Compact Mask Models for Optical Projection Lithography

Kompakte Maskenmodelle für die optische Projektionslithographie

Der technischen Fakultät
der Friedrich-Alexander-Universität
Erlangen-Nürnberg
zur Erlangung des Grades

Doktorin der Ingenieurwissenschaften

vorgelegt von

Viviana Agudelo, M.Sc.
aus Bogotá, Kolumbien
Als Dissertation genehmigt von
der technischen Fakultät
der Friedrich-Alexander-Universität
Erlangen-Nürnberg

Tag der mündlichen Prüfung: 19.09.2014

Vorsitzende/r des Promotionsorgans: Prof.-Dr.-Ing. Marion Merklein
Gutachter: PD Dr. Andreas Erdmann
Prof. Dr. Andreas Fröba
Acknowledgements

I would like to acknowledge all the members of the Lithography Simulation Group at the Fraunhofer Institute IISB for their valuable contributions and discussions. Specially, I would like to thank PD Dr. Andreas Erdmann, Head of the group and supervisor of this work, who introduced me to the lithography world, without his encouragement this work would not be possible. I thank the School of Advanced Optical Technologies from the Friedrich-Alexander Universität Erlangen-Nürnberg and the Fraunhofer Institute IISB, which funded my doctoral studies. Last but not least, I would like to express my deepest gratitude and extend a big thank to my family, friends and husband for their love, support and patience.
Abstract

The transfer of micro and nano patterns into a photosensitive material has a large number of technological applications. One of this techniques is known as optical lithography and is widely used in the fabrication process of integrated circuits (IC). The exposure, as one of the most important steps of a lithography process, has a critical influence on the dimension of the features in the fabricated IC. A mask contains the pattern that has to be replicated into the photosensitive material, which is coated on the top of a semiconductor wafer. A light source illuminates the mask, where diffraction phenomena occur. Then, the diffracted light is guided by means of an optical system to create a demagnified image of the mask. Modeling and simulation allow a deeper understanding of the image formation, in particular at small scales in the range of few wavelengths and below.

One of the most important aspects for the image formation is the appropriate modeling of the light diffraction from the mask. When the mask features are larger than the wavelength of light, the scalar diffraction theory (Kirchhoff approach) yields sufficiently accurate results in the computation of the diffraction spectrum. With feature sizes smaller than or comparable to the wavelength, the scalar approximation exhibits a serious limitation. It does not account for the three-dimensional mask geometry and related mask topography effects. That is why a rigorous description of the light diffraction from the mask is required.

The propagation of the light through the mask can be rigorously computed using the Maxwell’s equations. The effort to accomplish a highly accurate description of the diffracted field, introduces a huge computational expense. As a consequence, innovative modeling techniques are challenged to compromise accuracy and speed in the computation of the diffracted field, as well as in the computation of the imaging. So-called compact mask models speed up the mask diffraction spectrum and imaging computation, considering the three-dimensional mask geometry and related mask topography effects. These compact mask models introduce methods to improve the accuracy of the Kirchhoff-based imaging model. This is done by means of a systematic modification of
the scalar diffraction spectrum or the mask geometry, in order to yield similar results as the fully rigorous simulations.

In this work, three novel compact mask models are formulated. These approaches are considered in the spatial frequency domain. First, a Jones pupil function is introduced in the projector to describe amplitude, phase and polarization effects, which are introduced by the mask (pupil filtering model). Second, a correction is performed directly on the scalar diffraction spectrum, to tune the diffraction orders that are captured by the pupil of the optical projection system (spectrum correction model). Finally, an artificial neural network approach is considered. The artificial neural networks are trained using the scalar diffraction spectrum as input and the rigorous spectrum as target. The outcome of this training process is a neural network capable of reproducing a diffraction spectrum that approximates the rigorous spectrum, which is obtained from electromagnetic field simulations.

The proposed compact mask models account for and compensate mask topography-induced effects even at image planes out of focus. This allows to preserve the accuracy of the image computation in lithography simulations, at a reasonable computational cost compared to the rigorous mask model.
Zusammenfassung


Einer der wichtigsten Aspekte bei der Berechnung der Abbildung einer lithographischen Maske ist eine geeignete Modellierungsmethode zur Beschreibung der Lichtbeugung an dieser. Sind die Strukturen auf der Maske größer als die Beleuchtungswellenlänge, ermöglicht die skalare Beugungstheorie (Kirchhoff-Näherung) eine ausreichend genaue Beschreibung der Vorgänge. Bei Strukturgrößen im Bereich oder unterhalb der Belichtungswellenlänge zeigt die skalare Annäherung jedoch deutliche Einschränkungen. Es werden keine dreidimensionalen Maskengeometrien und die damit verbundenen Topographieeffekte dargestellt.


Die betrachteten Kompaktmaskenmodelle ermöglichen die Beschreibung von Topographieeffekten, welche durch die Maske hervorgerufen werden und sind in der Lage, Fehler der skalaren Modelle auch in Bildebenen außerhalb des Fokus zu kompensieren. Damit lässt sich die Genauigkeit der Bildberechnung in der Lithographie im Vergleich zu skalaren Modellen deutlich verbessern. Der erforderliche Rechenaufwand ist gegenüber rigorosen Verfahren erheblich reduziert.
# Contents

In the image provided, the table of contents includes sections and subsections as follows:

## Contents

vi

## List of Figures

ix

## List of Tables

xix

## Nomenclature

xxiii

### 1 Introduction

1.1 Thesis Contributions ................................................ 2
1.2 Thesis Structure ................................................... 4

### 2 Lithography Process in Semiconductor Technology Fabrication

2.1 Basics .................................................................. 5
  2.1.1 Wafer Preparation ............................................. 6
  2.1.2 Oxidation ....................................................... 6
  2.1.3 Photoresist Coating .......................................... 6
  2.1.4 Pre-Bake ......................................................... 6
  2.1.5 Exposure ........................................................ 7
  2.1.6 Post Exposure Bake (PEB) .................................. 8
  2.1.7 Development .................................................... 8
  2.1.8 Hard Bake ........................................................ 8
  2.1.9 Etching ............................................................ 8
  2.1.10 Stripping ........................................................ 9
  2.1.11 Doping ........................................................... 9
  2.1.12 Deposition of Materials .................................... 10
  2.1.13 Metallization .................................................. 10
2.2 Components of a Lithographic Imaging System ..................... 10
  2.2.1 Illumination System .......................................... 11
    2.2.1.1 Illumination Sources ................................... 12
    2.2.1.2 Köhler Illumination .................................... 13
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.2</td>
<td>Mask</td>
<td>14</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Projection System</td>
<td>17</td>
</tr>
<tr>
<td>2.2.4</td>
<td>Photoresist</td>
<td>18</td>
</tr>
<tr>
<td>2.3</td>
<td>Aerial Image Formation</td>
<td>19</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Scalar Diffraction Theory</td>
<td>19</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Numerical Aperture</td>
<td>23</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Partial Coherence Factor and Off-Axis Illumination</td>
<td>24</td>
</tr>
<tr>
<td>2.3.4</td>
<td>Resolution in Optical Lithography</td>
<td>27</td>
</tr>
<tr>
<td>2.3.5</td>
<td>Abbe Formulation</td>
<td>28</td>
</tr>
<tr>
<td>2.3.6</td>
<td>Hopkins Assumption</td>
<td>30</td>
</tr>
<tr>
<td>2.3.7</td>
<td>Polarization</td>
<td>32</td>
</tr>
<tr>
<td>2.4</td>
<td>Image Analysis</td>
<td>35</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Critical Dimension (CD)</td>
<td>35</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Image log-slope (ILS)</td>
<td>35</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Depth of Focus</td>
<td>36</td>
</tr>
<tr>
<td>2.4.4</td>
<td>Process Window</td>
<td>37</td>
</tr>
<tr>
<td>2.5</td>
<td>Optical Proximity Correction (OPC)</td>
<td>38</td>
</tr>
</tbody>
</table>

3 Rigorous and Compact Mask Models 40

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Rigorous Mask Models</td>
<td>41</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Finite-Difference Time-Domain (FDTD)</td>
<td>41</td>
</tr>
<tr>
<td>3.1.2</td>
<td>Waveguide Method</td>
<td>44</td>
</tr>
<tr>
<td>3.2</td>
<td>Mask-Topography Induced Effects</td>
<td>45</td>
</tr>
<tr>
<td>3.3</td>
<td>Compact Mask Models from Literature</td>
<td>47</td>
</tr>
<tr>
<td>3.3.1</td>
<td>QUASI-3D: Domain Decomposition Method (DDM)</td>
<td>48</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Boundary Layer Model (BLM)</td>
<td>49</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Pulses model (PM)</td>
<td>50</td>
</tr>
<tr>
<td>3.4</td>
<td>Proposed Mask Models</td>
<td>51</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Pupil Filtering Model (PFM)</td>
<td>51</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Spectrum Correction Model (SCM)</td>
<td>55</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Artificial Neural Network (ANN) Approach</td>
<td>57</td>
</tr>
</tbody>
</table>

4 Evaluation and Comparisons of Mask Models from Literature 63

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Simulation Conditions</td>
<td>63</td>
</tr>
<tr>
<td>4.2</td>
<td>Rigorous and Scalar Model for 2D Features (Lines and Spaces)</td>
<td>67</td>
</tr>
<tr>
<td>4.3</td>
<td>Full 3D, Quasi3D Rigorous and Scalar Models for 3D Features (Contact Holes)</td>
<td>79</td>
</tr>
<tr>
<td>4.4</td>
<td>Model Calibration for Compact Mask Models</td>
<td>88</td>
</tr>
</tbody>
</table>
CONTENTS

4.5 Boundary Layer Model (BLM) ................................................. 90
   4.5.1 Model Performance Compared to the Rigorous Model with the HA 91
   4.5.2 Model Performance Compared to the Rigorous Model without the HA ................................................. 94

4.6 Pulses Model (PM) ............................................................... 97
   4.6.1 Model Performance Compared to the Rigorous Model with the HA 98
   4.6.2 Model Performance Compared to the Rigorous Model without the HA ............................................................. 101

5 Evaluation and Comparisons of the Proposed Compact Mask Models 105
   5.1 Pupil Filtering Model (PFM) ................................................. 105
      5.1.1 Model Performance Compared to the Rigorous Model with the HA 106
      5.1.2 Model Performance Compared to the Rigorous Model without the HA ................................................. 111
   5.2 Spectrum correction model (SCM) ........................................ 115
      5.2.1 Model Performance Compared to the Rigorous Model with the HA 115
      5.2.2 Model Performance Compared to the Rigorous Model without the HA ................................................. 120
   5.3 ANN Approach as an Alternative Compact Mask Model 125
      5.3.1 Hidden Units and Hidden Layers of the ANN 126
      5.3.2 ANN Validation 126
      5.3.3 L/S with Variable Duty Cycle 127
      5.3.4 Model Performance Compared to the Rigorous Model with the HA 128
      5.3.5 Model Performance Compared to the Rigorous Model without the HA ................................................. 132
      5.3.6 Shifted Lines 137
      5.3.7 L/S with Constant Duty Cycle 138
      5.3.8 Contacts 139
      5.3.9 Line Ends 142
      5.4 Comparisons Between the Models 144

6 Conclusions 148

References 155
# List of Figures

2.1 A standard sequence of the resist processing. (a) Film deposition, (b) spin-coating, (c) pre-bake, (d) exposure, (e) post-exposure bake, (f) development, (g) hard bake, (h) etching and (i) stripping. ............................... 7

2.2 Selective doping of certain regions of the semiconductor by means of ion implantation. (top) Stripping and (bottom) ion implantation. ....................... 10

2.3 Schematic representation of an optical projection system. ......................... 11

2.4 Köhler illumination: Each point source can be treated as generating a coherent, linearly polarized plane wave of spatial frequency determined by the position of the point source relative to the optical axis. ......................... 13

2.5 Types of masks. (a) Binary mask, (b) alternating phase shift mask with etched substrate and (c) attenuated phase shift mask. From top to bottom, each plot represents the mask, the mask transmission function, the electrical field and the intensity distribution. ................................. 16

2.6 Other relevant types of PSM. (a) Tritone mask and (b) chromeless mask. 16

2.7 A schematic imaging projection system. ........................................... 17

2.8 Representation of the diffraction generated by a double slit configuration. The magnitude of the diffraction pattern squared (intensity) for a single slit is represented by the dotted line envelope, the solid line represents the diffraction pattern for two slits. ................................. 20

2.9 The diffracting aperture lies in the plane ($\xi, \eta$). Diffracted field lies in the plane ($x, y$). ....................................................... 21

2.10 Binary mask (COG) with a rectangular aperture of 700 nm, $\lambda = 193$ nm. (a) Intensity of the near field obtained with the scalar model. (b) Calculated near field intensity using a rigorous method (Waveguide). TM x-polarization component. ................................. 22

2.11 The numerical aperture (NA) with respect to the object space and the image space. ....................................................... 23

2.12 Partial coherent imaging system. ....................................................... 25
2.13 The degree of partial coherence $\sigma$ determines the tradeoff between a sharp slope and a low overshoot space. Intensity of the image corresponding to line and spaces of 10 $\mu$m linewidth, NA=0.28, $\lambda$ = 0.436 $\mu$m. Plot from [1].

2.14 Schematic of (a) On-axis illumination and (b) Off-axis illumination.

2.15 Idealized Top-down view of the lens aperture showing the diffraction pattern of a line/space space pattern when partially coherent illumination is used: (a) $\sigma = 0.25$ and (b)$\sigma = 0.5$

2.16 Common source schemes for advanced illumination used in optical lithography: (a) Circular illumination. (b) Annular illumination. (c) Quadrupole illumination. (d) Dipole illumination.

2.17 Imaging system: Source plane, object plane and image plane.

2.18 Representation of the (a) TE- and (b) TM-polarization states.

2.19 Representation of an electric field linearly polarized. The k vector point in the direction of the propagation of the wave.

2.20 Representation of an electric field linearly polarized. The k vector point in the direction of the propagation of the wave.

2.21 Polarization effects at high NAs. Plot from [2].

2.22 Critical dimension (CD) definition: The width of the image at a given threshold value.

2.23 Image Log-Slope as a metric to evaluate image quality in a lithographic process.

2.24 Process window for a periodic contact holes layout (AttPSM), 62 nm size. ArFi, CD of 62 nm, CQUAD illumination, unpolarized light. The central curve represent the process for the required CD, upper curve shows the process with 10 % deviation from the CD and bottom curve -10 % of deviation from the CD.

2.25 Optical proximity corrections (OPC) techniques. (a) Biasing, (b) Assisted features, (c) Hammer head and (d) Serif.

3.1 Yee cell for the electric and magnetic field representation.

3.2 Slicing of the mask in the Waveguide method. The blue color represents the absorber and the purple represents the substrate. The complete simulated area is divided in layers that are homogeneous in z-direction.
3.3 Diffraction efficiency effects for a standard MoSi-type AttPSM with lines/spaces for vertical incidence. NA=1.2, ArFi (n_{immersion} =1.44). Dipole illumination with y-polarization, \( \sigma = 0.8/0.96 \) and opening angle=40\(^\circ\). Analysis for the 0th (upper figure) and the 1st (bottom figure) diffraction orders. Rigorous simulation for (blue) y-polarization, (green) x-polarization and (black) Kirchhoff model. The diffraction efficiencies depend on the state of polarization. Plot from [3].

3.4 Phase effects described for a standard MoSi-type AttPSM with lines/spaces for vertical incidence. Same simulations conditions stated in Figure 3.3. Rigorous simulation for (blue) y-polarization, (green) x-polarization and (black) Kirchhoff model. The phase difference between the 0th and 1st diffraction orders depends on the polarization state. Plot from [3].

3.5 Top view of a 3D mask with a post absorber (light blue) in the center. The 3D mask is decomposed into two dimensional cuts (cut 1-6) with constant dielectric properties in x- and y-direction, respectively. Plot from [4].

3.6 The modified thin mask emulates the physical effects predicted by rigorous models.

3.7 Modification of the thin mask transmission by adding perturbations to the edges in form of delta functions.

3.8 Representation of the pupil plane.

3.9 Functional form of the Zernike polynomials for the first radial orders.

3.10 Wave deformation \( W(\rho, \varphi) \) expressed in term of Zernike polynomials corresponding to \( Z_i = 0.25 \). Plot from Andreas Erdmann, Lectures in Lithography, FAU Erlangen Graduate School of Advanced Optical Technologies SAOT, 2012.

3.11 Spectra observed from an ideal pupil without a modulating function. Spectrum corresponding to: (upper left) vertical illumination, (upper right;bottom left and right) other oblique illumination angles.

3.12 The spectrum modification is performed independently for all spectra corresponding to all relevant angles of illumination. The scalar spectrum is transformed with a calibrated function to emulate the rigorous spectrum corresponding to each angle of illumination.

3.13 Representation of an ANN.

3.14 Weighted connections within a three hidden layer ANN. The propagation of the signal is from left to right in this diagram.

3.15 Back propagation of the error \( \delta \). The error propagation is from right to left in this diagram.
4.1 Mask layouts used in this work. (left) lines and spaces (L/S), (center) contact and (right) line ends. Periodic boundary conditions are used for the computation of the diffraction spectrum. ................................. 64

4.2 Dipole illumination. $\theta = 30^\circ$, $\sigma_{in.} = 0.7$ and $\sigma_{out.} = 0.9$. $y$-polarization. With the Hopkins assumption (HA) only the spectrum corresponding to the normally incident angle of illumination is rigorously computed. Without the HA, different spectra corresponding to the representative angles of illumination are rigorously computed. ................................. 66

4.3 CQuad illumination. $\theta = 30^\circ$, $\sigma_{in.} = 0.5$ and $\sigma_{out.} = 0.7$. $xy$-polarization. ................................. 66

4.4 Schematic representation of the sampling of the source. ................................. 66

4.5 Cross sections at the best focus position for the analyzed mask models. CD=22nm. The Metrology line to extract the cross section is located at the coordinates $(x_1=\text{Pitch}, y_1=0)$; $(x_2=\text{Pitch}, y_2=0 \text{ nm})$. (a) Pitch=80 nm and (b) Pitch=330 nm. The rigorous model without the HA is represented by the blue curves. The rigorous model with the HA is represented by the green curves and the scalar model is represented by the red curves. ................................. 69

4.6 Cross sections at the best focus position for the analyzed mask models. CD=32nm. The Metrology line to extract the cross section is located at the coordinates $(x_1=\text{Pitch}, y_1=0)$; $(x_2=\text{Pitch}, y_2=0 \text{ nm})$. (a) Pitch=80 nm and (b) Pitch=330 nm. The rigorous model without the HA is represented by the blue curves. The rigorous model with the HA is represented by the green curves and the scalar model is represented by the red curves. ................................. 70

4.7 Cross sections at the best focus position for the analyzed mask models. CD=45nm. The Metrology line to extract the cross section is located at the coordinates $(x_1=\text{Pitch}, y_1=0)$; $(x_2=\text{Pitch}, y_2=0 \text{ nm})$. (a) Pitch=90 nm and (b) Pitch=330 nm. The rigorous model without the HA is represented by the blue curves. The rigorous model with the HA is represented by the green curves and the scalar model is represented by the red curves. ................................. 71

4.8 Process windows for lines and spaces (L/S) of 22 nm linewidth with pitches of: (a) 90 nm, (b) 110 nm, (c) 150 nm, (d) 210 nm, (e) 270 nm, (f) 290 nm, (g) 310 nm and (h) 330 nm. ................................. 72

4.9 Process windows for lines and spaces (L/S) 32 nm linewidth with pitches of: (a) 90 nm, (b) 110 nm, (c) 150 nm, (d) 210 nm, (e) 270 nm, (f) 290 nm, (g) 310 nm and (h) 330 nm. ................................. 73

4.10 Process windows for lines and spaces (L/S) 45 nm linewidth with pitches of: (a) 90 nm, (b) 110 nm, (c) 150 nm, (d) 210 nm, (e) 270 nm, (f) 290 nm, (g) 310 nm and (h) 330 nm. ................................. 74
4.11 Evaluation results for a linewidth of 22 nm. Red curves represent the scalar model, green curves represent the rigorous model with the HA and the blue curves represent the rigorous model without the HA. See Tables 4.3 and 4.4 for imaging conditions.

4.12 Evaluation results for a linewidth of 32 nm. Red curves represent the scalar model, green curves represent the rigorous model with the HA and the blue curves represent the rigorous model without the HA. See Tables 4.3 and 4.4 for imaging conditions.

4.13 Evaluation results for a linewidth of 45 nm. Red curves represent the scalar model, green curves represent the rigorous model with the HA and the blue curves represent the rigorous model without the HA. See Tables 4.3 and 4.4 for imaging conditions.

4.14 Time consumption curves for all the analyzed CD targets. In the rigorous model without the HA, two representative angles of illumination are considered. (a) CD target 22nm, (b) CD target 32 nm and (c) CD target 45 nm.

4.15 PW analysis for a contact size of 65 nm. Red curves represent the scalar model, black curves represent the Quasi-3D model with the HA, magenta curves represent the Quasi-3D model without the HA (9 angles of illumination), cyan curves represent the Quasi-3D model without the HA (49 angles of illumination), yellow curves represent the rigorous model with the HA, green curves represent the rigorous model without the HA (9 angles of illumination) and the blue curves represent the rigorous model without the HA (49 angles of illumination). (a) pitch=130 nm, (b) pitch=145 nm, (c) pitch=155 nm, (d) pitch=160 nm, (e) pitch=205 nm, (f) pitch=330 nm.

4.16 PW analysis for a contact size of 90 nm. Red curves represent the scalar model, black curves represent the Quasi-3D model with the HA, magenta curves represent the Quasi-3D model without the HA (9 angles of illumination), cyan curves represent the Quasi-3D model without the HA (49 angles of illumination), yellow curves represent the rigorous model with the HA, green curves represent the rigorous model without the HA (9 angles of illumination) and the blue curves represent the rigorous model without the HA (49 angles of illumination). (a) pitch=180 nm, (b) pitch=285 nm, (c) pitch=290 nm, (d) pitch=300 nm, (e) pitch=305 nm, (f) pitch=360.
4.17 Time consumption through pitch for 65 nm contact size. The scalar model (red), the Quasi-3D model with the HA (black), the Quasi-3D model without the HA (9 angles of illumination; magenta), the Quasi-3D model without the HA (49 angles of illumination; cyan), the rigorous model with the HA (yellow), the rigorous model without the HA (9 angles of illumination; green) and the rigorous model without the HA (49 angles of illumination; blue).  

4.18 Time consumption through pitch for 90 nm contact size. The scalar model (red), the Quasi-3D model with the HA (black), the Quasi-3D model without the HA (9 angles of illumination; magenta), the Quasi-3D model without the HA (49 angles of illumination; cyan), the rigorous model with the HA (yellow), the rigorous model without the HA (9 angles of illumination; green) and the rigorous model without the HA (49 angles of illumination; blue).  

4.19 (a) Mask layout of an adapted SRAM cell and (b) aerial image with different feature sizes and pitches. The Quasi-3D model is used to compute the spectrum using the Hopkins assumption.  

4.20 Representation of the thin mask modification with additional material around the mask features. Transmission and phase values for air ($t_1, \varphi_1$), absorber material ($t_2, \varphi_2$) and surrounding material ($t_3, \varphi_3$).  

4.21 PW analysis for L/S 45 nm linewidth. The BLM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.  

4.22 PW analysis for L/S 45 nm linewidth. The BLM model is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm.  

4.23 PW analysis for L/S 45 nm linewidth. The BLM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.  

4.24 PW analysis for L/S 45 nm linewidth. The BLM model is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm.  

4.25 PW analysis for L/S 45 nm linewidth. The PM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.
LIST OF FIGURES

4.26 PW analysis for L/S 45 nm linewidth. The PM model is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm. .................................................. 100

4.27 PW analysis for L/S 45 nm linewidth. The PM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm. ...................... 102

4.28 PW analysis for L/S 45 nm linewidth. The PM is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm. .................................................. 103

5.1 Optical projection system: Regions that can be manipulated in order to enhance the scalar model. ................................................................. 106

5.2 Pupil filters associated to $J_{xx}$ and $J_{yy}$ Jones matrix components for the compensation of the mask-induced effects predicted by the rigorous model without the HA. $J_{xx}$ component (a) amplitude and (b) phase. $J_{yy}$ component (a) amplitude and (b) phase. ............................................. 108

5.3 PW analysis for L/S 45 nm linewidth. The PFM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm. ...................... 109

5.4 PW analysis for L/S 45 nm linewidth. The PFM model is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm. .................................................. 110

5.5 Pupil filters associated to $J_{xx}$ and $J_{yy}$ Jones matrix components for the compensation of the mask-induced effects predicted by the rigorous model without the HA. $J_{xx}$ component (a) amplitude and (b) phase. $J_{yy}$ component (a) amplitude and (b) phase. ............................................. 112

5.6 PW analysis for L/S 45 nm linewidth. The PFM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330. ...................... 113

5.7 PW analysis for L/S 45 nm linewidth. The PFM is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm. .................................................. 114
5.8 Correction filters associated to TE- and TM-polarization components for the compensation of the mask-induced effects predicted by the rigorous model with the HA. TM component (a) amplitude and (b) phase. TE component (a) amplitude and (b) phase. ................................................................. 117

5.9 PW analysis for L/S 45 nm linewidth. The SCM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm. ......................... 118

5.10 PW analysis for L/S 45 nm linewidth. The SCM model is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm. .............................. 119

5.11 Correction filters associated to TE- and TM-polarization components for the spectrum corresponding to the representative angle $\varphi_1$. TM component (a) amplitude and (b) phase. TE component (a) amplitude and (b) phase. 121

5.12 Correction filters associated to TE- and TM-polarization components for the spectrum corresponding to the representative angle $\varphi_2$. TM component (a) amplitude and (b) phase. TE component (a) amplitude and (b) phase. 122

5.13 PW analysis for L/S 45 nm linewidth. The SCM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm. ......................... 123

5.14 PW analysis for L/S 45 nm linewidth. The SCM model is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm. .............................. 124

5.15 Comparison of ANN, scalar and reference models in terms of amplitude and phase of the diffraction spectrum. (a) amplitude difference for the TE-polarization component, (b) phase difference for the TE-polarization component, (c) amplitude difference for the TM-polarization component and (d) phase difference for the TM-polarization component. .... 129

5.16 PW analysis for L/S 45 nm linewidth. The ANN approach is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm. ......................... 130

5.17 PW analysis for L/S 45 nm linewidth. The ANN approach is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm. .............................. 131

5.18 Schematic representation of the representative angles for x-parallel dipole illumination. ................................................................. 132
5.19 Input and output vectors of the ANN. ........................................ 133
5.20 Real and imaginary part of the 0th and 1st diffraction orders. ANN
approach compared to the reference. (a) TE-polarization component of the il-
illumination angle corresponding to ($\vartheta_1$, $\varphi_1$). (b) TE-polarization component
of the illumination angle corresponding to ($\vartheta_2$, $\varphi_2$). (c) TM-polarization
component of the illumination angle corresponding to ($\vartheta_1$, $\varphi_1$). (d) TM-
polarization component of the illumination angle corresponding to ($\vartheta_2$, $\varphi_2$). 134
5.21 PW analysis for L/S 45 nm linewidth. The ANN approach is calibrated for
the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150
nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm. ............... 135
5.22 PW analysis for L/S 45 nm linewidth. The ANN approach is validated for
different pitches that are not used in the calibration. (a) pitch=100 nm,
(b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260
nm, (f) pitch=315 nm. ............................................................. 136
5.23 Cross-section at a focus 0.0 nm for negative shifts. (a) 45.0 nm shift corre-
sponds to a trained shift. (b)-12.5 nm and (c)-2.5 nm shifts belong to the
cross-validation data. ............................................................. 137
5.24 Cross-section at a focus 0.0 nm for positive shifts. (a)17.5 nm, (b)22.5 nm
and (c)32.5 nm shifts belong to the cross-validation data. Other imaging
parameters are describe in Tables 4.3 and 4.4 ................................ 137
5.25 Spectrum amplitude and phase comparisons for TM- and TE-polarization,
duty-cycle 1:1. Blue squares represent the training data composed of phase
and amplitude. Yellow diamonds represent the training data composed
of phase, amplitude and pitch. Cyan circles represent the training data
composed of amplitude and pitch. Green triangles represent the training
data composed of phase and pitch. Red triangles represent the training data
composed only of the pitch value. Black triangles represent the references.
TE-polarization component (a) amplitude and (b) phase. TM-polarization
component (c) amplitude and (d) phase. Other important data of this test
case is described in Chapter 4, Section 4.1 Simulation parameters. .......... 139
5.26 Spectrum amplitude and phase amplitude comparisons for TM- and TE-polarization, duty-cycle 1:3. Blue squares represent the training data composed of phase and amplitude. Yellow diamonds represent the training data composed of phase, amplitude and pitch. Cyan circles represent the training data composed of amplitude and pitch. Green triangles represent the training data composed of phase and pitch. Red triangles represent the training data composed only of the pitch value. Black triangles represent the references. TE-polarization component (a) amplitude and (b) phase. TM-polarization component (c) amplitude and (d) phase. Other important data of this test case is described in Chapter 4, in Section 4.1 Simulation parameters.

5.27 White dotted lines indicate the data extracted from the scalar and rigorous spectrum to conform the pool of training and cross-validation data for contact holes. (a) TE spectrum amplitude for vertical incidence and (b) TE spectrum phase for vertical incidence.

5.28 Spectrum amplitude and phase from $-3^{rd}$ to $3^{rd}$ diffraction orders. TE-polarization component spectrum (a) amplitude and (b) phase. TM-polarization component spectrum (c) amplitude and (d) phase.

