INTERFEROMETRIC LIGHT MICROSCOPY FOR WAFER DEFECT INSPECTION AND THREE-DIMENSIONAL OBJECT RECONSTRUCTION

BY

RENJIE ZHOU

DISSEPTION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering in the Graduate College of the University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:
Associate Professor Lynford L. Goddard, Chair
Associate Professor Gabriel Popescu
Professor Paul Scott Carney
Professor Minh N. Do
Professor John A. Rogers
ABSTRACT

The first topic of the dissertation is semiconductor wafer defect inspection. We developed a highly sensitive defect inspection system based on laser interferometric microscopy, called epi-illumination diffraction phase microscopy (epi-DPM), to measure both the phase and amplitude of the scattered field from the wafer. To detect deep sub-wavelength defects, we further reduced the noise in the images by developing an image post-processing method, called 2DISC (2nd-order image difference, image stitching, and convolution). With the 2DISC method, we examined a 22 nm node intentional defect array (IDA) wafer. The results showed that we can reliably detect defects down to areas as small as 20 nm by 100 nm. Moving forward, we adapted our epi-DPM system to inspect a densely patterned 9 nm node IDA wafer, a significantly more difficult task. To improve our system’s sensitivity, we replaced the old 532 nm solid-state laser with a new 405 nm diode laser which has 10x smaller noise. Using the 2DISC method, we were able to detect defects down to 15 nm by 90 nm.

In order to further increase our system’s sensitivity, we proposed several approaches. The first approach utilizes precision z-scans to produce a three-dimensional (3D) wafer image, and looks at different cross-sectional planes in this 3D image. The second approach utilizes a spatial light modulator (SLM) to make a dark-field filter to selectively filter the image, such that we maintain the defect signal and suppress the wafer’s underlying structure. Therefore, after increasing the power of the incident light, the CCD (charge-coupled device) camera’s full dynamic range can be used to measure the signal coming from the defect. For a preliminary demonstration, we used an inverse filter in a bright-field imaging system to high-pass the imaging beam at the Fourier plane to remove the low frequency laser speckle noise. Our experimental results demonstrated detection sensitivity improvements with this inverse filter dark-field imaging. In order to find the best filter for the dark-field inspection system, we also worked on simulating dark-field and
bright-field imaging for the 22 nm and 9 nm node IDA wafers. The third approach is using a white-light interferometric imaging system, which is expected to have better image contrast due to its speckle free nature. Recently, we finished building a white-light epi-illumination DPM (epi-wDPM) system. This system has the modalities of performing 3D scanning and dark-field filtering. A real-time inspection software and a table-top clean-room have also been built for this system. We have been working on implementing this system for 9 nm node wafer inspection, with the goal to break the previous defect detection limit.

The second topic of the dissertation is solving the inverse scattering problem for 3D object reconstruction in quantitative phase imaging (QPI). Three-dimensional optical reconstruction typically uses a laser light source, suffering from speckle noise, and the measurements are usually done in the far zone by scanning the laser angles. Building on the ideas of laser diffraction tomography, we developed a new tomographic reconstruction method, called white-light diffraction tomography (WDT). WDT is speckle free and reconstructs the depth dimension by simply scanning through the focus of an object. The depth resolution of WDT is limited by the coherence property of the white-light source. With this approach we successfully reconstructed cells in 3D with resolution beyond the diffraction limit. The wavevector space method used in WDT can be applied to solve the inverse scattering problem in systems under the Fresnel approximation, such as optical coherence tomography (OCT) and angle-resolved low-coherence interferometry (aLCI). We have also proposed using our wavevector space method to calculate optical trapping force, lens focusing, and light diffraction by an aperture, to better understand optical resolution, and to solve the light diffusion and time-reversal problems in thick tissue.
To my Mother and my Father, for their unconditional love
ACKNOWLEDGMENTS

I would like to acknowledge everyone who helped me over the years. Due to the page limit, it is likely that I will omit someone’s name, but please forgive me if that happened. First of all, my gratitude goes to my Ph.D. dissertation adviser, Prof. Lynford Goddard, for his meticulous guidance throughout my entire doctoral study here. He has always been prompt in giving me thoughtful feedback on my research. There were numerous unforgettable nights when we worked together to catch up on report and conference abstract deadlines. He has provided me a research environment with rationality and freedom. Prof. Goddard is also highly supportive of my career development by providing me opportunities to mentor undergraduates and present at major conferences. The most important thing I learned from him is being professional.

Next, I want to acknowledge my co-adviser, Prof. Gabriel Popescu, for his passion and stimulation to my research in general. There were so many evening discussions on calculations, from which I learned new ways of treating those seemingly intangible optical problems. The key point I learned from him is how to be effective and productive in research. In addition, I want to thank Dr. Chris Deeb and Dr. Timothy Crimmins from Intel for the discussions on my research through monthly teleconference meetings in the past three years. They have always been positive about my research progress. I am also indebted to my previous academic advisers. They are Prof. Joseph W. Haus, Prof. Qiwen Zhan, Prof. Peter E. Powers, Prof. Nasser Peyghambarian, and Prof. Shi Wei. The training I received from them was indispensable for my Ph.D. study at the University of Illinois at Urbana-Champaign (UIUC). I want to also acknowledge my Ph.D. committee members for devoting their time. They are Prof. Paul Scott Carney, Prof. Minh N. Do, and Prof. John A. Rogers. At the same time, I acknowledge my Ph.D. qualifying exam committee members Prof. J. Gary Eden, Prof. Shun Lien Chuang, and Prof. Kanti Jain for their advice on my Ph.D. study.
Now, I want to thank the current and previous Photonic Systems Lab (PSL) and Quantitative Light Imaging (QLI) lab group members. In particular, Chris Edwards helped me starting my experiments. I also enjoyed the fruitful discussions and conversations with him. Thanks also go to Steve McKeown for the miscellaneous help he offered. I also enjoyed the research collaborations with Taewoo Kim, through which we have made a high impact. I will always remember the conversations with Mikhail Kandel on some crazy research ideas. Let me also thank Dr. Basanta Bhaduri for the joyful conversations. And Dr. Gannavarp Rajshekhar, soon to be on the faculty at Indian Institute of Technology (IIT) Kanpur, for his kindness to me. In addition, I would like to thank my undergraduate mentees Casey Bryniarski, Marjorie Dallmann, Melissa Crawford, Austin Benjamin, Naphat Lertratanakul, and Kathy Qiu for their research contributions. Casey has always been proactive in his assigned research and helped me by proofreading my dissertation.

Let me also thank my best friend Yün-Long Han at UIUC for his personal care. I really enjoyed our discussions on everything which lead to new perspectives toward life. I want to also express my sincere gratitude to my friends from the church; special thanks go to the Ms. Winnie Hung and Mr. John Hung family in Dayton for their prayers and personal care. I should also thank my old friends (in chronological order): Lizhen Yang, Qingfeng Meng, Chengzhi Xie, Xu Zhou, Yuan Qin, Xiaodan Jin, Sicong Ma, Wei Han and his wife Wen Cheng, Zhi Wu and his wife Lei Jiang, Shenglin Ye, Yi Zhang and his wife Shan Yu, Yuan Zhao, Xida Sun and his wife, Pi-Cheng (James) Law, and Yuan Dai for their support. I cannot omit acknowledging the help from my friend Dr. Jiangrong (Frank) Cao and his wife Genhan Chen.

Finally, I would like to seriously give my thanks to Xiaojun (Tracy) Yang, my girlfriend, for her continuing support, tolerance, personal care, and love over the past 5 years. Over the past 3 years we were separated geographically, but our hearts remained united. Besides, I should also thank your parents for the support. At the end, let me give my special and great gratitude to my family members. No matter what I do
or where I go, my mother is the one who has always supported me unconditionally. I also thank my father for his vision on insisting that I study abroad and for setting high standards for me to achieve.
TABLE OF CONTENTS

LIST OF ABBREVIATIONS .............................................................................................................. xi

CHAPTER 1: INTRODUCTION ......................................................................................................... 1
  1.1 Light microscopy .................................................................................................................... 1
  1.2 Semiconductor defect inspection ....................................................................................... 3
  1.3 3D optical tomography ..................................................................................................... 5
  1.4 Dissertation statement ....................................................................................................... 7

CHAPTER 2: QUANTITATIVE PHASE IMAGING ........................................................................... 10
  2.1 Introduction ....................................................................................................................... 10
  2.2 Off-axis QPI methods ....................................................................................................... 12
  2.3 Phase shifting QPI methods ............................................................................................. 13
  2.4 Common-path QPI methods ............................................................................................. 14
  2.5 White-light QPI methods ................................................................................................. 15
  2.6 Summary ......................................................................................................................... 18

CHAPTER 3: WAFER MANUFACTURING AND METROLOGY .................................................... 19
  3.1 Semiconductor industry ................................................................................................... 19
  3.2 Integrated device manufacturing ..................................................................................... 22
  3.3 Wafer metrology ............................................................................................................. 26
  3.4 Optical metrology ........................................................................................................... 30
  3.5 Summary ......................................................................................................................... 34

CHAPTER 4: WAFER INSPECTION TOOLS ............................................................................... 35
  4.1 Laser interferometric inspection system .......................................................................... 35
  4.2 White-light interferometric inspection system ................................................................. 44
  4.3 Summary ......................................................................................................................... 60

CHAPTER 5: 22 NM NODE WAFER INSPECTION ...................................................................... 61
  5.1 Introduction ....................................................................................................................... 61
  5.2 22 nm wafer imaging ....................................................................................................... 62
  5.3 Image post-processing ..................................................................................................... 64
  5.4 Sensitivity and detectability ............................................................................................. 68
LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2DISC</td>
<td>2\textsuperscript{nd}-order image difference, image stitching, and convolution</td>
</tr>
<tr>
<td>AFM</td>
<td>atomic force microscopy</td>
</tr>
<tr>
<td>aLCI</td>
<td>angle-resolved low-coherence interferometry</td>
</tr>
<tr>
<td>CCD</td>
<td>charge-coupled device</td>
</tr>
<tr>
<td>CD-AFM</td>
<td>critical-dimension atomic force microscopy</td>
</tr>
<tr>
<td>CD-SAXS</td>
<td>critical-dimension small-angle x-ray scattering</td>
</tr>
<tr>
<td>CD-SEM</td>
<td>critical-dimension scanning electron microscopy</td>
</tr>
<tr>
<td>CMOS</td>
<td>complementary metal–oxide–semiconductor</td>
</tr>
<tr>
<td>CT</td>
<td>computed tomography</td>
</tr>
<tr>
<td>CUDA</td>
<td>compute unified device architecture</td>
</tr>
<tr>
<td>DHM</td>
<td>digital holographic microscopy</td>
</tr>
<tr>
<td>DPM</td>
<td>diffraction phase microscopy</td>
</tr>
<tr>
<td>DSA</td>
<td>directed self-assembly</td>
</tr>
<tr>
<td>DUV</td>
<td>deep ultraviolet</td>
</tr>
<tr>
<td>epi-DPM</td>
<td>epi-illumination diffraction phase microscopy</td>
</tr>
<tr>
<td>epi-wDPM</td>
<td>epi-illumination white-light diffraction phase microscopy</td>
</tr>
<tr>
<td>EUV</td>
<td>extreme ultraviolet</td>
</tr>
<tr>
<td>FBG</td>
<td>fiber Bragg grating</td>
</tr>
<tr>
<td>FFU</td>
<td>fan filter unit</td>
</tr>
<tr>
<td>FIB</td>
<td>focused ion beam</td>
</tr>
<tr>
<td>FinFET</td>
<td>fin-type field-effect transistor</td>
</tr>
<tr>
<td>FOV</td>
<td>field of view</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>FPM</td>
<td>Fourier phase microscopy</td>
</tr>
<tr>
<td>FTIR</td>
<td>Fourier transform infrared spectroscopy</td>
</tr>
<tr>
<td>GUI</td>
<td>graphical user interface</td>
</tr>
<tr>
<td>HAR</td>
<td>high aspect ratio</td>
</tr>
<tr>
<td>HPM</td>
<td>Hilbert phase microscopy</td>
</tr>
<tr>
<td>IC</td>
<td>integrated circuit</td>
</tr>
<tr>
<td>IDA</td>
<td>intentional defect array</td>
</tr>
<tr>
<td>ISAM</td>
<td>interferometric synthetic aperture microscopy</td>
</tr>
<tr>
<td>ITRS</td>
<td>international technology roadmap for semiconductors</td>
</tr>
<tr>
<td>LCI</td>
<td>low-coherence interferometry</td>
</tr>
<tr>
<td>LED</td>
<td>light-emitting diode</td>
</tr>
<tr>
<td>LER</td>
<td>line edge roughness</td>
</tr>
<tr>
<td>LWR</td>
<td>line width roughness</td>
</tr>
<tr>
<td>LIDAR</td>
<td>light detection and ranging</td>
</tr>
<tr>
<td>MuGFET</td>
<td>multiple gate field-effect transistor</td>
</tr>
<tr>
<td>NA</td>
<td>numerical aperture</td>
</tr>
<tr>
<td>NIR</td>
<td>near-infrared</td>
</tr>
<tr>
<td>NSOM</td>
<td>near-field scanning optical microscopy</td>
</tr>
<tr>
<td>OCT</td>
<td>optical coherence tomography</td>
</tr>
<tr>
<td>ODT</td>
<td>optical diffraction tomography</td>
</tr>
<tr>
<td>OQM</td>
<td>optical quadrature microscopy</td>
</tr>
<tr>
<td>PALM</td>
<td>photoactivated localization microscopy</td>
</tr>
<tr>
<td>PR</td>
<td>photo resist</td>
</tr>
<tr>
<td>PSF</td>
<td>point spread function</td>
</tr>
<tr>
<td>PSNR</td>
<td>peak signal to noise ratio</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>QPI</td>
<td>quantitative phase imaging</td>
</tr>
<tr>
<td>ROC</td>
<td>receiver operating characteristic</td>
</tr>
<tr>
<td>SEM</td>
<td>scanning electron microscopy</td>
</tr>
<tr>
<td>SLD</td>
<td>superluminescent diode</td>
</tr>
<tr>
<td>SLIM</td>
<td>spatial light interference microscopy</td>
</tr>
<tr>
<td>SLM</td>
<td>spatial light modulator</td>
</tr>
<tr>
<td>SNR</td>
<td>signal to noise ratio</td>
</tr>
<tr>
<td>STED</td>
<td>stimulated emission depletion</td>
</tr>
<tr>
<td>STORM</td>
<td>stochastic optical reconstruction microscopy</td>
</tr>
<tr>
<td>TEM</td>
<td>transmission electron microscopy</td>
</tr>
<tr>
<td>TSV</td>
<td>through-silicon via</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>WDT</td>
<td>white-light diffraction tomography</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

1.1 Light microscopy
The compound optical microscopy was invented more than 400 years ago. In the late 17th century, Dutch scientist Antonie van Leeuwenhoek (1632-1723) used his secretly made high magnification microscope to look at bacteria, spermatozoa, etc., which marked the beginning of microbiology [1]. However, the concept of optical resolution, i.e. the finest structure that can be resolved with light, was not well understood until hundreds of years later, when German physicist Ernst Abbe theoretically and experimentally established the resolution limit in 1873 [2]. Abbe pointed out that the resolution is diffraction limited to $0.5\lambda/NA$ in a microscope, which can only be explained with wave optics theory. Coincidentally, in the same year James Clerk Maxwell published his seminal paper “A Treatise on Electricity and Magnetism” [3], which became the now famous Maxwell’s equations and the foundation of wave optics. (Actually Maxwell had already published a series of papers on this topic many years earlier. The treatise paper summarized them with 11 equations. Around 1884, Oliver Heaviside and others simplified the 11 equations into four which became the standard [3].)

In addition to the diffraction limit, signal contrast is another important factor that limits the resolution in imaging. In the 1930s, Frits Zernike invented the phase contrast microscope, enabling us to look at transparent specimens [4, 5], which would otherwise not be possible to view under a bright-field microscope. In 1948, Dennis Gabor proposed a lensless imaging method that reconstructs the wavefront of an object, known today as holography [6]. Realizing its similarity to the synthetic aperture radar problem, holography gain popularity twenty years later [7] as lasers emerged as coherence light sources. In the 1990s,
advances in CCD devices and computers enabled holograms to be recorded on a CCD camera instead of the photorefractive materials, and to be reconstructed digitally with reasonable quality [8-12].

