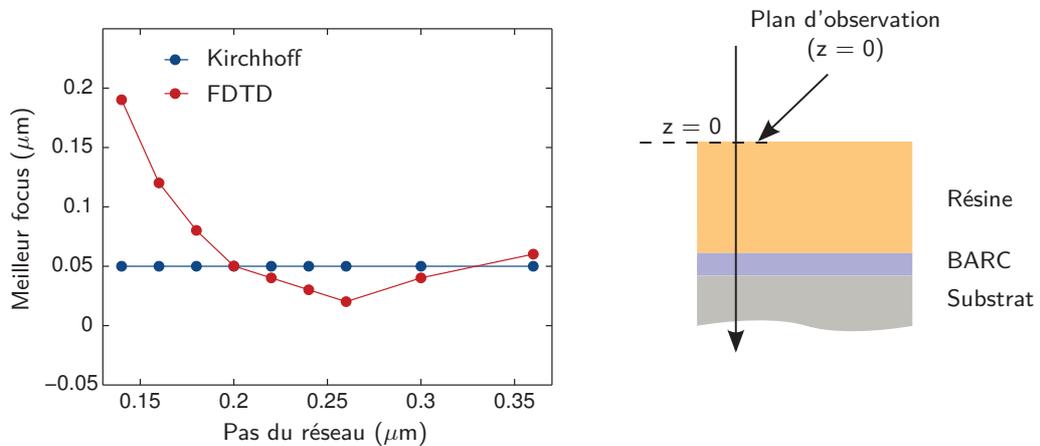


Figure 4.18 – Courbes de la position du meilleur focus pour des réseaux de lignes de pas allant de 140 nm à 360 nm calculées par la méthode rigoureuse FDTD et l'approximation de Kirchhoff. Les profils de l'intensité lumineuse sont calculés dans un unique plan d'observation fixé à $z = 0$.

4.4.4.3 Profondeur de champ

Nous avons simulé dans une plage de defocus de -250 nm à 250 nm les intensités lumineuses dans la résine pour des réseaux de lignes de pas allant de 140 nm à 360 nm. Pour une valeur donnée du defocus, l'intensité lumineuse est calculée de façon à prendre la moyenne de tous les profils d'intensité simulés à différentes positions du plan d'observation dans la résine. À partir de ces courbes d'intensité, nous avons évalué la latitude d'exposition (EL) en fonction de la profondeur de champ (DOF) pour les différents réseaux. La figure 4.19 illustre deux exemples de courbes EL/DOF obtenues pour les pas de réseau 140 nm et 260 nm avec la méthode rigoureuse FDTD et l'approximation de Kirchhoff.

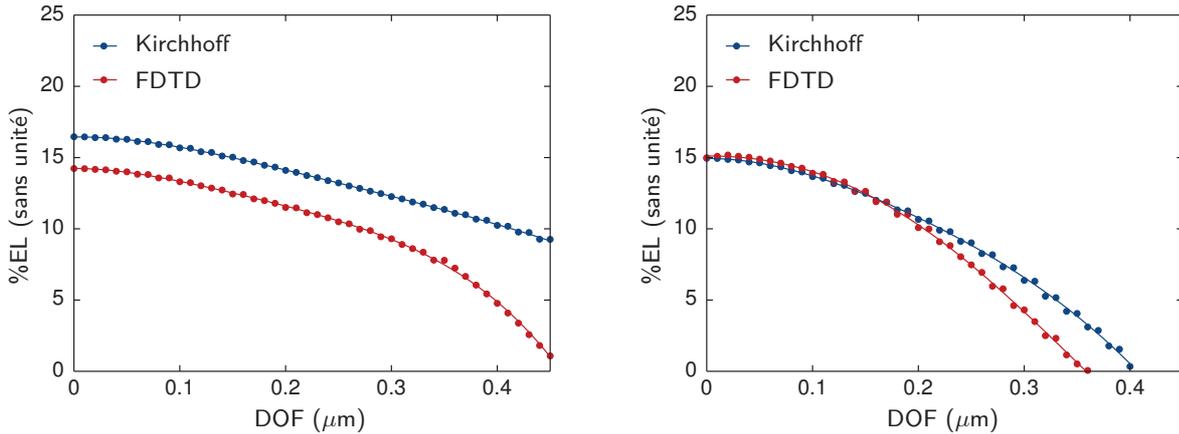


Figure 4.19 – Exemples de courbes EL/DOF calculées avec l'approximation de Kirchhoff et la méthode rigoureuse FDTD pour deux réseaux de lignes de pas 140 nm (à gauche) et 260 nm (à droite).

Dans le but de comparer les valeurs de DOF autorisées du procédé, nous avons calculé les différences ΔDOF entre l'approximation de Kirchhoff et la méthode rigoureuse FDTD pour des valeurs fixes de la latitude d'exposition ($\text{EL} = 8\% \pm 2\%$). Le choix de 8% est justifié par le fait que nous avons négligé la diffusion dans la résine lors des simulations des intensités lumineuses. Dans le cas réel où une diffusion est appliquée, une valeur de $\text{EL} = 6\%$ est souvent utilisée. Pour notre cas, $\text{EL} = 8\%$ est mieux appropriée pour comparer les résultats de simulation avec la réalité. La figure 4.20 montre l'allure de ΔDOF (Kirchhoff – FDTD) pour $\text{EL} = 6\%$, 8% et 10% en fonction du pas de réseau. D'après cette figure, l'approximation de Kirchhoff surestime la profondeur de champ DOF (de 150 nm) pour le pas de réseau égal à 140 nm. Pour les pas de réseau compris entre 160 nm et 300 nm, les différences en DOF restent faibles (moins de 40 nm) avec une légère surestimation de la DOF avec l'approximation de Kirchhoff. Cependant, le comportement est inversé pour le pas 360 où nous pouvons constater une sous-estimation (de 50 nm) de la DOF par l'approximation de Kirchhoff. La validité de cette dernière pour estimer la DOF est donc limitée aux pas de réseau compris entre 160 nm et 300 nm. Elle ne peut donc être utilisée pour les réseaux très denses (<160 nm) ou relâchés (>300 nm) et même pour les lignes isolées.

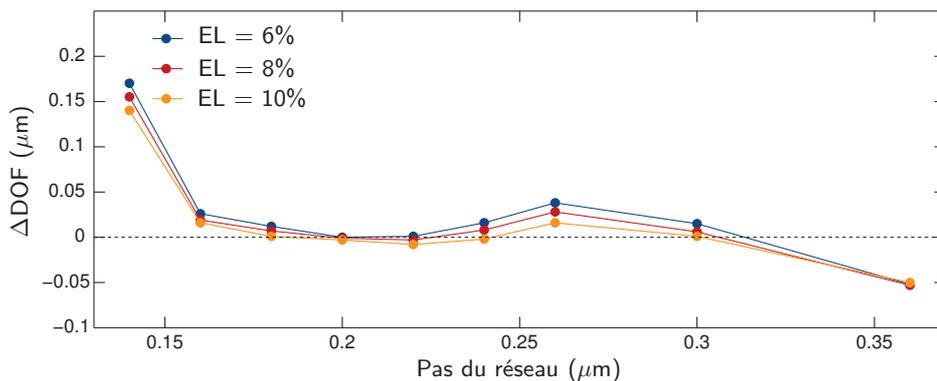


Figure 4.20 – Courbes de ΔDOF (Kirchhoff – FDTD) à $\text{EL} = 6\%$, 8% et 10% pour des réseaux de lignes de CD égal à 75 nm et de pas allant de 140 nm à 360 nm.

4.4.4.4 Effets de proximité

En appliquant un modèle à seuil constant aux profils d'intensité, nous avons estimé les valeurs de CD plaquette pour chacun des réseaux. Pour l'approximation de Kirchhoff, comme pour la

méthode FDTD, le seuil d'extraction a été fixé de manière à avoir un CD égal à 65 nm sur la plaquette (voir figure 4.21).

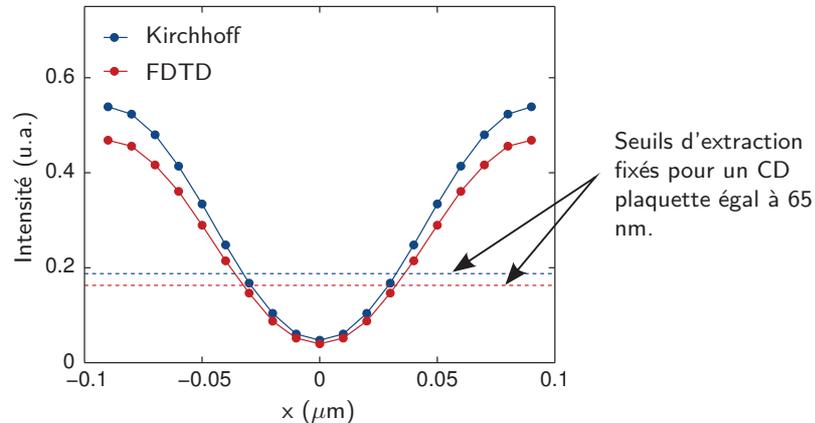


Figure 4.21 – Profils de l'intensité lumineuse dans la résine pour un réseau de lignes (CD 75 nm) de pas égal à 180 nm, calculés par l'approximation de Kirchhoff et la méthode FDTD.

Afin d'estimer les effets de proximité, nous avons calculé les dimensions sur la plaquette avec l'approximation de Kirchhoff et la méthode FDTD. La figure 4.22 montre que les effets de proximité sont similaires pour les deux méthodes. Néanmoins, les différences de CD sont grandes (entre 5 nm et 16 nm) pour tous les pas de réseau, à l'exception de ceux qui sont proches du pas de calibration des seuils d'extraction (ie. 160 nm, 180 nm et 200 nm). Nous verrons plus tard que le modèle à seuil constant est insuffisant pour garantir une bonne estimation du CD en utilisant un masque topographique.

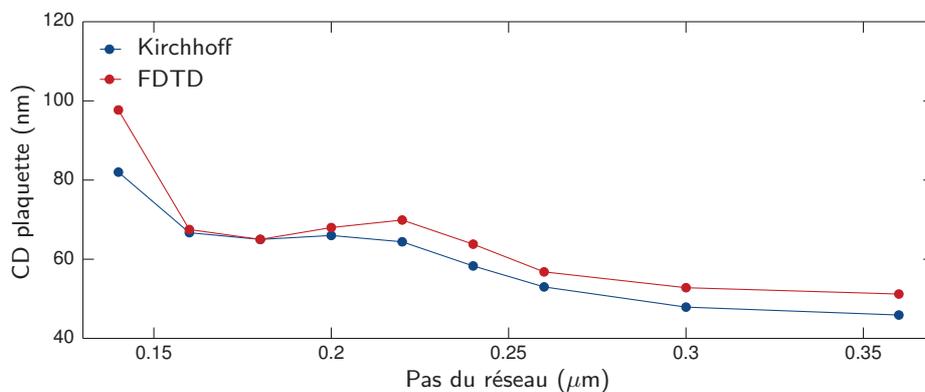


Figure 4.22 – Courbes du CD plaquette en fonction du pas de réseau pour des réseaux de lignes de pas allant de 140 nm à 360 nm et de CD masque égal à 75 nm, calculées par l'approximation de Kirchhoff et la méthode FDTD pour la partie optique et un modèle à seuil constant pour la partie résine.

4.4.4.5 Effets sur le seuil d'impression des SRAF

Afin d'éviter leurs impressions dans la résine, les motifs diffractants non résolus (SRAF) sont dessinés avec une résolution bien inférieure à celle du scanner. Les SRAF disparaissent après les étapes d'exposition et du développement et les mesures des dimensions critiques pouvant servir dans la calibration des modèles empiriques ne peuvent donc pas avoir lieu (cas de la figure 4.23). De plus, dans le cas où les SRAF s'impriment, les mesures que l'on peut effectuer ne sont pas utilisables puisqu'elles contiennent un taux de bruit élevé dû à la rugosité des lignes (cas de la figure 4.24). Pour toutes ces raisons, la prédiction de l'impression des SRAF par un modèle empirique devient problématique pour la technologie actuelle.

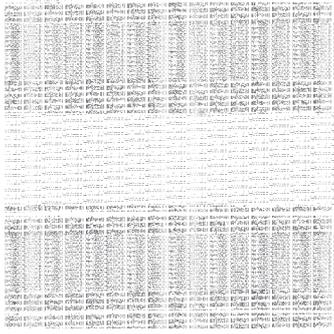


Figure 4.23 – Image au SEMCD d'une ligne appartenant à un réseau de pas égal à 296 nm avec des dimensions égales à 81.5 nm sur le masque (à l'échelle de la plaquette) et 64.55 nm dans la résine. Les SRAF mesurent 31.5 nm sur le masque (à l'échelle de la plaquette) et ne présentent que quelques traces non mesurables dans la résine.

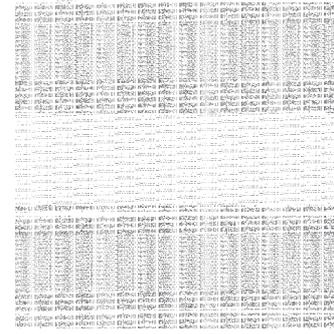


Figure 4.24 – Images au SEMCD d'une ligne appartenant à un réseau de pas égal à 348 nm avec des dimensions égales à 99 nm sur le masque (à l'échelle de la plaquette) et 73.21 nm dans la résine. Les SRAF mesurent 34 nm sur le masque (à l'échelle de la plaquette) et 35 nm (± 2 nm) dans la résine.

Dans le cas des modèles à seuil variable VT5, le seuil est fonction des paramètres I_{max} , I_{min} , Slope et Factor qui sont calculés sur des sites de contrôle positionnés sur les motifs principaux. Les simulations ne peuvent donc pas être réalisées pour les SRAF. Même dans le cas des modèles gaussiens (Brion) et à opérateur différentiel (CM1) où l'on introduit les termes permettant de modéliser la diffusion initiale de l'acide et de la base, la partie de la courbe de l'intensité lumineuse située entre deux seuils d'extraction s_1 et s_2 est exclue de la convolution. Les images latentes peuvent donc être mal simulées au niveau des SRAF (voir figure 4.25).

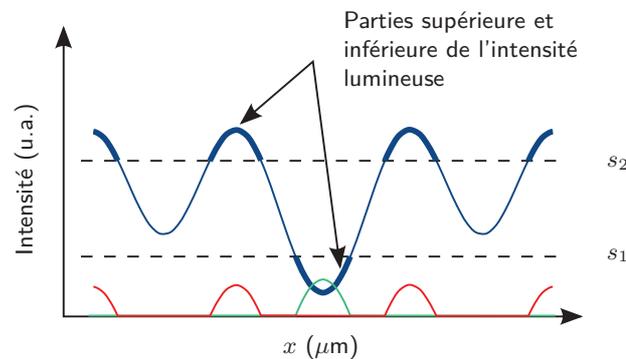


Figure 4.25 – Extraction des parties supérieure et inférieure de l'intensité lumineuse dans le cas des modèles gaussiens (Brion) et à opérateur différentiel (CM1). La partie de l'intensité située au niveau des motifs diffractants non résolus est exclue de la convolution.

Il existe une technique couramment utilisée en traitement OPC permettant de corriger la taille des SRAF afin d'éviter leur impression dans la résine. Cette technique permet de simuler, par l'intermédiaire d'un modèle optique simple, l'intensité lumineuse au niveau des SRAF et de rectifier (par un processus itératif) la taille des SRAF afin de garantir à ce que le minimum d'intensité I_{min} des SRAF soit supérieur au seuil de référence du modèle. Seulement, les simulations de l'intensité lumineuse sont effectuées en utilisant l'approximation de Kirchhoff. Comme dans le cas des motifs de petites dimensions et les réseaux de lignes de petits pas, les effets topographiques du masque deviennent importants et l'approximation de Kirchhoff ne sera plus appropriée pour de telles simulations. En effet, la figure 4.26 illustre l'exemple de l'intensité lumineuse d'une ligne isolée protégée par quatre SRAF simulée avec l'approximation de Kirchhoff et la méthode FDTD. Nous pouvons clairement remarquer sur cette figure que le minimum d'intensité des SRAF prédit par la méthode FDTD est plus faible que dans le cas de l'approximation de Kirchhoff. Un seuil d'extraction fixé entre les deux minima locaux de l'intensité lumineuse entraîne deux prédictions différentes de l'impression des SRAF. Autrement dit, le

risque d'impression des SRAF est mieux prédit par la méthode FDTD que par l'approximation de Kirchhoff.

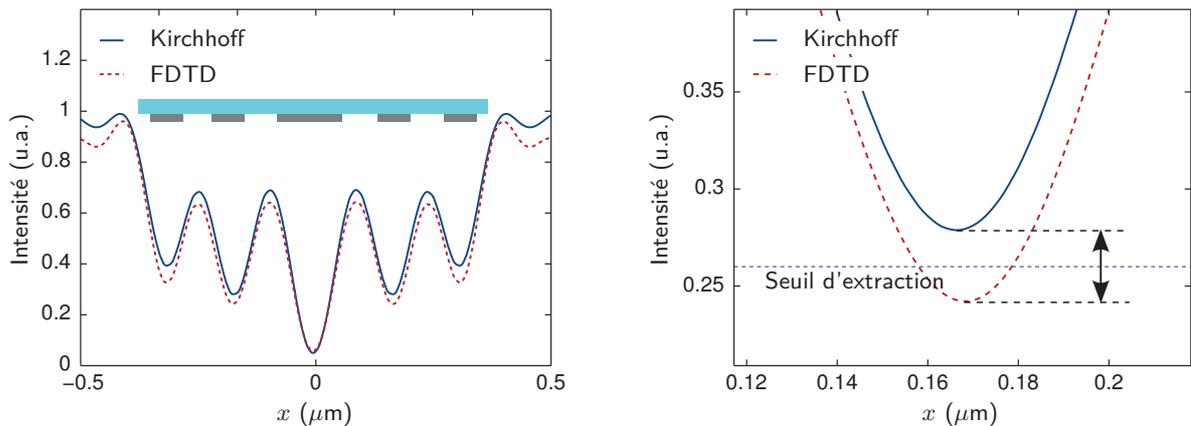


Figure 4.26 – Profils de l'intensité de l'image aérienne d'une ligne isolée protégée avec des SRAF, calculés en utilisant l'approximation de Kirchhoff et la méthode rigoureuse FDTD (à gauche) et présence de risque d'impression des motifs SRAF prédit par la méthode FDTD (à droite).

Dans le but d'évaluer l'impact des effets topographiques du masque sur l'impression des SRAF, nous avons dessinés des réseaux de lignes de pas allant de 300 nm à 1 μm auxquels nous avons inséré des SRAF dont le positionnement et la taille sont détaillés dans le tableau 4.2. Pour chacun des réseaux, nous avons calculé l'intensité lumineuse dans la résine de façon à prendre la moyenne de plusieurs profils de l'intensité simulés à différentes positions du plan d'observation. Les simulations ont été réalisées en utilisant un masque PSM avec 6% de transmission et un décalage de phase de 180 degrés, une source d'illumination de type Quasar⁶ (σ_{in} : 0.6, σ_{out} : 0.9, angle d'ouverture : 30 degrés), une ouverture numérique ON égale à 1.05 et une position du meilleur focus fixée à 35 nm à partir de la surface de la résine. Afin de viser un CD égal à 65 nm sur la plaquette, nous avons appliqué des corrections géométriques aux lignes des réseaux en utilisant la méthode FDTD pour les simulations. Finalement, nous avons simulé, par l'approximation de Kirchhoff et la méthode FDTD, les profils de l'intensité lumineuse des réseaux corrigés et extrait les dimensions sur la plaquette. Les deux seuils d'extraction (Kirchhoff et FDTD) ont été calibrés pour le réseau de lignes de pas égal à 180 nm et un CD plaquette égal à 65 nm.

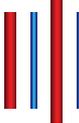
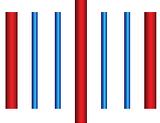
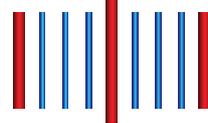
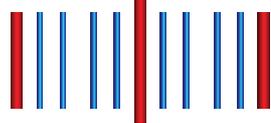
Pas (nm)	de 300 à 400	500	600	800 et 1000
Détails	 1 barre centrée, taille = 30 nm (34 nm pour pas = 400 nm)	 2 barres, espace = 105 nm, taille = 34 nm	 2 barres + 1 centrée, espace = 105 nm, taille = 34 nm	 4 barres, espace1 = 105 nm, taille1 = 34 nm, espace2 = 100 nm, taille2 = 34 nm

Tableau 4.2 – Détails sur le positionnement et la taille des motifs diffractants non résolus SRAF dessinés pour des réseaux de lignes de pas allant de 300 nm à 1000 nm.

6. voir définition page 10

Nous avons tracé sur la figure 4.27 le budget normalisé en seuil d'impression⁷ en fonction du pas de réseau. D'après cette figure, les premières SRAF insérées pour les réseaux de pas entre 300 nm et 400 nm présentent un risque élevé d'impression avec la méthode FDTD, contrairement à l'approximation de Kirchhoff. Les autres pas de réseau présentent un risque moins élevé d'impression par la méthode FDTD pour les premières SRAF. Par contre, les secondes SRAF ne présentent aucun risque d'impression pour les deux méthodes.

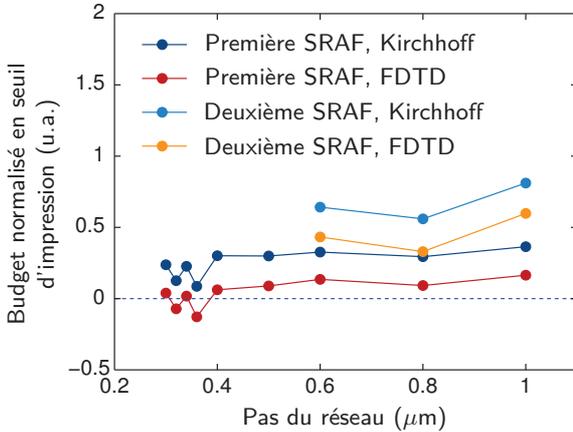


Figure 4.27 – Budget normalisé en dose d'impression des SRAF pour des réseaux de lignes de pas allant de 300 nm à 1000 nm (voir 4.2).

4.4.5 Discussion des résultats

À la lumière des résultats de simulation, l'approximation de Kirchhoff n'est pas capable de prendre en compte les effets de la diffraction au niveau du masque. Les effets topographiques du masque résultent en la dégradation du contraste de l'intensité lumineuse, le décalage de la position du meilleur focus et une mauvaise prédiction des effets de proximité ainsi que de l'impression des motifs diffractants non résolus (SRAF). Les erreurs de simulation sont pénalisantes lorsque les motifs sont petits (cas d'une ligne isolée de petite dimension) ou lorsque les espaces entre-eux sont réduits (cas des réseaux de lignes de pas plus petits que la longueur d'onde). Le calcul des critères de performance du procédé lithographiques (CD, DOF, MEEF, etc) peut ainsi être altéré. Ce qui peut être plus ou moins pénalisant quant à la simulation de l'intensité lumineuse et au calcul des distances de déplacement des bords des motifs lors d'un traitement OPC par exemple.

Il existe, néanmoins, des méthodes semi-rigoureuses pour remplacer l'approximation de Kirchhoff. Dans la suite de ce chapitre, nous allons présenter le principe de ces méthodes que nous évaluerons ensuite par rapport au calcul rigoureux de la méthode FDTD.

4.5 Méthodes semi-rigoureuses

4.5.1 Méthode de décomposition de domaine DDM

La méthode de décomposition de domaine (de l'anglais "*Domain Decomposition Method*" ou DDM) est une technique qui permet de résoudre un problème mathématique dans un domaine quelconque. Le principe de base de cette méthode est de décomposer le domaine concerné en plusieurs sous-domaines élémentaires indépendants et de résoudre le problème dans chacun de ces sous-domaines. L'origine de cette méthode remonte à 1869 quand H. A. Schwarz proposa une

7. Le budget normalisé en seuil d'impression est défini par

$$\frac{I_{\text{minSRAF}} - \text{seuil}}{\text{seuil}}$$

technique itérative pour la construction de fonctions harmoniques dans des régions de formes irrégulières pouvant être exprimées comme l'union de sous-régions de formes simples (comme des rectangles et des sphères). La technique de Schwarz fut la première fois adoptée en 1936 par Sobolev [110], puis par d'autres scientifiques [111–117].

En lithographie optique, la méthode DDM a été introduite en 2001 par Adam et Neureuther [118, 119] dans le but de développer des modèles simplifiés pouvant décrire rigoureusement les effets de transition au niveau des bords d'un masque. Elle a été implémentée en 2005 dans le logiciel de traitement OPC, Calibre™, et est la propriété de Mentor Graphics [120].

Le principe de cette méthode consiste à calculer la valeur du champ électrique en un point donné du masque en additionnant tous les champs électriques provenant des bords des structures, supposés indépendants les uns des autres, auxquels on soustrait les composantes continues. Le calcul est tronqué aux bords se situant uniquement dans un rayon égal à la moitié du diamètre optique de simulation à partir du point considéré. La figure 4.28 présente deux schémas décrivant la manière d'additionner ces champs électriques dans le cas d'une onde polarisée en TE (figure 4.28 (a)) et en TM (figure 4.28 (b)). La valeur finale du champ électrique au point M est alors égale à sa valeur idéale à laquelle on additionne les écarts par rapport à leurs profils idéaux des champs électriques provenant des bords. Nous pouvons écrire mathématiquement cette idée comme suit :

$$E(M) = E_{\text{idéal}} + \sum_{\text{bords}} E_{\text{rigoureux}} - E_{\text{idéal}} \quad (4.11)$$

où $E_{\text{idéal}}$ désigne le champ idéal de Kirchhoff et $E_{\text{rigoureux}}$ la composante TE ou TM du champ électrique rigoureux.

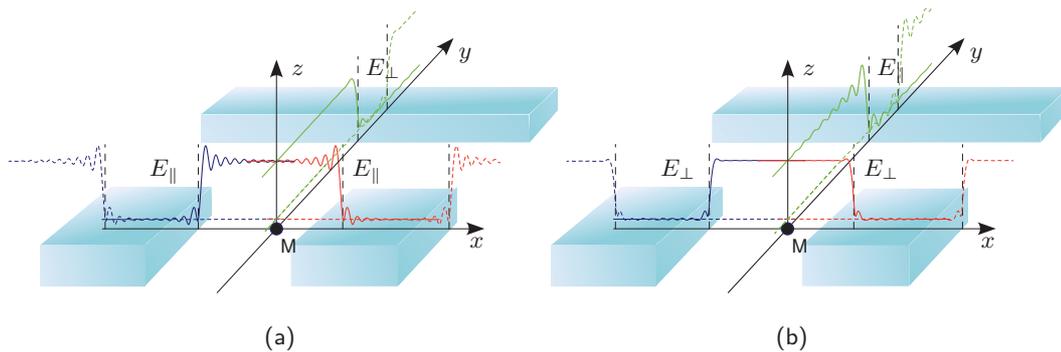


Figure 4.28 – Principe de la méthode DDM : le champ électrique en un point M donné du masque se calcule par la somme de tous les champs électriques provenant des bords des structures qui sont dans le rayon optique de simulation auxquels on soustrait les composantes continues. Les cas (a) et (b) correspondent aux modes TE et TM respectivement. Les profils des champs dessinés en traits continus proviennent des bords se situant dans l'angle de vision du point M, ceux qui sont en traits discontinus représentent le cas contraire.

4.5.2 Méthode de compensation des effets de diffraction M3D

Cette méthode est celle qu'utilise le logiciel commercial Tachyon™ de la société Brion pour une modélisation électromagnétique rapide des effets de la topographie. Elle a été mise en place par Liu et al. [43] en 2007. Son principe est basé sur une décomposition dans le plan de Fourier du champ électromagnétique issu d'un masque topographique en un champ idéal de Kirchhoff et en d'autres termes permettant de décrire la diffraction au niveau des bords et des coins. Selon cette méthode, l'expression de la fonction de transmission du masque dans le plan de Fourier est donnée par l'équation suivante :

$$t_{3D}(\vec{f}) = c(\vec{f})t(\vec{f}) + c_H t_H(\vec{f}) + c_V t_V(\vec{f}) + c_C t_C(\vec{f}) \quad (4.12)$$

avec $\vec{f} = (f_x, f_y)$ les fréquences spatiales, $t(\vec{f})$ la transformée de Fourier du masque fin utilisant l'approximation de Kirchhoff, $t_H(\vec{f})$, $t_V(\vec{f})$ et $t_C(\vec{f})$ les transformées de Fourier d'un bord horizontal, d'un bord vertical et d'un coin respectivement.

Les coefficients $c(\vec{f})$, $c_H(\vec{f})$, $c_V(\vec{f})$ et $c_C(\vec{f})$ peuvent être considérés comme des facteurs de correction permettant de modifier la fonction de transmission du masque fin au masque topographique. Ces coefficients sont calculés à partir de simulations rigoureuses et dépendent de :

- la topographie et des matériaux constituant l’empilement multicouche du masque (épaisseurs des couches et indices optiques n et k),
- la polarisation de la source,
- les angles d’incidence de la lumière.

Selon Liu et al., l’équation 4.12 est une manière plus générale dans la modélisation du champ électromagnétique couvrant ainsi la méthode DDM [119] et le modèle de couche limite BL⁸ [121]. En effet, elle est capable de prendre en compte les effets de la diffraction au niveau des coins. De plus, elle permet de résoudre le problème de l’incidence normale de la lumière.

4.5.3 Évaluation par rapport au calcul rigoureux

4.5.3.1 Définition des conditions de simulation

Pour le besoin de l’évaluation des deux méthodes citées ci-dessus, nous nous intéressons à comparer par rapport à la méthode rigoureuse FDTD, les résultats de simulation (comme par exemple les courbes de Bossung et EL/DOF) de plusieurs réseaux de lignes de différents pas. Étant donné que différents logiciels commerciaux (EM-SuiteTM, CalibreTM, TachyonTM) sont utilisés pour simuler l’intensité lumineuse, il n’est donc pas certain de modéliser les effets de la réflexion dans la résine de la même façon. Par conséquent, nous avons volontairement annulé ces effets en prenant une épaisseur infinie de la résine. Les plans focal et d’observation sont respectivement positionnés à 35 nm et à 65 nm dans la résine à partir de la surface comme le montre le schéma de la figure 4.29. Finalement, le masque est un PSM à 6% de transmission et à décalage de phase égal à 180 degrés. L’illumination est de forme Quasar⁹ (σ_{in} : 0.6, σ_{out} : 0.9, angle d’ouverture : 30 degrés) non polarisée et l’ouverture numérique ON est égale à 1.05.

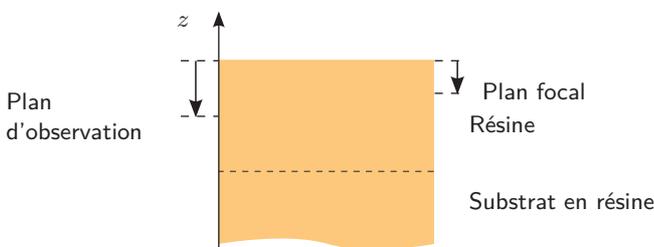


Figure 4.29 – Empilement multicouche de la résine utilisé dans les simulations pour l’évaluation des méthodes DDM et M3D.

Dans la continuité de l’approximation de Kirchhoff, le calcul de l’image aérienne avec la méthode DDM reste basé sur les mêmes vieilles techniques de Hopkins, à savoir [122–125] :

- l’incidence normale de la source d’illumination,
- le calcul des tables optiques basées sur les matrices TCC,
- la décomposition des matrices TCC en une série finie de noyaux de convolution,
- le calcul de l’image aérienne à partir d’un nombre fini des noyaux de convolution.

À l’heure actuelle, les modèles OPC du nœud technologique 45 nm utilisent un nombre de noyaux compris entre 10 et 20 avec l’approximation de Kirchhoff. Pour des besoins de comparaison, nous nous intéressons aux cas suivants :

- une troncature des tables TCC aux 20 premiers noyaux de convolution,
- une troncature des tables TCC aux 64 premiers noyaux de convolution,
- pas de troncature des tables TCC (Hopkins).

8. voir les méthodes heuristiques, page 111.

9. voir définition page 10

Bien qu'il soit plus précis que les autres, ce dernier cas ne peut être utilisé dans un environnement de production puisqu'il suggère un temps de calcul élevé et une capacité de calcul importante. Le but ici est d'être dans les mêmes conditions de calcul pour les trois logiciels commerciaux. Pour les méthodes M3D et FDTD, nous avons utilisé une incidence normale pour rester dans ces mêmes conditions. Finalement, nous avons fixé le diamètre optique de simulation à $2.048 \mu\text{m}$ pour les méthodes DDM et M3D, ce qui correspond au cas nominal de ce paramètre.

Nous avons simulé, à différentes positions du plan focal (defocus allant de -150 nm à 150 nm), les profils de l'intensité lumineuse dans la résine de plusieurs réseaux de lignes de pas allant de 140 nm à 500 nm et une ligne isolée. Ensuite, nous avons normalisé les profils de l'intensité lumineuse calculés par les trois méthodes par rapport à l'intensité du champ clair du masque. Puis, pour chacune des méthodes approchées, nous avons calculé cinq seuils d'extraction correspondant aux différentes conditions en dose du procédé. Les seuils calculés correspondent à des dimensions dans la résine égales à 65 nm , $65 \text{ nm} \pm 5\% \text{CD}$ et $65 \text{ nm} \pm 10\% \text{CD}$ pour une ligne appartenant au réseau de pas égal à 180 nm , comme le montre la figure 4.30.

Finalement, pour des besoins de confidentialité, nous ne représenterons que les résultats de simulation avec une troncature des tables optiques à 64 noyaux et nous désignerons ces deux méthodes semi-rigoureuses par les lettres A et B.

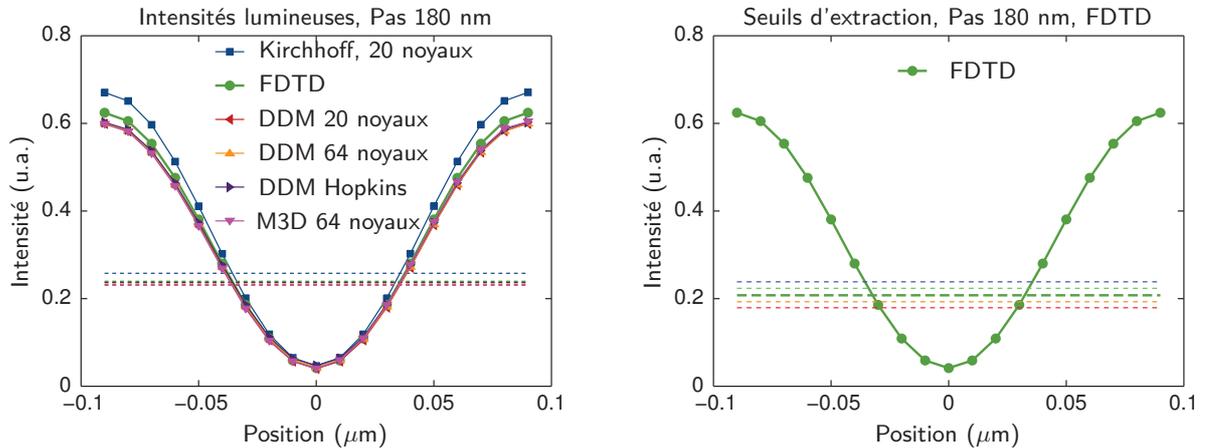


Figure 4.30 – Profils normalisés de l'intensité lumineuse dans la résine d'une ligne appartenant à un réseau de pas égal à 180 nm , obtenus par les différentes méthodes approchées (DDM, M3D et Kirchhoff) ainsi que la méthode rigoureuse FDTD (à gauche) et calcul, dans le cas de la méthode FDTD, des seuils d'extraction correspondant à des dimensions dans la résine égales à 65 nm , $65 \text{ nm} \pm 5\% \text{CD}$ et $65 \text{ nm} \pm 10\% \text{CD}$ (à droite).

4.5.3.2 Fenêtres de procédé et courbes de Bossung

Nous avons calculé les fenêtres de procédé et les courbes de Bossung pour un réseau de lignes denses de pas égal à 140 nm . Sur la figure 4.31, nous présentons les résultats obtenus dans le cas de la méthode A avec des tables optiques TCC tronquées à 64 noyaux. D'après cette figure, un premier constat concerne le décalage en dose entre les différentes fenêtres de procédé. Ce décalage signifie que les seuils d'extraction (ie. les doses nominales) calculés par ces méthodes sont différents, et ce malgré le fait que l'on ait normalisé les profils de l'intensité lumineuse par rapport à l'intensité du champ clair du masque. Ceci se traduit par des différences entre les profils de l'intensité calculés par les différentes méthodes. Notons également que les courbes de Bossung obtenues par la méthode A sont plus proches du calcul rigoureux FDTD qu'avec l'approximation de Kirchhoff. Néanmoins, les résultats ne coïncident pas parfaitement avec le calcul rigoureux et les erreurs en CD peuvent s'élever à quelques nanomètres pour certaines doses d'exposition.

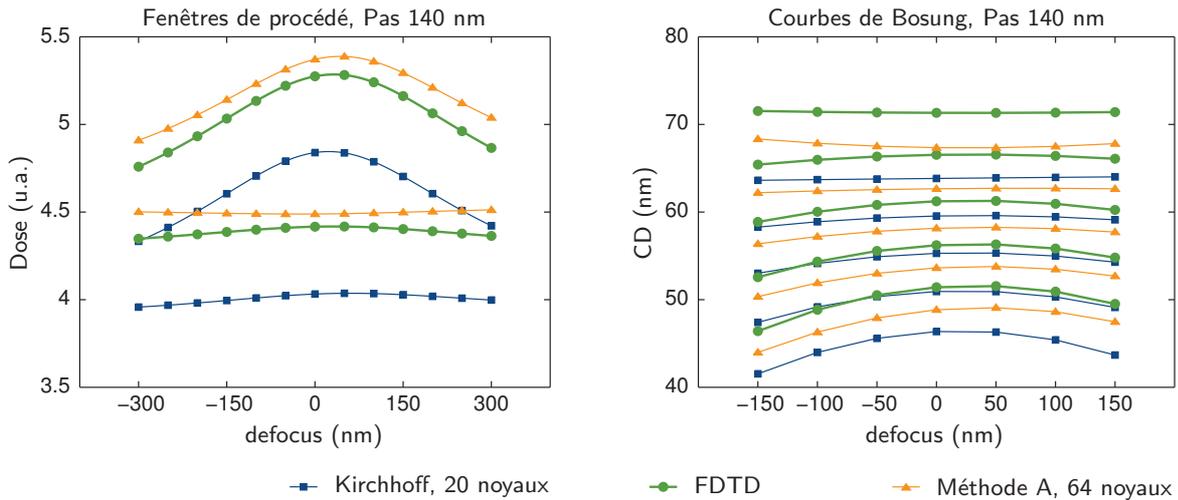


Figure 4.31 – Fenêtres de procédé (à gauche) et courbes de Bossung (à droite) calculées pour un réseau de lignes denses de pas égal à 140 nm par la méthode A (avec troncature des tables optiques à 64 noyaux), l'approximation de Kirchhoff ainsi que la méthode rigoureuse FDTD.

Pour une comparaison directe, nous avons présenté sur la figure 4.32 les résultats obtenus dans le cas de la méthode B avec une troncature des matrices TCC à 64 noyaux. Comme nous pouvons le remarquer, les fenêtres de procédé calculées par cette méthode sont décalées en dose d'exposition. En ce qui concerne les courbes de Bossung, les résultats obtenus par la méthode B sont plus proches du calcul rigoureux que dans le cas de l'approximation de Kirchhoff. Néanmoins, comme dans le cas de la méthode A, les valeurs de CD simulées par la méthode B ne coïncident pas parfaitement avec les résultats de l'outil rigoureux FDTD.

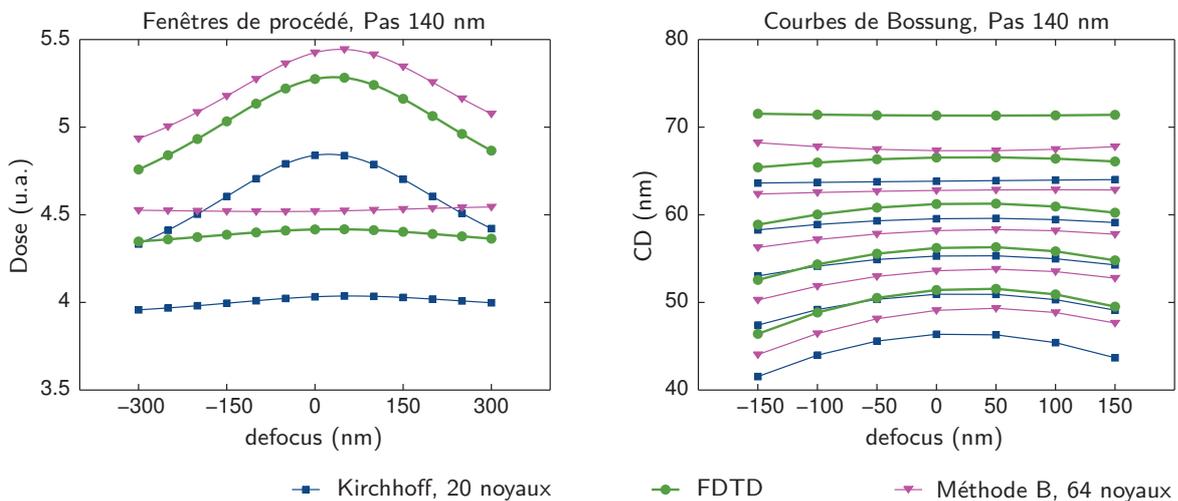


Figure 4.32 – Fenêtres de procédé (à gauche) et courbes de Bossung (à droite) calculées pour un réseau de lignes denses de pas égal à 140 nm par la méthode B (avec troncature des tables optiques à 64 noyaux), l'approximation de Kirchhoff ainsi que la méthode rigoureuse FDTD.

4.5.3.3 Position du meilleur focus

À partir des fenêtres de procédé obtenues par les différentes méthodes et pour différents pas de réseau, nous avons calculé la position du meilleur focus qui permet de maximiser la profondeur de champ¹⁰. D'après la figure 4.33, les méthodes approchées A et B, contrairement à l'approximation de Kirchhoff, permettent de mieux recentrer les fenêtres de procédé en focus

10. Cette position du plan focal n'est autre que le centre de la fenêtre de procédé

pour la plupart des structures. Il existe, néanmoins, des cas où les fenêtres de procédé sont légèrement décalées en focus, en l'occurrence les pas de réseau 140 nm et 300 nm.

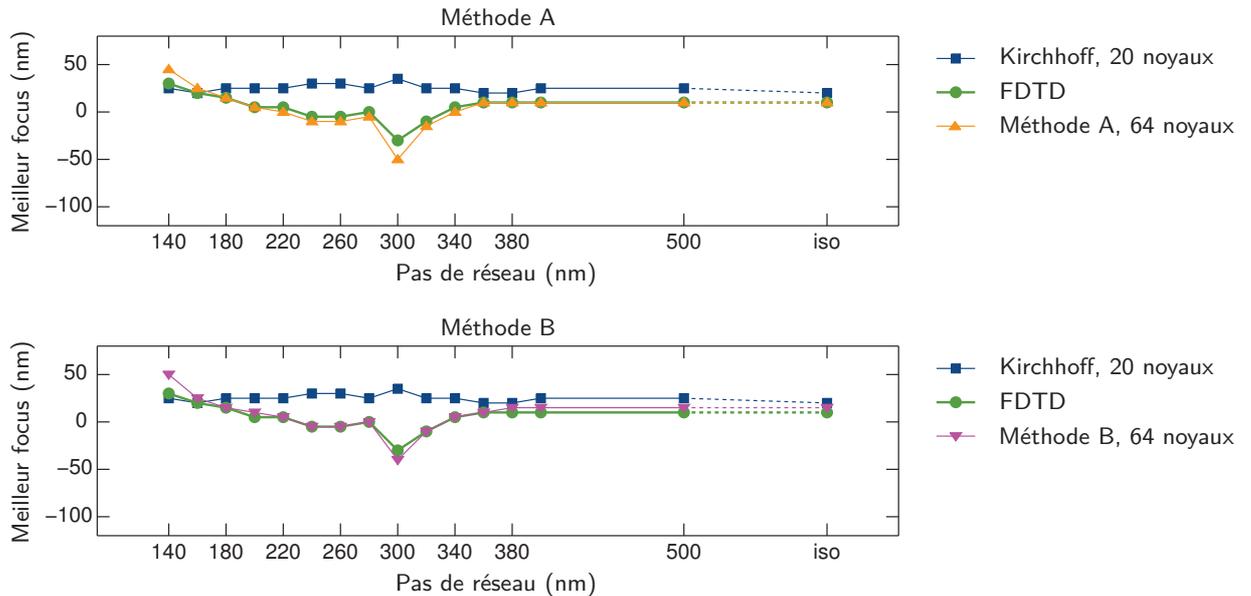


Figure 4.33 – Tracés de la position du meilleur focus en fonction du pas de réseau prédits par les méthodes approchées A et B (avec une troncature des tables optiques à 64 noyaux), l'approximation de Kirchhoff ainsi que la méthode rigoureuse FDTD.