5.29 Aerial images of a contact hole array, contact size of 90 nm in a dense configuration. (a) Aerial image with the ANN, (c) Aerial image with the reference and (c) difference between two of them.

5.30 White dotted lines indicate the data extracted from the scalar and rigorous spectrum to conform the pool of training and cross-validation data for line ends. (a) TE spectrum phase for vertical incidence (reference). (b) TM spectrum phase for vertical incidence (reference).

5.31 Aerial images of line ends. Linewidth 100 nm, pitch 200 nm and gap of 28 nm. (a) 125 diffraction orders in each direction (Ref.). (b) 5 diffraction orders in each direction. (c) 11 diffraction orders in each direction. (d) 17 diffraction orders in each direction.

5.32 Comparisons of the compact models in terms of best-focus-through-pitch for L/S, 45 nm linewidth and pitches from 90 nm to 330 nm.

5.33 Comparisons of the compact models in terms of CD-through-pitch for L/S, 45 nm linewidth and pitches from 90 nm to 330 nm.

5.34 Comparisons of the compact models in terms of threshold-to-size-through-pitch for L/S, 45 nm linewidth and pitches from 90 nm to 330 nm.
List of Tables

2.1 Wavelengths and sources applied in lithography. F$_2$ excimer laser were never introduced in manufacturing. ........................................... 12

4.1 Mask materials investigated in this work. Optical parameters are given for a wavelength of 193 nm. .......................................................... 64

4.2 Alpha test: Imaging and illumination conditions. All simulations were done for a wavelength of 193 nm and a 4X reduction factor. ..................... 65

4.3 Other mask layouts used in this thesis. The illumination and imaging conditions are described. The wavelength corresponds to 193 nm and the reduction factor is 4X. .......................................................... 65

4.4 The numerical settings for the simulations in this work. All units have a 4X reduction factor. The source sampling is given for $\sigma=1$. .................. 65

4.5 Aerial images and cross sections at the best focus position for the analyzed mask models. CD=65 nm, pitch=130 nm. The Metrology line to extract the cross section is located at the central contact at the coordinates $(x_1=-65$ nm, $y_1=0$ nm); $(x_2=65$ nm, $y_2=0$ nm). ................................................ 80

4.6 Aerial images and cross sections at the best focus position for the analyzed mask models. CD=65 nm, pitch=330 nm. The Metrology line to extract the cross section is located at the central contact at the coordinates $(x_1=-65$ nm, $y_1=0$ nm); $(x_2=65$ nm, $y_2=0$ nm). ................................................ 81

4.7 Aerial images and cross sections at the best focus position for the analyzed mask models. CD=90 nm, pitch=180 nm. The Metrology line to extract the cross section is located at the central contact at the coordinates $(x_1=-90$ nm, $y_1=0$ nm); $(x_2=90$ nm, $y_2=0$ nm). ................................................ 82

4.8 Aerial Images and Cross sections at the best focus position for the analyzed mask models. CD=90 nm, pitch=360 nm. The Metrology line to extract the cross section is located at the central contact at the coordinates $(x_1=-90$ nm, $y_1=0$ nm); $(x_2=90$ nm, $y_2=0$ nm). ................................................ 83

4.9 Optimized values for the BLM. Reference with the HA. ...................... 91

4.10 Optimized values for the BLM. Reference without the HA. ................. 94
4.11 Optimized values for the BLM. Reference with the HA. .................. 98
4.12 Optimized values for the BLM. Reference without the HA. ............... 101

5.1 Optimized values for the PFM. Reference with the HA. .................. 107
5.2 Optimized values for the PFM. Reference without the HA. ............... 111
5.3 Optimized parameters for the SCM. Reference with the HA. .............. 116
5.4 Optimized parameters for the SCM. Reference without the HA. The sub-
scripts 1 and 2 make reference to the spectrum corresponding to each re-
presentative angle of illumination: $\varphi_1 = 0.196027$ and $\varphi_2 = -0.196027$. 120
5.5 Representative angles for x-parallel dipole illumination .................. 133
5.6 Linewidth and pitches used in each duty-cycle. .............................. 138
5.7 Qualitative comparison of the proposed models. “L” in the red cell in-
dicates a limited behavior. “A” in the gray cell represents an acceptable
behavior. “O” in the green cell stands for an optimum behavior. “N.A.” in-
dicates that the calibration time factor does not apply to the scalar model,
since a calibration process does not take place in this case. ............... 147
Nomenclature

Greek Symbols

\( \lambda \)  wavelength

\( \sigma \)  partial coherence factor

Other Symbols

\( \iota \)  unit imaginary number \( \sqrt{-1} \)

Acronyms

1D One-Dimensional

2D Two-Dimensional

3D Three-Dimensional

AltPSM Alternating Phase-Shifting Mask

ANN Artificial Neural Network

ARC Anti-Reflecting Coating

ArFi Argon-Fluoride excimer laser - immersion

AttPSM Attenuated Phase-Shifting Mask

BARC Bottom Anti-Reflecting Coating

BLM Boundary Layer Model

CAD Computer-Aided Design

CAR Chemically Amplified Resist

CD Critical Dimension

CMP Chemical-Mechanical Polishing
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG</td>
<td>Chrome on Glass</td>
</tr>
<tr>
<td>CPU</td>
<td>Central Processing Unit</td>
</tr>
<tr>
<td>CVD</td>
<td>Chemical-Vapor Deposition</td>
</tr>
<tr>
<td>DDM</td>
<td>Domain Decomposition Method</td>
</tr>
<tr>
<td>DoF</td>
<td>Depth of Focus</td>
</tr>
<tr>
<td>DoP</td>
<td>Degree of Polarization</td>
</tr>
<tr>
<td>DUV</td>
<td>Deep Ultra-violet</td>
</tr>
<tr>
<td>EMF</td>
<td>Electro-Magnetic Field</td>
</tr>
<tr>
<td>EUV</td>
<td>Extreme Ultra-violet</td>
</tr>
<tr>
<td>FDTD</td>
<td>Finite-Difference Time-Domain</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithm</td>
</tr>
<tr>
<td>HA</td>
<td>Hopkins Assumption</td>
</tr>
<tr>
<td>IC</td>
<td>Integrated Circuit</td>
</tr>
<tr>
<td>ILS</td>
<td>Image Log-Slope</td>
</tr>
<tr>
<td>L/S</td>
<td>Lines and Spaces</td>
</tr>
<tr>
<td>MoSi</td>
<td>Molybdenum Silicide</td>
</tr>
<tr>
<td>NA</td>
<td>Numerical Aperture</td>
</tr>
<tr>
<td>NILS</td>
<td>Normalize Image Log-Slope</td>
</tr>
<tr>
<td>OAI</td>
<td>Off-Axis Illumination</td>
</tr>
<tr>
<td>OMOG</td>
<td>Opaque Molybdenum Silicide On Glass</td>
</tr>
<tr>
<td>OPC</td>
<td>Optical Proximity Correction</td>
</tr>
<tr>
<td>PAG</td>
<td>Photo-Acid Generator</td>
</tr>
<tr>
<td>PEB</td>
<td>Post-Exposure Bake</td>
</tr>
<tr>
<td>PFM</td>
<td>Pupil Filtering Model</td>
</tr>
<tr>
<td>PM</td>
<td>Pulses Model</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>PSM</td>
<td>Phase-Shifting Mask</td>
</tr>
<tr>
<td>PW</td>
<td>Process Window</td>
</tr>
<tr>
<td>QUASI-3D</td>
<td>Quasi-Three-dimensional</td>
</tr>
<tr>
<td>RCWA</td>
<td>Rigorous coupled-wave analysis</td>
</tr>
<tr>
<td>RET</td>
<td>Resolution Enhancement Technologies</td>
</tr>
<tr>
<td>SCM</td>
<td>Spectrum Correction Model</td>
</tr>
<tr>
<td>SRAM</td>
<td>Static Random-Access Memory</td>
</tr>
<tr>
<td>TE</td>
<td>Transverse Electric</td>
</tr>
<tr>
<td>TM</td>
<td>Transverse Magnetic</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra-violet</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

The word lithography comes from the Greek “lithos” that means “stone” and “graphia” that means ”to write”. That is, literally to write in a stone. Many different lithographic techniques have been developed through the history. Specially, the styles ”to write” in the ”stone” have significantly evolved in the last century. The type of lithography that concerns this research, is the one that uses light to ”write” and a photosensitive material as ”stone”. Engraving patterns with light has numerous applications in the fabrication of micro/nano structures. For instance, optical lithography or photolithography is recognized as one of the principal manufacturing technologies in the semiconductor industry nowadays. The continued improvements in optical lithography in the past decades, have been enabling the Moore’s Law \[5\], according to which the density of the chip, measure in number of components, should double roughly every two years while maintaining a nearly constant chip price. Considerable technological challenges regarding materials and fabrication techniques are faced day to day in order to keep the trend of the Moore’s Law.

Current semiconductor devices are being mass produced with 22-nm minimum feature sizes; by 2015 these devices will have 16-nm minimum feature sizes. Optical lithography has been, and will continue being for the near future, the key technology that allows this trend.

Optical lithography must assure the miniaturization of the patterns that can be printed on the semiconductor wafer. In this printing technique, deep ultraviolet lasers are used together with high numerical aperture (NA) optics with a near-diffraction-limited performance, as discussed in the following chapters. In the optical lithographic technique, the light travels from the source to the mask, where the diffraction phenomenon occur. Afterwards, the mask–containing the pattern to be printed–is imaged by an optical projection system into the photosensitive material coated on the top of the semiconductor wafer. During the exposure step, many different physical and chemical effects occur. In this work, mainly the optical effects are the subject of study. Specially, the diffraction of the light due to the presence of the mask and the image formation of the pattern to be
Optical lithography simulation has been developing together with the industry of fabrication in the last years. Modeling and simulation offer an alternative to prevent technical mistakes in the manufacturing process as well as unnecessary costs. Shrinking the size of the features on the wafer brings new challenges also for the accurate description of the physical and chemical phenomena that occur at such small scales. When the features sizes on the mask are larger than the wavelength used in the lithographic process, the scalar diffraction theory (Kirchhoff approach) is sufficient to model the diffraction spectrum of the mask. Nowadays, the presence of mask features on the nano scale produces phase, intensity and polarization effects that can only be described with a rigorous modeling of the light diffraction from the mask and vector imaging computation of the imaged pattern.

In the Kirchhoff approach the mask is considered to be infinitesimally thin and the diffracted light is computed by means of the scalar diffraction theory. In contrast, rigorous electromagnetic field (EMF) modeling takes the complete information of the topography of the mask and its material properties into account to solve Maxwell’s equations for the propagation of light in a three-dimensional space. Several well known methods are used to solve Maxwell’s equations such as Finite-difference time-domain (FDTD) [6] or Rigorous coupled-wave analysis (RCWA) [7]. The additional effort to achieve a highly accurate description of the diffracted field has a penalty in terms of computational cost. Different authors [4, 8, 9] have studied and formulated the optimization of the computing speed of the rigorous method using decomposition techniques. In this case, a 3D problem is replaced by 2D and 1D computations. Composition of the spectra of all 2D and 1D computations produces the spectrum of the 3D mask.

Compact mask models [10, 11] and imaging modeling in lithographic simulations have been pushed to their limits to obtain a good compromise compared to the accurate but computationally expensive models. Different methods based on the thin mask modification [12, 13] and frequency domain approaches [14–16] have been reported to approximate results of EMF rigorous simulations by an enhanced Kirchhoff model. In this thesis, different alternatives of correction techniques in the mask spectrum domain are proposed, to improve the accuracy of the Kirchhoff model, but keeping the simplicity and speed of its computation.

1.1 Thesis Contributions

The most important contribution of this thesis is the formulation of three novel compact models for the computation of the far field diffraction of the photomask. The resulting spectrum from these models can be incorporated directly into the vector imaging computation to obtain a considerable accurate description of the imaging of the projection printed on the wafer.
An important aspect of mask modeling is the consideration of spatial coherence. Modern lithographic methods use off-axis illumination (OAI) techniques. The mask is illuminated by a spectrum of mutually incoherent plane waves. Using the Hopkins assumption[17], the rigorously computed diffraction spectrum is evaluated only for vertical incidence. Diffraction orders of obliquely incident angles are approximated by shifting that spectrum accordingly. Without the Hopkins assumption, the diffraction spectrum of each illumination angle is computed by rigorous EMF models.

In this research different compact mask models are formulated, calibrated and tested to improve the accuracy of the less rigorous but more efficient scalar model. In the pupil filtering model (PFM), a Jones pupil function is introduced in the projector pupil plane to describe amplitude, phase and polarization effects, which are induced by the mask. This model can be thought as a spectrum filter since it modifies the diffraction spectrum orders that contribute to the image formation. In an OAI system, the diffraction orders corresponding to an oblique illumination angle are placed at certain frequencies in the pupil filter. These positions in the frequency domain change when the angle of illumination changes. The functional form of the pupil filter does not adapt according to each illumination angle. That is, the pupil filter is independent from the illumination source, in the sense that the calibration of the model does not depend on the number of illumination angles considered.

In the spectrum correction model (SCM), a correction process performed directly on the scalar spectrum is used to tune the diffraction orders that get into the pupil of the optical projection system. In an OAI system, this approach has a higher flexibility compared to the PFM, since the filtering process is applied independently to each spectrum corresponding to each incidence angle. In that way, a specialized tuning of the diffraction orders can be done to reproduce the topographic effects predicted by the EMF modeling. However, more computation time is needed since different filters have to be calibrated.

Finally, the application of the artificial neural network (ANN) approach is considered. The ANN is introduced as a tool to retrieve the rigorous diffraction spectrum having as input the scalar diffraction spectrum. During the learning process stage, the ANN is provided with input and output vectors containing the amplitude and phase information of the scalar and the rigorous spectra respectively. These vectors are compared until the termination criteria of the learning process is reached. Afterwards, the ANN is ready to produce a diffraction spectrum that approximates the rigorous one, given the training conditions. The ANN approach operates as well as a filter on the spectrum. However, some differences arise with respect to the PFM and SCM. One is the way how the scalar spectrum is modified to obtain the rigorous spectra. The other is the calibration procedure.
All the compact models studied require two conditions. First, a computation of the scalar diffraction of the mask. Second, a calibration process of the required filters or parameters. This process is performed through comparisons with a rigorous model and will be discussed in detail in the simulation results.

Additionally to the proposed models, other compact mask models from literature are implemented in order to be compared with the proposed models. In the spatial domain, the Boundary layer model [12] is considered. The bright features of the thin mask are surrounded with a semi-transparent region with a certain width, transmission and phase. Alternatively, the bright mask features of the Kirchhoff mask are modified by adding delta functions to the edges of the absorber (pulses model)[13].

The development of compact mask models has a direct impact in the development of optical proximity correction (OPC) techniques as well as in the diffraction computation of large masks, where the computation speed is a critical issue.

1.2 Thesis Structure

The foundations of the lithographic process, applied in the semiconductor fabrication, are presented in Chapter 2. Details of phenomena that are specially relevant to this work such as diffraction and image formation, are also presented in this Chapter. Since a main topic in this thesis is the mask modeling, a complete chapter is devoted to this end (Chapter 3). This chapter includes the physical and mathematical details about rigorous modeling, as well as the description of known compact mask models. In addition, Chapter 3 contains the most significant contribution of this thesis, which is the formulation of the three novel compact mask models.

The simulation results are presented in Chapters 4 and 5. In the first part of Chapter 4, the rigorous model is compared with compact mask models based on the domain decomposition methods (DDM). In the final part of Chapter 4, known compact mask models such as the boundary layer and the pulses model are described and compared to the rigorous model. Chapter 5 contains two important parts. In the first part, the simulations with the novel proposed compact mask models are presented. In the last part, the comparisons of all the analyzed compact mask models with the scalar and the rigorous model are summarized. Finally, Chapter 6 contains the discussion, conclusions and proposes possible future work.
Chapter 2

Lithography Process in Semiconductor Technology Fabrication

In this chapter the fundamentals of lithographic process are presented. Since the main topic of study in this thesis is mask modeling, a greater effort will be directed into explain the optical diffraction principle and how its description affects the image formation. All other steps involved in the manufacturing process will be mentioned and explained for the sake of completeness.

2.1 Basics

The fabrication of IC for the semiconductor industry involves multiple physical and chemical processes performed on a silicon wafer. Certainly, silicon is the most common material used for the IC fabrication. Silicon is an inexpensive and abundant element as well as a natural semiconductor. Its electrical conductivity can be manipulated to behave like a metal or an insulator. By the addition of impurities or by interaction with an electric field, a semiconductor material can be very useful for the fabrication of devices that manipulate an electrical power input. This is the very basic principle of a transistor that can function as a switch or an amplifier. Inside an IC, millions of transistors are connected in order to execute the most complex tasks such as the computations performed by the microprocessors.

There is not a unique sequency of steps in the IC manufacturing. Furthermore, the fabrication steps may be repeated several times in order to complete a chip. In general, the major processing steps involved in the IC fabrication, may fall into one of the following categories [18]: Layer deposition, patterning and doping.
2.1.1 Wafer Preparation

The first step in the fabrication must be the preparation of the wafer [19]. Once the silicon slices are obtained, they are lapped using a coarse abrasive suspended in an alkaline solution. Then, Chemical Mechanical Polishing (CMP) will remove damage caused by lapping, producing a mirror-finished surface. CMP combines mechanical action with chemical etching to smooth the surface of the wafer. Therefore, CMP is not only use in the preparation of the wafer but also in further steps of the IC fabrication as a planarization technique.

Wafer contamination can affect its physical properties and performance. Cleaning of the wafer is necessary at many steps of the IC fabrication to avoid the attaching of contaminants to the wafer surface.

2.1.2 Oxidation

A silicon dioxide (SiO$_2$) film is deposited on the silicon wafer (Figure 2.1(a)). SiO$_2$ films are utilized in the manufacture of IC for both interlayer dielectrics (insulators) and transistor gate oxides [20–24]. These films are grown either by thermal oxidation of silicon [25], or by thermal [26, 27] or plasma-enhanced chemical vapor deposition (CVD) [28]. The use of plasma-enhanced CVD is preferred, since the deposition temperature can be considerably lower than in other methods.

2.1.3 Photoresist Coating

On the top of the (SiO$_2$), a light-photosensitive material (polymer) called photoresist is coated (Figure 2.1(b)). Sometimes wafers are primed with an adhesive material to assist the resist coating. A small portion of the photoresist is applied to the center of the wafer. The wafer is then spun at high speed—between 1500 and 1800 rotations per minute (rpm)—to spread the material over the entire surface in a thin, uniform layer. The resist thickness is a function of spin speed, polymer concentration and intrinsic viscosity [29].

2.1.4 Pre-Bake

After photoresist application, the wafers are ”soft baked” (Figure 2.1(c)) at temperatures around 70-100 °C. This soft bake causes the remaining solvents to evaporate, removal of stresses and promotes adhesion of the resist layer to the wafer. After the pre-bake the resist thickness is typically in ranges between 0.1 and 2 µm.
2.1.5 Exposure

A mask containing the pattern to be printed in the photoresist, is exposed to a light source (Figure 2.1(d)). Typically, a deep ultraviolet source of 150-193 nm wavelength (DUV) or extreme ultraviolet source of 10-13.5 nm (EUV) is used. Characteristics of the light source will be further described, when the basic components of a lithographic imaging system are outlined in Section 2.2. The light is diffracted by the mask and then imaged towards the photoresist by an imaging projection system. The illumination and the projection systems must assure that the light is delivered to the photoresist with the proper intensity, directionality, spectral characteristics, and uniformity across the wafer, allowing a nearly perfect transfer or printing of the mask image onto the photoresist.

(a) Film deposition  (b) Spin-coating  (c) Pre-bake

(d) Exposure  (e) PEB  (f) Development

(g) Hard-bake  (h) Etching  (i) Stripping

Figure 2.1: A standard sequence of the resist processing. (a) Film deposition, (b) spin-coating, (c) pre-bake, (d) exposure, (e) post-exposure bake, (f) development, (g) hard bake, (h) etching and (i) stripping.
2.1.6 Post Exposure Bake (PEB)

The exposed wafer undergoes a bake treatment (Figure 2.1(e)), typically performed at 110°C for 1-2 min. A PEB process carried out near the softening point of the photoresist reduces mechanical stress formed during the exposure of thick resist films. The reduction of the stress turns into a better adhesion of the photoresist. The PEB treatment is also required because the reactions initiated during exposure might have not been completed. The PEB promotes the thermally activated diffusion of carboxylic acid formed during exposure from the photoactive compound [29]. This diffusion step smoothes the standing waves created during monochromatic exposure especially in case of highly reflective substrates [30]. Otherwise, the standing wave patterns would be replicated into the resist profile reducing the spatial resolution of the resist and the desired aspect ratio.

2.1.7 Development

The exposure to light and the PEB generate both physical and chemical transformations to the photoresist. During the development process, the regions that were exposed to light differentiate from those that were not. Then, the image formed in the photoresist is transformed into a relief (Figure 2.1(f)). The wafers are uniformly covered with a developer solution to remove either the exposed or the unexposed areas of the photoresist, depending on the type of photoresist used. With positive photoresist, the irradiated regions become soluble after the PEB. Under the development treatment, the photoresist is removed in the exposed parts. On the other hand, negative photoresist becomes insoluble after the PEB, making that the developer solution removes only the unexposed regions. During the development step, a “negative” of the mask pattern is formed.

2.1.8 Hard Bake

A third bake treatment takes place after the development (Figure 2.1(g)). Its purpose is the evaporation of solvents and to harden the photoresist. The heat takes the water out of the photoresist and increases the etch-resistant properties. Nominal hard bake temperatures and times are from 130 to 200°C for 30 min in a convection oven [18].

2.1.9 Etching

An etching process is necessary to transfer the pattern to the substrate. Etch involves both chemical and mechanical mechanisms for removal of the material not protected by the photoresist (Figure 2.1(h)). The two major techniques of etching are wet etching and dry etching [18][29]. In the wet etching technique, the wafer is immersed in a chemical solution to remove the silicon dioxide layer or any other deposited material. This is a low
cost and easy to implement process. It exhibits a high etching rate and a good selectivity for most materials. However, it is inadequate for defining feature sizes of less than 1 μm, presents a potential of chemical handling hazards and wafer contamination issues. The dry etching technique is often used in semiconductors fabrication, since it can be performed with anisotropic directionality, which avoids the undercutting of the photoresist. The dry etching technology may use one of the following methods:

- Physical sputtering: Physical bombardment with an ion milling machine.
- Plasma Etching: Plasma-assisted chemical reaction.
- Reactive Ion Etching (RIE): Chemical reaction together with ion bombardment.

The dry etching is capable of defining features smaller than 100 nm. However, it is an expensive and hard to implement technique with a low throughput, poor selectivity and contains potential radiation damage.

2.1.10 Stripping

After etching, the target pattern is fixed to the top layer of the wafer. The photoresist has then completed its task and is no longer needed. Traditionally, the photoresist material is removed from the surface (Figure 2.1(i)) by wet chemical processing [18].

2.1.11 Doping

Ion implantation is the most widely applied doping technique in modern semiconductor technology since it allows laterally well defined dopant profiles [19]. Selective doping of certain regions of the semiconductor occurs after the stripping of the photoresist in the patterning process (see Figure 2.2). Ionized dopant atoms are accelerated in a high-vacuum environment by an electrostatic field, strike the surface and penetrate into the wafer. The penetration depth is adjusted by the acceleration energy. Regions of the semiconductor that are not covered by photoresist are exposed to a dopant impurity. A p-type dopant like boron creates mobile holes, vacancies inside the material where an electron could go. n-type dopants such as phosphorous, arsenic and antimony create excess mobile electrons. The interface between p-type regions and n-type regions of silicon is called a p-n junction and is one of the basic structures for the functionality of semiconductor devices.

After the ion implantation, a thermal processing step is necessary for two reasons. First, the implanted dopant atoms stay on interstitial sites inside the crystal lattice and are thus electrically inactive. Second, the implanted ions displace silicon atoms and therefore cause implantation damage. A subsequent thermal treatment is necessary to
compensate for the implantation damage and to activate the dopants by moving them on substitutional lattice locations. This *annealing process* limits the possible resolution of the implanted profile. Rapid thermal annealing at a very high temperature allows to repair the damage with minimal dopant movement.

![Figure 2.2: Selective doping of certain regions of the semiconductor by means of ion implantation. (top) Stripping and (bottom) ion implantation.](image)

2.1.12 Deposition of Materials

The CVD technique is used to deposit additional material on the wafer surface. These layers may be applied at various stages during the manufacturing process. For instance, new layers for further junction formation or to form an insulating layer between two or more conductive layers. CVD is commonly used to deposit layers of polycrystalline silicon, silicon dioxide, and silicon nitride on the substrate [19].

2.1.13 Metallization

In order to make a semiconductor device functional, a metallization process is required [18]. This makes the connection of different parts inside the device and the electrical terminals for the chip. Metal is deposited by vacuum evaporation, sputtering or CVD techniques.

2.2 Components of a Lithographic Imaging System

Optical projection printing is the standard patterning method used in semiconductor fabrication. The lithographic patterning represents one of the most critical steps in the
determination of the circuit smallest critical dimension. The transfer of the pattern is performed by a projection system, which is shown in Figure 2.3. This can be divided into four parts: the illumination system, the photomask (also called reticle), the projection system and the silicon wafer. The wafer is composed of many layers including the photoresist spun on the top, which is sensitive to the light exposure.

![Figure 2.3: Schematic representation of an optical projection system.](image)

### 2.2.1 Illumination System

The illumination system must provide a uniform illumination of the entire reticle. For this purpose, arrangements of various optical elements, such as lenses, apertures, filters and mirrors are used. Highly monochromatic radiation is desired because high quality refractive optics (reflective for EUV technology) can be designed to operate nearly in aberration-free conditions. Intense illumination (high power) guarantees high throughput since the exposure time per wafer decreases at a high power. Light sources for high-volume production must be very intense and of very narrow bandwidth, this determines the available wavelengths of operation.
2.2.1.1 Illumination Sources

Table 2.1 shows the light sources applied in optical lithography for the past decades. Since the introduction of the optical lithography technique into the semiconductor fabrication, the shrinking of the wavelength— to meet the requirements of smaller feature sizes in the chip—has been a trend. To his end, high-pressure mercury Hg plasma arc discharge lamps emitting ultraviolet light were used in the past. The distribution of the emitted spectrum depended on the partial pressures of the Hg. In the early 1970s until the mid-1980s, the Hg-lamp spectrum was filtered to select the g-line at 436 nm and the i-line at 365 nm [31].

Table 2.1: Wavelengths and sources applied in lithography. F\textsubscript{2} excimer laser were never introduced in manufacturing.

<table>
<thead>
<tr>
<th>Wavelength (nm)</th>
<th>Light Source</th>
<th>Year of Introduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>436</td>
<td>Hg arc lamp (g-line)</td>
<td>1970</td>
</tr>
<tr>
<td>365</td>
<td>Hg arc lamp (i-line)</td>
<td>1984</td>
</tr>
<tr>
<td>248</td>
<td>KrF excimer laser</td>
<td>1989</td>
</tr>
<tr>
<td>193</td>
<td>ArF excimer laser</td>
<td>1999</td>
</tr>
<tr>
<td>157</td>
<td>F\textsubscript{2} excimer laser</td>
<td>-</td>
</tr>
<tr>
<td>13.5</td>
<td>Xe or Sn plasma sources</td>
<td>2015?</td>
</tr>
</tbody>
</table>

In 1989, the first excimer laser was used in the semiconductor production [32]. The deep UV (DUV) era began with the introduction of the KrF excimer laser at 248 nm. Subsequently, ArF at 193 nm excimer laser were introduced in 1999 [33]. In an excimer laser, the noble gas, e.g., Krypton (Kr) or Argon (Ar), is excited and reacts with a molecule composed of two identical atoms like Fluorine (F)\textsuperscript{1}. Excimers have a bandwidth of about 1 nm and emit strongly in a multimode fashion with relatively low spatial coherence. This is a crucial advantage for lithography applications, because it avoids or at least decreases the problem of “speckle”\textsuperscript{2}. The high-power pulses of excimer lasers allow extremely short exposure times (10-20 ns), which increases the throughput.

The Xe or Sn plasma sources for extreme UV (EUV) lithography are either discharge-produced or laser-produced [35–37]. Discharge-produced plasma is obtained by ionization of Sn vapor. Laser-produced plasma takes place when microscopic droplets of molten Sn are heated by a powerful laser. EUV sources at 13.5 nm are not introduced yet to the mass production of semiconductor devices. Critical technological and economical challenges such as the source life time and power, optics and the strict requirements for the photoresist and mask, prevent this technology from being used in the full-chip production.

\textsuperscript{1}A molecule with two atoms of the same element is called dimer. The name “excimer” combines the two words “excited” and “dimer.”

\textsuperscript{2}The mutual interference of multiple waves, whose components of phase and amplitude are random give as a result a random intensity pattern called “speckle” [34].
Nowadays, state-of-the-art optical lithography systems use the 193 nm ArF excimer laser together with liquid immersion techniques (ArFi).

2.2.1.2 Köhler Illumination

The image formation by a projection system requires that the object (photomask) is uniformly illuminated. Exposure systems used in optical lithography implement the Köhler configuration [38, 39] because it provides uniform illumination from a source that in general is non-uniform. Köhler illumination utilizes a collector lens in front of the light source (Figure 2.4). This lens focuses the light at the aperture stop (diaphragm) before the condenser. Then, the condenser lens forms an image of the illumination source at the entrance pupil of the objective lens, and the mask is at the exit pupil of the condenser lens. In the case of highly non-uniform sources, a diffuser may be added to the optical system to further improve uniformity at the focal plane.