Since the 2000s, image sensor technologies have made significant advancements. Nowadays, high-quality CCD and CMOS (complementary metal–oxide–semiconductor) cameras are available at reasonable prices. An imaging field called quantitative phase imaging (QPI) was born and it has steadily gained popularity [13]. Due to its noninvasive nature, QPI is now an up-and-coming modality for biomedical imaging [9, 13-15]. QPI reconstructs the wavefront of the electromagnetic field which contains the sample topography and index information, providing a stable and universally applicable source of intrinsic contrast. Most recently, efforts have been made to make portable QPI systems by integrating it with cell phones or smart phones [16, 17].

Most of the QPI techniques are based on the interference of a reference field with a signal field. Two interference configurations are typically used for QPI: off-axis geometry and on-axis geometry. Off-axis geometry based QPI, which typically uses a Mach-Zehnder interferometer geometry with two separate beam paths, is a single-shot phase reconstruction method. This method is generally called digital holographic microscopy (DHM) [9, 10]. In 2006, Gabriel Popescu and his colleagues at the GR Harrison Spectroscopy Laboratory at MIT developed a common-path geometry DHM, called diffraction phase microscopy (DPM) to study red blood cell dynamics [18]. Due to the common-path geometry in DPM, the vibrational noise between the two inference beams can be significantly reduced to achieve nanometer depth sensitivity [18-21]. Most recently, an epi-illumination (reflection-based) diffraction phase microscopy (epi-DPM) was developed and used for material science applications [22-24]. Since 2011, we have adapted the epi-DPM system for semiconductor wafer defect inspection for 22 nm and 9 nm node wafer inspection [25-29].
The second QPI technique is based on on-axis interferometry, which realizes phase reconstruction by using four phase-shifted images [11, 12]. In 2011, the Quantitative Light Imaging Laboratory led by Popescu at the University of Illinois invented an on-axis interferometry microscope, called the spatial light interference microscopy (SLIM) [15]. SLIM combines the concepts of phase-shifting and phase-contrast, and it has been demonstrated to be a highly sensitive microscope for biomedical imaging [15, 30-32]. SLIM uses a white-light illumination, and it has a strong sectioning effect, making it suitable for three-dimensional (3D) tomographic reconstruction. Previously, Mir et al. experimentally measured the 3D point-spread function (PSF) of the SLIM system, and then used it to reconstruct the 3D structure of cells [33]. Building upon these previous experimental results, we developed an inverse scattering model to physically describe the system [34, 35]. The model is built by extending Emil Wolf’s diffraction tomography to white light [36]. Using the obtained inverse scattering solutions, we were able to reconstruct 3D cell structures with resolution beyond the diffraction limit both transversely and axially. This method can be used to image cells in 3D for very long time periods (e.g. days) [37].

1.2 Semiconductor defect inspection

Detection of inline killer defects in patterned wafers is a grand challenge to the semiconductor industry. An isolated defect buried in a nano-pattern can substantially affect the functionality of the whole integrated circuit [38-40]. Thus, detecting defects during device fabrication is critical for maintaining a high yield [41-43]. Scanning electron microscopy (SEM) and atomic force microscopy (AFM) based metrology methods have been developed for wafer defect inspection [44, 45]. Despite advances made in recent years, SEM is still low throughput and could be destructive [46-48] and AFM still has extremely low throughput [49-52]. These characteristics make them unsuitable for inline defect inspection. Optical inspection methods are usually non-destructive. Optical scatterometry and dark-field based defect inspection tools have been developed in the past. These methods typically use focused light and scan the sample, thus limiting the
throughput. Optical microscopy systems have high throughput due to their relatively large field of view (FOV). But, the resolution of an optical system is diffraction limited to approximately $0.5\lambda/NA$ [53]. To detect deep sub-wavelength defects, one approach has been focused on improving the resolution, such as developing ultra-violet and high-order harmonic laser sources [54-58], or using very high numerical aperture lenses [59-63]. These methods typically have a small FOV because of the available cameras at these wavelengths, and can still potentially damage a wafer during fabrication due to the short wavelength, high pulse energy [64-67], and physical contact with the wafer.

However, the ultimate limiting factor for defect detection is noise [16, 68-70] which decreases image contrast and destroys sensitivity. If we have a very sensitive microscope system, detection of a single fluorescence molecule or nanoparticle is possible [69, 71]. Fluorescence microscopy is a highly sensitive method, which laid the foundation for many superresolution methods [72-75]. Unfortunately, there is no direct access to fluorescence for semiconductor materials. Furthermore, point-scanning methods, which have low throughput, are impractical for inspecting large-area wafers. To overcome these issues, we built a highly sensitive inspection system using an epi-DPM with laser illumination. This system uses a common-path interferometer and image post-processing techniques targeted to mitigate the effects of characteristic noise sources. Using this system, we have successfully detected defects with sizes down to 20 nm wide by 100 nm long by 110 nm tall in a 22 nm node patterned wafer.

After inspection of the 22 nm node wafer, we moved to detecting defects in a 9 nm node wafer, a significantly more challenging problem. This wafer has defect sizes that are only 10 nm wide and the pitches are 2x smaller compared with the 22 nm node wafer, making the defect detection really difficult. Thus, the system’s sensitivity needed to be enhanced tremendously to successfully detect the defects. We first replaced our 532 nm solid-state laser with a 405 nm diode laser that has 10x better power stability and inserted a 405 nm narrow-band filter in front of the camera. With the 2DISC method, we have detected
parallel bridge defects in both the phase and amplitude images. However, the optimum focus position becomes very critical for the 9 nm node wafer. Thus, performing a scan in the z direction is expected to also improve the sensitivity. We built a 3D stage with 8 nm scanning step resolution and synchronized it with the CCD camera image acquisition in NI LabVIEW. With this integration, we can now perform 3D sample scanning as well as real-time data collection and processing using a LabVIEW GUI.

To further improve the sensitivity, we proposed building an epi-illumination white-light DPM system (epi-wDPM) with Fourier space selective filtering which we call dark-field imaging. White-light imaging systems are speckle free, thus have higher sensitivity compared with lasers. Thus, the epi-wDPM system is expected to further enhance the defect signal contrast. In epi-wDPM, the Fourier space filtering is achieved by using a spatial light modulator (SLM). The SLM has two purposes, and the first one is to make a reference beam. The other purpose is to selectively filter the signal beam coming from the wafer’s underlying structure, such that the CCD camera’s full dynamic range can be used for the wafer signal. For the preliminary demonstration, we used a 50 micron diameter opaque aperture to act as the dark-filter. When measuring a parallel bridge defect, 50% sensitivity improvement was achieved as compared to bright field. After that, we built the epi-wDPM system with the SLM filter and performed the system validation. The system we built has a spatial coherence of about the diffraction spot, which is sufficient for wafer structures. We are currently using this system for inspecting the 9 nm wafer. There are issues with the epi-wDPM system for wafer defect inspection, and we are working to solve them.

1.3 3D optical tomography

A transparent object illuminated by an electromagnetic field generates a scattering pattern that carries specific information about its internal structure. Inferring this information from measurements of the scattered fields, i.e., solving the inverse scattering problem is called optical tomography. It uses the transmitted and diffracted light to reconstruct the 3D structure of the sample. The development of optical
tomography dates back to the early works on x-ray diffraction. When x-ray diffraction emerged in the 1940s, it was initially used for analyzing crystal structures [76], and then became widely adopted for intracellular structures studies [77]. In the 1960s, x-ray computed tomography (CT), a technique based on the absorption along the projection path, was developed [78]. However, it is well known that x-ray CT suffers from the phase problem [79]. James Fienup proposed an iterative phase retrieval method using the modulus of the Fourier transform (F.T); however, the uniqueness of the solution cannot be guaranteed [80, 81].

Inspired by x-ray CT, in 1969 Wolf published the seminal paper on optical diffraction tomography (ODT) [36]. Recently, ODT has been successfully used for 3D reconstruction of transparent objects, and the experiments are typically done by scanning the incident beam or rotating the object at different angles and measuring the scattered field in the far field [82-85]. ODT usually uses a laser as the light source; thus the images suffer from the well-known speckle noise [86] which degrades the image contrast and resolution. Recently, Christian Depeursinge’s Microvision and Micro-Diagnostics Group at the École Polytechnique Fédérale de Lausanne (EPFL) in Switzerland demonstrated a synthetic aperture method to remove laser speckle for achieving superresolution [87].

Another tomographic reconstruction method is called optical coherence tomography (OCT), which was successfully demonstrated for biomedical imaging in 1991 by James Fujimoto’s Laser Medicine and Medical Imaging Group at MIT, and soon it became widely adopted for biomedical imaging [88]. OCT uses the coherence gating effect, and the axial resolution is determined by the coherent length of the light source. OCT typically uses a very low numerical aperture (NA) objective to measure the cross-correlation between the reference and the signal beam. The object’s axial dimension is resolved by scanning the reference mirror over a broad range. The imaging depth of OCT is limited by the scattering which attempts to match the refractive indices of tissue components [89]. OCT typically has poor lateral resolution due to
the use of very low NA objectives. In 2006, the Biophotonics Imaging Laboratory led by Stephen Boppart at the University of Illinois in collaboration with P. Scott Carney developed the interferometric synthetic aperture microscopy (ISAM) [90] to improve the lateral resolution of OCT. They used the diffraction tomography theory instead of the widely used low-coherence interferometry theory.

Recently, we proposed a new tomographic reconstruction method using the concepts from OCT and ODT which we call white-light diffraction tomography (WDT). WDT is an imaging method with a relatively simple setup, without the need of measuring the sample in the far zone at different angles. It uses white light as the light source, thus it suffers no speckle and has high image contrast. Like OCT, the axial resolution of WDT is limited by the coherence properties of the light source. The distinguishing property to OCT is that WDT scans the object through focus and uses high numerical aperture objectives. WDT is realized with the SLIM system. We developed WDT reconstruction theory by extending Wolf’s diffraction tomography to white light. The theory we developed for WDT can be also used for solving inverse scattering in imaging systems under the Fresnel approximation.

1.4 Dissertation statement
This dissertation consists of two main topics. The first topic is on semiconductor wafer defect inspection which comprises the majority of the dissertation work. The second topic is focused on developing optical tomographic methods and the related applications. The connection is that both topics utilize quantitative phase imaging and scattering theory to address societally important problems.

The wafer inspection work is covered from Chapter 3 to Chapter 7 and the related Appendix sections. In Chapter 3, we provide an overview of the semiconductor field inducing the semiconductor industry, wafer fabrication process, and wafer defect metrology. This chapter serves as the motivation and background for the wafer inspection work. In Chapter 4, we describe the wafer inspection tools that we have developed, including their system designs and the working principles. In Chapter 5, the work on 22
nm node wafer defect inspection is discussed exclusively. In this chapter, we demonstrate detection of deep sub-wavelength defects using an image post-processing method called 2DISC. In Chapter 6, we discuss 9 nm node wafer inspection. This wafer was first inspected using our system with the 2DISC method, and one type of defect was detected. To detect more types of defects and smaller defects, a more sensitive technique is needed, motivating us to build a white-light interferometric system with dark-field imaging modality. Chapter 7 is on the outlook for wafer inspection, where we summarize the wafer inspection work and propose new methods to be pursued in the future. Supplementary material on wafer defect inspection is provided in Appendix A.

The work on optical tomography is covered from Chapter 8 to Chapter 11. In Chapter 8, we review inverse scattering, laser diffraction tomography, and low coherence interferometry (LCI). In Chapter 9, a new optical tomography method called white-light diffraction tomography (WDT) is introduced. This chapter includes a full mathematical description of the theory, the calculation of the 3D point spread function (PSF), and a 3D reconstruction example of red blood cells. WDT theory uses the wavevector space method. In Chapter 10, we apply WDT theory to solve the inverse scattering problem using low-coherence light. We also discuss the 3D reconstruction for imaging systems under the Fresnel approximation. Both the forward and the backward 3D reconstruction models are calculated in the frequency-domain and the time-domain. In Chapter 11 and Appendix B, several wavevector space method based applications are proposed, including refocusing in QPI, the optical trapping force calculation, the lens equation calculation, and the diffraction integral calculation. Finally, Chapter 12 summarizes the key conclusions of the dissertation.

It is worth mentioning that Appendix C focuses on CO$_2$ sensing using a 2 µm Tm-Ho co-doped fiber laser and has involved collaboration with the Department of Agriculture and Biological Engineering at
UIUC for developing a CO₂ sensor for monitoring corn storage. In addition, this project has brought several undergraduate students to work with us in our labs to obtain hands-on experience for scientific research.
CHAPTER 2: QUANTITATIVE PHASE IMAGING

The main focus of the dissertation is on using interferometric light microscopy or QPI to solve two practical and challenging problems: detecting 10 nm defects in densely patterned silicon wafers (metrology) and reconstructing 3D biological structures in biomedicine (biology). Quantitative phase imaging (QPI) is an optical microscopy method that is usually based on two-beam interferometry. Thus, it is beneficial to discuss QPI before entering the main themes of the dissertation. We recently published a review on the topic metrology meets biology in QPI [91], discussing applications that bridge biology and metrology.

For the best performance in each application, a specialized interferometric light microscopy system is needed. In wafer inspection, the structures are reflective, and high throughout and high sensitivity are required. A common path digital holography microscopy, called epi-illumination diffraction phase microscopy (epi-DPM), was developed to fulfill the task. Both a laser and a white-light illumination system were developed in order to find the best solution. In 3D biological imaging, the samples are typically transparent and miniature. A white-light interferometric system that uses phase-contrast and phase-shifting techniques, called spatial light interference light microscopy (SLIM), was used. One may argue that SLIM can be also used for wafer defect inspection and vice versa. Yes, it is true; the systems are interchangeable and the solution is certainly not unique. No matter what methods are used, solving the problem is the key. Moreover, knowing the advantages as well as the limitations of different interferometric imaging methods is also an important mission. In the rest of this chapter, we review QPI.

2.1 Introduction

Since the 1990s, image sensor technologies have made significant advancements, giving birth to a new field called quantitative phase imaging (QPI) [13, 92]. QPI, usually based on interferometric measurements,
provides quantitative amplitude and phase information associated with the electromagnetic field. Therefore, through QPI, one can reconstruct the complex field at the detector. Recent developments in QPI have brought high sensitivity to the technique, with the phase sensitivity down to 1 mrad, corresponding to 1 nm thickness changes, or conversely $10^{-4}$ refractive index changes in transparent biological structures, assuming a refractive index contrast of 0.03 [13, 21]. Importantly, QPI is a label-free method, without any fluorescence markers or dyes, and keeps the specimen intact. Therefore, it is ideal for biomedical imaging, where the live cells are sensitive to subtle perturbations in the environment [8-12, 24]. Recently, various applications of QPI have been applied for non-invasive study of cell dynamics [19, 20, 32, 93-95], blood testing [19, 21, 96-105], cell growth [105-107], and 3D cell imaging [14, 33, 34, 84, 87, 106]. Due to the non-destructive and high throughput nature, QPI has also been applied to material characterizations [22, 27, 108].

Most QPI techniques are based on two-beam interferometry, where a reference beam and a signal beam form an interference signal. In the general case, the measured intensity on a camera detector is

$$I(x, y; \tau) = I_{\text{sig}}(x, y; \tau) + I_{\text{ref}}(x, y; \tau) + 2|\Gamma_{12}(x, y; \tau)| \cos \left( \beta x + \phi_{\text{ref}}(\tau) - \phi_{\text{sig}}(x, y) \right),$$

(2.1)

where $I_{\text{sig}}(x, y; \tau)$ and $I_{\text{ref}}(x, y; \tau)$ are the intensity of the signal and the reference beam, respectively, $\Gamma_{12}(x, y; \tau)$ is the cross-correlation between the reference and signal field, $\beta$ is the spatial frequency shift, or carrier frequency, between the two beams, $\phi_{\text{ref}}(\tau) = \omega \tau$ is the time-varying phase of the reference beam, and $\phi_{\text{sig}}(x, y)$ is the phase of the signal beam. There are two different interferometric techniques typically used to retrieve the signal field phase, $\phi_{\text{sig}}(x, y)$: off-axis interferometry and phase-shifting interferometry.
2.2 Off-axis QPI methods

In off-axis interferometry, the angle between the two beams is non-zero, which creates a spatial frequency shift $\beta$. Also, $\phi_{ref}(\tau)$ is fixed as a constant. Therefore, a fringe pattern is created with a period determined by the carrier frequency, $\beta$:

$$I(x, y; \tau) = I_{sig}(x, y; \tau) + I_{ref}(x, y; \tau) + 2|\Gamma(12)(x, y; \tau)|\cos(\beta x + \phi_{ref} - \phi_{sig}(x, y)),$$

(2.2)

Since the propagation of the two interference beams can be precisely described by diffraction theory, it allows for numerical reconstruction of the complex image field. In experiments, this type of interferometer is often realized in a Michelson or Mach-Zehnder interferometer, followed by the Fourier or Hilbert transform technique on the acquired interferogram to recover the quantitative phase information [8-10, 24, 109].