4.5.3.4 Profondeur de champ

En ce qui concerne la profondeur de champ, nous avons présenté sur la figure 4.34 les valeurs de DOF obtenues par les différentes méthodes pour une latitude en dose d'exposition EL fixée à 8%. D'après cette figure, les méthodes approchées A et B permettent de mieux estimer la DOF, contrairement à l'approximation de Kirchhoff. Néanmoins, les pas de réseau 140 et 300 nm présentent une surestimation de la DOF de plus de 100 nm par rapport à la méthode FDTD.

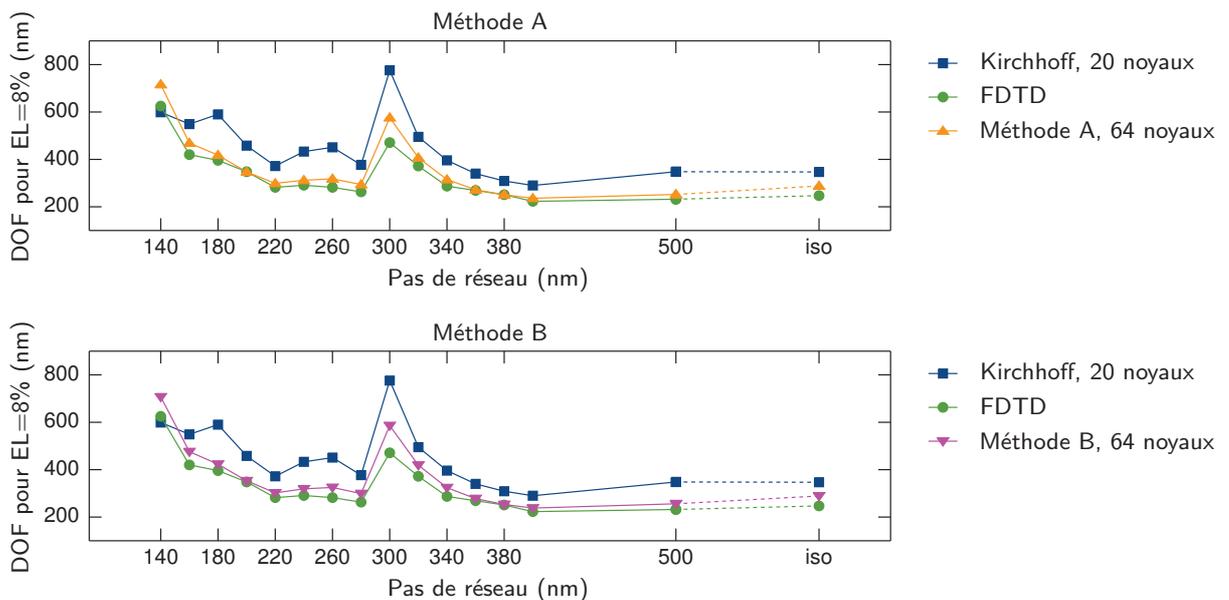


Figure 4.34 – Tracés de la profondeur de champ DOF pour une latitude en dose d'exposition EL fixée à 8% en fonction du pas de réseau pour les méthodes approchées A et B (avec une troncature des tables optiques à 64 noyaux), l'approximation de Kirchhoff ainsi que la méthode rigoureuse FDTD.

4.5.3.5 Effets de proximité

Avec les différentes méthodes approchées, nous avons calculé, en fonction du pas de réseau, les dimensions critiques dans la résine pour différentes positions du plan focal. Sur la figure 4.35, nous présentons, en fonction du pas de réseau et du décalage en focus, les valeurs de ΔCD (écarts avec la méthode FDTD) obtenues dans le cas de l'approximation de Kirchhoff avec 20 noyaux et les méthodes A et B avec 64 noyaux. D'après cette figure, nous pouvons clairement remarquer que les méthodes approchées A et B permettent de réduire les écarts en CD avec la méthode rigoureuse FDTD et de corriger une partie des erreurs focales que l'approximation de Kirchhoff est incapable de maîtriser. Cependant, l'erreur en dimensions critiques dépend du pas de réseau. Les méthodes A et B ne permettent donc pas de corriger tous les effets topographiques du masque. Elles permettent seulement une réduction partielle de ces erreurs.

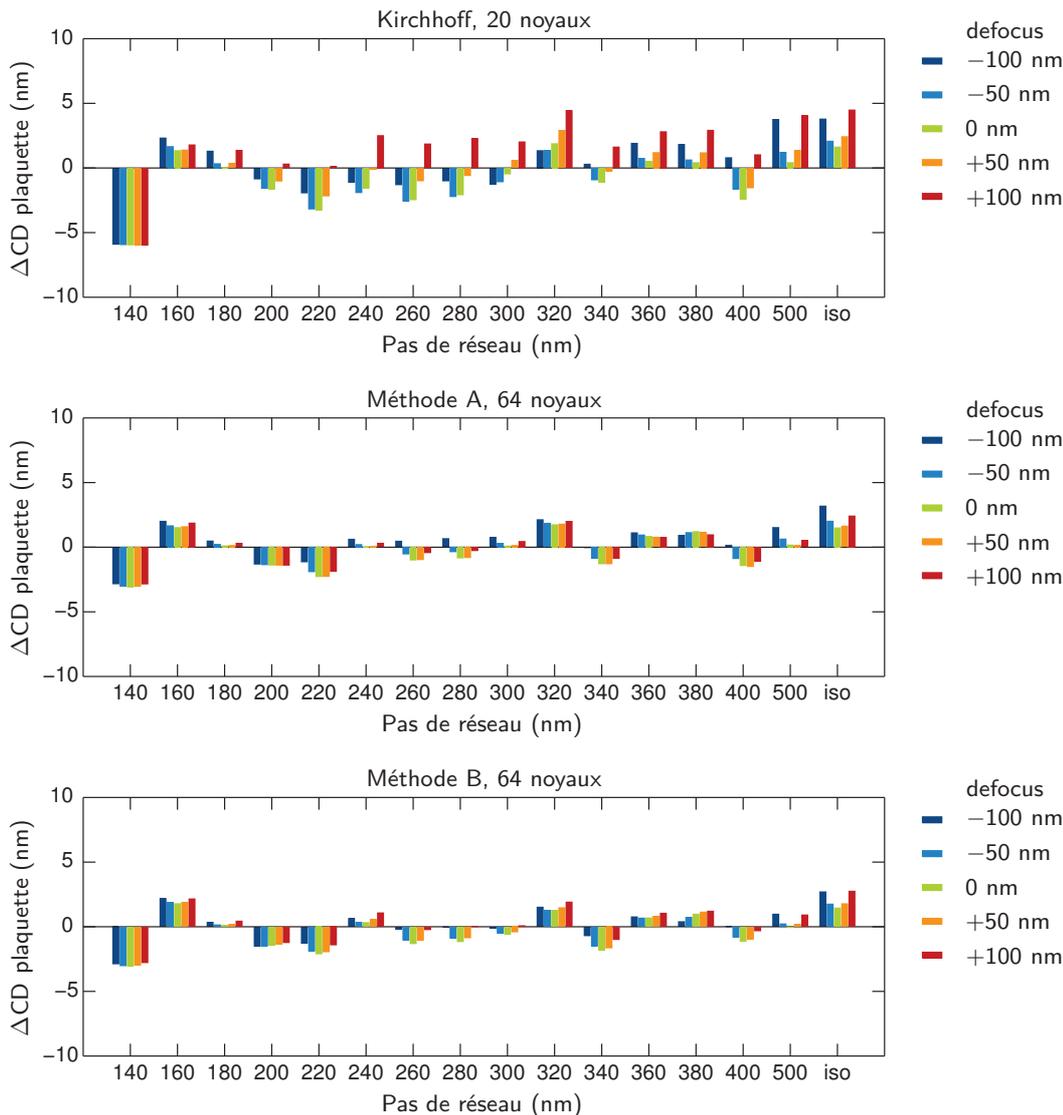


Figure 4.35 – Tracés des erreurs en CD calculées, par rapport à la méthode rigoureuse FDTD, avec l'approximation de Kirchhoff (avec 20 noyaux) et les méthodes approchées A et B (avec 64 noyaux) en fonction du pas de réseau et pour différentes valeurs de la position du plan focal.

4.5.3.6 Prédiction de l'impression des SRAF

Évaluons maintenant la prédiction de l'impression des motifs diffractants non résolus. Pour ce faire, nous avons calculé les seuils d'impression des SRAF correspondant aux minima locaux de l'intensité lumineuse dans la résine. Ensuite, nous avons calculé le budget normalisé en seuil

d'impression¹¹ comme étant la différence entre le minimum local de l'intensité lumineuse au niveau des SRAF et le seuil d'extraction des dimensions critiques dans la résine. Puis, nous avons normalisé le budget par rapport au seuil d'extraction afin d'avoir une meilleure prédiction de l'impression des SRAF en terme de probabilité. Par exemple, un budget normalisé supérieur à 5% nous permet d'écarter le risque d'impression des SRAF avec une certitude de 95%. En revanche, un budget normalisé inférieur à cette valeur permet d'accentuer le risque d'impression.

La figure 4.36 présente le budget normalisé en fonction du pas de réseau dans le cas d'une exposition à la dose nominale, c'est-à-dire, celle qui nous donne un CD sur plaquette égal à 65 nm pour une ligne appartenant à un réseau de pas égal à 180 nm. D'après cette figure, la méthode FDTD prédit un fort risque d'impression des SRAF pour les réseaux de lignes de pas allant de 320 nm à 380 nm. Cependant, les budgets normalisés en seuil d'impression calculés par les méthodes approchées A et B sont positifs et ne permettent pas de confirmer l'impression des SRAF à 100%. Néanmoins, pour une certitude de 95%, ces méthodes permettent de prédire l'impression des SRAF pour les pas de réseau entre 340 nm et 380 nm.

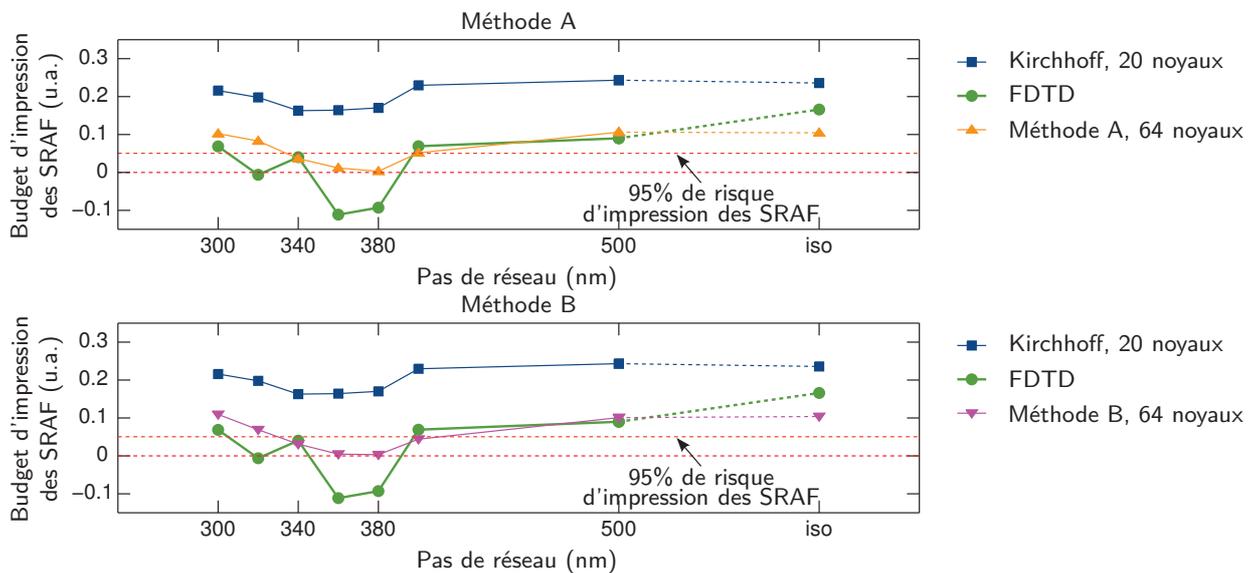


Figure 4.36 – Budgets normalisés d'impression des SRAF calculés par les méthodes approchées A et B (avec 64 noyaux) en fonction du pas de réseau dans le cas d'une exposition à la dose nominale, c'est-à-dire, celle qui donne un CD plaquette égal à 65 nm pour une ligne appartenant à un réseau de pas égal à 180 nm.

4.5.4 Discussion des résultats

À partir des données de simulation recueillies, nous avons évalué les méthodes approchées DDM et M3D par rapport à la méthode rigoureuse FDTD. Les principaux axes de cette évaluation sont les fenêtres de procédé (la profondeur de champ, la position du meilleur focus, etc.), les effets de proximité optique et la prédiction de l'impression des SRAF. Mais rappelons brièvement les objectifs de cette évaluation. Tout d'abord, il s'agit d'évaluer la précision des différentes méthodes approchées quant à la modélisation des effets topographiques du masque non seulement dans les conditions nominales du procédé lithographique mais aussi dans les conditions décalées en focus et en dose. Ensuite, le calcul de la position du meilleur focus nous permettra de mieux situer les erreurs focales liées à la topographie du masque. Car, comme nous l'avons déjà vu, la topographie du masque entraîne des erreurs focales pouvant déstabiliser la position du meilleur focus lorsque l'on travaille avec un unique plan d'observation (ce qui est déjà utilisé dans les

11. Le budget normalisé en seuil d'impression est défini par

$$\frac{I_{\min\text{SRAF}} - \text{seuil}}{\text{seuil}}$$

modèles de traitement OPC). Finalement, la prédiction des motifs diffractants non résolus devient problématique pour les nœuds technologiques 45 nm et au delà. Il devient donc nécessaire d'évaluer la capacité des méthodes approchées à prédire le risque d'impression des SRAF.

En ce qui concerne la précision des méthodes approchées à simuler les dimensions critiques, les résultats obtenus pour différents réseaux de lignes montrent que les méthodes DDM et M3D permettent de réduire les erreurs liées aux effets de la topographie du masque et d'améliorer ainsi l'estimation des dimensions critiques des motifs dans la résine à travers la fenêtre de procédé. Elles permettent également de se rapprocher de la méthode rigoureuse FDTD dans l'estimation de la profondeur de champ et de la position du meilleur focus, ce qui permet de réduire les erreurs focales. Néanmoins, l'efficacité de ces méthodes dépend fortement du pas de réseau considéré. Car, comme nous l'avons déjà vu dans le cas des réseaux de lignes de pas 140 nm et 300 nm, la précision des méthodes approchées DDM et M3D se dégrade quant à l'estimation de la profondeur de champ et de la position du meilleur focus. Ces deux pas de réseau sont particulièrement intéressants car ils présentent les plus petits espacements entre deux lignes successives¹². Les effets optiques dans des réseaux de lignes très proches ne sont donc pas correctement modélisés par une simple somme des champs électriques provenant des bords comme dans le cas des méthodes DDM (dans le domaine spatial) et M3D (dans le domaine fréquentiel). En effet, sur la figure 4.37, nous avons tracé dans le cas d'un réseau de lignes de pas 140 nm les courbes de l'amplitude et de la phase de la composante TE du champ électrique que nous avons reconstruit à la manière de la méthode DDM ainsi que celles obtenues par une simulation directe avec la méthode FDTD. D'après cette figure, nous pouvons clairement remarquer qu'il y a des différences au niveau de l'amplitude du champ électrique proche qui s'accroissent aux centres des lignes et des espaces du réseau. Ces différences d'amplitude entraînent des différences au niveau des profils de l'intensité lumineuse dans la résine. Ce qui explique les résultats obtenus par les méthodes DDM et M3D pour les petits pas de réseau.

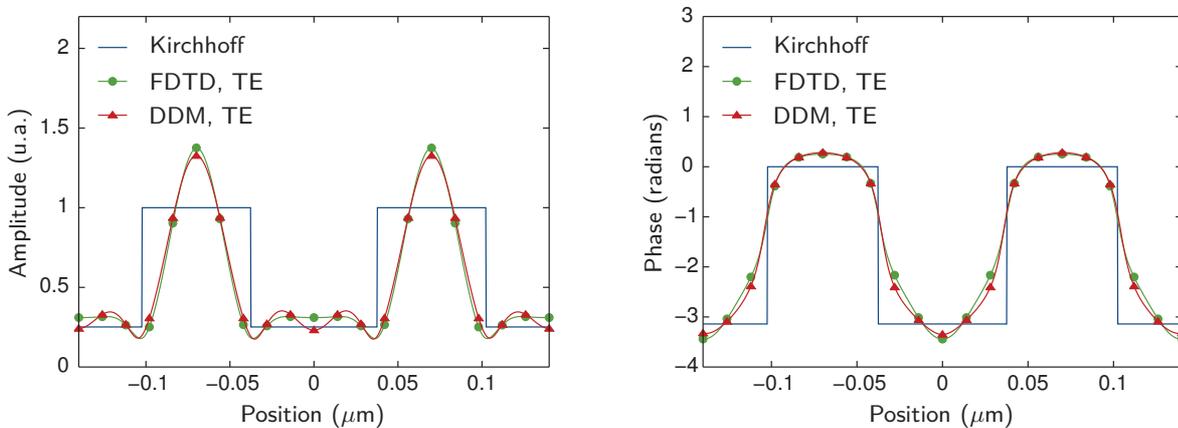


Figure 4.37 – Courbes de l'amplitude (à gauche) et de la phase (à droite) de la composante TE du champ électrique proche d'un masque topographique, calculées par l'approximation de Kirchhoff, la méthode rigoureuse FDTD et la méthode approchée DDM.

Quant à la prédiction de l'impression des SRAF, les méthodes DDM et M3D ne nous permettent pas d'affirmer le risque d'impression avec une certitude absolue. Puisque, comme nous l'avons montré, les budgets normalisés d'impression calculés par ces méthodes approchées pour certains pas de réseau sont légèrement supérieurs à ceux prédits par la méthode FDTD. Cependant, la prédiction de l'impression des SRAF est nettement améliorée par rapport à l'approximation de Kirchhoff.

12. Le pas 140 nm est le plus petit pas de réseau ne présentant pas de motifs diffractants non résolus SRAF et l'espace entre deux lignes successives est le plus petit (de l'ordre de 77 nm). Quant au pas 300 nm, c'est le plus petit pas de réseau présentant des SRAF et qui présente le plus petit espace entre une ligne du réseau et une barre SRAF voisine (de l'ordre de 90 nm)

4.6 Conclusion

Ainsi, dans ce chapitre, nous avons étudié les limites de validité de l'approximation de Kirchhoff, méthode qui, en raison de sa rapidité, est largement utilisée dans les logiciels de traitement OPC actuels pour la modélisation du champ électrique en sortie du masque. Dans cette étude, nous avons comparé par rapport au calcul rigoureux les résultats de simulation de plusieurs structures 1D. Au niveau du champ proche, nous avons montré que la topographie du masque entraîne une diminution de l'amplitude de l'ordre zéro et augmente celles des ordres ± 1 . De plus, les erreurs engendrées par l'approximation de Kirchhoff au niveau des spectres de Fourier sont d'autant plus élevées que les dimensions des motifs sont réduites par rapport à la longueur d'onde et l'ouverture numérique est élevée. Au niveau de la plaque, les différences en amplitude et en phase des ordres qui sont captées par la pupille se traduisent par des écarts plus ou moins importants au niveau des profils de l'intensité lumineuse. Les erreurs en intensité sont pénalisantes car elles nous conduisent à des calculs erronés de la profondeur de champ, de la position du meilleur focus et du contraste, mais aussi au niveau des dimensions critiques des motifs dans la résine et de la prédiction de l'impression des SRAF.

Nous avons également évalué les deux méthodes approchées semi-rigoureuses pour la modélisation des effets topographiques du masque : la méthode de décomposition de domaine (DDM) et celle du masque tridimensionnel (M3D). Ces méthodes ont été récemment développées et sont actuellement disponibles dans les logiciels de traitement OPC que nous avons utilisés pendant cette thèse. Les principes de fonctionnement de ces deux méthodes présentent quelques similitudes. En effet, en écrivant le champ proche rigoureux d'un bord diffractant en la somme d'un champ scalaire de Kirchhoff et d'autres fonctions erreurs de transmission causées par la topographie, on arrive à reconstruire le champ électrique "rigoureux" d'un motif de géométrie quelconque, par une somme pondérée des contributions des bords adjacents. Les résultats de l'évaluation de ces méthodes à travers la fenêtre de procédé pour différentes structures 1D sont plutôt satisfaisants. En effet, les méthodes DDM et M3D permettent de réduire les erreurs focales et conduisent à une meilleure estimation de la profondeur de champ ainsi que de la position du meilleur focus. Cependant, la précision du calcul dépend des dimensions des motifs. Car comme nous l'avons montré, les effets topographiques du masque dans des réseaux de lignes très proches (l'espace entre deux lignes est de l'ordre de 80 nm) ne sont pas correctement modélisés par une simple somme des champs électriques provenant des bords. Les réseaux de pas 140 nm et 300 nm sont en effet deux cas rencontrés où les erreurs au niveau de l'intensité lumineuse sont plus importantes que dans les autres cas de réseaux. D'autres paramètres de modélisation peuvent également intervenir dans la précision de ces méthodes, comme par exemple le nombre de noyaux utilisés dans les tables optiques. En effet, afin d'obtenir de meilleures précisions, il est conseillé d'utiliser un nombre élevé de noyaux (par exemple 64 noyaux). Ce qui rend ces méthodes assez gourmandes en ressources informatiques.



Chapitre 5

Correction des effets topographiques du masque

Bien qu'elles permettent d'améliorer la précision dans la modélisation des effets topographiques du masque, les méthodes semi-rigoureuses industrielles que nous avons présentées et évaluées dans le chapitre précédent sont assez gourmandes en ressources informatiques et nécessitent des temps de calcul plus élevés. En revanche, l'avantage des anciens algorithmes utilisant l'approximation de Kirchhoff c'est qu'ils sont rapides et les techniques existantes (comme l'optique de Fourier ou la création et la troncature des tables optiques) sont mieux adaptées à l'approche scalaire où une unique valeur de transmission et de phase permet de définir la fonction du masque. Afin de garder l'approximation de Kirchhoff dans la modélisation du masque, il existe des méthodes heuristiques rapides permettant de capturer une partie des effets topographiques, agissant ainsi comme une fonction de transfert intermédiaire entre l'optique de Kirchhoff et le modèle empirique de la partie résine.

Parmi les techniques envisagées, la correction des effets topographiques par un second modèle empirique placé en amont de la partie résine permet d'absorber une partie des erreurs causées par les nombreuses approximations de la partie optique, et plus particulièrement celle de Kirchhoff. Cependant, le modèle empirique résultant sera plus complexe et sa stabilité deviendra compromise.

Sommaire

5.1	Introduction	111
5.2	Méthodes heuristiques	111
5.2.1	Correction par adaptation de la dose	112
5.2.2	Correction par un modèle empirique	113
5.2.2.1	Introduction d'un facteur de dose	114
5.2.2.2	Régression quadratique	115
5.2.2.3	Optimisation des modèles	117
5.2.2.4	Précision et stabilité	118
5.2.3	Amélioration de la prédiction de l'impression des SRAF	119
5.2.3.1	Correction par réduction de la taille des SRAF	119
5.2.3.2	Détermination des spécifications des dimensions des SRAF	120
5.2.3.3	Correction par modification de la transmission et de la phase des SRAF	122
5.3	Méthode de modélisation par décomposition multi-niveaux du masque	
	fin	123
5.3.1	Principe	124
5.3.1.1	Expression du champ scalaire effectif	124
5.3.1.2	Décomposition multi-niveaux du champ électrique	125
5.3.2	Évaluation par rapport au calcul rigoureux	128
5.3.2.1	Définition des conditions de simulation	128
5.3.2.2	Fenêtres de procédé et courbes de Bossung	131
5.3.2.3	Position du meilleur focus	132
5.3.2.4	Profondeur de champ	132
5.3.2.5	Effets de proximité	132
5.3.2.6	Prédiction de l'impression des SRAF	134
5.3.2.7	Ressources nécessaires et temps d'exécution	134
5.3.2.8	Intégration en traitement OPC	135
5.4	Conclusion	136

5.1 Introduction

La précision dans la modélisation OPC est devenue d'une importance capitale afin de relever les défis de la feuille de route de l'ITRS. Mais au delà de la précision, les modèles OPC doivent être rapide afin de réduire les coûts quant à la préparation et à la fabrication des masques. À l'heure actuelle, les industriels n'ont qu'un objectif : trouver un compromis entre la précision et le temps de cycle. Certes, l'introduction dans les logiciels de traitement OPC de méthodes semi-rigoureuses pour la modélisation des effets topographiques du masque permet de gagner en précision. Mais ces méthodes sont plus gourmandes en temps et en ressources informatiques sans pour autant être exactes. De plus, les algorithmes utilisant l'approximation de Kirchhoff sont rapides et les techniques développées dans ce domaine (par exemple les matrices TCC) ne sont utilisables que dans le cadre de la théorie scalaire utilisant l'optique de Fourier.

Dans ce chapitre, nous proposons différentes façons de corriger les effets topographiques du masque, tout en gardant l'approximation de Kirchhoff qui permet de limiter le temps de calcul dans la modélisation de la partie optique. Parmi les méthodes proposées, nous exploiterons celle permettant de réduire les erreurs liées aux effets topographiques du masque par l'intermédiaire d'un second modèle empirique. Nous montrerons aussi que pour garantir une précision adéquate, il est nécessaire d'augmenter la complexité du modèle en rajoutant des termes additionnels, ce qui peut entraîner une dégradation de sa stabilité. Enfin, pour garantir la stabilité numérique du modèle empirique, nous introduirons une nouvelle méthode approchée par décomposition multi-niveaux du masque fin, moins coûteuse en temps de calcul et permettant de modéliser les effets topographiques directement dans la partie optique. Nous présenterons les limites de cette méthode par la comparaison des résultats de simulation avec l'outil rigoureux (FDTD).

5.2 Méthodes heuristiques

Dès lors que le problème des effets topographiques du masque commence à se poser, les ingénieurs commencent à chercher des techniques de correction dites heuristiques ou intuitives. Ce sont des méthodes approximatives pratiques qui peuvent être rapidement réalisables et fournir des solutions qui ne sont pas nécessairement optimales. L'utilisation de telles méthodes nous permet d'approcher les effets topographiques du masque afin de réduire les erreurs optiques liées à l'approximation de Kirchhoff. Étant des techniques approchées, leurs domaines de validité sont limités car ils dépendent des paramètres du procédé et de la complexité du circuit à traiter. Néanmoins, elles sont largement exploitées. D'ailleurs, nous pouvons citer quelques exemples de la littérature :

- Correction par réduction des dimensions des motifs sur le masque (biais) et/ou par modification de la transmission du masque [126].
- Correction par rajout de bords diffractants autour des motifs du masque [121, 127, 128].
- Prise en compte des effets topographiques dans la partie résine du modèle [78, 129].

Bien que ces méthodes soient faciles à mettre en place, elles nécessitent une étude préalable afin d'optimiser les multiples paramètres qui sont en jeu. En effet, la réduction des dimensions des motifs sur le masque nécessite une évaluation des dimensions à enlever par bord en fonction de la taille de la structure et de la distance la séparant des motifs voisins. La modification de la transmission des motifs nécessite un traitement appliqué au dessin du masque et un découpage en zones de différentes transmissions. Le modèle de couche limite BL (de l'anglais "*Boundary Layer*") qui consiste à rajouter des bords diffractants aux motifs, comme le montre la figure 5.1, nécessite l'optimisation de sa taille et de sa fonction imaginaire de transmission (fonction de phase). Finalement, l'utilisation d'un modèle empirique afin de compenser les effets topographiques du masque requiert des modèles de complexité supérieure dégradant ainsi leur stabilité.

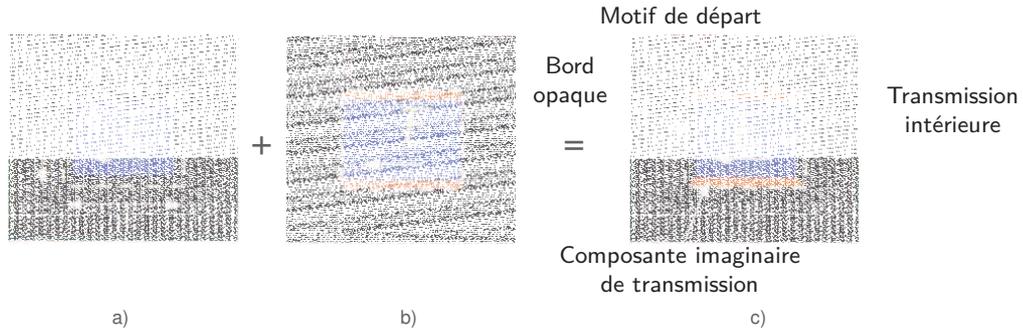


Figure 5.1 – Présentation du modèle de couche limite : a) composante réelle de transmission, b) composante imaginaire de transmission, c) résultat de la superposition des deux composantes réelle et imaginaire. Sources : [121, 127, 128].

Dans la suite, nous allons exploiter la méthode permettant de réduire les effets topographiques du masque par l'intermédiaire d'un modèle à seuil variable. Nous montrerons que pour garantir une précision adéquate du modèle, il est nécessaire d'augmenter sa complexité en rajoutant d'autres termes. Finalement, nous exploiterons la méthode de la réduction des dimensions, puis de la modification de la transmission et de la phase, afin d'améliorer la prédiction du seuil d'impression dans la résine des motifs diffractants non résolus (SRAF).

5.2.1 Correction par adaptation de la dose

Nous allons essayer de reproduire, dans le cas de l'approximation de Kirchhoff, la baisse de contraste de l'intensité lumineuse par la méthode FDTD observée dans des cas particuliers de motifs. Pour cela, nous avons simulé l'intensité lumineuse dans la résine d'une région du masque contenant deux blocs logique et SRAM de superficies respectives $31 \times 29 \mu\text{m}^2$ et $16 \times 14 \mu\text{m}^2$ (à l'échelle de la plaquette). Les simulations ont été effectuées avec l'approximation de Kirchhoff et la méthode FDTD sur des sites de contrôle que nous avons positionnés sur les bords des motifs perpendiculairement aux contours de la résine. Pour chacun des sites, nous avons calculé par la méthode des moindres carrés le meilleur coefficient de proportionnalité permettant de réduire l'écart entre le profil d'intensité simulé par la méthode FDTD et celui simulé par l'approximation de Kirchhoff multiplié par ce même coefficient. La figure 5.2 illustre un exemple de simulations montrant que nous pouvons pratiquement obtenir une coïncidence parfaite entre les deux profils de l'intensité lumineuse dans la résine en multipliant celui obtenu avec l'approximation de Kirchhoff avec un facteur de dose égal à 0.9062.

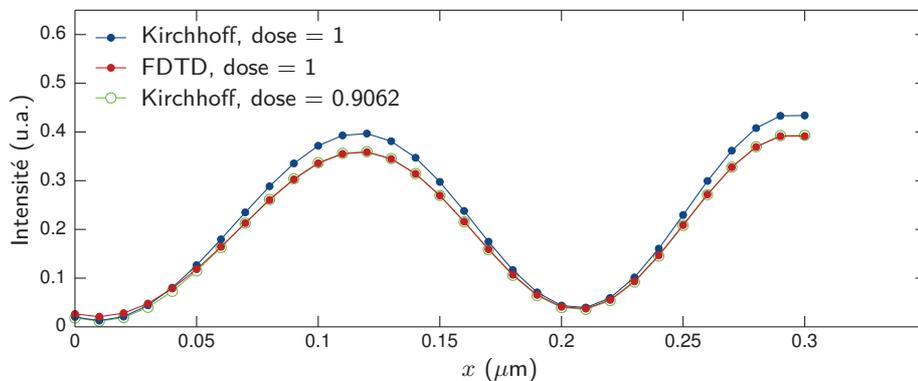


Figure 5.2 – Exemple de profils de l'intensité de l'image aérienne calculés en utilisant l'approximation de Kirchhoff et la méthode rigoureuse FDTD. Le profil corrigé est calculé en multipliant l'intensité obtenue avec l'approximation de Kirchhoff par un facteur de dose égal à 0.9062. Ce profil coïncide bien avec celui obtenu avec la méthode FDTD.

Après avoir éliminé les sites redondants¹, nous avons tracé sur la figure 5.3 l’histogramme des facteurs de dose calculés pour tous les sites de contrôle. D’après cette figure, la distribution du facteur de dose est de forme gaussienne centrée sur la valeur 0.9 et présentant un écart type égal à 0.0263. Ce qui nous donne un intervalle de dose qui s’étend de 0.82 ($0.9 - 3\sigma$) à 0.98 ($0.9 + 3\sigma$). La limite supérieure du facteur de dose correspond à des structures où les effets topographiques sont négligeables, c’est-à-dire, ayant des dimensions bien supérieures à la longueur d’onde λ qui est égale à 193 nm. Quant à la limite inférieure du rapport de dose, celle-ci correspond à des structures critiques où les effets topographiques du masque sont importants. Ces structures sont donc prioritaires dans la calibration des modèles empiriques.

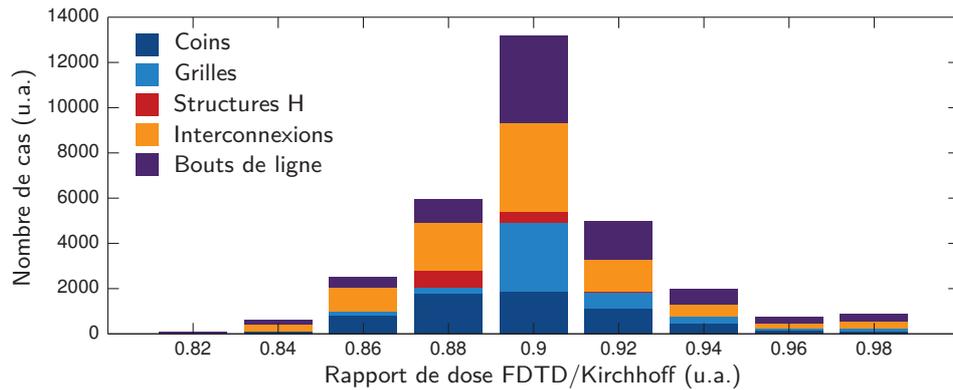


Figure 5.3 – Rapport de dose FDTD/Kirchhoff calculé de manière à coïncider l’image aérienne obtenue par Kirchhoff avec celle obtenue avec la méthode FDTD. Dans les deux cas, l’image aérienne est calculée sur les mêmes sites de simulation qui sont positionnés de manière perpendiculaire au contour de la résine. L’écart type de la distribution est égal à 0.0263.

Vu l’étendue de l’intervalle de dose, le fait d’appliquer une valeur unique du facteur de dose aux profils de l’intensité lumineuse simulés par l’approximation de Kirchhoff ne nous permet donc pas de modéliser les effets topographiques du masque d’une manière efficace. Car ceci se ramène à une simple modification du seuil d’extraction du modèle empirique utilisé. Nous allons donc essayer de réduire l’écart entre les dimensions critiques simulés par les deux méthodes en utilisant un modèle à seuil variable.

5.2.2 Correction par un modèle empirique

Nous avons créé deux modèles à seuil variable (VT5) en introduisant ou non les effets topographiques du masque dans la partie optique. La figure 5.4 présente les valeurs des coefficients c_k des termes des modèles obtenus. D’après cette figure, nous pouvons clairement remarquer que le coefficient c_0 correspondant au seuil du modèle optique pur est plus faible avec un masque topographique que pour l’approximation de Kirchhoff. Le rapport des deux seuils est égal à 0.84 qui est inférieur à la moyenne 0.9 du facteur de dose. Ceci est dû au fait que nous avons utilisé des motifs critiques pour la calibration des deux modèles. Concernant les autres termes, nous remarquons que les coefficients des termes croisés Slope \times Imax et Slope \times Imin sont plus faibles avec un masque topographique. Ce qui implique que le modèle est plus proche à l’optique pur et est par conséquent plus stable.

1. Ce sont les sites qui présentent les mêmes valeurs de Imax, Imin, Slope et Factor

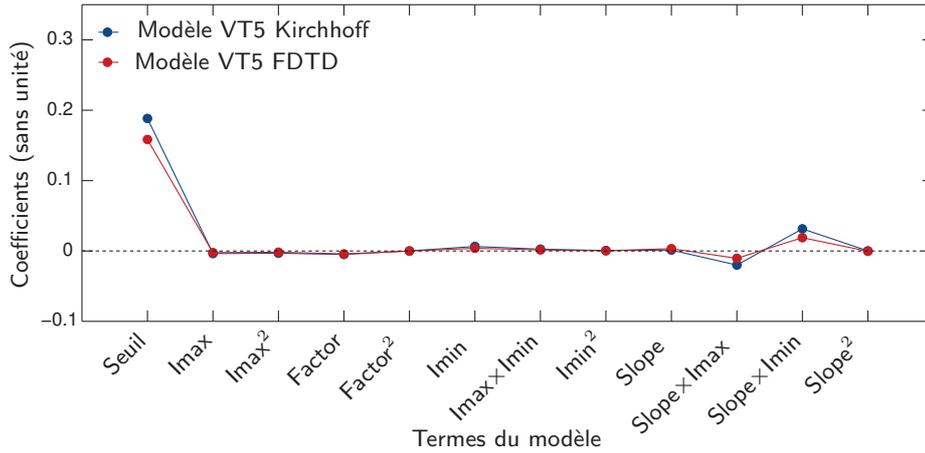


Figure 5.4 – Tracé des coefficients des modèles VT5 créés en introduisant ou pas les effets topographiques du masque dans la partie optique.

Le but de la correction est de trouver une méthode permettant d'utiliser le modèle empirique, obtenu en introduisant des effets topographiques du masque, avec l'intensité lumineuse simulée avec l'approximation de Kirchhoff. Cela consiste à trouver une fonction de transfert qui permet de rapprocher le profil de l'intensité lumineuse simulée avec l'approximation de Kirchhoff à celui obtenu avec un masque topographique, comme le montre le schéma de la figure 5.5

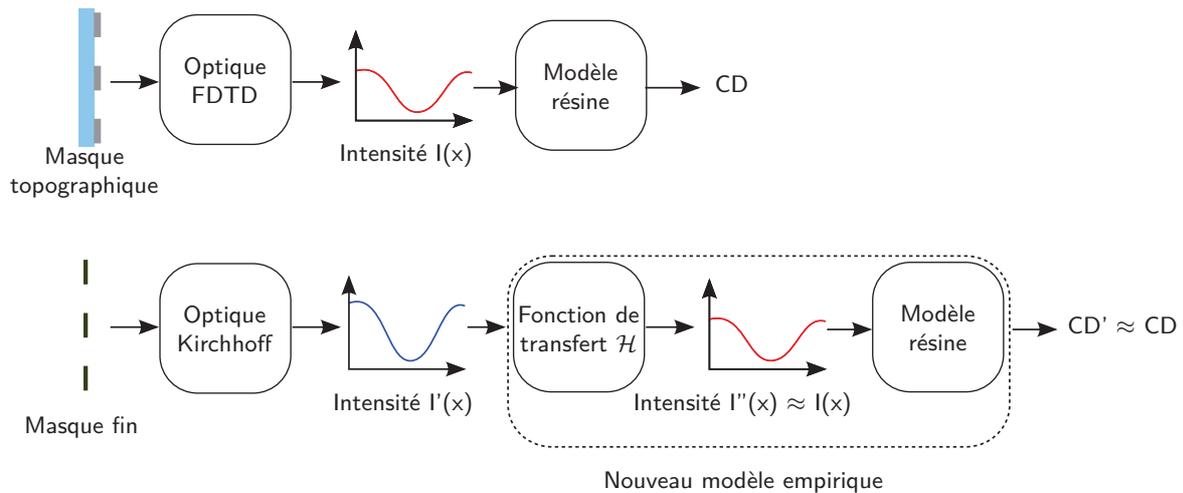


Figure 5.5 – Méthode permettant d'inclure les effets topographiques dans la partie résine du modèle. Une fonction de transfert permet de rapprocher l'intensité lumineuse calculée par l'approximation de Kirchhoff à celle obtenue avec la méthode FDTD.

5.2.2.1 Introduction d'un facteur de dose

L'introduction d'un facteur de dose égal à 0.9 dans la partie optique du modèle créé avec l'approximation de Kirchhoff est similaire au fait de multiplier les paramètres I_{max} , I_{min} , et $Slope$ par 0.9 (la courbure $Factor$ n'est pas affectée puisque son calcul est effectué de la même manière). Nous avons donc appliqué le modèle VT5 créé avec un masque topographique aux profils de l'intensité lumineuse obtenus avec l'approximation de Kirchhoff et un facteur de dose égal à 0.9. Dans ce cas, notre fonction de transfert \mathcal{H} , appliquée à I_{max} , I_{min} , $Slope$ et $Factor$, s'exprime ainsi :

$$\mathcal{H} \begin{pmatrix} I_{max} \\ I_{min} \\ Slope \\ Factor \end{pmatrix} = \begin{pmatrix} 0.9 I_{max} \\ 0.9 I_{min} \\ 0.9 Slope \\ Factor \end{pmatrix} \tag{5.1}$$

Dans le tableau 5.1, nous avons récapitulé les trois modèles OPC de cette étude. Le modèle A inclut les effets topographiques dans la partie optique et utilise un modèle empirique 1 qui est supposé ne pas prendre en compte les effets topographiques. Le modèle B utilise l'approximation de Kirchhoff et tous les effets topographiques sont gérés dans le modèle empirique 2. Finalement, le modèle C est celui que l'on veut évaluer. Il inclut une partie des effets topographiques du masque en injectant un facteur de dose égal à 0.9 dans sa partie optique qui est gérée par l'approximation de Kirchhoff. Son modèle empirique est celui du modèle A et est donc supposé ne pas inclure les effets topographiques. Par cette démarche, nous allons vérifier si le facteur de dose de 0.9 est suffisant ou pas afin de prendre en compte la majorité des effets topographiques du masque, mais aussi évaluer l'efficacité de la correction. Pour ce faire, nous avons calculé les écarts ΔEPE entre les modèles A et C et les comparer à ceux obtenus entre les modèles A et B.

Modèle OPC	Optique	Résine
A	FDTD	Modèle 1
B	Kirchhoff	Modèle 2
C	Kirchhoff, dose = 0.9	Modèle 1

Tableau 5.1 – Descriptif des modèles OPC étudiés

La figure 5.6 présente les distributions de ΔEPE calculées pour plus de $30 \cdot 10^3$ sites de contrôle. D'après cette figure, le modèle C permet de mieux recentrer la distribution autour de la valeur zéro mais dégrade la dispersion des erreurs. En effet, la moyenne et l'écart type de la distribution de ΔEPE entre les modèles A et C sont de -0.12 nm et 1.16 nm respectivement, tandis que ceux obtenus entre les modèles A et B sont de -0.89 nm et 0.81 nm. Nous concluons donc qu'une simple modification de la dose dans la partie optique ne suffit pas pour permettre à la partie résine d'absorber la majorité des effets topographiques du masque.

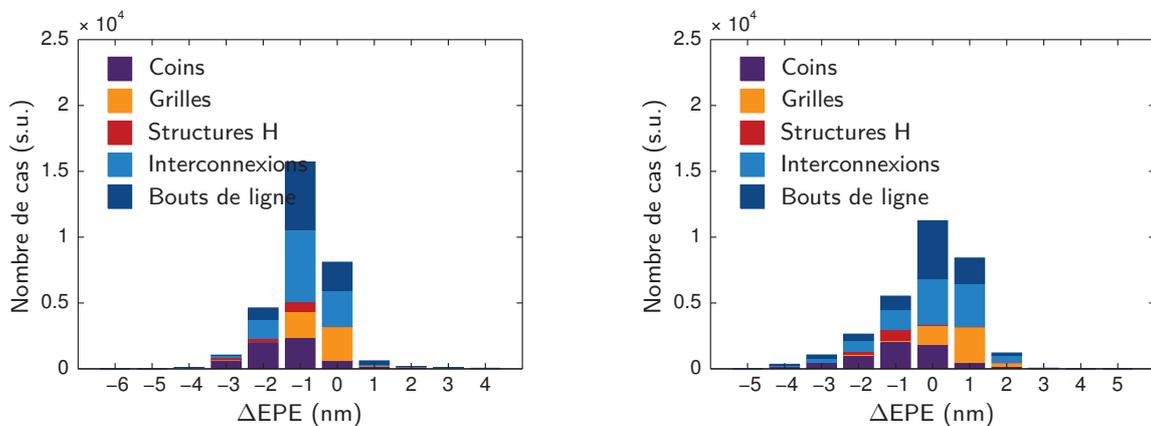


Figure 5.6 – Distributions des écarts ΔEPE entre les modèles A et B (moyenne -0.89 nm, écart type 0.81 nm) (à gauche) et les modèles A et C (moyenne -0.12 nm, écart type 1.16 nm) (à droite). Les modèles A, B et C sont tels que définis par le tableau 5.1.

5.2.2.2 Régression quadratique

Dans le but de réduire l'écart entre les dimensions critiques simulées par l'approximation de Kirchhoff et la méthode FDTD, nous avons calculé les courbes de régression à partir des paramètres extraits sur plus de $30 \cdot 10^3$ sites de contrôle. La figure 5.7 montre que l'on peut trouver des corrélations simples de type quadratique entre ces différents paramètres. Les valeurs des coefficients de détermination R^2 , proches de 1, montre que la qualité de la régression est satisfaisante pour 95% des sites au minimum.