![Figure 2.4: Köhler illumination](image)

Each point source can be treated as generating a coherent, linearly polarized plane wave of spatial frequency determined by the position of the point source relative to the optical axis.

Each point source can be treated as generating a coherent, linearly polarized plane wave of spatial frequency determined by the position of the point source relative to the optical axis [38]. In other words, using Köhler illumination each point in the mask area is illuminated by the entire source so that irradiance variations across the source do not affect the mask illumination. Alternatively, the source cross section can be imaged on the entrance pupil of the projection lens, in which case each point source produces spherical waves converging toward one point on the entrance pupil plane. Given that, the distance between the object plane and the entrance pupil in real lithography scanners is at least
several wavelengths, the curvature of the spherical wavefronts that illuminate the mask can be assumed negligible over several resolution units, allowing their approximation as plane wavefronts [40].

2.2.2 Mask

The photomask contains the information about the pattern to be printed, encoded as a variation of transmittance and/or phase. The layout of this pattern is generated by converting the circuit diagram with the help of computer-aided design (CAD) tools into geometrical shapes. The mask containing one chip pattern or a matrix of several chip patterns is also called reticle.

The fabrication of the mask is in many steps similar to the patterning on the wafer [41]. Basically, a photomask is a glass plate covered with a layer of absorber. Then, a layer of the anti-reflective coating (ARC) is used to prevent the internal reflection of light in the photoresist. Finally, the photoresist is coated on top of the ARC layer.

Currently, there are two techniques to write the patterns on the mask [19]. These techniques are known as electron-beam writers and laser-beam writers. In any technique, both mask design and data preparation have to be performed to transfer the pattern to the writing system. Since complex mask layouts with strict resolution requirements are demanded, the electron-beam technique is preferred over the laser-beam [18].

Once the mask layout has been patterned in the photoresist, a development process takes place. This is completely equivalent to the development process in the wafer-patterning explained in 2.1.7. Subsequently, an etching of the material not protected by the photoresist is carried out (2.1.9) as well as a stripping of the photoresist (2.1.10). A highly controlled etch process is required, when phase shift masks are fabricated, since precise amounts of glass material have to be removed. The mask fabrication process is not finalized until strict inspection and repair steps are carried out. The photomask bears the pattern that will be transferred to thousand of wafers, thus its quality and durability are of critical importance since defects on the mask will be reproduced on the wafer.

Lithographic projection systems are designed to reduce the size of the image compared to the size of the mask, by a factor of 4. Therefore, the mask is drawn 4 times the required size on the wafer. Depending on their operation principle, photomasks can be divided into two broad categories: conventional binary or chrome on glass (COG) masks and phase-shifting masks (PSM) [41]. Figure 2.5 shows a comparison between a binary mask and two typical phase-shifting masks.

A binary or COG mask (Figure 2.5(a)) consists of a transparent substrate (mask blank) covered with a thin opaque film that describes the desired pattern. Light can either pass unobstructed through the transparent substrate or be completely blocked if it hits on
an area that is protected by the opaque film. This characteristic binary behavior of the transmission of the mask is responsible for its name. The mask blank for DUV lithography typically consists of fused silica glass that has excellent transmission at $\lambda = 248$ nm. For $\lambda = 193$ nm and $\lambda = 157$ nm, quartz glass is also used. The opaque film is typically made from chromium (Cr) or opaque molybdenum silicide on glass (OMOG) and has a thickness on the order of 60 nm to 100 nm.

Phase-shifting masks (PSM) modulate both amplitude and phase of the electromagnetic field propagating through them and improve image contrast by inducing destructive interference of the fields with opposite phases. This phase modulation can strongly increase the attainable resolution. The phase-shifting principle was introduced in 1982 by Levenson [42]. There are many different types of PSMs depending on the way, how the phase modulation is achieved.

In the alternating phase-shifting mask (AltPSM) (Figure 2.5(b)), a 180° phase change is created by etching into the substrate to a specified depth. Repeating the etched space on alternate open areas creates a periodic phase shifting structure that improves the imaging of the features defined by the opaque absorbers. The phase difference on the AltPSM leads to destructive interference, resulting in sharp dark areas in the image. The binary mask image has not as high contrast as the AltPSM, due to the lack of phase interaction. Because of the etching process, the AltPSM manufacturing is more complex than the COG fabrication process. The difference in the amount of material removed, $d_{\text{etch}}$, is such that the path length difference between light passing through the different material regions is half of the wavelength in air. The theoretical etch depth is given by:

$$d_{\text{etch}} = \frac{\lambda}{2 [n_{\text{substrate}}(\lambda) - n_{\text{air}}(\lambda)]}$$  \hspace{0.5cm} (2.1)

where $\lambda$ is the wavelength, $n_{\text{substrate}}$ is the refractive index of the glass substrate at the exposure wavelength and $n_{\text{air}}$ the refractive index of air.

The attenuated phase-shifting mask (AttPSM; also half tone mask) (Figure 2.5(c)) is obtained by using a partial transmitting absorber such as molybdenum silicide (MoSi), with 6% intensity transmission and 180° phase. The refractive index of this material has a complex value of $n_{\text{MoSi}} = 2.442 - 0.586i$. The partially transmitting region contributes—by means of the destructive interference between those regions and the main features—to improve the image contrast. In this case, the mask manufacturing is not dramatically different from the COG mask, although thickness control of the absorber region is a critical factor.

Other PSMs are shown in Figure 2.6. A tri-tone mask (Figure 2.6(a)) is also a type of AttPSM but contains an additional layer of chrome on top of the phase shift layer. In this way, the contrast is increased in the photoresist.
A chromeless mask (Figure 2.6(b)) does not use opaque films. The phase shift is achieved by trenches, which are directly etched into the glass substrate. This type of masks share the same etching process challenges as the AltPSM. However, the inspection of defects in these masks is indeed more difficult due to the lack of material contrast.

Figure 2.5: Types of masks. (a) Binary mask, (b) alternating phase shift mask with etched substrate and (c) attenuated phase shift mask. From top to bottom, each plot represents the mask, the mask transmission function, the electrical field and the intensity distribution.

Figure 2.6: Other relevant types of PSM. (a) Tritone mask and (b) chromeless mask.
2.2.3 Projection System

For industrial high volume semiconductor production, the most common image forming technique is the projection printing. The entire mask or at least a large portion is imaged at once. Figure 2.7 shows schematically the projection system components. Several lenses (up to 40), apertures, filters and other optical elements assure the quality of the formed image. In addition, the optical system provides a reduction factor R of 4X of the pattern engraved in the mask.

![Schematic imaging projection system](image)

Figure 2.7: A schematic imaging projection system.

The mask is uniformly illuminated. The propagation of the light through the mask is perturbed due to the presence of apertures of different sizes and shapes on the reticle. This is explained by the diffraction theory which describes how light propagates in a medium, including the effects of obstacles and boundaries. The diffracted light is collected by the entrance pupil and is recombined by the projection optics to be imaged onto the resist surface. Higher frequency components are filtered out by the lens. Therefore, the final aerial image (see Aerial Image Formation 2.3) is a partial reproduction of the original pattern.

If the imperfections that may be introduced by the optical elements are corrected with a very high quality material and strict design specifications. Then, the optical system is

---

1 The quality of the image refers to the correction of optical aberrations or deviations of the wavefront from its ideal shape.
said to be diffraction-limited since is free of aberrations or apodization effects but its resolution is limited by the diffraction phenomenon [43].

2.2.4 Photoresist

The photoresist is applied as a thin film to the substrate. Once the aerial image is defined, it has to be transferred into the photoresist to generate a replica of the mask directly into the wafer surface. Photoresists can be divided in two main branches, positive or negative. Positive resists become removable in developer solution upon exposure of light, while negative resists remain after the development process.

A photoresist may be manufactured for specific applications. It can be tuned to respond to certain conditions of the exposing source. All resists contain at least the 3 following components [18]: Polymer, solvent, and sensitizers. The polymer contributes to the photosensitive properties by changing its structure in reaction to the irradiance (polymerization or photo-solubilization). The largest component by volume in a photoresist is the solvent. The solvent allows the resist to be coated to the wafer as a thin layer by spinning. Chemical sensitizers are added to the resist to induce certain reactions in the polymer.

Chemically amplified resist (CAR) are composed of a base resin, a dissolution inhibitor and a photo-acid generator (PAG). The exposure of the photoresist to DUV light generates acid from the PAG. During a subsequent PEB, the photo-generated acid catalyzes a thermally induced reaction that makes the exposed photoresist regions soluble in the developer [44][45]. Meanwhile, the diffusion of acid enables the removal of standing waves originated from the interference of internal reflected waves. A step further to prevent reflections inside the photoresist that may generate undesired effects on the pattern, is the utilization of a bottom antireflective coating (BARC) [46]. This layer is placed just below to the photoresist to reduce reflections from the substrate.

The 3D resist image (bulk image) after development serves as a guide pattern through which series of etching, doping, deposition and metallization steps result in a functional semiconductor device [47]. The photoresist plays an important role in the final resolution that is obtained in the transferred pattern. Regardless of the quality of the aerial image, poor resist contrast may degrade the resolution attainable with a particular lithography system. The resist contrast depends on the resist material as well as on the resist process parameters, which need to be carefully calibrated.

It is important to point out that in this work, the critical dimension (CD) measurements (see Section 2.4) are the result of an aerial-image-based analysis. The bulk image analysis is not carried out and the resist for the purpose of this research is treated as an intensity threshold detector.
2.3 Aerial Image Formation

The aerial image corresponds to the intensity distribution obtained in the image plane of the optical projection system. The image formation is a fundamental optical phenomena that involves the study of the propagation of light through different media. The theory of Fourier optics and Maxwell’s equations–either in its scalar or vector form–provide a strong physical framework to describe the image formation by an optical projection system. In this section, the details about the imaging theory are described. In addition, important characteristics and parameters related to the optical imaging system are explained.

2.3.1 Scalar Diffraction Theory

Light can be described as a transverse electromagnetic wave with coupled electric and magnetic fields traveling through time and space \[38\]. In classical physics, the diffraction phenomenon is described as the bending of waves around small obstacles and the spreading out of waves passing small openings. In general, diffraction theory describes how light propagates in free space or through a medium, including the effects of boundaries. Diffraction occurs whenever the light encounters an obstacle, but its effects are more pronounced when the wavelength \(\lambda\) of the propagating light is roughly similar to the dimension of the diffracting object.

In optical lithography, this phenomenon is observed when the light from the source is propagating through the mask, which consist of several apertures of different sizes and shapes. Figure 2.8 shows the schematic principle of diffraction when a double slit configuration is placed in front of a propagating plane wave.

In the Figure 2.8, \(m\) represents the diffraction order being the zero order the light that passes through the slit and is not bent. To either side of the zero order are two peaks called the +1 and -1 diffraction orders. These peaks occur at spatial frequencies of \(\pm1/d\), where \(d\) is the distance between the two slits. In this study, the mask pattern is the diffracting object of interest and the distance separating the apertures is called pitch. The mask diffraction spectrum –for apertures much larger than the wavelength– is fairly well described by the scalar diffraction theory.

The scalar diffraction theory collects an extensive research and observations of several mathematicians and physicist such as Sommerfeld, Rayleigh, Huygens and Fresnel. But it was not until late XIX century that Gustav Kirchhoff rigorously described the solution of the homogeneous wave equation at any arbitrary point in the field, in terms of the values of the solution and its first normal derivatives on an arbitrary closed surface enclosing the observation point. This theorem constitutes the basis of the Scalar Diffraction Theory \[48][49\].

The derivation of scalar diffraction considers the light as an electromagnetic wave with
coupled electric and magnetic fields traveling through homogeneous space. Neglecting polarization, the diffracted field can be described by a scalar function $U(x, y)$ representing the electric field amplitude, according to the Huygens-Fresnel principle

$$U(x, y) = \frac{1}{j\lambda} \int\int_{\Sigma} U(\xi, \eta) \frac{e^{jkro_1}}{r_{01}^2} d\xi d\eta$$  \hspace{1cm} (2.2)

$r_{01}$ is the distance from the point $P_o$ to the point $P_1$ as shown in Figure 2.9. $U(\xi, \eta)$ represents the object field. The light used in lithography exhibits a strong monochromacity, the time dependence of the field $U(x, y)$ is a harmonic one and satisfies the Helmholtz equation

$$(\nabla^2 + k^2) U = 0$$  \hspace{1cm} (2.3)

where $k$ is the wave vector with magnitude $|k|^2 = k = \frac{2\pi}{\lambda}$ and $\lambda$ the wavelength.

The object field $U(\xi, \eta)$ satisfies the Kirchhoff boundary conditions:

1. Across the surface $\Sigma$ the field distribution $U(\xi, \eta)$ is the same as it would be in the absence of an aperture $A$.

2. Out of the surface $\Sigma$ the object field $U(\xi, \eta) = 0$

$^1$Since the electric field is considerably more effective at exerting forces and doing work on charges than the magnetic field, the electric field $E$ is referred to as the optical field [50]
Figure 2.9: The diffracting aperture lies in the plane \((\xi, \eta)\). Diffracted field lies in the plane \((x, y)\).

The Cartesian coordinates of \(P_o\) and \(P_1\) are \((x, y, z)\) and \((\eta, \xi, 0)\) respectively. The Euclidean distance between these two points is given by

\[
    r_{01} = \sqrt{z^2 + (x - \xi)^2 + (y - \eta)^2} \quad (2.4)
\]

Applying a binomial expansion over \(r_{01}\):

\[
    r_{01} \approx z \left[ 1 + \frac{1}{2} \left( \frac{x - \xi}{z} \right)^2 + \frac{1}{2} \left( \frac{y - \eta}{z} \right)^2 \right] \quad (2.5)
\]

where the distance \(z\) satisfies that

\[
    z^3 \gg \frac{\pi}{4\lambda} \left[ (x - \xi)^2 + (y - \eta)^2 \right]_{\text{max}}^2. \quad (2.6)
\]

Substituting Equation 2.5 in Equation 2.2:

\[
    \mathcal{U}(x, y) = \frac{e^{jkz}}{j\lambda z} \int \int_{\Sigma} \mathcal{U}(\xi, \eta) e^{\frac{j}{\lambda z}[(x-\xi)^2+(y-\eta)^2]} d\xi d\eta \quad (2.7)
\]

which is the Fresnel approximation for the scalar diffraction. If the dimensions of the diffracting object are small compared to the distance between the diffracting plane and the diffracted field, that is if
the quadratic factor in the Fresnel integral in 2.7 is approximately unity over the entire aperture $A$ and the diffracted field is expressed as

$$U(x, y) = \frac{e^{jkz}e^{j\frac{\pi}{2}}}{j\lambda z} \int_\Sigma U(\xi, \eta)e^{-j\frac{\pi}{2}\left[x\xi+y\eta\right]}d\xi d\eta$$

which is the Fraunhofer approximation or diffraction in the far field. Aside from a phase factor, the Equation 2.10 is the Fourier transform of $U(\xi, \eta)$, thus the diffracted field can be written as

$$U(x, y) = \frac{e^{jkz}e^{j\frac{\pi}{2}}}{j\lambda z} \mathcal{F}\{U(\xi, \eta)\} \bigg|_{f_x = x/\lambda z; f_y = y/\lambda z}$$

Figure 2.10 shows the calculation of the mask near field using the scalar model and a rigorous computation (see Chapter 3) for a rectangular aperture. In the scalar model, the light passes unperturbed through the aperture and is blocked by the obstacle (Figure 2.10(a)). In Figure 2.10(b), the Waveguide method (Section 3.1.2) is used to compute the near field of an equivalent topographic mask. The edges of the absorber perturb the light and the diffraction effects can be clearly observed. The near field is extracted at 0.02 $\mu$m from the mask.

Figure 2.10: Binary mask (COG) with a rectangular aperture of 700 nm, $\lambda = 193$ nm. (a) Intensity of the near field obtained with the scalar model. (b) Calculated near field intensity using a rigorous method (Waveguide). TM x-polarization component.
2.3.2 Numerical Aperture

The numerical aperture (NA) of a lens (or an optical system) characterizes its ability of gathering light. The space at the entrance side of the lens is called object space, and the space at the exit side of the lens is called image space. The NAs on the object and image spaces are defined as:

\[ NA_o = n_o \sin(\theta_o) \quad (2.11) \]
\[ NA_i = n_i \sin(\theta_i) \quad (2.12) \]

where \( n_o \) is the refractive index of the medium between the object and the lens, \( n_i \) is the refractive index of the medium between the lens and the image. \( \theta_o \) and \( \theta_i \) are the half-angle of the maximum cone of light that can enter and exit the lens respectively, as shown in Figure 2.11.

![Figure 2.11: The numerical aperture (NA) with respect to the object space and the image space.](image)

The magnification of the system \( M \) is the ratio of the numerical apertures on the side of the object and the side of the image:

\[ M = \frac{NA_o}{NA_i} \quad (2.13) \]

\( M \) usually takes values \( M = 1/4 \) and \( M = 1/5 \) for 4X or 5X reduction systems, respectively. Nowadays, values of \( NA_i \) larger than 0.7 are commonly used in optical lithography.
projection. Higher values of NA allow more oblique rays to enter to the optical system, producing a more highly resolved image.

### 2.3.3 Partial Coherence Factor and Off-Axis Illumination

An important aspect in a lithographic system is the coherence. As mention in previous sections, it is important that the light is strongly coherent in the time domain (temporally coherent) because of the required monochromaticity of the light source. The illumination is said to be spatially coherent, when the phase of two points in the wavefront are highly correlated for all times [43]. That is, the spatial coherence is an indicator of the uniformity of the phase in a wavefront.

The concept of partial (spatial) coherence in optics was formally introduced by Hopkins in 1951 [51] with the contributions of Zernike [52] and Van Cittert [53]. The coherence in the spatial domain, is an adjustable parameter that has a critical impact on the imaging performance. The amount of partial coherence is determined by the ratio of the numerical apertures $NA_c$ and $NA_p$ (Figure 2.12):

$$\sigma = \frac{NA_c}{NA_p}, \quad (2.14)$$

where $\sigma$ is the partial coherence factor. $NA_c$ is the numerical aperture of the illumination source defined as:

$$NA_c = n_c \sin(\alpha_{max}), \quad (2.15)$$

where $n_c$ is the refractive index of the material surrounding the condenser and $\alpha_{max}$ is the half-angle of the illumination cone. $NA_p$ defines the numerical aperture of the optical projection system.

$$NA_p = n_p \sin(\beta_{max}). \quad (2.16)$$

In this way, the partial coherent factor is finally expressed as:

$$\sigma = \frac{n_c \sin(\alpha_{max})}{n_p \sin(\beta_{max})}, \quad (2.17)$$

The impact of $\sigma$ on the image formation is demonstrated in Figure 2.13 by showing the image intensity near a simple knife-edge [1]. The extreme case $\sigma = \infty$ corresponds to incoherent illumination and gives the smoothest profile. Decreasing $\sigma$ increases the edge slope (contrast) and also the side-lobes while decreasing the intensity minimum near the edge. Approaching to this minimum contributes to the linewidth fidelity and profile quality. If $\sigma$ is further reduced to values as low as 0.2 in order to decrease the intensity
Figure 2.12: Partial coherent imaging system.

Figure 2.13: The degree of partial coherence $\sigma$ determines the tradeoff between a sharp slope and a low overshoot space. Intensity of the image corresponding to line and spaces of 10 $\mu$m linewidth, NA=0.28, $\lambda = 0.436$ $\mu$m. Plot from [1].

minimum, the side-lobes generation becomes excessive and extends laterally indicating that proximity effects between adjacent features are more likely. The extreme case $\sigma = 0$ refers to an ideally coherent point source yielding the sharpest slope but intolerable side-lobes generation. Hence, there must be a trade-off between a sharp edge and the side-lobe generation in order to choose $\sigma$. In practical lithography typical values for $\sigma$ range between 0.3-0.9.
In a spatially coherent illumination system, a normally incident plane wave illuminates the mask. Hence, the diffraction pattern is centered at the entrance pupil (Figure 2.14(a)). For partially coherent illumination, the extended source is considered as a set of several plane waves illuminating the mask at different angles. By illuminating the mask at an off-axis angle (Figure 2.14(b)), different orders that were blocked before by the entrance pupil may contribute to the image formation. In that way, the image can be formed by interference of the 0th diffraction order and either one of the ±1 orders. This increases the resolution by allowing more diffraction orders into the projection pupil, which would otherwise not pass through the optical system.

The impact of shifting the diffraction pattern within the entrance pupil may be difficult to generalize. In some cases, a diffraction order is blocked or allowed to pass, this depends on the incidence angle of illumination, the mask pattern and the NA. If one off-axis angle of illumination causes a shift in the diffraction pattern, a range of off-axis angles will cause a range of shifts, as shown in Figure 2.15. By introducing an extended range of incidence angles, the number of diffraction orders within the numerical aperture of the lens is increased. Different illumination schemes (Figure 2.16) can be achieved by introducing an aperture between the light source and the condenser optics. Symmetric light sources such as circular (Figure 2.16(a)) and annular (Figure 2.16(b)) provide directional uniformity and guarantee the same printing quality for features in all orientations. More advanced illumination configurations such as the quadrupole (Figure 2.16(c)) or the dipole (Figure 2.16(d)), can further improve image quality of mask patterns with specific symmetry at the expense of some directional uniformity.

Figure 2.14: Schematic of (a) On-axis illumination and (b) Off-axis illumination.
2.3.4 Resolution in Optical Lithography

The smallest resolvable feature or critical dimension (CD) determines the resolution of a lithographic imaging system and is given by \[ CD = k_1 \frac{\lambda}{NA} \] \[ (2.18) \]

where \( \lambda \) is the wavelength, \( NA \) the numerical aperture of the objective lens and \( k_1 \) is a process-related factor that strongly depends on optics and resist parameters. Values of \( k_1 \) between 0.3-1.0 are typical, with a trend toward decreasing to values approaching 0.2 \[57\]. In order to increase the resolution (decrease the critical dimension) any combination of reducing \( k_1 \), decreasing \( \lambda \) and increasing \( NA \) may be applied. Decreasing \( \lambda \) arbitrarily is not possible for numerous practical reasons. At a given exposure wavelength, the power of the light source has to satisfy the throughput requirements. Suitable optical materials for

---

1Ernst Abbe (1873) and Lord Rayleigh (1879) were the greatest contributors to the definition of the resolution limit concept in optics. Nowadays, this limit is often called the Rayleigh criterion.
lenses and mask substrates and also resist materials are not easy to develop at arbitrary wavelengths. The physical limit of NA is one, as it is impossible to increase resolution by further refraction, additional resolution is obtained by inserting an immersion medium with a higher index of refraction between the lens and the image. If the refraction index of the immersion medium is 1.44 an NA of 1.35 can be reached. This is state-of-the-art lithography at 193 nm wavelength in the ArF-immersion systems.

In a system of multiple apertures, like a grating, the ultimate resolution is determined by the minimum resolvable distance between two features, that is the period \( p \) of the grating. Under partial coherent illumination, the smallest resolvable period \( p_{\text{min}} \) is given by the following expressions [58]:

\[
 p_{\text{min}} = \begin{cases} 
 \frac{\lambda}{NA} & \text{for } \sigma = 0 \text{ coherent illumination} \\
 \frac{\lambda}{1 + \sigma NA} & \text{for } 0 < \sigma < 1 \text{ partially coherent illumination} \\
 \frac{\lambda}{2 NA} & \text{for } \sigma = \infty \text{ incoherent illumination}
\end{cases}
\]  

(2.19)

### 2.3.5 Abbe Formulation

The study of the image formation considering the coherence of light, was first introduced by Ernst Abbe in 1873 [59]. Years later, Van Cittert [53] and Zernike [52] showed how the coherence of the light can be considered as a state that is able to propagate. The formulation presented in this section reviews parts of the analysis of Wolf [38] and Goodman [60] about the illumination coherence and its implications on the image formation.

The illumination used in optical lithography is temporally coherent (monochromatic). In addition, all the points in the light source are mutually (spatially) incoherent. Each point in the source is considered individually and the image intensity produced by the light from that single point is calculated. Afterward, all image intensity contributions from the points that constitute the source are added. Addition (integration) of the image intensity distributions is possible since the original source is assumed to be spatially incoherent.

Considering the scheme in Figure 2.17, the object (mask) in the plane \((\xi, \eta)\) is illuminated by a light source located at the plane \((u, v)\) and imaged at the plane \((x, y)\). For a Köhler illumination system, each point of the source corresponds to a plane wave illuminating the mask with an angle of incidence that depends on the position \((u, v)\). A point source at coordinates \((u, v)\) emits light represented by a time-varying phasor \( A_s(u,v,t) \). The propagating light reaches the object and passes through it, producing a time-varying phasor \( A'_o(\xi, \eta; u,v,t) \).

An optical system can be represented by an amplitude response function \(^1\). This function denotes the response of an imaging system to a point source or point object.

---

\(^1\)also called impulse response or commonly point-spread function (PSF)
The Fourier Transform of the impulse response, called the transfer function of the system, indicates the effects of the system in the frequency domain \([49, 61, 62]\). In particular, the transfer function \(F(\xi, \eta; u, v)\) of the illuminating system is useful to express the time-varying phasor \(A'_o(\xi, \eta; u, v; t)\) as follows \([60]\):

\[
A'_o(\xi, \eta; u, v; t) = F(\xi, \eta; u, v)A_s(u, v; t - \delta_1)t_o(\xi, \eta),
\]

where \(\delta_1\) is the time delay produced by the separation between the source plane \((u, v)\) and the object plane \((\xi, \eta)\). \(t_o\) is the transmittance amplitude of the object. The final propagation reaching the image plane can be written as \([63]\):

\[
A_i(x, y; u, v, t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} K(x, y; \xi, \eta)A'_o(\xi, \eta; u, v; t - \delta_2)d\xi d\eta,
\]

where \(K(x, y; \xi, \eta)\) is the transfer function\(^1\) of the imaging system and \(\delta_2\) is the time delay produced by the separation between the object plane \((\xi, \eta)\) and the image plane \((x, y)\).

Substituting Equation 2.20 in Equation 2.21:

\[
A_i(x, y; u, v, t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} K(x, y; \xi, \eta)F(\xi, \eta; u, v)A_s(u, v; t - \delta_1 - \delta_2)t_o(\xi, \eta)d\xi d\eta. \tag{2.22}
\]

Therefore, the intensity produced by a source point in the image plane can be expressed as:

\(^1K(x, y; \xi, \eta)\) is also called amplitude point-spread function
\[ I(x, y; u, v) = \langle |A(x, y; u, v, t)|^2 \rangle \]
\[ = \iiint_{-\infty}^{+\infty} K(x, y; \xi_1, \eta_1)K^*(x, y; \xi_2, \eta_2)F(\xi_1, \eta_1; u, v)F^*(\xi_2, \eta_2; u, v) \]
\[ \langle A_s(u, v; t - \delta_1 - \delta_2)A_s^*(u, v; t - \delta'_1 - \delta'_2) \rangle t_o(\xi_1, \eta_1)t_o^*(\xi_2, \eta_2)d\xi_1d\eta_1d\xi_2d\eta_2 \] (2.23)

Under monochromatic conditions, the time delay differences satisfies that:
\[ |\delta_1 + \delta_2 - \delta'_1 - \delta'_2| \ll \tau_c, \] (2.24)

where \( \tau_c \) is the coherence time [43]. The previous relation causes that the time-average quantity in Equation 2.23 reduces to the intensity \( I_s(u, v) \) of the source. In order to obtained the intensity produced by all the points of the source, an integration over the source is done:

\[ I(x) = \iiint_{-\infty}^{+\infty} I_s(u, v)K(x, y; \xi_1, \eta_1)K^*(x, y; \xi_2, \eta_2) \]
\[ F(\xi_1, \eta_1; u, v)F^*(\xi_2, \eta_2; u, v)t_o(\xi_1, \eta_1)t_o^*(\xi_2, \eta_2)d\xi_1d\eta_1d\xi_2d\eta_2dudv \] (2.25)

knowing \( I_s, F, K \) and \( t_o \), it is possible to compute the final intensity distribution in the image plane.

### 2.3.6 Hopkins Assumption

H. H. Hopkins introduced a different approach for the computation of the intensity distribution in the image plane [51][64]. The implicit integration over the source is omitted and the effects of the illumination source are represented by a mutual intensity function describing the illumination incident on the object. In general, the mutual intensity function can be written as [38]:
\[ J(r_1, r_2) = \langle U(r_1)U(r_2) \rangle, \] (2.26)

where \( U(r_1) \) and \( U(r_2) \) represent the complex amplitude distribution of the optical wave in the spatial coordinates \( r_1 \) and \( r_2 \).