In 1967, Goodman and Lawrence first demonstrated numerical reconstruction of the complex signal field from a hologram, which has been widely known as digital holography [110]. Later on, Takeda et al. proposed a fast Fourier transform method to analyze the fringe pattern to reconstruct the sample topography [111]. As large CCD detector arrays became available in 1990s, Schnars et al. demonstrated the first lensless off-axis digital holography system with much improved image quality [8]. In the following years, digital holography was adapted in a microscopy system by several groups, which became a new microscopic technique called digital holographic microscopy (DHM) [10, 112]. The early DHM systems were based on a scattering measurement, which required a deconvolution of the Fresnel propagation kernel to retrieve the image complex field. However, spatial sampling and phase discontinuity problems existed in these systems. For live cell imaging, the optimum detection plane is the image plane. Thus, in 2005 Popescu et al. proposed an off-axis QPI method called Hilbert phase microscopy (HPM), in which the CCD camera is placed exactly at the image plane. HPM uses the complex analytic signal, calculated from
the measured intensity (real-valued), to retrieve the phase of a transparent object. Therefore, HPM is a single shot technique, which is suitable for studying cell dynamics and morphology [109, 113, 114].

2.3 Phase shifting QPI methods

In phase-shifting interferometry, the angle between the reference and the scattered beam is zero, while the reference field phase \( \phi_{\text{ref}}(\tau) \) is modulated several times, in order to obtain quantitative phase images. Therefore, Eq. (2.1) becomes,

\[
I(x, y; \tau) = I_{\text{sig}}(x, y; \tau) + I_{\text{ref}}(x, y; \tau) + 2 \left[ \Gamma_{12}(x, y; \tau) \right] \cos \left( \phi_{\text{ref}}(\tau) - \phi_{\text{sig}}(x, y) \right).
\]

The reference phase \( \phi_{\text{ref}}(\tau) \) can be shifted through either changing the delay \( \tau \) or the frequency \( \omega \) of the illumination light. By combining four interferograms at \( \phi_{\text{ref}}(\tau) = 0, \pi/2, \pi, 3\pi/2 \), the phase of the sample field is determined uniquely as,

\[
\phi_{\text{sig}}(x, y) = \tan^{-1} \left[ \frac{I(x, y; 3\pi/2) - I(x, y; \pi/2)}{I(x, y; 0) - I(x, y; \pi)} \right].
\]

In 1993, the group led by Graham Dunn at King’s College, London, invented a technique called digitally recorded interference microscopy with automatic phase-shifting (DRIMAPS) based on a Mach-Zehnder interferometer setting [115-117]. In this technique, the phase shift is added to the reference by horizontally sliding an optical wedge. Another experimental setup, called optical quadrature microscopy (OQM), was developed in 1998 at Northeastern University. OQM uses phase shifts generated by polarization waveplates. It was initially used to determine the sign of the Doppler velocity [118], and later on applied to cell counting in embryos [119, 120]. To improve the phase imaging speed, several groups have recently developed QPI methods that can simultaneously generate four phase shift images. This type of system has enabled fast dynamics imaging [121, 122].
2.4 Common-path QPI methods

In 2004, Popescu et al. developed a common-path phase-shifting QPI technique called Fourier phase microscopy (FPM). FPM combines the concepts of phase-shifting and Zernike’s phase-contrast in one setup to quantitatively measure the phase of transparent biological structures. In the first FPM demonstration, a super-luminescent diode (SLD) was used as the illumination. To ensure full spatial coherence on the CCD camera, the SLD is spatially filtered through a single mode fiber. A Fourier lens is used to separate the reference (unscattered) and signal (scattered) fields. The reference field is focused onto a spatial light modulator (SLM) where four phase shifts are created sequentially. The phase contrast images are captured at the image plane and used to reconstruct the signal field phase image using Eq. (2.4). It should be pointed out that the phase associated with the object is related to the image field. Thus, it is different than the signal field, due to the absorption through the phase ring inside the objective lens. And the image phase is calculated as,

\[
\phi_{\text{sig}}(x, y) = \tan^{-1} \left\{ \frac{\beta(x, y) \sin \left[ \phi_{\text{sig}}(x, y) \right]}{1 + \beta(x, y) \sin \left[ \phi_{\text{sig}}(x, y) \right]} \right\},
\]

(2.5)

where \( \beta(x, y) = \left| \frac{E_{\text{sig}}(x, y)}{E_{\text{ref}}(x, y)} \right| \). FPM has high temporal stability, without the need of active stabilization. FPM has been used for measuring nanoscale membrane fluctuations and cell growth [104, 106].

In 2006, Popescu et al. developed a new type of off-axis transmission QPI method with a common-path geometry, called diffraction phase microscopy (DPM) [18]. The first DPM was built using laser illumination, which was filtered through a single mode fiber to ensure full spatial coherence [123]. Using the laser as a light source, an image is formed through a commercial microscope at the image plane, where the DPM module is attached. The DPM module consists of a 4f lens system that grants access to the Fourier plane of the field. A diffraction grating is placed at the microscope image plane to generate multiple
diffraction orders. These diffraction orders are spatially separated in the Fourier plane, where a physical pinhole filter or an SLM is located. This filter low-passes the 0th-order diffracted beam to generate a reference beam and passes the 1st-order diffracted beam without filtering to carry the signal from the sample. These two beams are then combined at the image plane through the second Fourier lens to form an interferogram on the detector. Since these two beams are traveling in a common-path geometry, any noise induced by the vibration of the optical elements is minimized, enabling path length sensitivity below 1 nm. Furthermore, DPM is a single-shot technique where the acquisition speed is limited only by the acquisition rate of the camera. Therefore, this system is suitable for studying nanometer scale fluctuations in a highly dynamic sample. As an example, red blood cell membrane dynamics and other mechanical properties have been studied using DPM [19-21]. More recently, DPM has been extended to image samples in a reflection geometry [22, 27, 108]. Over the past several years, many other common-path QPI methods have also been proposed. For example, V. Mico et al. have demonstrated common-path QPI by spatially multiplexing the sample [124, 125]. QPI has also been realized by simply scanning the object through the objective focus and reconstructing the phase image using the transport-of-intensity equation (TIE) technique, without involving interferometry [126, 127].

2.5 White-light QPI methods
2.5.1 wDPM and wHPM
Laser based QPI techniques suffer from laser speckle, due to the high coherence of the laser, which degrades the image quality. To mitigate this problem, several white-light QPI techniques have been developed in Popescu’s lab recently. In 2012, Bhaduri et al. developed a white-light version of DPM (wDPM) to study cell morphology and dynamics [128]. In wDPM, an SLM is used as the spatial Fourier filter to ensure full coherence of the reference field at the CCD camera. Spatial and temporal sensitivity have been significantly improved in wDPM compared with laser DPM, which has enabled it to better study of cell morphology...
and dynamics [128]. Following up this white-light method, Edwards et al. studied the spatial coherence properties in wDPM for optimization of the method [129]. In 2013, a white-light version of FPM (wFPM) was also developed by Bhaduri et al. and used to study cell membrane dynamics [130].

2.5.2 Spatial light interference microscopy (SLIM)

![Spatial light interference microscopy setup](image)

In 2011, a white-light phase-shifting QPI method, called spatial light interference microscopy (SLIM), was demonstrated [30]. Similar to FPM, SLIM is also based on a commercial phase-contrast microscope with a white-light illumination generated by a halogen lamp and filtered by a ring-shaped annulus. In SLIM, as
illustrated in Fig. 2.1, the back focal plane of the phase contrast objective lens is relayed to the Fourier plane of the 4f system, located at the output port of a commercial microscope. At this plane, an SLM adds four different phase delays, 0, π/2, π and 3π/2, to the reference field, generating four different interferograms at the detector plane. Combining these four interferograms, the phase delay through the objective can be reconstructed according to Eqs. (2.4) and (2.5).

Since SLIM is based on a commercial microscope, it can utilize all the microscope peripherals such as atmosphere control, temperature control, or high-precision stages. Furthermore, due to the low-coherence illumination and the common-path geometry, SLIM ensures high temporal and spatial sensitivity, 0.03 nm and 0.3 nm in optical path length, respectively. Therefore, with this high sensitivity and stability, SLIM is very suitable for imaging unlabeled live cells for a long period of time, since the growth happens over days and the changes in live cells are on the scale of one femtogram [131, 132]. Needless to say, it can also be coupled with other microscopy modalities, such as fluorescence microscopy or differential interference contrast microscopy, in order to obtain more specific information about the sample. Recently, a SLIM system has been improved in its speed to acquire 12 frames per second (fps), making it suitable for fast dynamics studies [32].

Another advantage of SLIM is the inherent depth-sectioning effect due to the low-coherence illumination. Similar to OCT, the coherence gating effect makes it suitable for 3D tomographic reconstruction of the sample refractive index distribution. Previously, Wang et al. have shown that SLIM is capable of imaging live cells in 3D with an axial resolution of 1.34 μm [133]. Following studies by Mir et al., also have shown that, with a sparsity constraint based deconvolution, SLIM can reconstruct the helical subcellular structures in *E. coli* cells [33]. Building upon these experimental results, Kim et al. have developed a physical model to describe the 3D imaging principle of the SLIM system, by extending Wolf’s
diffraction tomography to a broadband source, and further improved the tomographic imaging using SLIM [34, 36]. This result is described in Chapter 9 of this dissertation.

2.6 Summary

Here we summarize the attributes (acquisition speed, field of view, temporal sensitivity, and spatial sensitivity) of different QPI techniques. Off-axis techniques are fast due to their single shot nature. Phase-shifting techniques maintain a wide field of view, since they do not require additional magnification. Off-axis is better than phase-shifting for acquisition speed, but phase-shifting is better for field of view. For both configurations, the phase is independent of the intensity variations across the pupil. Compared with non-common path, common-path techniques have higher temporal sensitivity due to their insensitivity to any time-varying noise such as vibrational noise. Unlike laser techniques, white-light techniques are speckle free and offer excellent spatial sensitivity due to their low temporal coherence. QPI techniques can be categorized by these four attributes. Hybrid QPI techniques have been developed to include as many of these beneficial attributes as possible. For example, wDPM is an off-axis, common-path, white-light technique, and SLIM is a phase-shifting, common-path, and white-light technique. It is interesting to see whether all four attributes can actually be possessed by a single QPI method. One possibility would result from the simultaneous generation and capture of four phase-shifting images in SLIM. Actually, this is possible and has already been demonstrated previously [121, 122].

Overall, quantitative phase imaging enables one to measure the full complex field from light scattering by measuring the amplitude and phase of the field. Recent advances in light sources, CCD and CMOS detectors, and computers have brought new QPI techniques to provide highly sensitive phase measurement in both 2D and 3D, with a diffraction limited resolution. All these developments have motivated this dissertation research on wafer defect inspection and 3D object reconstruction by developing highly sensitive and high-speed QPI techniques.
CHAPTER 3: WAFER MANUFACTURING AND METROLOGY

Before introducing wafer defect inspection and metrology, it is instructive to have an overview of the semiconductor industry. In this way, we know what role metrology is playing in the semiconductor field. This is also the motivation of the dissertation research of defect inspection. This chapter is organized in the following way. We first survey the semiconductor industry and especially discuss the wafer inspection and metrology field. This is followed by a brief introduction of the semiconductor manufacturing process where metrology is involved ubiquitously. After that, we exclusively review the wafer metrology field, including the current methods, future trends, and challenges. At the end, we focus on optical metrology and discuss different optical metrology methods, which naturally leads to the dissertation focus on developing optical defect metrology tools using interferometric light microscopy.

3.1 Semiconductor industry

The semiconductor industry started in the late 1950s in the San Francisco Bay area, later nicknamed as the Silicon Valley. It was incubated with the founding of the Fairchild Semiconductor Company by the traitorous eight, who had left the Shockley Semiconductor Laboratory in 1957. In the past over 57 years, the semiconductor field has grown enormously to become a very competitive high-tech industry. Today, there are hundreds of thousands of companies related to the semiconductor industry worldwide. A chart showing the global market growth in the semiconductor industry from 1992 to 2017 is shown in Fig. 3.1, where the 2013-2017 values are forecasted [134]. Over 20 years from 1992 to 2012, the total market has grown fivefold from $64 billion to $313 billion. By 2015, the total market is estimated to be around $400 billion, despite the recent economic recession. In 2013, the leading semiconductor companies (mixed by integrated device and fabless manufactures) by sales were: Intel Corp, Samsung Electronics, Qualcomm,
Micron Technology, SK Hynix, Toshiba Semiconductor, Texas Instruments, Broadcom, and STMicroelectronics [135]. The competition among those top semiconductor companies has intensified in recent years, due to the thriving mobile device industry.

Fig. 3.1. Worldwide semiconductor market history and forecast.

Fig. 3.2. A pyramid representation of the semiconductor industry market.

Semiconductor device manufacturing is highly interdisciplinary, involving the development of a series of precision equipment devices such as lithography systems, chemical vapor deposition systems, metrology
systems, microscopy systems, spectroscopic systems, etc. Figure 3.2 is a pyramid chart showing the market share in semiconductor devices, semiconductor equipment, wafer inspection, and metrology, from the bottom to top for the year 2004 [136]. The total semiconductor device market is $175 billion with semiconductor equipment taking about 17% of the market with a total capital of $30 billion. Moreover, in the equipment market, wafer inspection represent 10% of the equipment market with a capital of $3 billion, which is a relatively large market.

![Pie Chart of Semiconductor Manufacturing Equipment Market Shares in 2013](image)

**Fig. 3.3.** Semiconductor manufacturing equipment market shares in 2013.

Today, the major semiconductor equipment supplier companies are ASML, Applied Materials, Tokyo Electron, and KLA Tencor. The market share in 2013 of the leading equipment companies is shown in a pie chart in Fig. 3.3 [137]. In 2013, the total market in equipment is about $35 billion, up by $5 billion from 2004. Among the companies listed in the pie chart, KLA Tencor, Applied Materials, and Hitachi High-
Tech are the leaders in wafer metrology. We also show a pie chart exclusively for the metrology market for the same year in Fig. 3.4 [138]. In this chart, KLA Tencor takes about half of the whole market, estimated at $5 billion, which is a 50% increased from 2004. KLA Tencor specializes in developing optical and E-beam defect metrology tools. Thus, from the market point of view, optical metrology is very important and still has a lot of room for research and development.

![Pie chart showing market shares in 2013.](image)

Fig. 3.4. Metrology equipment market shares in 2013.

### 3.2 Integrated device manufacturing

So far, we have looked over the market of wafer metrology. It will be also helpful to review the manufacturing process of integrated devices and identify where metrology is physically presented and what type of role it is playing. The references are mostly coming from the class notes of ECE444 at the UIUC, Intel (from sand to sand [139]), Applied Materials (Silicon Wafer Processing [140]), and the metrology and inspection chapter in the Handbook of Thin Film Deposition Processes and Techniques (Chapter 6 in [141]).
Over the past 40 years, the feature size of transistors has been continuously shrinking from over 10 µm to 22 nm, following Moore’s law. Transistors are the major building blocks of integrated circuits (ICs). In the following, we review the fabrication process of IC devices. The IC device fabrication takes place on a Si wafer. The production of Si wafers starts from quartz, which mostly consists of silicon (Si) in the form silicon dioxide (SiO₂). Si is the second most frequent element in the earth crust, just after oxygen. To make semiconductor devices, Si with an impurity level less than one part per billion is needed. The purification of Si starts from the chemical reaction: SiO₂ + 2 C → Si + 2 CO, in an electric arc furnace at over 1900 °C. After the extraction, a chemical reaction takes place that changes Si to the form of trichlorosilane (SiHCl₃) for a later purification, called the fraction distillation process. This is followed by the Siemens process (or chemical vapor deposition process) with the chemical reaction: SiHCl₃ + H₂ → Si + 3HCl, which yields electronic grade Si. At the end, bulk single crystal Si is grown, as a cylindrical ingot, in a heated crucible, through the Czochralski process. An image of the Si ingot is shown in Fig. 3.5 at the top-right corner.