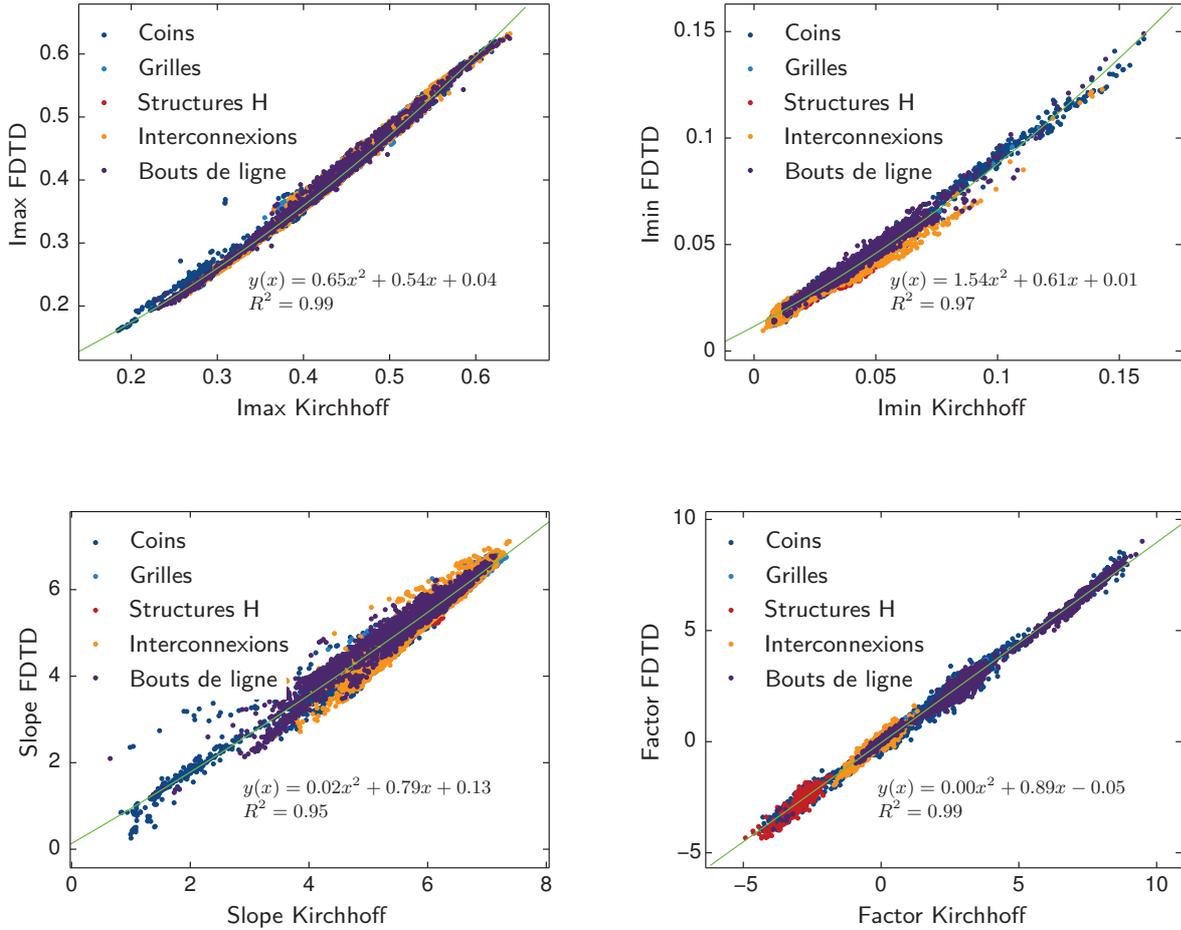


Figure 5.7 – Régression quadratique des paramètres $\{I_{max}, I_{min}, Slope \text{ et } Factor\}$ simulés par l'approximation de Kirchhoff et la méthode FDTD.

En remplaçant les valeurs des paramètres $\{I_{max}, I_{min}, Slope \text{ et } Factor\}$ par leurs équations de régression, nous pouvons inclure une partie des effets topographiques du masque dans la partie résine du modèle. Notre fonction de transfert \mathcal{H} , appliquée à I_{max} , I_{min} , $Slope$ et $Factor$, s'exprime désormais de la façon suivante :

$$\mathcal{H} \begin{pmatrix} I_{max} \\ I_{min} \\ Slope \\ Factor \end{pmatrix} = \begin{pmatrix} 0.65 I_{max}^2 + 0.54 I_{max} + 0.04 \\ 1.54 I_{min}^2 + 0.61 I_{min} + 0.01 \\ 0.02 Slope^2 + 0.79 Slope + 0.13 \\ 0.00 Factor^2 + 0.89 Factor - 0.05 \end{pmatrix}$$

Nous obtenons donc un nouveau modèle VT5 d'ordre quatre, que l'on appelle modèle D. La figure 5.8 présente la distribution des différents ΔEPE entre les modèles A et D. D'après cette figure, la distribution admet une moyenne égale à -0.03 nm et un écart type égal à 1.03 nm. Nous obtenons donc de meilleurs résultats avec ce modèle par rapport au modèle C. Cependant, la dispersion de l'erreur reste plus importante que dans le cas du modèle B.

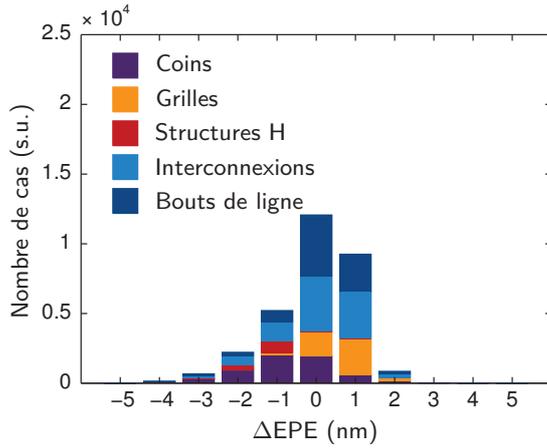


Figure 5.8 – Distribution des écarts ΔEPE entre les modèles A et D (moyenne -0.03 nm, écart type 1.03 nm)

5.2.2.2.1 Récapitulatif des résultats

Récapitulons les résultats trouvés avec ces deux démarches :

- La modulation de la dose (en injectant un facteur de dose égal à 0.9 dans l'approximation de Kirchhoff) ne corrige les effets topographiques que partiellement ; ce qui impacte la précision du modèle empirique qui devrait absorber une partie de ces effets.
- L'augmentation de la complexité du modèle en introduisant les termes quadratiques n'est pas suffisante dans la modélisation des effets topographiques, car elle ne fait que recentrer l'erreur en CD mais dégrade la précision du modèle empirique (qui sera plus dispersée).

Ces résultats montrent que l'on ne peut pas séparer les effets optiques de la résine et qu'il faut s'orienter plutôt vers l'optimisation conjointe de ces deux parties. Nous pensons donc à re-optimiser les coefficients c_k des modèles C et D en se basant sur les informations apportées par les $30 \cdot 10^3$ sites simulés.

5.2.2.3 Optimisation des modèles

Nous avons calculé le seuil donné par le modèle A (voir tableau 5.1) pour tous les sites de contrôle que nous disposons (environ $30 \cdot 10^3$ sites). Afin d'optimiser les coefficients du modèle, nous avons à résoudre un système linéaire à m équations (m étant le nombre de sites) que nous écrivons sous cette forme matricielle :

$$\begin{pmatrix} 1 & T_{1,1} & \dots & T_{1,n-1} & T_{1,n} \\ \vdots & & \ddots & & \vdots \\ 1 & T_{m,1} & \dots & T_{m,n-1} & T_{m,n} \end{pmatrix} \begin{pmatrix} c_0 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} S_1 \\ \vdots \\ S_m \end{pmatrix} \quad (5.2)$$

où n désigne le nombre des termes du modèle, S_j les seuils donnés pas le modèle A, c_i les coefficients des termes à optimiser et $\{T_{i,j}\}$ la valeur du $i^{\text{ème}}$ terme calculée pour le $j^{\text{ème}}$ site de contrôle. En posant des contraintes limitant la divergence de chacun des coefficients (par exemple c_0 doit être compris entre 0.1 et 0.2 et les coefficients c_k des autres termes doivent être inférieurs à 0.02 en valeur absolue), nous avons pu résoudre, par la méthode des moindres carrés le système d'équations pour les modèles C et D. La figure 5.9 présente les distributions des différences ΔEPE entre ces modèles optimisés et le modèle A de référence dont la partie optique inclut les effets topographiques du masque. D'après cette figure, nous remarquons une nette amélioration des résultats avec ces modèles optimisés. En effet, la moyenne et l'écart type de la distribution de ΔEPE entre les modèles A et C optimisé sont de 0.02 nm et 0.60 nm respectivement. Le modèle D optimisé permet de gagner un peu de précision par rapport au modèle C avec une distribution un peu moins dispersée et un écart type égal à 0.58 nm.

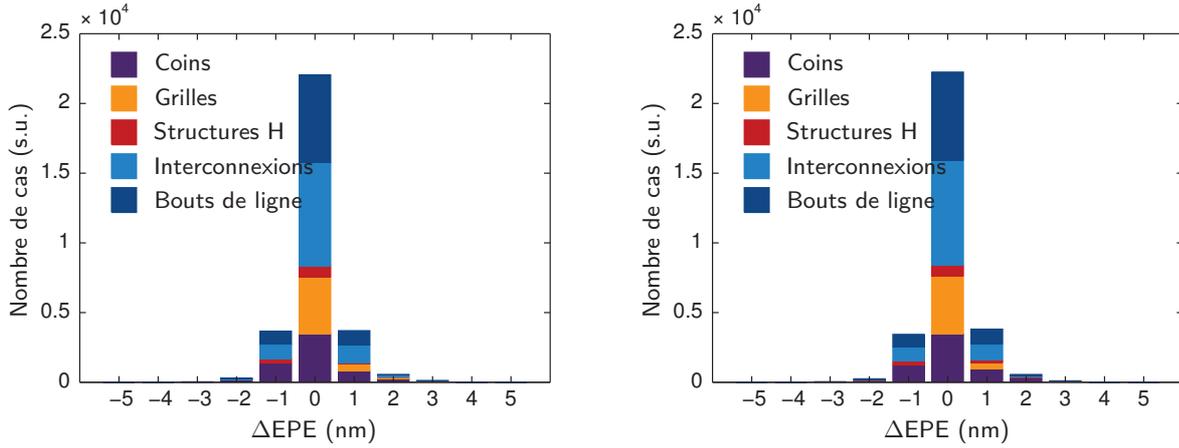


Figure 5.9 – Distribution des écarts ΔEPE entre les modèles A et C avec des coefficients optimisés (moyenne 0.02 nm, écart type 0.60 nm) (à gauche) et les modèles A et D avec des coefficients optimisés également (moyenne 0.03 nm, écart type 0.58 nm) (à droite)

5.2.2.4 Précision et stabilité

Dans l'objectif d'étudier la précision et la stabilité du modèle C, nous avons optimisé les coefficients c_k des termes empiriques en faisant varier les contraintes qui y sont imposées de 0.01 à 0.1 (en valeur absolue). La figure 5.10 présente les coefficients des modèles obtenus. D'après cette figure, il est clair que les coefficients du modèle tendent à diverger lorsque la valeur de la contrainte imposée est plus élevée. Les termes les plus sensibles (ceux que nous avons marqués avec une flèche) sont I_{max} , I_{max}^2 , I_{min} , I_{min}^2 . Les termes croisés $I_{max} \times I_{min}$ et $Slope \times I_{min}$ présentent une moindre sensibilité. La divergence des coefficients des termes empiriques du modèle entraîne la dégradation de sa stabilité. Car comme nous l'avons déjà montré, dans le chapitre 2 de cette thèse, le modèle s'éloigne de l'optique pur et devient dépendant des paramètres du procédé.

En ce qui concerne la précision des modèles obtenus, nous avons répertorié dans le tableau 5.2 les valeurs des moyennes et des écarts-types des distributions ΔEPE entre le modèle A et le modèle C optimisé. Il est clair qu'à partir d'une certaine configuration des contraintes imposées, la capacité du modèle à absorber les effets topographiques tend à disparaître. Le modèle atteint alors une certaine saturation et sa précision (en terme d'écart type), bien qu'elle soit améliorée, devient constante.

	$\ c_k\ \leq$	0.01	0.02	0.03	0.05	0.08	0.1
Moyenne (nm)		0.06	0.02	0.03	0.03	0.03	0.03
Écart type (nm)		0.71	0.60	0.59	0.59	0.59	0.59

Tableau 5.2 – Moyennes et écarts-types des distributions ΔEPE entre le modèle A et le modèle C que nous avons optimisé en imposant des contraintes aux coefficients c_k des termes empiriques.

Dans une autre étude, nous avons volontairement augmenté la complexité du modèle en prenant les termes d'ordre 4 (l'exemple du modèle D). Cependant, nous avons obtenu des résultats similaires en terme de précision. Il convient donc à dire qu'un modèle empirique n'est pas capable d'absorber complètement les effets topographiques. Ce résultat a été confirmé par Michael C. Lam et Konstantinos Adam dans leurs travaux sur la correction des effets topographiques du masque par l'intermédiaire d'un modèle empirique [129].

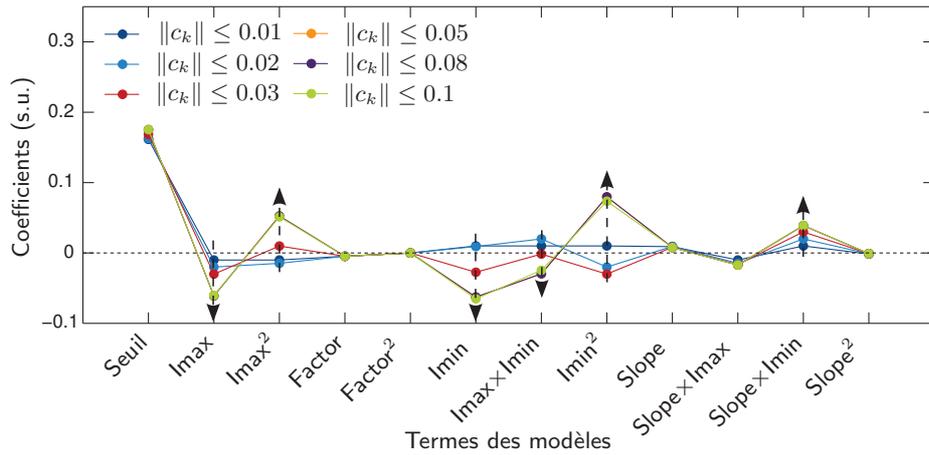


Figure 5.10 – Tracés des coefficients c_k des termes empiriques des modèles VT5 obtenus après optimisation. Les contraintes que nous avons imposées aux coefficients c_k varient de 0.01 à 0.1 (en valeur absolue).

5.2.3 Amélioration de la prédiction de l'impression des SRAF

5.2.3.1 Correction par réduction de la taille des SRAF

Nous allons maintenant essayer de corriger la taille des SRAF pour qu'un modèle utilisant l'approximation de Kirchhoff puisse détecter leur impression. Pour ce faire, nous avons pris le cas d'un réseau de lignes de pas égal à 360 nm (CD égal à 97 nm) et simulé, par l'approximation de Kirchhoff et la méthode FDTD, les intensités lumineuses dans la résine en faisant varier les dimensions des SRAF de 22 nm à 34 nm.

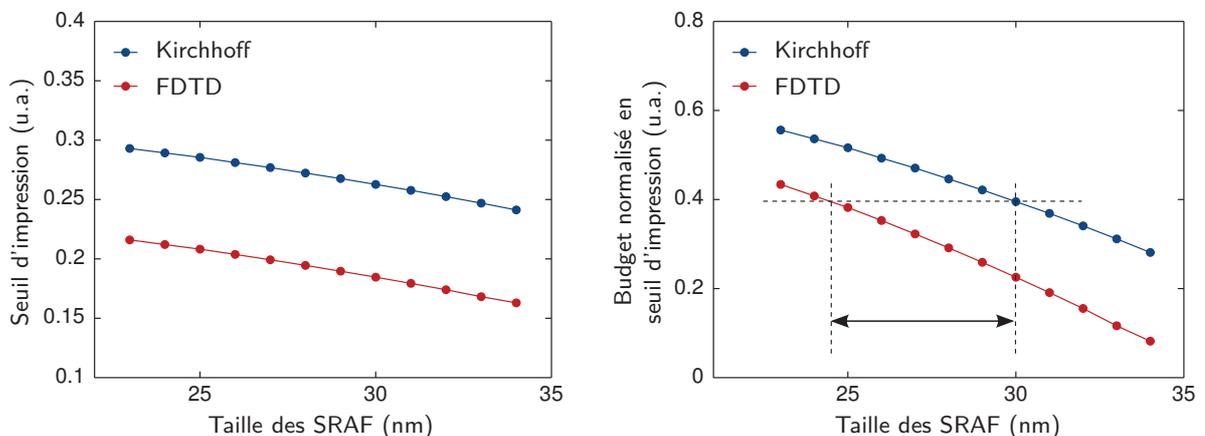


Figure 5.11 – Courbes du budget normalisé en seuil d'impression des motifs diffractants non résolus en fonction de leurs tailles (à l'échelle de la plaquette), calculées avec l'approximation de Kirchhoff et la méthode rigoureuse FDTD.

Sur la figure 5.11, nous avons tracé le seuil d'impression (I_{\min}) des SRAF en fonction de leurs dimensions respectives. Puis par un modèle à seuil constant², nous avons calculé le budget normalisé en seuil d'impression³ que nous avons tracé sur cette même figure. Il est clair que si

2. Les seuils d'extraction pour l'approximation de Kirchhoff et la méthode FDTD ont été calculés pour viser un CD égal à 65 nm pour le réseau de lignes de pas égal à 180 nm.

3. Le budget normalisé en seuil d'impression est défini par

$$\frac{I_{\min_{\text{SRAF}}} - \text{seuil}}{\text{seuil}}$$

nous souhaitons avoir une prédiction de l'impression des SRAF par l'approximation de Kirchhoff similaire à celle de la méthode FDTD, il faut dessiner les SRAF avec une réduction de taille d'environ 5 à 6 nm sur le masque. Cette réduction nous permet d'avoir un budget en seuil d'impression prédit par l'approximation de Kirchhoff équivalent à celui prédit par la méthode FDTD. Il est à noter ici que les courbes de budget en seuil d'impression de cette figure sont calculées pour les deux seuils d'extraction que nous avons obtenus avec nos paramètres et peuvent donc changer pour d'autres seuils et paramètres.

5.2.3.2 Détermination des spécifications des dimensions des SRAF

Nous allons déterminer l'intervalle de variation des dimensions des motifs diffractants non résolus (SRAF) qui respectent les spécifications suivantes :

- Erreurs au niveau du masque : les dimensions d'une ligne dans un réseau de pas égal à 360 nm (CD_{masque}) varient de 95 nm à 99 nm sur le masque.
- Variabilité du procédé : les dimensions acceptées de la ligne dans la résine pour ce même réseau de pas égal à 360 nm ($CD_{\text{plaquette}}$) doivent être comprises entre 62 nm et 68 nm.

Nous allons étudier les cas suivants :

- Cas nominal en dose défini par rapport au réseau de lignes de pas égal à 180 nm : $CD_{\text{masque}} = 77$ nm (masque) et $CD_{\text{plaquette}} = 65$ nm (résine).
- Le plus mauvais cas en dose défini par rapport au réseau de lignes de pas égal à 180 nm : $CD_{\text{masque}} = 75$ nm (c'est-à-dire une réduction de 2 nm sur le masque) et $CD_{\text{plaquette}} = 67$ nm (c'est-à-dire une augmentation de 2 nm dans la résine).

Pour les simulations de cette étude, nous avons utilisé un masque PSM à 6% de transmission et à décalage de phase égal à 180 degrés, une ouverture numérique ON égale à 1.05, une illumination non polarisée de type Quasar⁴ ($\sigma_{in} : 0.6$, $\sigma_{out} : 0.9$, angle d'ouverture : 30 degrés). Nous avons calculé les profils d'intensité de façon à prendre la moyenne de plusieurs profils à différentes positions du plan d'observation et pour un plan focal fixé à 35 nm à partir de la surface de la résine. Les simulations ont été réalisées avec la méthode FDTD uniquement en faisant varier les dimensions des lignes de 93 nm à 99 nm et la taille des SRAF de 20 nm à 34 nm. Suivant les deux cas cités ci-dessus, nous avons calculé deux seuils d'extraction que nous avons appliqués aux profils de l'intensité lumineuse du réseau de ligne de pas 360 nm afin de calculer les dimensions dans la résine. En plus des dimensions sur la plaquette, nous avons calculé le seuil d'impression des SRAF.

5.2.3.2.1 Cas d'une diffusion nulle dans la résine

Dans ce cas, nous n'avons introduit aucune diffusion dans la résine. Sur la figure 5.12, nous avons tracé les dimensions sur le masque en fonction de la taille des SRAF pour les deux cas de dose. Cette figure montre trois zones :

- La zone A correspond aux cas où les SRAF s'impriment dans la résine.
- La zone B correspond aux cas où les dimensions dans la résine sont en dehors des spécifications.
- La zone C correspond aux cas contraires qui respectent les spécifications des dimensions au niveau du masque.

Le domaine le plus intéressant est donc celui de la zone C qui respecte toutes les spécifications ainsi que la non impression des SRAF dans la résine. Notons ici que, pour le cas nominal en dose, les dimensions des SRAF doivent être inférieures à 26.25 nm afin d'éviter leur impression, alors qu'elles doivent être inférieures à 24.5 nm dans le plus mauvais cas en dose.

4. voir définition page 10

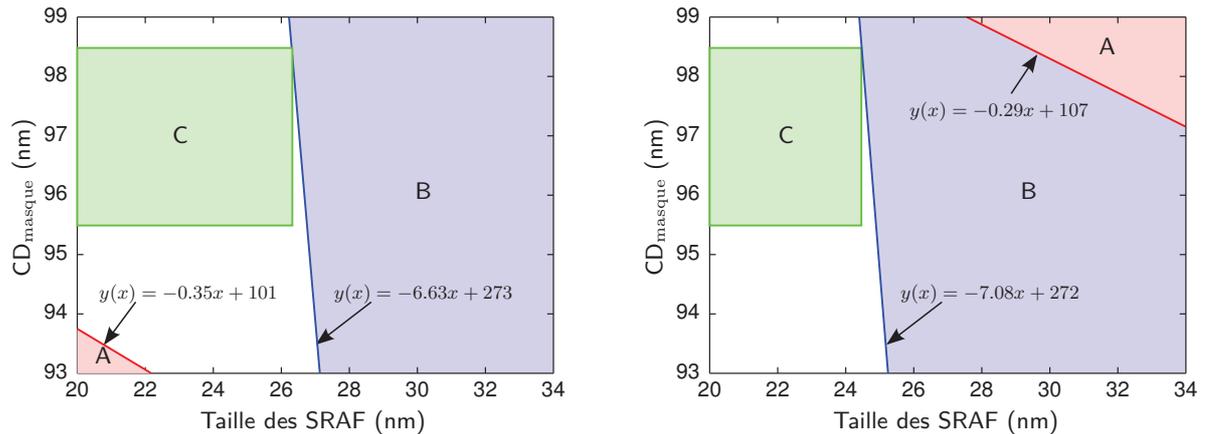


Figure 5.12 – CD masque (à l'échelle de la plaquette) en fonction de la taille des SRAF (à l'échelle de la plaquette) pour le cas nominal (à gauche) et le plus mauvais cas en dose (à droite). Les simulations ont été effectuées sans diffusion dans la résine. La zone A désigne les cas où les dimensions sur la plaquette sont en dehors des spécifications. La zone B désigne les cas où les SRAF s'impriment dans la résine. Finalement, la zone C correspond aux cas contraires et où les dimensions sur le masque respectent les spécifications.

5.2.3.2.2 Cas d'une diffusion non nulle dans la résine

Afin de mieux rapprocher nos résultats de simulation à la réalité, nous avons introduit une diffusion dans la résine avec des longueurs de diffusion égales à 15 nm, 20 nm et 25 nm. Nous avons donc convolué les profils de l'intensité lumineuse avec des fonctions gaussiennes d'écart types correspondant aux trois longueurs de diffusion. La figure 5.13 correspond aux résultats trouvés pour une longueur de diffusion égale à 15 nm. D'après cette figure, les dimensions des SRAF doivent être inférieures à 29.25 nm dans le cas nominal et à 27 nm dans le pire cas en dose. Contrairement au cas précédent, la diffusion permet de relâcher les contraintes au niveau des dimensions des SRAF et les intervalles de spécifications sont plus grands.

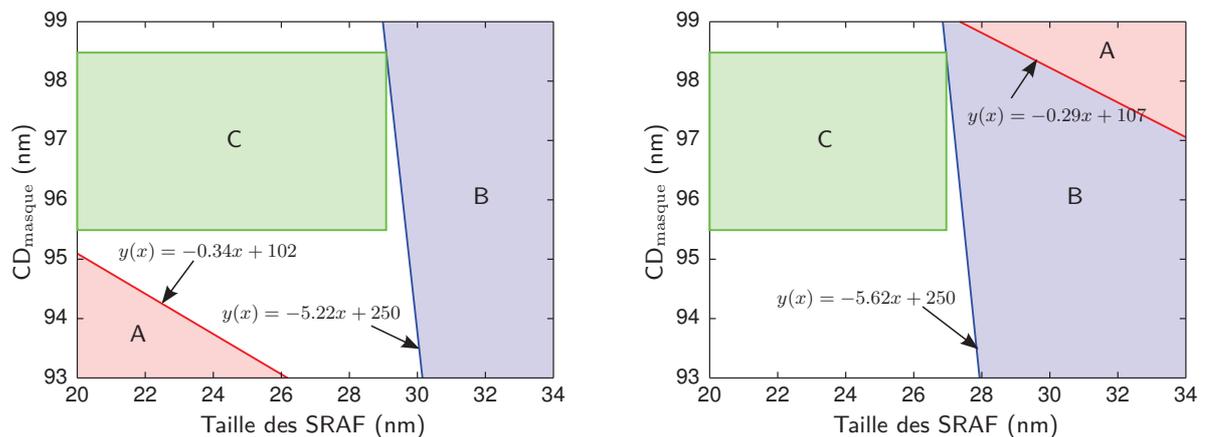


Figure 5.13 – CD masque (à l'échelle de la plaquette) en fonction de la taille des SRAF (à l'échelle de la plaquette) pour le cas nominal (à gauche) et le plus mauvais cas en dose (à droite). Les simulations ont été effectuées avec une diffusion de 15 nm dans la résine. La zone A désigne les cas où les dimensions sur la plaquette sont en dehors des spécifications. La zone B désigne les cas où les SRAF s'impriment dans la résine. Finalement, la zone C correspond aux cas contraires et où les dimensions sur le masque respectent les spécifications.

Le tableau 5.3 résume les résultats obtenus pour les trois cas de la longueur de diffusion. Si nous comparons les valeurs des simulations trouvées dans ce tableau avec les résultats expérimentaux, nous pouvons conclure que le cas le plus proche à la réalité correspond à une diffusion comprise entre 15 nm et 20 nm. En effet, les traces de résine observées sur la figure 5.14 cor-

respondent à la limite des spécifications du cas nominal en dose avec une longueur de diffusion égale à 15 nm. Elles peuvent également correspondre à la limite des spécifications du pire cas en dose avec une longueur de diffusion égale à 20 nm.

	Cas nominal	Pire cas
Sans diffusion	20–26.25	20–24.5 nm
Diffusion à 15 nm	20–29.25	20–27 nm
Diffusion à 20 nm	23–31.25	20–29.25 nm
Diffusion à 25 nm	27–34	20–31.75 nm

Tableau 5.3 – Spécifications par rapport aux dimensions (en nm) des SRAF pour le pas de réseau égal à 360 nm. Ces spécifications respectent les variations maximales autorisées au niveau du masque comme au niveau de la plaquette et garantissent à ce que les SRAF ne s'impriment pas dans la résine.

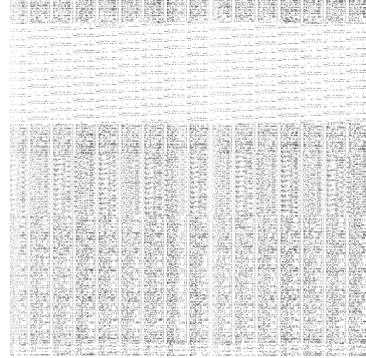


Figure 5.14 – Image au SEMCD d'une ligne d'un réseau de pas égal à 360 nm. La ligne mesure 96 nm sur le masque (à l'échelle de la plaquette) et 67.2 nm dans la résine. Les SRAF qui commencent à s'imprimer dans la résine mesurent 29.5 nm sur le masque (à l'échelle de la plaquette).

5.2.3.3 Correction par modification de la transmission et de la phase des SRAF

Nous avons simulé avec la méthode FDTD l'intensité lumineuse d'un réseau de lignes de pas égal à 360 nm et de dimensions égales à 65 nm avec des SRAF de 25 nm sur le masque. Nous avons ensuite calculé le budget en seuil d'impression des SRAF normalisé⁵ par rapport au seuil d'extraction (donné par le pas de réseau 180 nm avec des lignes de dimensions égales à 65 nm et 55 nm sur le masque et sur la plaquette respectivement). Nous avons effectué les mêmes simulations et calculs en utilisant l'approximation de Kirchhoff et en faisant varier la taille, la transmission et la phase des SRAF. Le tableau 5.4 récapitule les intervalles de variation de chacun des paramètres.

Paramètre	Intervalle de variations
Taille (nm)	de 20 à 40
Transmission (normalisée)	de 0 à 0.5
Phase (degrés)	de 140 à 220

Tableau 5.4 – Intervalles de variations de la taille, la transmission et la phase des SRAF pour les simulations avec l'approximation de Kirchhoff.

En fixant la phase à 180 degrés, la figure 5.15 présente la différence (en valeur absolue) des budgets normalisés calculés avec l'approximation de Kirchhoff et la méthode FDTD en fonction de la taille et de la transmission des SRAF. Nous pouvons clairement remarquer qu'il existe des couples transmission et taille des SRAF où les budgets sont égaux (c'est-à-dire que la prédiction de l'impression des SRAF par l'approximation de Kirchhoff est pratiquement similaire à celle de la méthode FDTD). Pour une taille fixée des SRAF, nous pouvons donc obtenir une valeur de transmission optimale réduisant les différences de budget. La figure 5.15 montre que la transmission optimale est égale à 0.12 pour une SRAF de 35 nm de taille et à 0.27 pour une SRAF de 30 nm.

5. Le budget normalisé en seuil d'impression est défini par

$$\frac{I_{\text{minSRAF}} - \text{seuil}}{\text{seuil}}$$

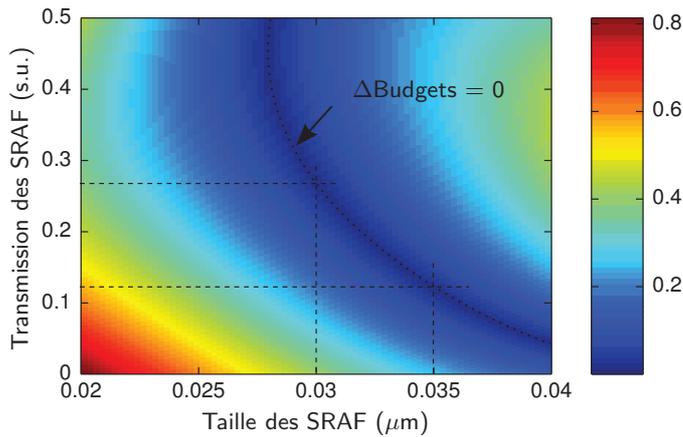


Figure 5.15 – $|\Delta Budgets|$ (Kirchhoff-FDTD) normalisés en dose d'impression des SRAF en fonction de leurs tailles et de leurs transmissions (la phase étant fixée à 180 degrés).

Concernant la phase des SRAF, la figure 5.16 montre que ce cas est moins intéressant que celui de la transmission. En effet, le minimum de l'écart des budgets ne varie pas beaucoup en fonction de la taille des SRAF pour différentes valeurs de la phase. L'exemple tracé sur cette figure montre que pour, une transmission fixée à 0.12, la taille des SRAF permettant de minimiser l'écart des budgets est située entre 35.5 nm et 38 nm pour des variations de la phase de ± 40 degrés.

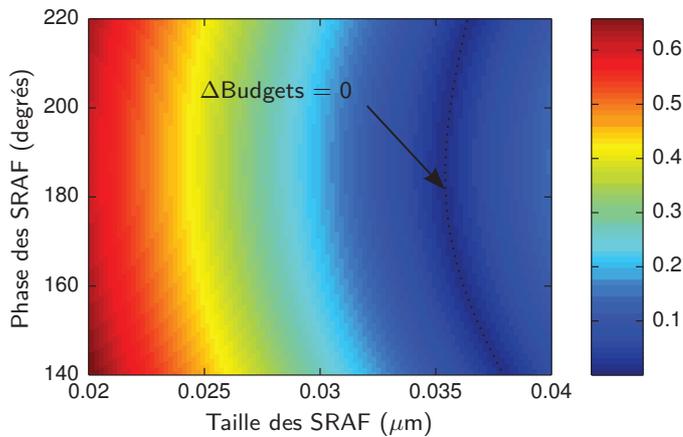


Figure 5.16 – $|\Delta Budgets|$ (Kirchhoff-FDTD) normalisés en dose d'impression des SRAF en fonction de leurs tailles et de leurs phases (la transmission étant fixée à 0.12).

5.3 Méthode de modélisation par décomposition multi-niveaux du masque fin

Le problème des méthodes optiques approchées DDM et M3D que nous avons présentées précédemment se situe lorsque l'on veut obtenir :

- Une précision de calcul de l'intensité lumineuse satisfaisante afin de réduire les erreurs liées aux effets topographiques du masque.
- Des temps de simulation en traitement OPC et en vérification post-OPC équivalents à ceux que l'on peut obtenir avec l'approximation de Kirchhoff.

Cependant, malgré tous les efforts effectués, il nous est impossible d'arriver à des temps d'exécution équivalents à ceux de l'approximation de Kirchhoff. Car, même si l'on utilise des tables TCC réduites à 20 noyaux, les temps d'exécution sont multipliés par un facteur allant de deux à trois fois selon la complexité du dessin du masque traité. Dans l'objectif de pallier ce problème, nous avons développé une troisième méthode approchée permettant, à la fois, de profiter

de la rapidité des algorithmes de calcul, déjà existants, de l'approximation de Kirchhoff et de permettre de corriger une partie des erreurs liées aux effets topographiques du masque.

Dans la suite, nous commencerons par présenter le principe de base de cette méthode. Puis nous effectuerons une évaluation afin de dégager les principaux avantages et inconvénients et nous comparerons les résultats de simulation avec ceux que nous avons obtenus pour les méthodes DDM et M3D.

5.3.1 Principe

Le principe de cette technique est inspiré des méthodes DDM et M3D et du modèle de couche limite BL⁶ ("Boundary Layer") de Tirapu et al. [121] qui consiste à rajouter des bords aux motifs du masque avec différentes valeurs de transmission et de phase. Notre méthode de modélisation par décomposition multi-niveaux du masque permet d'approcher la fonction de transmission d'un bord diffractant par une fonction en escalier et de préserver l'énergie du signal. Ainsi, le masque fin est redéfini par la juxtaposition de plusieurs niveaux – considérés comme "masques fins" – définis chacun par une fonction unique de transmission et de phase.

5.3.1.1 Expression du champ scalaire effectif

Comme nous l'avons décrit dans le premier chapitre de ce mémoire, une onde électromagnétique, polarisée ou non, peut être décomposée en deux composantes TE et TM de polarisation rectiligne, chacune d'elles étant perpendiculaire à l'autre. Le champ électrique réel est donc simplement décrit par la somme de ses deux composantes vectorielles TE et TM, contrairement au cas de l'approximation de Kirchhoff où il est uniquement présenté par une fonction scalaire. La transformation de l'ensemble formé par les composantes TE et TM du champ par une unique composante scalaire nécessite l'expression d'une unique fonction d'amplitude et de phase, que nous appelons champ électrique effectif.

Le calcul du champ électrique effectif est simple. Afin de calculer l'amplitude du champ effectif, nous allons utiliser l'expression de l'intensité totale qui s'écrit :

$$\begin{aligned} I &= \alpha I_{TE} + (1 - \alpha) I_{TM} \\ &= \alpha \vec{E}_{TE} \cdot \vec{E}_{TE}^* + (1 - \alpha) \vec{E}_{TM} \cdot \vec{E}_{TM}^* \end{aligned} \quad (5.3)$$

$$= \vec{E}_{eff} \cdot \vec{E}_{eff}^* \quad (5.4)$$

α étant un facteur compris entre 0 et 1 qui désigne le pourcentage de lumière polarisée en TE. Notons que pour une lumière non polarisée, le facteur α vaut 1/2. En exprimant les composantes TE et TM du champ électrique réel et la composante scalaire du champ électrique effectif par :

$$E_{TE} = A_{TE} e^{i\varphi_{TE}}, \quad E_{TM} = A_{TM} e^{i\varphi_{TM}} \quad \text{et} \quad E_{eff} = A_{eff} e^{i\varphi_{eff}},$$

et en remplaçant ces expressions dans les equations 5.3 et 5.4, nous obtenons :

$$A_{eff} = \sqrt{\alpha A_{TE}^2 + (1 - \alpha) A_{TM}^2}. \quad (5.5)$$

Cette dernière expression n'est vraie que si l'incidence est normale, ce qui est bien notre cas. En ce qui concerne la phase, nous allons commencer par exprimer les composantes vectorielles TE et TM dans un repère cartésien :

$$\vec{E}_{TE} = A_{TE} e^{i\varphi_{TE}} \begin{pmatrix} -\sin \varphi \\ \cos \varphi \\ 0 \end{pmatrix}, \quad \vec{E}_{TM} = A_{TM} e^{i\varphi_{TM}} \begin{pmatrix} \cos \theta \cos \varphi \\ \cos \theta \sin \varphi \\ -\sin \theta \end{pmatrix},$$

θ étant l'angle que fait le vecteur d'onde \vec{k} avec l'axe optique z et φ celui que fait le plan d'incidence (porté par \vec{z} et \vec{k}) avec l'axe des x . Ainsi, en tenant compte de l'incidence normale

6. voir les méthodes heuristiques en page 111.

de l'onde ($\theta = 0$) et l'égalité suivante exprimant le champ effectif en fonction des composantes vectorielles TE et TM :

$$\vec{E}_{eff} = \sqrt{\alpha} \vec{E}_{TE} + \sqrt{1-\alpha} \vec{E}_{TM} \quad \text{avec} \quad \vec{E}_{eff} = A_{eff} e^{i\varphi_{eff}} \begin{pmatrix} \sqrt{1-\alpha} \\ \sqrt{\alpha} \\ 0 \end{pmatrix},$$

nous arrivons au système d'équations suivant :

$$\sqrt{1-\alpha} A_{eff} e^{i\varphi_{eff}} = -\sqrt{\alpha} A_{TE} e^{i\varphi_{TE}} \sin \varphi + \sqrt{1-\alpha} A_{TM} e^{i\varphi_{TM}} \cos \varphi \quad (5.6)$$

$$\sqrt{\alpha} A_{eff} e^{i\varphi_{eff}} = \sqrt{\alpha} A_{TE} e^{i\varphi_{TE}} \cos \varphi + \sqrt{1-\alpha} A_{TM} e^{i\varphi_{TM}} \sin \varphi. \quad (5.7)$$

En élevant au carré ces deux dernières expressions et en les sommant ensemble, nous arrivons à l'égalité suivante :

$$A_{eff}^2 e^{i2\varphi_{eff}} = \alpha A_{TE}^2 e^{i2\varphi_{TE}} + (1-\alpha) A_{TM}^2 e^{i2\varphi_{TM}}.$$

Or, dans les zones claires du masque, c'est-à-dire en l'absence de motifs, les amplitudes des composantes TE et TM sont égales à celle d'un champ scalaire de Kirchhoff. L'égalité suivante reste donc valable :

$$e^{i2\varphi_{eff}} = \alpha e^{i2\varphi_{TE}} + (1-\alpha) e^{i2\varphi_{TM}}.$$

L'expression finale de la phase effective est donc la suivante :

$$\varphi_{eff} = \arg(\pm \sqrt{\alpha e^{2i\varphi_{TE}} + (1-\alpha) e^{2i\varphi_{TM}}}) \quad (5.8)$$

Afin d'enlever l'ambiguïté sur le signe devant la racine carrée, nous travaillerons plutôt avec $\Delta\varphi_{eff}$, $\Delta\varphi_{TE}$ et $\Delta\varphi_{TM}$ qui désignent les différences de phase avec le champ idéal de Kirchhoff :

$$\varphi_{eff} = \arg(\sqrt{\alpha e^{2i\Delta\varphi_{TE}} + (1-\alpha) e^{2i\Delta\varphi_{TM}}}) + \varphi_{idéal} \quad (5.9)$$

Ainsi le champ scalaire effectif s'écrit :

$$E_{eff} = \sqrt{\alpha A_{TE}^2 + (1-\alpha) A_{TM}^2} e^{i \arg(\pm \sqrt{\alpha e^{2i\varphi_{TE}} + (1-\alpha) e^{2i\varphi_{TM}}})}.$$

5.3.1.2 Décomposition multi-niveaux du champ électrique

Les maxima et minima locaux du champ électrique d'un bord diffractant sont les lieux des interférences constructives et destructives de l'onde diffractée. Les pseudo-oscillations observées au niveau de l'amplitude et de la phase du champ électrique sont le résultat d'une alternance de contributions positives et négatives des différentes zones de Fresnel. En effet, considérons de nouveau le schéma de la figure 4.7 et prenons deux points P et Q qui représentent deux extréma locaux consécutifs du champ. Ils sont donc séparés d'une distance avoisinant la moitié de la longueur d'onde λ (voir démonstration page 84). Comme il s'agit d'un maximum et d'un minimum d'intensité, l'onde reçue en ces deux points provient de deux zones de Fresnel de phases opposées. D'après l'équation 4.7, nous avons l'égalité suivante :

$$b(x_Q) = b(x_P) + \frac{\lambda}{2}.$$

L'expression de l'onde reçue au point Q s'écrit :

$$\begin{aligned} U(Q) &= j\lambda A \frac{e^{jk(r+b(x_Q))}}{r+b(x_Q)} K_1 \\ &= - \left(j\lambda A \frac{e^{jk(r+b(x_P))}}{r+b(x_P)} K_1 \right) \left(\frac{r+b(x_P)}{r+b(x_Q)} \right) \\ &= -U(P) \left(\frac{r+b(x_P)}{r+b(x_Q)} \right). \end{aligned}$$

Ainsi, nous avons montré que les amplitudes de l'onde reçue en deux extréma consécutifs du champ proviennent des contributions de deux zones de Fresnel successives de signes opposés. C'est donc sur la base de cette alternance de contributions positives et négatives que nous avons effectué la décomposition multi-niveaux du champ électrique. Afin d'isoler les différentes contributions, nous nous sommes basés sur les points d'inflexion de l'amplitude du champ électrique. En effet, les points d'inflexion représentent les lieux où les contributions de deux zones de Fresnel s'annulent mutuellement.

Pour calculer l'amplitude et la phase des différents niveaux, nous nous sommes basés sur l'égalité de Parseval. Cette égalité (parfois appelée également Théorème de Parseval ou Identité de Rayleigh) stipule que l'énergie totale d'un signal est égale à la somme des contributions de ses harmoniques. Il s'agit donc de la conservation de l'énergie entre les espaces réel et fréquentiel, qui s'exprime ainsi :

$$E(f) = \int |f(x)|^2 dx = \frac{1}{2\pi} \int |TF\{f\}(X)|^2 dX$$

avec f une fonction quelconque et $TF\{f\}$ sa transformée de Fourier. Ainsi, si l'on considère les deux fonctions $g(x)$ et $r(x)$ (voir figure 5.17) définies par :

$$g(x) = \begin{cases} f(x) & x_1 \leq x \leq x_2 \\ 0 & \text{ailleurs} \end{cases} \quad \text{et} \quad r(x) = \begin{cases} 1 & x_1 \leq x \leq x_2 \\ 0 & \text{ailleurs} \end{cases},$$

le problème se résume à trouver la constante A permettant de conserver l'énergie entre $g(x)$ et la fonction définie par $A \times r(x)$, c'est-à-dire :

$$\int |TF\{g\}(X)|^2 dX = \int |A TF\{r\}(X)|^2 dX.$$

En première approximation, nous ne considérons que le point central du spectre de fréquence nulle. Nous avons donc :

$$\begin{aligned} A &\approx \left| \frac{TF\{g\}(X)}{TF\{r\}(X)} \right|_{X=0} \\ &= \frac{\int g(x) dx}{d} \end{aligned} \quad (5.10)$$

avec d la distance séparant les points d'abscisses x_1 et x_2 . La constante A n'est donc rien d'autre que la hauteur du rectangle de largeur d et de surface égale à l'aire située sous la courbe $g(x)$, comme le montre la figure 5.17 ci-dessous.