The time-varying phasor \( A_i(x, y; t) \) of the light arriving to the image plane can be represented in terms of the time-varying phasor \( A_o(\xi, \eta; t) \) of the light incident on the object plane by [60]:

\[ A_i(x, y; t) = A_o(\xi, \eta; t) \]

\[ K(x, y; \xi, \eta) = \int_{-\infty}^{+\infty} F(x, y; \xi, \eta; u, v)F^*(x, y; \xi, \eta; u, v)du dv \]

\[ t_0(\xi, \eta) = \frac{1}{\tau_c} \]
\[ A_i(x, y; t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} K(x, y; \xi, \eta)t_o(\xi, \eta)A_o(\xi, \eta; t-\delta)d\xi d\eta, \tag{2.27} \]

where \( K(x, y; \xi, \eta) \) is the transfer function of the imaging system and \( \delta \) is the time delay between the object and image plane. The intensity in the image plane is obtained like follows:

\[ I(x, y) = \langle | A_i(x, y; t) |^2 \rangle = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} K(x, y; \xi_1, \eta_1)K^*(x, y; \xi_2, \eta_2)t_o(\xi_1, \eta_1)t_o^*(\xi_2, \eta_2) \]

\[ \langle A_o(\xi_1, \eta_1; t-\delta_1)A_o^*(\xi_2, \eta_2; t-\delta_2) \rangle d\xi_1 d\eta_1 d\xi_2 d\eta_2, \tag{2.28} \]

under the monochromatic assumption, the following relation is fulfilled:

\[ | \delta_1 - \delta_2 | \ll \tau_c, \tag{2.29} \]

and the time-average in the Equation 2.28 is the mutual intensity distribution incident on the object.

\[ J_o(\xi_1, \eta_1; \xi_1, \eta_1) = \langle A_o(\xi_1, \eta_1; t-\delta_1)A_o^*(\xi_2, \eta_2; t-\delta_2) \rangle. \tag{2.30} \]

Then, the final intensity distribution can be written as:

\[ I(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} K(x, y; \xi_1, \eta_1)K^*(x, y; \xi_2, \eta_2)t_o(\xi_1, \eta_1)t_o^*(\xi_2, \eta_2) \]

\[ J_o(\xi_1, \eta_1; \xi_1, \eta_1)d\xi_1 d\eta_1 d\xi_2 d\eta_2. \tag{2.31} \]

The previous equation is referred to as the Hopkins formulation. This formalism is often applied for the numerical solutions to the image formation in lithographic software. Comparing the expressions for the calculation of the image intensity obtained by Abbe (Equation 2.25) and Hopkins (Equation 2.31), six integrals are explicitly written in the first case and just 4 in the second case. However, the mutual intensity distribution \( J_o \) requires the computation of 4 integrals that are reduced to 2 in the case of incoherent illumination. That is, the same number of integrals have to be computed in the Abbe and the Hopkins formulations.
2.3.7 Polarization

The propagating light is a transversal wave, which means that the electric and magnetic field vectors are orthogonal to the propagation direction. The direction in which the electric field is oscillating in space and time, gives the polarization of the electromagnetic wave. This is an important consideration since the amount of interference between two fields is dictated by the amount of polarization of the fields. If they are oscillating orthogonal to each other they do not interfere. Otherwise, if they are oscillating parallel to each other, interference takes place.

The polarization state can be classified according to the direction of the electric field with relation to the incidence plane of illumination. (Transverse Electric) polarized light is characterized by its electric field being in a plane perpendicular (i.e., transverse) to the incidence plane (Figure 2.18(a)). TM (Transverse Magnetic) polarized light refers to a polarization state where the magnetic field is in a plane perpendicular to the incidence plane (Figure 2.18(b)).

![Plane of incidence with illumination, TE polarization](image1.png)

**Figure 2.18:** Representation of the (a) TE- and (b) TM-polarization states.

The simplest type of polarization is called linear polarization, shown in Figure 2.19. At an instant in time, the electric field $E$ is always pointing in the same direction (in this case positive and negative directions in y-axis).

Other polarization states are represented in the Figure 2.20. When circular polarized, the electric field rotates following the path of a circle keeping a constant magnitude (Figure 2.20(b)) at a given point in space. The elliptical polarization refers to a state between the linear and the circular polarization state, where the electric field follows the path of an ellipse (Figure 2.20(c)) at a given point in space. The unpolarized state refers to an electric field which polarization state is not determined at a given point in space (Figure 2.20(d)). The polarization state for the resulting superposition of random polarized waves can be represented by a statistical average.

The degree of linear polarization (DoP) is a measure to report the amount of polarized light in an electromagnetic wave. If the electric field is written in term of its components...
Figure 2.19: Representation of an electric field linearly polarized. The k vector point in the direction of the propagation of the wave

Figure 2.20: Representation of an electric field linearly polarized. The k vector point in the direction of the propagation of the wave

$E_x$ and $E_y$, the degree of linear polarization can be written as [65]:

$$\text{DoP} = \frac{I_x - I_y}{I_x + I_y},$$

(2.32)

where $I_x$ and $I_y$ are the intensities corresponding to the component $E_x$ and $E_y$ respectively. If the light is completely polarized DoP=1, for unpolarized light DoP=0. If the light is circularly polarized, the DoP can be defined similarly as in Equation 2.32. Unpolarized light can be described as an incoherent superposition of two orthogonal polarization states [66]. When one of the two dominates, the total effect is partially polarized light.

In general, partially polarized light can be described as a superposition of a completely unpolarized component, and a completely polarized one:

$$P_{\text{total}} = P_{\text{pol}} + P_{\text{unpol}}$$

(2.33)
The polarization state of the illumination arriving at the photomask has a high impact on the image formation in the lithographic device. The contrast of the mask image depends on the orientation of the incoming polarization [2]. At high NA, the consideration of the polarization becomes critical. In Figure 2.21, the contrast image for vertical lines and spaces of 80 nm linewidth is evaluated for different polarization states. The TE polarization produces the best image contrast, this polarization is parallel to the orientation of the features. When the polarization angle $\alpha$ moves towards the TM polarization (perpendicular to the lines and spaces), the image contrast decreases drastically for high NA values. Then, features with determined orientations require illumination with specific polarization to be properly imaged. In practice, masks are composed of features oriented in different directions. Therefore, the illumination source has to provide different polarization directions at the same time (unpolarized light) to assure the proper imaging of all the features in the mask.

![Figure 2.21: Polarization effects at high NAs. Plot from [2].](image)

Several authors have studied the impact of the polarization phenomenon at high NAs for different types of masks [67–69]. It has been observed that the phase and the amplitude of the mask spectrum are polarization-dependent. At the same time, binary (CoG) and AttPSM masks have been observed to have a wire-grid polarizer behavior \(^1\). The pitch on the mask provides a DoP modulation for the 0th, -1st and 1st diffraction orders.

The TE and TM polarization states can be used to formulate the image formation with unpolarized light. Each of the TE and TM polarization components associated to each point in the source, can be treated separately and propagated through the illumination and imaging optics [70]. The two mutually incoherent images arising from the two

\[^1\text{A wire-grid polarizer is capable of converting an unpolarized beam into one with a single linear polarization. The polarization components that are not perpendicular to the orientation of the wires are blocked (absorbed and reflected).}\]
mutually incoherent source polarizations are then averaged to result in the final image for an unpolarized light source [71].

2.4 Image Analysis

The critical dimension (CD) control constitutes one of the most important aspects in the measurement of the lithographic process performance. In this section some metrics that express CD control [58] are explained.

2.4.1 Critical Dimension (CD)

The CD is defined as the target feature size to be obtained with a lithographic imaging system. Figure 2.22 illustrates the definition of the CD. At a given threshold value \( I_{th} \) and focus position, the width of the image is measured. For the purpose of this work, the resist acts as an intensity threshold detector. This threshold value depends on the chemical composition of the resist. Below the threshold, the resist can not “sense” the image.

![Image of critical dimension (CD) definition](image)

Figure 2.22: Critical dimension (CD) definition: The width of the image at a given threshold value.

2.4.2 Image log-slope (ILS)

A measure of the contrast of the image can be given in terms of the steepness of the image in the transition from bright to dark(Figure 2.23). This metric is called the image log-slope (ILS) and is defined by [58]:

\[
ILS = \frac{1}{I} \frac{dI}{dx} = \frac{d \ln(I)}{dx}.
\]  

(2.34)
the ILS is measured in a determined edge in the image. If a feature size is $w$ a normalized image log-slope can be formulated:

$$NILS = w \frac{d \ln(I)}{dx}.$$  \hspace{1cm} (2.35)

![Diagram of Mask and Image with ILS](Image)

Figure 2.23: Image Log-Slope as a metric to evaluate image quality in a lithographic process.

**2.4.3 Depth of Focus**

The depth of focus (DoF) indicates the range of focus variation that can be tolerated, before the image resolution and contrast decrease beyond the stated specifications. Images observed in planes out of focus become blurred and loose definition. This quantity is closely related to the resolution criteria given by Rayleigh and is referred to as the Rayleigh depth of focus [72]:

$$DoF = k_2 \frac{\lambda}{NA^2},$$  \hspace{1cm} (2.36)

where $k_2$ is another process-dependent factor (similar to $k_1$) with values typically in the range of 0.4 to 1.0 [73]. $k_2$ is usually experimentally determined. The DoF depends on the size of the features to be imaged, normally it is specified for the smallest features. At large NA values, Equation 2.36 does not properly predict the DoF observed in the lithographic process. Therefore, a more accurate DoF criterion is used [74]:

$$DoF = k_3 \frac{\lambda_n}{\sin^2(\theta_n/2)},$$  \hspace{1cm} (2.37)

where $k_3$ is a technology factor and $\lambda_n$ is the wavelength in the imaging medium. The angle $\theta_n$ is the emerging angle from the lens corresponding to one pitch on the mask. The DoF of the optical system must more or less match the thickness of the photoresist to
provide good image quality throughout the resist thickness. The process window (PW) is optimized to assure the largest possible DoF and threshold latitude (see Section 2.4.4).

### 2.4.4 Process Window

The process window (PW) represents a collection of process parameters that allow the printing of the mask pattern under certain target conditions. The lithographic process window is typically defined as the plot of exposure dose vs. focus positions to control critical dimension (CD) variation within a range of ±10%. The maximal inscribed rectangle (or ellipse) in the plots gives the process window [75]. If the process window analysis is performed using the aerial image, the threshold intensity instead of the exposure dose may be used. The metric can go beyond the CD (e.g., sidewall angle) as can the process variables (e.g., overlay, temperature). The final process window for a design is given by the overlap of process windows for each target CD existing in the design [76][77]. In Figure 2.24 the features of a typical process window can be observed. The best focus position and best threshold value are those located at the center of the rectangle. The process window has 3 curves. The upper curve represents a CD variation of 10%. The middle curve represents the CD with no variation and the bottom curve represents a CD variation of -10%. That means that for a lithographic process performed under specific conditions, the CD has a tolerance of 10%.

![Process Window Diagram](image)

**Figure 2.24:** Process window for a periodic contact holes layout (AttPSM), 62 nm size. ArFi, CD of 62 nm, CQUAD illumination, unpolarized light. The central curve represents the process for the required CD, upper curve shows the process with 10% deviation from the CD and bottom curve -10% of deviation from the CD.

The large area of the inscribed rectangle in Figure 2.24 assures a high tolerance of the process in terms of threshold intensity and focus position. The threshold required to
produce the proper CD is denominated Threshold-to-Size.

2.5 Optical Proximity Correction (OPC)

Resolution Enhancement Technologies (RET) are used to compensate for the limitations presented in the lithographic process. For instance, RET are used to improve the contrast and resolution of the image and to improve the lithographic PW. The different mask alternatives (PSM), OAI illumination and optical proximity correction (OPC) techniques are known as RET.

The work for the lithographers would be much more easier (but less interesting), if the mask could go just as the way it is directly to the semiconductor fabrication process. However, the nonlinear nature of a high-resolution lithography process assures that what is observed on the mask is not exactly the same that is obtained in the photoresist.

The proximity effects are the variations of the CD or distortion of the shape of a feature as a function of the proximity to other nearby features. Optical proximity effects refer to those proximity effects that are presented during the optical image formation. In general, these effects can not be always attributed to an optical phenomenon. Several authors contributed to the analysis and correction of the proximity effects in optical lithography [78–84].

One of the simplest OPC correction consist of changing the size of the mask feature from its initial dimension depending on the position of the nearest features. This technique is known as mask biasing and is illustrated in Figure 2.25(a). Alternatively, a different type of OPC is applied to isolated features to simulate a dense environment as shown in Figure 2.25(b). The Scattering Bars or Assist Features used in this process are sub-resolution geometries, that are inserted into available space near the main feature to improve image contrast and resolution through focus. These features do not print on the wafer because all the high spatial frequencies are filtered out by the pupil of the optical system. Modifications to the original mask pattern with the addition of geometries such as Hammer Heads and serifs are illustrated in Figures 2.25(a) and 2.25(b). These features can compensate for corner rounding and line-end shortening.

The designed mask used in the lithographic tool to obtain the desired features, with the correct CD and shape, implies an optimization process that involves the modeling and computation of the complete optical lithographic process. A design of the mask is presented, then the diffraction spectrum is computed. Next, the aerial image formed by the projection system is obtained. This process is repeated several times, until the image formation produced by the designed mask matches the target requirements within the range of tolerance. Such optimization task requires a large number of highly accurate image simulations. Hence, mask and imaging models that preserve the accuracy but speed
Figure 2.25: Optical proximity corrections (OPC) techniques. (a) Biasing, (b) Assisted features, (c) Hammer head and (d) Serif.

up the computation time are required.

Next chapter is devoted to the modeling aspect of the mask diffraction within the image computation and the formulation of compact mask models.
In the earlier years of lithography simulation, where mask features were comparably large to the wavelength of light, the scalar diffraction theory (Kirchhoff approach) yielded sufficiently accurate results in the computation of the diffraction spectrum. This simplified theory of how light behaves in the vicinity of an obstacle was good enough to describe the diffraction spectrum resulting from the mask. The race of fabricating smaller chips in the semiconductors technology has brought as a consequence the continuous scaling to smaller features sizes on the mask. The scalar theory then becomes invalid and fails to fully describe the physical effects that occur when the features on the mask are comparable to or even smaller than the wavelength of the light \[85–87\].

The vectorial theory of the electromagnetic fields considers the polarization state of the incoming light and the coupling of the fields at the interfaces of the medium where the light is propagating. The rigorous description of the light diffraction from the mask in optical lithography can be accomplished using several well known methods such as Finite-difference time-domain (FDTD) [6] or Rigorous coupled-wave analysis (RCWA) [7] among others, to solve Maxwell’s equations for the propagation of light. The additional effort to achieve a highly accurate description of the diffracted field is obtained at a huge computational cost.

The so-called compact mask models\[10][11] stand as an alternative to obtain reasonably good results compared to rigorous models with the advantage of having a higher efficiency. In the formulation of a compact mask model, physical effects are modeled in order to describe the phenomena predicted by a rigorous approach. However, an optimization process is required to calibrate such compact models. Every compact mask model depends on some variables that will be adjusted through an optimization process where the merit function may depend on the error of the aerial or bulk image, the error of the diffraction spectrum, the Process Window values or the CD measurements, all together or just a combination of some of them. Precisely, accuracy of a compact mask model depends on how well the optimization variables and the merit function are chosen.
3.1 Rigorous Mask Models

Rigorous mask models are based in the rigorous computation of the electromagnetic field (EMF) as it propagates through the mask. (EMF) modeling takes the complete information of the topography of the mask and its material properties into account to solve Maxwell’s equations for the three-dimensional space. In this section, the basis of two techniques used to rigorously compute the diffraction spectrum from the mask are explained. The FDTD method is a traditional numerical algorithm to solve the Maxwell’s Equations. There are several established and optimized routines to perform FDTD computations. For many years up to now, FDTD has been one of the most popular techniques for the solution of EMF problems. The RCWA method is a frequency-domain algorithm. That is, the propagating modes of the light are computed instead of the time-space dependent field distribution. This leads to significantly shorter computation times compared to the FDTD method. An adaptation and improvement of the RCWA –called the Waveguide method [88]– has been implemented by the Lithography Simulation Group at the Fraunhofer Institute IISB. This algorithm is used in this work as the reference model to compare the proposed models.

3.1.1 Finite-Difference Time-Domain (FDTD)

The well-known Maxwell’s equations in differential form describe the relationship between electric and magnetic fields in the space and time:

\[ \nabla \cdot D = \rho \]  
\[ \nabla \cdot B = 0 \]  
\[ \nabla \times E = -\frac{\partial B}{\partial t} \]  
\[ \nabla \times H = J + \frac{\partial D}{\partial t} \]

Where \( E \) is the electric field, \( B \) the magnetic flux density, \( D = \varepsilon E \) the electric displacement and \( H = B/\mu \) the magnetic field strength. Equations 3.1 and 3.2 are denominated the Gauss law for electrostatic and magnetostatic respectively. The Gauss law for electrostatic states that for the generation of an electric field there must exist an electric charge causing this field. In the case of magnetostatic, the Gauss law can be interpreted as the absence of magnetic mono-poles. Equations 3.3 and 3.4 are known as the Faraday’s law of induction and the Ampère’s law with Maxwell’s correction respectively. The Faraday’s law of induction and the Ampère’s law describe how an electric field can generate a magnetic field, and vice versa. The Faraday’s law of induction states that any variation in time
of the magnetic flux density produces an electric field. On the other hand, the Ampère’s law explains the nature of the magnetic field strength. The presence of a current density $\mathbf{J}$ (flux of electric charge) generates magnetic field as well as any variation in time of the electric field.

Light is defined as a dual entity, being at the same time an electromagnetic wave and a particle. For our purposes we are entirely interested in the electromagnetic wave description given by the solution of the Maxwell’s equations for the propagation of light. The solution of Maxwell's equations means the computation of the fields using the material parameters and some initial and boundary conditions. In the computational domain there are neither current densities nor free charges, that means $\mathbf{J} = 0$ and $\rho = 0$. With these considerations the Maxwell’s equations reduce to:

\[
\begin{align*}
\nabla \cdot \mathbf{D} &= 0 \quad (3.5) \\
\nabla \cdot \mathbf{B} &= 0 \quad (3.6) \\
\n\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \quad (3.7) \\
\n\nabla \times \mathbf{H} &= \frac{\partial \mathbf{D}}{\partial t} \quad (3.8)
\end{align*}
\]

if Equations 3.5 and 3.6 are satisfied by the initial conditions, then they are valid automatically at any further time instant. Thus the system of equations to solve consist of Equations 3.7 and 3.8. The analytical solution of this system is not straightforward, that is why numerical methods are generally applied. The first and still extensively applied method is the Finite Difference Time Domain (FDTD) method constructed by K. Yee in 1966[6]. The method defines a rectangular mesh with the step-sizes $\Delta x$, $\Delta y$ and $\Delta z$ for the electric field and another staggered grid with $\Delta x/2$, $\Delta y/2$ and $\Delta z/2$ for the magnetic field in the computational domain. The building blocks of this mesh are the so-called Yee-cells (see Figure 3.1).

The FDTD method employs finite differences as approximations to both the spatial and temporal derivatives that appear in Maxwell’s equations (specifically Amperes and Faradays laws). Considering the Taylor series expansions of the function $f(x)$ expanded about the point $x_0$ with an offset of $\Delta x/2$, the central-difference approximation is defined by:

\[
\left. \frac{df(x)}{dx} \right|_{x=x_0} = \frac{f(x_0 + \Delta x/2) - f(x_0 - \Delta x/2)}{\Delta x} \quad (3.9)
\]

Yee’s method uses Equation 3.9 to express the derivatives of the electric and magnetic fields in time and space. From the Equations 3.7 and 3.8 is straightforward to derive the set of scalar equations to be solved by the FDTD method. As an illustrative example,
let us consider a one-dimensional space where there are only variations in the x direction. Assuming that the electric field only has a z component, the resulting scalar equations can be written as:

\[ \frac{\mu}{\Delta t} \frac{\partial H_y}{\partial t} = \frac{\partial E_z}{\partial x} \]  
(3.10)

\[ \frac{\mu}{\Delta t} \frac{\partial E_z}{\partial t} = \frac{\partial H_y}{\partial x} \]  
(3.11)

The next step is to replace the derivatives in terms of finite differences. To do this, space and time need to be discretized. The following notation will be used to indicate the location where the fields are sampled in space and time:

\[ E_z(x, t) = E_z(m\Delta x, q\Delta t) = E_{z}^{q}[m] \]  
(3.12)

\[ H_y(x, t) = H_y(m\Delta x, q\Delta t) = H_{y}^{q}[m] \]  
(3.13)

where \( \Delta x \) and \( \Delta t \) are the sample sizes of space and time respectively. \( m \) and \( q \) represent the discrete steps in space and time. Using this notation, the solution to the scalar system of equations provided by the FDTD method is as follows:

\[ H_{y}^{q+\frac{1}{2}} \left[ m + \frac{1}{2} \right] = H_{y}^{q-\frac{1}{2}} \left[ m + \frac{1}{2} \right] + \frac{\Delta t}{\mu \Delta x} (E_{z}^{q}[m + 1] - E_{z}^{q}[m]) \]  
(3.14)
\[ E_{z}^{q+1}[m] = E_{z}^{q}[m] + \frac{\Delta t}{\epsilon \Delta x} \left( H_{y}^{q+\frac{1}{2}} \left[ m + \frac{1}{2} \right] - H_{y}^{q-\frac{1}{2}} \left[ m - \frac{1}{2} \right] \right) \] (3.15)

Equations 3.14 and 3.15 are the description of the future values of \( H_{y} \) and \( E_{z} \) respectively. The future values depend on the past values of the magnetic and electric fields. The entire notation and scalar equations can be generalized to electric and magnetic fields varying in the \( y \) and \( z \) directions.

### 3.1.2 Waveguide Method

The foundations of the Waveguide method were given by Burckhardt [89] in 1966 for the rigorous analysis of the diffraction at a hologram formed in a "thick" photographic emulsion. Further contributions to the method were done by Yuan [90], Moharam [7], Nyyssonen [91], Lucas [92–94] and many other authors [4] [88]. Based on the real mask 3D geometry, the simulated area is divided in layers which are homogeneous in \( z \)-direction. This is called the waveguide assumption (see layers in Figure 3.2).

![Figure 3.2: Slicing of the mask in the Waveguide method. The blue color represents the absorber and the purple represents the substrate. The complete simulated area is divided in layers that are homogeneous in \( z \)-direction.](image)

The Waveguide method is a frequency-domain method where the electromagnetic field is expressed as a superposition of some basis functions. Generally, the electromagnetic fields and dielectric function of all layers are described by Fourier series. Inside each layer, electromagnetic field distributions or modes are obtained from the solution of the corresponding eigenvalue problem. The application of appropriate boundary conditions
to couple the resulting fields for all layers produces the reflected and transmitted plane waves at the top and bottom side of the mask.

All Fourier series are truncated according to the number \( m \) of modes taken into account, that is the number of reflected and transmitted plane waves in Figure 3.2. Considering the description from above, the Waveguide method can be characterized by the following properties:

- Good convergence.
- The mask geometry usually corresponds to the required slicing.
- The memory consumption and the computing time is quadratic both in the number number \( m \) of modes in x- and y-direction. Additionally, it increases linearly with the number of slices with nonconstant optical properties along x and y.
- Because of the periodic boundary conditions in lateral direction, the mask is always regarded as a periodic structure.
- Only stationary fields are of interest, no simulations in the time domain required since the Waveguide method is operating in the frequency domain. This turns out in fast simulations compared to other rigorous methods such as the FDTD.

### 3.2 Mask-Topography Induced Effects

The scalar mask model ignores diffraction and polarization effects shown by the rigorous model. However, it provides surprisingly accurate results when the dimension from mask features are much larger than the wavelength \((kd \gg 1\text{, with } k\text{ being the wavevector and } d\text{ the aperture size})\) and the diffraction is evaluated at distances several wavelengths from the aperture \((z > 2d^2/\lambda)\). Another reason for the relative good performance of the scalar model is the filtering of high spatial frequencies by the projection pupil. Evanescent and high orders, which do not pass the pupil, are not important for the description of the image formation.

The rigorous modeling of the diffraction of a real 3D mask with feature sizes comparable with the wavelength brings as a consequence the observation of physical effects neglected before by the scalar model [95–97]. Some of these mask-topography-induced effects are:

- Polarization dependence due to the different boundary conditions for the electric and magnetic fields [3]: In Figure 3.3, the diffraction efficiency from a MoSi-type AttPSM with a pattern of lines and spaces is shown. The scalar model (black curves)
Figure 3.3: Diffraction efficiency effects for a standard MoSi-type AttPSM with lines/spaces for vertical incidence. NA=1.2, ArFi \( (n_{\text{immersion}} = 1.44) \). Dipole illumination with \( y \)-polarization, \( \sigma = 0.8/0.96 \) and opening angle=40°. Analysis for the 0th (upper figure) and the 1st (bottom figure) diffraction orders. Rigorous simulation for (blue) \( y \)-polarization, (green) \( x \)-polarization and (black) Kirchhoff model. The diffraction efficiencies depend on the state of polarization. Plot from [3].

predicts a sharp cut-off where certain diffraction orders disappear. Additionally, a polarization independent and constant diffraction efficiency \( \eta \) is observed with the scalar model. On the other hand, rigorous mask modeling (blue curves) predicts non-constant and polarization dependent diffraction efficiency \( \eta \).

- Phase effects: Rigorous mask modeling predicts mask induced phase effects which can be interpreted as aberrations in the aerial image [98]. In Figure 3.4, a polarization dependent behavior in the phase difference is shown. When the linewidth is
smaller than 150 nm this difference starts to become important.

- Transmission imbalance between etched and unetched openings in the glass substrate [68, 99, 100]: This effect is specially critical for alternating phase shift masks (AltPSM) since they are based on the modulation of both amplitude and phase of the electromagnetic field propagating through them.

- Cross-talk[71, 101, 102]: Energy cross-coupling between neighboring apertures in (AltPSM).

Consequently, rigorous electromagnetic field simulations have become necessary in the spectrum computation of 3D masks but extremely resource and time-consuming. The high CPU-time consumption limits the rigorous approaches and raises the demand for accurate and fast but still simple physical and numerical models.

### 3.3 Compact Mask Models from Literature

The intention of a compact mask model is to reproduce with a high degree of accuracy the results obtained with a rigorous method but with the advantage of a high computational performance. Compact mask models are helpful in optical proximity correction techniques and in the inspection and analysis of mask-induced effects.
3.3.1 QUASI-3D: Domain Decomposition Method (DDM)

The simulation of large mask areas with a rigorous method is non-practical for lithography purposes due to the large computational time. The domain decomposition method (DDM) simplifies the three dimensional computation by converting it into a 2D and 1D tasks. Excellent accuracy in the calculation of the near scattered fields of more than 99% (in a normalized mean square error sense) compared with the fully rigorous mask model has been achieved with this technique. The DDM was originally proposed in 2001 by Adam [101]. In 2008, Shao [4] introduced an optimized electromagnetic field (EMF) solver based on the Waveguide method with a decomposition technique. This algorithm is used in this thesis for the comparisons with the rigorous model.

The idea behind the decomposition technique is to separate diffraction effects from mask edges along x- and y-direction. A full three dimensional electromagnetic field simulation is split into several two dimensional and one dimensional simulations. In Figure 3.5, the domain decomposition method (DDM) applied in 2D layouts (3D masks) is shown. The 3D mask is decomposed into three cuts in x-direction (cut 1, cut 2 and cut 3) and 3 cuts in y-direction (cut 4, cut 5 and cut 6). The electromagnetic fields in all cuts are calculated with the Waveguide method under the assumption that the dielectric properties of each cut vary only in y-direction (for cuts 1 to 3) or vary only in x-direction (for cuts 4 to 6). Derived from this assumption, the waveguide computation of each cut is converted from 3D to 2D, which contributes to the speed up of the required numerical simulations. The transmissions of the mask at all cross-areas of the cuts (1D cuts) are calculated with the Waveguide method. Finally, the composition of the diffraction spectra of all the cuts and of the 1D cuts transmissions produces the diffraction spectrum of the 3D mask.

Figure 3.5: Top view of a 3D mask with a post absorber (light blue) in the center. The 3D mask is decomposed into two dimensional cuts (cut 1-6) with constant dielectric properties in x- and y-direction, respectively. Plot from [4].
3.3.2 Boundary Layer Model (BLM)

The Kirchhoff approximation neglects the polarization and edge diffraction from the real (thick) mask. The thick mask is replaced by an ideal transmission function. In order to obtain the object field (near field) on the exit surface of the mask, the incoming field is multiplied by the transmission function. The far field is then obtained by a Fourier Transformation of that object field.

The Boundary Layer model (BLM) replaces the thick mask with a modified thin mask (Figure 3.6) to emulate the edge effects predicted by rigorous models. Every feature in the mask is surrounded width a semitransparent material of certain width, transmission and phase. The width of the boundary layer accounts for intensity corrections, while the imaginary transmission coefficient corrects phase deviations of the scalar case. This model was originally proposed by Tirapu-Azpiroz [12] to reproduce the predictions of rigorous models using the Hopkins assumption for the vector imaging formulation.

![Mask transmission](image1)

(a) Thin mask.

![Mask transmission](image2)

(b) Modified mask.

Figure 3.6: The modified thin mask emulates the physical effects predicted by rigorous models.

The BLM has its theoretically foundation on the well-established scalar diffraction theory. The proposed model, consists of a sophisticated version of Kirchhoff approximation, simply adding a boundary layer to every edge in the mask. The BLM accurately describes the thick mask effects, incorporating effects of electromagnetic coupling due to the high numerical aperture, and accurately compensates for phase errors even at planes out of focus[103]. The contribution of this formulation can be observed in the improvement of the accuracy of the aerial image computation in photolithography simulations at a reasonable computational cost.
3.3.3 Pulses model (PM)

Following the proposal from previous publications [13][104][105], the edge diffraction effects described by the rigorous models can be mimicked by means of the scalar diffraction theory with a special modification. This is the foundation of the Pulses model (PM) where the transmittance function \( T(x) \) of the thin mask is transformed by adding perturbations in the form of \( \delta \)-functions,

\[
T_{\text{pert}}(x) = A e^{i\varphi} [\delta(x - d/2) + \delta(x + d/2)],
\]

(3.16)

where \( A e^{i\varphi} \) is the value of the edge pulse and \( d \) is the duty cycle, defined as:

\[
d = \frac{C D}{\text{pitch}}
\]

(3.17)

Figure 3.7: Modification of the thin mask transmission by adding perturbations to the edges in form of delta functions.