Single Si wafers are produced by slicing the ingot with a saw followed by lapping and polishing. In addition, an epitaxy single crystal layer of Si, acting as a barrier layer for the subsequent device processing, is also formed on the wafer single crystal substrate with the CVD process. The diameter of silicon wafers used in most of today’s fabs is 300 mm or 12 inches. When Intel made the first IC devices, the wafers were 2 inches. As the wafer sizes increases, the production cost per chip decreases quadratically. In the future, 450 mm wafers will be used by Intel, with the development enabled by ASML. At the unpatterned wafer level, wafer inspection has already begun, with methods such as measuring the wafer surface roughness, inspecting surface particles, micro-scratches, and crystalline defects. Electronic tests are also performed at this level such as measuring the resistivity with the 4-point probe and conductive flaws with eddy-current measurements.
The IC device processing starts with an epitaxy wafer. Figure 3.5 shows the major steps involved. In the following, we will identify the metrology steps concurrent with each processing step. The first step is oxidation which is used to create a thick SiO$_2$ layer to prevent unwanted parasitic junctions in the devices. The oxidation of Si is achieved by oxygen diffusion, governed by Fick’s law of diffusion, in a furnace at about 1100 °C.

Oxidation is the starting point of photolithography. Initially, a thin coating of photoresist (PR) is applied on the Si wafer surface on a spinner to ensure uniform PR coverage. Afterward, the wafer is exposed to the ultraviolet (UV) light patterned by a mask. The mask is aligned with a precise mask aligner. Any further photolithography requires the masks to be overlaid precisely with the previous mask; thus, optical overlay metrology tools are required.

After the UV exposure, the exposed PR becomes soluble (or the unexposed region become soluble, depending on the tone of the PR). Thus, the wafer is developed in a solvent to dissolve the patterned regions. Lithography can be also achieved with an e-beam instrument which is also based on changing the solubility
of the resist. At this stage, many metrology systems are involved such as SEM, optical overlay tools, and defect inspection tools to ensure the lithography pattern satisfies the design rule.

Following the resist development, the wafer undergoes etching, in order to remove the SiO$_2$ in the patterned region. There are many etching methods such as wet etching (usually etches isotropically) with HF or dry etching (usually etches anisotropically) with an overall neutral ionized plasma gas. Isotropic wet etching can also be achieved with metal-assisted chemical etching [142]. In the etching process, the metrology tools typically used are a reflectometer to monitor the removal rate and etch selectivity, SEM and AFM to measure the etch profile such as the line edge roughness (LER) and line width roughness (LWR), and defect inspection systems to measure the pattern defects and particulate contamination. After etching, impurities are doped into Si to alter its conductance to form junctions, through a process called ion implantation, using either high kinetic energy or diffusion with heat (a relatively cheap and simple method).

At the end of the processing, plasma ashing is used to remove the PR. The fabrication of transistors is almost done. To make faster and smaller chips, multiple layers of transistors are actually fabricated, which can be realized by repeating all the above processing steps. Additional processing steps are electroplating of copper (deposit of a thin layer of copper ions on the transistor surface) and polishing the excessive copper. Multiple copper layers are also created as wires to interconnect the transistors at different layers. The wires are formed in an architecture that is determined by the function of the device. Now the on wafer IC processing is completed. The following procedures are slicing the wafer into individual dies and packaging. During these processes, many tests are performed mostly on functionality and reliability. Obviously, these processes involve using many metrology tools. The correlation results of metrology to yield is managed by the fab wide data management system, which is the “brain” of the metrology tools used during the whole manufacturing process.
3.3 Wafer metrology

Metrology is the science of measurement. Over the past years, semiconductor device feature size has shrunk to the nanoscale, thus, extremely high-precision measurement tools need to be developed. Due to the highly serialized nature of semiconductor manufacturing, metrology is becoming more and more important to ensure that wafers have not been damaged by previous processing steps. If defects are presented in a die, it can actually fail the whole circuit. Thus, it is essential to develop inline metrology tools to maximize the yield during the manufacturing.

3.3.1 Metrology challenges

Detection of in-line killer defects in patterned wafers has been a grand challenge, especially as the semiconductor industry moves to beyond the 11 nm node, according to the 2013 International Technology Roadmap for Semiconductors (ITRS) report on metrology and a previous discussion from Intel Corp [143]. A detailed discussion on metrology is presented in the metrology section of the 2007 ITRS report and in Chapter 6 of reference [141]. The challenges come from the following aspects. Firstly, the node process dimension continues to shrink below the 12 nm (half pitch size), which therefore requires detecting defects smaller than 1 nm, which is significantly beyond the resolution limit of current microscope systems. This requires development of x-ray, EUV, SEM, AFM, focused ion beam (FIB), and optical microscopy with high sensitivity and resolution to be able to measure the critical dimensions of the features.

Secondly, new patterning methods are currently used and being pursued. Currently, 193 nm immersion lithography is the prevailing technology for manufacturing integrated circuits. To enable the 22 nm half-pitch process, double patterning, triple patterning, or even quadruple patterning UV lithography are being pursued. As the number of patterning steps increases, the number of metrology steps also scales up, which can potentially decrease the productivity. In the future, extreme UV (EUV) lithography, x-ray, nanoimprint, and directed self-assembly (DSA) patterning methods will also be investigated as the node process moves
beyond the 12 nm half-pitch. These new lithography methods will force the metrology industry to revolutionize their current metrology methods. For examples, in EUV lithography, there will be a demand in mask pattern defect inspection (such as film thickness variation, non-uniformity, reflectivity, etc.); in DSA, metrology tools need to be able to measure the size, location, and the alignment of the patterned contact holes that are also material dependent (even worse, those materials have very similar properties).

Thirdly, new materials are being pursued as the process dimension decreases, such as graphene-based circuits, high- and low-\(\kappa\) dielectric films gates, III-V film stack channels, and new doping materials for ultra-shallow junctions. In the material development, metrology tools need to be able to characterize both the physical and electrical properties such as the film porosity and thickness for both the surface layers and the buried layers, material contamination, current leakage and conductance in the gate.

Finally, complex 3D structures are also being developed, such as FinFET and MuGFET transistors and 3D interconnects, enabled by new lithography methods and new materials. Moving to 3D structures is accompanied with dimension down scaling. As discussed earlier, most of the metrology challenges come from measuring and inferring the film stack structures with limited measuring information. Thus, the major challenge facing the metrology industry is developing tools to measure the critical dimensions in 3D, even in buried structures.

So far, we have discussed the future trends and the related challenges in metrology. As metrology is a crucial component in the semiconductor manufacturing process, the metrology industry needs to proactively work on developing tools to satisfy the chip manufacturers’ standard. More importantly, they need to investigate for the future, and get ready for at least two generations ahead of the current technology.
3.3.2 Current metrology methods

Here, let us briefly review the metrology methods. There are many ways to categorize the metrology methods such as based on the wafer structure and inspection source or its working principle. Here we will use the inspection working principle to categorize metrology systems.

SEM is one of the most commonly used systems for defect metrology. SEM measures the scattered secondary electrons emitted by atoms on the sample surface, as a result of the electron beam excitation. SEM can be used to measure the surface topography as well as the material composition. In the metrology field, CD-SEM has been developed to measure the critical dimensions (CDs) of the smallest features, including line width roughness on the top, middle, and bottom of the feature; line edge roughness, sidewall roughness, angle, and profile; and depth or height of the feature. It is one the most commonly used tools for CD measurement. A discussion on critical dimension metrology is presented in reference [44].

Transmitted electron microscopy (TEM) (or a scanning type TEM, called STEM) is another metrology tool that uses the electron beams. TEM measures the transmitted electrons interacting within a thin specimen to provide internal structure and material composition information. A TEM system can achieve resolution down to 1 Å, which is an order of magnitude better than SEM. However, TEM has not been widely used in metrology due to its invasive nature, coming from the sample preparation and physical damage to the structure. Similar to SEM, focused ion beam (FIB) that uses ions instead of electrons is also used for CD metrology. The integration of FIB and SEM has been achieved in metrology tools.

Atomic force microscopy (AFM) is another widely used CD metrology tool. AFM is a point scanning method which uses a cantilever to sense the force, based on Hooke’s law, from the structure point by point. Based on the force measurement, AFM can render the surface topography and also identify different materials. In recent years, CD-AFM tools have become increasingly popular due to their advantages over CD-SEM as the critical dimension scales down. One of the advantage of CD-AFM is providing nondestructive measurements of the CD profile, where CD-SEM has a lot of trouble, even though tilted
CD-SEM can obtain some profile information. In addition, CD-SEM causes minor damage to CD structures, while non-contact AFM is nondestructive. CD-AFM certainly has many drawbacks, one of the most significant is the inspection speed. Even though video-rate AFM systems are being developed, the speed is still far below CD-SEM [52]. Another drawback of AFM is the complex tip characterization and tip-wear. Other drawbacks of AFM include the inability to measure high aspect ratio or dense structures (the tips cannot reach down). In fact, hybrid methods that combine CD-AFM and CD-SEM (and also optical CD tools) have been proposed to combine the advantages of two methods.

X-ray diffraction based techniques are also studied for metrology. In the early days, x-ray diffraction was mostly used for determining the crystal structures and later the DNA and protein structure. X-ray diffraction crystallography is based on the interference between the scattered x-rays coming from different planes in the crystal, which follows Bragg’s law of diffraction. X-ray diffraction provides the internal sample structure information such as particle size and shape distributions and material composition. The system typically uses wavelengths from 10 nm down to 0.01 nm; thus, the x-ray photons have very high energy, up to 10 keV, which can potentially ionize the sample. There are also many other issues associated with x-ray diffraction for CD measurement of small crystal structures [45]. However, CD metrology tools using small angle x-ray scattering (CD-SAXS) have some promising potential for thin film [45] and 3D metrology, revealing the buried structural information with high resolution and specificity, due to their nature of correlating the international structure information. In fact, over the past few years, Applied Materials has been working on developing such tools for FinFET measurement [144]. Extreme UV coherent diffraction imaging (EUV-CDI) is similar to CD-SAXS, but uses photons in the 10 nm to 124 nm range. Recently, the Kapteyn-Murnane group at the University of Colorado has developed a EUV-CDI tool that has reached a resolution of 24 nm. Their system is also capable of performing phase-contrast imaging and 3D tomographic reconstruction, which is promising for the metrology field [145]. However, this
method is still at the laboratory research level, and it has very low-throughput and is cost-ineffective. In the future, it may hold promise for EUV lithography mask inspection.

There are many other metrology methods based on spectroscopy including FTIR, Raman spectroscopy (or surface enhanced Raman). However, due to the limitations of this dissertation, we will discuss optical metrology as the last method, which is also the most important metrology method for the dissertation purpose. A review and comparison of different metrology methods is presented in reference [45]. Optical metrology is currently the most widely used method in industry. Any method that collect optical photons (reflected or scattered) and uses the properties of photons (polarization, phase, or intensity) is considered as an optical metrology method. In industry, optical metrology tools are categorized (ambiguously) as scatterometry, ellipsometry, polarimetry, profilometry, reflectometry, bright-field, dark-field, etc. In the current metrology systems, it is very common to integrate as many metrology methods as possible to be able to handle the highly challenging task. Developing optical metrology tools is the main interest of this dissertation. In Section 3.4, we exclusively discuss different optical metrology methods.

3.4 Optical metrology
Optical metrology methods are advantageous over other methods due to their nondestructive and high throughput nature. Moreover, optical physics have been thoroughly studied over the past hundreds to thousands of years. Optical methods are irreplaceable in inline metrology. Over the past decades, highly stable lasers at different wavelengths ranging all the way from DUV to IR or even broadband light (such as based on LED or super-continuum generation) have been developed. Those light sources have been implemented in many optical wafer metrology tools. In wafer metrology, the applications are usually task oriented. The metrology industry is currently working on developing optical metrology at 2x/1x node processes. The developed metrology tools need to be able to perform overlay metrology (measuring the errors from multiple patterning), CD metrology (characterize the smallest feature profile), thin film
metrology (measure the film thickness and refractive index), profiling metrology (measure the surface topography), mask defect metrology (characterize the lithography mask pattern), and dust inspection metrology (detect 10 nm scale dust on un-patterned wafers). In the following, we discuss different optical metrology methods.

Fig. 3.6. An illustration of scatterometry for multilayer thin film grating structure measurement.

Scatterometry is probably the most widely used method in optical CD metrology. As wafer patterns are moving to 3D with multilayer grating structures, a technique that can measure the CDs on both the surface and below is in demand. Scatterometry, a technique that collects the spectra of the scattered light due to the grating structures at different diffraction angles, provides a solution for optical CD measurement. An illustration of scatterometry is shown in Fig. 3.6. A collimated broadband (or scanning source) incident light illuminates the multilayer stack at an incident angle $\theta_i$, and is reflected at an angle $\theta_o$. At $\theta_i = \theta_o$, the detector collects the 0th-order diffraction light which is also unscattered. Due to the complex grating structure, different diffraction orders are also created, similar to the grating diffraction equation: $d(\sin \theta_l +$
\[ \sin(\theta_m) = m \lambda, \] where \( d \) is the grating period, \( m \) is the diffraction order, and \( \theta_m \) is the diffraction angle of the \( m^{th} \) order. In a multilayer grating stack with a high aspect ratio, the diffraction is more complex. Thus, a system that can capture the scattered spectra at different diffraction orders has also been implemented. And in order to obtain the layer CD parameters, a reference sample also needs to be measured. Therefore, each test sample measurement is compared with the reference to infer the layer structure. As the number of layer parameters increases (refractive index, thickness, and material composition), the fitting becomes inefficient and difficult. To solve this problem in practice, a data library with optimization tools is established.

In a scatterometer system, one can also incorporate ellipsometry but with incident light at the Brewster’s angle for optimum performance. An ellipsometer measures the change of polarization due the film thickness or refractive index change. The equation that describes ellipsometry is: \[ \rho = r_p/r_s = \tan \Psi e^{i\Delta}, \] where \( \rho \) is the ratio between the s and p polarization, \( \Psi \) and \( \Delta \) are the amplitude and phase parameters which are also the free parameters to be fitted to extract the thin film thickness, refractive index, or layer structures. To realize the measurement, a linear polarizer converting the light into linear polarization and an analyzer measuring the polarization ratio \( \rho \) are inserted in the input and output beam paths, respectively. By scanning the incident beam angle or the spectrum, curves for \( \Psi \) and \( \Delta \) are obtained. Through fitting the measured curves with the reference curves, one can infer the film information. Similar to scatterometry, as the number of film parameters increases, the fitting becomes complex and inaccurate, because a global optimization is usually difficult to obtain. Typically, ellipsometry is based on the Jones matrix formulation of polarization, which assumes fully polarized light. In reality, the samples under test may induce a depolarization effect, making the light partially polarized. Thus, a more accurate treatment of polarization using the Stokes-Mueller formulation can be adapted. In this way, light of any degree of polarization (or incoherent light) can be characterized for sample measurement.
Unlike scatterometry, bright-field metrology tools are imaging-based methods; thus, they are ideal for high throughout inspection due to the areal measurement. In a bright-field imaging system, both the reflected and scattered photons are collected and measured as intensity distributions. Thus, it can be used for either patterned or unpatterned wafer inspection. Light source stability and system noise isolation are very crucial to achieve nanometer scale defect inspection sensitivity. In addition to light intensity measurement, interferometry has been added to bright-field to measure the phase of the light. The phase information relates to the structure topography that enables profilometry. One of the most widely used profilometry tools is based on a white-light source and the phase-shifting technique develop by Zygo and Wyko Corporation in the 1990s. As discussed in Chapter 2, there have been many interferometric techniques developed over the past two decades. It should be emphasized that the focus of this dissertation is on developing highly stable, interferometric, and high throughput optical metrology tools to inspect densely patterned 2x/1x process node Si wafers. Chapter 4 to Chapter 7 will cover the development of the metrology tool and the experimental validation.