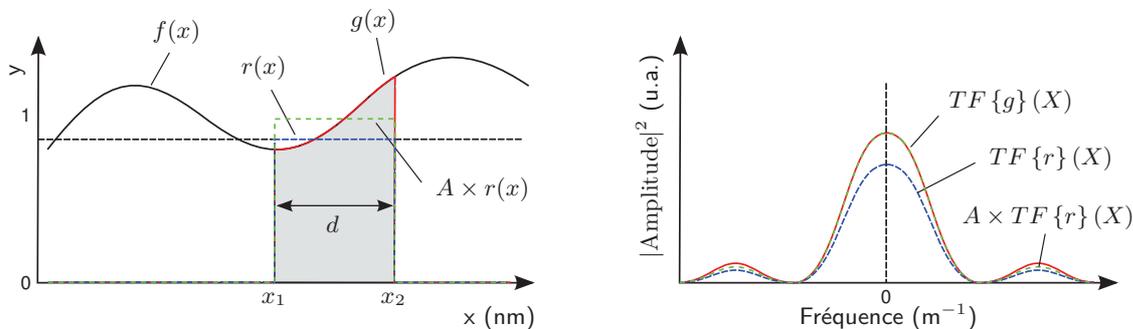


Figure 5.17 – Principe du calcul des niveaux d'énergie équivalente par le théorème de Parseval de la conservation de l'énergie : définition des fonctions $g(x)$ et $r(x)$ (à gauche) et tracés des spectres de Fourier des fonctions $g(x)$, $r(x)$ et $A \times r(x)$ (à droite). La constante A est telle que définie par l'équation 5.10.

La figure 5.18 présente les courbes de l'amplitude et de la phase du champ électrique effectif ainsi que les courbes en escalier correspondant à la décomposition multi-niveaux dans le cas

d'un bord diffractant isolé. Pendant cette décomposition, nous avons rencontré deux problèmes. Le premier concerne la phase du champ électrique qui oscille en quadrature par rapport à l'amplitude. Ainsi, les points d'inflexion obtenus au niveau de la phase représentent des extréma d'amplitude et réciproquement. Par ailleurs, une décomposition multi-niveaux de la phase entraîne un résultat différent que celui obtenu pour l'amplitude. Nous nous sommes donc contenté d'un seul découpage en prenant les points d'inflexion de l'amplitude et nous avons appliqué ce même découpage à la phase. Le deuxième problème concerne les deux parties des courbes de l'amplitude et de la phase localisées entre les deux extréma situés de part et d'autre du bord (abscisse x égale à zéro). Vu l'importance de ces deux parties au niveau de la diffraction, car elles constituent l'étape de transition entre champs clair et obscur, nous avons considéré un découpage plus fin de cette partie. Pour ce faire, nous avons inséré deux points à mi-distance entre l'unique point d'inflexion et les deux extréma de cette partie de la courbe. Les deux points insérés seront désormais considérés comme lieux de découpage de la même façon que les autres points d'inflexion.

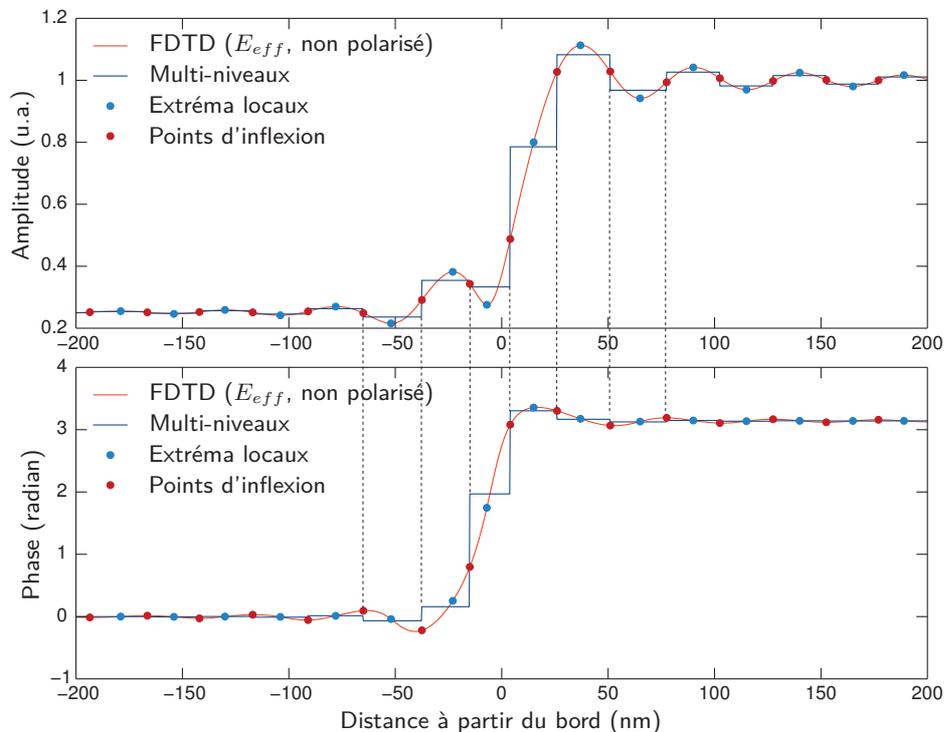


Figure 5.18 – Principe de la décomposition multi-niveaux de l'amplitude et de la phase du champ électrique effectif dans le cas d'un bord diffractant isolé.

Nous avons tracé sur les 5.19 et 5.20) les spectres de Fourier du champ électrique effectif dans les trois cas d'un calcul rigoureux, d'une décomposition multi-niveaux et d'un champ scalaire de Kirchhoff. D'après ces deux figures, les écarts en amplitude par rapport au calcul rigoureux ont été visiblement réduits avec la méthode des multi-niveaux. Nous arrivons à corriger l'amplitude de l'ordre zéro et réduire les erreurs des autres ordres pouvant être captés par la pupille du système optique. Grâce à cette décomposition multi-niveaux, nous allons pouvoir gagner en précision quant au calcul de l'intensité lumineuse dans la résine.

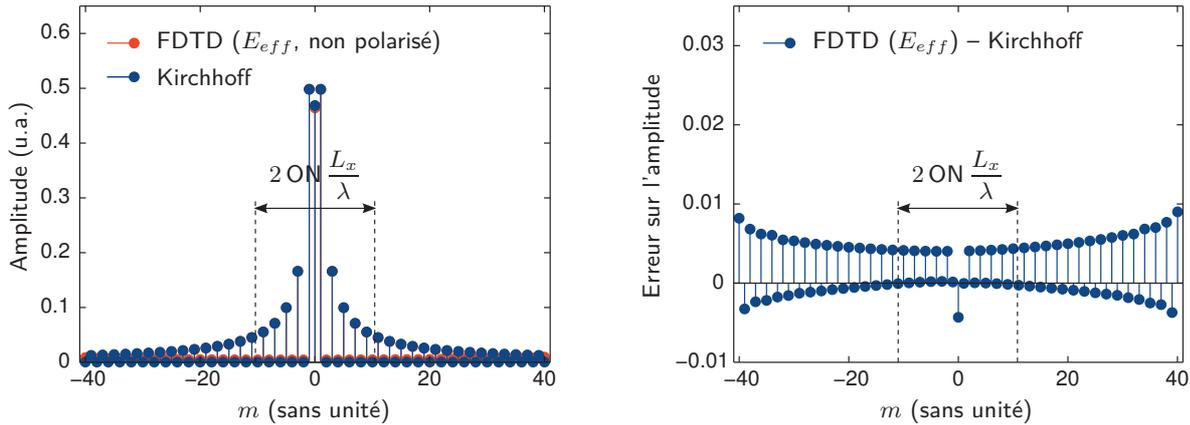


Figure 5.19 – Spectres de Fourier du champ électrique d'un bord diffractant isolé calculés avec l'approximation de Kirchhoff et la méthode FDTD (champ électrique effectif) (à gauche) et différences entre les ordres (à droite). m étant l'indice de Bragg défini par $\frac{k_x L_x}{2\pi}$, k_x et L_x étant la composante du vecteur d'onde et la dimension du masque selon l'axe x respectivement.

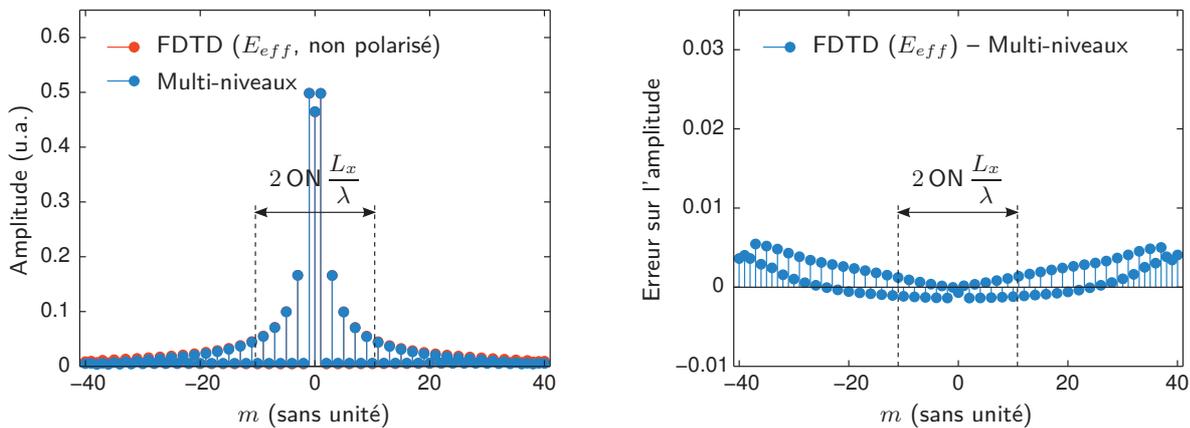


Figure 5.20 – Spectres de Fourier du champ électrique d'un bord diffractant isolé calculés avec la méthode multi-niveaux et la méthode FDTD (champ électrique effectif) (à gauche) et différences entre les ordres (à droite). m étant l'indice de Bragg défini par $\frac{k_x L_x}{2\pi}$, k_x et L_x étant la composante du vecteur d'onde et la dimension du masque selon l'axe x respectivement.

5.3.2 Évaluation par rapport au calcul rigoureux

5.3.2.1 Définition des conditions de simulation

Pour des besoins de comparaison, nous avons gardé les mêmes conditions de simulation que dans les cas des méthodes approchées DDM et M3D. Dans cette évaluation, nous avons utilisé des tables optiques TCC tronquées à 20 noyaux de convolution. Comme nous l'avons déjà vu, la méthode de décomposition multi-niveaux n'est autre que l'approximation de Kirchhoff que nous appliquons à plusieurs morceaux de "masques fins" ayant chacun une fonction scalaire de transmission et de phase. De toute évidence, afin d'obtenir de meilleurs résultats, il faut calculer les valeurs de transmission et de phase en partant du champ électrique simulé par la méthode FDTD pour chacun des réseaux de lignes. Dans ce cas, les niveaux que nous obtiendrons peuvent être différents selon le motif ou le réseau traité. Ceci nous conduit à créer un nombre élevé de calques⁷ dans le logiciel CalibreTM pour traiter un circuit quelconque. Ce qui est impossible à réaliser puisque l'on est limité par un certain nombre de calques visibles.

7. Dans le logiciel CalibreTM, le calque (de l'anglais "Layer") est un support imaginaire qui contient les éléments d'un circuit. Le calque est défini par une polarité (claire ou obscure) et une fonction unique de transmission et de

Dans l'objectif de résoudre ce problème, nous avons utilisé un nombre limité (inférieur ou égal à 10) de niveaux préalablement calculés à partir du champ électrique effectif d'un bord diffractant isolé. De tels niveaux sont partagés à nombre égal de part et d'autre du bord d'une structure.

La reconstitution de champ électrique d'un motif quelconque s'effectue de façon à symétriser les champs provenant des bords voisins opposés. Le tableau 5.5 présente un récapitulatif des différentes étapes suivies dans cette évaluation pour la décomposition multi-niveaux où l'on présente la manière de reconstituer le champ électrique dans les cas d'une ligne et d'un espace.

La figure 5.21 illustre deux exemples d'application de cette approche réalisés sous Matlab. Le premier cas est celui d'une ligne isolée de dimension 70 nm protégée par des SRAF de dimensions 30 nm. Le second est celui d'une tranchée isolée de même dimension avec des SRAF sous forme de tranchées de 30 nm. Comme le montre cette figure, le dessin initial du circuit (utilisé avec l'approximation de Kirchhoff) a été décomposé en un nombre de calques qui, définis, chacun, par une valeur particulière de transmission et de phase, permettent d'approcher le champ électrique rigoureux.

Description	Représentation schématique	
a) Calcul du champ électrique effectif d'un bord diffractant isolé.		
b) Décomposition multi-niveaux du champ électrique effectif obtenu.		
c) Troncature du champ électrique effectif afin de ne laisser qu'un nombre limité de niveaux.		
	Cas d'une ligne	Cas d'un espace
d) Application de la décomposition multi-niveaux aux bords des motifs.		
e) Élimination des niveaux qui se superposent et symétrisation du champ entre bords adjacents.		

Tableau 5.5 – Récapitulatif des étapes suivies pour la décomposition multi-niveaux des motifs du masque effectuée dans cette évaluation.

Comme nous pouvons bien le remarquer sur la figure 5.22, les premiers résultats de simulation de l'intensité lumineuse dans la résine par cette approche sont encourageants. Cette figure montre également que la méthode de décomposition multi-niveaux du masque, avec des tables TCC tronquées à 20 noyaux, permet d'atteindre une précision de calcul équivalente à celle de la méthode DDM avec des tables TCC non tronquées. En effet, pour un seuil fixé à 0.175, les valeurs mesurées des dimensions critiques CD sont de 67.5 nm, 70.5 nm et 70 nm pour l'approximation de Kirchhoff (20 noyaux), la méthode DDM (Hopkins) et la décomposition multi-niveaux (20 noyaux) respectivement. Comme nous pouvons également l'observer, les profils de l'intensité lumineuse au niveaux des SRAF sont pratiquement identiques à ceux de la méthode DDM, ce qui permet d'améliorer la prédiction de l'impression des SRAF. Nous nous intéressons donc à

phase et permet ainsi de représenter un niveau quelconque. La superposition des calques dits visibles permet de définir la fonction de transmission et de phase du masque.

évaluer la méthode de décomposition multi-niveaux par rapport à la méthode rigoureuse FDTD, et ce au niveau de la précision en CD, la position du meilleur focus, la profondeur de champ et la prédiction de l'impression des SRAF. Nous nous intéressons également à une comparaison des différentes méthodes au niveau du temps de cycle et des ressources nécessaires pour la simulation optique.

Cette approche, que nous allons évaluer dans la suite de ce chapitre, permet de traiter un masque complet. Cependant, elle présente deux inconvénients majeurs. Le premier consiste à utiliser le champ électrique provenant d'un bord isolé et l'appliquer pour construire le champ proche d'un motif quelconque. Comme nous l'avons vu dans le cas des méthodes approchées DDM et M3D, ceci constitue une source d'erreurs pour le calcul du champ proche dans le cas des réseaux de petits pas ou des motifs où l'espace entre deux lignes voisines est réduit. Le deuxième inconvénient vient de la symétrisation du champ provenant des bords. Cette manière de reconstituer le champ entre bords opposés voisins considère qu'il n'existe aucune interaction entre eux. En conséquence, cette approche ne permet pas de gérer les effets d'interaction optique.

Une meilleure approche serait donc de calculer les niveaux à partir du champ électrique rigoureux que l'on obtient par une simulation FDTD directe d'un motif donné. Par contre, l'inconvénient de celle-ci est son caractère spécifique et qu'elle est restreinte aux motifs qui présentent les mêmes caractéristiques géométriques du motif de départ. Néanmoins, elle peut être utilisée pour une correction ciblée.

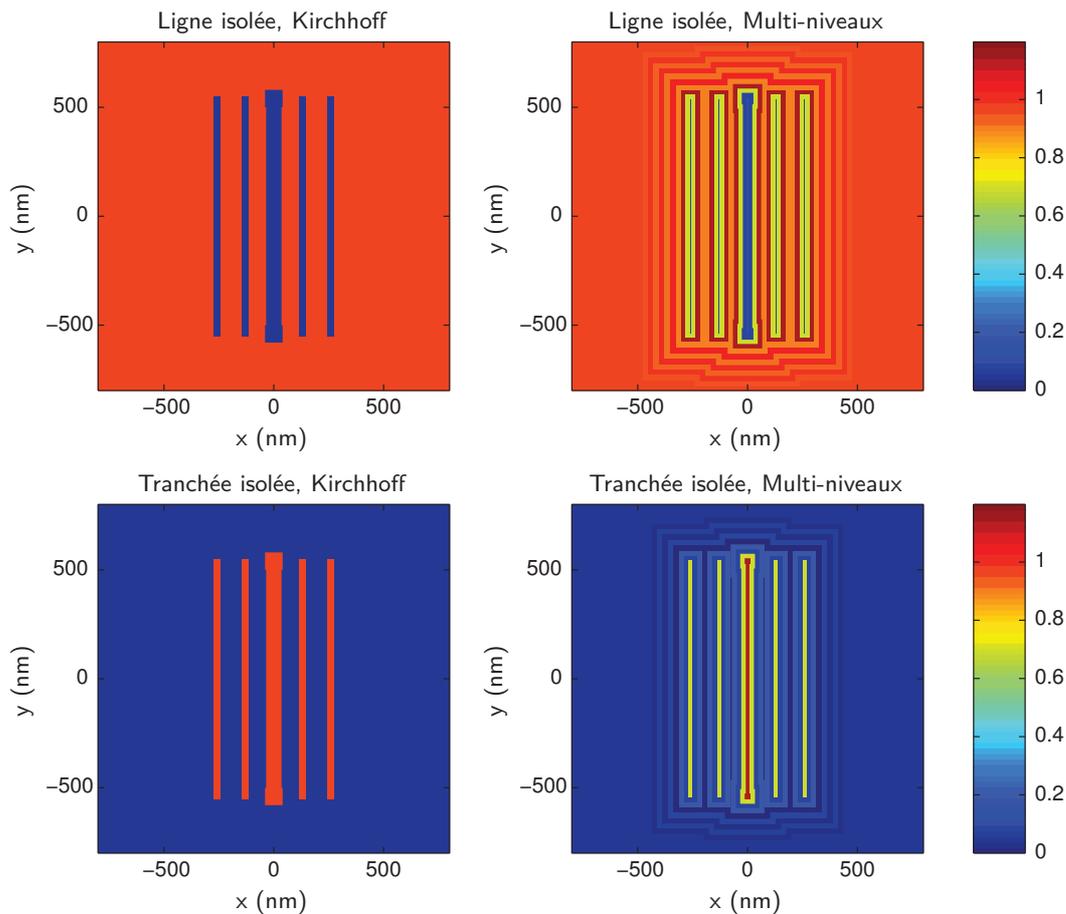


Figure 5.21 – Tracés des amplitudes du champ électrique calculées par l'approximation de Kirchhoff (à gauche) et la méthode de décomposition multi-niveaux (à droite) pour deux cas de motifs : une ligne isolée protégée par des SRAF (en haut) et une tranchée isolée protégée par des SRAF (en bas).

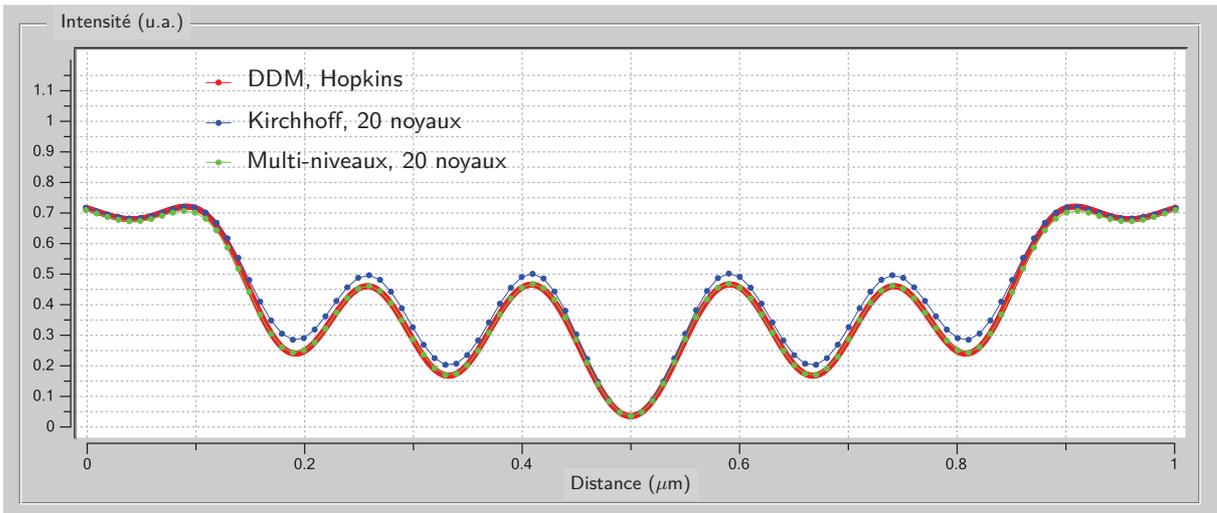


Figure 5.22 – Profils de l'intensité lumineuse dans la résine d'une ligne isolée (CD 97 nm) protégée par des SRAF (taille 34 nm) calculés par l'approximation de Kirchhoff (20 noyaux), la méthode DDM (Hopkins) et la méthode Multi-niveaux (20 noyaux), sous Calibre™ (Mentor Graphics). Le masque utilisé est un PSM à 6% de transmission et à décalage de phase égal à 180 degrés, l'illumination est non polarisée de type Quasar^s (σ_{in} : 0.6, σ_{out} : 0.9, angle d'ouverture : 30 degrés) et l'ouverture numérique ON est égale à 1.05.

5.3.2.2 Fenêtres de procédé et courbes de Bossung

Nous avons calculé avec la méthode de décomposition multi-niveaux les fenêtres de procédé et les courbes de Bossung pour un réseau de lignes de pas égal à 140 nm. D'après la figure 5.23, nous constatons que la méthode des multi-niveaux permet de recentrer en dose les fenêtres de procédés au niveau de celles obtenues avec la méthode FDTD. Par contre, il existe un léger décalage en focus qui peut entraîner des erreurs focales. Les courbes de Bossung sont également légèrement décentrées en focus. On arrive pourtant à obtenir des résultats très similaires à ceux obtenus avec les méthodes semi-rigoureuses DDM et M3D (voir chapitre précédent). Le décalage en dose observé au niveau des courbes de Bossung est dû au fait que l'on ait utilisé le champ électrique provenant d'un bord isolé et que l'on ait négligé les interactions bord à bord dans les réseaux de faibles pas. Ce qui est également le cas des méthodes DDM et M3D où l'on observe le même comportement pour ce pas de réseau.

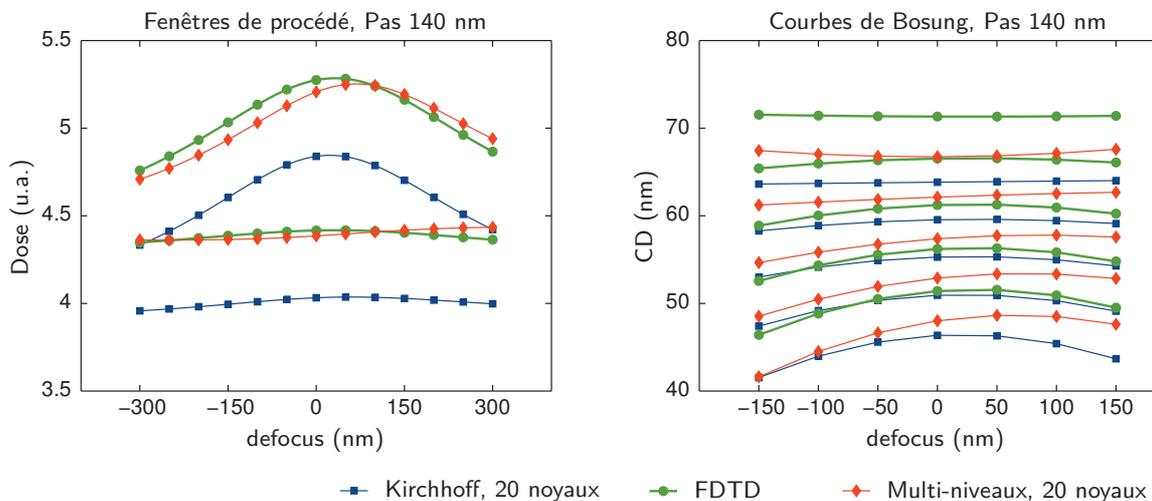


Figure 5.23 – Fenêtre de procédé (à gauche) et courbes de Bossung (à droite) calculées pour un réseau de lignes denses de pas égal à 140 nm par la méthode de décomposition multi-niveaux, l'approximation de Kirchhoff ainsi que la méthode rigoureuse FDTD.

5.3.2.3 Position du meilleur focus

Sur la figure 5.24, nous avons tracé les courbes du meilleur focus en fonction du pas de réseau calculées par la méthode des multi-niveaux ainsi que l'approximation de Kirchhoff et la méthode rigoureuse FDTD. D'après cette figure, nous pouvons clairement constater que les positions du meilleur focus extraites des fenêtres de procédé calculées par la méthode des multi-niveaux sont plus proches de celles obtenues avec la méthode rigoureuse FDTD. Ceci se traduit par le fait que la méthode des multi-niveaux permet de mieux recentrer en focus les fenêtres de procédé et les courbes de Bossung par rapport à la méthode FDTD. Néanmoins, les courbes obtenues possèdent les mêmes caractéristiques de celles des méthodes approchées DDM et M3D. À savoir, l'existence d'un décalage en focus pour les faibles pas de réseau (par exemple le pas 140 nm) et autour du pas 300 nm qui représente le début de l'insertion des SRAF.

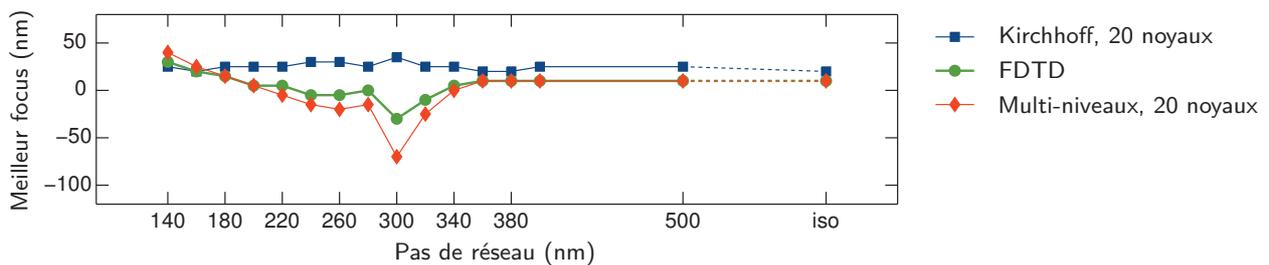


Figure 5.24 – Tracés du meilleur focus en fonction du pas de réseau prédits par la méthode de décomposition multi-niveaux et la méthode rigoureuse FDTD.

5.3.2.4 Profondeur de champ

Toujours dans le cadre de l'évaluation de la méthode des multi-niveaux, nous avons calculé la profondeur de champ en fonction du pas de réseau pour une latitude en dose égale à 8%. D'après la figure 5.25, la prédiction de la DOF avec la méthode des multi-niveaux est légèrement améliorée par rapport à l'approximation de Kirchhoff. Cependant, il n'existe pratiquement aucune amélioration pour les réseaux de lignes denses (à partir de 200 nm). Ceci traduit le comportement similaire avec l'approximation de Kirchhoff des fenêtres de procédé que l'on a déjà observé pour le pas de réseau 180 nm et constitue ainsi un inconvénient majeur de cette approche.

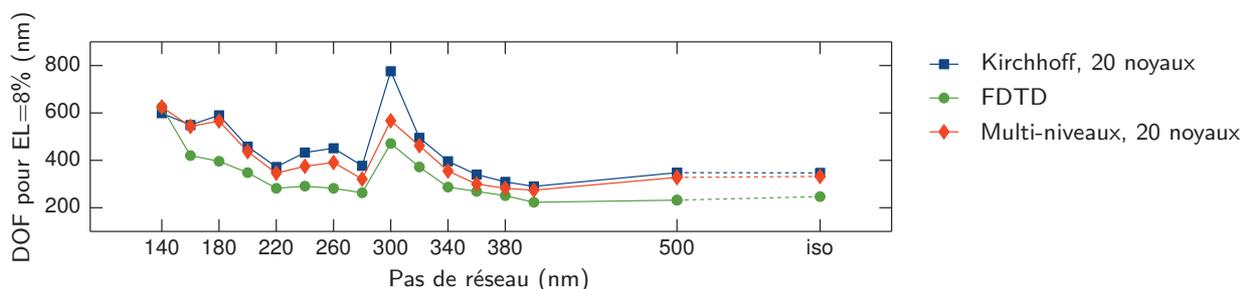


Figure 5.25 – Tracés de la profondeur de champ DOF pour une latitude en exposition fixée à 8% en fonction du pas de réseau prédits par la méthode de décomposition multi-niveaux et la méthode rigoureuse FDTD.

5.3.2.5 Effets de proximité

Nous avons tracé sur la figure 5.26 les valeurs de ΔCD (écarts avec la méthode FDTD) obtenues dans le cas de la méthode des multi-niveaux en fonction du pas de réseau et du décalage

en focus. D'après cette figure, la méthode des multi-niveaux permet de réduire les écarts en CD avec la méthode rigoureuse FDTD. Par ailleurs, comme dans les cas des méthodes DDM et M3D, l'erreur en dimensions critiques dépend du pas de réseau mais reste inférieure à 4 nm. La méthode des multi-niveaux permet ainsi d'améliorer la modélisation des effets topographiques du masque.

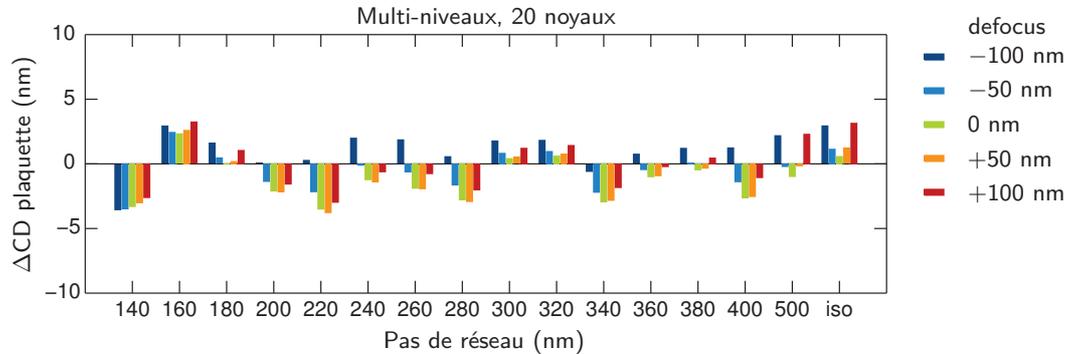


Figure 5.26 – Tracés des erreurs en CD calculées par la méthode de décomposition multi-niveaux (avec 20 noyaux) en fonction du pas de réseau pour différentes valeurs de la position du plan focal.

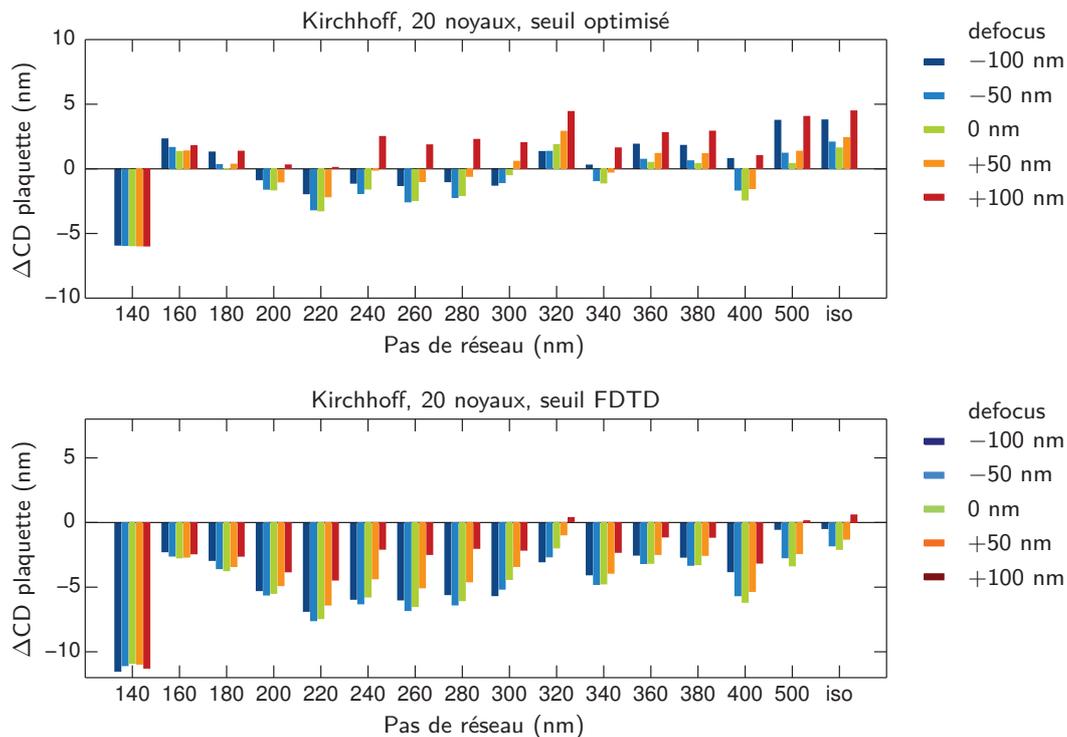


Figure 5.27 – Tracés des erreurs en CD calculées par l'approximation de Kirchhoff (avec 20 noyaux) en fonction du pas de réseau pour deux valeurs du seuil d'extraction : un seuil optimisé pour viser un CD égal à 65 nm pour le pas de réseau 180 nm et un second qui est celui de la méthode FDTD pour viser le même CD pour ce même pas de réseau.

Pour comparer les résultats de notre méthode des multi-niveaux avec ceux de l'approximation de Kirchhoff, nous avons tracé sur la figure 5.27 les valeurs de ΔCD obtenues pour deux seuils d'extraction. Le premier seuil est celui que l'on a optimisé avec l'approximation de Kirchhoff pour viser un CD égal à 65 nm pour une ligne appartenant à un réseau de pas 180 nm. Le second est celui obtenu avec la méthode FDTD pour viser le même CD avec ce même pas de réseau. D'après ces deux graphes, les écarts en CD liés aux effets topographiques du masque sont importants. De plus, l'optimisation du seuil ne permet pas d'absorber la plupart des erreurs focales que l'on peut facilement visualiser sur ces deux graphiques pour les pas de réseau situés

entre 200 et 340 nm⁹. La méthode des multi-niveaux, quant à elle, permet de réduire les erreurs liées aux effets de la topographie et de corriger une partie des erreurs focales (pour les pas de réseau compris entre 300 et 340 nm).

Finalement, nous avons retracé sur la figure 5.28 les erreurs en CD de la méthode A que nous avons évaluée dans le chapitre précédent. Comme nous pouvons bien le remarquer, notre méthode des multi-niveaux est capable d'égaliser la méthode A en terme de précision de calcul pour plusieurs pas de réseaux. Les différences que l'on peut visualiser pour les autres pas de réseau nécessitent un travail supplémentaire afin d'optimiser les paramètres de la méthode des multi-niveaux afin de compenser les erreurs de simulation dues au nombre limité de noyaux (20 noyaux contre 64 pour la méthode A) que l'on ait pris dans les tables optiques.

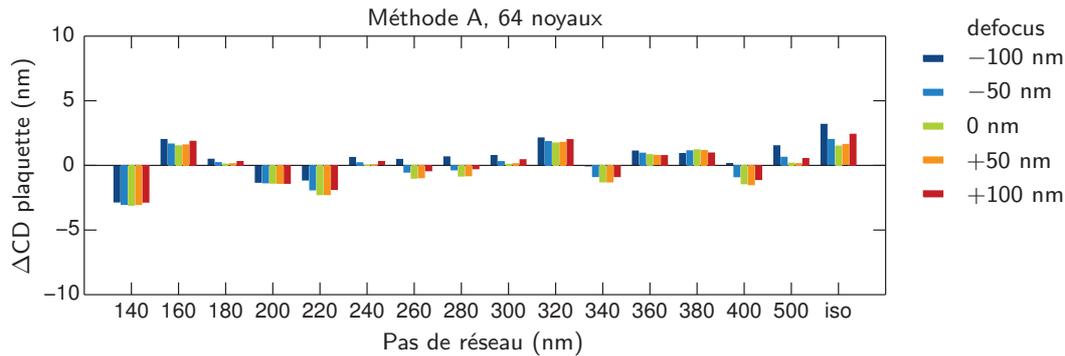


Figure 5.28 – Tracés des erreurs en CD calculées par la méthode A (avec 64 noyaux) en fonction du pas de réseau.

5.3.2.6 Prédiction de l'impression des SRAF

Évaluons maintenant la prédiction de l'impression des SRAF par la méthode des multi-niveaux. Pour ce faire, nous avons tracé sur la figure 5.29 le budget normalisé en seuil d'impression des SRAF en fonction du pas de réseau dans le cas d'une exposition à la dose nominale. D'après cette figure, la méthode des multi-niveaux permet d'améliorer la prédiction de l'impression des SRAF mais elle ne permet pas de confirmer ce risque avec une certitude absolue. Néanmoins, pour une certitude de 95%, elle permet de prédire un risque d'impression des SRAF pour les pas de réseau compris entre 340 et 380 nm. Dans ce cas, l'efficacité de la méthode multi-niveaux est similaire à celle des méthodes DDM et M3D.

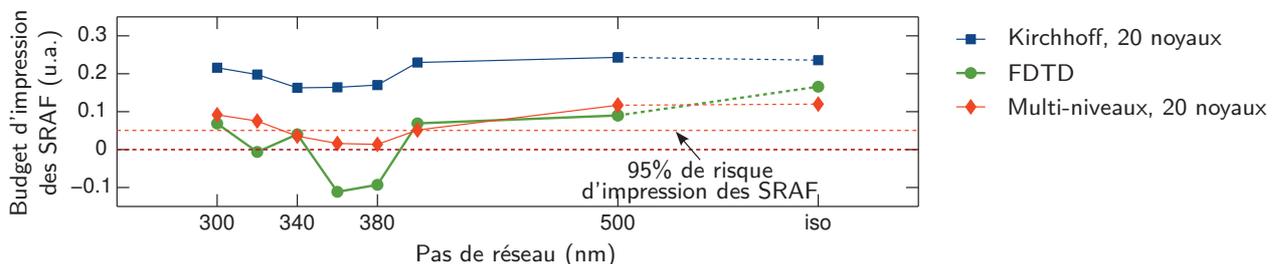


Figure 5.29 – Budgets normalisés en seuil d'impression des SRAF calculés par la méthode de décomposition multi-niveaux, l'approximation de Kirchhoff et la méthode rigoureuse FDTD en fonction du pas de réseau dans le cas d'une exposition à la dose nominale, c'est-à-dire, celle qui donne un CD plaquelette égal à 65 nm pour une ligne appartenant à un réseau de pas égal à 180 nm.

5.3.2.7 Ressources nécessaires et temps d'exécution

Pour chacune des méthodes approchées (Méthode A, Multi-niveaux et l'approximation de Kirchhoff), nous avons simulé¹⁰ les contours dans la résine d'une région d'un dessin ayant une

9. Il suffit de comparer les écarts en CD pour deux valeurs de signes opposés du defocus

10. Pour la partie résine, nous avons utilisé un modèle à seuil constant

superficie égale à $30 \times 30 \mu\text{m}^2$. Nous avons ensuite relevé la taille en espace mémoire requise pour créer les tables optiques du modèle. La figure 5.30 montre que la méthode A requiert entre 5.4 et 15.6 fois plus d'espace disque que l'approximation de Kirchhoff. Alors que la méthode des multi-niveaux utilise les mêmes tables optiques de l'approximation de Kirchhoff et ne rajoute aucun espace disque.

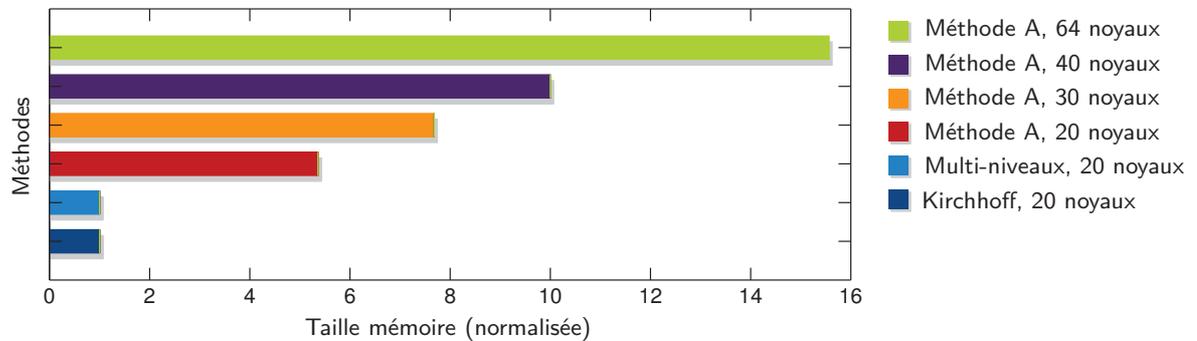


Figure 5.30 – Taille du modèle en mémoire, normalisée sur la base de 300 Mo, relevée pour différentes configurations des méthodes approchées et des tables optiques.

Regardons maintenant le temps requis pour la simulation. D'après la figure 5.31, la méthode A requiert 2 à 4.7 fois plus de temps que l'approximation de Kirchhoff pour la simulation tandis que la méthode des multi-niveaux permet pratiquement de conserver ces temps avec une petite augmentation de 23% due à l'utilisation des 10 niveaux d'amplitude et de phase dans la simulation. Cette dernière est donc avantageuse en terme de ressources nécessaires pour la simulation et en temps de cycle. Grâce à sa rapidité et à la manière simple de l'utiliser, cette méthode a fait l'objet d'un brevet [130].

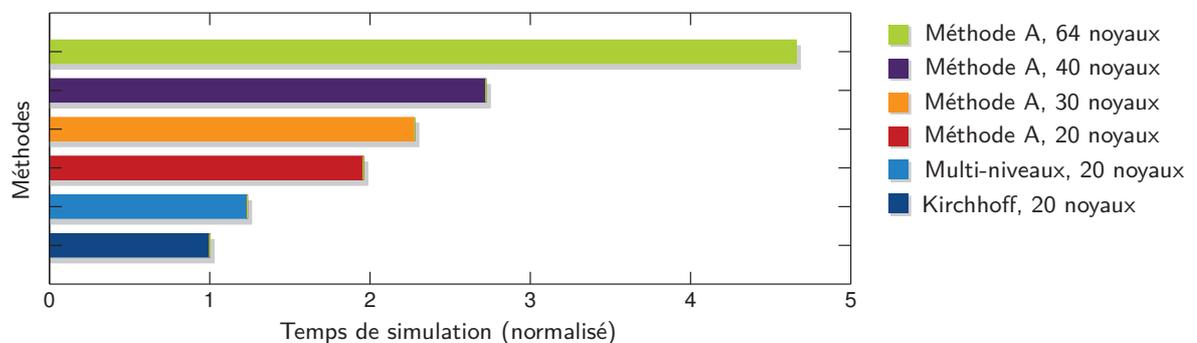


Figure 5.31 – Temps de simulation, normalisé sur la base d'une heure, obtenu pour différentes configurations des méthodes approchées et des tables optiques.

5.3.2.8 Intégration en traitement OPC

L'intégration de la méthode des multi-niveaux dans le processus de traitement par OPC est simple. En effet, nous avons repris le schéma de la figure 1.17 et nous avons inséré une étape supplémentaire entre le déplacement des bords des motifs et la simulation des contours dans la résine. Pendant cette étape, le dessin du masque que l'on veut simuler est remplacé par une succession de bords épais. Ces bords sont obtenus par une simple opération booléenne de négation entre le dessin de départ que l'on redimensionne d'une première distance d_n et celui redimensionné d'une seconde distance d_{n+1} . Le résultat de cette opération logique est un bord d'une épaisseur égale à $|d_{n+1} - d_n|$. Les valeurs des distances d_n sont calculées par une décomposition multi-niveaux du champ électrique effectif d'un bord diffractant. Les couples $[d_n, d_{n+1}]$ désignent les abscisses des points définissant chacun des morceaux du champ multi-niveaux. Ainsi, chacun des bords obtenus est un morceau de masque fin défini par un couple

unique de transmission et de phase. Le schéma de la figure 5.32 illustre les différentes étapes du processus d'un traitement OPC en utilisant la méthode des multi-niveaux dans la modélisation des effets topographiques du masque.