The PM describes physical effects that are originated by the features in the mask (spatial domain), but its implementation is can be directly performed in the scalar diffraction spectrum (frequency domain). The complex transmission of the modified thin mask can be written as:

\[
T'(x) = T(x) + T_{\text{pert}}(x).
\]

(3.18)

Where \( T(x) \) is the transmission of the thin mask:

\[
T(x) = \begin{cases} 
0, & -d/2 < x < -d/2 \\
T_a, & -1/2 < x < d/2 \text{ and } d/2 < x < 1/2
\end{cases}
\]

(3.19)
Substituting Equation 3.16 in 3.18:

\[ T'(x) = T(x) + A e^{i\phi} [\delta(x - d/2) + \delta(x + d/2)], \]  

(3.20)

The Fourier transform of Equation 3.20 is given by:

\[ \mathcal{F}(T'(x); m) = \mathcal{F}(T(x)) + 2Ae^{i\phi}\cos(\pi md), \]  

(3.21)

where \( m \) represents the diffraction order. Equation 3.21 shows that the PM can be easily implemented in the frequency domain by adding the second term to the scalar diffraction spectrum of the thin mask.

### 3.4 Proposed Mask Models

The following sections will be dedicated to explain in detail the new mask models, which are proposed, implemented and investigated in this thesis. All three proposed mask models are directed to manipulate the mask diffraction spectrum to obtain results comparable with rigorous methods. These models were calibrated against the Waveguide method as can be seen in the next chapter where the reader can find all the simulation results.

#### 3.4.1 Pupil Filtering Model (PFM)

The optical projection system is composed of several refracting elements. These elements modify the wavefront and the optical path of the incident illumination. For modeling and simulation purposes, all the effects of the imaging elements are represented by a transfer system with entrance and exit pupils. The finite extent of the entrance aperture can be associated to a pupil function defined as:

\[ P(f_x, f_y) = \begin{cases} 
1, & \sqrt{\sin(\theta_x)^2 + \sin(\theta_y)^2} \leq NA. \\
0, & \text{outside NA}. 
\end{cases} \]  

(3.22)

Where the frequency coordinates are denoted by \( f_x = \sin(\theta_x)/\lambda \) and \( f_y = \sin(\theta_y)/\lambda \). \( NA \) is the numerical aperture and \( \lambda \) is the wavelength. \( P(f_x, f_y) \), defined as in Equation 3.22, indicates that the light arriving to the pupil is ideally not modified in amplitude or phase. Figure 3.8 shows the representation of the pupil plane, its finite size affects the image formation since only a portion of the diffracted light is captured.

The pupil filtering model (PFM) has its foundations in the observations made of the masks-induced effects described by the rigorous EMF simulations[69][106]. At very high NA (>1.0) and extreme off-axis illumination angles, more pronounced mask topography effects can lead to polarization, phase and amplitude deviations between the thin mask.
approximation and the more rigorous models. The mask-induced phase effects can be interpreted as wavefront aberrations in the optical projection system. The description of the aberrations can be carried out through conventional wavefront description, specifically through Zernike polynomials.

The wavefront $W(\rho, \varphi)$ is expressed in terms of the Zernike polynomials, which are expressed in the polar coordinates of the pupil as follows[107, 108]:

$$W(\rho, \varphi) = \sum_{n}^{\infty} \sum_{m=-n}^{n} C_{nm} Z_{n}^{m}(\rho, \varphi),$$  (3.23)

The American National Standards Institute ANSI defines the Zernike polynomials as [109]:

$$Z_{n}^{m}(\rho, \varphi) = \begin{cases} N_{n}^{m} R_{n}^{m}(\rho) \cos(m \theta), & \text{for } m \geq 0. \\ -N_{n}^{m} R_{n}^{m}(\rho) \sin(m \theta), & \text{for } m < 0. \end{cases}$$ (3.24)

where $n$ is the radial order, $m$ is the azimuthal frequency. $N_{n}^{m}$ is the normalization factor, $R_{n}^{m}$ the radial component and $\cos(m \theta)$ and $\sin(m \theta)$ the azimuthal components. In Figure 3.9 the functional form of the first radial orders of the Zernike polynomials is presented. In Figure 3.10 the wave deformation corresponding to each radial order is illustrated.

The pupil function of the aberrated wavefront can be re-defined as:

$$P(\rho, \varphi) = \begin{cases} A(\rho, \varphi) e^{i2\pi W(\rho, \varphi)}, & \rho \leq NA. \\ 0, & \text{outside NA.} \end{cases}$$ (3.25)
Figure 3.9: Functional form of the Zernike polynomials for the first radial orders.

Figure 3.10: Wave deformation $W(\rho, \varphi)$ expressed in term of Zernike polynomials corresponding to $Z_i = 0.25$. Plot from Andreas Erdmann, Lectures in Lithography, FAU Erlangen Graduate School of Advanced Optical Technologies SAOT, 2012.
Where $A(\rho, \phi)$ represents the amplitude of the optical field inside the pupil and $W(\rho, \phi)$ the phase of the wavefront. $A(\rho, \phi)$ and $W(\rho, \phi)$ are expressed as follows:

$$A^{TE,TM}(\rho, \phi) = C_{1}^{TE,TM} + C_{4}^{TE,TM} (2\rho^2 - 1) + C_{9}^{TE,TM} (6\rho^4 - 6\rho^2 + 1), \quad (3.26)$$

$$W^{TE,TM}(\rho, \phi) = Z_{4}^{TE,TM} (2\rho^2 - 1) + Z_{5}^{TE,TM} (2\rho^2 \cos(2\phi)) + Z_{9}^{TE,TM} (6\rho^4 - 6\rho^2 + 1), \quad (3.27)$$

where the indices $TM$ and $TE$ denote the coefficients associated to the TM- and TE-polarization components. That is, the polarization effects are addressed in the PFM through the calibration of different coefficients for each polarization state. $C_{1}^{TE,TM}$, $C_{4}^{TE,TM}$ and $C_{9}^{TE,TM}$ correspond to the coefficients of the polynomial expressions in the amplitude. The importance of the functional form of the amplitude lies in the ability to enhance or weaken diffraction orders that contribute to the image formation, this is critical due to the consideration of off-axis illumination and its impact on the diffraction spectrum from the topographic mask.

The spherical aberration expressed in the polynomial $Z_{9}^{TE,TM} (6\rho^4 - 6\rho^2 + 1)$ is considered to be the principal cause of the shift in the best focus position through pitch[110]. However, in this study it has been found that defocus $Z_{4}^{TE,TM} (2\rho^2 - 1)$ and astigmatism $Z_{5}^{TE,TM} (2\rho^2 \cos(2\phi))$ aberrations may also contribute to describe the mask-induced effects reflected in the PWs analysis (see Sec 4.5.1).

The PFM emulates the mask topography effects to enhance the scalar diffraction model. The model introduces the mask-induced effects as aberration to the optical system in the vector image computation. The scalar spectrum is computed as usual and all the mask-topography effects are expressed as deformation of the wavefront in the pupil function.

The PFM also makes use of the Jones formalism to introduce the polarization effects in the image formation. In 1941, R.C. Jones introduces for the first time the mathematical formalism to analyze an optical system which is composed of retardation plates, partial polarizers, and plates with the ability of rotate the plane of polarization [111–116]. When light passes through any optically active element, the state of polarization and sometimes the intensity of the light are altered. Jones found the way to represent the effect of any optical element on the light as a linear operator acting upon the electric vector of the light wave. The operator is expressed in the convenient form of a two-by-two matrix called the *Jones matrix* [117]:

$$
\mathbf{J} = \begin{bmatrix} 
J_{xx} & J_{xy} \\
J_{yx} & J_{yy} 
\end{bmatrix} \quad (3.28)
$$

The four elements of the Jones Matrix completely describe the effects of a linear device on the state of polarization of the wave in such a way that the new state of polarization
can be represented by a lineal transformation of an initial state of polarization $\tilde{P}$:

$$\tilde{P}^\prime = \begin{bmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{bmatrix} \tilde{P}$$ (3.29)

It is important to point out that the state of polarization may be completely defined by stating the amplitudes and phases of the $x$ and $y$ components of the electric vector of the optical field. Therefore, Equation 3.29 is also applicable to describe states of the electric field before and after being transformed by an optically active element.

In order to take into account polarization dependent effects, two orthogonal polarization modes are considered, the TM and TE modes. The Jones matrix can express the total state of polarization in terms of two orthogonal polarization states as follows:

$$J = \begin{bmatrix} J_{xx} & 0 \\ 0 & J_{yy} \end{bmatrix}$$ (3.30)

Where the term $J_{xx}$ is associated with pure (TM) $x$-polarization and $J_{yy}$ with pure (TE) $y$-polarization. Using Equations 3.26 and 3.27 together with the described Jones-formalism, the scalar pupil function in Equation 3.25 can be generalized by a Jones pupil function:

$$J = \begin{bmatrix} A^{TM}(\rho, \phi)e^{i2\pi W^{TM}(\rho, \phi)} & 0 \\ 0 & A^{TE}(\rho, \phi)e^{i2\pi W^{TE}(\rho, \phi)} \end{bmatrix}$$ (3.31)

Each component of the matrix $J$ requires a calibration process of the corresponding amplitude $A(\rho, \phi)$ and phase $W(\rho, \phi)$ values in order to mimic the mask-induced effects. Afterwards, this effects are introduce to the vector imaging computation in order to improve the performance of the scalar model.

### 3.4.2 Spectrum Correction Model (SCM)

The spectrum correction model (SCM) can be considered as an extension of the (PFM). An important aspect of mask modeling is the consideration of spatial coherence. Modern lithographic methods use off-axis illumination (OAI) techniques (partial coherent illumination). The mask is illuminated by a spectrum of mutually incoherent plane waves.

Using the Hopkins assumption\[17\], the rigorously computed diffraction spectrum is evaluated only for vertical incidence. Diffraction orders of obliquely incident angles are approximated by shifting that spectrum accordingly. Without the Hopkins assumption, the diffraction spectrum of each illumination angle is calculated by rigorous EMF computation. This allows a high accuracy in terms of description of the diffracted field; however, the computational expense increases significantly. The proposed compact model, takes...
the scalar spectrum and corrects it according to comparisons with the respective rigorous spectra for representative incidence angles. In this way, a calibrated filter is produced per angle of illumination.

The SCM considers a set of functions expressed in terms the Zernike polynomials to modify the scalar spectrum. All calibrated functions are applied to the scalar spectrum to reproduce a vectorial spectrum with the corresponding components for each state of polarization necessary for the imaging formulation. In Figures 3.11 and 3.12 the SCM is illustrated. The spectrum modification is performed independently for all spectra corresponding to all relevant angles of illumination. This provides the flexibility to modify each spectrum according to a specific need.

Figure 3.11: Spectra observed from an ideal pupil without a modulating function. Spectrum corresponding to: (upper left) vertical illumination, (upper right;bottom left and right) other oblique illumination angles.

The spectrum correction model (SCM) modifies directly the scalar spectrum to compensate for the mask-induced effects. In this case, the Jones pupil contains the usual description of an entrance pupil without aberrations. The SCM performs a description of the aberration-like effects produced by the mask out the domain of the pupil, directly on the scalar spectrum.

The SCM differs from the PFM in the sense that the spectrum is modified right before
The spectrum modification is performed independently for all spectra corresponding to all relevant angles of illumination. The scalar spectrum is transformed with a calibrated function to emulate the rigorous spectrum corresponding to each angle of illumination.

entering to the pupil aperture. This allows to manipulate each spectrum corresponding to each illumination angle, while the pupil filtering is performed in the same functional way to all the spectra. Therefore, it can be expected that the correction of the diffraction spectra provides a higher flexibility to reproduce images, which are obtained with rigorous mask diffraction and imaging simulation without the Hopkins approach. On the other hand, the spectrum filtering includes more model parameters. This makes the model parameter calibration more difficult.

### 3.4.3 Artificial Neural Network (ANN) Approach

Up to now, the proposed models (PFM and SCM) modify the scalar spectrum with specific analytical functions described by Zernike expansions. Both models enable physical interpretation of the mask-induced effects. However, the PFM and the SCM are limited by the number of the Zernike orders used in the calibration. Additionaly, the number of parameters to optimize is increased in the SCM because of the consideration of independent filters per incident angle of illumination.

The ANN offers an alternative formulation to retrieve the characteristics of the rigorous spectrum. The combination of linear and non-linear behaviour of the learning process of the ANN, provides a higher flexibility for the modification of the scalar spectrum, which
does not rely on specific analytical functions. This section starts with a description of the used ANN. Afterwards, the application of these ANNs for the spectrum modification is explained.

An ANN is a processing algorithm that mimics the behavior of biological neural networks [118, 119]. It is composed of a large number of highly interconnected processing units (neurons) working together to accomplish specific tasks. A primary property of the neuron is that, in some basic, inexperienced state, it can be made to fire by incoming stimulation. In this analogy, learning consists of the control of the response to the stimulation by means of the modification of the connected system.

The ANN are typically organized in layers that are made up of a number of interconnected nodes. The first layer has input neurons, which communicate via synapses (connections) to one or more hidden layers. The hidden layers are linked to an output layer to produce a final response or output (see Figure 3.13). In the scope of this work, the input of the ANN is the scalar spectrum data and the output to be achieved is the rigorous diffraction spectrum.

![Figure 3.13: Representation of an ANN.](image)

The actual processing is done by the system of weighted connections. Each node contains an activation function that converts a neuron’s weighted input to its output.
activation. The ANN itself is defined by some form of learning process, which modifies the weight of the connections according to the input patterns that are presented to the ANN.

**Learning Process: Backpropagation of errors** There are numerous learning rules, the most widely used, and the one used in this study is the backward propagation of errors [120]. The procedure repeatedly adjusts the weights of the connections in the network to minimize a measure of the difference between the actual output of the system and the target output. As the algorithm’s name indicates, the errors propagate backwards from the output nodes to the inner nodes. During the backpropagation process the gradient of the error of the network regarding the network’s modifiable weights is calculated[121]. This gradient is used in a simple stochastic gradient descent algorithm to find weights that minimize the error. The backward propagation of errors usually allows quick convergence towards satisfactory local minima.

To illustrate the backpropagation algorithm, a three hidden layer ANN is used, as shown in Figure 3.14. The signal is propagating from left to right in the diagram. The hidden units perform two operations over the collection of inputs in the preceding layer. First a linear function over all the weighted connections and then a non-linear function to control the ”firing” of the neuron. The latter is called the activation function, in this case the sigmoid function \((1 + e^{-x})^{-1}\) tends to fire the neuron when the input \(x\) goes to \(\infty\) and shuts it down when it goes to \(-\infty\). Each function \(f_n\) is written following Equation 3.32.

\[
y_j = f_j(\text{net}) = \frac{1}{1 + e^{-\text{net}}}, \tag{3.32}
\]

\[
\text{net} = \sum_i w_{ji}x_i, \tag{3.33}
\]
where $i$ is the number of inputs to unit $j$, $j$ is the unit firing, $w_{ji}$ represents the weight of the corresponding connection from $i$ to $j$ and $y_j$ is the output of the units connected to $j$. At the end of the propagation of the signal, the final output is given by the function $f_6$:

$$f_6(x_6) = \frac{1}{1 + e^{-x_6}},$$

$$x_6 = w_{46} f_4(x_4) + w_{56} f_5(x_5),$$

$$x_5 = w_{15} f_1(x_1) + w_{25} f_2(x_2) + w_{35} f_3(x_3),$$

$$x_4 = w_{14} f_1(x_1) + w_{24} f_2(x_2) + w_{34} f_3(x_3),$$

$$x_3 = w_{13} x_1 + w_{23} x_2,$$

$$x_2 = w_{12} n_1 + w_{22} n_2,$$

$$x_1 = w_{11} n_1 + w_{21} n_2. \tag{3.34}$$

In the next step of the algorithm, the final output signal $y$ of the network is compared with the desired output value $z$ (the target). The difference $\delta = y - z$ is the error signal of the output layer. For the computation of the error signal $\delta$ of internal layers, the error signal is propagated back to all neurons in every step of the training.

Figure 3.15: Back propagation of the error $\delta$. The error propagation is from right to left in this diagram.

The weights’ coefficients $w_{ji}$ used to propagate errors back are equal to those used during the computation of the output value. Only the direction of data flow is changed (signals are propagated from output to inputs one after the other). This technique is used for all network layers.

The error $E$ of the actual output compared to the desired target is defined as

$$E = \frac{1}{2} \sum_c \sum_j (y_{c,j} - z_{c,j})^2 \tag{3.35}$$
where \( c \) is an index over the weighted connections entering to each unit, \( j \) is an index over the output units, \( y \) is the actual state of an output unit and \( z \) is the desired state. The error signal of the unit \( j \) in the output layer is

\[
\delta_j = \frac{\partial E}{\partial y_j} = y_j - z_j
\]  

(3.36)

The negative gradient for weight \( w_{ji} \) is defined as

\[
\Delta w_{ji} = -\frac{\partial E}{\partial w_{ji}}
\]  

(3.37)

The gradient descent method minimizes the error \( E \) by computing the negative partial derivative of \( E \) with respect to each weight in the network. This is the sum of the partial derivatives for each weighted connection in the network. For a weight \( w_{ij} \), the application of the chain rule results in the following derivative:

\[
\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial w_{ji}}
\]  

(3.38)

A second application of the chain rule over \( \frac{\partial y_j}{\partial w_{ij}} \) produces:

\[
\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial net} \frac{\partial net}{\partial w_{ji}}
\]  

(3.39)

Differentiating Equations 3.32 and 3.33 and substituting in Equation 3.39:

\[
\frac{\partial E}{\partial w_{ji}} = (y_j - z_j) y_j (1 - y_j) x_j
\]  

(3.40)

Equation 3.40 provides the updating of the weight \( w_{ji} \). Using gradient descent, one must chooses a learning rate, \( \alpha \). The change in weights after learning then is the product of the learning rate and the negative gradient

\[
\Delta w_{ji} = -\alpha [(y_j - z_j) y_j (1 - y_j) x_j]
\]  

(3.41)

the term \( y_j (1 - y_j) \) represents the derivative of the activation function that was specified as the sigmoid function. In general the gradient is written as

\[
\Delta w_{ji} = -\alpha [(y_j - z_j) f'(net) x_j]
\]  

(3.42)

For intermediate layers the expression is modified since the output of the hidden layers is unknown:

\[
\Delta w_{ji} = -\alpha \left[ \sum_k w_{ki} \delta_k f'(net) x_j \right]
\]  

(3.43)
According to the above formulation, the values of the signal errors for the ANN under consideration are expressed as follows:

\[ \delta = z - y, \]
\[ \delta_5 = w_{56}\delta, \]
\[ \delta_4 = w_{46}\delta, \]
\[ \delta_3 = w_{34}\delta_4 + w_{35}\delta_5, \]
\[ \delta_2 = w_{24}\delta_4 + w_{25}\delta_5, \]
\[ \delta_1 = w_{14}\delta_4 + w_{15}\delta_5, \]

(3.44)

The final step of the backpropagation algorithm is to update the weights. This is done by taking a gradient step and adding it to the actual weight as follows:

\[ w'_{ji} = w_{ji} - \alpha \left[ \sum_k w_{ki}\delta_k f'(net) x_j \right] \]

(3.45)

The learning process (updating of weights) continues its cycle (epochs) until the error between the current output and the target reaches a minimum or simply the termination criteria. However, there is not a real guarantee about the convergence, this may oscillate or reach a local minima.

The ANN approach has been extensively used in several mathematical, biological and physical scenarios such as pattern recognition, artificial intelligence and virtual reality, among others [122–125]. In this work, the ANN is introduced as a tool to retrieve the rigorous diffraction spectrum having as an input the scalar diffraction spectrum. The ANN is provided with the amplitude and the phase of the scalar spectrum of different mask layouts that will work as training patterns for the network. Afterwards, in a cross-validation step, the ANN is tested with patterns that have not been used in the training process to evaluate the performance and flexibility of the network. These patterns are “unknown” to the network but are expected to generate the response the network was trained for.

The data the ANN is fed with, can vary depending on the effects to be retrieved. For some cases, the phase and the amplitude of the scalar spectrum are enough information to obtain a good performance of the ANN. Without the Hopkins assumption or for complex mask geometries, additional input has to be provided to the ANN so that each entry has a unique signature during the learning process, meaning that preserving a training diversity helps the ANN not to become either over fed or over trained.
Chapter 4

Evaluation and Comparisons of Mask Models from Literature

In this thesis, the Waveguide method is used for the rigorous computation of the mask spectrum together with the Abbe imaging formulation without the Hopkins assumption (HA) [17] for the vector imaging computation. Specifically, this assumption refers to the diffraction spectrum calculation used by the Fraunhofer IISB lithography simulation software Dr.LiTHO [126]. Without the HA, each spectrum corresponding to a representative angle of illumination source is rigorously computed. This allows a high accuracy in terms of the description of the diffracted field. However, the computational expense increases significantly compared to image simulations with the HA, where the rigorously computed spectrum for the vertical incidence is shifted accordingly to each angle of illumination to account for the oblique illumination effect. Several point sources of illumination can represent the entire illumination shape [127]. The higher the number of representative point sources of illumination considered in the rigorous computation, the higher the accuracy in the model. This is a matter of trade off between the accuracy and the computation time.

In this chapter, the comparison of the rigorous model with compact mask models is shown.

4.1 Simulation Conditions

The \textit{alpha test} used for the comparisons of the proposed compact mask models, consist in a L/S layout with a linewidth of 45 nm. A pitch range from 90 nm to 330 nm is taken (see Table 4.2). Unless otherwise stated, the dimensions of the mask layouts and pitch are expressed in wafer units. Additional layouts are considered, in order to prove the capabilities of some of the mask models studied in this thesis (see Table 4.3).
All simulations in this work are done for water immersion lithography at a wavelength of 193 nm. Figure 4.1 shows the mask layouts contemplated in this study. A typical AttPSM (6% attenuation) is considered, the optical properties and mask materials are shown in Table 4.1. Illumination and imaging conditions are described in Table 4.3. The numerical settings such as source sampling, sampling (distance) of image points and Waveguide orders are shown in Table 4.4. The layouts, shown in Figure 4.1, were chosen for this study because they represent basic structures that can be used to construct more complex mask geometries. Moreover, the diffraction spectrum of such structures present a relatively low computation time compared with more complex structures, which is very beneficial for testing the proposed compact mask models.

![MoSi Absorber, Glass substrate](image)

Figure 4.1: Mask layouts used in this work. (left) lines and spaces (L/S), (center) contact and (right) line ends. Periodic boundary conditions are used for the computation of the diffraction spectrum.

<table>
<thead>
<tr>
<th>Material</th>
<th>Thickness</th>
<th>n</th>
<th>k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quartz substrate</td>
<td>20 nm</td>
<td>1.5595</td>
<td>0</td>
</tr>
<tr>
<td>MoSi absorber (6%)</td>
<td>68 nm</td>
<td>2.442</td>
<td>0.586</td>
</tr>
</tbody>
</table>

The illumination sources shown in Figures 4.2 and 4.3, are specified in order to capture the first diffraction orders for the most dense case. The smallest pitches produce diffracted orders at larger angles. Therefore, whenever the capture of the first orders for the smallest pitches is assured, the capture of the first orders for the larger pitches is assured as well. Figure 4.2 additionally represents the HA, in terms of the position of the illumination point sources for which the spectrum is rigorously computed with and without the HA.

In Figure 4.4, the schematic representation of the source sampling is shown. It is established that for $\sigma = 1$, the number of tangential points is 70 and the number of radial...
Table 4.2: Alpha test: Imaging and illumination conditions. All simulations were done for a wavelength of 193 nm and a 4X reduction factor.

<table>
<thead>
<tr>
<th>Mask layout</th>
<th>Lines and Spaces (L/S)</th>
<th>Target size (nm)</th>
<th>Pitch (nm)</th>
<th>NA</th>
<th>(n_{\text{immersion}})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>45</td>
<td>(90,330)</td>
<td>1.35</td>
<td>1.44</td>
</tr>
<tr>
<td><strong>Illumination</strong></td>
<td></td>
<td>dipole illumination</td>
<td>with (y)-polarization, (\sigma=0.7/0.9), opening angle=30°</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3: Other mask layouts used in this thesis. The illumination and imaging conditions are described. The wavelength corresponds to 193 nm and the reduction factor is 4X.

<table>
<thead>
<tr>
<th>Mask layout</th>
<th>Lines and Spaces (L/S)</th>
<th>Contacts</th>
<th>Line Ends</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target size (nm)</td>
<td></td>
<td>22,32</td>
<td>65,90</td>
</tr>
<tr>
<td>Pitch (nm)</td>
<td>(80,330)</td>
<td>130 and 330;180 and 360</td>
<td>200</td>
</tr>
<tr>
<td>NA</td>
<td></td>
<td>1.35</td>
<td></td>
</tr>
<tr>
<td>(n_{\text{immersion}})</td>
<td></td>
<td>1.44</td>
<td></td>
</tr>
<tr>
<td><strong>Illumination</strong></td>
<td></td>
<td>dipole illumination</td>
<td>CQuad illumination</td>
</tr>
<tr>
<td></td>
<td></td>
<td>with (y)-polarization, (\sigma=0.7/0.9), opening angle=30°</td>
<td>with xy-polarization, (\sigma=0.5/0.7), opening angle=30°</td>
</tr>
</tbody>
</table>

Table 4.4: The numerical settings for the simulations in this work. All units have a 4X reduction factor. The source sampling is given for \(\sigma=1\).

<table>
<thead>
<tr>
<th>Source Sampling</th>
<th>Image Sampling</th>
<th>Waveguide Orders</th>
</tr>
</thead>
<tbody>
<tr>
<td>70 points (tangential direction)</td>
<td>20 points (radial direction)</td>
<td>0.1 nm</td>
</tr>
</tbody>
</table>

points is 20 equally distributed. For smaller \(\sigma\) values, the density of the sampling points is kept constant in each direction.

The method of evaluation for each compact mask model is an aerial-image-based analysis, in which the PW corresponding to each model is compared to the PW of the rigorous model. OPC techniques are not performed in the mask layout, since the purpose of this investigation is to obtain the most approximate PW compared to the PW of the rigorous model, rather than maximize the lithographic performance for a specific layout.
Figure 4.2: Dipole illumination. $\theta = 30^\circ$, $\sigma_{in.} = 0.7$ and $\sigma_{out.} = 0.9$. $y$-polarization. With the Hopkins assumption (HA) only the spectrum corresponding to the normally incident angle of illumination is rigorously computed. Without the HA, different spectra corresponding to the representative angles of illumination are rigorously computed.

Figure 4.3: CQuad illumination. $\theta = 30^\circ$, $\sigma_{in.} = 0.5$ and $\sigma_{out.} = 0.7$. $xy$-polarization.

Figure 4.4: Schematic representation of the sampling of the source.
4.2 Rigorous and Scalar Model for 2D Features (Lines and Spaces)

The rigorous model based on the waveguide method with and without the HA (3D-w. HA and 3D-w/o HA, respectively) is analyzed and compared with the Kirchhoff model. A dipole illumination with y-polarization (Figure 4.2) is used. The mask layout corresponds to a periodic array of lines and spaces (L/S) (Figure 4.1).

In Figures 4.5-4.7, the cross sections across the corresponding metrology line are extracted from the aerial images. The most isolated lines correspond to those with the largest pitch of 330 nm. On the other hand, the most dense lines correspond to those with the smallest pitch of 80 nm and 90 nm. The incorporation of the physical properties of the MoSi absorber to the rigorous simulation has a strong impact on the intensity profile. This can be seen in the rigorous model with (green curves) and without (blue curves) the HA, which presents the lowest intensity in all the cases compared to the scalar model (red curves). Isolated configurations exhibit intensity side lobes in the vicinity from the absorber. This is due to the interaction of a higher number of diffraction orders compared to the dense configurations. The rigorous model with and without the HA yield similar results in terms of contrast for a CD of 32 nm and 45 nm in an isolated configuration. In general, the differences between the rigorous and the scalar models are much more critical for dense distributions and small CDs.

The rigorous and the scalar model are compared in terms of process windows (PW) and CD measurements for layouts of L/S with linewidths of 22 nm, 32 nm and 45 nm, as shown in Figures 4.8, 4.9 and 4.10, respectively. The rigorous model is analyzed with and without the HA. With the HA, 2 point sources at different angles of illumination are rigorously computed. Each point source is located at the center of each dipole, as indicated in Figure 4.3. Without the HA only the vertical angle of illumination is rigorously computed.

In Figure 4.8-4.10, the PW analysis shows that the scalar model (red curves) presents a symmetry in terms of focus position and bigger intensity (threshold) values than the rigorous model with (green curves) and without (blue curves) the HA. The blue curves in all cases present a strong shift of the focus position for different threshold values. This can be understood as a mask topography effect, interpreted as an aberration in the image formation. The introduction of this observed effects to the scalar model, are the foundation of this thesis. The rigorous model with the HA (green curves) can predict the effects introduced by the mask but not with the same accuracy as without the HA. In all cases, the scalar model predicts large intensity thresholds compared to the rigorous models, this is because the model neglects light diffraction from the mask edges, which is closely related to the intensity loss in the image.