There are many other optical metrology methods used in industry, such as reflectometry based on the time of flight principle (such as in radar) which is ideal for high aspect ratio structure measurement, for example, the through-silicon via (TSV). Briefly speaking, in reflectometry, the waves that are reflected from the top and the bottom of the structure create an interference pattern encoding the depth information of the TSV. Over the past decades, many near-field and far-field superresolution methods have also been developed, such as near-field scanning optical microscopy (NSOM), stimulated emission depletion microscopy (STED), structured-illumination microscopy, stochastic optical reconstruction microscopy (STORM), and photo activated localization microscopy (PALM). The far-field super-resolution methods are typically based on injecting artificial-fluorescence dyes into biological structures. We believe that intrinsic fluoresce, Raman scattering, multi-photon absorption, or parametric down conversion may also be
used as the endogenous contrast to achieve super-resolution imaging. In the near future, the metrology industry may also turn to develop superresolution imaging metrology tools as pointed out by T. Crimmins from Intel [143].

3.5 Summary
In conclusion, we have briefly reviewed the semiconductor industry and manufacturing process, in particular, the metrology industry. It should be clear by now that metrology plays a crucial and irreplaceable role in the semiconductor field. It ensures continuing the future node process development and the new lithography system development. Knowing the importance of metrology, we have discussed the metrology challenges, future trends, and current metrology methods. At the end, we reviewed optical metrology methods that leads to the topic of the dissertation on developing interferometric optical metrology tools for defect inspection of densely patterned Si wafers. Chapter 4 to Chapter 7 will be focused on the interferometric inspection system design, validation, numerical simulation, and experimental data analysis.
CHAPTER 4: WAFER INSPECTION TOOLS

In this chapter, we will describe the inspection tools that we have developed. The first inspection system will be the laser based epi-illumination diffraction phase microscopy system, or laser epi-DPM for short. The first version of this system uses a 532 nm frequency doubled solid-state laser which has successfully detected 20 nm defects in a 22 nm node wafer. The system was later updated by using a 405 nm diode laser that has much higher power stability and image resolution. This update enabled defect inspection in a densely patterned 9 nm node wafer. Following that, a white-light based inspection system has also been developed. The first system is just a bright-field white-light imaging system with scanning modalities in 3D. This system has been able to significantly improve the sensitivity of the 9 nm node wafer defect inspection. To add phase to the inspection, we have worked on developing a white-light epi-DPM system, called epi-wDPM. The system development is done and the sample test verification is also performed. However, when applying it for 22 nm and 9 nm node wafer inspection, there have been some issues which we will discuss to motivate directions for future research.

4.1 Laser interferometric inspection system
Lasers source are highly coherent, thus, ideal for interferometry. With recent technology advances, lasers with different features have been realized to satisfy the need in different applications. In this section, we develop an epi-DPM system with 532 and 405 nm laser illumination for patterned wafer defect inspection. The system design, working principle, and verification will be covered in this section.
4.1.1 532 nm laser based system design

![Diagram of epi-DPM system](image)

Fig. 4.1. Epi-illumination DPM (epi-DPM) system. This system uses a common-path interferometer geometry. The pinhole filter is used for filtering one copy of the beam after the grating into a reference beam. The signal beam and reference beam interfere at the CCD camera to form the interferogram.

The 532 nm laser epi-DPM wafer inspection system is illustrated in Fig. 4.1. The system uses a 532 nm frequency-doubled Nd:YAG laser as the illumination source. The laser is first coupled into a single-mode fiber for filtering out the high-order transverse modes. After the fiber, the beam is collimated and passes through a rotating diffuser to reduce the laser speckle noise. Then, the laser goes through a linear polarizer and half-wave plate at 532 nm, which are used to control the angle of the laser polarization relative to the wafer structure. After that, the beam enters an inverted microscope (Zeiss Axiovert 200 M) and is normally incident on the wafer. The reflected beam exits the microscope and is incident on a blazed diffraction grating where multiple orders of the beam are created. We select two orders, specifically the +1st-order and
the 0th-order, to pass through a 4f system which has a 75 mm focal length first lens and a 400 mm focal length second lens. The +1st-order beam has the highest diffraction intensity and is low-pass filtered through a 10 µm pinhole to serve as the reference beam, and the 0th-order beam serves as the signal beam. The two beams interfere at the CCD camera plane where the interferogram is captured. The common-path geometry and the rotating diffuser are used to physically reduce the noise. From the interferogram, the reflected wave amplitude and phase are retrieved. The retrieval process is discussed in more detail in Section 4.1.2. The phase image contains information about the sample’s topography (or surface profile) and the amplitude image quantifies the variation in the sample’s reflection coefficient.

4.1.2 Phase and amplitude retrieval

In epi-DPM, the formation of an interferogram is the result of the interference between 0th-order (signal) and 1st-order (reference) beams. When the optical path difference between the two orders is a half-integer or full-integer multiple of wavelengths, destructive or constructive interference will occur and dark or bright fringes will form, respectively. The wavefront of each beam is distorted by the topography of the sample. However, in the Fourier plane, the +1st-order is filtered down to its zero k-vector component by the pinhole. In other words, it only contains the directly reflected (unscattered) component, which has a flat wavefront at the CCD plane. Thus, there will be a spatially dependent optical path difference between the two beams, which causes the location of the fringes in the image to vary spatially. By analyzing the fringe pattern, we can obtain sample topography (phase) information. The fringe contrast is determined by the relative strengths of the signal and reference beams. Thus, it is related to the ratio of the reflection coefficient (amplitude) of the structure at each spatial position to the average reflection coefficient from the sample. Here we mathematically describe the phase and amplitude retrieval algorithm. A complete description of DPM is presented in a recent review paper [24].
At the camera plane, the reference beam which is low-pass filtered by the pinhole has a uniform distribution. It can be written as

$$U_0(x, y) = |A_0| e^{i\Phi_0},$$  \hspace{1cm} (4.1)$$

where $A_0$ and $\Phi_0$ are constants. The signal beam $C(x, y)$ which contains the sample information is modulated by the diffraction grating and becomes

$$U_1(x, y) = C(x, y) e^{i\beta x},$$  \hspace{1cm} (4.2)$$

where $\beta = 2\pi / \Lambda$, and $\Lambda$ is the grating period in the image plane; $C(x, y)$ contains the reflection amplitude and phase information from the sample. The irradiance due to the interference between the reference and signal beams is given by

$$I(x, y) = |U_0(x, y) + U_1(x, y)|^2$$

$$= |A_0|^2 + |C(x, y)|^2 + |A_0|C^*(x, y) e^{-i\beta x} e^{i\Phi_0} + |A_0|C(x, y) e^{i\beta x} e^{-i\Phi_0}.$$  \hspace{1cm} (4.3)$$

Equation (4.3) can be used to retrieve the sample information $C(x, y)$ through the Hilbert or Fourier transform method [22, 146]. Here we use the Fourier transform approach. In the Fourier domain, Eq. (4.3) becomes

$$\mathcal{F}\{I(x, y)\} = |A_0|^2 \delta(k_x, k_y) + C(k_x, k_y) \otimes C^*(-k_x, -k_y)$$

$$+ |A_0|C^*(-k_x, -k_y) e^{i\Phi_0} \otimes \delta(k_x + \beta, k_y)$$

$$+ |A_0|C(k_x, k_y) e^{-i\Phi_0} \otimes \delta(k_x - \beta, k_y),$$  \hspace{1cm} (4.4)$$

where $\otimes$ denotes two-dimensional convolution over $k_x$ and $k_y$. We see that the second, third, and fourth terms on the right side of Eq. (4.4) will create three orders that each contain the sample information. The second term is the 0th-order, but it is overlapped with the first term, making it unsuitable to retrieve the sample information $C(x, y)$. Using the sifting property of the Dirac delta function [7], we see that the third (-1st-order) and fourth term (+1st-order) are shifted along $k_x$ by $-\beta$ and $\beta$, respectively. Thus we can select
either of them to retrieve the sample information. We select the +1st-order by multiplying Eq. (4.4) with a circle function centered at \( k_x = \beta, k_y = 0 \). The radius of this circle is determined by the cut-off frequency of the objective lens of the microscope. Note that the grating period needs to be chosen such that the 0th-order is not overlapped with the +1st-order. After selecting the +1st-order, we shift it back to the origin, and obtain

\[
|A_i| \hat{C}(k_x, k_y) e^{-i \phi_i}. \tag{4.5}
\]

Now we apply the inverse Fourier transform to Eq. (4.5) to obtain the spatial domain sample information \( C(x, y) \). Since the sample is illuminated by a plane wave, if we only consider reflection from the wafer surface, we can simply write

\[
C(x, y) = |R(x, y)| e^{i \phi(x, y)}, \tag{4.6}
\]

where \( R(x, y) \) is the reflection coefficient, and \( \phi(x, y) \) is the phase due to sample surface height \( h(x, y) \) variation. In reflection mode, \( \phi(x, y) \) and \( h(x, y) \) are related by

\[
h(x, y) = \frac{\phi(x, y) \lambda}{4\pi}. \tag{4.7}
\]

We tested our system performance and characterized the spatial and temporal noise. We found that with the diffuser we have 2x smaller noise value. The spatial noise of our system is 5.18 nm and the temporal noise is 0.85 nm; see more details in Appendix A.1.

### 4.1.3 Experimental validation

We use a 40x objective with a numerical aperture (NA) of 0.9 (Zeiss EC Plan-Neofluar 40x/0.9 Pol) and the system has a FOV of 30 µm by 27 µm. We test our system on a silicon wafer sample. This wafer has 22 nm node patterns and is used later for defect inspection. On this wafer, there are letter makers that can be used for our system validation. The letter features are about 100 nm high and 2 µm wide. We first measure the A letter in an interferogram, as shown in Fig. 4.2(a). The fringe contrast in the interferogram
is essential for phase retrieval. Thus, we analyze the fringe contrast by selecting a region outside the pattern region as indicated by a red box. We plot the intensity profile in the horizontal direction after averaging in the vertical direction. In this plot, good fringe visibility is observed. Another way to check the system performance is by transferring the interferogram into the Fourier space, as shown in Fig. 4.2(c), where three orders are well separated from each other as desired. As we discussed in Section 4.1.2, the signal is contained in the +1st-order as shown in Fig. 4.2(d). By inverse Fourier transforming the +1st-order, the phase and therefore the surface profile can be retrieved. In Fig. 4.3, we show the retrieved surface profile of the letter A as the first image in the 2D image array, where surface profiles of several other letters are also measured and presented. This indicates that we can also use this system to measure the patterned wafer structures.

Fig. 4.2. Experimental validation of the system in the Fourier space by measuring letter structures on a 22 nm node silicon wafer. (a) An interferogram of a test letter. (c) Intensity profile of a selected region to shown the fringe contrast. (c) Magnitude of the Fourier spectrum in log scale. (d) The +1st-order of the Fourier spectrum which contains the test structure information.
4.1.4 405 nm laser based system

The 532 nm solid-state laser has power instability issues which prevented us from detecting smaller defects on a 9 nm node densely patterned wafer. Thus, we decided to use a more stable 405 nm diode laser (QPHOTONICS, QFLD-405-20S). This laser is measured to have a spectrum linewidth of about 2 nm at the imaging operation current of 60 mA. To further improve the laser power stability, we added a Faraday rotator based isolator after the laser, which can prevent back reflection of the light into the laser cavity. In
addition, we inserted a 405 nm narrow-band filter in front of the camera to remove the ambient room light. To characterize the noise improvement, we used the method illustrated in Appendix A.1 for the 532 nm laser noise measurements. In Fig. 4.4 we show the noise histogram with and without the isolator. Figure 4.4(a) and (c) are the spatial noise histogram without and with the isolator, respectively. Figure 4.4(b) and (d) are the temporal noise histogram without and with the isolator, respectively. From the histograms, we conclude that we have achieved spatial noise improvement from 6.92 nm to 1.65 nm and temporal noise improvement of 0.19 nm to 0.13 nm. When compared with the 532 nm laser with diffuser (refer to Appendix A.1), the overall spatial noise is reduced by 5.6x and temporal noise by 11.2x.

![Noise characterization of the 405 nm diode laser.](image)

Fig. 4.4. Noise characterization of the 405 nm diode laser. (a) The spatial noise histogram without the isolator. (b) The spatial noise with the isolator. (c) The temporal noise without the isolator. (d) The temporal noise with the isolator.
The Fourier plane filter is another issue of the 532 nm laser DPM system, where we usually use a homemade cardboard to roughly block the ambient light coming into the interference beam orders. To mitigate this problem, in the 405 nm laser system, we implemented a machined filter that exactly fits into the interference beam orders at this wavelength. With this filter, we were able to obtain interferograms with high fringe visibility. We compare the coupling of the ambient light for the 532 nm laser and the 405 nm laser in Fig. 4.5, where the magnitude of their interferogram Fourier spectra are plotted in log scale on a flat test sample. The Fourier components of the 532 nm laser, as shown in Fig. 4.5 (a), are not distinct due to the ambient light coupling, while for the 405 nm laser the orders are very clear. Notice that in Fig. 4.5 (b) the center of the +1<sup>st</sup>-order is not aligned with its maximum value which is due to the optical alignment error, which is estimated to be about 2%.

![Fig. 4.5](image)

Fig. 4.5. Comparison of the interferogram spectra from the 532 nm laser and the 405 nm laser. (a) 532 nm laser epi-DPM interferogram spatial spectrum. (b) 405 nm laser epi-DPM interferogram spatial spectrum.

After checking the interferogram spectra, we image the letter and number markers on a 9 nm node wafer, as we did for the 532 nm laser. The surface profiles for the letters are plotted in Fig. 4.6, where feature
height of about 50 nm is measured. The surface profile measurement needs to be validated with an alpha-step surface profiler.

![Surface profile measurement of letters and numbers on a 9 nm node silicon wafer.](image)

**Fig. 4.6.** Surface profile measurements of letters and numbers on a 9 nm node silicon wafer.

### 4.2 White-light interferometric inspection system

In contrast to lasers, white-light sources have low spatial and temporal coherence. The spatial coherence is determined by the spreading of the wavevector $k$ in space, while the temporal coherence is determined by the spreading of the wavelength $\lambda$. A halogen lamp emits light at all directions and wavelengths randomly, thus, it is spatial and temporally incoherent. As we know, incoherent light imaging systems are speckle free due to low coherence, allowing for highly sensitive measurements. To realize interference with incoherent light, spatial filtering through a pinhole is usually done to improve the spatial coherence. The optical path delay between the two interference beams also needs to be compensated within the temporal coherence length to ensure interfering. In this section, we propose a white-light interferometry system based on the laser epi-DPM principle, called epi-wDPM. We also propose adding 3D scanning and selective Fourier plane filtering to the epi-wDPM system to further enhance the defect sensitivity.

#### 4.2.1 White-light bright-field inspection

The white-light interferometric system is built upon a white-light bright-field microscopy system. It was expected that if we can detect defects in the intensity images in the bright-field system, then we could add interferometry to retrieve the phase images and improve detectability. In Fig. 4.7, we illustrate the white-
light bright-field inspection system. The system is based on a Zeiss Z2 microscope which uses a halogen lamp illumination that covers the spectrum from 400 nm to 1100 nm as the white-light source. The microscope is equipped with an XYZ (3D) sample scanning stage to move the sample across the field of view. In this system, a white-light source (a halogen lamp here but we use a white-light LED later for the interferometric system) is first imaged onto the aperture diaphragm which selects the portion of the filament source for Köhler illumination. (Notice that in Köhler illumination for reflection, there is no condenser aperture but just an aperture diaphragm, since there is no condenser lens but the objective lens in the place of the condenser.) The back focal plane of the objective is the conjugate plane of the aperture diaphragm. Thus, plane waves at different angles illuminate the sample, ensuring uniform intensity on the sample. Köhler illumination is advantageous to critical illumination where the filament is directly imaged on to the sample, resulting in an undesired nonuniform intensity modulation coming from the light filaments.

Fig. 4.7. The white-light bright-field microscopy system.