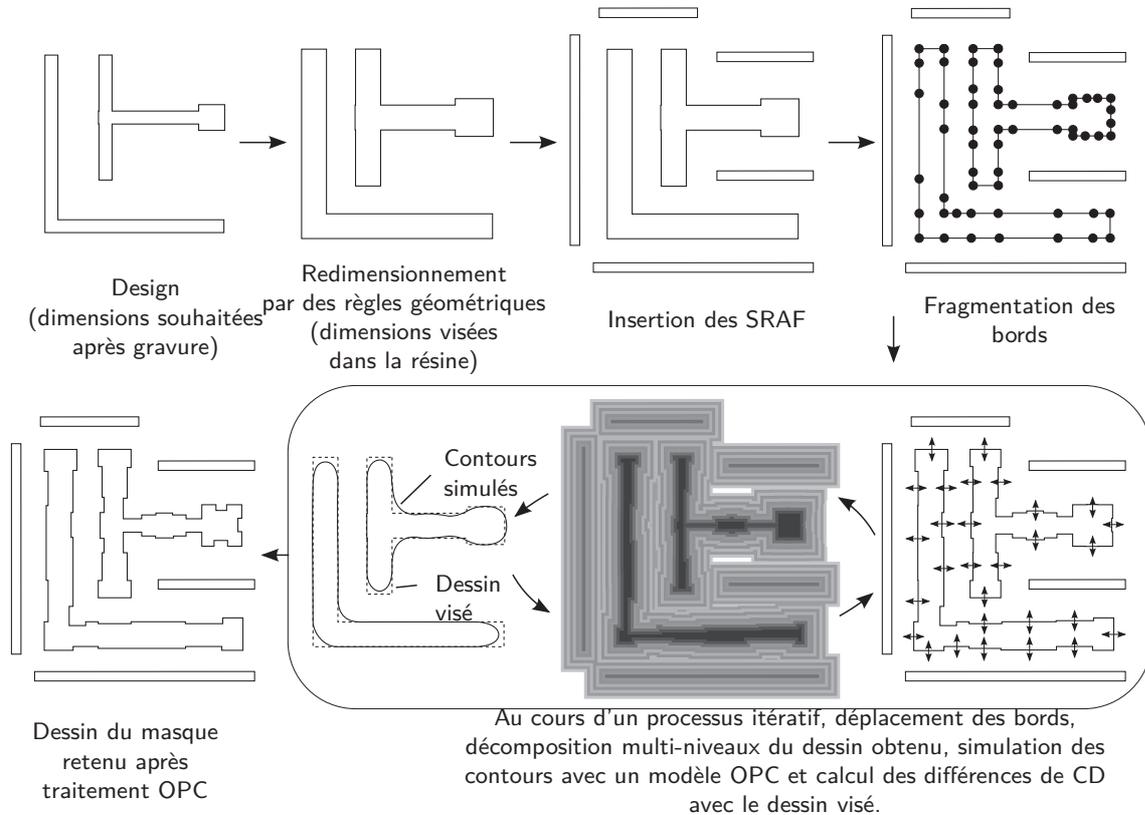


Figure 5.32 – Intégration de la méthode des multi-niveaux dans le cadre d'un traitement OPC du dessin d'un circuit intégré.

Finalement, le modèle empirique utilisé dans la partie résine doit être également ajusté avec la méthode des multi-niveaux ou l'une des méthodes semi-rigoureuses que nous avons rencontrées. Dans le cas de la méthode des multi-niveaux, le dessin des structures du masque qui servent aux mesures de calibration doit être remplacé par une succession de bords ayant chacun une transmission et une phase.

5.4 Conclusion

Les méthodes heuristiques permettent de gagner un temps précieux dans la modélisation des effets topographiques du masque en gardant l'approximation de Kirchhoff dans la modélisation de la partie optique. Dans ce chapitre, nous avons évalué quelques méthodes heuristiques trouvées dans la littérature et/ou développées dans cette thèse. Pour la correction des dimensions critiques des motifs, nous avons exploité la méthode qui permet de compenser les erreurs liées aux effets topographiques par l'intermédiaire d'un modèle empirique. À ce stade, nous avons montré que la simple introduction d'un facteur de dose dans la partie optique ne nous permet pas de modéliser les effets topographiques du masque de manière efficace. Puis, l'utilisation d'un modèle à seuil variable d'ordre 4 nous a permis de réduire les écarts en dimensions critiques entre le calcul rigoureux et l'approximation de Kirchhoff. Cependant, ce modèle, bien qu'il soit plus complexe, n'est pas capable de capturer tous les effets topographiques du masque. Nous avons donc procédé à l'optimisation de ses paramètres. Les résultats obtenus après optimisation sont très satisfaisants. En effet, nous arrivons grâce à ce modèle à recentrer la moyenne de l'erreur (de -0.89 nm à 0.03 nm), mais aussi à réduire l'écart type de cette distribution (de 0.81 nm à 0.59nm).

Dans l'objectif d'améliorer la prédiction de l'impression des motifs diffractants non résolus (SRAF, nous avons exploité une deuxième méthode qui consiste à modifier la taille, la transmission et/ou la phase des motifs afin de rapprocher les profils de l'intensité lumineuse à ceux calculés par l'outil rigoureux FDTD. Nous avons montré qu'il suffit de réduire la taille des SRAF de quelques nanomètres ou d'augmenter leur transmission afin d'atteindre des budgets en seuil d'impression équivalents à ceux calculés avec la méthode FDTD. Cette méthode est donc rapide en modélisation mais nécessite une étude préalable afin d'optimiser les paramètres en jeu.

Finalement, une dernière partie de ce chapitre a été consacrée à décrire une nouvelle méthode approchée que nous avons développée au cours de cette thèse. Cette méthode permet de modéliser les effets topographiques du masque directement dans la partie optique. Elle consiste à rapprocher le champ proche rigoureux par une fonction en escalier. Les différents niveaux de cette fonction seront donc considérés comme la succession de plusieurs masques fins définis chacun par une fonction scalaire de transmission et de phase. La décomposition multi-niveaux du champ rigoureux nécessite donc de trouver une unique fonction scalaire pour remplacer les deux composantes TE et TM du champ rigoureux. Pour ce faire, nous nous sommes basés sur les expressions vectorielles des composantes TE et TM et celle du champ résultant afin d'aboutir à une unique fonction scalaire que nous appelons "champ électrique effectif".

La décomposition multi-niveaux du champ électrique effectif n'était pas évidente à réaliser. Le problème consiste à trouver, parmi une infinité de façons, la meilleure manière qui permet de décrire le signal en entrée tout en réduisant le nombre de niveaux créés. Pour ce faire, nous nous sommes basés sur des phénomènes physiques (les zones de Fresnel) afin de justifier le choix des points d'inflexion comme étant les lieux de découpage. Ensuite, nous nous sommes basés sur l'égalité de conservation de l'énergie de Parseval afin de calculer les transmissions et les phases des niveaux créés de façon à préserver le maximum d'information dans le signal en escalier.

Comme dans le cas des méthodes approchées DDM et M3D, nous avons évalué la méthode des multi-niveaux par rapport à la méthode rigoureuse FDTD et ce en terme de précision de calcul, de capacité à prédire l'impression des SRAF et enfin en terme de rapidité d'exécution. Les résultats obtenus jusque là sont satisfaisants puisque l'on arrive à conserver les temps de calcul, à réduire les erreurs en CD liées aux effets topographiques du masque et d'améliorer ainsi l'estimation des dimensions critiques des motifs dans la résine. De plus, la méthode des multi-niveaux permet d'absorber une partie des erreurs focales et de recentrer les fenêtres de procédé en dose et en focus. Finalement, cette méthode permet d'améliorer la prédiction du risque de l'impression des SRAF et présente ainsi des résultats similaires à ceux des méthodes DDM et M3D.

Contrairement aux méthodes DDM et M3D où la précision peut être uniquement améliorée en augmentant le nombre de noyaux utilisés dans les tables optiques, la méthode des multi-niveaux permet d'atteindre la même précision que la méthode FDTD si les niveaux que l'on applique à un motif donné étaient directement créés à partir d'une simulation rigoureuse. Cette approche, comme nous l'avons déjà noté, ne permet pas de traiter un circuit entier car les niveaux créés sont spécifiques à une géométrie particulière qui est définie par le motif de départ. En revanche, elle peut être utilisée pour effectuer une correction ciblée.



Conclusion générale

Les retards accumulés dans le développement des techniques de lithographie de nouvelle génération NGL (de l'anglais "*Next Generation Lithography*") et de la lithographie optique en extrême ultraviolet EUV suggèrent l'utilisation des méthodes existantes pour les nœuds technologiques actuels. Le maintien de la lithographie optique à 193 nm n'a été possible qu'avec l'amélioration du procédé de photolithographie. C'est grâce aux techniques d'amélioration de la résolution (RET) et aux procédés de doubles expositions qu'il est possible de résoudre des dimensions plus petites que la longueur d'onde d'exposition [131]. La correction des effets de proximité optique (OPC) est une technique d'amélioration de la résolution devenue aujourd'hui indispensable afin de continuer à suivre le rythme de la miniaturisation prédit par la loi de Moore. C'est par une modélisation fine du procédé photolithographique que les modèles de traitement OPC actuels permettent de prédire le moindre risque de défaillance du circuit traité. Aujourd'hui, les modèles OPC doivent relever les défis en terme de précision et les exigences, encore plus strictes, de la feuille de route de l'ITRS. De plus, le traitement OPC d'un circuit doit être rapide afin de réduire les coûts qu'impliquent la préparation du dessin à imprimer sur les masques. C'est dans ce cadre de la modélisation OPC que s'inscrit ce travail de thèse.

Dans une première partie de ce mémoire, nous avons expliqué, par le biais d'une comparaison des résultats de simulation et des mesures expérimentales, qu'un modèle simple (à l'instar du modèle à seuil constant) n'est plus suffisant pour garantir une précision respectant les spécifications de fabrication. L'augmentation de la complexité du modèle permet ainsi de réduire les écarts entre la simulation et l'expérience. Cependant, les modèles avancés incluent de nombreux termes additionnels qui risquent de dégrader leur stabilité numérique. En essayant de chercher une solution pour évaluer la stabilité d'un modèle, nous avons dressé un état de l'art des méthodes existantes et discuté leurs avantages et inconvénients. Nous avons également présenté qualitativement les principales causes d'instabilité des modèles avancés. Nous avons montré que les principales causes d'instabilité des modèles sont liées aux paramètres physiques et empiriques d'une part, et à la qualité et au nombre des mesures expérimentales de calibration d'autre part.

Jusqu'ici, la majorité des méthodes proposées dans la littérature pour l'étude de la stabilité ont pour principal objectif l'évaluation de la sensibilité du procédé photolithographique plutôt que l'étude de la stabilité numérique des modèles. La difficulté de trouver une méthode compatible avec la définition adoptée de la stabilité nous a permis de développer une nouvelle méthode perturbative. Elle consiste à quantifier la sensibilité d'une structure donnée (en terme de dimensions critiques) par rapport à des petites variations d'un paramètre du procédé. Par le biais de cette nouvelle méthode, nous avons évalué la stabilité de différents modèles utilisés pour les nœuds technologiques 65 et 45 nm. En étudiant la stabilité des modèles lorsque l'on injecte une petite perturbation à la position du plan focal ou de l'épaisseur de résine, nous avons montré qu'il existe un fort lien entre l'instabilité des modèles pour des motifs particuliers et les effets optiques mal simulés par la partie optique. Cette étude nous a également montré que la précision d'un modèle est toutefois souvent mise en balance avec la stabilité du modèle lui-même. En effet, nous avons montré qu'un modèle plus simple est globalement plus stable que d'autres modèles complexes. Dans cette partie, à côté de l'unique métrique de la précision utilisée aujourd'hui pour qualifier un modèle OPC, nous avons apporté une autre métrique, devenue aujourd'hui indispensable, qui est le critère de la stabilité.

La dégradation de la précision des modèles simples dans la modélisation des contours dans la résine pour les petites structures est, depuis longtemps, considérée en étroite relation avec les effets chimiques dans la résine qui sont plus importants lorsque les dimensions des motifs sont réduites. Cette logique nous semble juste dans la mesure où la modélisation de l'intensité lumineuse reste exacte. Car, comme dans tout modèle OPC actuel, le point de départ de la modélisation des effets dans la résine est l'intensité lumineuse initiale calculée par la partie optique. Il en résulte qu'une mauvaise modélisation de l'intensité de départ rend la tâche plus difficile aux modèles empiriques qui, par une complexité supérieure, permettent toutefois d'absorber une partie des erreurs optiques. Aujourd'hui, les modèles OPC incluent de nombreuses approximations visant à accélérer le calcul de l'intensité lumineuse à travers le système optique. À titre d'exemple, nous pouvons citer le cas des sources d'illumination simplifiées ou symétrisées, l'absence d'aberrations optiques, l'idéalisation du masque ou alors l'incidence normale de l'onde incidente. Pour la cohérence de cette thèse, nous n'avons présenté que l'étude des effets liés à la topographie du masque. L'étude d'autres paramètres a fait l'objet de publications [109, 132].

Ainsi, dans une seconde partie de ce mémoire, nous avons étudié les limites de validité de l'approximation de Kirchhoff, méthode qui, jusqu'à présent, est la plus utilisée dans les modèles OPC afin de décrire la fonction du masque en raison de sa simplicité et de sa rapidité. L'approximation de Kirchhoff décrit le masque par une unique fonction scalaire de transmission et de phase et consiste donc à une idéalisation du champ proche. Dans cette étude, nous avons montré, à travers des simulations, que cette approximation n'est plus valide pour les structures de petites dimensions où les effets topographiques sont devenus importants. Les erreurs en amplitude et en phase des ordres captés par la pupille de projection se traduisent par des écarts de plus en plus importants au niveau de l'intensité lumineuse lorsque les dimensions sont plus faibles et que l'ouverture numérique est plus élevée. Les erreurs en intensité peuvent être plus ou moins pénalisantes quant à l'estimation des dimensions critiques des motifs et de la profondeur de champ, mais aussi dans la prédiction de l'impression des motifs diffractants non résolus (SRAF).

Nous avons également évalué deux méthodes approchées semi-rigoureuses pour la modélisation des effets topographiques du masque : la méthode de décomposition de domaine (DDM) et celle du masque tridimensionnel (M3D). Les résultats de simulation en terme de précision de calcul sont satisfaisants. En effet, ces méthodes permettent de réduire les erreurs focales et conduisent à une meilleure estimation de la profondeur de champ ainsi que de la position du meilleur focus. Cependant, la précision est impactée par le nombre de noyaux utilisés dans la décomposition des matrices TCC constituant les tables optiques. De plus, les temps d'exécution sont multipliés par un facteur allant de 2 à 4 fois selon la complexité du circuit traité. Pour pallier ce problème, nous avons proposées différentes façons de corriger les effets topographiques du masque, tout en gardant l'approximation de Kirchhoff dans la modélisation de la partie optique. Parmi les méthodes proposées, nous avons exploité celle permettant de réduire les erreurs liées aux effets topographiques du masque par l'intermédiaire d'un second modèle empirique. Nous avons montré que pour garantir une précision adéquate, il est nécessaire d'augmenter la complexité du modèle en rajoutant des termes additionnels.

Enfin, pour garantir la stabilité numérique du modèle empirique, nous avons introduit une nouvelle méthode approchée, permettant d'inclure les effets topographiques directement dans la partie optique. Cette nouvelle méthode consiste à une décomposition multi-niveaux du masque fin. Chaque niveau est défini par une fonction scalaire de transmission et de phase agissant ainsi comme un masque fin. La juxtaposition des différents niveaux permet d'approcher le profil du champ électrique rigoureux calculé par la méthode FDTD. Les résultats de l'évaluation de cette méthode en terme de précision de calcul, de capacité à prédire l'impression des SRAF et de rapidité d'exécution sont satisfaisants. En effet, elle permet d'absorber une partie des erreurs focales, de recentrer les fenêtres de procédé en dose et en focus et améliore ainsi l'estimation des dimensions critiques et la prédiction de l'impression des SRAF tout en conservant les temps d'exécution de l'approximation de Kirchhoff.

Perspectives

Dans cette thèse, nous avons développé deux méthodes dans deux contextes différents. La première est une méthode perturbative qui permet d'évaluer la stabilité d'un modèle empirique vis-à-vis des paramètres du procédé et de détecter les motifs présentant une forte sensibilité du modèle. Nous avons appliqué la méthode perturbative afin d'étudier la stabilité de différents modèles empiriques avancés. Dans cette étude, nous avons choisi le modèle à seuil constant comme référence. Cependant, la partie optique de ce dernier utilise l'approximation de Kirchhoff dans la modélisation de la fonction du masque. Il serait donc intéressant d'inclure les effets topographiques dans le modèle de référence et de réétudier la stabilité des modèles empiriques utilisés. Bien que l'on ait visé le niveau de la grille pour les nœuds 65 nm et 45 nm, la méthode perturbative reste valable pour l'étude de la stabilité d'autres modèles créés pour différents niveaux et pour d'autres nœuds technologiques. De manière plus générale, elle pourrait servir à étudier le comportement de n'importe quel modèle où des effets imprévisibles peuvent se présenter. À titre d'exemple, nous pouvons citer les modèles qui servent dans la modélisation des mouvements des charges électriques.

La deuxième méthode développée dans cette thèse est la décomposition multi-niveaux du champ électrique pour la modélisation des effets topographiques des masques utilisés en photolithographie. La manière adoptée dans la décomposition multi-niveaux est basée sur le champ électrique provenant d'un bord diffractant. Il serait donc intéressant de réétudier et d'évaluer cette méthode en prenant le champ rigoureux calculé directement par la méthode FDTD. Il est intéressant également d'intégrer cette méthode directement dans le processus de correction par OPC afin d'apporter une correction ciblée pour les structures les plus critiques. Mais au delà du contexte lithographique, une application intéressante de cette méthode serait de pouvoir modéliser par morceaux la température après l'étape de recuit qui sert à l'activation des porteurs de charge lors de l'implantation ionique. Grâce à sa rapidité et à la manière simple pour désigner les niveaux par différents calques dans les logiciels de dessin, cette application est d'ores et déjà en cours d'évaluation pour la modélisation des flux thermiques.



Annexes

Annexe A

Éléments de la photolithographie

Cette annexe a pour but de présenter les principaux éléments d'un système de photolithographie, d'expliquer la formation d'image au sein d'un outil de photolithographie par projection et de présenter les différents critères de performance du procédé photolithographique.

A.1 Éléments d'un système de photolithographie

Un système de photolithographie par projection est, de façon générale, constitué de quatre parties principales : la source d'illumination, le masque, l'optique de projection et la plaquette de silicium sur laquelle une résine photosensible est déposée. L'éclairage utilisé dans cette chaîne est de Köhler. Un schéma simplifié de celui-ci est représenté sur la figure A.1.

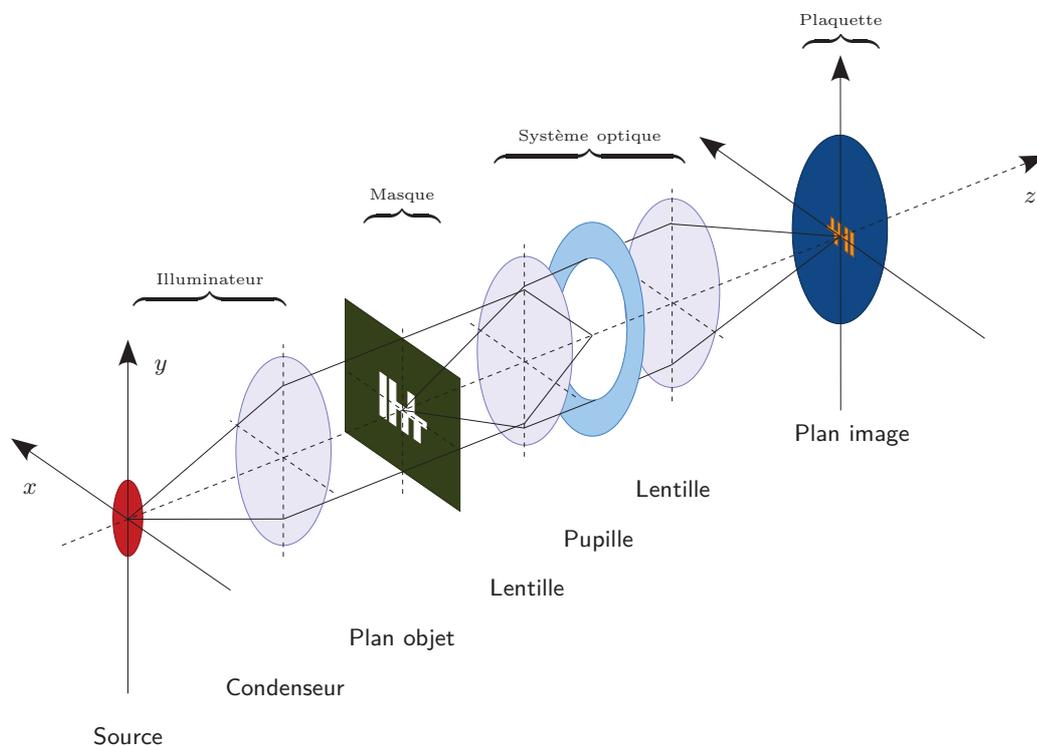


Figure A.1 – Schéma d'un système d'illumination de Köhler.

Dans ce type d'éclairage, l'image de la source d'illumination est projetée dans la pupille de l'optique de projection afin de la séparer du masque. L'uniformité de l'éclairage sur toute la surface du masque est assurée grâce à un système de lentilles, appelé le condenseur, permettant de projeter l'image de la source à l'infini au niveau du plan du masque. Les différents ordres

de diffraction captés par l'optique de projection contribuent à la formation de l'image finale des éléments du masque dans la résine.

A.1.1 L'illumination

L'illumination constitue l'ensemble des dispositifs permettant d'éclairer le masque. Elle est composée d'une source lumineuse et d'un système de diaphragmes et de lentilles, appelé l'illuminateur, permettant de filtrer la lumière et de la projeter uniformément sur la surface du masque. La source lumineuse est une lampe à décharge ou un laser émettant un rayonnement ultraviolet qui permet d'insoler une résine photosensible et modifier ses propriétés de solubilité. L'illumination peut être définie par une longueur d'onde d'exposition λ , une polarisation et une forme géométrique.

La source peut être cohérente lorsqu'elle est ponctuelle, ou partiellement cohérente lorsqu'elle est étendue. Les sources étendues émettent une multitude de faisceaux non cohérents, c'est-à-dire, n'ayant pas forcément une relation de phase entre-eux. Comme le montre le schéma de la figure A.2, le facteur de cohérence partielle de la source σ est défini par le quotient du diamètre de la source dans le plan de la pupille de l'optique de projection, \varnothing_s , par celui de la pupille elle-même, \varnothing_p . Il s'écrit donc ainsi :

$$\sigma = \frac{\varnothing_s}{\varnothing_p}. \quad (\text{A.1})$$

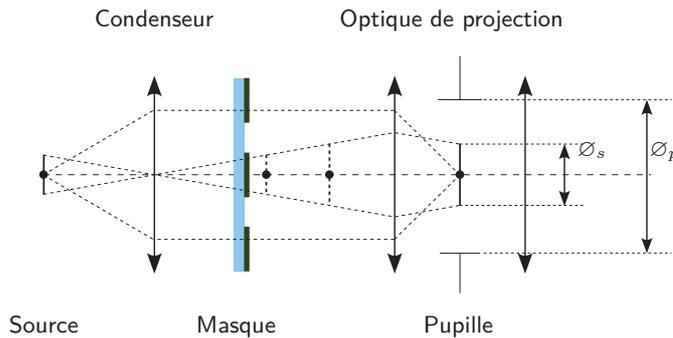


Figure A.2 – Définition du facteur de cohérence partielle dans une illumination du type Köhler. Dans le cas d'une source conventionnelle, le facteur de cohérence partielle est le rapport entre le diamètre de la source dans le plan de la pupille \varnothing_s et celui de la pupille elle-même \varnothing_p .

Dans le cas d'une lumière partiellement cohérente, le facteur de cohérence partielle σ est strictement compris entre 0 et 1.

A.1.2 Le masque

Le masque est le support sur lequel les éléments du circuit sont dessinés. Il est composé d'un substrat transparent épais (de l'ordre de quelques millimètres) sur lequel une fine couche (de l'ordre de quelques dizaines à quelques centaines de nanomètres) d'un matériau absorbant à la longueur d'onde est déposée. Suivant le facteur de réduction du système optique (4 ou 5), les éléments du circuit sont gravés dans cette couche absorbante avec des dimensions quatre ou cinq fois supérieures à celles que l'on veut obtenir sur la plaquette.

Comme dans le cas des circuits intégrés, le transfert des motifs sur le masque nécessite une étape de lithographie. La technique couramment utilisée est une lithographie sans masque (de l'anglais “*maskless*”) permettant, par l'intermédiaire d'un faisceau d'électrons, d'insoler une résine sensible aux électrons. Le dessin du circuit, préalablement découpé en des petits rectangles élémentaires, est transféré à la machine d'écriture. Puis, le faisceau d'électrons, contrôlé par ordinateur, permet d'insoler chacune de ces zones rectangulaires avec une dose bien déterminée. En

dépôt de sa faible vitesse d'écriture, cette technique permet d'atteindre une excellente résolution (de l'ordre de 10 nm pour des équipements de 50keV à 100keV).

L'une des propriétés d'un masque de photolithographie est sa polarité. On parle de masque en champ clair lorsque les motifs sont définis dans la couche absorbante et de masque en champ obscur lorsqu'ils y sont ajourés (voir figure A.3). Le choix de la polarité du masque dépend de plusieurs éléments tels que la forme de la source d'illumination, la taille des motifs du circuit, la tonalité positive ou négative de la résine ainsi que ses propriétés résolvantes.

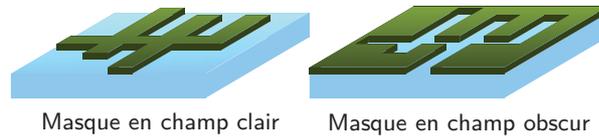


Figure A.3 – Polarité d'un masque utilisé en lithographie optique : a) Masque en champ clair, b) Masque en champ obscur.

Le masque est un élément important en photolithographie car il comporte le dessin du circuit à reproduire sur la plaquette. La qualité des motifs obtenus dans la résine dépend de la qualité du masque. La qualification d'un masque par les industries du semiconducteur est une étape importante de la technologie car elle permet de contrôler certaines spécifications notamment sur :

- les dimensions critiques des motifs gravés sur le masque,
- le déphasage pour les masques à décalage de phase (l'épaisseur du substrat gravé),
- les angles des bords des motifs,
- le contrôle des défauts (comme les motifs manquants ou l'élargissement et l'amincissement localisés des lignes),
- et la résistance à la chauffe (en exposition) et à la contamination chimique en salle blanche (en stockage).

A.1.3 L'optique de projection

L'optique de projection est composée d'un système de lentilles qui permet de collecter une partie des ordres diffractés au niveau du masque et de les projeter dans le plan de la résine. La reconstitution de l'image des éléments du masque s'effectue alors par l'interférence des ondes collectées dans la résine. On appelle image aérienne, l'intensité lumineuse formée au niveau de la plaque dans l'air (ou le liquide d'immersion), c'est-à-dire en l'absence de la résine.

L'optique de projection se caractérise par son facteur de réduction et son ouverture numérique. Le facteur de réduction définit le rapport entre les dimensions des motifs sur plaquette et leurs dimensions sur le masque. À l'heure actuelle, les outils de photolithographie à 193 nm utilisent un facteur de réduction égal à 4, c'est-à-dire que la taille des motifs gravés sur le masque est 4 fois supérieure à celle dans la résine. L'utilité d'une telle technique est de relâcher les contraintes au niveau du masque pour éviter toute perte d'information liée au procédé de fabrication.

Au sein du système optique représenté sur la figure A.1, il est possible de trouver plusieurs diaphragmes. Le diaphragme d'ouverture est celui qui limite, en entrée et en sortie du système optique, l'ouverture angulaire du faisceau lumineux participant à la formation de l'image. La pupille d'entrée est l'objet fictif qui a pour image le diaphragme d'ouverture. Le faisceau traversant ce diaphragme d'ouverture est dit "utile" car il représente la fraction des rayons lumineux issus de la source qui participe à la formation de l'image. Du côté image, la pupille de sortie est définie comme étant l'image du diaphragme d'ouverture.

La pupille est donc un élément virtuel qui définit l'angle maximal d'un ordre diffracté avec l'axe optique pouvant être capté par l'optique de projection. C'est elle qui définit l'ouverture

numérique du système optique. Du côté image, celle-ci est définie par :

$$ON_i = n_i \sin \alpha_i, \quad (\text{A.2})$$

n_i étant l'indice du milieu dans lequel les faisceaux se propagent en sortie de l'optique de projection (air ou liquide d'immersion) et α_i l'angle maximal d'obliquité que peut faire un rayon projeté avec l'axe optique. De la même façon, on exprime l'ouverture numérique, côté objet par :

$$ON_o = n_o \sin \alpha_o, \quad (\text{A.3})$$

où n_o désigne l'indice de réfraction du milieu dans lequel baigne l'objet¹ et α_o le demi-angle maximal du cône de lumière créé par tous les faisceaux émanant de la source et atteignant un point du masque.

L'ouverture numérique maximale d'un outil lithographique est limitée par les facteurs géométriques des lentilles constituant l'optique de projection, en l'occurrence, leurs diamètres et leurs courbures. Nous verrons, dans le cas d'une lithographie à immersion, que l'on peut augmenter l'ouverture numérique en insérant un liquide en sortie du système optique permettant de régler le problème de la réflexion totale de la lumière entre les deux milieux.

Un dernier aspect de l'optique de projection concerne les aberrations optiques et la qualité de polissage des lentilles. Les techniques de projections actuelles, par répétition et balayage (de l'anglais "*step and scan*"), permettent de réduire les aberrations des lentilles en les moyennant sur toute la surface de l'optique de projection. Les derniers outils lithographiques (X1900i de la société ASML) sont équipés de systèmes de correction d'aberrations pour un maximum de performance.

A.1.4 L'empilement multicouche de la résine

A.1.4.1 La résine photosensible

La résine photosensible est un élément essentiel en photolithographie. C'est elle qui permet de transférer les motifs du masque sur la plaquette. La résine photosensible est un matériau organique photosensible habituellement utilisé à des longueurs d'onde dans le spectre ultraviolet. L'interaction de la lumière ultraviolette avec la résine photosensible permet de modifier ses propriétés physico-chimiques pour la rendre soluble dans un développeur aqueux. Outre la photosensibilité, les résines utilisées en photolithographie doivent avoir des propriétés de viscosité et d'adhérence pour permettre un étalement uniforme sur toute la surface du substrat. Elles doivent aussi avoir des propriétés mécaniques et thermiques leur permettant de résister aux traitements des étapes ultérieures tels que la gravure ou l'implantation ionique.

Dans les années passées, une résine photosensible positive communément utilisée avec les longueurs d'onde de 436 et 365 nm est basée sur un mélange de DNQ (diazonaphtoquinone) et d'une matrice Novolac (phénolformaldéhyde) soluble dans un milieu basique. Après exposition, le DNQ se transforme en un acide rendant la matrice Novolac très soluble dans les zones exposées (voir figure A.4). De plus, il joue le rôle d'inhibiteur de dissolution rendant la matrice Novolac moins soluble dans les zones non exposées. C'est cette différence de solubilité entre les zones exposées ou non de la résine qui permet la formation de l'image.

1. L'objet ici est le masque. Une valeur typique de n_o serait donc égale à 1 qui est l'indice de réfraction de l'air ambiant.

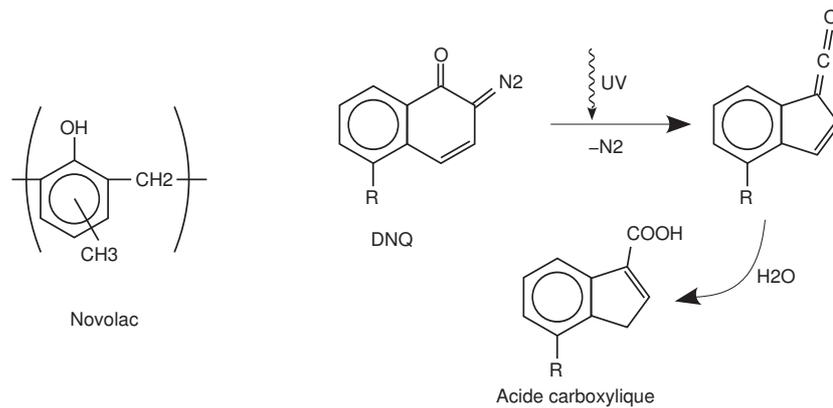


Figure A.4 – Le composant DNQ réagit avec la lumière ultraviolette pour donner un acide carboxylique permettant de rendre la résine très soluble dans les endroits exposés. Source [133]

La forte absorption des résines à base de Novolac et DNQ dans l’ultraviolet profond a fait obstacle à leur utilisation pour les technologies opérant avec les longueurs d’onde 248 et 193 nm. Car, une absorption élevée de la résine conduit à une quantité d’acide formée plus importante dans la partie supérieure de la résine créant ainsi des flancs inclinés après développement. C’est pourquoi, des résines moins absorbantes et plus photosensibles à ces longueurs d’ondes ont été développées. Ces résines sont dites à amplification chimique. Elles sont composées :

- d’un polymère, pratiquement insoluble dans un développeur basique aqueux par le greffage partiel d’un groupement protecteur (inhibiteur de dissolution) sur sa chaîne principale,
- d’un composé photo-générateur d’acide (PAG) qui se transforme en un acide sous un rayonnement ultraviolet de longueur d’onde 284 ou 193 nm,
- et d’une base, qui lors de l’insolation, consomme en partie l’acide formé afin de contrôler sa concentration et limiter sa diffusion

L’acide photo-généré réagit avec le groupement inhibiteur de dissolution et le retire de la chaîne polymère (déprotection ou déblocage) qui devient alors localement soluble dans une solution aqueuse basique. Ensuite, un recuit après exposition PEB (de l’anglais “*Post Exposure Bake*”) est nécessaire pour déclencher un cycle catalytique permettant de régénérer la fonction acide (voir figure A.5). L’acide régénéré contribue à la déprotection d’une autre chaîne de polymère. Ainsi, une seule molécule d’acide photo-générateur PAG peut déclencher entre 800 et 1200 réactions de déprotection [134]. C’est pour cette raison que l’on parle d’amplification chimique.

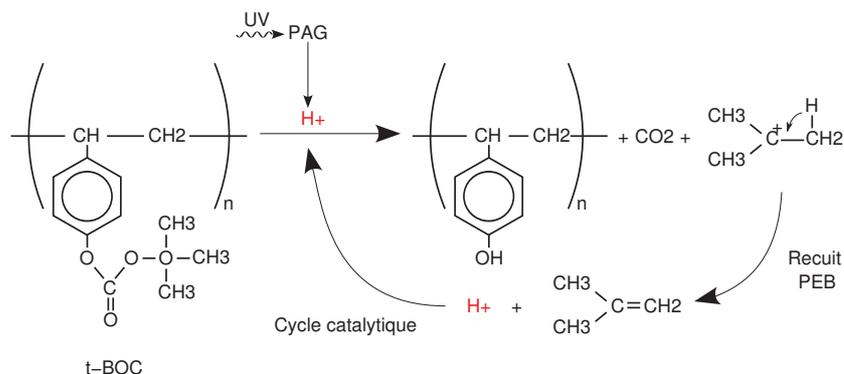


Figure A.5 – Réaction de déprotection du polymère poly(*p*-hydroxystyrène) protégé par un *t*-BOC

Malgré les nombreux avantages des résines à amplification chimique, elles admettent une limitation majeure : la diffusion physico-chimique. En effet, pendant la réaction de déprotection, il est possible que l’acide photo-généré dans les zones exposées se déplace par diffusion vers les zones non exposées de la résine. C’est le résultat de la formation d’un gradient de concentration

entre les deux zones. La migration de l'acide photo-généré entraîne un biais² entre la distribution de l'énergie d'exposition et la distribution finale obtenue. Ce biais dépend essentiellement des propriétés physico-chimiques de la résine et des conditions de recuit, mais il est largement indépendant des conditions d'exposition (focus, dose) et donc il ne peut être éliminé en améliorant le procédé d'exposition. Il est néanmoins possible de réduire la diffusion en augmentant la taille de l'espèce diffusante ou en réduisant la température utilisée pendant le recuit PEB. Une autre approche pour réduire cette migration est d'incorporer dans la résine un additif basique (Quencher) qui consomme la partie diffusante de l'acide afin de contrôler sa concentration.

La diffusion dans le film de résine diminue le contraste de l'image latente et dégrade le contrôle des dimensions. Outre la diffusion chimique, la rugosité des bords des lignes LER (de l'anglais "*Line Edge Roughness*") est un paramètre important à prendre en compte. Pour les structures de petites dimensions, la rugosité des lignes peut être une fraction de la dimension mesurée. Alors que les règles de dessin des circuits intégrés sont de plus en plus petites, la rugosité LER des lignes reste inchangée. La rugosité des résines à amplification chimique demeure un sérieux problème pour le contrôle des dimensions pour les nœuds technologiques 90 nm et au delà. Depuis 2001, la feuille de route de l'ITRS (de l'anglais "*International Technology Roadmap for Semiconductors*") spécifie que la rugosité LER doit être inférieure à 5% du CD³ mesuré.

Finalement, l'utilisation d'un liquide d'immersion, dans le cas de la photolithographie à 193 nm, impose aux résines photosensibles, notamment celles à amplification chimique, d'avoir la propriété d'être imperméable vis-à-vis du liquide afin d'empêcher l'acide de diffuser dans celui-ci. Pour les résines dépourvues de cette propriété, d'autres solutions peuvent être envisagées, comme par exemple l'utilisation d'une couche superficielle d'un matériau imperméable et transparent à la longueur d'onde.

A.1.4.2 Les couches antireflet

Arrivant dans la résine, l'onde subit un nombre de réflexions par les surfaces séparant la résine des autres couches de l'empilement (voir figure A.6-a)). Ces réflexions multiples contribuent à la formation d'ondes stationnaires dans la résine. En effet, l'onde stationnaire est le résultat de la superposition de deux ondes progressives de sens de propagation opposés, comme le montre le schéma de la figure A.6-b). Dans la résine, l'onde stationnaire crée des maxima et des minima d'énergie qui traduisent des concentrations non homogènes de l'acide photo-généré.

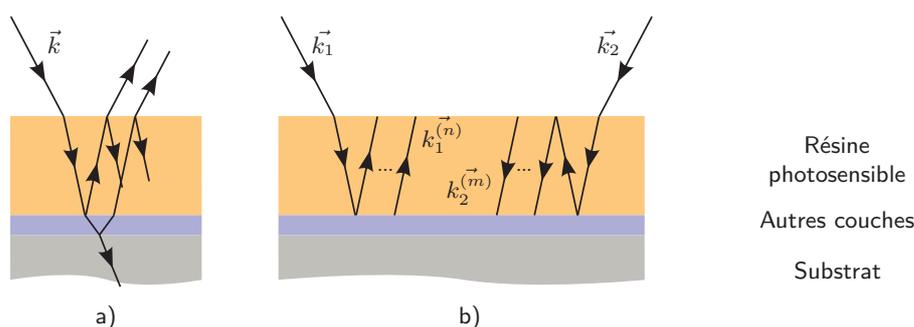


Figure A.6 – a) Représentation schématique du phénomène de réflexions multiples entraînant la formation d'ondes stationnaires dans la résine. L'onde arrivant dans ce multicouche subit de multiples réflexions à travers les surfaces séparant les différentes couches. b) L'onde stationnaire est le résultat de la superposition de deux ondes progressives de sens de propagation opposés.

Après développement, le profil final de la résine présente des ondulations latérales (voir figure A.7) pouvant représenter une fraction de la dimension du motif. Comme dans le cas de la

2. En microélectronique, le biais (de l'anglais "*bias*") est par définition le retrait que l'on mesure au niveau des dimensions des motifs entre deux étapes du procédé, par exemple la lithographie et la gravure.

3. Le CD (de l'anglais "*Critical Dimension*") est un acronyme communément utilisé pour désigner la dimension d'une ligne sur masque ou sur plaquette.

rugosité LER des lignes, les ondes stationnaires entraînent une dégradation du contraste ainsi que le contrôle des dimensions.



Figure A.7 – Profil d'une ligne de résine après développement : a) sans ondes stationnaires, b) avec ondes stationnaires.

Afin de supprimer les ondes stationnaires ou d'atténuer leurs amplitudes, une couche anti-reflective ARC (de l'anglais “*Anti Reflective Coating*”) peut être utilisée. Cette couche est un revêtement de quelques dizaines à quelques centaines de nanomètres formé d'un matériau organique ou inorganique. Elle assure une différence de phase de 180 degrés entre l'onde réfléchie et celle retransmise dans la résine, permettant ainsi de créer des interférences destructives au sein de la résine et d'atténuer les ondes stationnaires. De plus, elle permet d'absorber une partie de l'onde afin d'atténuer son amplitude quand elle sera retransmise dans la résine. Les couches anti-reflectives peuvent être insérées entre la résine et les autres couches de l'empilement, cas des couches dites BARC (de l'anglais “*Bottom Anti Reflective Coating*”), ou alors à la surface de la résine, cas des couches dites TARC (de l'anglais “*Top Anti Reflective Coating*”). L'avantage des couches du type BARC est que l'onde incidente arrivant dans la résine n'est pas atténuée par une autre couche absorbante superficielle, ce qui permet d'avoir un meilleur contraste de l'intensité lumineuse. C'est pourquoi, elle est majoritairement utilisée dans le milieu industriel.

A.2 Formation de l'image

Dans le système d'illumination Köhler, schématisé sur la figure A.1, l'image du masque est rejetée à l'infini dans le plan de la pupille de sortie de l'optique de projection. Le cas de la diffraction à l'infini nous met donc dans le régime de Fraunhofer. En sortie du système optique, l'image est reconstruite à partir des ordres diffractés au niveau du masque. Sachant qu'une partie de ces ordres ne parviennent pas à passer à travers l'optique de projection, l'image finale est dégradée par rapport à l'objet.

A.2.1 Cas d'une source cohérente

Nous allons maintenant nous intéresser à la diffraction au niveau du masque pour une source ponctuelle. Considérons le schéma de la figure A.8. Le masque comportant un réseau de lignes de pas p est illuminé par une onde plane monochromatique provenant de l'illuminateur. L'angle que fait l'onde incidente avec l'axe optique est égal à θ_i . Arrivant sur le masque, l'onde subit une diffraction par le réseau et les rayons diffractés se propagent dans les directions définies par les vecteurs d'onde \vec{k}_m tels que :

$$\vec{k}_m = \vec{k}_i + m \frac{2\pi}{p} \vec{x}, \quad m \in \mathbb{Z} \quad (\text{A.4})$$

avec

$$\vec{k}_i = \frac{2\pi}{\lambda} \begin{pmatrix} \sin \theta_i \\ 0 \\ \cos \theta_i \end{pmatrix}, \quad \vec{k}_m = \frac{2\pi}{\lambda} \begin{pmatrix} \sin \theta_m \\ 0 \\ \cos \theta_m \end{pmatrix}$$

λ étant la longueur d'onde. En projetant cette dernière équation sur l'axe des abscisses, nous obtenons la formule de Bragg, donnant les angles θ_m en fonction de θ_i et du pas de réseau p :

$$\sin \theta_m = \sin \theta_i + m \frac{\lambda}{p}, \quad m \in \mathbb{Z} \quad (\text{A.5})$$

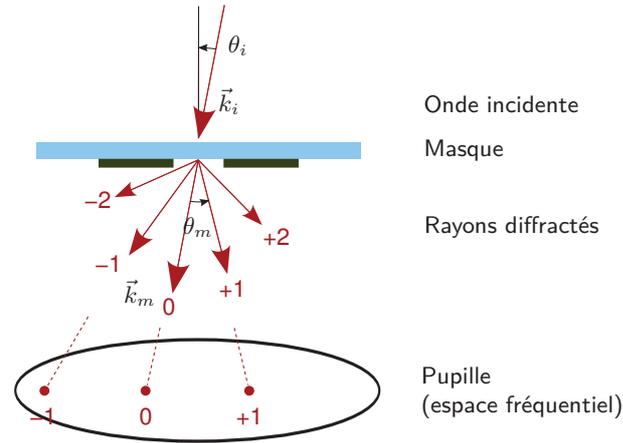


Figure A.8 – Diffraction par un masque comportant un réseau de lignes de pas p pour une onde monochromatique plane faisant un angle θ_i avec l'axe optique (cas d'une source ponctuelle).

La théorie de Fraunhofer peut être décrite avec la formulation mathématique de Fourier. En effet, la fonction du masque, que l'on peut définir par l'amplitude du champ électrique par exemple, peut être décomposée en une infinité de fonctions harmoniques traduisant les différentes fréquences spatiales du réseau, c'est-à-dire, l'ensemble $\left\{\frac{m}{p}\right\}_{m \in \mathbb{Z}}$. Les premières harmoniques de basses fréquences (m petit) comportent la partie importante du signal d'entrée. Les autres harmoniques contiennent l'information sur les détails de l'image. Nous avons tracé, sur la figure A.9-a), l'amplitude du champ électrique d'un réseau de lignes de taille 40 nm et de pas 90 nm et en b) la reconstitution de ce réseau par une somme partielle des harmoniques.

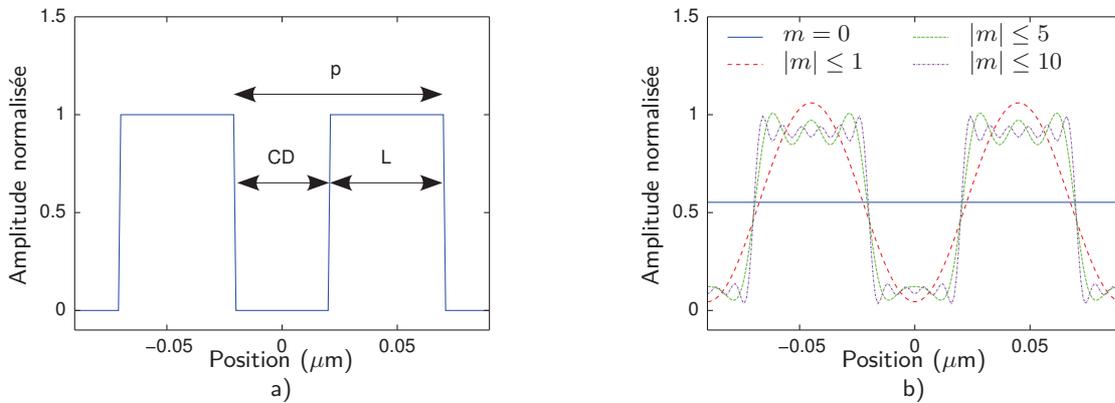


Figure A.9 – a) Amplitude du champ électrique d'un réseau de lignes de taille 40 nm et de pas 90 nm. b) Reconstitution de l'image du réseau à partir d'une somme partielle des harmoniques que contient le signal d'entrée. L'image ressemble de plus en plus au signal d'entrée que l'ordre m est grand.