For the smallest feature size (22 nm; Figure 4.8), a considerable difference between the
models for all the pitches from 90 nm to 330 nm is observed. For 32 nm linewidth (Figure 4.9) the rigorous model with the HA gets closer to the rigorous model without the HA. For the largest linewidth (45 nm; Figure 4.10) the process windows for the compared models have a small overlap, but still the shift on the focus described by the rigorous model is notorious. If the linewidth grows until the size is much larger than the wavelength $\lambda$ used in the lithography process, the difference between the models should get smaller and then the use of rigorous models is not required. That is, of course, not the case of the present study where the features in the mask are smaller or at least comparable to the wavelength. The optical effects observed due to this fact, have an important impact in the modeling of the optical lithographic process.
Figure 4.5: Cross sections at the best focus position for the analyzed mask models. CD=22 nm. The Metrology line to extract the cross section is located at the coordinates \((x_1=-\text{Pitch}, y_1=0)\); \((x_2=\text{Pitch}, y_2=0 \text{ nm})\). (a) Pitch=80 nm and (b) Pitch=330 nm. The rigorous model without the HA is represented by the blue curves. The rigorous model with the HA is represented by the green curves and the scalar model is represented by the red curves.
Figure 4.6: Cross sections at the best focus position for the analyzed mask models. CD=32 nm. The Metrology line to extract the cross section is located at the coordinates \((x_1=-\text{Pitch}, y_1=0); (x_2=\text{Pitch}, y_2=0 \text{ nm})\). (a) Pitch=80 nm and (b) Pitch=330 nm. The rigorous model without the HA is represented by the blue curves. The rigorous model with the HA is represented by the green curves and the scalar model is represented by the red curves.
Figure 4.7: Cross sections at the best focus position for the analyzed mask models. CD=45 nm. The Metrology line to extract the cross section is located at the coordinates \((x_1=-\text{Pitch}, y_1=0); (x_2=\text{Pitch}, y_2=0 \text{ nm})\). (a) Pitch=90 nm and (b) Pitch=330 nm. The rigorous model without the HA is represented by the blue curves. The rigorous model with the HA is represented by the green curves and the scalar model is represented by the red curves.
Figure 4.8: Process windows for lines and spaces (L/S) of 22 nm linewidth with pitches of: (a) 90 nm, (b) 110 nm, (c) 150 nm, (d) 210 nm, (e) 270 nm, (f) 290 nm, (g) 310 nm and (h) 330 nm.
Figure 4.9: Process windows for lines and spaces (L/S) 32 nm linewidth with pitches of: (a) 90 nm, (b) 110 nm, (c) 150 nm, (d) 210 nm, (e) 270 nm, (f) 290 nm, (g) 310 nm and (h) 330 nm.
Figure 4.10: Process windows for lines and spaces (L/S) 45 nm linewidth with pitches of: (a) 90 nm, (b) 110 nm, (c) 150 nm, (d) 210 nm, (e) 270 nm, (f) 290 nm, (g) 310 nm and (h) 330 nm.
So far, the cross sections and PW versus pitch have been analyzed. The PW analysis is as important as the analysis of the imaging through pitch. Therefore, the next step is to study the variation of typical lithographic performance parameters versus pitch. The CD, best-focus position and threshold-to-size through pitch for 22 nm, 32 nm and 45 nm linewidths are observed in Figures 4.11, 4.12 and 4.13, respectively. As a general statement, it can be said that the rigorous model with the HA lies between the most rigorous and the scalar model for all parameters.

To obtain the CD value for every pitch, an extraction of a cross section of the aerial image is required. The aerial image is located at the best focus value yielded by the PW analysis. The cross section is taken from this image, with the intensity threshold corresponding to the most dense (smallest pitch) layout. This condition is imposed due to the higher requirements that the smallest pitch presents in terms of image formation. Once the intensity threshold of the most dense layout is found, by means of the PW
Figure 4.12: Evaluation results for a linewidth of 32 nm. Red curves represent the scalar model, green curves represent the rigorous model with the HA and the blue curves represent the rigorous model without the HA. See Tables 4.3 and 4.4 for imaging conditions.

analysis, the image formation for the rest of the pitches is roughly assured. This choice reproduces the conditions of a practical case, where the mask is composed of several pitches and only one intensity threshold value is used. The CD through pitch behavior for all the linewidths is shown in Figures 4.11(a), 4.12(a) and 4.13(a). For all cases, the rigorous model with the HA is closer to the scalar model than to the most rigorous model.

The trend of the variation of best-focus position through pitch is very similar for the rigorous model with and without HA, as shown in Figures 4.11(b)-4.13(b). On the other hand, the scalar model can not described the focus variation, having a value near to 0 for the complete range of focus. This is an expected behavior, since the scalar model does not take into account the 3D topography of the mask neither the aberration-like effects that this provokes.

The threshold-to-size value is the intensity threshold that yields the required target CD. The threshold-to-size results confirm that as a result of the modeling of the absorbing
Figure 4.13: Evaluation results for a linewidth of 45 nm. Red curves represent the scalar model, green curves represent the rigorous model with the HA and the blue curves represent the rigorous model without the HA. See Tables 4.3 and 4.4 for imaging conditions.

material on the mask, the most rigorous model has the lowest threshold-to-size values through the complete range of pitches for all the linewidths (Figures

In Figures 4.14(a)-4.14(c), the computational cost for all the analyzed target CDs is shown. The 22 nm and 32 nm CD targets, were computed with a 4x Dual-Core AMD Opteron 8218 @2.6GHz, (SuSE Linux 11.1), 32 GB RAM. The 45 nm CD target, was computed with a 2 x Intel Quad Core Xeon E5405 @ 2.0 GHz, EM64T, (SuSE Linux 11.1), 48 GB RAM.

As expected, the most accurate model, e.g. the rigorous model without the HA, is the most time consuming among the studied models. More angles of illumination involve more computation time. The scalar model takes just a few ms of computation time, since it is based on the computation of Fourier Transforms, which are quite optimized calculations nowadays. The rigorous model with the HA offers a good compromise in terms of time, speeding up the computation by a factor of 1.7 approximately.
The computation time of the most rigorous model, based on the Waveguide method, depends on the number of diffraction orders. At the same time, the number of diffraction orders is proportional to the pitch value, that is for higher pitch values more diffraction orders have to be computed. This dependence is clearly observed in Figures 4.14(a)-4.14(c), where the time consumption increase for higher pitch values.

In this thesis, the low-computational-cost characteristic of the scalar model is used to propose compact mask models based on a fast model that describe the effects predicted by the most rigorous model with considerable accuracy.
4.3 Full 3D, Quasi3D Rigorous and Scalar Models for 3D Features (Contact Holes)

In order to test the capabilities of the Quasi-3D model, 2D features are considered. A CQUAD illumination with xy-polarization (Figure 4.3). The mask is composed of a periodic array of contact holes (Figure 4.1) with target CDs of 65 nm and 90 nm. Other important imaging parameters are indicated in Tables 4.3 and 4.3.

In Tables 4.5-4.8 the aerial images at the best focus position, resulting from the PW analysis, are observed. The scalar model, the Quasi-3D model with and without the HA and the rigorous model with and without the HA are compared. Without the HA, 9 and 49 rigorously computed spectra are considered. This is done in order to observe the impact that the number of spectra has on the model performance. For the 65 nm contact size in a dense configuration (pitch 130 nm; Table 4.5), all the aerial images present a low contrast level, but still some information about the difference between the models can be extracted. The cross sections obtained with the Kirchhoff approach (red curve) show the highest NILS among the models. However, the most accurate model is clearly demonstrating that this lithographic task requires more effort in the OPC to obtain images of good contrast.

In Table 4.6, the aerial images at the best focus for the 65 nm contact size and pitch=330 nm, are shown. This layout represents an almost isolated configuration, since the pitch has a value of more than 5 times the size of the contact. Cross sections comparison shows that the scalar approach predicts again the highest NILS among the models. The rigorous model and the Quasi-3D model without the HA present small differences between them. The rigorous model and the Quasi-3D model with the HA exhibit similarities in their tendencies but a higher intensity threshold value compared to the rigorous model and the Quasi-3D model without the HA.

In Table 4.7, the aerial images at the best focus for the 90 nm contact size and pitch=180 nm, are shown. The cross sections show that in general the contrast is increased and the rigorous model and the Quasi-3D model with and without the HA present the same tendency with some small differences in the threshold intensity values.

In Table 4.8, the aerial images at the best focus for the 90 nm contact size and pitch=360 nm, are shown. The rigorous model and the Quasi-3D model without the HA are almost indistinguishable. A small difference in terms of threshold intensity is found between the rigorous model and the Quasi-3D model with the HA.
Table 4.5: Aerial images and cross sections at the best focus position for the analyzed mask models. CD=65 nm, pitch=130 nm. The Metrology line to extract the cross section is located at the central contact at the coordinates $(x_1=-65 \, \text{nm}, \, y_1=0 \, \text{nm}); \, (x_2=65 \, \text{nm}, \, y_2=0 \, \text{nm})$.

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Image 1</th>
<th>Image 2</th>
<th>Image 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D-w/o HA (49)</td>
<td><img src="image1" alt="Image" /></td>
<td><img src="image2" alt="Image" /></td>
<td><img src="image3" alt="Image" /></td>
</tr>
<tr>
<td>3D-w/o HA (9)</td>
<td><img src="image4" alt="Image" /></td>
<td><img src="image5" alt="Image" /></td>
<td><img src="image6" alt="Image" /></td>
</tr>
<tr>
<td>3D-w. HA (1)</td>
<td><img src="image7" alt="Image" /></td>
<td><img src="image8" alt="Image" /></td>
<td><img src="image9" alt="Image" /></td>
</tr>
</tbody>
</table>

### Scalar Cross Sections

![Scalar Cross Sections](image10)

![Scalar Cross Sections](image11)

### Cross Sections

![Cross Sections](image12)

![Cross Sections](image13)
Table 4.6: Aerial images and cross sections at the best focus position for the analyzed mask models. CD=65 nm, pitch=330 nm. The Metrology line to extract the cross section is located at the central contact at the coordinates \((x_1=-65 \text{ nm}, y_1=0 \text{ nm}); (x_2=65 \text{ nm}, y_2=0 \text{ nm})\).
Table 4.7: Aerial images and cross sections at the best focus position for the analyzed mask models. CD=90 nm, pitch=180 nm. The Metrology line to extract the cross section is located at the central contact at the coordinates \((x_1=-90 \text{ nm}, y_1=0 \text{ nm})\); \((x_2=90 \text{ nm}, y_2=0 \text{ nm})\).

<table>
<thead>
<tr>
<th>Model</th>
<th>Aerial Image</th>
<th>Cross Sections</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D-w/o HA (49)</td>
<td><img src="image" alt="3D-w/o HA (49)" /></td>
<td><img src="image" alt="Cross Sections" /></td>
</tr>
<tr>
<td>3D-w/o HA (9)</td>
<td><img src="image" alt="3D-w/o HA (9)" /></td>
<td><img src="image" alt="Cross Sections" /></td>
</tr>
<tr>
<td>3D-w. HA (1)</td>
<td><img src="image" alt="3D-w. HA (1)" /></td>
<td><img src="image" alt="Cross Sections" /></td>
</tr>
<tr>
<td>Quasi-3D-w/o HA (49)</td>
<td><img src="image" alt="Quasi-3D-w/o HA (49)" /></td>
<td><img src="image" alt="Cross Sections" /></td>
</tr>
<tr>
<td>Quasi-3D-w/o HA (9)</td>
<td><img src="image" alt="Quasi-3D-w/o HA (9)" /></td>
<td><img src="image" alt="Cross Sections" /></td>
</tr>
<tr>
<td>Quasi-3D-w. HA (1)</td>
<td><img src="image" alt="Quasi-3D-w. HA (1)" /></td>
<td><img src="image" alt="Cross Sections" /></td>
</tr>
</tbody>
</table>
Table 4.8: Aerial Images and Cross sections at the best focus position for the analyzed mask models. CD=90 nm, pitch=360 nm. The Metrology line to extract the cross section is located at the central contact at the coordinates $(x_1=-90 \text{ nm}, y_1=0 \text{ nm})$; $(x_2=90 \text{ nm}, y_2=0 \text{ nm})$. 

![Aerial Images and Cross sections](image-url)
In Figure 4.15, the PW analysis for a 65 nm contact hole size is observed. The HA has been applied in the rigorous model and in the Quasi-3D model. Without the HA rigorously computed spectra corresponding to 49 and 9 angles of illumination have been considered. For the smallest pitch (130 nm; Figure 4.15(a)) the rigorous and the Quasi-3D models describe the shift in the focus position, but present differences in terms of threshold latitude and best threshold. However, the Quasi-3D model without the HA (49 and 9 angles of illumination) does not describe properly the imaging through focus. The scalar model, as observed with the L/S layouts, present the a symmetry around the zero defocus value. The rigorous model without the HA (49 angles of illumination) is similar to the rigorous model without the HA (9 angles of illumination). Regarding the larger pitches (Figures 4.15(b)-4.15(f)), the scalar model (red curves) has always a different tendency with respect to the other models, for all the observed pitches. The rigorous and the Quasi-3D models without the HA (blue, green, cyan and magenta curves) exhibit the same behavior. Additionally, the rigorous and Quasi-3D models with the HA have always higher threshold values compared to the other models but similar tendencies between them.

In Figure 4.16, the PW analysis for a 90 nm contact hole size is shown. As expected, a critical difference is observed for the smallest pitch (180 nm; Figure 4.16(a)), between the scalar model and the rest of the models. Two general statements can be suggested from this PW analysis. First, a focus shift effect is well described by the rigorous and Quasi-3D models with and without the HA. In contrast, the scalar model present a symmetry around the zero defocus value. Second, the rigorous and Quasi-3D models with and without the HA present a smaller percentage difference between them – in terms of focus-shift and threshold values – in comparison to the previous smaller contact size analyzed.

Computational times for 65 nm CD and 90 nm CD are observed in Figures 4.17 and 4.18 respectively. The high performance of the scalar model and the Quasi-3D model is clearly demonstrated. The rigorous model without the HA (49 angles of illumination) is the most time consuming, taking 7359.2 min for the 90 nm contact size (360 nm pitch) and 5888.7 min for the 65 nm contact size (330 nm pitch). The general observation is that for higher pitches each model turns out to be more time consuming, except for the scalar model, where the computation time is too short to observe a significant variation versus pitch.
Figure 4.15: PW analysis for a contact size of 65 nm. Red curves represent the scalar model, black curves represent the Quasi-3D model with the HA, magenta curves represent the Quasi-3D model without the HA (9 angles of illumination), cyan curves represent the Quasi-3D model without the HA (49 angles of illumination), yellow curves represent the rigorous model with the HA, green curves represent the rigorous model without the HA (9 angles of illumination) and the blue curves represent the rigorous model without the HA (49 angles of illumination). (a) pitch=130 nm, (b) pitch=145 nm, (c) pitch=155 nm, (d) pitch=160 nm, (e) pitch=205 nm, (f) pitch=330 nm.
Figure 4.16: PW analysis for a contact size of 90 nm. Red curves represent the scalar model, black curves represent the Quasi-3D model with the HA, magenta curves represent the Quasi-3D model without the HA (9 angles of illumination), cyan curves represent the Quasi-3D model without the HA (49 angles of illumination), yellow curves represent the rigorous model with the HA, green curves represent the rigorous model without the HA (9 angles of illumination) and the blue curves represent the rigorous model without the HA (49 angles of illumination). (a) pitch=180 nm, (b) pitch=285 nm, (c) pitch=290 nm, (d) pitch=300 nm, (e) pitch=305 nm, (f) pitch=360.
Figure 4.17: Time consumption through pitch for 65 nm contact size. The scalar model (red), the Quasi-3D model with the HA (black), the Quasi-3D model without the HA (9 angles of illumination; magenta), the Quasi-3D model without the HA (49 angles of illumination; cyan), the rigorous model with the HA (yellow), the rigorous model without the HA (9 angles of illumination; green) and the rigorous model without the HA (49 angles of illumination; blue).

Figure 4.18: Time consumption through pitch for 90 nm contact size. The scalar model (red), the Quasi-3D model with the HA (black), the Quasi-3D model without the HA (9 angles of illumination; magenta), the Quasi-3D model without the HA (49 angles of illumination; cyan), the rigorous model with the HA (yellow), the rigorous model without the HA (9 angles of illumination; green) and the rigorous model without the HA (49 angles of illumination; blue).

An additional test is performed with a mask layout containing different feature sizes and pitches. In Figure 4.19, the mask and the aerial image of an adapted static random-access memory (SRAM) cell is observed. The spectrum is computed using the Hopkins as-
The computational time for the spectrum corresponding to the TE-polarization component is 3 s with a memory consumption of 3.6 MB. The computation of the spectrum corresponding to the TM-polarization component takes 2 s with a memory consumption of 4.2 MB.

![Mask layout of an adapted SRAM cell and aerial image](image)

**Figure 4.19**: (a) Mask layout of an adapted SRAM cell and (b) aerial image with different feature sizes and pitches. The Quasi-3D model is used to compute the spectrum using the Hopkins assumption.

The aerial-imaged-based PW analysis for different structures, sizes and pitches have demonstrated that, in general, the Quasi-3D model with or without the HA provides a good compromise in terms of effectiveness (computation time) and accuracy compared to the most rigorous model.

## 4.4 Model Calibration for Compact Mask Models

The Quasi-3D model is a simplification of the diffraction spectrum calculation, but still requires the computation of the Maxwell’s Equations. The compact models studied in the following sections are based on the enhancement of the Kirchhoff model. This method is based on the scalar diffraction theory and its implementation is straightforward. However, the compact models require a calibration with a reference model.

The calibration consists on the optimization of the required parameters for each model. A genetic algorithm (GA) is suitable for this task, since the search space of the variables to optimize is more or less known. The GA uses this search space to find an approximate solution to the optimization problem. Heuristic approaches based on principles of nature selection[128] are applied. The algorithm is initialized with a population of $n$ chromosomes, which represent a potential solution to the problem. Each compact mask model has representative parameters that need to be optimized. In the case of the BLM,
the chromosomes are represented by the width \( w \), transmission \( t_3 \) and phase \( \varphi_3 \) of the surrounding layer. In the PM, the transmission \( A \) and the phase \( \phi \) are the values to be optimized. The Zernike coefficients are the values to be obtained in the PFM. Finally, in the SCM, the polynomial coefficients that modify the mask spectrum are the chromosomes to be optimized.

After the initial population is set (by a random construction), an evaluation of a fitness function \( F(x) \) is performed over the entire population. Then a new population is created, based on the following steps [125]:

- **Selection**: Select two parent chromosomes (two candidate solutions) from a population according to their fitness. This selection can be performed according to probabilistic rules.

- **Crossover**: With a crossover probability, recombine two parents to form a new offspring. This will produce two children that will have different percentages of the genetic information corresponding to each parent, depending of the type of recombination used. In this case, two-point crossover recombination is used. That is, two points are selected on each parent chromosome. Everything between the two points is exchanged between the parents, producing the new offspring.

- **Mutation**: With a mutation probability, alter the new offspring at a random position selected in the chromosome. This is done to introduce what in evolution theory is called variation. Natural selection will only cause evolution if there is enough genetic diversity in a population.

A new generation is iteratively created until a termination criteria is reached. For instance, a local minimum of the fitness function, number of generations, time consumption, etc.

The optimum parameters needed for each compact mask model are extracted from the optimization process described above. The fitness function \( F \) that determines the best parameters is constructed from comparisons of the compact model to be analyzed with the most rigorous model. It has been shown that the PW analysis and the CD measurements are good indicators to evaluate each model. Therefore, the fitness function is written in terms of these variables like follows:

\[
F = \frac{1}{N} \sum_{n=1}^{N} \left[ \left( \frac{ThLat_{test} - ThLat_{ref}}{ThLat_{max}} \right)^2 + \left( \frac{Th_{test} - Th_{ref}}{Th_{max}} \right)^2 \\
+ \left( \frac{f_{test} - f_{ref}}{f_{max}} \right)^2 + \left( \frac{cd_{test} - cd_{ref}}{cd_{max}} \right)^2 \right],
\]

(4.1)
where \( N \) is the number of pitches used in the optimization process. \( ThLat \) is the threshold latitude, \( Th \) is the best threshold, \( f \) is the best focus position and \( cd \) is the CD target. \( ref \) stands for the reference model and \( test \) for each alternative compact mask model. The intention of using several pitches in the optimization process is to increase the prediction capacity of each model. The introduction of different pitches results in a higher complexity regarding the optimization process, since the number of diffraction orders that contribute to the image formation is directly related to the pitch value. Therefore, it is not a trivial task to find optimized parameters for each model, which are suitable for small and large pitches simultaneously. The use of a single pitch only in the optimization process turns out in a straightforward optimization. It is important to emphasize the fact that just a few pitch values are taken into account in the calibration process. The cross-validation step of the compact model, consists in taking mask layouts corresponding to pitches that were not included in the calibration process to evaluate the predictability of the model.

### 4.5 Boundary Layer Model (BLM)

The BLM is based on the scalar model. The only requirement is to modify the features in the thin mask with a material of certain values of transmission and phase (Figure 4.20). The task then is to optimize the values of the width of the surrounding layer \( w \), the transmission \( t_3 \) and the phase \( ϕ_3 \).

![Figure 4.20: Representation of the thin mask modification with additional material around the mask features. Transmission and phase values for air \((t_1,ϕ_1)\), absorber material \((t_2,ϕ_2)\) and surrounding material \((t_3,ϕ_3)\)](image-url)
Alpha test is used. Other important specifications are found in Tables 4.3 and 4.4. For this model, the calibration process uses the information of 9 pitches: 90 nm, 120 nm, 150 nm, 180 nm, 210 nm, 240 nm, 270 nm, 300 nm and 330 nm.

4.5.1 Model Performance Compared to the Rigorous Model with the HA

The optimized values found with this reference are observed in Table 4.9. The total width value that has to be added to the line is optimized. Half of this value is added to the left side of the line and half to the right side. In that way, the linewidth changes from 45 nm to 52.9 nm. The transmission value of the material is close to one and the phase is either greater than 180° or smaller than 0.

<table>
<thead>
<tr>
<th>width (nm)</th>
<th>transmission (a.u.)</th>
<th>phase (rad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.91579</td>
<td>0.99872</td>
<td>4.99017</td>
</tr>
</tbody>
</table>

The PW analysis for pitches used in the calibration process, is shown in Figure 4.21. In Figure 4.22 the PW analysis for pitches that are not used in the calibration process is shown. The results confirm that – for such a simple mask model implementation – the agreement with the rigorous model using the HA is remarkable. The difference between the calibrated pitches and the non-calibrated pitches can not be noticed. The BLM predicts the results of the rigorous model in all the range of pitches without exceptions. However, when the pitch becomes smaller, the implementation of the model has to be carefully considered, since the proximity effects become larger and overlapping of the assist layers may occur.
Figure 4.21: PW analysis for L/S 45 nm linewidth. The BLM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.
Figure 4.22: PW analysis for L/S 45 nm linewidth. The BLM model is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm.
4.5.2 Model Performance Compared to the Rigorous Model without the HA

The optimized values found with this reference are observed in Table 4.10. Half of the width value is added to the left side of the line and half to the right side. In that way, the linewidth is changed from 45 nm to 65.09 nm.

<table>
<thead>
<tr>
<th>width (nm)</th>
<th>transmission (a.u.)</th>
<th>phase (rad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20.0954</td>
<td>0.99686</td>
<td>5.33927</td>
</tr>
</tbody>
</table>

Table 4.10: Optimized values for the BLM. Reference without the HA.

In Figures 4.23 and 4.24 the PW analysis for the considered test case is shown. The reference without the HA takes into account two representative angles of illumination. The BLM is calibrated using the pitches observed in Figure 4.23. Additionally a validation with pitches that are not used in the calibration process is shown in Figure 4.24.

The BLM for the smallest pitches of 90 nm (Figure 4.23(a)) and 100 nm (Figure 4.24(a)) does not reproduce the mask-induced effects predicted by the rigorous model in terms of intensity threshold. However, the focus shift is very well described by the model for the 100 nm pitch as shown in Figure 4.24(a). Taking into consideration multiple pitch values in the calibration process, has influence on the threshold differences between the compared models for small and large pitches. Individual optimizations using only the smallest or the largest pitches produced a considerable improvement. However, as mentioned before, the aim of this analysis is to observe the degree of predictability of the model for several pitches. Therefore, the calibration process with the utilization of only one or few pitches is irrelevant, at least for this study.

In Figures 4.23(b)-4.23(f) a good agreement is found between the BLM and the rigorous model. A small difference in terms of threshold values is found for 150 nm and 180 nm pitches.

In spite of the behavior of the smallest pitches, the BLM yielded very good results. This can be observed in Figure 4.24, where the BLM reproduce the rigorous model tendency even for pitches that were not used in the calibration process.
Figure 4.23: PW analysis for L/S 45 nm linewidth. The BLM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.
Figure 4.24: PW analysis for L/S 45 nm linewidth. The BLM model is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm.
The BLM does not reproduce the threshold effects described by the rigorous model without the HA for the smallest pitch. This can be explain because when more oblique angles are computed rigorously, like in the case of the rigorous model without the HA, a cross-talk between the boundary layers may become evident, specially when layouts with dense configurations are considered. The effect of oblique illumination has a strong impact on the smallest pitches and is very well described by the rigorous model without the HA.

It is important to highlight that the consideration of different angles of illumination in the reference without the HA, does not bring any additional effort to the implementation of the BLM compared to the reference with the HA. The reference for the calibration process is replaced and no further changes have to be introduced to the BLM, since the method is based on modifications in the spatial domain and is independent from the illumination source.

The time consumption for the BLM calibration depends not only on the number of parameters to optimize but also on the parameters of the genetic algorithm, such as the number of generations, the size of the search space of the variables, the type of selection, crossover and mutation. Additionally, the time consumption is highly influenced by the combination of number of nodes and processors per node\(^1\) used in the parallelization of the genetic algorithm. Therefore, is not a simple task to quantify the time consumption, this can vary between 1 to 2 hours (using 2 nodes; 1 process per node), with 9 pitches for the calibration. Once the width of the surrounding layer \(w\), the transmission \(t_3\) and the phase \(\varphi_3\) values are found, the computation time of the spectrum is the same of the scalar model observed in the Figure 4.14(c).

### 4.6 Pulses Model (PM)

In the previous chapter, the expression for the modified transmittance used in the PM was shown. In order to increase control over the transmission variation in the calibration process, a correction factor is introduced to the model like follows:

\[
T'(x) = A' e^{i\varphi'} T(x) + A e^{i\varphi} [\delta(x - d/2) + \delta(x + d/2)],
\]

\[\text{(4.2)}\]

The diffraction spectrum of the modified mask is given by the Fourier transform of Equation 4.2:

\[
\mathcal{F}(T'(x); m) = \mathcal{F}(A' e^{i\varphi'} T(x)) + 2 A e^{i\varphi} \cos(\pi m d),
\]

\[\text{(4.3)}\]

\(^1\)Node is a common term for the computational elements that refers to a distributed-memory parallel machine. Each node contains its own memory and one or more processors.
where \( m \) represents the diffraction order and \( d = CD/pitch \) the duty cycle. The parameters to optimize in this case are the amplitude \( A \) and the phase \( \varphi \) that represent the magnitude and the phase of the edge pulse. Additionally, \( A' \) and \( \varphi' \) representing the value of the correction factor, have to be included as well in the calibration process.

One important aspect to consider in the calibration process of this model is the polarization. Unlike the BLM, the PM has a direct dependence with the illumination source. Therefore, polarization has to be considered in the calibration process, taking different parameters for each polarization state. \( A_{TM} \) and \( A_{TE} \) represent the amplitude value of the edge pulse for TM and TE polarization components. In the same way, \( \varphi_{TM} \) and \( \varphi_{TE} \) represent the phase value of the edge pulse for TM and TE polarization components.

### 4.6.1 Model Performance Compared to the Rigorous Model with the HA

In the Table 4.11 the optimized values of the calibration process are shown. For the correction factor \( A' = 0.956 \) and \( \varphi' = -0.75823 \) rad are obtained. The polarization effects are expressed in the edge pulses values \((A, \varphi)\) assigned to the TE and TM polarization components.

<table>
<thead>
<tr>
<th>Polarization</th>
<th>( A ) (a.u)</th>
<th>( \varphi ) (rad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM</td>
<td>0.006</td>
<td>0.57</td>
</tr>
<tr>
<td>TE</td>
<td>-0.006</td>
<td>-2.46</td>
</tr>
</tbody>
</table>

The PW analysis (Figure 4.25) for the calibrated pitches shows that for the smallest pitches, the PM cannot compensate the large threshold shift and considerable best focus shift between the Kirchhoff and the rigorous model. However, the performance significantly improves for larger pitches.
Figure 4.25: PW analysis for L/S 45 nm linewidth. The PM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.
Figure 4.26: PW analysis for L/S 45 nm linewidth. The PM model is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm.
For pitches that were not used in the calibration process (Figure 4.26), the results are satisfactory taking into account that the model was calibrated using only 9 pitches. In general, the PM resembles with some limitations the behavior of the rigorous model with the HA.