In white-light bright-field imaging, the spatial coherence of the light source is an important factor to consider, especially for improving the image contrast in the context of the modulation transfer function [7],
which is crucial for quantitative measurement of the surface profile in interferometry [129] (this reference is a complete study of spatial coherence for white-light interferometry with a transmission white-light DPM system as an example). By adjusting the condenser aperture size, the spatial coherence length $l_s$ of the light source on the sample can be also controlled, described by the relation $l_s = 0.61\lambda/\text{NA}_{\text{con}}$, where $\text{NA}_{\text{con}}$ is the condenser numerical aperture at the back focal plane of the objective. To achieve the best contrast, $l_s$ needs to be sufficiently larger such that the illumination on the sample field of view acts like a plane wave. This is important for phase imaging, which we will discuss in Section 4.2.2. The temporal coherence is another important factor to consider that relates to the image contrast with speckle. The temporal coherence length is inverse proportional by the spectrum bandwidth. Smaller temporal coherence length illumination suffers less speckle. The spatial coherence also affect the speckle similarly like the temporal coherence, and this was confirmed with a super-continuum light source [129] (with small temporal coherence length and large spatial coherence length, but has speckle like a regular laser). The 4f system after the intermediate image plane is used to relay and magnify the sample pattern onto the camera plane.

For wafer inspection, we use a 40x, NA = 0.9 objective with the aperture diaphragm NA closed down to about 0.1. The 4f system uses a 60 mm focal length achromatic camera lens (Canon EF-S 60 mm f/2.8 Macro USM) as lens1 and the lens2 is a regular Thorlabs achromatic lens with 150 mm focal length. The field of view of the system is about 45 $\mu$m by 34 $\mu$m. We mount the wafer on the XYZ scanning sample stage. We can scan the wafer in the transverse (XY) or axial (Z) directions to improve the detectability, and the experimental results on inspecting a 9 nm node wafer are presented in Section 6.4. In Section 4.2.2, we discuss the white-light interferometer design, but it should be noted that the Zeiss Axiovert inverted microscope was used throughout the dissertation. Only the white-light bright-field inspection used the Zeiss Z2 microscope (it actually belongs to a different project), and it was only used for the purpose of testing white-light inspection.
4.2.2 Interferometer design

As mentioned above, the white-light bright-field microscopy system is based on the Zeiss Axiovert inverted microscope. The system configuration is still the same as shown in Fig. 4.7, except that the white-light source and the Köhler illumination are built outside the microscope. In order to perform white-light bright-field interferometry, or epi-wDPM here in particular, we need a white-light source and an interferometry design. The white-light source we decided to use is a cold white-light LED (Thorlabs MCWHL5 source, cold meaning the spectrum is close to blue) with 800 mW maximum output power. The spectrum of this light source is plotted in Fig. 4.8 (data from Thorlabs). The spectrum has an abnormal and non-symmetric profile. When measuring the sample surface profile, we need to know the mean wavelength of the light source. Using the spectrum data, we determined the source has a mean wavelength at 503 nm. It is also necessary to measure the temporal coherence length, which limits the maximum surface height.
measurement and also relates to speckle effect. For such a non-symmetric spectrum, we need to determine the power equivalent spectrum bandwidth $\Delta \lambda = |\lambda_{\text{max}} - \lambda_{\text{min}}|$ and use $l_c = \frac{\lambda_{\text{max}} - \lambda_{\text{min}}}{\Delta \lambda}$ to calculate the temporal coherence length. The $\lambda_{\text{max}}$ and $\lambda_{\text{min}}$ values are calculated to be 532.2 nm and 473.2 nm, thus $l_c$ is found to be $\sim 4.3$ µm.

Fig. 4.9. The interferometer system for epi-wDPM.

To make an interferometer, we need to create a reference beam and a signal beam. This is done with a diffraction grating and a projector-based spatial light modulator (SLM), as shown in Fig 4.9. We use a 110 lpm grating at the intermediate image plane which creates multiple orders of the image. And in the Fourier plane of the 4f system, the SLM passes the $+1^{\text{st}}$-diffraction order and low-passes the $0^{\text{th}}$-diffraction order by
projecting a black (block) and white (transmission) pattern from a PowerPoint slide on a laptop. Notice that, to avoid creating astigmatism for the image in the laser epi-DPM system, we low-pass the +1st-diffraction order and let the 0th-diffraction order be the image order, which always passes through the center of the lens. However, in epi-wDPM the +1st-order would spread out all the colors in the SLM plane due to dispersion, making it impossible to low-pass filter it. Thus, the epi-wDPM system uses the 1st-order as the image order and suffers from the astigmatism.

To ensure the two beams interfere with each other in the image plane requires precise overlaying of the two beams with each other within the temporal coherence length (lasers are more tolerant due to the orders of magnitude larger temporal coherence length). This is done by first passing the two beam orders without filtering, essentially two signal beams. Then we adjust the camera position until the structure features in the two signal orders overlap with each other on the camera. After the overlay is done, we use the low-pass filter pattern on the SLM as illustrated in Fig. 4.9. The 0th-order filter needs to be sufficiently small such that it approaches a plane wave on the image plane for making an ideal reference beam. As discussed in Section 4.2.1, adjusting the aperture diaphragm is also important since it relates to the spatial coherence. Thus, in phase retrieval, the aperture diaphragm and the SLM filtering are equally important, as described in reference [129]. According to this reference, the measured phase is the true phase subtracting the low-pass filtered phase due to the filter and illumination, which can be expressed as

\[
\phi(r) = \arg[S(r)] - \arg\left[\mathcal{T}_{ill}(r)\tilde{T}_{ref}(r) \otimes S(r)\right],
\]

where \(S\) is the sample reflection function, \(T_{ref}\) is the SLM reference pinhole aperture function (tilde denotes Fourier transform), and \(T_{ill}\) is the illumination mutual intensity function. By assuming a shift-invariant illumination field, we can defined \(T_{ill}\) as

\[
T_{ill}(r - r') = \left\langle U_{ill}(r, t)U_{ill}^\ast(r, t)\right\rangle_t,
\]

where \(r\) is the position vector and \(t\) is the time delay.
where $U_{ill}$ is the sample illumination field selected by the aperture diaphragm $T_{ap}$, as

$$U_{ill}(r,t) = \hat{U}_s(r,t) \otimes \hat{T}_{ap}(r). \quad (4.10)$$

Note that $T_{ill}$ and $T_{ref}$ are not independent with each other, a smaller aperture diaphragm makes it easier to make a smaller reference beam filter, this is reflected by the product $T_{ill}(r)\hat{T}_{ref}(r)$ in Eq. (4.8). In Eq. (4.8), the first term $\arg[S(r)]$ is the true phase of the sample. Notice that the true phase may not be the surface profile of the object, depending on the internal object structure. The second term $\arg\left[T_{ill}(r)\hat{T}_{ref}(r) \otimes S(r)\right]$ is the low-pass filtered version of the true phase. Thus, the overall phase is high-pass filtered. Ideally, we need both $T_{ill}(r)$ and $T_{ref}(r)$ to be uniform (both the aperture diaphragm and the SLM pinhole filter to be delta functions), then the second term becomes a constant that can be removed through calibration. However, this is impractical, because almost all the power of the light source will be blocked, resulting in a noisy interference image. In reality, the aperture diaphragm will have a certain size that compromises the coherence and power, and the SLM pinhole filter will have a certain size to balance the fringe contrast and power. To determine the coherence area of the system, we need to measure the aperture diaphragm numerical aperture, which can be done in its conjugate plane on the SLM, as shown in Fig. 4.8 and Fig. 4.9. On the sample stage, we can use a diffusive objective to fill the resolution circle in the 1st-diffraction order. The white-light source selected by the aperture diaphragm will also show up in this diffraction order. Thus, by measuring the ratio $\sigma$, defined as the resolution circle diameter over the light source circle diameter, the spatial coherence length can be determined as $d_{air}/\sigma$, where $d_{air}$ is the Airy spot (or diffraction spot) diameter, $d_{air} = 1.22\lambda_{mean}/NA_{obj}$. For a fully coherent source $\sigma$ is 0, thus infinite coherence, whereas for a completely incoherent source $\sigma$ equals to 1 that corresponds to a coherence length of the diffraction spot. Thus, in white-light bright-field imaging, with no light source filtering, the objective lens determines the coherence length.
In most cases, having spatial coherence length to cover the field of view is difficult to obtain [129]. Practically, we can just achieve full spatial coherence over the sample. For cell imaging, spatial coherence length of ~ 10 μm is usually sufficient, while for patterned wafer defect inspection a coherence length of a diffraction spot is sufficient. Thus, for the wafer inspection experiments, the coherence is not an issue for phase imaging. However, it is still desirable to have relatively good image contrast through reducing the aperture diaphragm NA at the expense of light power.

There is one more thought about Eq. (4.8) that relates to the halo effect in phase contrast or quantitative phase image. We think the true phase in Eq. (4.8) is actually not the true phase, which should also incorporate the illumination, i.e., the halo is already created by the aperture diaphragm (or condenser aperture in transmission) and it is carried in the phase. The reason that halo is carried in phase is that we only see the halo in phase imaging not in bright-field intensity. Filtering the reference beam, even down to a delta function, cannot remove the halo, because it is buried in the signal beam as a phase modulation. The only way to remove it is by using a small enough aperture diaphragm (or condenser aperture in transmission), such that the illumination on the sample is as uniform as a laser. This reasoning is based on the measurement results from reference [129], where a small enough condenser aperture was also very necessary to ensure the correct phase.

4.2.3 3D scanning automation and real-time image processing

For wafer defect inspection, a precise and automated XYZ 3D sample scanning stage is necessary. It is also desired to have a real-time inspection system that can integrate the camera image acquisition, XYZ stage control, and image post-processing into a single graphic user interface (GUI). Over the past year, we have been able to implement such a system in NI LabVIEW as illustrated in Fig. 4.10. The XYZ stage is assembled from three separate actuator based translation stages (Thorlabs, ZST225B) with 25 mm travel distance and 8 nm minimum step size. In addition to the XYZ stages, we also have tip, tilt, and rotation
correction. The XYZ stage is fitted onto the Zeiss Axiovert inverted microscope with a custom machined stage. A setup image for the XYZ stage is shown in Fig. 4.11. The old microscope scanning stage, which has only two translation axes and 250 nm minimum step size, is removed. This system is capable of performing wafer scans with specified step size, direction, and range. It can also perform sequential scans in different directions.

3D translation stage

Real-time processing

Camera

Fig. 4.10. Illustration of the real-time defect inspection system.

In the real-time processing LabVIEW program, the camera control uses the NI Vision module and the Hamamatsu video capture library. The XYZ stage motors are controlled with the Thorlabs software through LabVIEW. The image processing is done mostly with the NI CUDA module (such as performing Fourier transforms) with the most efficient image retrieval algorithm. CUDA is designed for parallel computation in GPU by NVIDIA. In order to use the NI CUDA module, an NVIDIA certified GPU is installed.
Fig. 4.11. The installed XYZ sample scanning stage.

Fig. 4.12. XYZ stage test with different acceleration and velocity.

In Fig. 4.12, we test the stage motor response time at different acceleration and velocity. In the plot, the green curve is the backward stepping, the red curve is the forward stepping, and the white curve is the theoretical minimum. The backward stepping is slower since the stepper goes past the set point and then moves forward (reduces hysteresis errors). Although the spec-sheet maximum speed is 1 mm/s, we can only get to 0.75 mm/s, due to increased friction from the motor (unknown reasons). In summary, it takes
about 200 ms for 0 \mu m < d < 20 \mu m, 450 ms for d = 100 \mu m, and 2000 ms for d = 1 mm, where d is the step size. Note that it even takes 125 ms to make a step of d = 0, where the 125 ms is mostly the time of the command message (there are two command messages for each step: ask the motor to move and check if it is done).

The image post-processing takes about 210 ms which includes retrieving the phase and amplitude image and computing the histogram and the fringe visibility curve. The camera (Hamamatsu C4742-80-12AG) image acquisition takes about 125 ms. Thus, the overall time for one image capture takes about 500 ms for small scanning steps which corresponds to two frame/s. The real-time processing program can also be used for system alignment in DPM. In fact, a LabVIEW automatic pinhole alignment program has also been demonstrated in a different DPM system. By integrating the pinhole alignment and the real-time processing, the whole DPM system can be automated, which will be used in the shared DPM facility in the Beckman Institute starting the Fall of 2014 (co-developed by the Photonic Systems Lab and the Quantitative Light Imaging Lab).

4.2.4 Table-top cleanroom
During the past years, we found that keeping the wafer clean under the microscope has been a big challenge. Over time, dust particles, usually charged, get trapped onto the wafer structure, preventing us from detecting defects down to 10 nm size. Therefore, it is necessary find a solution to keep the wafer clean (or find ways to clean the dust particles that accumulate on the wafer). Several methods have been proposed in the past. The first proposed method was mounting the wafer in a vacuum chamber, but it was thought to be complicated and not cost-effective. The second proposed method was to cover the microscope with a plastic bag, and for a long time this was practiced and seemed to be helpful over a short period of time. However, after a month, the wafer would still get dusty. To keep the wafer clean over a long period of time, we
proposed building a table-top cleanroom to prevent the dust particles from coming into the system in the first place.

A modular cleanroom, as shown in Fig. 4.13 (image from Terra Universal, specialized in designing cleanrooms), is generally used for device fabrication outside a cleanroom, but it is very expensive, about $100k. But for our case, we only need a 4’ by 6’ size cleanroom (probably class 10,000) over the microscope system; thus, a table-top cleanroom is sufficient. To build a table-top cleanroom, a fan filter unit (NEW Gordon Phantom 771129 Cleanroom Filter Fan Unit, efficiency 99.9995%), static dissipative boards (UHMW plastic from Cope Plastics), and frame structure are needed, as shown on the top of Fig. 4.13. The fan filter unit exchanges air between the chamber and the ambient with a HEPA filter and a fan. The fan maintains a positive air pressure inside the chamber. The static dissipative boards and the frame structure forms the chamber; thus, dust particles which are usually charged will not be attracted to the chamber.
A lot of machining work was done with the help of the Electrical and Computer Engineering (ECE) workshop at UIUC to make the chamber fit onto the table. Also, a special power supply (277 Volts, 1 phase, 60 Hz) was needed and installed with the help from the electrician in the Beckman Institute. After cleaning the chamber and testing the fan filter unit, we moved the inspection system in place. Fig. 4.14 shows the table-top cleanroom post-completion.

![Image](image_url)

Fig. 4.14. Table-top cleanroom post-completion.

4.2.5 System validation

So far, we have discussed all the aspects of the white-light inspection system. In this section, we will show the system testing results. In the epi-wDPM system that we constructed, an iris is used as the aperture diaphragm. To tradeoff between beam power for contrast, the iris is closed down to ~ 1 mm (the minimum) to select the light source. As discussed earlier in Section 4.2.2, the size of the aperture is crucial for coherence, but for wafer inspection coherence is not an issue (features are much smaller than the diffraction spot). However, we still need a small aperture diaphragm to improve the image contrast, and we found experimentally that a 1 mm aperture is ideal. In the Fourier plane, we gradually shrunk the size of the reference beam pinhole filter on the SLM. As discussed earlier, ideally we need the beam to be a plane wave, at least across the camera sensor area, which requires the size of the pinhole to be around 10 µm. However, as the pinhole size gets smaller, the reference beam power decreases significantly, resulting in
poor fringe visibility on the camera. To maintain relatively good visibility, we have to use a pinhole of about 200 µm. However, this induces a high-pass filtering effect in the phase image as we described earlier. This high-pass filtering effect is notorious for large sample structures, but can be actually helpful in defect inspection where structures are tiny and defects are sharp. The high-pass filtering provides edge detection in the optical domain.

Before imaging test samples, we first check the fringe contrast and the Fourier spectrum of the fringes, as shown in Fig. 4.15. Figure 4.15(a) is an interferogram taken on a flat sample, Fig. 4.15(b) is a zoomed-in portion of the interferogram, Fig. 4.15(c) is a normalized fringe profile plot along the horizontal dimension of the zoomed-in interferogram with averaging in the vertical dimension, and Fig. 4.15(d) is the log scale plot of the interferogram Fourier spectrum magnitude, where we see three beam orders are clearly separated as expected. From the analysis, we conclude that good fringe contrast has been achieved.