Nous pouvons clairement remarquer que l'image ressemble de plus en plus au signal d'entrée lorsque le nombre d'harmoniques augmente. De plus, l'image ne se forme que lorsque les harmoniques d'ordres $m = \pm 1$ existent. D'une manière équivalente, l'image ne se forme à la sortie d'un système optique que lorsque les ordres $m = \pm 1$ sont captés par la pupille. D'après l'équation A.5, l'angle θ_m augmente avec la valeur de m . La pupille fait donc passer les ordres de basses fréquences (m petit) et bloque les hautes fréquences (m grand). Elle agit ainsi comme un filtre passe-bas.

L'équation A.5 nous indique aussi que l'angle θ_m est plus grand lorsque le pas de réseau p est petit, et réciproquement. Cela veut dire que la résolution, définie par le plus petit pas de réseau pouvant être imprimé, est limitée par les ordres de diffraction captés par la pupille. Comme nous l'avons indiqué, la formation de l'image nécessite qu'au moins les deux premiers ordres ± 1

soient captés. La résolution s'exprime donc par la capacité du système optique à collecter les premiers ordres de diffraction.

Pour une lumière cohérente, l'image d'un réseau de pas p dans le plan de la pupille est un spectre discret de fonctions Dirac positionnées aux points $\left\{\frac{m}{p}\right\}_{m \in \mathbb{Z}}$. Un exemple de ce spectre est illustré sur la figure A.10, où l'on a un réseau de lignes de taille 40 nm et de pas égal à 90 nm.

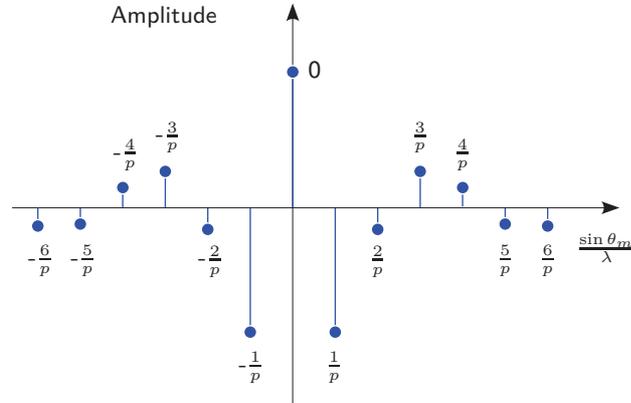


Figure A.10 – Spectre de diffraction d'un réseau de lignes de taille 40 nm et de pas p égal à 90 nm dans le cas d'une source ponctuelle.

Les amplitudes des différents ordres sont données par une décomposition en série de Fourier du signal. Dans le cas d'une incidence normale ($\theta_i = 0$), les amplitudes sont données par :

$$c_m = \frac{1}{p} \int_{-p/2}^{p/2} f(x) e^{-i2\pi \frac{mx}{p}} dx = \begin{cases} \frac{L}{p} & m = 0 \\ \frac{L}{p} \operatorname{sinc}\left(m\pi \frac{L}{p}\right) & m \neq 0 \end{cases}$$

L étant la taille de la fente ($L = p - CD$) et $f(x)$ la fonction du masque qui s'annule en tout point du motif et vaut 1 ailleurs. En incidence oblique ($\theta_i \neq 0$), comme l'équation A.5 l'indique, les positions des ordres seront décalées de la quantité

$$\frac{\sin \theta_i}{\lambda}.$$

A.2.2 Cas d'une source partiellement cohérente

Dans le cas d'une source étendue, le raisonnement est similaire, en admettant que tout point de la source est indépendant, c'est-à-dire, agissant comme une source ponctuelle. Les ordres de diffraction se présentent sous forme de cônes (voir figure A.11) et chacune de leurs enveloppes forme dans le plan de la pupille l'image de la source de départ. Le rayon de l'enveloppe de diffraction est donnée par :

$$\sin \phi = \sigma \times ON, \quad (\text{A.6})$$

σ et ON étant respectivement le facteur de cohérence partielle de la source et l'ouverture numérique de l'optique de projection.

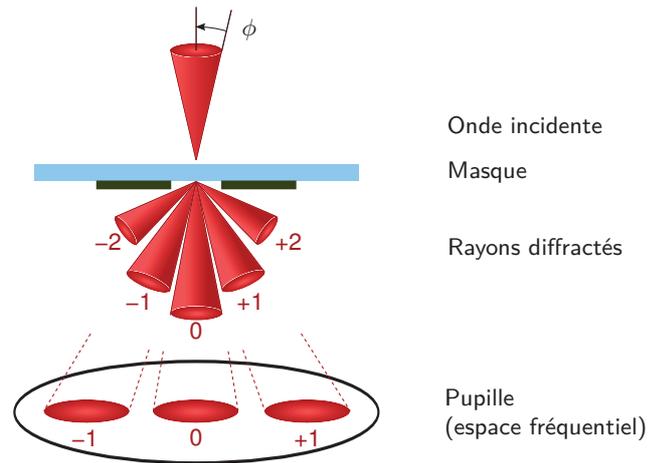


Figure A.11 – Diffraction par un masque comportant un réseau de lignes de pas p éclairé par plusieurs ondes planes incidentes faisant un cône d'angle solide égal à ϕ (cas d'une source étendue).

Le spectre de fréquences ainsi formé dans le cas de l'exemple précédent (c'est-à-dire le réseau de lignes de taille 40 nm et de pas p égal à 90 nm) est obtenu par produit de convolution du spectre discret de la figure A.10 par la fonction de l'image de la source dans le plan de la pupille (définie par son facteur de cohérence partielle σ). Le résultat de convolution pour une source étendue conventionnelle est donné par la figure A.12.

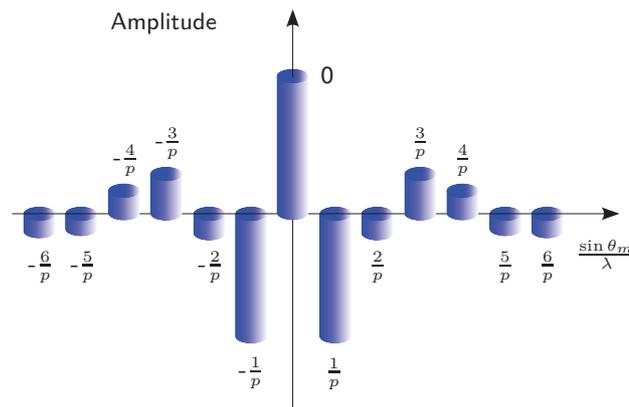


Figure A.12 – Spectre de diffraction d'un réseau de lignes de taille 40 nm et de pas p égal à 90 nm dans le cas d'une source étendue.

L'image est formée lorsque deux faisceaux issus d'un même point de la source interfèrent. On peut ainsi obtenir une interaction entre deux ordres (ordres -1 et 0 ou ordres 0 et 1) ou entre trois ordres (-1, 0 et 1) ou plus selon le pas du réseau. L'intensité finale est la somme de toutes les intensités résultant des interactions entre les différents ordres.

Le contraste de l'image finale est amélioré lorsque le nombre d'interactions entre les ordres est élevé. Dans le cas d'une lumière partiellement cohérente, on parle de surface de recouvrement des ordres. Une surface de recouvrement définit la région de la pupille où deux faisceaux issus d'un même point de la source interagissent ensemble. Nous avons schématisé sur la figure A.13 les différents cas de recouvrement des ordres diffractés dans l'espace fréquentiel pour un réseau de lignes à pas relâché ($> \lambda/ON$), semi-dense ($\approx \lambda/ON$) et dense ($< \lambda/ON$), éclairé avec une source conventionnelle.

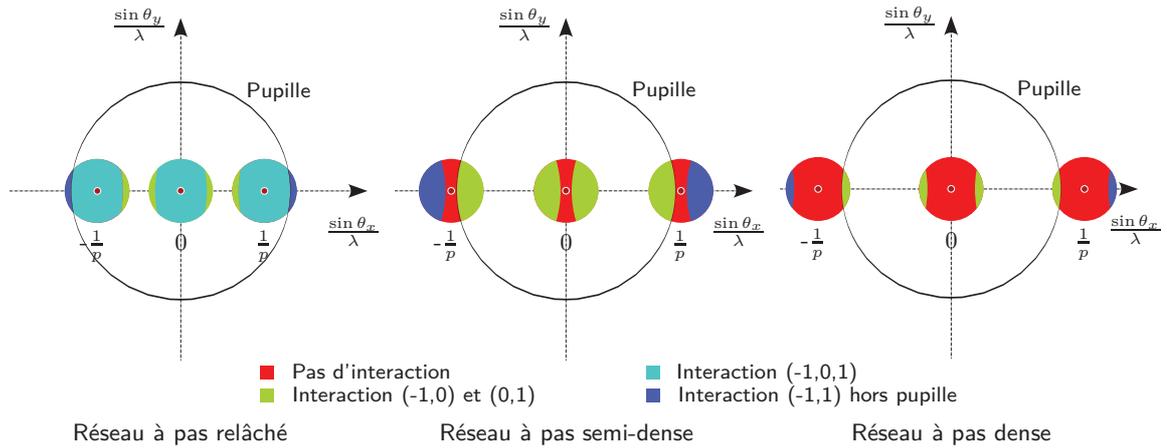


Figure A.13 – Interactions entre les ordres en fonction du pas du réseau pour une source conventionnelle. Les différents cas de surfaces de recouvrement sont représentés avec des couleurs différentes.

D’après la figure A.13, nous pouvons clairement remarquer que plus le pas de réseau est dense, moins (en nombre) de surfaces de recouvrement nous obtenons. De plus, les interactions à trois ondes (-1,0,1) et à deux ondes (-1,1) disparaissent progressivement avec les pas de réseau denses. Au final, il ne reste que les interactions (-1,0) et (0,1) avec des surfaces de recouvrement de plus en plus faibles.

La figure A.14 présente les différents cas de recouvrement des ordres diffractés dans l’espace fréquentiel pour une source annulaire. Contrairement au cas précédent, cette illumination permet d’augmenter les surfaces de recouvrement des ordres, et plus particulièrement des ordres (-1,0) et (0,1) pour les réseaux de pas semi-denses et denses.

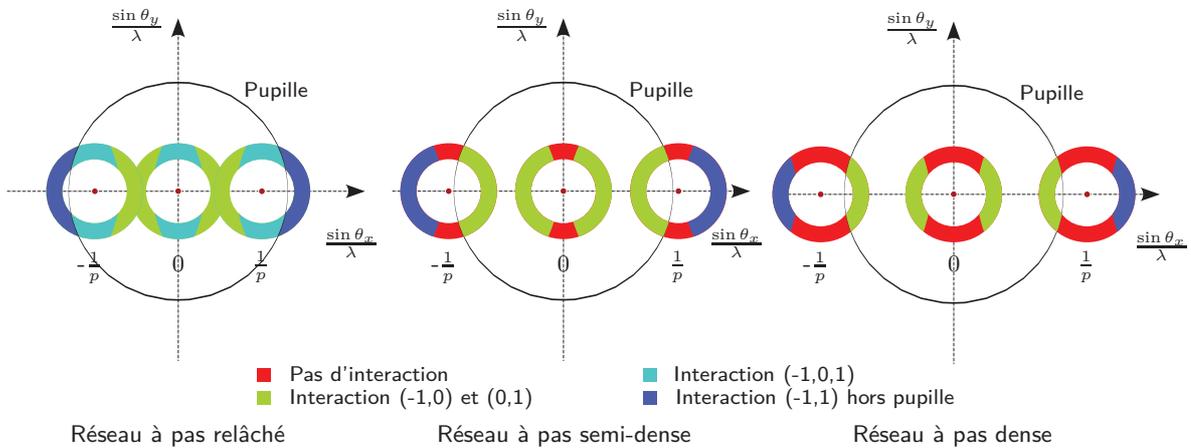


Figure A.14 – Interactions entre les ordres en fonction du pas du réseau pour une source annulaire. Les différents cas de surfaces de recouvrement sont représentés avec des couleurs différentes.

Nous pouvons prédire, en fonction du pas de réseau, de l’ouverture numérique ON et du facteur de cohérence partielle de la source σ les différentes interactions possibles. Pour les deux types d’illuminations citées précédemment, nous avons récapitulé sur le tableau A.1 les différents cas d’interactions que l’on peut envisager pour un pas de réseau donné. Nous verrons, dans la suite de ce chapitre, que ce type d’éclairage, dit “hors axe”, permet d’améliorer la résolution d’un système de lithographie optique.

	Source conventionnelle	Source annulaire
Aucune d'interaction (seule la source est captée)	$p < \frac{\lambda}{ON(1+\sigma)}$	$p < \frac{\lambda}{ON(1+\sigma_{out})}$
Interactions à deux ordres (-1,0) et (0,1) + la source (ordre 0)	$\frac{\lambda}{ON(1+\sigma)} < p < \frac{\lambda}{ON}$	$\frac{\lambda}{ON(1+\sigma_{out})} < p < \frac{\lambda}{ON\sqrt{1-\sigma_{in}^2}}$
Interactions à deux et trois ordres (-1,0), (0,1) et (-1,0,1) + la source (ordre 0)	$\frac{\lambda}{ON} < p < \frac{\lambda}{ON\sqrt{1-\sigma^2}}$	$\frac{\lambda}{ON\sqrt{1-\sigma_{in}^2}} < p < \frac{\lambda}{ON\sqrt{1-\sigma_{out}^2}}$
Interactions à deux et trois ordres (-1,0), (0,1) et (-1,0,1)	$\frac{\lambda}{ON\sqrt{1-\sigma^2}} < p < \frac{\lambda}{ON(1-\sigma)}$	$\frac{\lambda}{ON\sqrt{1-\sigma_{out}^2}} < p < \frac{\lambda}{ON(1-\sigma_{out})}$
Interactions à trois ordres ou plus	$p > \frac{\lambda}{ON(1-\sigma)}$	$p > \frac{\lambda}{ON(1-\sigma_{out})}$

Tableau A.1 – Récapitulation des différents cas d'interaction entre les ordres de diffraction pour un réseau de lignes de pas p donné, dans les deux cas d'illuminations conventionnelle et annulaire.

A.3 Critères de performances en photolithographie

A.3.1 La résolution

La limite de résolution est un critère permettant d'affirmer si l'on arrive à distinguer deux objets éloignés très rapprochés. Elle est donnée par le critère de Rayleigh qui permet d'exprimer l'ouverture angulaire limite permettant d'observer distinctement l'image de deux points à travers une ouverture (voir figure A.15). Dans le cas d'une ouverture circulaire, la figure de diffraction obtenue, connue sous le nom de tâche d'Airy, est formée d'un disque lumineux central entouré de plusieurs anneaux lumineux. Selon Rayleigh, deux images sont discernables lorsque le maximum central de la première coïncide avec le premier minimum de la seconde. La séparation angulaire critique entre deux sources, correspondant au critère de Rayleigh s'écrit ainsi :

$$\sin \theta_c = 1.22 \frac{\lambda}{D}, \tag{A.7}$$

λ étant la longueur d'onde et D le diamètre de l'ouverture. D'après l'équation A.7, on remarque que plus le diamètre D est petit plus l'angle θ_c est grand.

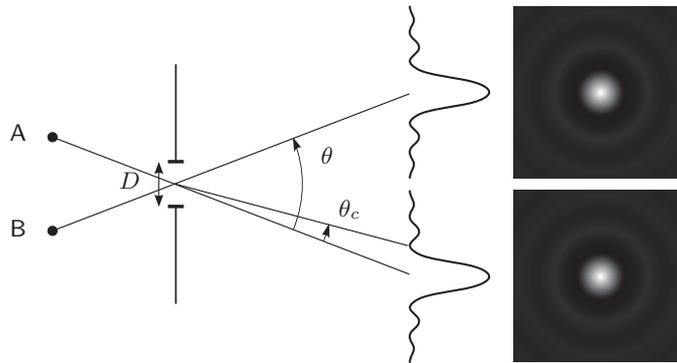


Figure A.15 – Montage simplifié pour la démonstration du critère de Rayleigh.

Dans un instrument d'optique, l'angle θ_c est limité par l'ouverture numérique ($ON = \sin \theta$). Suivant le critère de Rayleigh, le diamètre minimal de la fente pouvant être résolu par le système optique correspond à un angle θ_c correspondant à deux fois l'ouverture numérique ON :

$$D_{min} = 1.22 \frac{\lambda}{2 \times ON} = 0.61 \frac{\lambda}{ON} \tag{A.8}$$

En photolithographie, on introduit un facteur k_1 permettant de prendre en compte les paramètres du procédé lithographique (type d'illumination, cohérence de la source, type du masque, type de résine...). La résolution de l'outil lithographique est définie par rapport au plus petit pas de réseau :

$$p_{lim} = k_1 \frac{\lambda}{ON}, \quad (\text{A.9})$$

La limite inférieure du facteur k_1 est égale à 0.5, ce qui correspond au fait que seulement deux ordres diffractés (0,1) ou (-1,0) puissent être captés par la pupille ($2 \times ON$). Dans la littérature, nous pouvons rencontrer des valeurs de k_1 de l'ordre de 0.3; la résolution étant exprimée cette fois ci par rapport à la dimension minimale résolvable d'un motif (CD) et non pas par rapport au pas du réseau. Dans ce cas, la limite inférieure du facteur k_1 est égale à 0.25.

A.3.2 La profondeur de champ

La profondeur de champ DOF (de l'anglais "*Depth Of Focus*") est une grandeur importante en lithographie optique. Elle permet de définir l'intervalle de variations de la position du plan focal dans lequel l'image obtenue reste bien définie. Nous allons utiliser le critère de Rayleigh afin de déterminer la profondeur de champ théorique pour un système optique. Considérons le schéma de la figure A.16. Deux rayons provenant de l'optique de projection interfèrent en deux positions différentes. La différence de chemins optiques entre ces deux rayons est égale à $d(1 - \cos \theta)$ avec d le décalage en focus selon l'axe optique.

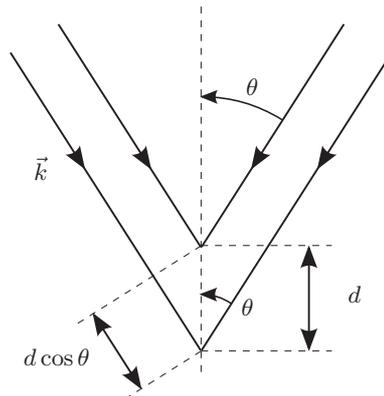


Figure A.16 – Détermination de la profondeur de champ théorique en calculant la différence de chemins optiques entre deux rayons.

Selon Rayleigh, l'image reste nette lorsque la différence de chemins optiques est inférieure au quart de la longueur d'onde. La profondeur de champ DOF théorique est donc définie ainsi :

$$\text{DOF} = 2 \times d_{\lambda/4} = \frac{\lambda}{2(1 - \cos \theta)}$$

Pour θ petit, l'expansion de $\cos \theta$ donne :

$$\cos \theta = \sqrt{1 - \sin^2 \theta} = 1 - \frac{\sin^2 \theta}{2} + \dots$$

En remplaçant $\sin \theta$ par l'ouverture numérique ON, on obtient :

$$\text{DOF} = \frac{\lambda}{ON^2} \quad (\text{A.10})$$

En photolithographie, on introduit un facteur k_2 afin de prendre en compte les différents paramètres d'exposition (cohérence de la source d'illumination, type de masque, aberrations optiques ...) et de l'empilement multicouche de la résine (épaisseurs et indices de réfraction

et d'extinction, longueur de diffusion chimique...). L'expression de la profondeur de champ devient :

$$\text{DOF} = k_2 \frac{\lambda}{\text{ON}^2} \quad (\text{A.11})$$

Dans la pratique, la profondeur de champ DOF est calculée en fixant une variation maximale des dimensions des motifs (CD), par exemple $\pm 10\%$ du CD.

A.3.3 La latitude en énergie

Comme dans le cas de la profondeur de champ DOF, la latitude en énergie EL (de l'anglais "Exposure Latitude") est l'intervalle de variation de la dose d'exposition dans lequel l'image reste bien définie. La latitude en énergie EL est donc un critère de tolérance sur la dose. En pratique, elle est définie par rapport à une variation, de $\pm 10\%$ par exemple, du CD nominal.

A.3.4 La fenêtre de procédé

Les plages de variations possibles en focus et en dose de l'ensemble des structures d'un circuit forment ce que l'on appelle la fenêtre du procédé. Ces plages permettent de respecter les spécifications correspondant à une variation de CD et de nous donner une information sur les capacités du procédé lithographique en terme de résolution. La courbe de la figure A.17 montre un exemple d'une fenêtre de procédé pour un type particulier de motif. La fenêtre de procédé globale est donnée par la superposition de toutes les fenêtres de procédé des motifs du circuit. En pratique, elle est calculée en prenant un motif dense moins sensible aux variations de focus qu'aux variations de la dose, et un motif isolé qui possède des sensibilités opposées.

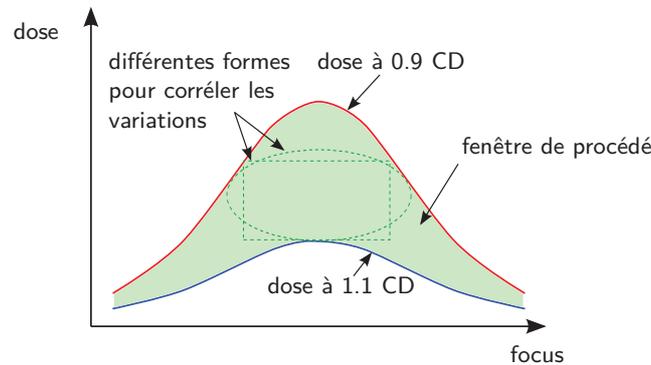


Figure A.17 – Courbe donnant les plages de variations en focus et en dose respectant une variation de la taille du motif égale à $\pm 10\%$ du CD nominal.

Afin de corrélérer les dérives systématiques en focus et en dose du procédé lithographique, nous pouvons utiliser un rectangle (voir figure A.17) donnant la variation en focus (largeur du rectangle) en fonction de la variation en dose (hauteur du rectangle). Le rectangle représente les erreurs maximales en focus et en dose pouvant avoir lieu simultanément sur la plaquette. Pour une distribution d'erreur entièrement arbitraire, nous pouvons utiliser une fonction gaussienne de probabilité de forme suivante :

$$\rho(\Delta E, \Delta F) = \frac{1}{2\pi\sigma_E\sigma_F} e^{-\left(\frac{\Delta E^2}{2\sigma_E^2} + \frac{\Delta F^2}{2\sigma_F^2}\right)},$$

ΔE et ΔF étant les erreurs en dose et en focus respectivement. Les solutions de telle équation sont des ellipses d'axes principaux σ_E et σ_F :

$$\frac{\Delta E^2}{(3\sigma_E)^2} + \frac{\Delta F^2}{(3\sigma_F)^2} = 1$$

Les variations simultanées de l'erreur en dose et de l'erreur en focus (largeur et hauteur du rectangle, ou axes principaux de l'ellipse) décrivent une courbe bien connue, la courbe EL/DOF. Un exemple de cette courbe est présenté sur la figure A.18. La courbe EL/DOF permet de séparer l'espace de variations du focus et de la dose en deux régions selon que la dimension du motif (CD) respecte ou non les spécifications. La superposition de plusieurs courbes EL/DOF calculées pour différents types de structures permet de mieux définir les limites de la fenêtre de procédé.

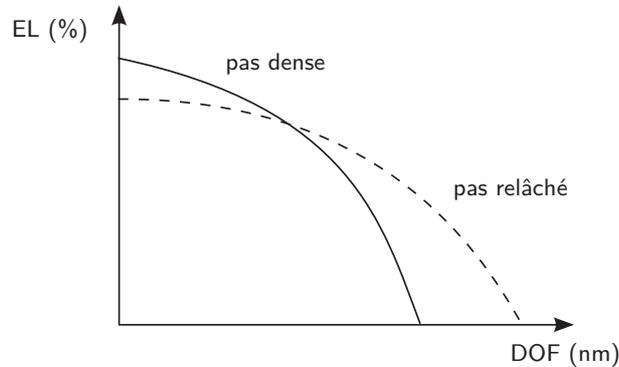


Figure A.18 – Exemples de courbes EL/DOF pour deux réseaux de lignes à pas dense et relâché.

Une autre méthode permettant de quantifier les performances du procédé lithographique est d'utiliser les courbes de Bossung. Ces courbes sont caractéristiques d'un motif particulier et permettent d'obtenir le CD pour une condition de dose donnée en fonction du décalage en focus (que l'on note defocus). Un exemple de courbes de Bossung est présenté sur la figure A.19.

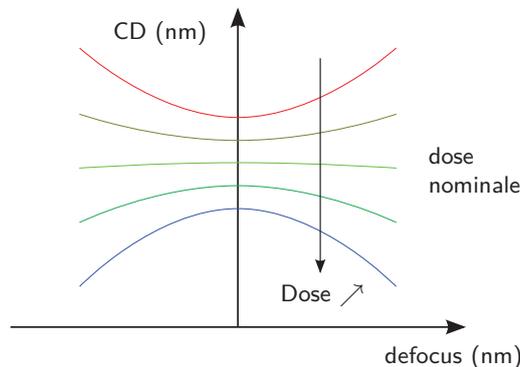


Figure A.19 – Exemple de courbes de Bossung donnant le CD d'un motif en fonction du décalage en focus pour différentes valeurs de la dose d'exposition.

On peut déterminer la sensibilité du motif aux variations de la dose et du focus en observant l'allure des courbes de Bossung. En effet, une courbure importante du profil de ces courbes correspond à une sensibilité élevée aux variations du focus. En revanche, un espacement important entre elles correspond à une sensibilité élevée aux variations de la dose.

A.3.5 Le facteur d'erreur du masque

Le facteur d'erreur du masque MEEF (de l'anglais “*Mask Error Enhancement Factor*”) est un autre paramètre permettant de qualifier un procédé lithographique. Ce paramètre permet de quantifier la sensibilité du procédé aux variations des dimensions sur le masque. Il est défini comme décrit dans l'équation suivante :

$$\text{MEEF} = \frac{\partial \text{CD}_{\text{Résine}}}{\partial \text{CD}_{\text{Masque}}} \times M, \quad (\text{A.12})$$

M étant le facteur de réduction du système optique. Le MEEF permet ainsi d'évaluer la linéarité des dimensions dans la résine en fonction de celles du masque. Dans le cas d'un transfert idéal, le procédé est linéaire et la valeur du MEEF est égale à 1. Néanmoins, les aberrations, les effets de proximité optique et la diffusion chimique dans la résine entraînent une non linéarité du transfert. C'est pourquoi le MEEF est fonction des paramètres du procédé lithographique (cohérence de la source d'illumination, type de masque, conditions d'expositions, type de résine...), du type de motif à résoudre et de l'environnement local du motif sur le masque.



Annexe B

État de l’art sur la stabilité des modèles empiriques

Cette annexe a pour but de présenter les différentes méthodes rencontrées dans la littérature pour l’étude de la stabilité des modèles empiriques utilisés dans la correction des effets de proximité optique OPC.

B.1 Méthode d’un seul facteur à la fois

La méthode d’un seul facteur à la fois OFAT (de l’anglais “*one-factor-at-a-time*”) est une technique qui s’intègre dans le cadre des plans d’expérience. Elle permet d’étudier l’influence de plusieurs paramètres sur la réponse d’un modèle en ne faisant varier qu’un seul paramètre à la fois. L’inconvénient majeur de la méthode OFAT est qu’elle ne permet pas de prendre en compte les interactions entre les paramètres étudiés. De plus, le nombre d’expériences requis pour parvenir à une combinaison optimale des paramètres est élevé.

Dans un contexte lithographique, la méthode OFAT a été employée par K. Elian et al. [65] pour identifier les paramètres sensibles de la résine dans le but d’en comprendre leurs limitations. Elian a étudié l’influence de treize paramètres sur plusieurs réponses telles que le CD, le MEEF, la position de l’isofocale, la latitude d’exposition et la profondeur de champ. Les variations d’un paramètre donné étaient de $\pm 20\%$ autour de sa valeur nominale. Les résultats obtenus avec cette méthode sont en accord avec les plans d’expérience factoriels où l’on fait varier plusieurs paramètres à la fois.

B.2 Méthode de la sensibilité aux paramètres

La méthode de la sensibilité aux paramètres PS (de l’anglais “*Parameter Sensitivity*”) a été introduite par B.S. Ward [66,67] dans le but d’améliorer la régression d’un modèle empirique en traitement OPC par rapport aux mesures expérimentales de dimensions critiques [67]. Elle est basée sur une analyse statistique de la sensibilité du modèle aux variations d’un paramètre donnée pour différents types de structures. Ce raisonnement est justifié du fait que la sensibilité du modèle dépend du type de motif à imprimer. En effet, dans la majorité des cas rencontrés, les structures 2D sont plus sensibles aux variations de focus que celles 1D. Nous pouvons également noter que, dans le cas des motifs 1D, la sensibilité du modèle au focus dépend du pas de réseau. Il existe des configurations qui sont pratiquement insensibles aux variations du focus¹. Il est alors raisonnable de bien choisir les motifs qui vont servir à la calibration du modèle, c’est-à-dire, ne garder que les motifs qui présentent une forte sensibilité.

1. Ce sont des motifs denses présentant un seuil d’exposition proche de l’isofocale.

Le principe de la méthode de la sensibilité aux paramètres est assez simple et intuitif. Il suffit de regarder de combien varie le CD prédit par le modèle si l'on fait varier un paramètre donné. La sensibilité du modèle est calculée par l'intermédiaire de cette formule :

$$S_{Param} = \frac{\partial CD}{\partial Param}.$$

En pratique, une discrétisation de cette formule est utilisée :

$$S_{Param} = \frac{CD(\Delta+) - CD(\Delta-)}{2\Delta},$$

Δ étant une petite perturbation du paramètre et $CD(\Delta+)$ et $CD(\Delta-)$ les valeurs de CD calculées par le modèle suite à une variation du paramètre de $\pm\Delta$.

Le terme de sensibilité S_{Param} est calculé en fonction du motif et du procédé lithographique utilisé. Il est ensuite utilisé pour filtrer les mesures expérimentales ou leur attribuer des poids en fonction de la sensibilité. Il est possible d'enlever les points de mesures les moins sensibles dans le but de réduire le nombre de motifs de calibration et ainsi le bruit de mesure ; ce qui améliore la régression du modèle. Au niveau de la métrologie, cette méthode peut être employée afin de réduire le nombre de mesures expérimentales qui servent à la calibration du modèle sans pour autant affecter sa couverture, ce qui réduit efficacement le temps consacré aux mesures [67].

Dans cette étude de la stabilité, l'avantage de la méthode de la sensibilité aux paramètres est qu'elle nous permet d'avoir une estimation rapide de la sensibilité du modèle par rapport aux variations d'un paramètre donné et par type de motif (deux simulations suffisent pour un paramètre donné). Par contre, elle présente deux inconvénients majeurs : La valeur de la sensibilité obtenue ne nous permet pas de juger si le modèle est stable. Il est alors nécessaire de comparer les valeurs calculées avec celles obtenues avec un modèle de référence (sous l'hypothèse qu'il soit stable). Ensuite, il nous est impossible de comparer les valeurs de la sensibilité obtenues pour des paramètres différents. Il nous est donc impossible de conclure si le modèle est plus sensible aux variations d'un paramètre donné qu'à celles d'un autre.

B.3 Concept des modèles MSM

Le concept de l'approche MSM (de l'anglais "*Manufacturing Sensitivity Models*") a été introduit par L.S. Melvin et al. [68] pour l'étude de la sensibilité du procédé lithographique, et sert à qualifier par modélisation sa robustesse. Un procédé est considéré robuste si les variations des dimensions critiques causées par une perturbation d'un paramètre sont nulles ou acceptables. Dans le cas contraire, le procédé est considéré instable.

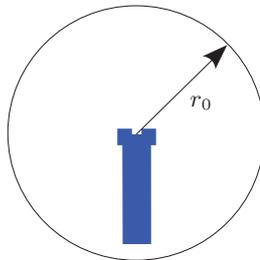


Figure B.1 – Calcul de l'intégrale de l'intensité lumineuse dans la résine dans un rayon r_0 autour d'un bout de ligne. Source : [68]

L'approche MSM consiste à effectuer l'intégrale de l'intensité dans une zone de rayon r_0 autour du motif comme l'illustre le schéma de la figure B.1. Le rayon de la zone définit la

distance maximale d'interaction entre deux motifs distants et est de l'ordre du diamètre optique de simulation.

$$P(r_0) = \int_0^{2\pi} \int_0^{r_0} I(r, \theta) r dr d\theta \quad (\text{B.1})$$

La sensibilité de $P(r_0)$ par rapport aux variations du paramètre considéré est ensuite examinée :

$$I_{\Delta} = \frac{\partial P(r_0)}{\partial \text{Param}}. \quad (\text{B.2})$$

Si la quantité I_{Δ} est approximativement nulle (très faible), le motif est considéré robuste. Dans le cas contraire, le motif est supposé instable par rapport au paramètre considéré. Dans le cas du focus, Melvin et al. ont montré qu'il est possible d'utiliser l'approche MSM afin d'identifier les cas de structures sensibles aux variations de focus [68]. La figure B.2 illustre un exemple d'application de cette technique dans le cas d'un circuit ayant subi un traitement OPC. Les croix désignent les sites pouvant présenter une forte sensibilité aux variations de focus, et donc des problèmes d'impression dans la résine. Un deuxième exemple est illustré sur la figure B.3, où l'on a dessiné la courbe de I_{Δ} en fonction de l'espace entre deux motifs consécutifs d'une matrice de contacts carrés de dimensions $90 \text{ nm} \times 90 \text{ nm}$. Les discontinuités observés sur cette courbe désignent les configurations où l'on peut insérer des motifs diffractants non résolus.

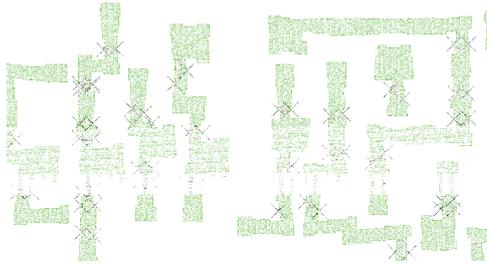


Figure B.2 – Application de l'approche MSM à un design entier. Les croix correspondent aux lieux qui présentent une forte sensibilité aux variations de focus. Source : [68].

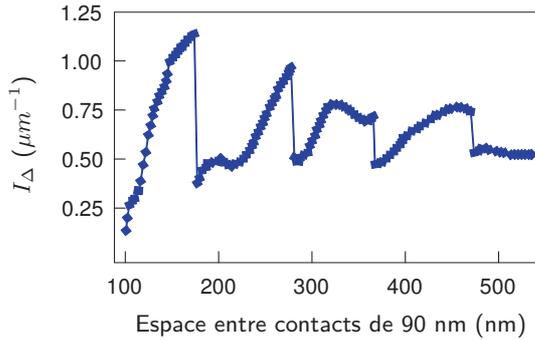


Figure B.3 – Variations I_{Δ} en fonction de l'espace. Les valeurs de I_{Δ} sont calculées pour des matrices de contacts carrés de dimension $90 \text{ nm} \times 90 \text{ nm}$. Le pire cas se produit pour un espace égal à 180 nm . Les discontinuités désignent les configurations où l'on peut insérer des motifs diffractants non résolus. Source : [68].

Les résultats publiés par Melvin et al. [68] dans le cas du paramètre focus ne nous semblent pas être en accord avec l'équation B.1 précédemment citée. En effet, bien que l'on fait varier le focus, l'énergie reçue par le disque de rayon r_0 est théoriquement constante. $P(r_0)$ admet donc une valeur constante quelque soit la valeur du focus. La valeur de I_{Δ} est donc nulle. Ce problème peut être rectifié en calculant la somme des différences en valeurs absolues et non pas la différence des sommes, comme telle la formule suivante décrit :

$$\partial P(r_0) = \int_0^{2\pi} \int_0^{r_0} |\partial I(r, \theta)| r dr d\theta$$

Nous avons également détecté un autre problème par rapport au rayon d'interaction r_0 . En effet, l'estimation de I_{Δ} sur un site particulier peut être biaisée si des motifs situés à l'intérieur

du cercle de rayon r_0 sont potentiellement sensibles. La figure B.4 illustre un exemple de dysfonctionnement de cette technique où l'on peut constater que la valeur de I_Δ calculée pour un site peu sensible aux variations du focus² peut être assez élevée et le site est donc jugé sensible.

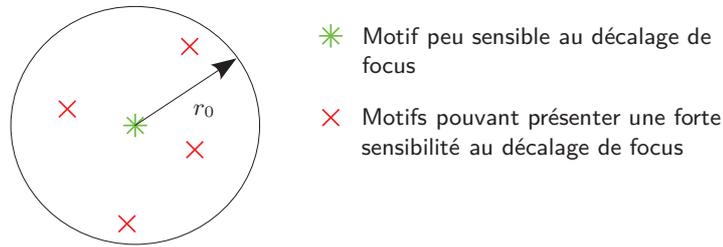


Figure B.4 – Exemple d'un motif peu sensible aux variations de focus entouré d'un ou plusieurs motifs pouvant être très sensibles. La valeur de I_Δ calculée pour ce motif peut être assez élevée.

Pour remédier à ce problème, nous pouvons réduire le rayon r_0 du cercle entourant le site pour réduire l'interaction avec les motifs sensibles voisins ; Ce qui nous permet d'avoir une meilleure estimation de I_Δ . Il faut donc que le rayon soit de l'ordre de la dimension du motif ou de la distance qui le sépare avec son proche voisin. Nous pouvons également annuler toute interaction avec les motifs voisins ($r_0 = 0$), cas déjà proposé par Melvin et al. [70, 71] dans un premier lieu. L'expression proposée par Melvin est décrite par :

$$I_\Delta = \frac{\partial I(x, y)}{\partial Param}.$$

Si l'on veut garder un rayon égal au diamètre optique de simulation, nous pouvons introduire dans l'équation B.1 un terme de pondération afin d'attribuer des poids aux contributions de chaque motif qui se trouve à l'intérieur du cercle. Nous pensons qu'une équation de ce type nous permet de mieux estimer la quantité I_Δ :

$$\partial P(r_0) = \int_0^{2\pi} \int_0^{r_0} |\partial I(r, \theta)| e^{-\frac{r^2}{2\sigma^2}} r dr d\theta, \quad (B.3)$$

où σ est l'écart type de la fonction gaussienne de pondération (distance d'interaction entre motifs).

Finalement, bien que son objectif principal soit l'étude de la sensibilité du procédé lithographique, cette méthode peut être adaptée pour l'analyse de la stabilité d'un modèle empirique. En effet, la comparaison directe des valeurs de I_Δ calculées par le modèle empirique (image latente) et la partie optique seule (intensité lumineuse) nous permet de détecter les motifs où les effets de la résine sont importants.

B.4 La méthode des motifs sensibles à l'imagerie

La méthode des motifs sensibles à l'imagerie OSP (de l'anglais “*Optically Sensitive Patterns*”) est une technique qui permet la détection des structures sensibles aux variations du focus. Cette technique a été publiée par J.E. Jung et al. [72]. Son principe est assez simple ; Il consiste à évaluer la différence d'intensité dans la résine dans les conditions nominale et décalée en focus, comme illustrée sur la figure B.5.

². Nous pouvons imaginer que le site admet un seuil d'exposition proche de l'isofocale. Il est donc peu sensible aux variations du paramètre focus.

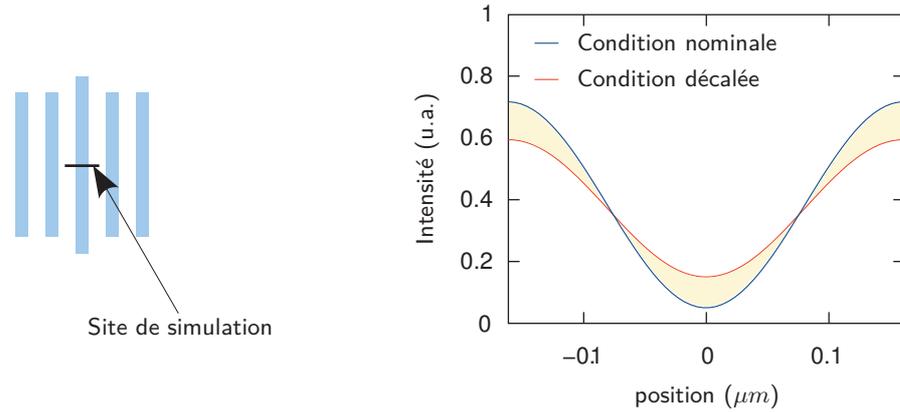


Figure B.5 – Schéma illustrant le calcul, par la méthode des motifs sensibles à l'imagerie, de l'écart entre les intensités dans la résine pour deux conditions différentes du procédé. La surface de la zone située entre les deux courbes constitue la quantité à évaluer.

Cette différence est calculée le long d'un site de simulation. Elle peut être calculée en utilisant la distance des moindres carrés de la manière suivante³ :

$$S = \int_{\text{site}} (I_{\text{Condition nominale}} - I_{\text{Condition décalée}})^2 \quad (\text{B.4})$$

Selon Jung et al., la méthode des motifs sensibles à l'imagerie permet d'évaluer la sensibilité d'un motif aux variations du focus au niveau de la partie optique uniquement. Elle a été principalement utilisée dans le but de détecter les structures non robustes, c'est-à-dire très sensibles à la qualité de l'imagerie [72]. Le but de cette détection est de calibrer le modèle empirique avec les données expérimentales des structures sensibles au focus afin d'améliorer la régression.

Cette méthode n'a pas été employée pour étudier la stabilité des modèles empiriques. Cependant, elle nous est particulièrement intéressante dans la mesure où nous pouvons l'étendre à l'évaluation des différences d'intensité au niveau des images latentes. Une comparaison des résultats obtenus (S) au niveau de l'optique (images aériennes) et du modèle empirique (images latentes) nous permettra de quantifier la sensibilité de ce dernier par rapport à l'imagerie.



3. L'expression analytique qui permet de calculer cette différence n'est pas spécifiée par Jung et al. [72].

Annexe C

Présentation de modèles empiriques avancés

Cette annexe a pour but de présenter les modèles empiriques avancés qui sont actuellement disponibles dans les logiciels de traitement OPC.

C.1 Modèle à opérateur différentiel

Le modèle CM1 est basé sur l'application d'un seuil constant à l'intensité de l'image latente $R(x, y)$. Cette intensité est calculée par une combinaison linéaire de termes M_i selon l'équation :

$$R(x, y) = \sum_i c_i M_i(x, y). \quad (\text{C.1})$$

Les termes $M_i(x, y)$ sont des opérateurs mathématiques qui s'appliquent à l'intensité lumineuse initiale dans la résine $I(x, y)$ selon l'équation :

$$M_i(x, y) = \left((\nabla^k I_{\pm b}(x, y))^n \otimes G_{s,p}(x, y) \right)^{\frac{1}{n}}. \quad (\text{C.2})$$

où

$G_{s,p}(x, y)$ est la famille de fonctions noyaux de Gauss-Laguerre définies par :

$$G_{s,p}(x, y) = \frac{1}{\pi s^2} e^{-\frac{x^2+y^2}{s^2}} L_p\left(2\frac{x^2+y^2}{s^2}\right), \quad L_p(x) = \frac{e^x}{p!} \frac{d^p(e^{-x}x^p)}{dx^p}$$

L_p est le polynôme de Laguerre d'ordre p et $I_{\pm b}(x, y)$ est la partie supérieure ou inférieure de l'intensité lumineuse initiale dans la résine que l'on peut exprimer par :

$$I_{+b}(x, y) = \begin{cases} I(x, y) - b & \text{si } I(x, y) \geq b \\ 0 & \text{sinon} \end{cases} \quad \text{et} \quad I_{-b}(x, y) = \begin{cases} 0 & \text{si } I(x, y) \geq b \\ b - I(x, y) & \text{sinon} \end{cases}$$

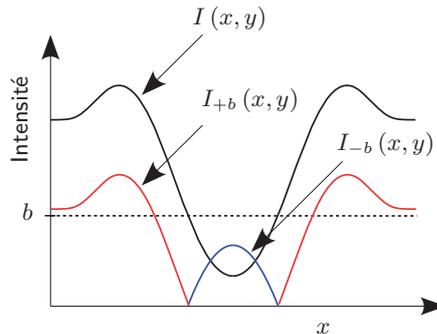


Figure C.1 – Parties supérieure et inférieure de l'intensité lumineuse dans la résine tronquée à un seuil b (cas d'une ligne isolée).

L'équation C.2 montre qu'un terme $M_i(x, y)$ peut inclure une opération différentielle d'ordre k (allant de 0 à 2), appliquée à la partie supérieure ou inférieure de l'intensité lumineuse initiale tronquée à un seuil b , suivi d'un produit de convolution par une fonction noyau de Gauss-Laguerre $G_{s,p}(x, y)$. Les parties supérieure $I_{+b}(x, y)$ et inférieure $I_{-b}(x, y)$ de l'intensité lumineuse initiale permettent de modéliser les réactions de neutralisation de l'acide et de la base respectivement.