### 4.6.2 Model Performance Compared to the Rigorous Model without the HA

Without the HA two representative angles of illumination are considered, then two pairs of \((A, \varphi)\) values are calibrated for each state of polarization. In addition, the \((A', \varphi')\) corresponding to the correction factor are also incorporated. That makes a total of 10 parameters to optimized for the PM model when the reference without the HA is used. In the Table 4.12, the calibrated values for the TM and TE polarization states are given. The subscripts 1 and 2 represent the angles of illumination. The amplitude and phase of the correction factor are \(A' = 0.941, \varphi' = 2.46964\).

| Table 4.12: Optimized values for the BLM. Reference without the HA. |
|------------------|------------------|------------------|------------------|
| Polarization     | \(A_1\) (a.u)   | \(\varphi_1\) (rad) | \(A_2\) (a.u)   | \(\varphi_2\) (rad) |
| TM               | -0.01859        | 0.652            | -0.01677        | 0.486              |
| TE               | -0.01827        | 0.984            | 0.01898         | 2.931              |

The PW analysis for the calibrated (Figure 4.27) and non-calibrated (Figure 4.28) pitches show that in the case where more representative angles of illumination are considered, the PM presents higher difficulties to retrieve the rigorous model behavior. This is due to the consideration of more parameters in the calibration process, which makes the optimization task more difficult. Compared to the case where the reference with the HA is used, the PM model is not able to completely reproduce the behavior through focus of the rigorous model without the HA. However, an important correction to the scalar model is indeed observed.
Figure 4.27: PW analysis for L/S 45 nm linewidth. The PM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.
Figure 4.28: PW analysis for L/S 45 nm linewidth. The PM is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm.
In terms of computation time for the calibration process, the PM presents an elevated cost compare with the BLM since more parameters are required. This can be compensated with the use of more nodes in the parallel computation, but still results expensive in terms of resources.
Chapter 5

Evaluation and Comparisons of the Proposed Compact Mask Models

The proposed models in this work are based on the modification of the conventional, and rather inaccurate, scalar model, but with the capability of incorporating topographic mask-induced effects and polarization dependencies of the field transmitted by the photomask. Analyzing the optical projection system (Figure 5.1), there are different regions that can be directly manipulated to improve the performance of the scalar (Kirchhoff) model. First, the thin mask can be modified in the spatial domain. Second, the scalar diffraction spectrum of the mask can be modified in the frequency domain. A special case of frequency domain modification is the pupil filtering, in which the mask-induced effects are represented in the pupil plane of the projection system. An additional modification in the projection system can be carried out in the illumination source, but is not considered in this study. The modification of the illumination source is performed in many resolution enhancement techniques (RET), but this is out of the scope of this work.

In order to validate the proposed models, the same procedure is followed as in the previous analysis of the compact mask models from literature. Two references with and without the HA are evaluated. The test case used to compare the proposed compact mask models, is the alpha test defined in Section 4.1. Illumination, imaging and numerical conditions are described in Tables 4.2 and 4.4. 9 pitches are used to calibrate the model and the rest are taken into account for the cross-validation procedure.

5.1 Pupil Filtering Model (PFM)

In order to approximate to the rigorous model, many aberration components can be introduced to the model, although this might be in this thesis, the Waveguide method is used for the rigorous computation of the mask spectrum together with the Abbe imaging
formulation without the Hopkins assumption (HA) [17] for the vector imaging computation. Specifically, this assumption refers to the diffraction spectrum calculation used by the Fraunhofer IISB lithography simulation software Dr.LiTHO [126]. Without the HA, each spectrum corresponding to a representative angle of illumination source is rigorously computed. This allows a high accuracy in terms of the description of the diffracted field. However, the computational expense increases significantly compared to image simulations with the HA, where the rigorously computed spectrum for the vertical incidence is shifted accordingly to each angle of illumination to account for the oblique illumination effect. Several point sources of illumination can represent the entire illumination shape [127]. The higher the number of representative point sources of illumination considered in the rigorous computation, the higher the accuracy in the model. This is a matter of trade off between the accuracy and the computation time.

In this chapter, the comparison of the rigorous model with compact mask models is shown. limiting since the number of parameters to optimized in the calibration process is increased. For this investigation, a total of 12 parameters are considered, 6 per polarization component: 3 coefficients for the amplitude and 3 coefficients for the phase of the wavefront.

5.1.1 Model Performance Compared to the Rigorous Model with the HA

The calibrated parameters using the rigorous model with the HA are shown in Table 5.1. The mayor contribution to the phase compensation is done by the spherical aberration coefficient $Z_9$ in both Jones components $J_{xx}$ and $J_{yy}$. The defocus $Z_4$ and astigmatism $Z_5$
coefficients have also a non-negligible contribution. In terms of amplitude, the constant factor has the largest contribution to the compensation followed by the coefficients $C_4$ and $C_9$, corresponding to the quadratic polynomial and the highest degree polynomial, respectively.

In Table 5.1, the optimized values for the PFM are shown. Reference with the HA.

<table>
<thead>
<tr>
<th>Jones term</th>
<th>$C_1$</th>
<th>$C_4$</th>
<th>$C_9$</th>
<th>$Z_4$ (mλ)</th>
<th>$Z_5$ (mλ)</th>
<th>$Z_9$ (mλ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J_{xx}$</td>
<td>0.916</td>
<td>0.197</td>
<td>-0.056</td>
<td>-9.1</td>
<td>7.3</td>
<td>24.1</td>
</tr>
<tr>
<td>$J_{yy}$</td>
<td>1.033</td>
<td>-0.165</td>
<td>0.125</td>
<td>-6.1</td>
<td>-8.3</td>
<td>16.5</td>
</tr>
</tbody>
</table>

In Figure 5.2, the amplitude and phase pupil functions corresponding to the $J_{xx}$ and $J_{yy}$ are shown. The amplitude function acts ultimately as a weight factor of the phase function. At those spatial frequencies where the amplitude is small, the phase value will not contribute as much as if the amplitude would have a considerable value. With this in mind, it can be said that the amplitude function of the $J_{xx}$ term (Figure 5.2(a)) gives preference to the diffraction orders entering to the pupil at very oblique angles of incidence. On the other hand, the amplitude function of the $J_{yy}$ term (Figure 5.2(c)) gives preference to the diffraction orders entering to the pupil at smaller oblique angles of incidence.

The phase of the pupil function corresponding to $J_{xx}$ (Figure 5.2(b)) appears to have a rotation of roughly $90^\circ$ with respect to the phase of the pupil function corresponding to $J_{xx}$ (Figure 5.2(d)) in the ($f_x,f_y$) pupil plane. For normally incident light, the spectrum of vertical y-parallel L/S lies entirely on the $f_x$ axis. With the HA, only the spectrum corresponding to normally incidence is computed. The spectra corresponding to the rest of the oblique illumination points that compose the source are obtained by a shift of the rigorously computed spectrum. In the vector imaging computation, the pupil filters $J_{xx}$ and $J_{yy}$ are applied to the resulting spectra.

The shifting of the rigorously computed spectra may result in spectral information along the $f_y$ axis. This may explain the functional form of the $J_{xx}$ phase, which contains important modulation information along the $f_y$ axis. The $J_{yy}$ phase has a strong modulation in the $f_x$ axis, this will cause strong filtering over the spectrum that lies near the $f_x$ axis.

The introduction of the obtained pupil filters into the vector image computation turns in an enhancement of the scalar model, as shown in Figures 5.3 and 5.4. In Figure 5.3, the PW analysis for calibrated pitches shows the improvement obtained with the pupil filters $J_{xx}$ and $J_{yy}$ applied to the scalar spectrum. In Figure 5.4, a cross-validation with non-calibrated pitches is done. The PFM is represented by the green curves with triangle markers. The pitch-dependence of the best focus position observed in the rigorous model.
Figure 5.2: Pupil filters associated to $J_{xx}$ and $J_{yy}$ Jones matrix components for the compensation of the mask-induced effects predicted by the rigorous model without the HA. $J_{xx}$ component (a) amplitude and (b) phase. $J_{yy}$ component (a) amplitude and (b) phase.

with the HA (blue curves with diamond markers) is reproduced with success by the PFM. The scalar model (red curves with circle markers) is plotted to facilitate the visualization of the shift in the focus position obtained by the PFM. For pitches greater than 180 nm the correction in terms of best focus position and intensity threshold is remarkable. However, for pitches smaller than 180 nm the threshold correction is questionable. This is attributed to the fact that several pitches are used in the calibration process. A pupil filter that introduces properly the mask-induced effects for smaller pitches has limitations to reproduce the mask-induced effects produced larger pitches and vice-versa. The optimization process finds a point in between for the dense and isolated lines.
Figure 5.3: PW analysis for L/S 45 nm linewidth. The PFM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.
Figure 5.4: PW analysis for L/S 45 nm linewidth. The PFM model is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm.
5.1.2 Model Performance Compared to the Rigorous Model without the HA

It is important to remember that in this case, two representative angles of illumination are considered. That is, the calibration process should somehow take into account different filters corresponding to each representative angle. This does not make sense physically speaking, since the pupil of the projection system describes the wavefront deformation of a total incident optical field. Calibrating two different pupil filters may result in an interesting academic task but does not reproduce the practical conditions, where a single pupil represents the wave deformation of the entire optical projection system. Therefore, the consideration of more pupil filters according to the number of representative angles is omitted in the PFM. This topic will be introduced in the next section where the details of the spectrum correction model (SCM) are given. For now, just a set of amplitude and phase filters are obtained for both of the representative angles of illumination.

The calibrated parameters using the rigorous model without the HA are shown in Table 5.2. The $J_{yy}$ term contains a larger spherical aberration coefficient $Z_9$ compared to the $J_{xx}$ term. In this optimization process, the astigmatism coefficient $Z_5$ for the $J_{xx}$ and $J_{yy}$ terms turns out to be larger than the spherical aberration coefficient. Among all the aberration contributions for the $J_{xx}$ and $J_{yy}$ terms, the astigmatism is the largest one. Unlike the previous case with the HA, where the largest contribution to the compensation of the mask-induced effects, was the spherical aberration. The optimization for the amplitude coefficients yielded the smallest coefficients for the quadratic polynomial, and the largest values for the constant factor in both $J_{xx}$ and $J_{yy}$ terms.

<table>
<thead>
<tr>
<th>Jones term</th>
<th>$C_1$ (a.u)</th>
<th>$C_2$ (a.u)</th>
<th>$C_3$ (a.u)</th>
<th>$Z_4$ (m(\lambda))</th>
<th>$Z_5$ (m(\lambda))</th>
<th>$Z_9$ (m(\lambda))</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J_{xx}$</td>
<td>0.995</td>
<td>-0.271</td>
<td>-0.1767</td>
<td>-9.1</td>
<td>-63</td>
<td>-11.6</td>
</tr>
<tr>
<td>$J_{yy}$</td>
<td>0.981</td>
<td>-0.1566</td>
<td>-0.0664</td>
<td>6.2</td>
<td>-77.3</td>
<td>36.1</td>
</tr>
</tbody>
</table>

Figure 5.5 shows the amplitude and phase pupil filters for the $J_{xx}$ and $J_{yy}$ terms of the Jones matrix. A strong modulation of the phase in both $f_x$ and $f_y$ axis is observed. Regarding the amplitude, high amplitude values are presented along the complete pupil, specially for the $J_{yy}$ component. The $J_{xx}$ term exhibits a stronger modulation than the $J_{yy}$ term. In this case, the modulation oscillates from lower amplitudes in the border of the pupil to higher amplitudes towards the center of the pupil.

Figures 5.6 and 5.7 show the PW analysis for calibrated and non-calibrated pitch values respectively. Figures 5.6(c)-5.6(f) show that for pitches larger than 150 nm the PFM corrects the threshold and focus values of the scalar model. For pitches smaller than 150 nm (Figures 5.6(a)-5.6(b) and 5.7(b)), the correction is possible only in terms of
focus. However, Figure 5.7(a) shows that for a non-calibrated and 100 nm pitch there is not correction at all. This analysis shows that when more spectra corresponding to oblique angles of illumination are rigorously computed, the situation becomes more complex for the compact model when attempting to reproduce the mask-induced effects for small pitches. This situation is addressed in a different way within the spectrum correction model (SCM), shown in the next section.
Figure 5.6: PW analysis for L/S 45 nm linewidth. The PFM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.
Figure 5.7: PW analysis for L/S 45 nm linewidth. The PFM is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm.
In terms of computation time of the model calibration, the PFM requires at least 8 nodes and 4 processors per node in order to make the computation in a period of approximately 6 to 8 hours. However, it is important to highlight that the calibration process takes into account 9 pitches in a range between 90 nm to 330 nm. After the optimization process, the pupil filters obtained have the capability of reproduce the mask-induced effects for any pitch corresponding to this range, with the limitations that have been already mentioned before.

5.2 Spectrum correction model (SCM)

Like in the case of the PFM, the calibration procedure determines the coefficients of the polynomials representing the amplitude and the phase for the SCM. This is achieved by a comparison between the scalar and the rigorous spectrum. After the calibrated coefficients are obtained, the spectra corresponding to the TE- and TM-polarization components are introduced to the vector imaging computation to calculate the aerial image.

5.2.1 Model Performance Compared to the Rigorous Model with the HA

In Table 5.3, the calibrated coefficients are shown. The coefficients of the Zernike polynomials have arbitrary units (a.u.) because the application of the filters is performed before the application of the pupil, where the normalization of the polynomials takes place. The value of the coefficients $C_1$ for both TE and TM components is strongly reduced, compared to the $C_1$ values of the pupil filters. This remarkable difference corresponds to the energy and polarization corrections performed inside the pupil by the imaging computation algorithm. In the PFM these corrections are taking into account during the optimization process of the Zernike polynomials coefficients. On the other hand, in the SCM, the corrections are performed after the calibration of the coefficients. In Figure 5.8, the phase and amplitude of the correction filters corresponding to the TE and TM components are shown. The plots are not limited by the circular pupil as in the case of the pupil filters. This represents, as mentioned before, the application of the filters directly on the scalar spectrum, before the application of the pupil. The amplitude of the TM component (Figure 5.8(a)), has its larger values in the central region and negative values at large angles. This corresponds to a 180° phase shift for diffraction orders at oblique angles of incidence. The phase of the TM component (Figure 5.8(b)) is radially symmetric, with its maximum values at very oblique angles. On the other hand, the amplitude of the TE component (Figure 5.8(c)) has negative values in the central region and maximum values at large angles. The phase of the TE component (Figure 5.8(d))
presents a strong modulation in the horizontal direction. In Figures 5.9 and 5.10, the PW analysis is shown. The scalar model is represented by the red curves and circle markers, the rigorous model is represented by the blue curves and diamond markers and the SCM is represented by the green curves and the triangle markers. In Figure 5.9, the SCM for the calibrated pitches exhibit a good agreement with the rigorous model. In the same way, the cross-validation pitches (Figure 5.10) show that the SCM approximates to the rigorous model. In general, the PW analysis shows that the SCM successfully retrieves the best focus position and threshold intensity described by the rigorous model for the complete pitch range from 90 nm to 330 nm. Additionally, a considerable improvement compared to the PFM for the smallest pitches is obtained. This can be clearly observed for the calibrated pitches of 90 nm (Figure 5.9(a)) and 120 nm (Figure 5.9(b)), as well as for the non-calibrated pitches of 100 nm (Figure 5.10(a)) and 130 nm (Figure 5.10(b)).

<table>
<thead>
<tr>
<th>Polarization</th>
<th>$C_1$</th>
<th>$C_4$</th>
<th>$C_9$</th>
<th>$Z_4$ (a.u.)</th>
<th>$Z_5$ (a.u.)</th>
<th>$Z_9$ (a.u.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM</td>
<td>0.181</td>
<td>-0.406</td>
<td>0.013</td>
<td>-0.0087</td>
<td>-0.0032</td>
<td>0.0192</td>
</tr>
<tr>
<td>TE</td>
<td>0.069</td>
<td>0.732</td>
<td>-0.093</td>
<td>-0.0148</td>
<td>-0.014</td>
<td>0.0143</td>
</tr>
</tbody>
</table>
Figure 5.8: Correction filters associated to TE- and TM-polarization components for the compensation of the mask-induced effects predicted by the rigorous model with the HA. TM component (a) amplitude and (b) phase. TE component (a) amplitude and (b) phase.
Figure 5.9: PW analysis for L/S 45 nm linewidth. The SCM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.
Figure 5.10: PW analysis for L/S 45 nm linewidth. The SCM model is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm.
5.2.2 Model Performance Compared to the Rigorous Model without the HA

The consideration of more rigorous computed spectra produces a better description of the oblique illumination effects. This is also challenging for the compact mask model since more data in the calibration process have to be taken into account. Specifically in this case, the SCM which is dealing directly with the spectrum information, a set of parameters per spectrum for each polarization component needs to be considered. That makes a total of 24 parameters to calibrate, as shown in Table 5.4. Two representative angles are considered: $\varphi_1 = 0.196027$ and $\varphi_2 = -0.196027$, the subscripts 1 and 2 make reference to the spectrum corresponding to each representative angle of illumination. The higher computation cost that requires this amount of data compared to the previous models, is compensated by a better description of the mask-induced effects. With the SCM, better results are expected for calibrated and non-calibrated pitches in comparison to the PFM. The improvement should be the result of calibrating two independently filters per polarization state corresponding to each representative illumination angle. From the point of view of the optical projection system, this model seeks to emulate an entrance pupil that is moving according to the location in the frequency space of the diffraction spectrum that is analyzed. In contrast, in the PFM, the diffraction orders are entering at different positions to pupil, which is in a fixed position.

Table 5.4: Optimized parameters for the SCM. Reference without the HA. The subscripts 1 and 2 make reference to the spectrum corresponding to each representative angle of illumination: $\varphi_1 = 0.196027$ and $\varphi_2 = -0.196027$.

<table>
<thead>
<tr>
<th>Polarization</th>
<th>$C_1$</th>
<th>$C_2$</th>
<th>$C_3$</th>
<th>$Z_4$ (a.u.)</th>
<th>$Z_5$ (a.u.)</th>
<th>$Z_9$ (a.u.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM$_1$</td>
<td>0.3592</td>
<td>-0.1858</td>
<td>0.0419</td>
<td>-0.01294</td>
<td>-0.02407</td>
<td>-0.0087</td>
</tr>
<tr>
<td>TE$_1$</td>
<td>-0.2939</td>
<td>0.2440</td>
<td>0.0896</td>
<td>-0.00867</td>
<td>-0.01597</td>
<td>-0.00296</td>
</tr>
<tr>
<td>TM$_2$</td>
<td>-0.2093</td>
<td>0.1364</td>
<td>-0.0053</td>
<td>0.00853</td>
<td>-0.00555</td>
<td>0.00369</td>
</tr>
<tr>
<td>TE$_2$</td>
<td>0.0588</td>
<td>0.2495</td>
<td>0.0013</td>
<td>-0.00671</td>
<td>0.01452</td>
<td>0.01342</td>
</tr>
</tbody>
</table>

In Figures 5.11 and 5.12, the amplitude and phase of the correction filters applied to the spectra corresponding to the each illumination angle, are shown. For $\varphi_1$, the phase of the TM (Figure 5.11(b)) and TE (Figure 5.11(d)) components have a strong modulation in the vertical direction. This is enhanced for the TM component, with the large weights given by the amplitude (Figure 5.11(a)) and attenuated for the TE component by the small amplitude values (Figure 5.11(c)). In the case of $\varphi_2$, both amplitude values of the TE and TM (Figures 5.12(a) and 5.12(c)) components, enhance the phase values at large angles. The TM phase (Figure 5.12(b)) presents a slight modulation in the horizontal direction. In contrast, the TE phase (Figure 5.12(d)) exhibits the modulation along the
vertical direction.

Figure 5.11: Correction filters associated to TE- and TM-polarization components for the spectrum corresponding to the representative angle $\varphi_1$. TM component (a) amplitude and (b) phase. TE component (a) amplitude and (b) phase.
Figure 5.12: Correction filters associated to TE- and TM-polarization components for the spectrum corresponding to the representative angle $\varphi_2$. TM component (a) amplitude and (b) phase. TE component (a) amplitude and (b) phase.
Figure 5.13: PW analysis for L/S 45 nm linewidth. The SCM model is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.
Figure 5.14: PW analysis for L/S 45 nm linewidth. The SCM model is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm.
In Figures 5.13 and 5.14, the PW analysis is shown. In general, the SCM model (green curves with triangle markers) shows a correction of the scalar model (red curves with circle markers) for most of the cases. Additionally, a considerable improvement—compared to the PFM—for the smallest calibrated (Figures 5.13(a) and 5.13(b)) and non-calibrated (Figures 5.14(a) and 5.14(b)) pitches is achieved. The PW of the SCM for large pitches like 210 nm (Figure 5.13(e)), 260 nm (Figure 5.13(f)), 315 nm (Figure 5.13(f)) and 330 nm (Figure 5.14(e)), do not report data for the largest defocus values. This limitation of the SCM can be attributed to the large number of parameters to calibrate. In this case, the search interval for each of the 24 parameters has to be carefully chosen.

In terms of calibration time, the SCM requires more processing time when the rigorous model without the HA is used, since 24 parameters have to be calibrated. When the rigorous model with the HA is used, the processing time is in the same order of the PFM. With 8 nodes and 4 processors per node, the computation of the calibration process takes approximately from 5 to 6 hours. Depending on the available resources, this computation time can be speed up to approximately from 4 to 5 hours.

5.3 ANN Approach as an Alternative Compact Mask Model

The proposed models evaluated up to now (PFM and SCM) have exhibited a correction of the scalar model. However, certain limitations either for the smallest or the largest pitches have been observed. A considerable improvement can be expected when more Zernike orders are added to the calibration procedure. However, this will result in higher complexity of the optimization and higher time consumption, limiting the performance of the model. Instead of considering analytical functions to act as correction filters, the concept of the ANN is introduced. The main differences between the ANN approach and the other proposed models, are related to the calibration procedure and the fitness function.

In the PFM and the SCM, the calibration procedure is based on the comparisons of the PW values and the CD target. When the difference between the tested and the target values has reached a minimum, the parameters that produce this convergence are chosen as the optimized parameters. In the ANN, the fitness function does not depend on the PW values or the CD target. Instead, the fitness function is determined by the difference between the output spectrum and the target spectrum. During the calibration procedure, the parameters that are optimized correspond to the weights of the connections between the layers of neurons. That is, the calibration of this model is completely determined by the learning process of the ANN.

To address the polarization dependence of the diffraction spectrum, the ANN approach
considers two different neural networks to account for the TE and TM polarization effects. The rigorous spectrum is constructed from the combination of the output of the TE and TM ANNs. This is an important consideration for the vector imaging formulation, to reproduce the behavior of the projection imaging system.

5.3.1 Hidden Units and Hidden Layers of the ANN

There is no precise guidance on how to select the number of units in each layer or the size of the ANN. Therefore, designing such a structure is usually a trial and error procedure. In most situations there is no need of having more than two hidden layers. With too few hidden units, a high training error due to under-fitting may be obtained. In contrast, the use of too many hidden units results in a low training error but an over-fitting and high variance [129][130]. There are numerous “rules of thumb” to choose the number of hidden units. Some of them relate the number of input units to the number of output units. Others express the number of hidden units in terms of the number of training cases.

In this work, a GA is utilized to find the number of hidden units and hidden layers for the ANN[131][132] for the L/S test case. The rest of the cases are addressed with a trial and error procedure.

The objective function of the GA, is a combined function of the R-squared value of the linear regression of the target (rigorous spectrum) vs. the output and the absolute value of the maximum error of the outputs with respect to the targets. The goal of the GA algorithm is to minimize the error and maximize the R-squared value.

The ANN is trained and cross-validated in each generation of the GA. The training is performed with spectrum data corresponding to arbitrary pitches, but the cross-validation step is performed with spectrum data corresponding to pitches that can be critical in the performance of the ANN. For instance, those pitches, for which new diffraction orders contribute to the image formation are considered. The optimization process yielded 2 to 4 hidden layers and a number of hidden units between 3 to 30 times the input units.

5.3.2 ANN Validation

The quality of the approach is evaluated by the resulting aerial images in terms of lithographic PWs and CD measurements. Additionally, the amplitude and phase of the spectrum obtained with the ANN are compared with the values obtained with the rigorous computation.

The effectiveness of the ANN learning process is mainly given by the appropriate selection of the training patterns and the topology of the network. In order to accomplish a satisfactory training, it is necessary to provide a unique signature to each training pattern. This aspect is studied for different mask layouts in the following sections.
Different test cases are evaluated from simple 1D structures like L/S to 2D structures such as contact holes and line ends. The ANN targets for the L/S are computed with and without the HA, whereas the targets for contact hole and line ends have been computed only without the HA. The amount of data handled in the ANN training varies from case to case. With the Hopkins assumption, a single spectrum per target is used for the training process of the ANN. On the other hand, without the HA different rigorously computed spectra per target are considered.

For L/S three different scenarios are chosen. The first case is a constant linewidth with a variable duty cycle. This is the same configuration proposed in the alpha test (See Table 4.2). In this case, the scalar spectrum varies for every pitch. Particularly, comparisons with targets computed with and without the HA are presented for this case. In the second case, a position shift in the space domain is applied to a dense L/S configuration. In this scenario, the spectrum amplitude remains constant for every shift performed in the mask layout. Then, the proper signature of the pattern is contained in the phase of the spectrum. Finally, a set of layouts with a constant duty cycle is considered. The amplitude and phase of the diffraction spectrum remain constant for combinations of linewidth and pitch values that produce a constant duty cycle. This results in a lack of signature for every training pattern. Therefore, different input data components are evaluated and compared to identify the most suitable for the ANN training process.

Contact holes and line ends are analyzed with a different approach. For a range of illumination angles, the corresponding diffraction spectra for the training process are considered. Without the HA, the volume of data to be handled by the ANN increases. In order to keep up the efficiency of the ANN approach, only one mask layout per ANN is considered in this test case.

2D test cases are considered to evaluate the capability of the ANN approach. However, they are not used to perform comparisons with the other proposed models. 2D tests were performed as well with the other models, but are not presented in this thesis. 2D structures require some additional effort in terms of OPC and computation time that are not suitable when making a proof-of-concept study. This may be included in a phase of future work.

5.3.3 L/S with Variable Duty Cycle

The alpha test is used. 90 nm to 330 nm pitches constitute a configuration of dense, semi-dense and isolated features. Other imaging conditions are described in Tables 4.3 and 4.4.
5.3.4 Model Performance Compared to the Rigorous Model with the HA

In the Hopkins approach, the diffraction spectrum is computed only for the normal incidence. Therefore, the training of the ANN involves one rigorously computed spectrum target per scalar spectrum input for every pitch value. Two independent ANNs are trained to account for polarization effects. The ANNs for TE- and TM-polarization states are trained for pitches in the range of 90 nm to 330 nm every 30 nm and cross-validation step is performed every 5 nm. In general, the ANNs for both TE- and TM-polarization states exhibit a good behavior in terms of amplitude and phase difference of the 0\textsuperscript{th} and 1\textsuperscript{st} diffraction orders (Figure 5.15). The blue curves represent the target, the red curves represent the Kirchhoff approach and the green curves represent the results of the ANN. In Figure 5.15, the ANN accuracy is demonstrated for both trained and non-trained spectrum inputs corresponding to different pitch values. All the relevant diffraction orders are used to train the ANN. But only the 0th and 1st diffraction orders are compared in the plots. The mask induced effects that need to be mimicked with the ANN approach, can be already observed in the diffraction efficiency and phase difference represented by these 2 orders.
Figure 5.15: Comparison of ANN, scalar and reference models in terms of amplitude and phase of the diffraction spectrum. (a) amplitude difference for the TE-polarization component, (b) phase difference for the TE-polarization component, (c) amplitude difference for the TM-polarization component and (d) phase difference for the TM-polarization component.

The aerial images used for the PW analysis are obtained, as for the previous models, using the Abbe formulation of partially coherent imaging. Figures 5.16 and 5.17 show the PW analysis for trained and non-trained pitches, respectively. In both cases, a very good match with the reference PW is obtained. It is worthwhile to mention that the limitations demonstrated by the PFM and the SCM, specially for the smallest pitches, are overcome by the ANN approach. The success of the learning process of the ANN is reflected in its capability to reproduce the EMF effects, regardless the value of the pitch.

The dimension of the input and output data to be analyzed by the ANN depends on the number of diffraction orders that enters the pupil of the optical system. For coherent light, e.g. light normally incident on a periodic structure, the number of propagating diffracted orders are given by Equation 5.1.
Figure 5.16: PW analysis for L/S 45 nm linewidth. The ANN approach is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.

\[ m = \left\lfloor \frac{p \cdot NA}{\lambda} \right\rfloor, \]  

(5.1)

where \( p \) is the pitch, \( NA \) is the numerical aperture of the optical system and \( \lambda = 193 \text{ nm} \) the wavelength, \( \lfloor x \rfloor \) represent the floor function and gives the largest integer less than or equal to \( x \). Taking into account that the input spectrum is based on scalar computations and the light source is partially coherent, additional considerations of the number of orders
that are captured be the pupil are given by the Equation 5.2.

\[ m = \left\lfloor \frac{p\, NA \, (1 + \sigma)}{\lambda} \right\rfloor, \]  

(5.2)

where \( \sigma \) is the partial coherence factor. The dimension of the input data changes for different pitch values. As the dimension of every input of the ANN is expected to be constant, the largest possible dimension given by the largest pitch value (330 nm) is taken.
as the dimension of all the ANN inputs. The patterns with lower number of propagating
diffraction orders will be padded with zeros to match the largest dimension. This can be
thought of as an application of a filter to the spectrum data. For the smallest pitches,
less orders will be considered during the learning process.

An alternative to avoid padding with zeros is to group the layouts that contribute
with the same number of diffraction orders. Then, different ANNs for each subset of
pitch values are trained. Excellent results were obtained applying this technique, but the
predictability and performance of the approach are compromised since more ANNs are
trained for particular tasks.