Fig. 4.15. Fringe contrast analysis of the epi-wDPM system. (a) An interferogram on a flat sample region. (b) A zoomed-in portion of the interferogram. (c) The intensity along the horizontal dimension with the vertical dimension averaged. (d) The Fourier spectrum of the interferogram.
With good fringe contrast, we started to perform sample tests on a photo-lithography silicon sample which has different structures with about 25 nm designed height. We first test large structures such as a 40 μm squares as shown in Fig. 4.16. The normalized amplitude (Fig. 4.16 (a)) and surface profile (Fig. 4.16 (b)) images both show only the edge of the structure, which is due to a lack of spatial coherence as we expected. Then, we tested samples on smaller structure features where almost full coherence coverage might be achieved. In Fig. 4.17, we measure a horizontal alignment sample in amplitude and phase in Fig. 4.17 (a) and (b), and a vertical alignment sample in Fig. 4.17 (c) and (d), where 2 μm wide lines are gradually dislodged starting from the center. The full structures start to show up, indicting almost sufficient spatial coherence. Finally, we measure the amplitude and surface profile of grating structures with linewidth around the diffraction spot size, as shown in Fig. 4.18. Two different gratings are measured, one with about 300 lpm as shown in Fig. 4.18 (a) and (b), and one with 700 lpm as shown in Fig. 4.18 (c) and (d). The measured heights of the structures from the surface profile images are close to the designed 25 nm, but not exactly correct. Several reasons may account for this, such as misalignment and lack of spatial coherence in the system. It is also possible that the samples were not fabricated perfectly. To validate the epi-wDPM measurement data, SEM and alpha-step measurements are needed.
Fig. 4.17. Test measurement of two alignment structures. (a) Amplitude image of the horizontal alignment sample. (b) Surface profile of the horizontal alignment sample. (c) Amplitude image of the vertical alignment sample. (d) Surface profile of the vertical alignment sample.

Fig. 4.18. Test measurement of two grating structures. (a) Amplitude image of the 300 lpm grating sample. (b) Surface profile of the 300 lpm grating sample. (c) Amplitude image of the 700 lpm grating sample. (b) Surface profile of the 700 lpm grating sample.
4.2.6 Selective Fourier filtering

The SLM sitting in the Fourier plane can also be used for selected signal beam filtering depending on the spectrum of the light and the structure of the sample. For example, with a broadband source, on the SLM the +1st-order will spread out the colors. Therefore, by changing the projected +1st-order pattern, we can change the spectrum of the light, similar to using a physical color filter (this is demonstrated experimentally). We have found that by using the SLM to select the blue portion (around the 450 nm peak of the white-light LED) we can improve the phase image contrast (more effectively remove the amplitude modulation of the grating on the camera). The SLM filtering possibility has enabled us to do spectroscopic imaging, as demonstrated earlier with manual tuning without using a SLM [147]. Spectroscopic imaging will be helpful for wafer defect inspection, since the wafer grating structure is wavelength dependent. We can also filter out the periodic structures on the sample and maintain only the feature signal of interest such as the defect in wafer inspection. Performing structural filtering may require light sources with high power, such that the signal after filtering is well above the noise level on the camera for the best noise suppression.

4.3 Summary

So far, we have presented the laser and the white-light interferometric system for wafer defect inspection. The system design, working principle, and experimental validations are presented in great detail. Both defect inspection systems have their own limitations. Laser interferometric systems like epi-DPM are easy to build but suffer laser speckle and cannot be used for spectroscopic imaging. On the other hand, white-light systems like epi-wDPM are more difficult to build, if quantitative phase is required, but offer better image contrast and can also be used for spectroscopic image.
CHAPTER 5: 22 NM NODE WAFER INSPECTION

5.1 Introduction

Recently, the semiconductor industry has shrunk their critical dimension down to 22 nm. Thus, detecting defects in a 22 nm node wafer has become critical. We developed a defect inspection system based on an epi-illumination diffraction phase microscopy (epi-DPM) with a 532 nm laser as illumination for 22 nm node wafer inspection. This system, as described in Section 4.1, measures an interferogram, which is used to retrieve both the amplitude and phase of the scattered field from the sample using a Fourier transform method. However, in order to detect sub-wavelength defects in the 22 nm node wafer, we need to reduce the residual noise in the images and amplify the defect to background signal contrast. To overcome this problem, we first collect a sequence of amplitude and phase images by translating the wafer in one direction with 0.75 micron steps. Then, we use the scanning images to produce 2nd-order difference images to reduce spatial noise and also produce a tripole pattern for the defect. Next, we stitch the 2nd-order difference images together to produce a stretched panoramic image to reduce temporal noise and system calibration errors. Finally, the stretched image is convolved with a matched tripole pattern to extract the defect and eliminate the background signal from the wafer’s underlying structure. We call this image post-processing method 2DISC, which uses 2nd-order image difference, image stitching, and convolution. Applying 2DISC for our collected images, we examined a 22 nm node intentional defect array (IDA) wafer. The results show that we can reliably detect defects down to about 20 nm × 160 nm for different types of defects. We verified our detection with a scanning electron microscopy (SEM). The detection limits are studied using receiver operating characteristic curves.
5.2 22 nm wafer imaging

A defect-free portion of the wafer is shown in Fig. 5.1(a). This pattern is periodic and consists of parallel lines made of polysilicon on a silicon substrate. The pattern consists of two different lines that are 22 nm in width, 120 nm or 260 nm in length, and 110 nm in height. The lines are arranged to form a 0.8 µm by 0.8 µm square unit cell containing eight lines. The unit cell is repeated in a rhombic lattice pattern to form a two-dimensional array with an area of 100 µm by 100 µm. Figure 5.1(b) shows a zoomed-in portion of the defect-free pattern with the locations of four different types of defects indicated. Parallel (red) and perpendicular (green) bridge defects are shown in Fig. 5.1(c) and (d), respectively; an isolated (light blue) defect is shown in Fig. 5.1(e), and a line extension (dark blue) defect is shown in Fig. 5.1(f). Each defect is
located in the center of the two-dimensional array and only one defect type is printed per array. In adjacent arrays, various sizes of these defects are printed by varying the linewidth of the defect. In the following sections, we use the 532 nm laser-based epi-DPM defect inspection system and a comprehensive image post-processing method for finding the defects.

The wafer inspection system uses a 40x objective with a numerical aperture (NA) of 0.9 with FOV of 30 µm by 27 µm. According to the Rayleigh criterion, the diffraction limited resolution is calculated to be about $1.22\lambda/(NA_{\text{objective}}+NA_{\text{condenser}}) = 720$ nm since $NA_{\text{condenser}} = 0$ for collimated illumination. Thus, individual features of the patterned wafer will not be resolvable, but rather blurred by the point spread function (PSF) of the system. Since the features under inspection are deep sub-wavelength, the system is still far from detecting defects in the pattern. Any subtle residual noise present in the system can greatly affect the image contrast. Also, slight system calibration errors such as defocus, sample-tilt, aberration, non-uniformity or time-varying illumination intensity can also degrade the image quality substantially.

Here, we show an example of phase and amplitude retrieval in the wafer structure using the method described above. Figure 5.2(a) is an interferogram image of the wafer with a 160 nm by 20 nm parallel bridge defect in the center. The amplitude of the Fourier transform of the interferogram plotted on a log scale is shown in Fig. 5.2(b). In this figure, we can identify the three orders of the signal. The $+1^{\text{st}}$-order is selected and moved to the origin for phase and amplitude retrieval. The retrieved phase and amplitude image are shown in Fig. 5.2(c) and (d), respectively. Note that in these single frame images, a single reference image is first collected and used to remove the common noise background. In both images, we see the wafer’s underlying structure; however, the defect cannot be identified in a single frame image due to the noise in the system.
Fig. 5.2. An example of phase and amplitude image retrieval algorithm in the wafer. (a) Interferogram image. (b) The amplitude of the Fourier transform of (a) in log scale. (c) The retrieved phase image. (d) The retrieved amplitude image. The color scale image values are in arbitrary units. In the center of each image, there is a 20 nm by 160 nm parallel bridge defect. Due to the noise, it is not detectable in a single frame amplitude or phase image.

5.3 Image post-processing

In order to detect the defect, we collect a sequence of images by translating the wafer in the direction parallel to the underlying line structure. With the scanning images, we propose an image post-processing method called 2DISC to remove different types of noise and system imperfections and then extract the defect signal.

Here, we explain the 2DISC method for removing noise and system imperfection. We define system imperfection as the noise in the image profile due to non-uniformity in the illumination source, illumination path, or camera response. We will illustrate the 2DISC method using the array region that has a 20 nm by 160 nm size parallel bridge defect (Fig. 5.1(c)). First, a sequence of 90 interferogram frames with adjacent
frame horizontal translation steps of 0.75 µm is collected. The choice of translation direction and step size are explained in Appendix A.3. We calculate the amplitude and phase for each image from the captured interferograms. As described previously, in a single amplitude or phase image, we cannot detect the defect due to the different types of residual noise. The spatial distribution of noise can be decomposed into time-variant and time-invariant sources. Time-invariant noise is mostly due to the laser speckles created at the optical component surfaces. Since this noise is also shift-invariant, especially when the sample is translated a small step, we can remove this noise by using the 1st-order difference, which is defined as: $F_{n+1}(x,y) - F_n(x,y)$, where $F_n(x,y)$ denotes the $n$th phase or amplitude frame of the scan. However, it was found that the 2nd-order difference yields better noise reduction and image contrast. See Appendix A.3 for a systematic study of the difference order. The 2nd-order difference image frame $n$ is obtained as

$$F_{n+1}(x,y) - 2F_n(x,y) + F_{n-1}(x,y).$$

Figure 5.3 shows three different 2nd-order difference amplitude frames in an example array, the location of the defect is marked by the red rectangular box. In those images, the shift-invariant noise is removed, but due to time-variant noise, we still cannot clearly detect the 20 nm by 160 nm defect. Averaging in time is a simple way to remove time-variant noise in the images. However, instead of measuring for a long period of time at a single location, we translate the wafer and measure the defect at different locations in the image. Due to the large overlap of the adjacent frames, we obtain equivalent time-averaging by stitching together all of the scan images to produce a panoramic average image, see Appendix A.2. Translation and stitching with averaging further reduces system non-uniformity noise because the defect is measured at many different locations in the laser illumination and on the camera.

Figure 5.3(d) is the panoramic amplitude image of the example array. Similarly, Fig. 5.3(e) is the panoramic phase image. Phase and amplitude images give different information about the defect. In Fig. 5.3(d) and (e), the defect at the center can be clearly identified. However, the signal from the underlying
structure reduces detectability. This signal can be complicated depending on the structure. Thus, if we want to extract just the defect information that is buried in the structure, we need to develop a suitable extraction algorithm.

Fig. 5.3. Illustration of the 2DISC method for defect detection. A defect array with a 20 nm by 160 nm parallel bridge defect in the center is used as an example. (a)-(c) Example 2nd-order difference amplitude image frames showing the defect moving right to left across the field of view as the wafer is translated to the right. (d)-(e) Panoramic 2nd-order difference amplitude and phase images, respectively. These are the 2DISC images prior to the convolution step. (f)-(g) Full 2DISC amplitude and phase images, respectively. For each image, the location of the defect is marked by a red rectangular box.

After the 2nd-order difference and image stitching steps, the defect will appear as a tripole pattern of +1 -2 +1, see Fig. 5.4 for a schematic description. To detect this pattern, we convolve the final panoramic image with a matched tripole pattern. The matched tripole pattern \( M(x, y) \) has a form of

\[
M(x, y) = T(x + ds, y) - 2T(x, y) + T(x - ds, y),
\]

(3.2)
where \( T(x, y) \) is a test function and \( ds = 0.75 \, \mu \text{m} \) is the separation between the poles. \( T(x, y) \) should be optimized to obtain the maximum defect signal and reject the underlying structure. We designed \( T(x, y) \) as a Gaussian function. Convolution serves as a low-pass filter. We picked the standard deviation of the Gaussian to be 360 nm. This choice is large enough to filter out high-frequency components in the 2\(^{nd}\)-order difference image due to the wafer’s underlying structure, but small enough that it does not filter out the tripole pattern of the defect signal that has spacing \( ds = 0.75 \, \mu \text{m} \). See Appendix A.3 for a detailed discussion. In the end, we obtain a defect pattern matched panoramic image, i.e. the 2DISC image. Figure 3.3(f) and (g) are the 2DISC amplitude and phase images of the example array, respectively. In both images, we can see the defect signal contrast has significantly improved as compared with Fig. 5.3(d) and (e). For a quantitative study, we can compute the peak-signal to noise ratio (PSNR) which is defined as

\[
PSNR = 20 \times \log_{10} \left( \frac{|D_{\text{max}} - D_{\text{min}}|}{\sigma} \right),
\]

where \( D_{\text{max}} \) and \( D_{\text{min}} \) are the maximum and minimum signal strength around the defect region, respectively and \( \sigma \) is the standard deviation of the whole panoramic image. The defect region is chosen to be a 4 µm by 4 µm rectangle. The term \(|D_{\text{max}} - D_{\text{min}}|\) is a measure of the defect tripole pattern contrast. Since \(|D_{\text{max}} - D_{\text{min}}|\) is always larger than \( \sigma \), there is a non-zero baseline of around 12-16 dB for PSNR even if the image contains only random noise. Thus, only the difference in PSNR should be used, e.g. for comparing amplitude versus phase images or quantifying the improvement from adding a post-processing step. For the amplitude image, the PSNR increased by 7.0 dB from Fig. 5.3(d) to Fig. 5.3(f). For the phase image, the PSNR increased by 8.7 dB from Fig. 5.3(e) to Fig. 5.3(g). The 2DISC amplitude image in Fig. 5.3(f) has slightly larger PSNR (1.8 dB) than the 2DISC phase image in Fig. 5.3(g). To validate that the overall approach of the 2DISC method with 0.75 µm step size is a good choice, we present a systematic study of PSNR for difference orders from 0\(^{th}\) to 4\(^{th}\)-order in Appendix A.3. The study also examines the noise reductions from adding the
image stitching step and the convolution step as well as the effects of changing the step size and the test function width.

![Fig. 5.4. Illustration of the 2nd-order difference image stitching. (a) A schematic description of the sequence of 2nd-order difference images showing the expected tripole pattern when the defect is in the field of view. (b) The stitched 2nd-order difference image.](image)

### 5.4 Sensitivity and detectability

The PSNR is a single number that can be useful to quickly evaluate the defect detection method. However, to study the improvement in defect detectability, ROC curves [148] are used since they quantify the tradeoff between sensitivity and specificity of the defect classifier algorithm. Here, we will continue to discuss the previous 20 nm by 160 nm defect example. We first study how well we can find a defect in any single image frame. For the original image frames, we partition each frame into 4 µm by 4 µm submatrices, and then calculate |max-min| for each submatrix. Next, we classify the submatrix as defect containing if this difference is above a threshold. By varying the threshold, we generate an ROC curve.
Fig. 5.5. Study of defect detectability with the 2DISC method using ROC curves. (a) The blue, red, green, and black curves are for the original amplitude single images, the 2nd-order difference with convolution (2DC), the full 2DISC panoramic image, and random guessing, respectively. (b) Corresponding ROC curves for the phase images. From (a) and (b), we find the 2DISC image without stitching, the false alarm rate is 0.4% for amplitude and 3% for phase at 90% detection probability; while for the full 2DISC panoramic amplitude and phase images, there is 100% detection without any false alarms.

The blue curves in Fig. 5.5(a) and (b) are the ROC curves for original amplitude and phase single frame images, respectively. On each figure, a random guess line is also plotted (black curve). The original image frames are similar to random guessing or sometimes even worse, possibly because the defect reduces the $|\text{max-min}|$ variation from the underlying pattern. For comparison, we calculate the ROC for the 2nd-order difference images with tripole convolution (2DC). These are shown in Fig. 5.5(a) and (b) as red curves. Both curves are now well above the random guessing line. We achieve 90% detection probability with a false alarm rate of 0.4% for the amplitude image and 3% for the phase image. The lower false alarm rate for the amplitude image is attributed to the difference in the way the image post-processing reduces additive and multiplicative noise and to the difference in the sensitivity of amplitude and phase to these types of noises. Finally, we studied the defect detection using the full 2DISC panoramic images; see the green curves
in Fig. 5.5(a) and (b). As expected (refer to Fig. 5.3(f) and (g)), we achieved 100% detection with no false alarms, which means we are able to accurately find the defect in the 70 µm by 27 µm panoramic image.