Les opérations différentielles d'ordre $k=1$ et $k=2$ ne peuvent avoir lieu que dans le cas où $b=0$, cas où les fonctions $I_{+0}(x, y) = I(x, y)$ (correspondant à la concentration de l'acide) et $I_{-0}(x, y) = 0$ (correspondant à celle de la base) sont continues et différentiables. Dans ce cas particulier, il est intéressant de noter que la différentielle d'ordre 1 est le gradient de l'intensité lumineuse initiale qui n'est autre que la pente locale au point (x, y) . La différentielle d'ordre 2 est le Laplacien de l'intensité et sert à calculer la courbure locale du motif au point (x, y) .

Le produit de convolution par des fonctions noyaux de Gauss-Laguerre sert à modéliser plusieurs types de diffusion dans la résine qui peuvent se produire durant les étapes de recuit et du développement. En effet, les fonctions noyaux $G_{s,p}$ forment une famille de fonctions orthogonales entre-elles au sens de \mathcal{L}_2 . Il est donc clair que toute fonction, utilisée pour décrire la diffusion, peut s'écrire comme combinaison linéaire des fonctions noyaux $G_{s,p}(x, y)$.

La mise à la puissance n de la différentielle de l'intensité et le calcul de sa racine $n^{\text{ième}}$ permettent d'approcher certaines propriétés de l'intensité lumineuse telles que son maximum I_{\max} et son minimum I_{\min} . Par conséquent, le modèle CM1 peut être analogue à un modèle VT5 de premier ordre pour une combinaison donnée de n , k , s et p pour chaque terme $M_i(x, y)$.

C.2 Modèle à convolution étendue

Dans le cas du modèle de Brion, le calcul des contours de la résine s'effectue par la résolution de l'équation $R(x, y) = 0$ où R s'exprime suivant l'équation suivante :

$$\begin{aligned}
 R = & c_A A + c_{M_{av}} M_{av} + c_{A_p} A_p \otimes G_{A_p} + c_{B_p} B_p \otimes G_{B_p} + c_{A_m} A_m \otimes G_{A_m} + c_{B_n} B_n \otimes G_{B_n} \\
 & + c_{AG_1} A \otimes G_{AG_1} + c_{AG_2} A \otimes G_{AG_2} + c_{MG_1} M \otimes G_{MG_1} + c_{MG_2} M \otimes G_{MG_2} + c_{Slope} Slope \\
 & + theta - thresh
 \end{aligned}
 \tag{C.3}$$

où

- $c_A, c_{M_{av}} \dots c_{AG}$ et $theta$ sont des coefficients à optimiser,
- A est l'intensité normalisée de l'intensité lumineuse,
- M est l'image du masque,
- A_p et B_p sont les parties supérieure et inférieure de l'intensité lumineuse tronquées au seuil b_0 ,
- A_m est la partie supérieure de l'intensité lumineuse tronquée au seuil b_m ,
- B_n est la partie inférieure de l'intensité lumineuse tronquée au seuil b_n ,
- M_{av} est la moyenne locale de l'image du masque M dans un rayon autour du point (x, y) ,
- Slope est la pente locale de l'intensité lumineuse au point (x, y) .

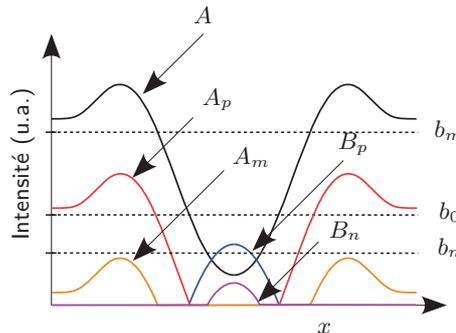


Figure C.2 – Extraction des parties A_p , B_p , A_m et B_n à partir de l'intensité lumineuse A .

Il est intéressant d'expliquer la contribution de chaque terme dans la modélisation des effets physiques qui peuvent se produire dans la résine durant les étapes du procédé. D'abord, la convolution de l'intensité lumineuse initiale A par les gaussiennes G_{AG_1} et G_{AG_2} permet de modéliser deux effets différents causés par la diffusion de l'acide photogénéré. Ces effets peuvent avoir lieu tout au début du recuit où la concentration de l'acide est du même profil que l'intensité lumineuse de départ. Ensuite, la convolution des parties supérieure et inférieure de l'intensité A_p et B_p respectivement par les gaussiennes G_{A_p} et G_{B_p} servent à modéliser la diffusion de l'acide et de la base pendant le recuit. La concentration de la base B_p est celle obtenue après neutralisation de l'acide. Le produit de convolution de la partie A_m par la gaussienne G_{A_m} sert à modéliser les effets de diffusion tridimensionnelle dans les endroits où la concentration de l'acide est élevée. D'une façon similaire, la convolution de B_n par la gaussienne G_{B_n} permet de modéliser des effets comparables causés par une concentration élevée de la base.

En plus de la modélisation des effets de la diffusion chimique, le modèle gaussien de Brion inclut des termes permettant de prendre en compte les effets à courte, moyenne et longue distance pendant l'étape du développement. En effet, le terme M_{av} sert à modéliser les effets à longue distance (de l'anglais "*long range loading effects*") dans la résine, tandis que les convolutions de l'image du masque M par les gaussiennes G_{MG_1} et G_{MG_2} permettent de modéliser les effets à courte distance (de l'anglais "*short range effects*") et à distance moyenne (de l'anglais "*medium range effects*") causés par la variation de la densité locale sur le masque.

Finalement, le terme *thresh* est le seuil d'intensité permettant d'extraire les contours depuis la surface résine obtenue. On peut toutefois spécifier que ce terme n'est pas le seuil constant qu'on a l'habitude d'appliquer au modèle optique pur (intensité lumineuse de départ), mais qu'il s'agit seulement du seuil d'extraction des contours à partir de la surface latente $R(x, y)$. Il existe toutefois un coefficient c_t qui n'intervient pas dans l'équation C.3 mais qui sert à définir le seuil constant du modèle optique pur et de fixer les intervalles de variation des seuils *thresh*, b_0 , b_m et b_n pendant l'optimisation des paramètres.



Annexe D

Théorie scalaire de la diffraction

Cette annexe a pour but de présenter la théorie scalaire de la diffraction et son utilisation pour calculer l'onde électromagnétique diffractée par un objet dans un voisinage proche.

D.1 Historique

Le phénomène de la diffraction fut la première fois observé par l'italien Francesco Maria Grimaldi (1618–1663) qui remarqua la présence de lumière dans l'ombre géométrique d'un objet éclairé par un faisceau lumineux traversant une petite ouverture. Grimaldi remarqua que l'ombre de l'objet apparaît plus large que ce que l'optique géométrique pouvait prédire et est entourée d'une série de franges colorées. Ce phénomène fut de nouveau observé et confirmé par Robert Hook (1635-1703). Hook était le premier à penser que la lumière consiste en une propagation de vibrations rapides et que, dans un milieu homogène, chaque vibration crée une sphère qui grandit continuellement. Cette idée fut adoptée puis reformulée par Christian Huygens (1623-1695) qui énonça le fameux principe, portant son nom, le principe d'Huygens. Ce principe consiste à dire que chaque point d'un milieu subissant une perturbation lumineuse peut être vu comme le centre d'une nouvelle perturbation qui se propage sous la forme d'ondes sphériques se recombinaient de manière à ce que leur enveloppe reconstruise, après un certain temps, le front d'onde de départ.

Au cours du 18^{ème} siècle, le principe d'Huygens subit une forte opposition et ne fut jamais utilisé pour expliquer le phénomène de diffraction. Et ce n'est qu'en 1818 qu'Augustin Jean Fresnel (1788-1827) réussit à employer la construction de l'enveloppe de Huygens et le principe d'interférence de Young pour expliquer la nature rectiligne de la propagation de la lumière et pour comprendre le phénomène de diffraction. Fresnel étudia la diffraction de la lumière par des objets simples (bord droit, ouvertures rectangulaire et circulaires). Quelques années plus tard, les analyses de Fresnel ont été mises sous forme mathématique par Gustav Kirchhoff (1882). Cependant, les problèmes de diffraction en optique furent parmi les plus difficiles à résoudre et les solutions rigoureuses étaient rares. Des méthodes approchées ont été proposées pour simplifier les calculs selon le cas étudié. Parmi ces méthodes, la théorie scalaire de Huygens-Fresnel était la plus adéquate pour traiter la plupart des problèmes.

D.2 Principe de Huygens-Fresnel

Ce principe fut énoncé par Huygens en 1678, puis amplement expliqué par Fresnel en 1818 en se basant sur le phénomène d'interférence. Selon Huygens, chaque point d'un front d'onde est considéré comme le centre d'une perturbation donnant naissance à des ondelettes sphériques. Celles-ci se propagent et leurs enveloppes se recombinaient, un instant plus tard, pour reconstruire le front d'onde. Fresnel a complété ce principe en expliquant que la reconstitution du front d'onde se fait par interférence des ondelettes secondaires d'Huygens. Ce principe est illustré sur la figure D.1 (a) où l'on voit clairement la superposition des ondelettes sphériques secondaires.

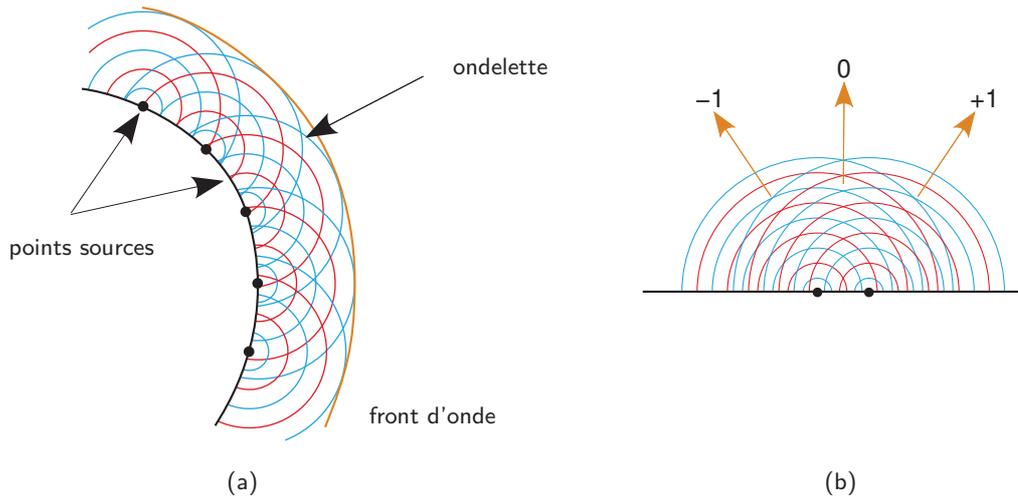


Figure D.1 – (a) Schéma représentatif du principe de Huygens : chaque point source du front d'onde émet des ondelettes sphériques qui se propagent et se superposent pour reconstruire le front d'onde après un certain temps. (b) Application du principe de Huygens pour comprendre l'interférence de la lumière : les lignes blanches représentent les lieux d'intensité maximale (interférence constructive), les lignes noires sont les lieux d'intensité minimale (interférence destructive).

Le principe d'Huygens permet aussi de comprendre l'interférence de la lumière. En effet, la figure D.1 (b) présente l'interférence entre des ondelettes sphériques provenant de deux points sources différents : les interférences constructives résultent de la superposition de deux ondelettes de même phase et d'une manière identique, les interférences destructives se créent lorsque les ondelettes sont de phases opposées.

D.3 Zones de Fresnel

Dans cette partie, nous allons vérifier que le principe de Huygens-Fresnel décrit correctement la propagation de la lumière. Considérons une onde sphérique de centre P_0 se propageant dans le vide et soit S son enveloppe sphérique à un instant quelconque et r la distance parcourue par le front d'onde depuis P_0 . Nous allons exprimer l'onde à un point P supposé à l'extérieur de l'enveloppe S (voir figure D.2).

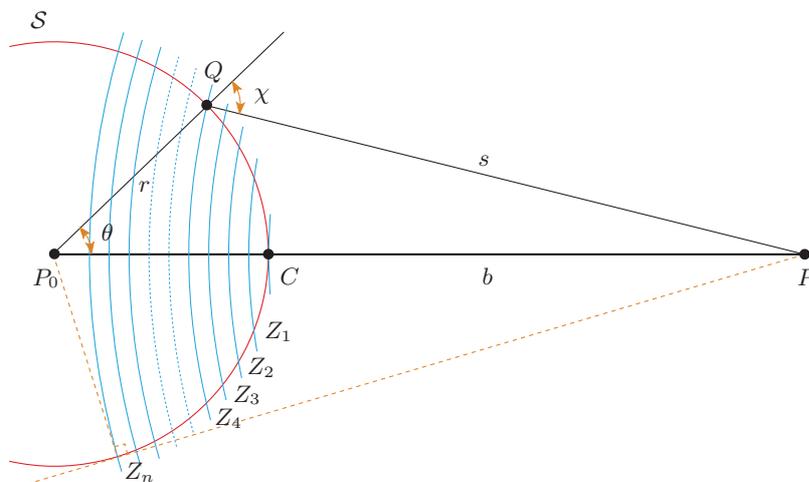


Figure D.2 – Décomposition du front d'onde sphérique en une série finie de zones de Fresnel.

En éliminant le facteur temporel $e^{j\omega t}$, l'amplitude de l'onde à un point Q du front d'onde peut être simplement présentée par $A \frac{e^{jkr}}{r}$, A étant une constante. La contribution d'un élément

de surface dS de l'onde reçue en P s'exprime, selon le principe de Huygens-Fresnel :

$$dU(P) = K(\chi) A \frac{e^{jkr}}{r} \frac{e^{jks}}{s} dS \quad (D.1)$$

où $s = QP$ et $K(\chi)$ est un facteur d'inclinaison qui a été introduit par Fresnel pour tenir compte de l'anisotropie dans la distribution de l'énergie diffractée et de l'absence de diffraction arrière, χ étant l'angle entre les deux vecteurs portés par r et s respectivement. Selon Fresnel, $K(\chi)$ vaut 1 dans la direction originale de propagation, c'est-à-dire quand χ est égal à 0, puis décroît rapidement quand χ augmente pour s'annuler complètement lorsque l'onde n'est plus vue par le point P , c'est-à-dire, quand QP est tangent au front d'onde \mathcal{S} ($\chi = \frac{\pi}{2}$). Alors, il suffit de faire la somme des contributions élémentaires pour déterminer l'amplitude au point P , nous écrivons alors :

$$U(P) = A \frac{e^{jkr}}{r} \iint_{\mathcal{S}} K(\chi) \frac{e^{jks}}{s} dS \quad (D.2)$$

Afin d'évaluer l'expression de l'équation D.2, nous allons utiliser les zones de Fresnel. Considérons des sphères de centre le point P et de rayons $b, b + \frac{\lambda}{2}, b + \lambda, b + 3\frac{\lambda}{2}, \dots, b + p\frac{\lambda}{2}, \dots$, où $b = CP$, C étant le point d'intersection de P_0P et le front d'onde sphérique \mathcal{S} . On note $Z_1, Z_2, Z_3, \dots, Z_p, \dots$, les zones d'intersection des sphères avec le volume défini par \mathcal{S} (cf. figure D.2). En supposant que b et r sont grands devant la longueur d'onde λ , le facteur d'inclinaison $K(\chi)$ varie très peu avec χ pour les éléments de surface d'une même zone Z_p et peut être assimilé à une valeur unique K_p .

En écrivant

$$s^2 = r^2 + (r + b)^2 - 2r(r + b) \cos(\theta)$$

et

$$s ds = r(r + b) \sin \theta d\theta,$$

nous pouvons donc exprimer l'élément de surface dS par :

$$dS = r^2 \sin \theta d\theta d\phi = \frac{r}{r + b} s ds d\phi.$$

La contribution de la $p^{\text{ème}}$ zone à $U(P)$ vaut :

$$U_p(P) = 2\pi A \frac{e^{jkr}}{r + b} K_p \int_{b+(p-1)\frac{\lambda}{2}}^{b+p\frac{\lambda}{2}} e^{jks} ds$$

ou bien,

$$U_p(P) = \frac{2\pi}{jk} A \frac{e^{jk(r+b)}}{r + b} K_p e^{jk\frac{p\lambda}{2}} \left(1 - e^{-jk\frac{\lambda}{2}}\right).$$

En substituant k par $\frac{2\pi}{\lambda}$, nous pouvons écrire $e^{jk\frac{p\lambda}{2}} \left(1 - e^{-jk\frac{\lambda}{2}}\right) = e^{jp\pi} (1 - e^{-j\pi}) = 2(-1)^p$.
D'où

$$U_p(P) = 2j\lambda A \frac{e^{jk(r+b)}}{r + b} K_p (-1)^{p+1}$$

En sommant sur toutes les zones, l'amplitude de l'onde vue en P s'exprime :

$$U(P) = \sum_{p=1}^n U_p(P) = 2j\lambda A \frac{e^{jk(r+b)}}{r + b} \sum_{p=1}^n K_p (-1)^{p+1}$$

Appelons Σ la somme $\sum_{p=1}^n K_p (-1)^{p+1} = K_1 - K_2 + K_3 - \dots + (-1)^{n+1} K_n$. Nous pouvons réécrire cette série de deux façons différentes :

– Première façon :

$$\Sigma = \frac{K_1}{2} + \left(\frac{K_1}{2} - K_2 + \frac{K_3}{2} \right) + \left(\frac{K_3}{2} - K_4 + \frac{K_5}{2} \right) + \cdots + \begin{cases} \frac{K_n}{2} & \text{si } n \text{ est impair} \\ \frac{K_{n-1}}{2} - K_n & \text{si } n \text{ est pair} \end{cases}$$

– Deuxième façon :

$$\Sigma = \frac{K_1}{2} - \frac{K_2}{2} - \left(\frac{K_2}{2} - K_3 + \frac{K_4}{2} \right) - \left(\frac{K_4}{2} - K_5 + \frac{K_6}{2} \right) - \cdots - \begin{cases} -\frac{K_{n-1}}{2} + K_n & \text{si } n \text{ est impair} \\ -\frac{K_n}{2} & \text{si } n \text{ est pair} \end{cases}$$

Si l'on suppose que le facteur d'inclinaison K_p de la zone Z_p est supérieur à la moyenne arithmétique des facteurs des zones voisines, c'est-à-dire

$$K_p > \frac{K_{p+1} + K_{p-1}}{2},$$

alors, compte tenu de la première façon, nous pouvons dire que :

$$\begin{cases} \Sigma < \frac{K_1}{2} + \frac{K_n}{2} & \text{si } n \text{ est impair} \\ \Sigma < \frac{K_1}{2} + \frac{K_{n-1}}{2} - K_n & \text{si } n \text{ est pair} \end{cases} \quad (\text{D.3})$$

De la même façon, nous pouvons écrire, compte tenu de la deuxième façon :

$$\begin{cases} \Sigma > \frac{K_1}{2} - \frac{K_2}{2} - \frac{K_{n-1}}{2} + K_n & \text{si } n \text{ est impair} \\ \Sigma > \frac{K_1}{2} + \frac{K_2}{2} - \frac{K_n}{2} & \text{si } n \text{ est pair} \end{cases} \quad (\text{D.4})$$

Comme les facteurs K_p ne varient que très peu d'une zone à une autre, nous pouvons considérer que les termes des deux relations D.3 et D.4 sont pratiquement égaux. Par conséquent, nous pouvons approcher la valeur de Σ à :

$$\begin{cases} \Sigma = \frac{K_1}{2} + \frac{K_n}{2} & \text{si } n \text{ est impair} \\ \Sigma = \frac{K_1}{2} - \frac{K_n}{2} & \text{si } n \text{ est pair} \end{cases} \quad (\text{D.5})$$

Cette relation est aussi obtenu si l'on considère que :

$$K_p < \frac{K_{p+1} + K_{p-1}}{2}.$$

Pour la dernière zone Z_n , l'angle χ étant $\pi/2$, le facteur d'inclinaison K_n vaut 0. On en déduit alors que :

$$\Sigma = \frac{K_1}{2} \quad (\text{D.6})$$

L'amplitude de l'onde reçue en P s'écrit simplement :

$$U(P) = j\lambda A \frac{e^{jk(r+b)}}{r+b} K_1 = \frac{U_1(P)}{2} \quad (\text{D.7})$$

D.4 Théorème intégral de Helmholtz-Kirchhoff

Le principe de Huygens-Fresnel a été traduit mathématiquement par Kirchhoff en ré-écrivant l'équation de l'onde sous forme d'une intégrale approchée exprimant la solution de l'équation d'onde, en un point arbitraire, en fonction de ses valeurs et de ses dérivées partielles en tout point d'une surface fermée entourant ce point.

Si l'on considère une onde scalaire monochromatique se propageant dans le vide

$$V(r, t) = U(r) e^{j\omega t},$$

la partie spatiale $U(r)$ de cette onde doit satisfaire l'équation d'onde de Helmholtz

$$(\nabla^2 + k^2)U = 0 \quad (\text{D.8})$$

où $k = \frac{\omega}{c}$, c étant la vitesse de la lumière dans le vide.

Si l'on prend une deuxième onde U' qui, aussi, satisfait l'équation d'onde de Helmholtz

$$(\nabla^2 + k^2)U' = 0 \quad (\text{D.9})$$

En multipliant les équations D.8 par U' et D.9 par U et en soustrayant les résultats, nous pouvons écrire :

$$U'\nabla^2 U - U\nabla^2 U' = 0 \quad (\text{D.10})$$

En utilisant le théorème de Green¹, et si l'on considère que U et U' sont continues et différentiables en tout point d'un volume \mathcal{V} et de la surface \mathcal{S} entourant ce volume, nous pouvons écrire :

$$\iiint_{\mathcal{V}} (U\nabla^2 U' - U'\nabla^2 U) d\mathcal{V} = - \iint_{\mathcal{S}} \left(U \frac{\partial U'}{\partial n} - U' \frac{\partial U}{\partial n} \right) d\mathcal{S}$$

avec \vec{n} orienté dans le sens entrant à la surface \mathcal{S} .

Compte tenu de l'équation D.10, on en déduit que :

$$\iint_{\mathcal{S}} \left(U \frac{\partial U'}{\partial n} - U' \frac{\partial U}{\partial n} \right) d\mathcal{S} = 0 \quad (\text{D.11})$$

Si l'on choisit U' de la forme $U'(s) = e^{jks}/s$ (onde sphérique), on obtient :

$$\iint_{\mathcal{S}} U \frac{\partial}{\partial n} \left(\frac{e^{jks}}{s} \right) - \frac{e^{jks}}{s} \frac{\partial U}{\partial n} d\mathcal{S} = 0 \quad (\text{D.12})$$

Cependant, la fonction $U'(s)$ que l'on a choisie présente une singularité en $s = 0$, nous allons donc entourer le point P en $s = 0$ par une sphère de rayon r que l'on va tendre vers 0 et étendre l'intégration dans le volume situé entre la surface \mathcal{S} et celle de la sphère que l'on note \mathcal{S}' (cf figure D.3).

1. Le théorème de Green se déduit de celui de Gauss qui stipule que le flux d'un vecteur \vec{q} à travers une surface fermée est égal à l'intégrale de la divergence de ce vecteur sur le volume délimité par cette surface :

$$\iiint_{\mathcal{V}} (\vec{\nabla} \cdot \vec{q}) d\mathcal{V} = \iint_{\mathcal{S}} (\vec{n} \cdot \vec{q}) d\mathcal{S}$$

où \vec{n} est la normale orientée dans le sens sortant de la surface \mathcal{S} entourant le volume \mathcal{V} . En remplaçant \vec{q} par $f\vec{\nabla}g$ puis par $g\vec{\nabla}f$ (f et g étant deux fonctions complexes quelconques continues et différentiables), et en faisant une soustraction des deux équations, nous arrivons facilement au théorème de Green à deux fonctions :

$$\iiint_{\mathcal{V}} (f\nabla^2 g - g\nabla^2 f) d\mathcal{V} = \iint_{\mathcal{S}} \left(f \frac{\partial g}{\partial n} - g \frac{\partial f}{\partial n} \right) d\mathcal{S}$$

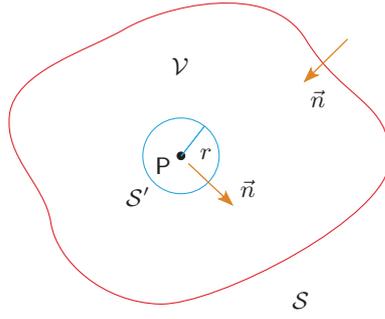


Figure D.3 – Schéma indiquant la région d'intégration utilisée pour le calcul de l'intégrale de Kirchhoff.

L'équation D.12 devient :

$$\left\{ \iint_S + \iint_{S'} \right\} U \frac{\partial}{\partial n} \left(\frac{e^{jks}}{s} \right) - \frac{e^{jks}}{s} \frac{\partial U}{\partial n} dS = 0 \quad (\text{D.13})$$

Or

$$\frac{\partial U}{\partial n} = \left(\vec{n} \cdot \frac{\vec{r}}{r} \right) \frac{\partial U}{\partial r} = \frac{\partial U}{\partial r} \text{ et } dS = r^2 d\Omega, \quad (\text{D.14})$$

$d\Omega$ étant l'angle solide élémentaire. Nous obtenons l'égalité suivante :

$$\iint_{S'} U \frac{\partial}{\partial n} \left(\frac{e^{jks}}{s} \right) - \frac{e^{jks}}{s} \frac{\partial U}{\partial n} dS = \iint_{\Omega} \left(U \frac{e^{jkr}}{r} \left(jk - \frac{1}{r} \right) - \frac{e^{jkr}}{r} \frac{\partial U}{\partial n} \right) r^2 d\Omega$$

Cette intégrale tend vers $-4\pi U(P)$ quand r tend vers 0. On en déduit donc l'expression du théorème intégral de Helmholtz- Kirchhoff :

$$U(P) = \frac{1}{4\pi} \iint_S \left(U \frac{\partial}{\partial n} \left(\frac{e^{jks}}{s} \right) - \frac{e^{jks}}{s} \frac{\partial U}{\partial n} \right) dS \quad (\text{D.15})$$

Cette relation est fondamentale dans la théorie scalaire de diffraction car elle nous permet d'exprimer l'onde U en un point d'observation P en fonction des conditions aux limites calculées sur une surface quelconque autour du point P .

D.5 Formule de Fresnel-Kirchhoff

Le théorème intégral de Helmholtz-Kirchhoff (équation D.15) peut être utilisé pour étudier les effets de diffraction d'une fente se situant dans un plan opaque supposé infini (cf. figure D.4). D'abord, on considère que la fente est grande par rapport à la longueur d'onde utilisée et est petite par rapport aux distances des points P et P_0 . Ensuite, l'intégration sur la surface est segmentée en trois parties différentes : \mathcal{A} , \mathcal{B} et \mathcal{C} . La partie \mathcal{A} définit l'ouverture de la fente, \mathcal{B} est la partie opaque du plan objet et \mathcal{C} est la demi-sphère centrée sur le point d'observation P .

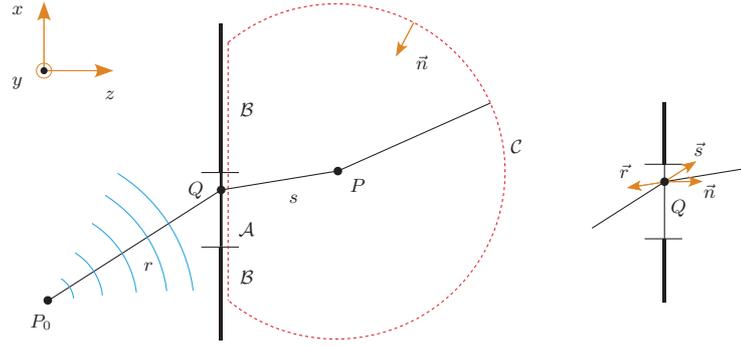


Figure D.4 – Application du théorème intégral de Helmholtz-Kirchhoff dans le cas d'une fente située dans un plan opaque.

Si l'on applique D.15 à ce problème, nous pouvons ré-écrire cette équation sous la forme :

$$U(P) = \frac{1}{4\pi} \left[\iint_{\mathcal{A}} + \iint_{\mathcal{B}} + \iint_{\mathcal{C}} \right] \left\{ U \frac{\partial}{\partial n} \left(\frac{e^{jks}}{s} \right) - \left(\frac{e^{jks}}{s} \right) \frac{\partial U}{\partial n} \right\} dS \quad (\text{D.16})$$

La résolution de cette équation d'onde peut se faire en appliquant les conditions aux limites de Kirchhoff aux parties de surface \mathcal{A} et \mathcal{B} et la condition de radiation de Sommerfeld à la partie \mathcal{C} :

$$\begin{cases} \text{en } \mathcal{A} & U(r) = U_0(r), \quad \frac{\partial U(r)}{\partial n} = \frac{\partial U_0(r)}{\partial n} \\ \text{en } \mathcal{B} & U(r) = U_0(r), \quad \frac{\partial U(r)}{\partial n} = 0 \\ \text{en } \mathcal{C} & \lim_{r \rightarrow \infty} r \left(\frac{\partial U(r)}{\partial n} - ikU(r) \right) = 0 \end{cases} \quad (\text{D.17})$$

où

$$U_0(r) = \frac{Ae^{jkr}}{r} \quad \text{et} \quad \frac{\partial U_0(r)}{\partial n} = \frac{Ae^{jkr}}{r} \left(jk - \frac{1}{r} \right) \cos(\vec{n}, \vec{r})$$

sont les valeurs relatives à l'onde incidente, A étant une constante.

Les conditions aux limites de Kirchhoff consistent à une idéalisation de la distribution du champ réel de l'objet puisque l'on assume que le champ et ses dérivées partielles définis en \mathcal{A} sont égaux à ceux calculés en l'absence de l'objet et qu'ils sont nuls en tout point de \mathcal{B} . Bien qu'elles paraissent raisonnables, ces hypothèses ne sont valides que pour des distances s et r grandes devant la longueur d'onde λ . Le domaine de validité de ces hypothèses dépend de la longueur d'onde utilisée ainsi que des dimensions de l'ouverture. La condition de radiation de Sommerfeld est valide dans la mesure où l'amplitude de l'onde calculée décroît quand r tend vers l'infini aussi rapidement qu'une onde sphérique.

En tenant compte des hypothèses de Kirchhoff et de Sommerfeld, nous pouvons négliger les contributions des parties \mathcal{B} et \mathcal{C} . Ainsi, l'onde calculée en P s'écrit :

$$U(P) = \frac{jA}{2\lambda} \iint_{\mathcal{A}} \frac{e^{jk(r+s)}}{rs} (\cos(\vec{n}, \vec{r}) - \cos(\vec{n}, \vec{s})) dS \quad (\text{D.18})$$



Annexe E

La méthode FDTD

Cette annexe a pour but de présenter la méthode des différences finies dans le domaine temporel et son application pour résoudre l'équation d'onde en électromagnétisme.

E.1 Historique

La méthode des différences finies dans le domaine temporel (de l'anglais "*Finite-Difference Time-Domain*" ou simplement FDTD, acronyme inventé par Taflové en 1980) est une technique de modélisation numérique qui permet de résoudre les équations de Maxwell de l'électromagnétisme dans l'espace et dans le temps. Elle fut la première fois mise en place en 1966 par Yee qui discrétisa les équations de Maxwell sur une matrice cubique définissant l'échantillonnage du milieu dans lequel l'onde se propage. Ensuite, Taflové et Brodwin (1975) ont étudié la stabilité de l'algorithme de Yee et mis en place un critère de stabilité numérique. Quelques années plus tard, Taflové (1980) a inventé l'acronyme FDTD et publié les premiers modèles valides décrivant la pénétration d'ondes électromagnétiques stationnaires dans une cavité métallique. Mur (1981) a publié les premières conditions d'absorption aux limites de seconds ordres appliquées à la méthode FDTD. En 1983, Taflové et Umashankar ont développé les premiers modèles de diffraction d'ondes électromagnétiques en champs proche et lointain. Puis, en 1988, Sullivan et al. ont réalisé le premier modèle FDTD en 3D pour décrire l'absorption d'une onde électromagnétique par le corps humain.

Depuis les années 1990, les techniques basées sur la méthode FDTD se sont montrées très utiles dans la modélisation de nombreux problèmes d'électromagnétisme traitant la propagation de l'onde électromagnétique et son interaction avec la matière. À l'origine, la méthode FDTD fut utilisée à des fins militaires pour la simulation de la propagation des ondes radios dans les radars et les antennes. Ensuite, comme le résumait Taflové et Hagness [80], son domaine s'est étendu à de nombreuses applications dans le domaine médical (la détection du cancer) comme dans celui de l'électronique (la modélisation des éléments de circuits électroniques) ou alors dans le domaine spatial (la modélisation de phénomènes géophysiques à très basses fréquences dans l'ionosphère). L'intérêt pour les solveurs des équations de Maxwell basés sur la méthode FDTD a augmenté de façon exponentielle depuis les années 1990 (cf. figure E.1). En 2006, on estime à 2000 le nombre des publications parues dans la littérature portant sur la méthode FDTD.

En lithographie, la méthode FDTD est utilisée dans de nombreux logiciels de simulation pour la modélisation des effets topographiques du masque. Cependant, cette modélisation reste toutefois limitée à des masques de petites tailles car les algorithmes basés sur la méthode FDTD requièrent une puissance et un temps de calcul importants.

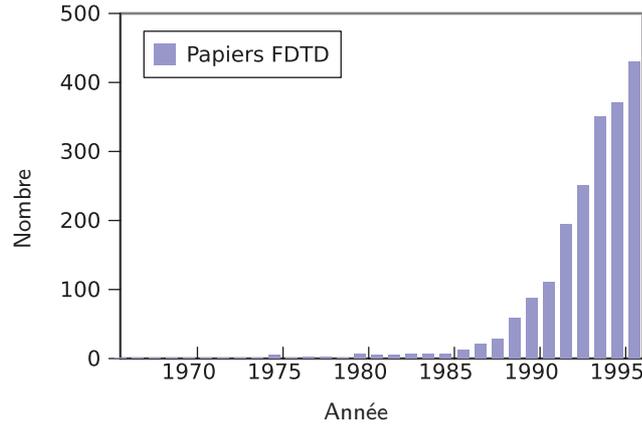


Figure E.1 – Nombre de papiers annuels publiés autour de la méthode FDTD depuis sa création par Yee en 1966 jusqu'à 1996. Source : Shlager et Shneider, "A Survey of the Finite-Difference Time-Domain Literature", Chapitre 1 de "Advances in Computational Electrodynamics : The Finite-Difference Time-Domain Method", A. Taflove, ed., Norwood, MA : Artech House, 1998.

E.2 Principe de la méthode FDTD

La méthode FDTD consiste à déterminer les champs électrique \vec{E} et magnétique \vec{H} en tout point d'une grille cubique dans la région de simulation. Les caractéristiques électromagnétiques des matériaux définissant cette région doivent être connues (indices de réfraction n et d'extinction k). Outre les caractéristiques et la région de simulation, il faut spécifier une source, comme par exemple, une onde incidente, un courant électrique ou un champ électrique.

Les résultats de simulation par la méthode FDTD sont les champs électrique \vec{E} et magnétique \vec{H} calculés en tout point de la grille du domaine et dans le temps. Selon la géométrie du motif, quelques dizaines à quelques centaines d'itérations sont requises afin d'aboutir à un état stationnaire où les champs \vec{E} et \vec{H} convergent vers une solution unique qui sera retenue.

E.2.1 Approximation des différences finies

Considérons une fonction continue $f(x, t)$. À un instant quelconque t_n , cette fonction prend les valeurs f_{i-1}^n , f_i^n et f_{i+1}^n en $x = x_{i-1}$, x_i et x_{i+1} :

$$f_{i-1}^n = f(x_{i-1}, t_n) = f(x_i - \Delta x, t_n) \quad (\text{E.1})$$

$$f_i^n = f(x_i, t_n) \quad (\text{E.2})$$

$$f_{i+1}^n = f(x_{i+1}, t_n) = f(x_i + \Delta x, t_n) \quad (\text{E.3})$$

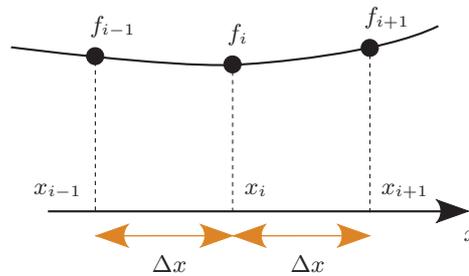


Figure E.2 – Schéma d'une fonction 1D continue

Nous pouvons exprimer f_{i-1}^n et f_{i+1}^n en séries de Taylor autour de $x = x_i$:

$$f_{i-1}^n = f_i^n - \Delta x \left. \frac{\partial f}{\partial x} \right|_{x_i, t_n} + \frac{(\Delta x)^2}{2} \left. \frac{\partial^2 f}{\partial x^2} \right|_{x_i, t_n} - \frac{(\Delta x)^3}{6} \left. \frac{\partial^3 f}{\partial x^3} \right|_{x_i, t_n} + O((\Delta x)^4) \quad (\text{E.4})$$

$$f_{i+1}^n = f_i^n + \Delta x \left. \frac{\partial f}{\partial x} \right|_{x_i, t_n} + \frac{(\Delta x)^2}{2} \left. \frac{\partial^2 f}{\partial x^2} \right|_{x_i, t_n} + \frac{(\Delta x)^3}{6} \left. \frac{\partial^3 f}{\partial x^3} \right|_{x_i, t_n} + O((\Delta x)^4) \quad (\text{E.5})$$

La somme des équations E.4 et E.5 nous permet d'obtenir :

$$f_{i+1}^n + f_{i-1}^n = 2f_i^n + (\Delta x)^2 \left. \frac{\partial^2 f}{\partial x^2} \right|_{x_i, t_n} + O((\Delta x)^4). \quad (\text{E.6})$$

On peut alors exprimer la dérivée seconde de la fonction f par rapport à la variable spatiale x :

$$\left. \frac{\partial^2 f}{\partial x^2} \right|_{x_i, t_n} = \frac{f_{i-1}^n - 2f_i^n + f_{i+1}^n}{(\Delta x)^2} + O((\Delta x)^2) \quad (\text{E.7})$$

De la même façon, on exprime la dérivée seconde de la fonction f par rapport au temps :

$$\left. \frac{\partial^2 f}{\partial t^2} \right|_{x_i, t_n} = \frac{f_i^{n-1} - 2f_i^n + f_i^{n+1}}{(\Delta t)^2} + O((\Delta t)^2) \quad (\text{E.8})$$

E.2.2 Application à l'onde électromagnétique

La méthode FDTD repose sur la résolution directe des équations de Maxwell sous leur forme différentielle :

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (\text{E.9})$$

$$\vec{\nabla} \times \vec{H} = \vec{j} + \frac{\partial \vec{D}}{\partial t} \quad (\text{E.10})$$

$$\vec{\nabla} \cdot \vec{D} = \rho \quad (\text{E.11})$$

$$\vec{\nabla} \cdot \vec{B} = 0 \quad (\text{E.12})$$

où \vec{D} est le déplacement (ou flux) de charge électrique (en $C.m^{-2}$), \vec{E} est le champ électrique (en $V.m^{-1}$), \vec{B} est la densité de champ magnétique (en $Wb.m^{-2}$), \vec{H} est le champ magnétique (en $A.m^{-1}$), ρ est la densité de charge électrique (en $C.m^{-3}$) et \vec{j} est le vecteur densité de courant (en $A.m^{-2}$). Deux autres relations lient \vec{D} à \vec{E} et \vec{B} à \vec{H} sont :

$$\vec{D} = \varepsilon \vec{E} \quad (\text{E.13})$$

$$\vec{B} = \mu \vec{H} \quad (\text{E.14})$$

où ε et μ représentent respectivement la permittivité et la perméabilité du milieu dans lequel l'onde se propage. Dans le cas d'un matériau isotrope, non dispersif et sans sources, ces relations se traduisent, dans un repère cartésien (x,y,z) par :

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} \right) \quad (\text{E.15})$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z} \right) \quad (\text{E.16})$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \right) \quad (\text{E.17})$$

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu} \left(\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \right) \quad (\text{E.18})$$

$$\frac{\partial E_x}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \right) \quad (\text{E.19})$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} \right) \quad (\text{E.20})$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\varepsilon} \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right) \quad (\text{E.21})$$

Les variations spatiales des composantes du champ H régissent l'évolution temporelle des composantes du champ E et réciproquement. La réduction de ce système dans un plan (x,y) permet de le diviser en deux sous systèmes indépendants. L'un fait intervenir les composantes du champ électrique confinées dans ce plan, c'est-à-dire E_x et E_y , et la composante du champ magnétique H_z normale à ce plan, et l'autre fait intervenir les composantes restantes, c'est-à-dire E_z , H_x et H_y . Le premier cas est généralement appelé le mode TM (Transverse Magnétique) et le second cas le mode TE (Transverse Électrique).

L'algorithme de Yee permet de résoudre les équations de Maxwell en utilisant les différences finies. Chaque composante du champ électrique (en 3D) est entourée par quatre composantes du champ magnétique et vice-versa (voir figure E.3).

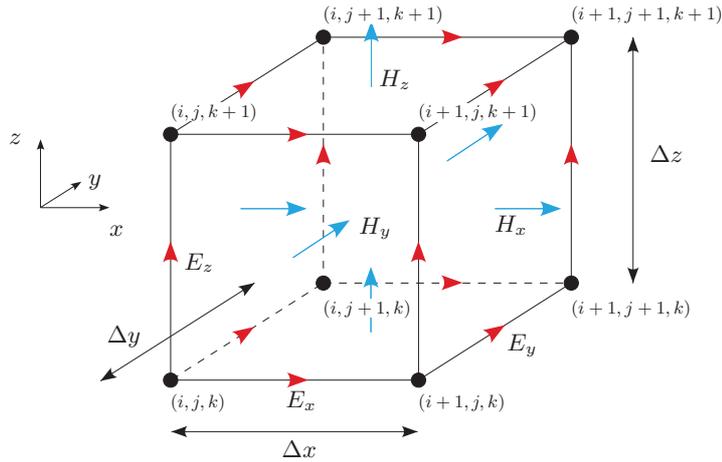


Figure E.3 – Représentation des composantes des champs électrique et magnétique dans un réseau cubique de Yee.

La discrétisation se fait de la manière suivante :

- $[i, j, k] = (i\Delta x, j\Delta y, k\Delta z)$, où i, j et k sont des entiers et $\Delta x, \Delta y$ et Δz sont les incréments selon x, y et z respectivement.
- $t_n = n\Delta t$, où n est un entier.
- $f(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = f_{i,j,k}^n$.

À un instant t_n quelconque, l'expression de la première dérivée partielle par rapport à x de la fonction f est donnée par :

$$\frac{\partial f}{\partial x}(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = \frac{f_{i+\frac{1}{2},j,k}^n - f_{i-\frac{1}{2},j,k}^n}{\Delta x} + O((\Delta x)^2) \quad (\text{E.22})$$

De la même manière, la première dérivée partielle de la fonction f par rapport au temps t s'exprime :

$$\frac{\partial f}{\partial t}(i\Delta x, j\Delta y, k\Delta z, n\Delta t) = \frac{f_{i,j,k}^{n+\frac{1}{2}} - f_{i,j,k}^{n-\frac{1}{2}}}{\Delta t} + O((\Delta t)^2) \quad (\text{E.23})$$

L'algorithme proposé par Yee utilise de manière astucieuse cette discrétisation dans les équations de Maxwell. Il propose tout d'abord une discrétisation des composantes du champ magnétique \vec{H} sur des grilles décalées par rapport à celles du champ électrique \vec{E} , de manière à conserver une répartition des composantes. La grille des composantes H_x est décalée d'un demi pas spatial suivant les axes y et z , celle de H_y est décalée suivant x et z et finalement celle de H_z est décalée suivant x et y . Un décalage temporel d'un demi pas est aussi introduit entre le calcul des champs \vec{E} et \vec{H} . Ce décalage est clairement observé sur la figure E.3.