5.3.5 Model Performance Compared to the Rigorous Model without the HA

Without the HA a set of representative angles over the illumination source are consid-
ered. The diffraction spectrum is rigorously computed for the corresponding angle of
illumination. For a dipole illumination, one representative angle is considered for each
pole. The illumination direction is defined by a pair of angles. One is the incidence angle
$\varphi$ of illumination with respect to the optical axis and the other is the azimuth angle $\vartheta$ of
illumination in the source plane. In this particular case, the L/S are y-parallel, that is,
perpendicular to the optical axis. Therefore, the dipole illumination is x-parallel and the
representative angles are chosen to lie on the x-axis as shown in Figure 5.18. The values
of the considered angles are presented in Table 5.5.

![Figure 5.18: Schematic representation of the representative angles for x-parallel dipole illumination.](image)

The ANN has to be trained to account for the effect of oblique incidence angles of
illumination. For this purpose, the input vector for the ANN training process has to
include the information corresponding to each angle of incidence. In this way, the input
vector to the ANN contains the scalar diffraction spectrum and the pair of angles that define each representative illumination direction. Additionally, the target becomes the combination of the rigorous spectra for the relevant incidence angles (Figure 5.19).

The ANNs for TE- and TM-polarization states are trained for pitches in the range of 90 nm to 330 nm every 30 nm, and a cross-validation step is performed every 5 nm. The results of the training and cross-validation process are shown in Figure 5.20. Here, the real and imaginary part of the 0th and 1st diffraction orders of the ANN approach and the reference are compared. Accurate results are obtained along the complete range of pitches.

Figures 5.21 and 5.22 show further comparisons in terms of PW, for trained and non-trained pitches, respectively. The focus shift and symmetry, obtained with the ANN approach, reproduce the behavior of the reference model. The simulations without the HA show much more pronounced focus effects (mask-induced aberrations) compared to the simulations with the HA. The ANN approach reproduces this behavior in a very accurate way, regardless the value of the pitch.

The consideration of a larger number of representative angles of incidence implies a higher volume of data to be taken into account in the training process of the ANN. The number of units in the output vector increases for a larger number of incidence angles since the rigorous spectrum for each representative angle is considered. Therefore, a more suitable representation of the input and output vectors of the ANN has to be studied when the reference is computed without the HA for a large number of incidence angles.

In general, the performance of the ANN approach is outstanding compared to the previous models. The PW analysis exhibits an accurate behavior compared with the rigorous model. This is a remarkable result since important properties of the PWs, such as best focus and threshold, are not included in the objective function for the calibration procedure of the ANN.
Figure 5.20: Real and imaginary part of the 0th and 1st diffraction orders. ANN approach compared to the reference. (a) TE-polarization component of the illumination angle corresponding to $(\theta_1, \varphi_1)$. (b) TE-polarization component of the illumination angle corresponding to $(\theta_2, \varphi_2)$. (c) TM-polarization component of the illumination angle corresponding to $(\theta_1, \varphi_1)$. (d) TM-polarization component of the illumination angle corresponding to $(\theta_2, \varphi_2)$. 
Figure 5.21: PW analysis for L/S 45 nm linewidth. The ANN approach is calibrated for the presented pitches. (a) pitch=90 nm, (b) pitch=120 nm, (c) pitch=150 nm, (d) pitch=180 nm, (e) pitch=210 nm, (f) pitch=330 nm.
Figure 5.22: PW analysis for L/S 45 nm linewidth. The ANN approach is validated for different pitches that are not used in the calibration. (a) pitch=100 nm, (b) pitch=130 nm, (c) pitch=170 nm, (d) pitch=190 nm, (e) pitch=260 nm, (f) pitch=315 nm.
5.3.6 Shifted Lines

For this test case, the references are computed with the Hopkins assumption. A spatial shift of a L/S layout is performed. A linewidth of 45 nm and pitch of 90 nm are considered. The shift varies in the range of one period in the mask from $-45$ nm to 45 nm. The ANN is expected to reproduce the effect of shifting the line in one mask period. For this purpose, training and cross-validation data include the spectrum phase of the shifted lines and the shift value itself. The spectrum amplitude is neglected since the distribution of energy remains constant for shifts performed in one period of the mask.

Figure 5.23: Cross-section at a focus 0.0 nm for negative shifts. (a) 45.0 nm shift corresponds to a trained shift. (b)-12.5 nm and (c)-2.5 nm shifts belong to the cross-validation data.

Figure 5.24: Cross-section at a focus 0.0 nm for positive shifts. (a)17.5 nm, (b)22.5 nm and (c)32.5 nm shifts belong to the cross-validation data. Other imaging parameters are describe in Tables 4.3 and 4.4

In Figure 5.23(a) a perfect matching between the ANN outcome for trained data and the rigorous model is observed. The rest of the plots in Figures 5.23 and 5.24 represent the behavior of non-trained data. Figures 5.23(b) and 5.24(b) show a slight deviation of the ANN outcome from the rigorous model in contrast with Figure 5.24(c) where the ANN exhibits a poor performance. From this results it can be inferred that for an appropriate description of shifted lines using the ANN approach, different ANNs must be trained for smaller ranges of shifting.
5.3.7 L/S with Constant Duty Cycle

In the Kirchhoff model, the scaled energy distribution of the diffracted field and the phase remain constant for a constant line/space ratio. That means that the amplitude and the phase inputs are constant for every pitch. This scenario represents a great challenge for the ANN approach. Training the ANN with constant input yields marginal results in the learning process. The definition of an input with a unique signature for the training process then becomes critical.

Tests are conducted to provide the best choice of input to feed the ANN. In this section, the results for duty-cycles 1:1 and 1:3 are presented and compared against the reference computed with the HA. Additionally, the behavior of the network for different signatures of the input is analyzed. Training and cross-validation data of the ANN is grouped according to the number of diffraction orders collected in the entrance pupil of the optical projection system (Table 5.6). Training data includes pitch values with a step size of 5 nm in the pitch range. Cross-validation step is performed for a larger number of pitches with a step size of 0.5 nm in the pitch range.

Table 5.6: linewidth and pitches used in each duty-cycle.

<table>
<thead>
<tr>
<th>Duty-cycle</th>
<th>linewidth (nm)</th>
<th>pitch (nm)</th>
<th>N. of diff. orders</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:1</td>
<td>45.0 - 75.0</td>
<td>90.0 - 150.0</td>
<td>3</td>
</tr>
<tr>
<td>1:2</td>
<td>50.0 - 75.0</td>
<td>150.0 - 225.0</td>
<td>5</td>
</tr>
<tr>
<td>1:3</td>
<td>57.0 - 75.0</td>
<td>228.0 - 300.0</td>
<td>7</td>
</tr>
</tbody>
</table>

For a duty-cycle of 1:1 and 1:3, the amplitude and phase difference of the 0th and 1st diffraction orders of the spectrum for the TM- and TE- polarization components are plotted in Figures 5.25(a)-5.25(d) and Figures 5.26(a)-5.26(d). The ANN shows a poor performance when it is trained with constant spectrum phase and amplitude values only, as shown by the blue curves. In this case, the output of the ANN oscillates around an average point in the interval of pitches, hardly retrieving the behavior of the target (black curves). Adding the pitch value to the input data (yellow curve) improves the behavior of the ANN for the TM- and TE-polarization components. The spectrum phase and pitch value input (green curves) shows a better approximation to the target compared to the input composed of the spectrum amplitude and the pitch (cyan curves). Finally, if only the pitch is given as input to the network (red curves), the ANN accuracy is high as shown in Figures 5.25(a)-5.25(d) and Figures 5.26(a)-5.26(d). However, this is not a practical solution since a mask layout is not described strictly by just one pitch value. The second best performance is provided by the input composed of the spectrum phase and pitch value. For practical purposes, this represents a more suitable solution for a constant duty-cycle layout.
Figure 5.25: Spectrum amplitude and phase comparisons for TM- and TE-polarization, duty-cycle 1:1. Blue squares represent the training data composed of phase and amplitude. Yellow diamonds represent the training data composed of phase, amplitude and pitch. Cyan circles represent the training data composed of amplitude and pitch. Green triangles represent the training data composed of phase and pitch. Red triangles represent the training data composed only of the pitch value. Black triangles represent the references. TE-polarization component (a) amplitude and (b) phase. TM-polarization component (c) amplitude and (d) phase. Other important data of this test case is described in Chapter 4, Section 4.1 Simulation parameters.

5.3.8 Contacts

For contact holes the references are computed without the HA. This requires rigorous computation of the spectrum for all the representative illumination incidences. In the ANN approach, only a few angles of illumination and the scalar spectrum are used to trained the network. In the cross-validation step, it is possible to retrieve the rest of the spectra corresponding to the different angles of illumination. From 49 angles of illumination, only 9 angles are used for the training process, then in the cross-validation phase all the 49 spectra corresponding to the total illumination angles are obtained with the ANN. The input of the ANN is composed of the scalar spectrum together with a pair of angles. One is the incidence angle of illumination with respect to the optical axis and the other is the azimuth angle of illumination in the source plane.

For the training of the ANN in the case of 2D structures, more diffraction orders
Figure 5.26: Spectrum amplitude and phase amplitude comparisons for TM- and TE-polarization, duty-cycle 1:3. Blue squares represent the training data composed of phase and amplitude. Yellow diamonds represent the training data composed of phase, amplitude and pitch. Cyan circles represent the training data composed of amplitude and pitch. Green triangles represent the training data composed of phase and pitch. Red triangles represent the training data composed only of the pitch value. Black triangles represent the references. TE-polarization component (a) amplitude and (b) phase. TM-polarization component (c) amplitude and (d) phase. Other important data of this test case is described in Chapter 4, in Section 4.1 Simulation parameters.

need to be taken into account in order to reconstruct an aerial image similar to the one obtained with the reference method. The directions to extract the spectrum diffraction data for the training process are represented in Figure 5.27 as white dotted lines. Due to the symmetry of the contacts array, it is possible to consider only 2 directions in the spectrum for training purposes. In the reconstruction of the spectrum obtained with the ANN, additional data of the diffraction spectrum are retrieved either by rotation or reflection of the extracted directions. The training of more data leads to a larger time in the ANN learning process. The lithographer has to cope with the decision of how much data should be included in the training process in order to keep the ratio performance vs. accuracy as balanced as possible.

The results of the training process and the cross-validation step for the horizontal direction of the extracted data are shown in Figure 5.28. The TE and TM- polarization
Figure 5.27: White dotted lines indicate the data extracted from the scalar and rigorous spectrum to conform the pool of training and cross-validation data for contact holes. (a) TE spectrum amplitude for vertical incidence and (b) TE spectrum phase for vertical incidence.

components of the spectrum vs. the No-Hopkins orders\(^1\) show a good agreement with the reference data for the non-trained and the trained orders. In Figures 5.28(a) and 5.28(b) the phase and amplitude of the spectrum for the TE polarization are shown. The obtained results indicate that the ANN can retrieve both the amplitude and the phase of diffraction orders between \(-3\) and \(3\) with reasonable accuracy. The ANN is sensitive enough to recognize very small differences in the order of \(10^{-3}\). A similar behavior is observed for the case of TM polarization in Figures 5.28(c) and 5.28(d). When these spectra are combined to reproduce the final aerial image, a good agreement is obtained in comparison with the reference, as shown in Figure 5.29. The remaining difference between the aerial images observed in Figure 5.29(c) can be attributed to the assumptions made to extract the training data. For a higher correlation between the images more data of the reference and scalar spectrum need to be considered.

\(^1\)In this context, No-Hopkins orders refers to the number of spectra that are considered corresponding to the representative angles of illumination.
5.3.9 Line Ends

Line ends require 2D simulations that involve a larger amount of data to process. This results in a much more intense effort in the ANN training compared to the 1D cases. A mask layout with 100 nm linewidth, 200 nm pitch and 28 nm gap is considered, see Figure ???. The references for this case are computed without the Hopkins approximation. 4 different ANN are trained for the $x$ and $y$ axis of the TE- and TM- polarization components.
Additional information from the diffraction spectrum has to be extracted to constitute the pool of data for the training process. In Figures 5.30(a) and 5.30(b), the directions to extract the spectrum diffraction data are represented as white dotted lines.

Figure 5.30: White dotted lines indicate the data extracted from the scalar and rigorous spectrum to conformed the pool of training and cross-validation data for line ends. (a) TE spectrum phase for vertical incidence (reference).  (b) TM spectrum phase for vertical incidence (reference).

9 spectra corresponding to different incidence angles are used in the training process of each ANN. 49 spectra of the remaining incidence angles are obtained in the cross-validation step. The comparison of the difference of the 0th and 1st diffraction orders showed a good performance for both trained and non-trained data. Aerial images are shown in Figure 5.31. In Figure 5.31(a) the rigorous aerial image is shown. Figures 5.31(b)-5.31(d) show the aerial images obtained with the ANN approach. When more diffraction orders per direction are considered in the training process, the aerial image retrieved with the ANN approach gets closer to the aerial image computed with the reference model. This results in a considerable growth of the time consumption in the training process since more data is included.

In general, the calibration times of the ANNs used in this test cases study, were remarkably short. A learning process, including 5000 epochs, lasted approximately 15 min. Additional time is required when the topology of the ANN is optimized.
Figure 5.31: Aerial images of line ends. Linewidth 100 nm, pitch 200 nm and gap of 28 nm. (a) 125 diffraction orders in each direction (Ref.). (b) 5 diffraction orders in each direction. (c) 11 diffraction orders in each direction. (d) 17 diffraction orders in each direction.

5.4 Comparisons Between the Models

In order to compare the models, the alpha test described in Section 4.1 is used. In general, the ANN approach exhibits a better performance compared with the rest of the compact mask models\cite{133}. In Figure 5.32 the capability of the ANN to predict best focus shifts can be observed. The scalar (Kirchhoff) model does not report any useful information about the best focus behavior. The ANN approach shows the best performance followed by the boundary layer and the spectrum modification model. The PFM shows a similar behavior to the reference model but it is less accurate than the ANN. The PM predicts a focus behavior different from the scalar model but differs also significantly from the reference model.

The CD-through-pitch behavior can be observed in Figure 5.33. The results of the PM vary significantly from the rigorous model, while the ANN reproduces with a high accuracy its behavior. The rest of the models are found in between the BLM and the PM. The behavior of the SCM is similar to the reference model, only for pitches smaller than 160 nm. In contrast, the ANN approach reproduces the behavior of the reference model...
Figure 5.32: Comparisons of the compact models in terms of best-focus-through-pitch for L/S, 45 nm linewidth and pitches from 90 nm to 330 nm.

Figure 5.33: Comparisons of the compact models in terms of CD-through-pitch for L/S, 45 nm linewidth and pitches from 90 nm to 330 nm.

for the complete range of pitches.

All models exhibit a good behavior in terms of threshold-to-size-through-pitch (Figure 5.34). The investigated models present a very good threshold-to-size correction for semi-dense configurations from 150 nm up to 330 nm. The most dense pitches are challenging
Figure 5.34: Comparisons of the compact models in terms of threshold-to-size-through-pitch for L/S, 45 nm linewidth and pitches from 90 nm to 330 nm.

for most of the models, except for the ANN approach and the BLM, which reproduce the reference model behavior through the complete range of pitches.

Table 5.7 qualitatively summarizes the advantages and limitations of the proposed models in terms of the analyzed lithographic performance. Each cell represents the quality of the model behavior. The letter “L” in red cell color indicates a limited behavior, the letter “A” in gray cell color represents an acceptable behavior, and the letter “O” in green cell color stands for an optimum behavior. N.A. stands for “not applicable”, implying that no calibration is required for the scalar model. The extendibility to complex layouts, 6th column in the Table 5.7, refers to the ability of the model to retrieve mask-induced effects from more complex mask layouts with 2D and 3D features. With the BLM, PM, PFM and SMC, this aspect was investigated in simulations with contact holes. With the ANN, 2D tests were done with contacts and line-ends. The BLM and PM have less parameters to calibrate, which results in a straightforward application of the models for 2D layouts. However, when the features are more complex than just contact holes, the effects of the boundary layers and pulses may get overlapped. The results also show that the PFM can be directly applied to 2D features without the requirement of more data in the calibration process. The SMC, presents limitations when extended to 2D features. This is due to the dependence of the model on the illumination source and the number of representative angles chosen to compute the spectrum. Consequently, as 2D features require more complex illumination shapes, the SCM has to be calibrated for a larger number of parameters for each spectrum. The ANN extendibility is limited because more
data is required for the learning process.

Table 5.7: Qualitative comparison of the proposed models. “L” in the red cell indicates a limited behavior. “A” in the gray cell represents an acceptable behavior. “O” in the green cell stands for an optimum behavior. ‘N.A.’ indicates that the calibration time factor does not apply to the scalar model, since a calibration process does not take place in this case.

<table>
<thead>
<tr>
<th>Model</th>
<th>B. Focus through pitch</th>
<th>Th. to size through pitch</th>
<th>Calibration Time through pitch</th>
<th>CD through pitch</th>
<th>Extendibility to complex layouts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalar</td>
<td>L</td>
<td>L</td>
<td>N.A.</td>
<td>A</td>
<td>O</td>
</tr>
<tr>
<td>BLM</td>
<td>A</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>A</td>
</tr>
<tr>
<td>PM</td>
<td>L</td>
<td>O</td>
<td>O</td>
<td>L</td>
<td>A</td>
</tr>
<tr>
<td>PFM</td>
<td>A</td>
<td>A</td>
<td>L</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>SCM</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>O</td>
<td>A</td>
</tr>
<tr>
<td>ANN</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>A</td>
</tr>
</tbody>
</table>

In the following and last chapter, the conclusions and discussion are presented. Some final statements about the proposed models are pointed out and the most important results are highlighted. Additionally, further alternatives are proposed, based on the findings of this study.
Chapter 6
Conclusions

Rigorous mask models are based on the solution of Maxwell’s equations for the electromagnetic fields in 3D structures. They are suitable to accurately describe the propagation of light through a 3D mask that contains features sizes comparable to the wavelength. However, their high computational requirements (memory and speed) make them inappropriate for tasks that involve complex structures in large area (thousands of µm) simulations or many computations of mask diffraction spectra for adjusting mask geometry parameters, such as in OPC techniques. In practice, the design of a mask that satisfies the requirements of a sufficient image quality has to go through an optimization process, where the accuracy and speed of the mask spectrum computation is an essential component.

Compact mask models offer an important alternative to speed up the imaging simulations by considering the mask optical material properties and the 3D mask geometry or mask-induced effects. These compact mask models improve the accuracy of the fast Kirchhoff-based imaging model, where the mask diffraction spectrum is computed using the scalar diffraction theory. The scalar spectrum is modified and enhanced to account for the mask-induced effects and used in the vectorial imaging computation to yield similar results as the fully rigorous simulations, where the spectrum is computed using the Maxwell’s equations. This provides a compromise between speed and accuracy that results beneficial for many techniques of practical interest in optical lithography, including optical resolution enhancement techniques.

For the evaluation of the mask models, a typical AttPSM (6% attenuation) is considered. Lines and spaces, contact holes and line-ends layouts were chosen for this study because they represent basic structures that are the main components of more complex mask geometries. An alpha test was defined and used for the comparison of the proposed mask models. It is a lines and spaces layout with a linewidth of 45 nm and pitch ranging from 90 nm to 330 nm. An aerial image-based analysis is performed in order to study each compact mask model. The process window (PW) corresponding to each model is
compared to the PW of the rigorous model. The PW analysis is a very convenient evaluation method, mainly because of two reasons. On the one hand, the PW allows the study of the image formation for a certain range of focus values. On the other hand, the impact of the mask-induced effects can be easily detected in the PW plots (e.g., focus shift, threshold difference, tilting of the PW). The values obtained from the PW analysis, i.e., the intensity threshold, the best focus value and the threshold latitude, are used in the fitness function, within the calibration procedure for the parameters of the compact mask models. The artificial neural network (ANN) approach uses an alternative merit function in the model calibration procedure. Here the spectrum represents the fitness function, which is optimized in the learning process of the ANN.

From this study, it can be concluded that all compact mask models improve the accuracy of the scalar model. The known compact mask models such as the boundary layer model (BLM) and the pulses model (PM), retrieve the mask-induced effects with some limitations. The BLM exhibits a remarkable behavior compared with the reference using the Hopkins assumption (HA). However, when the more accurate reference without the HA is used to compare the BLM, some important differences can be noticed, specially for the smallest pitches. The PM is extremely pitch-sensitive and its limitations can be demonstrated in both cases, using the HA and without the HA. Despite of this fact, the PM has an important advantage over the BLM, since its implementation is performed directly in the frequency domain, the cross-talk between the pulses for dense layouts is avoided.

The main accomplishment of this thesis is the formulation of new compact mask models that overcome the limitations observed in the BLM and PM. To this end, different approaches in the frequency domain were evaluated and compared. In the pupil filtering model (PFM), the mask-induced effects are interpreted as wavefront aberrations in the optical projection system. The description of the aberrations is carried out through the Zernike polynomials within a Jones pupil representation. In this way the mask-induced phase and polarization effects are combined to achieve an accurate description. The mayor contribution to the phase compensation is done by the spherical aberration coefficient $Z_9$ in both Jones components $J_{xx}$ and $J_{yy}$. It was also found that the defocus $Z_4$ and astigmatism $Z_5$ coefficients have a non-negligible contribution to the phase correction of the spectrum. In terms of the amplitude, the constant correction factor provides the largest contribution to the compensation. It describes the loss of light due to the scattering from the mask edges, which is not properly considered in the Kirchhoff model. The results obtained with this approach demonstrated that there is a satisfactory correction to the scalar model, specially for the largest pitches. This applies both to image simulations with and without the HA. However, for the smallest pitches using the reference without the HA, the corrections are not satisfactory. This happens because the optimized filters of
the PFM are described in terms of analytical functions and act globally on the spectrum, instead of adapting according to a specific value of the pitch or to the incidence angle of illumination. Therefore, this model is not flexible enough to cover the mask topography effects both for large and small pitches. The consideration of higher order aberration components, could provide an improvement of the behavior of the model, but would also impact the performance of the model in terms of the calibration time.

Based on the previous results, the spectrum correction model (SCM) arises as a new alternative to improve the behavior of the PFM. The SCM uses the same Zernike polynomials representation to filter the scalar spectrum. However, the modification of the scalar spectrum is done before the light gets into the entrance pupil of the optical projection system. Additionally, different filters are optimized to reproduce the spectra corresponding to all relevant illumination angles of the mask. This provides a higher flexibility to the model, which can be mainly noticed when the reference without the HA is used. In this case, a considerable improvement for the smallest pitches is obtained, compared to the PFM. The advantage that represents this type of filtering to the spectrum, is that the oblique illumination effects can be addressed by an independent calibration for each filter. This improves the description of the oblique illumination effects, but makes the SCM computationally more expensive than the PFM, since more information in the calibration process is required.

From the PFM and the SCM results, two important statements can be made. First, a compact mask model with higher flexibility is obtained at a higher computational cost. Second, analytical filters showed an acceptable behavior, considering that the mask-induced effects are retrieved for a large range of pitches. However, they may not represent an optimum solution to completely grasp all the mask-induced effects. A third solution is proposed, based on the concept of artificial neural networks (ANN). The combination of linear and non-linear characteristics of this artificial intelligence approach, provides higher flexibility to account for the mask-induced effects. For both references, i.e., with and without the HA, the ANN approach achieves a very good reproduction of mask-induced effects for the complete range of pitch values. In a very outstanding way, the ANN approach presents the highest performance and accuracy in comparison to the other compact mask models. Thus, the PW analysis exhibits an excellent agreement with the rigorous model. This is a remarkable outcome, since important figures of the PW, such as best focus and threshold, are not considered in the calibration procedure of the ANN. Additionally, in order to prove the capability of the method, other mask layouts were tested (e.g., contact holes and line-ends). The obtained results demonstrated the great adaptability that characterize the ANN. Robustness and predictability are the major advantages of the ANN approach over the rest of the analyzed compact models.

The calibration time determines the CPU time consumption for each compact mask
model. These times are difficult to compare though, due to the difference in data processing of each model. Regarding the calibration procedure, the ANN approach together with the BLM and PM are the fastest. The SCM and PFM are computationally more expensive since more parameters have to be optimized. However, once the pre-processing of data and calibration are performed, the image computation time for all the compact models is reduced in comparison with the rigorous method.

Comparisons in terms of best-focus-through-pitch, CD-through-pitch and threshold-to-size-through-pitch demonstrated a superior behavior of the ANN approach. For the training process, the ANN is provided with data from the scalar and rigorous diffraction spectrum. As the complexity of the layout increases (e.g., contact holes and line-ends), more input values are required. This results in a lower efficiency of the ANN approach.

Summarizing the most important aspects of the compact mask models, it can be stated that the BLM, the PM, the SCM and the PFM, are compact models based on physical effects and observations. On the one hand, the BLM and PM are models related to the description of local edge effects of the mask. On the other hand, the mask-induced effects expressed as aberrations, are the key of the foundation of the SCM and the PFM. In contrast to that, the ANN approach does not model any physical phenomenon but simply works as a filter to the scalar spectrum. That filter is set up in the training process and consists of a collection of activation functions and weighted connections.

The ANN approach must be addressed in a different way, when it comes to complex mask layouts. The accuracy of the model for basic mask layouts has been demonstrated. Hence, a decomposition in the mask domain into basic features may provide a promising alternative. With this in mind, it can be said that the ANN approach offers a new possibility for the compact mask models in lithography simulation.

**Future Research**

The word “flexibility” may be controversial in this study, since it has both positive and negative connotations in the outcome of a compact mask model. From one side, the flexibility of the model allows a higher predictability and robustness in the description of the mask-induced effects. On the other hand, flexibility has a higher price to pay in terms of computational costs. A compromise between this positive and negative aspects has to be found and it is indeed one of the reasons of this thesis.

It may be desirable that the compact mask model offers answers to questions that arise from the physical phenomena, involved in the propagation of light at micro and nano scales. Based on the results obtained in the previous simulations, the BLM and the PFM considered in one hybrid model, completely give answer to the local effects
generated by the mask features edges as well as to the wavefront deformation in the pupil of the optical projection system due to the 3D mask topography. The independence of the calibration procedure of these two models from the illumination source, is beneficial for the calibration time of such hybrid model. However, oblique illumination effects have to be covered by the compact mask model as well. For this purpose, the analytical function describing the pupil may be optimized by subsets of pixels in order to have a bigger control over the diffraction orders arriving at the pupil. That is, the effects of oblique illumination combined with those of the mask topography can be enhanced or weakened in a more selective way along the pupil of the optical projection system.

The extendibility of the ANN approach to complex mask layouts is challenging. It has been shown that for contact holes and line-ends, a bigger effort in the training process is required. More complex mask layouts are translated in more diffraction orders that are coupled. This increases the amount of data necessary to accomplish a successful training, since the learning process depends directly on the spectrum. To overcome this situation, different techniques may be applied. First, the learning process could be addressed in a different way. Compared to the PFM and the SCM, the ANN approach utilizes a different fitness function, which is expressed in terms of the target and the output spectrum. At a first glance, one way to reduce the amount of data to compare, is to contemplate a set of figures that represent the spectrum. That is, instead of using the complete spectrum during the ANN training, only a representative sample of it, is used. For example, a vector describing the position of the first diffraction orders, together with the information of phase and amplitude of those specific orders. In this way, the data that has to be processed is considerably reduced. A second strategy, could be the combination of spectrum-based and image-based fitness functions in the calibration process of the proposed mask models. If the spectrum figures and the PW data are used in the ANN training, the flexibility of the model is further increased. Consequently, great advantages may be obtained, since the calibration process will take into account the spectrum and the image simultaneously. This has to be carefully analyzed since one of the major advantages of the ANN approach, is the highly accurate results in terms of the PW. This was achieved without the use of the PW data in the fitness function. However, in the case where the spectrum is reduced to a set of representative figures, the introduction of the PW values in the fitness function could result in the improvement of the ANN training. Finally, a decomposition of the mask layout in basic features may be considered. Some work has been done in this direction but a final conclusion has not been drawn yet. The ANN has been trained with the diffraction spectrum of lines and spaces and contact holes at the same time, to retrieve the diffraction spectrum of line ends. Additional work is required to determine rules for the layout decomposition, that prove to be advantageous for the learning process.
List of Publications


Conference and Workshop Contributions

- 5th EOS Topical Meeting on Advanced Imaging Techniques, "Mask models for the imaging of contact holes in optical projection lithography". Engelberg, Switzerland 2010 (Poster).

- 8th Fraunhofer IISB Lithography Simulation Workshop, "Compact mask model Based on pupil filtering for optical projection lithography". Hersbruck, Germany 2010 (Presentation).

- SPIE Advanced Lithography Symposium, Optical Microlithography XXIV Conference, “Modeling of mask diffraction and projection imaging for advanced optical and EUV lithography”. San Jose, USA 2011 (Presentation).

- 9th Fraunhofer IISB Lithography Simulation Workshop, "Mask model optimization in spatial and frequency domains: Enhancement of the scalar model to retrieve mask-induced effects”. Hersbruck, Germany 2011 (Presentation).

- SPIE Advanced Lithography Symposium, Optical Microlithography XXV Conference, “Evaluation of various compact mask and imaging models for the efficient simulation of mask topography effects in immersion lithography”. San Jose, California, USA 2012 (Presentation).

- 10th Fraunhofer IISB Lithography Simulation Workshop, “Application of artificial neural networks to a compact mask model optimization”. Hersbruck, Germany 2012 (Presentation).


References


REFERENCES


REFERENCES


REFERENCES


REFERENCES


[110] M. Kempsell-Sears and B. Smith, “Modeling the effects of pupil-manipulated spherical aberration in optical nanolithography,” *J. Micro/Nanolith. MEMS MOEMS*, vol. 12, no. 1, pp. 013008(1)–013008(8), 2013. 54