En utilisant les deux relations E.22 et E.23, nous pouvons exprimer chacune des composantes des champs électrique et magnétique de la manière suivante :

$$E_x|_{i,j,k}^{n+\frac{1}{2}} = E_x|_{i,j,k}^{n-\frac{1}{2}} + \frac{\Delta t}{\varepsilon_{i,j,k}} \left(\frac{H_z|_{i,j+\frac{1}{2},k}^n + H_z|_{i,j-\frac{1}{2},k}^n}{\Delta y} - \frac{H_y|_{i,j,k+\frac{1}{2}}^n + H_y|_{i,j,k-\frac{1}{2}}^n}{\Delta z} \right) \quad (\text{E.24})$$

$$E_x|_{i,j,k}^{n+\frac{1}{2}} = E_x|_{i,j,k}^{n-\frac{1}{2}} + \frac{\Delta t}{\varepsilon_{i,j,k}} \left(\frac{H_z|_{i,j+\frac{1}{2},k}^n + H_z|_{i,j-\frac{1}{2},k}^n}{\Delta y} - \frac{H_y|_{i,j,k+\frac{1}{2}}^n + H_y|_{i,j,k-\frac{1}{2}}^n}{\Delta z} \right) \quad (\text{E.25})$$

$$E_y|_{i,j,k}^{n+\frac{1}{2}} = E_y|_{i,j,k}^{n-\frac{1}{2}} + \frac{\Delta t}{\varepsilon_{i,j,k}} \left(\frac{H_x|_{i,j,k+\frac{1}{2}}^n + H_x|_{i,j,k-\frac{1}{2}}^n}{\Delta z} - \frac{H_z|_{i+\frac{1}{2},j,k}^n + H_z|_{i-\frac{1}{2},j,k}^n}{\Delta x} \right) \quad (\text{E.26})$$

$$E_z|_{i,j,k}^{n+\frac{1}{2}} = E_z|_{i,j,k}^{n-\frac{1}{2}} + \frac{\Delta t}{\varepsilon_{i,j,k}} \left(\frac{H_y|_{i+\frac{1}{2},j,k}^n + H_y|_{i-\frac{1}{2},j,k}^n}{\Delta x} - \frac{H_x|_{i,j+\frac{1}{2},k}^n + H_x|_{i,j-\frac{1}{2},k}^n}{\Delta y} \right) \quad (\text{E.27})$$

$$H_x|_{i,j,k}^{n+1} = H_x|_{i,j,k}^n + \frac{\Delta t}{\mu_{i,j,k}} \left(\frac{E_y|_{i,j,k+\frac{1}{2}}^{n+\frac{1}{2}} + E_y|_{i,j,k-\frac{1}{2}}^{n+\frac{1}{2}}}{\Delta z} - \frac{E_z|_{i,j+\frac{1}{2},k}^{n+\frac{1}{2}} + E_z|_{i,j-\frac{1}{2},k}^{n+\frac{1}{2}}}{\Delta y} \right) \quad (\text{E.28})$$

$$H_y|_{i,j,k}^{n+1} = H_y|_{i,j,k}^n + \frac{\Delta t}{\mu_{i,j,k}} \left(\frac{E_z|_{i+\frac{1}{2},j,k}^{n+\frac{1}{2}} + E_z|_{i-\frac{1}{2},j,k}^{n+\frac{1}{2}}}{\Delta x} - \frac{E_x|_{i,j,k+\frac{1}{2}}^{n+\frac{1}{2}} + E_x|_{i,j,k-\frac{1}{2}}^{n+\frac{1}{2}}}{\Delta z} \right) \quad (\text{E.29})$$

$$H_z|_{i,j,k}^{n+1} = H_z|_{i,j,k}^n + \frac{\Delta t}{\mu_{i,j,k}} \left(\frac{E_x|_{i,j+\frac{1}{2},k}^{n+\frac{1}{2}} + E_x|_{i,j-\frac{1}{2},k}^{n+\frac{1}{2}}}{\Delta y} - \frac{E_y|_{i+\frac{1}{2},j,k}^{n+\frac{1}{2}} + E_y|_{i-\frac{1}{2},j,k}^{n+\frac{1}{2}}}{\Delta x} \right) \quad (\text{E.30})$$

Les itérations sur les entiers i , j et k à un instant quelconque $t_n = n\Delta t$ permettent de propager l'onde électromagnétique dans tout le domaine de simulation. De la même façon, les itérations sur n permettent de suivre les champs électrique et magnétique en fonction du temps. Dans la plupart des cas, le système d'équation se stabilise après quelques itérations et converge vers une solution unique.

E.3 Dispersion numérique de la méthode FDTD

L'échantillonnage du domaine de simulation entraîne l'apparition d'une dispersion non-physique des signaux qui se propagent sur la grille. Cette contrainte exige que le pas de grille est inférieur à la plus petite des longueurs d'onde pour laquelle nous voulons être précis. Le problème est de savoir quel est le meilleur échantillonnage à prendre pour garantir une précision adéquate du calcul. Souvent, la contrainte de 10 pas de grille par longueur d'onde est utilisée, ce qui veut dire que le pas doit être inférieur à $\lambda/10$. Dans quelques situations, où l'on veut obtenir des précisions élevées (par exemple le cas des radars), un échantillonnage de moins de $\lambda/20$ est exigé. Par contre, il est possible d'obtenir des résultats raisonnables si le pas est inférieur à la limite de Nyquist, c'est-à-dire, $\Delta x < \lambda/2$.

Dans le cas où le domaine est composé de plusieurs régions différentes (matériaux différents), le pas de grille global de simulation doit être inférieur à celui attribué à la plus petite des régions. Par exemple, la simulation d'un masque utilisé en photolithographie par la méthode FDTD nécessite l'utilisation d'un pas de grille inférieur à la longueur d'onde mais aussi à l'épaisseur de la plus fine couche de matériau composant le masque.

E.4 Critère de stabilité de la méthode FDTD

Pour que l'algorithme de la méthode FDTD soit stable, il doit bien décrire la propagation de l'onde dans le temps. De ce fait, il est important de bien choisir l'incrément temporel Δt pour que l'onde puisse se propager. On peut montrer que l'algorithme est stable si la condition suivante entre le pas temporel et le pas spatial est vérifiée [81] :

$$v\Delta t \leq \frac{1}{\sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2} + \frac{1}{(\Delta z)^2}}}, \quad (\text{E.31})$$

où v est la plus grande vitesse de propagation de l'onde dans les régions du domaine.



Liste des publications

1 Brevets

- 1/ **M. Saied** and E. Yesilada, "Simulation of the image projected by a mask," US PATENT, App. No. 12/839,817, Fil. date : 20 Jul 2010, Pub. No. US 2011/0022219 A1, U.S. Cl. 700/121 ; 703/2 ; 700/117, Pub. date : 27 Jan 2011. [Online]. Available : <http://www.freepatentsonline.com/y2011/0022219.html>

2 Publications premier auteur

- 2/ **M. Saied**, F. Foussadier, Y. Trouiller, J. Belledent, K. Lucas, I. Schanen, A. Borjon, C. Couderc, C. Gardin, L. LeCam, Y. Rody, F. Sundermann, J.-C. Urbani, and E. Yesilada, "Sensitivity of a variable threshold model toward process and modeling parameters," *Proceedings of SPIE*, vol. 6349, 2006.
- 3/ **M. Saied**, F. Foussadier, J. Belledent, Y. Trouiller, I. Schanen, C. Gardin, J. C. Urbani, P. K. Montgomery, F. Sundermann, F. Robert, C. Couderc, F. Vautrin, G. Kerrien, J. Planchot, E. Yesilada, C. Martinelli, B. Wilkinson, A. Borjon, L. LeCam, J. L. Di-Maria, Y. Rody, N. Morgana, and V. Farys, "Three-dimensional mask effects and source polarization impact on OPC model accuracy and process window," *Proceedings of SPIE*, vol. 6520, 2007.
- 4/ **M. Saied**, F. Foussadier, J. Belledent, Y. Trouiller, I. Schanen, E. Yesilada, C. Gardin, J. C. Urbani, F. Sundermann, F. Robert, C. Couderc, F. Vautrin, L. LeCam, G. Kerrien, J. Planchot, C. Martinelli, B. Wilkinson, Y. Rody, A. Borjon, N. Morgana, J. L. Di-Maria, and V. Farys, "3D mask modeling with oblique incidence and mask corner rounding effects for the 32nm node," *Proceedings of SPIE*, vol. 6730, 2007.
- 5/ **M. Saied**, F. Foussadier, Y. Trouiller, V. Farys, I. Schanen, C. Gardin, F. Sundermann, E. Yesilada, M.-K. Top, L. Perraud, A. Villaret, G. Kerrien, F. Robert, C. Martinelli, F. Vautrin, J.-C. Le-Denmat, J. Planchot, J.-L. Di-Maria, and B. Vandewalle, "Resist model stability and sensitivity study using small variation method," Oral presentation, Bacus, 2008.

3 Publications co-auteur

- 6/ K. Lucas, J. S. Gordon, W. Conley, **M. Saied**, S. Warrick, M. Pochkowski, M. D. Smith, C. West, F. Kalk, and J. P. Kuijten, "Optical issues of thin organic pellicles in 45-nm and 32-nm immersion lithography," *Proceedings of SPIE*, vol. 6349, 2006.
- 7/ Y. Trouiller, V. Farys, A. Borjon, J. Belledent, C. Couderc, F. Sundermann, J.-C. Urbani, Y. Rody, C. Gardin, J. Planchot, W. Conley, P.-J. Goirand, S. Warrick, F. Robert, G. Kerrien, F. Vautrin, B. Wilkinson, **M. Saied**, E. Yesilada, P. Montgomery, L. LeCam, and C. Martinelli, "32-nm SOC printing with double patterning, regular design, and 1.2 NA immersion scanner," *Proceedings of SPIE*, vol. 6520, 2007.

- 8/ F. Sundermann, Y. Trouiller, J.-C. Urbani, C. Couderc, J. Belledent, A. Borjon, F. Foussadier, C. Gardin, L. LeCam, Y. Rody, **M. Saied**, E. Yesilada, C. Martinelli, B. Wilkinson, F. Vautrin, N. Morgana, F. Robert, P. Montgomery, G. Kerrien, J. Planchot, V. Farys, and J.-L. D. Maria, "OPC structures for maskshops qualification for the CMOS65nm and CMOS45nm nodes," *Proceedings of SPIE*, vol. 6533, 2007.
- 9/ J. C. Urbani, J. D. Chapon, J. Belledent, A. Borjon, C. Couderc, J. L. Di-Maria, V. Farys, F. Foussadier, C. Gardin, G. Kerrien, L. LeCam, C. Martinelli, P. Montgomery, N. Morgana, J. Planchot, F. Robert, Y. Rody, **M. Saied**, F. Sundermann, Y. Trouiller, F. Vautrin, B. Wilkinson, and E. Yesilada, "Characterization of inverse SRAF for active layer trenches on 45-nm node," *Proceedings of SPIE*, vol. 6607, 2007.
- 10/ A. Borjon, J. Belledent, Y. Trouiller, C. Gardin, C. Couderc, Y. Rody, F. Sundermann, J.-C. Urbani, F. Foussadier, J. Planchot, E. Yesilada, P. Montgomery, B. Willkinson, **M. Saied**, C. Martinelli, G. Kerrien, L. LeCam, F. Vautrin, F. Robert, and P. Schiavone, "Analysis of the diffraction pattern for optimal assist feature placement," *Microelectronic Engineering*, vol. 84, no. 5-8, pp. 741 – 745, 2007, proceedings of the 32nd International Conference on Micro- and Nano-Engineering.
- 11/ C. Gardin, J. Belledent, Y. Trouiller, A. Borjon, C. Couderc, F. Foussadier, E. Yesilada, J.-C. Urbani, F. Sundermann, Y. Rody, **M. Saied**, J. Planchot, and F. Robert, "Etch modeling for model-based optical proximity correction for 65 nm node," *Microelectronic Engineering*, vol. 84, no. 5-8, pp. 770–773, 2007, proceedings of the 32nd International Conference on Micro- and Nano-Engineering.
- 12/ F. Sundermann, F. Foussadier, T. Takigawa, J. Wiley, A. Vacca, L. Depre, G. Chen, S. Bai, J.-S. Wang, R. Howell, V. Arnoux, K. Hayano, S. Narukawa, S. Kawashima, H. Mohri, N. Hayashi, H. Miyashita, Y. Trouiller, F. Robert, F. Vautrin, G. Kerrien, J. Planchot, C. Martinelli, J. L. Di-Maria, V. Farys, B. Vandewalle, L. Perraud, J. C. Le-Denmat, A. Villaret, C. Gardin, E. Yesilada, and **M. Saied**, "Model-based mask verification on critical 45nm logic masks," *Proceedings of SPIE*, vol. 7028, 2008.
- 13/ V. Farys, F. Robert, C. Martinelli, Y. Trouiller, F. Sundermann, C. Gardin, J. Planchot, G. Kerrien, F. Vautrin, **M. Saied**, E. Yesilada, F. Foussadier, A. Villaret, L. Perraud, B. Vandewalle, J. C. Le-Denmat, and M. K. Top, "Study of SRAF placement for contact at 45 nm and 32 nm node," *Proceedings of SPIE*, vol. 6924, 2008.
- 14/ M. K. Top, Y. Trouiller, V. Farys, D. Fuard, E. Yesilada, C. Martinelli, **M. Saied**, F. Foussadier, and P. Schiavone, "Outliers detection by fuzzy classification method for model building," *Proceedings of SPIE*, vol. 7272, 2009.
- 15/ F. Foussadier, E. Yesilada, J. C. Le-Denmat, Y. Trouiller, V. Farys, F. Robert, G. Kerrien, C. Gardin, L. Perraud, F. Vautrin, A. Villaret, C. Martinelli, J. Planchot, J. L. Di-Maria, **M. Saied**, and M. K. Top, , "Impact of modelisation pixel size on OPC consistency," *Proceedings of SPIE*, vol. 7274, 2009.
- 16/ D. Rio, C. Constancias, **M. Saied**, B. Icard, and L. Pain, "Study on line edge roughness for electron beam acceleration voltages from 50 to 5 kv," *Journal of Vacuum Science & Technology*, vol. 27, no. 6, pp. 2512–2517, 2009.

Bibliographie

- [1] A. Kahng and Y. Pati, "Subwavelength lithography and its potential impact on design and EDA," in *Design Automation Conference, 1999. Proceedings. 36th*, 1999, pp. 799–804. ix
- [2] T. Terasawa, "Embedded tutorial : subwavelength lithography," in *Proceedings of the 2000 Asia and South Pacific Design Automation Conference*, ser. ASP-DAC '00. ACM, 2000, pp. 295–300. ix
- [3] A. B. Kahng and Y. C. Pati, "Subwavelength optical lithography : challenges and impact on physical design," in *Proceedings of the 1999 international symposium on Physical design*, ser. ISPD '99. ACM, 1999, pp. 112–119. ix, x
- [4] T. Michael, U. Wilhelm, G. Aksel, and K. Winfried, "Semiconductor fabrication : Pushing deep ultraviolet lithography to its limits," *Nature Photonics*, vol. 1, no. 11, pp. 629–631, 2007. ix, x
- [5] A. K. Wong, "Resolution enhancement techniques in optical lithography," *SPIE Press*, 2001. x, 5
- [6] B. Streefkerk, J. Baselmans, W. G. van Ansem, J. Mulkens, C. Hoogendam, M. Hoogendorp, D. Flagello, H. Sewell, and P. Graeupner, "Extending optical lithography with immersion," *Proceedings of SPIE*, vol. 5377, pp. 285–305, 2004. x
- [7] L. W. Liebmann, "Layout impact of resolution enhancement techniques : impediment or opportunity?" in *Proceedings of the 2003 international symposium on Physical design*, ser. ISPD '03. New York, NY, USA : ACM, 2003, pp. 110–117. x
- [8] E. J. Walker, "Reduction of photoresist standing-wave effects by post-exposure bake," *IEEE Transactions on Electron Devices*, vol. ED, no. 22, pp. 464–466, 1975. 8, 24
- [9] Y.-C. Lin, A. J. Purdes, S. A. Saller, and W. R. Hunter, "Improvement of linewidth control with antireflective coating in optical lithography," *Journal of Applied Physics*, vol. 55, no. 4, pp. 1110–1115, 1984. 8
- [10] D. Hopper, K. Ghandehari, and M. V. Ngo, "Structure and method for reducing standing waves in a photoresist," US PATENT 7,070,911 B1, Jul 2006, App. No. 10/350,472 ; Fil. date : 23 Jan 2003. [Online]. Available : <http://www.google.com/patents?id=Nlh6AAAABAJ> 8
- [11] L. Rayleigh, "Investigations in optics with special reference to the spectroscope," *Philosophical Magazine Series 5*, vol. 8, no. 49, pp. 261–274, 1879. 9
- [12] G. S. Chua, C. J. Tay, C. Quan, and Q. Lin, "Improvement of Rayleigh criterion with duty ratio characterization for subwavelength lithography," *Journal of Vacuum Science Technology B : Microelectronics and Nanometer Structures*, vol. 22, no. 2, pp. 801–808, mar 2004. 9
- [13] D. Melville *et al.*, "Demonstrating the benefits of source-mask optimization and enabling technologies through experiment and simulations," *Proceedings of SPIE*, vol. 7640, pp. 1–18, 2010. 12
- [14] E. Hecht, *Optics, Fourth Edition*, 4th ed. Addison Wesley, 2002, ch. 9, pp. 386–387. 12
- [15] J. E. Webb, R. L. Maier, D. S. Goodman, and W. Conley, "Hyper-numerical aperture imaging challenges for 193 nm," *Proceedings of SPIE*, vol. 5754, pp. 69–79, 2005. 14

- [16] B. W. Smith, L. Zavyalova, and A. Estroff, "Benefiting from polarization — effects on high-NA imaging," *Proceedings of SPIE*, vol. 5377, pp. 68–79, 2004. 14
- [17] D. G. Flagello, S. Hansen, B. Geh, and M. Totzeck, "Challenges with Hyper-NA ($NA > 1.0$) Polarized Light Lithography for Sub $\lambda/4$ resolution," *Proceedings of SPIE*, vol. 5754, pp. 53–68, 2005. 14
- [18] W. de Boeij, G. Swinkels, N. L. Masson, A. Koolen, H. van Greevenbroek, M. Klaassen, M. van de Kerkhof, K. van Ingen Schenau, L. de Winter, M. Wehrens, S. Hansen, and C. Wagner, "Enabling the 45nm node by hyper-NA polarized lithography," *Proceedings of SPIE*, vol. 6154, 2006. 14
- [19] E. van Setten, W. de Boeij, B. Hepp, N. le Masson, G. Swinkels, and M. van de Kerkhof, "Pushing the boundary : low-k1 extension by polarized illumination," *Proceedings of SPIE*, vol. 6520, 2007. 14
- [20] K. Ozawa, B. Thunnakart, T. Kaneguchi, I. Mita, A. Someya, T. Nakashima, H. Nishinaga, Y. Mizuno, T. Matsuyama, and M. Hamatani, "Effect of Azimuthally Polarized Illumination Imaging on Device Patterns Beyond the 45-nm-Node," *Proceedings of SPIE*, vol. 6154, 2006. 14
- [21] M. D. Levenson, N. S. Viswanathan, and R. A. Simpson, "Improving resolution in photolithography with a phase-shifting mask," *IEEE Transactions on Electron Devices*, vol. ED-29, pp. 1828–1836, 1982. 14
- [22] Z. Cui, *Nanofabrication, Principles, Capabilities and Limits*. Springer Science + Business Media, LLC, 2008, ch. 2, pp. 34–38. 16
- [23] W. an Loong, C. wei Chen, Y. hui Chang, C. ming Lin, Z. Cui, and C. an Lung, " $tisixoy$ as an absorptive shifter for embedded phase-shifting mask in 248 nm and the modification of R-T method for the determination of shifter's n and k," *Microelectronic Engineering*, vol. 41/42, pp. 125–128, 1998. 16
- [24] C. ming Lin and W. an Loong, " $tisixny$ and $tisixoynz$ as embedded materials for attenuated phase-shifting mask in 193 nm," *Microelectronic Engineering*, vol. 46, pp. 93–96, 1999. 16
- [25] G. Dao, G. Liu, R. Hainsey, J. Farnsworth, Y. Tokoro, S. Kawada, T. Yamamoto, N. Yoshioka, A. Chiba, and H. Morimoto, "248 nm DUV MoSiON Embedded Phase-Shifting Mask for 0.25 Micrometer Lithography," *Bacus Photomask News*, vol. 11, no. 8, pp. 1–9, 1995. 16
- [26] K. K. Shih and D. B. Dove, "Thin film materials for the preparation of attenuating phase shift masks," *Journal of Vacuum Science and Technology*, vol. B12, pp. 32–36, 1994. 16
- [27] J. Ushioda, Y. Seki, K. Maeda, T. Ohfuji, and H. Tanabe, "Chromium fluoride attenuated phase-shifting mask for argon fluoride excimer laser lithography," *Journal of Applied Physics*, vol. 35, pp. 6356–6359, 1996. 16
- [28] P. F. Carcia, R. H. French, M. H. Reilly, M. F. Lemon, and D. J. Jones, "Optical superlattices — a strategy for designing phase-shift masks for photolithography at 248 and 193 nm : Application to AlN/CrN," *Applied Physics Letters*, vol. 70, no. 18, pp. 2371–2373, 1997. 16
- [29] P. F. Carcia, R. H. French, C. Torardi, G. Reynolds, and L. Dieu, "Thin films for phase-shift masks," *Vacuum and Thin Films*, pp. 14–21, September 1999. 16
- [30] A. Bourov, L. Litt, and L. Zavyalova, "Impact of flare on cd variation for 248nm and 193nm lithography systems," *Proceedings of SPIE*, vol. 4346, 2001. 17, 25, 39
- [31] K. Sato, K. Chiba, K. Hayasaki, and K. Kawano, "Mid-range flare measurement and modeling," *Proceedings of SPIE*, vol. 5754, 2005. 17, 25, 39
- [32] Y. J. Yun, J. Kim, and K. Kim, "Flare effect of different shape of illumination apertures in 193-nm optical lithography system," *Proceedings of SPIE*, vol. 6520, 2007. 17, 25, 39
- [33] T. Brunner, Z. Chen, K. Chen, and S. Scheer, "A new long range proximity effect in chemically amplified photoresist processes : "chemical flare"," *Proceedings of SPIE*, vol. 5753, 2005. 17, 25, 39

- [34] D. Sullivan, "Chemical flare long range proximity effects in photomask manufacturing with chemically amplified resists," *Proceedings of SPIE*, vol. 6349, 2006. 17, 25, 39
- [35] M. Drapeau, V. Wiaux, E. Hendrickx, S. Verhaegen, and T. Machida, "Double patterning design split implementation and validation for the 32nm node," *Proceedings of SPIE*, vol. 6521, 2007. 20
- [36] M. Hori, T. Nagai, A. Nakamura, T. Abe, G. Wakamatsu, T. Kakizawa, Y. Anno, M. Sugiura, S. Kusumoto, Y. Yamaguchi, and T. Shimokawa, "Sub-40nm half-pitch double patterning with resist freezing process," *Proceedings of SPIE*, vol. 6923, 2008. 20
- [37] C. Fonseca, M. Somervell, S. Scheer, W. Printza, K. Nafusb, S. Hatakeyamab, Y. Kuwahara, T. Niwa, S. Bernard, and R. Gronheid, "Advances and challenges in dual-tone development process optimization," *Proceedings of SPIE*, vol. 7274, 2009. 21
- [38] W. Shiu, H. Liu, J. Wu, T. Tseng, C. T. Liao, C. Liao, J. Liu, and T. Wang, "Advanced self-aligned double patterning development for sub-30-nm DRAM manufacturing," *Proceedings of SPIE*, vol. 7274, 2009. 22
- [39] Y. C. C. Bencher, H. Dai, "Gridded design rule scaling : taking the cpu toward the 16nm node," *Proceedings of SPIE*, vol. 7274, 2009. 22
- [40] T. K. L. A. Carlson, "Low-variability negative and iterative spacer processes for sub-30-nm lines and holes," *J. Micro/Nanolith. MEMS MOEMS*, vol. 8, no. 1, 2009. 22
- [41] H. H. Hopkins, "On the diffraction theory of optical images," *Proc. Roy. Soc.*, vol. A, no. 217, pp. 408–432, 1953. 24, 29
- [42] T. V. Pistor, A. R. Neureuther, and R. J. Socha, "Modeling oblique incidence effects in photomasks," *Proceedings of SPIE*, vol. 4000, no. 228, 2000. 24
- [43] P. Liu, Y. Cao, L. Chen, G. Chen, M. Feng, J. Jiang, H. yu Liu, S. Suh, S.-W. Lee, and S. Lee, "Fast and accurate 3D mask model for full-chip OPC and verification," *Proceedings of SPIE*, vol. 6520, p. 65200R, 2007. 24, 99
- [44] Y. S. Sohn, D. S. Sohn, and H. K. Oh, "Effect of temperature variation during post exposure bake on 193 nm chemically amplified resist simulation," *Microprocesses and Nanotechnology Conference, 2000 International*, pp. 100–101, July 2000. 25
- [45] F. H. Dill, "Optical lithography," *IEEE Transactions on Electron Devices*, vol. ED-22, no. 7, pp. 440–444, 1975. 25
- [46] F. H. Dill, W. P. Hornberger, P. S. Hauge, and J. M. Shaw, "Characterization of positive photoresist," *IEEE Transactions on Electron Devices*, vol. ED-22(7), no. 445 - 452, 1975. 25, 41
- [47] K. L. Konnerth and F. H. Dill, "In-situ measurement of dielectric thickness during etching or developing processes," *IEEE Transactions on Electron Devices*, vol. ED-22(7), no. 452 - 456, 1975. 25
- [48] F. H. Dill, A. R. Neureuther, J. A. Tuttle, and E. J. Walker, "Modeling projection printing of positive photoresists," *IEEE Transactions on Electron Devices*, vol. ED-22(7), no. 456 - 464, 1975. 25
- [49] C. A. Mack, "Advanced topics in lithography modeling," *Advances in Resist Technology and Processing III*, vol. 631, no. 276–285, 1986. 25
- [50] —, "Development of positive photoresists," *Journal of the Electrochemical Society, Solid-state Science and Technology*, vol. 134, no. 1, pp. 148–152, 1987. 25
- [51] —, "New kinetic model for resist dissolution," *Journal of the Electrochemical Society*, vol. 139, no. 4, p. 35, 1992. 25
- [52] M. A. Toukhy, S. G. Hansen, R. J. Hurditch, and C. A. Mack, "Experimental investigation of a novel dissolution model," *Advances in Resist Technology and Processing IX*, vol. 1672, pp. 286–298, 1992. 25

- [53] G. Arthur, C. A. Mack, and B. Martin, "A new development model for lithography simulation," *Olin Microlithography Seminar*, vol. Interface'97, pp. 55–66, November 1997. 25
- [54] C. A. Mack and G. Arthur, "Notch model for photoresist dissolution," *Electrochemical and Solid-State Letters*, vol. 1, no. 2, pp. 86–88, 1998. 25
- [55] S. Robertson and E. Pavelchek, "An improved notch model for resist dissolution in lithography simulation," *Proceedings of SPIE*, vol. 4345, 2001. 25
- [56] S. Postnikov *et al.*, "Critical dimension control in optical lithography," *Microelectronic Engineering*, vol. 69, pp. 452–458, 2003. 26
- [57] Y. Granik and N. B. Cobb, "New process models for OPC at sub-90-nm nodes," *Proceedings of SPIE*, vol. 5040, pp. 1166–1175, 2003. 27
- [58] W. Tawfic, M. Al-Imam, and G. E. Bailey, "Safe interpolation distance for VT5 resist model," *Proceedings of SPIE*, vol. 6730, 2007. 27
- [59] Y. Granik *et al.*, "Application of CM0 resist model to OPC and verification," *Proceedings SPIE*, vol. 6154, 2006. 28
- [60] —, "Towards standard process models for OPC," *Proceedings of SPIE*, vol. 6520, 2007. 28
- [61] F. Sundermann *et al.*, "OPC Structures for Maskshops Qualification for the CMOS65nm and CMOS45nm nodes," *Proceedings of SPIE*, vol. 6533, 2007. 39
- [62] D. Sohn *et al.*, "Analysis of the relation between exposure parameters and critical dimension by Response Surface Model," *Proceedings of SPIE*, vol. 4345, p. 973, 2001. 41, 57
- [63] S. K. Kim *et al.*, "Simulation parameter effects on critical dimension and sensitivity of 193 nm chemically amplified resist," *Microprocesses and Nanotechnology Conference*, vol. 2001 International, pp. 226–227, October 2001. 41
- [64] J. Schacht, K. Herold, and R. Zimmermann, "Calibration of OPC models for multiple focus conditions," *Proceedings of SPIE*, vol. 5377, 2004. 54
- [65] K. Elian *et al.*, "Resist parameter sensitivity analysis based on calibrated simulation for understanding resist limitations in next generation lithography," *Journal of Microelectronic Engineering*, vol. 83, pp. 1082–1086, 2006. 57, 159
- [66] B. S. Ward, "Improving HyperNA OPC using targeted measurements for model parameter extraction," *Proceedings of SPIE*, vol. 6730, 2007. 57, 159
- [67] —, "Evaluation of OPC test patterns using parameter sensitivity," *Proceedings of SPIE*, vol. 6924, 2008. 57, 159, 160
- [68] L. S. I. Melvin, D. N. Zhang, K. J. Strozewski, and S. Wolfer, "The use of the manufacturing sensitivity model forms to comprehend layout manufacturing robustness for use during device design," in *Proceedings of the 7th International Symposium on Quality Electronic Design*, ser. ISQED '06. Washington, DC, USA : IEEE Computer Society, 2006, pp. 485–490. 57, 160, 161
- [69] D. Zhang and L. Melvin, "Model based lithography verification using the new manufacturing sensitivity model," *Proceedings of SPIE*, vol. 6349, 2006. 57
- [70] L. S. Melvin, J. P. Shiely, and Q. Yan, "Use of optical defocus components to investigate and improve pattern spatial frequency characteristics for more robust layouts," *J. Vac. Sci. Technol. B*, vol. 23, no. 6, 2005. 57, 162
- [71] L. S. Melvin, Q. Yan, and W. F. Kielhorn, "Lithographic manufacturing robustness analysis for as drawn patterns," *Elsevier, Microelectronic Engineering*, vol. 83, no. pp. 1003–1007, 2006. 57, 162
- [72] J. E. Jung *et al.*, "Sem contour-based model OPC calibrated with optically sensitive patterns," *Proceedings of SPIE*, vol. 6925, 2008. 57, 162, 163

- [73] A. Abdo *et al.*, “The effect of calibration feature weighting on OPC optical and resist models – investigating the influence on model coefficients and on the overall model fitting,” *Proceedings of SPIE*, vol. 5992, 2005. 57
- [74] —, “The effect of OPC optical and resist model parameters on the model accuracy, run time, and stability,” *Proceedings of SPIE*, vol. 6349 (2), no. pp. 634923.1-634923.12, 2006. 57
- [75] M. Saied *et al.*, “Sensitivity of a variable threshold model towards process and modeling parameters,” *Proceedings of SPIE*, vol. 6349, 2006. 57
- [76] K. Patterson *et al.*, “Improving model - based OPC performance for the 65nm node through calibration set optimization,” *Proceedings of SPIE*, vol. 5756, p. 294, 2005. 64
- [77] A. Khoh *et al.*, “Maximizing test pattern coverage for OPC model build,” *Proceedings of SPIE*, vol. 6154, 2006. 64
- [78] H.-Y. Liu, Q. Zhao, J. F. Chen, J. Jiang, R. Socha, E. V. Setten, A. Engelen, J. Meessen, M. M. Crouse, M. Feng, W. Shao, H. Cao, Y. Cao, L. V. Look, J. Bekaert, G. Vandenbergh, and J. Finders, “Separable OPC models for computational lithography,” *Proceedings of SPIE*, vol. 7028, no. 1, p. 70280X, 2008. 79, 111
- [79] A. Taflove, *Computational electrodynamics : The Finite-Difference Time-Domain Method*. Artech House Publishers, 1995. 81
- [80] A. Taflove and S. C. Hagness, *Computational Electrodynamics : The Finite-Difference Time-Domain Method*, 3rd ed. Artech House Publishers, June 2005. 81, 174
- [81] K. Kunz and R. Luebbers, *The finite difference time domain method for electromagnetics*, 1st ed. CRC-Press, July 1993. 81, 179
- [82] M. G. Moharam and T. K. Gaylord, “Rigorous coupled-wave analysis of planar grating diffraction,” *Journal of the Optical Society of America*, vol. 71, no. 7, pp. 811–818, 1981. 81
- [83] —, “Rigorous coupled-wave analysis of grating diffraction - E-mode polarization losses,” *Journal of the Optical Society of America*, vol. 73, no. 4, pp. 451–455, 1983. 81
- [84] —, “Diffraction analysis of dielectric surface-relief gratings,” *Journal of the Optical Society of America*, vol. 72, no. 10, pp. 1385–1392, 1982. 81
- [85] —, “Three-dimensional vector coupled-wave analysis of planar-grating diffraction,” *Journal of the Optical Society of America*, vol. 73, no. 9, pp. 1105–1112, 1983. 81
- [86] K. Rokushima and J. Yamakita, “Analysis of anisotropic dielectric gratings,” *Journal of the Optical Society of America*, vol. 73, no. 7, pp. 901–908, 1983. 81
- [87] C. B. Burckhardt, “Diffraction of a plane wave at sinusoidally stratified dielectric grating,” *Journal of the Optical Society of America*, vol. 56, no. 11, pp. 1502–1509, 1966. 81
- [88] F. G. Kaspar, “Diffraction by thick, periodically stratified gratings with complex dielectric constant,” *Journal of the Optical Society of America*, vol. 63, no. 1, pp. 37–45, 1973. 81
- [89] K. Knop, “Rigorous diffraction theory for transmission phase gratings with deep rectangular grooves,” *Journal of the Optical Society of America*, vol. 68, no. 3, pp. 1206–1210, 1978. 81
- [90] P. Lalanne and G. M. Morris, “Highly improved convergence of the coupled-wave method for TM polarization,” *Journal of the Optical Society of America A*, vol. 13, no. 4, pp. 779–784, Apr 1996. 82
- [91] G. Granet and B. Guizal, “Efficient implementation of the coupled-wave method for metallic lamellar gratings in TM polarization,” *Journal of the Optical Society of America A*, vol. 13, no. 5, pp. 1019–1023, May 1996. 82
- [92] E. Silberstein, P. Lalanne, J.-P. Hugonin, and Q. Cao, “Use of grating theories in integrated optics,” *Journal of the Optical Society of America A*, vol. 18, no. 11, pp. 2865–2875, Nov 2001. 82

- [93] M. Palamaru and P. Lalanne, "Photonic crystal waveguides : Out-of-plane losses and adiabatic modal conversion," *Applied Physics Letters*, vol. 78, no. 11, pp. 1466–1468, 2001. 82
- [94] H. Benisty and P. Lalanne, "Out-of-plane losses of two-dimensional photonic crystal waveguides : Electromagnetic analysis," *Journal of Applied Physics*, vol. 89, no. 2, p. 1512–1514, 2001. 82
- [95] P. Lalanne, "Electromagnetic analysis of photonic crystal waveguides operating above the light-cone," *IEEE Journal of Quantum Electronics*, vol. 38, no. 7, pp. 800–804, 2002. 82
- [96] C. Sauvan, P. Lalanne, and J. P. Hugonin, "Slow-wave effect and mode-profile matching in photonic crystal microcavities," *Physical Review B*, vol. 71, p. 165118, Apr 2005. 82
- [97] C. Sauvan, G. Lecamp, P. Lalanne, and J. Hugonin, "Modal-reflectivity enhancement by geometry tuning in photonic crystal microcavities," *Optics Express*, vol. 13, no. 1, pp. 245–255, Jan 2005. 82
- [98] A. David, H. Benisty, and C. Weisbuch, "Fast factorization rule and plane-wave expansion method for two-dimensional photonic crystals with arbitrary hole-shape," *Physical Review B*, vol. 73, p. 075107, Feb 2006. 82
- [99] B. Chernov, M. Nevière, and E. Popov, "Fast fourier factorization method applied to modal analysis of slanted lamellar diffraction gratings in conical mountings," *Optics Communications*, vol. 194, no. 4–6, pp. 289–297, 2001. 82
- [100] H. Bremmer, "The W. K. B. approximation as the first term of a geometrical-optical series," *Communications on Pure and Applied Mathematics*, vol. 4, p. 105, 1951. 82
- [101] D. S. Jones, "The theory of electromagnetism," *New York : Macmillan Encyclopedia of Physics*, vol. 47, pp. 438–450, 1964. 82
- [102] C. J. Bouwcamp, "Diffraction theory," *Reports on progress in physics*, vol. 17, pp. 35–100, 1954. 82
- [103] J. Ripoll, V. Ntziachristos, R. Carminati, and M. Nieto-Vesperinas, "Kirchhoff approximation for diffusive waves," *The American Physical Society, Physical Review E*, vol. 64, no. 051917, 2001. 82
- [104] J. M. Soto-Crespo and M. Nieto-Vesperinas, "Electromagnetic scattering from very rough random surfaces and deep reflection gratings," *Journal of Optical Society of America*, vol. 6, no. 3, pp. 367–384, 1989. 82
- [105] N. Pinel and C. Bourlier, "Scattering from very rough layers under the geometric optics approximation : further investigation," *Journal of Optical Society of America*, vol. 256, no. 6, pp. 1293–1306, 2008. 82
- [106] A. V. Moroz and Z. I. Feizulin, "Possible extension of kirchhoff's approximation in problems of wave diffraction by transparent inhomogeneous objects," *Radiophysics and Quantum Electronics*, vol. 26, no. 7, 1983. 82
- [107] Z. L. Wang, F. Xu, Y. Q. Jin, and H. Ogura, "A double kirchhoff approximation for very rough surface scattering using the stochastic functional approach," *Radio Science*, vol. 40, 2005. 82
- [108] M. S. Yeung and E. Barouch, "Limitation of the kirchhoff boundary conditions for aerial image simulation in 157-nm optical lithography," *IEEE Electron Device Letters*, vol. 21, no. 9, 2000. 83
- [109] M. Saied *et al.*, "3d mask modeling with oblique incidence and mask corner rounding effects for the 32nm node," *Proceedings of SPIE*, vol. 6730, 2007. 84, 140
- [110] S. L. Sobolev, "L'algorithme de schwarz dans la théorie de l'élasticité," *Comptes rendus doklady de l'académie des sciences de l'URSS IV*, vol. XIII, no. 6, pp. 243–246, 1936. 99
- [111] I. Babuska, "The schwarz algorithm in partial differential equations of mathematical physics," *Czech. Math. J.*, vol. 83, no. 8, 1958. 99

- [112] K. Miller, "Numerical analogs of the Schwarz alternating procedure," *Numerische Mathematik* 7, vol. 7, no. 2, April 1965. 99
- [113] A. M. Matsokin and S. V. Nepomnyaschikh, "A schwarz alternating method in a subspace," *Soviet Mathematics* 29, vol. No. 10, no. 78-84, 1985. 99
- [114] M. Dryja and O. B. Widlund, "An additive variant of the schwarz alternating method for the case of many subregions," *Technical Report 339, Department of Computer Science, Courant Institute, New York University*, 1987. 99
- [115] P.-L. Lions, "On the schwarz alternating method. i," *First international symposium on domain decomposition methods for partial differential equations. SIAM*, 1988. 99
- [116] —, "On the schwarz alternating method. ii," *Second international symposium on domain decomposition methods for partial differential equations. SIAM*, 1989. 99
- [117] —, "On the schwarz alternating method. iii," *A variant for nonoverlapping subdomains. Domain decomposition methods for partial differential equations. SIAM*, 1990. 99
- [118] K. Adam and A. R. Neureuther, "Simplified models for edge transitions in rigorous mask modeling," *Proceedings of SPIE*, vol. 4346, pp. 331–344, 2001. 99
- [119] K. Adam, "Domain decomposition methods for the electromagnetic simulation of scattering from three-dimensional structures with applications in lithography," Ph.D. dissertation, University of California, Berkeley, 2001. 99, 100
- [120] —, "Modeling of electromagnetic effects from mask topography at full-chip scale," *Proceedings of SPIE*, vol. 5754, pp. 498–505, 2005. 99
- [121] J. Tirapu-Azpiroz, P. Burchard, and E. Yablonovitch, "Boundary layer model to account for thick mask effects in photolithography," *Proceedings of SPIE*, vol. 5040, no. 1611, 2003. 100, 111, 112, 124
- [122] T. L. C. Z. Y. C. J. Fung Chen, Hua-Yu Liu and R. Socha, "Development of a Computational Lithography Roadmap," *Proceedings of SPIE*, vol. 6924, 2008. 100
- [123] K. Adam, "Sum of Coherent Systems (SOCS) approximation based on object information," US PATENT 2007/0218176 A1, Sep 2007, App. No. 11/715,667 ; Fil. date : 8 Mar 2007. [Online]. Available : <http://www.freepatentsonline.com/y2011/0119643.html> 100
- [124] N. Cobb, "Sum of Coherent Systems Decomposition by SVD," Ph.D. dissertation, University of California at Berkeley, 1995. 100
- [125] C. Zuniga and E. Tejnil, "Heuristics for truncating the number of optical kernels in hopkins image calculations for model-based OPC treatment," *Proceedings of SPIE*, vol. 6520, 2007. 100
- [126] S. Kazuya, I. Masamitsu, and S. Takashi, "HP45 lithography in consideration of the mask 3D effect," *Proceedings of SPIE*, vol. 6283, no. 2, p. 62831F, 2006. 111
- [127] J. Tirapu-Azpiroz, "Analysis and modeling of photomask near-fields in sub-wavelength deep ultraviolet lithography with optical proximity corrections," Ph.D. dissertation, University of California, Los Angeles, 2004. 111, 112
- [128] J. Tirapu-Azpiroz and E. Yablonovitch, "Incorporating mask topography edge diffraction in photolithography simulations," *J. Opt. Soc. Am. A*, vol. 23, no. 4, pp. 821–828, April 2006. 111, 112
- [129] M. C. Lam and K. Adam, "Understanding the impact of rigorous mask effects in the presence of resist models," *Proceedings of SPIE*, vol. 6520, no. 141, 2007. 111, 118
- [130] M. Saied and E. Yesilada, "Simulation of the image projected by a mask," US PATENT 2011/0022219 A1, Jan 2011, App. No. 12/839,817 ; Fil. date : 20 Jul 2010. [Online]. Available : <http://www.freepatentsonline.com/y2011/0022219.html> 135
- [131] M. D. Jo Finders and S. Hsu, "Double Patterning Lithography : The Bridge Between low k1 ArF and EUV," *Microolithography World*, 2008. 139

- [132] M. Saied *et al.*, “Three-dimensional mask effects and source polarization impact on OPC model accuracy and process window,” *Proceedings of SPIE*, vol. 6520, 2007. 140
- [133] D. Roy, P. K. Basu, P. Raghunathan, and S. V. Eswaran, “DNQ–novolac photoresists revisited : ^1H and ^{13}C NMR evidence for a novel photoreaction mechanism,” *Magnetic Resonance in Chemistry*, vol. 41, pp. 84–90, 2003. 147
- [134] Z. Cui, *Nanofabrication, Principles, Capabilities and Limits*. Springer Science + Business Media, LLC, 2008, ch. 2, pp. 44–53. 147

Résumé

À l'heure actuelle, les modèles photochimiques utilisés dans la correction des effets de proximité optique (OPC) en photolithographie sont devenus complexes et moins physiques afin de permettre de capturer rapidement le maximum d'effets optiques et chimiques. La question de la stabilité de tels modèles purement empiriques est devenue d'actualité. Dans ce mémoire, nous avons étudié la stabilité des modèles photochimiques actuels en examinant les différentes causes d'instabilité vis-à-vis des paramètres du procédé. Dans la suite, nous avons développé une méthode perturbative permettant d'évaluer le critère de la stabilité. L'obtention de modèles simples et stables nous conduit à séparer les effets optiques des autres effets chimiques. De plus, les approximations utilisées dans la modélisation des systèmes optiques opérant à grande ouverture numérique entraînent des erreurs résiduelles pouvant dégrader la précision et la stabilité des modèles OPC. Ainsi, nous nous sommes intéressés à étudier les limites de validité de l'approximation de Kirchhoff, méthode qui, jusqu'à présent, est la plus utilisée dans la modélisation du champ proche d'un masque. D'autres méthodes semi-rigoureuses, permettant de modéliser les effets topographiques, ont été également évaluées. Ces méthodes approchées permettent de gagner en précision mais dégradent le temps de calcul. Nous avons ainsi proposé différentes façons de corriger les effets topographiques du masque, tout en gardant l'approximation de Kirchhoff dans la modélisation de la partie optique. Parmi les méthodes proposées, nous exploitons celle permettant de réduire les erreurs liées aux effets topographiques du masque par l'intermédiaire d'un second modèle empirique. Nous montrons que pour garantir une précision adéquate, il est nécessaire d'augmenter la complexité du modèle en rajoutant des termes additionnels. Enfin, pour garantir la stabilité numérique du modèle empirique, nous introduisons une nouvelle méthode approchée hybride rapide et précise, la méthode des multi-niveaux, permettant d'inclure les effets topographiques par décomposition multi-niveaux du masque fin et discuterons ses avantages et ses limites.

Mots-clé : photolithographie, effets de proximité optique, OPC, modélisation, précision, stabilité, résine, système optique, 193 nm, masque, champ proche, effets topographiques, effets 3D du masque, méthode perturbative, méthode multi-niveaux, Kirchhoff, DDM, M3D.

Abstract

At present, common resist models utilized in photolithography to correct for optical proximity effects (OPC) became complex and less physical in order to capture the maximum of optical and chemical effects in shorter times. The question on the stability of such models, purely empirical, become topical. In this thesis, we study the stability of existing OPC models by examining the origins of model instability towards process parameters. Thus, we have developed a perturbative method in order to evaluate the stability criterion. However, achieving stable and simple models needs a separation between optical and other chemical effects. Besides, multiple approximations, widely utilized in the modeling of optical systems operating at high numerical aperture, lead to residual errors which can degrade OPC model accuracy and stability. Thus, we were interested to study the limits of validity of the Kirchhoff approximation, a method which, so far, is the most commonly used in mask near-field modeling. Other semi-rigorous methods for mask topography effect modeling were also evaluated. These approximate methods can improve the accuracy but degrades the run time. We then suggested different techniques to correct for mask topography effects, while keeping the Kirchhoff approximation in the modeling of the optical part. Among them, we showed that errors due to mask topography effects can be partially captured by a second empirical model. However, in order to ensure a good accuracy, it is necessary to increase the model complexity by using more additional empirical terms. Finally, in order to achieve a numerical stability of the empirical model, we introduced a new hybrid fast and accurate method, the multi-level method, which allows us to correct for mask topography effects through a multi-level decomposition of the thin mask and discussed its advantages and drawbacks.

Keywords : photolithography, optical proximity effects, OPC, modeling, accuracy, stability, resist, optical system, 193 nm, mask, near-field, topography effects, 3D mask effects, perturbative method, multi-level method, Kirchhoff, DDM, M3D.