5.5 Detection results

![Defect detection results for different defect types using the 2DISC method.](image)

Fig. 5.6. Defect detection results for different defect types using the 2DISC method. Each sub-figure shows a 9 µm by 9 µm zoomed-in FOV phase image (top) and amplitude image (bottom). (a)-(c) Parallel bridge defects that are 160 nm by 55 nm, 160 nm by 35 nm, and 160 nm by 20 nm, respectively. (d)-(f) Perpendicular bridge defects that are 55 nm by 100 nm, 40 nm by 100 nm, and 20 nm by 100 nm, respectively. (g) Isolated dot defect that is 60 nm by 90 nm. (h) Perpendicular line extension defect that is 60 nm by 50 nm.
Fig. 5.7. SEM verification of the detected defects. (a)-(c) SEM image of parallel bridge defects which are 55 nm by 160 nm, 35 nm by 160 nm, and 20 nm by 160 nm, respectively. (d)-(f) SEM image of perpendicular bridge defects which are 55 nm by 100 nm, 40 nm by 100 nm, and 20 nm by 100 nm, respectively. (g) SEM image of isolated dot defect which is 90 nm by 60 nm. (h) SEM image of perpendicular line extension defect which is 60 nm by 50 nm. These SEM images correspond to the 2DISC images shown in Fig. 5.6.

The same technique used to generate Fig. 5.3(f) and (g) is also used for detection of other types of defects. For each 2DISC image, we show just a zoomed-in FOV (9 µm by 9 µm) around the defect in Fig. 5. The phase image is on the top of each sub-figure and the amplitude is on the bottom. For parallel bridge defects, we detected defects with sizes of 160 nm by 55 nm (Fig. 5.6(a)), 160 nm by 35 nm (Fig. 5.6(b)), and 160 nm by 20 nm (Fig. 5.6(c)). For perpendicular bridge defects, we detected defects with sizes of 55 nm by 100 nm (Fig. 5.6(d)), 40 nm by 100 nm (Fig. 5.6(e)), and 20 nm by 100 nm (Fig. 5.6(f)). We have detected 60 nm by 90 nm isolated dot defects (Fig. 5.6(g)) and 60 nm by 50 nm perpendicular line extension defects (Fig. 5.6(h)). Note that for the perpendicular bridge and line extension defects, the incident linear
polarization was rotated parallel to the defect, i.e. perpendicular to the lines in the underlying structure. This maximizes the defect signal and minimizes the signal from the underlying structure. For all other types of defects, the polarization was parallel to the lines. In this way, we achieve the best detection for each type of defect.

After collecting and processing all the data using the 2DISC method, the sample is imaged with the SEM for verification of the defect detection locations. We verified all the detected defects discussed in Fig. 5.6. The SEM images of the corresponding 2DISC images are shown in Fig. 5.7. Figure 5.7(a)-(c) show parallel bridge defects of 160 nm by 55 nm, 160 nm by 35 nm, and 160 nm by 20 nm sizes, respectively. Figure 5.7(d)-(f) show perpendicular bridge defects of 55 nm by 100 nm, 40 nm by 100 nm, and 20 nm by 100 nm sizes, respectively. Figure 5.7(g) shows the isolated dot defect with 60 nm by 90 nm size. Figure 5.7(h) shows the perpendicular line extension defect with 60 nm by 50 nm size. These SEM images show that the 2DISC method has successfully detected the defects in the pattern.

5.6 Summary
In summary, we developed a non-destructive defect inspection tool for 22 nm node wafer inspection. This tool has a large FOV using epi-DPM for image data collection and 2DISC for image post-processing. Epi-DPM is a common-path interferometer which is insensitive to mechanical vibration noise. The measured interferograms from epi-DPM are used to retrieve the phase and amplitude images from the sample. Using the 2DISC method, panoramic phase and amplitude defect detection images are created, where significant defect signal contrast is achieved. The defect detectability is studied using ROC curves. This inspection method has successfully detected different types of defects with sizes down to 20 nm by 100 nm due to significant suppression of noise. To detect even smaller defects, laser power stabilization and system installation in a cleanroom environment should be implemented.
CHAPTER 6: 9 NM NODE WAFER INSPECTION

6.1 Introduction

The semiconductor industry will launch the 14 nm node FinFET computer chips by the end of 2014, which is expected to significantly improve the current 22nm node predecessors, especially in power consumption. Continuing efforts are being made on lithography [58], and denser patterns with node sizes shrinking down to 10 nm will be possible around 2016, according to the International Technology Roadmap for Semiconductors (ITRS). This is posing critical challenges for optical defect inspection, see the discussion in Section 3.3.1.

Recently, considerable effort has been devoted to developing UV, deep-UV, or x-ray generated from high-order harmonic laser sources [55, 57, 149], to improve the optical defect inspection system resolution. However, the ultimate limiting factor for defect detection is the system noise and the sparsity of the pattern, not the diffraction limited resolution of the microscope [68-70, 150]. In the past, we developed a very low noise and highly sensitive optical metrology tool, which is based on a 532 nm laser epi-DPM system, see Section 4.1. Using this system, we successfully demonstrated detection of different types of defects down to 20 nm wide by 100 nm long or similar in a 22 nm node IDA wafer [27]. We retrieved both the phase and amplitude of the reflected light from the wafer surface and used the 2DISC image post-processing method to extract the wafer defect signal. The detailed description of the method is presented in Section 5.3.

In 2013, we adapted the epi-DPM system to inspect a 9 nm node IDA wafer. This wafer has defect sizes that are only 10 nm wide (40% the linewidth of the 22 nm node IDA wafer). Further, the pattern is more than 2x denser than the 22 nm node IDA wafer [149]. In order to detect defects in this 9 nm node wafer, the system’s sensitivity needed to be enhanced tremendously. Thus, we replaced our 532 nm solid-state
laser with a 405 nm diode laser which has a 10x better power stability (refer to Section 4.1.3), and inserted a 405 nm narrow-band filter in front of the camera. With the 2DISC method, we detected parallel bridge defects in the 9 nm node wafer in both phase and amplitude images at the optimum polarization and focus.

White-light based interferometric systems are more sensitive compared with lasers, which may also be used to further break the defect detection limit. At the end of 2013, we decided to develop a white-light interferometric inspection system called epi-wDPM. A detailed description of this system is presented in Section 4.2. In this chapter, we first show the 9 nm node wafer defect inspection results using a white-light bright-field system with in-plane and vertical through focus sample scanning. In bright-field, only intensity or amplitude images are collected. The in-plane scanning is used to produce a 2DISC image and the through focus scanning is used to produce a volume image. These images have shown significantly improved detection sensitivity which is a check point for the epi-wDPM system. Then, we spent some time to build the epi-wDPM system and test its performance. Recently, we started to use epi-wDPM to measure both the phase and amplitude images and repeated the measurements done with the bright-field system. Currently, the system can retrieve the phase and amplitude images in the 22 nm node IDA wafer with reasonable quality, but it is unable to detect defects in it. We are currently working on identifying the problems and solving them so that we can further deploy this system for 9 nm node IDA wafer inspection.

In Section 6.5, we propose dark-field inspection using selective Fourier space filtering with an SLM. We have carried Fourier spectrum simulations on the 22 nm node IDA wafer and the 9 nm node IDA wafer. The simulation results can guide us on designing the SLM based filter. We also simulated optical images and compared them with our experimental results. This comparison will be useful for studying the noise and estimating our detection detectability limit. For preliminary demonstration, we used the 405 nm laser epi-DPM system with an inverse filter to high pass the image at the Fourier plane to remove laser speckle noise. In this way, the CCD camera’s full dynamic range can be used for the wafer signal. Our results from
measuring a parallel bridge defect shows about 50% sensitivity improvement compared to the bright field. In addition to the Fourier space filtering, the SLM can be also used as a spectrum filter in the epi-wDPM system to select the desired spectrum from the broadband signal beam order. This SLM spectrum filtering method is demonstrated experimentally in Section 6.5 and it agrees well with the results that uses a physical spectrum filter.

6.2 9 nm node IDA wafer pattern

The 9 nm node IDA wafer design consists of one type of line structure which has a size of 9 nm by 270 nm. However, the actually width of line could be 15 nm due to imperfections in the fabrication process. The rectangular unit cell is 360 nm by 240 nm and contains two lines. It is repeated in a rhombic lattice pattern to form a two-dimensional array. Different types of intentional defects are printed to the wafer pattern for inspection testing. In Fig. 6.1, we simulate the wafer pattern with different defects. The pattern’s dimensional information is obtained from reference [149] and we also verified it with SEM. Figure 6.1(a)-(c) illustrates a parallel bridge defect, a perpendicular bridge defect, and an isolated dot defect. In Section 6.6, we use these simulated wafer patterns for dark-field inspection simulation study.

Fig. 6.1. Illustration of the 9 nm node IDA wafer structure and its defect. (a) A parallel bridge defect. (b) A perpendicular bridge defect. (c) An isolated defect. The unit cell is 360 nm in the horizontal direction and 240 nm in the vertical direction.
6.3 405 nm laser inspection using 2DISC method

Fig. 6.2. 2DISC phase and amplitude images using 405 nm laser epi-DPM for detecting a parallel bridge defect. (a) The panoramic 2nd-order difference phase image. (b) The image after convolution with the tripole pattern. (c) The panoramic 2nd-order difference amplitude image. (b) The image after convolution with the tripole pattern. The defect is marked by a red rectangle.

The 405 nm laser has better power stability compared with our previous 532 nm laser. During the inspection process, we first collect a sequence of interferograms in the wafer, by translating the wafer in-plane along the line structure, with an adjacent frame translation step of 0.75 µm. Then, a phase and an amplitude image are retrieved for each interferogram. From a single phase or amplitude image, it is impossible to discern the defect due to the residual noise. Therefore, we produce the 2DISC image post-processed from the 2nd-order difference scan images to improve the defect detectability. To demonstrate our detection, we inspect a parallel bridge defect array with a 15 nm by 90 nm designed size defect. Fig. 6.2(a) is the panoramic 2nd-order difference phase image. In this image, we can see the intentional defect at the center of the image. Due to the 2nd-order difference process, the defect signal will have a tripole pattern.
as: 1 -2 1, with a separation of 0.75 µm. Thus, in order to extract the defect signal, we convolve the image with a matched tripole pattern, which takes into account the point spread function of the system. After the convolution, we obtain the 2DISC detection image in Fig. 6.2(b) where we see improved defect signal contrast. We also calculated the panoramic amplitude image and the 2DISC amplitude image, shown in Figs. 6.2(c) and (d), respectively.

Fig. 6.3. SEM images in a parallel bridge array. (a) SEM image with the defect in the center enclosed by a red box. (b) A zoomed-in SEM image of the defect. The dimension of the pattern is also shown in this image.
After the optical inspection of the defect, we verify the detection with SEM. Figure 6.3 shows the SEM images taken from the same defect array which confirms our defect detection. Figure 6.3(a) shows an SEM image of the parallel bridge defect array where the defect is in the center, enclosed by a red box, and Fig. 6.3(b) is a zoomed-in SEM image of the defect. The dimension of the pattern is also shown in this image. The dimension of the pattern is also measured and shown in this zoomed-in SEM image. From the SEM image, we found that the actual intentional defect size is similar to the designed size, but the two lines it bridges are slightly larger than normal. We have also detected defects in other parallel defects arrays with 2DISC images and confirmed the results with SEM.

6.4 White-light bright-field inspection

Here we present white-light bright-field wafer defect inspection which only produces intensity or amplitude images. We inspect the same defect arrays as we did with the 405 nm laser epi-DPM system, these results are also verified with SEM. Both in-plane and vertical through focus wafer scans are performed for image post-processing to detect the defects. The in-plane (in x direction) scan is used to produce 2nd-order difference amplitude images for 2DISC post-processing and the vertical through focus (in z direction) scan is used to produce a volume image. Both methods can detect the defects; thus, can be used together to provide the defect information. In this section we show the results.

6.4.1 Inspection with 2DISC processing

On the white-light bright-field microscopy system, we scan the wafer horizontally with 0.5 micron step size and produce 2nd-order difference images with 1 micron separation to remove the common image background. In Fig. 6.4, two sample 2nd-order difference images are presented on the top. Due to the high image contrast in white light, the defects clearly show up in both images. To further remove the background noise, we stitch the 2nd-order difference images to make a panoramic image, as shown in Fig. 6.4(c). Then, we convolve the image with a matched tripole pattern to make the 2DISC image in Fig. 6.4(d). The
background in Fig. 6.4(c) and (d) becomes smoother compared with Fig. 6.4(a) and (b), indicating improved noise cancellation. However, in the 2DISC images, there are still vertical line structures which are due to the e-beam patterning error, which is confirmed with SEM measurement.

![Fig. 6.4. 2DISC amplitude image using the white-light bright-field microscopy system. (a)-(b) Sample 2nd-order difference amplitude images. (c) The panoramic 2nd-order difference amplitude phase image. (d) The 2DISC image obtained with convolution on the 2DISC image. The defect is marked by a black rectangle.](image)

### 6.4.2 Inspection with 3D scanning

In addition to scanning in x or y, we can also perform scanning in z on the same defect to improve the detectability by producing a volumetric wafer image. The through-focus scanning method was previously demonstrated by Barnes *et al.* at NIST, where they used a 193 nm deep UV light source in scatterfield microscopy [149]. For our z-scan wafer defect inspection experimental demonstration, we use a white-light illumination that covers the spectrum from 400 nm to 1100 nm. This broadband source is free of laser
speckle which greatly enhances the image quality. In the parallel bridge defect array, we first collect three z-stacks around the center where the defect is located. The z-step size for each stack is 50 nm, and the horizontal (in x) separation between the stacks is 1 µm. At each z step, we produced a 2nd-order difference image in x by using the corresponding images from the three stacks. At the end, a volumetric image is formed. To visualize the volume image, we show cross-section images cut at the defect center position (defined as the origin of the coordinate) at three different planes, i.e. z = 0 plane in Fig. 6.5(a), x = 0 plane in Fig. 6.5(b), and y = 0 plane in Fig. 6.5(c), respectively. The defect produces a unique and predominant feature in Fig. 6.5(b) and (c). Thus, a pattern recognition algorithm can be used to extract the defect from the wafer’s underlying structure. We also explore alternative ways to visualize the volume image in 3D. First, we partition the volumetric image into 4 µm by 2 µm by 0.1 µm sub-volumes. Second, we calculate the variance of each sub-volume. Third, we map the variance values back in 3D. To visualize the processed volumetric image in 3D, we present each sub-volume with a disk with its strength indicated by the diameter and color. The final 3D image is shown in Fig. 6.5(d), where the defect is clearly observed at the center in the 3D volume image.

Fig. 6.5 2nd-order x-difference z-scan detection results. (a) z = 0 plane image; (b) y = 0 plane image; (c) x = 0 plane image; (d) 3D visualization of the sub-volume processed image. In (a)-(c), the defect is marked by a red circle, and in (d) the defect is marked by a red arrow.
In Fig. 6.6, we show another 3D visualization method, developed by our undergraduate researcher Casey Bryniarski, using ParaView (a software specializes in 3D visualization). The left shows a transparent view of the volume image where the defect tripole pattern is clearly presented. The right-hand side shows three image slices at different z-scan locations with defect pattern remains. From these images, we can better visualize the defect and also determined its location in three dimensions.

6.5 White-light interferometric inspection

Here we show the wafer defect inspection results using the recently built epi-wDPM system. We are using this system to inspect both the 22 nm node IDA wafer and the 9 nm node IDA wafer. After a month of testing, we are still not able to obtain the expected detection results. The retrieved phase images are noisy.
and have very poor contrast. There are two key issues that account for this: the power of the CCD camera is too weak and an amplitude modulation coming from the diffraction grating dominates the wafer signal.

In order to make the epi-wDPM system work properly, two low-pass filters are used to both make a uniform reference beam (on the SLM filter, refer to Fig. 4.9) and a reasonable image contrast (on the aperture diaphragm, refer to Fig. 4.7). Two linear polarizers, which have partial transmission ~ 70%, are also used to make the SLM to work. Although we have tried to balance the tradeoff between adequate power for the interfering beams with good image contrast and spatial coherence, this has not helped to improve the detection sensitivity. In the 405 nm laser epi-DPM inspection system, the camera was typically set at 100 ms exposure and 0 contrast gain (Hamamatsu C4742-80-12AG). However, in epi-wDPM, we have to set the camera at 300 ms exposure and 100 contrast gain. The other key issue is the diffraction grating. Ideally, a pure phase grating should be used for creating diffraction orders for the interference. However, the approximate 110 lpm diffraction phase grating we need is not accessible in the market. Thus, we have to use a 110 lpm amplitude grating from Edmund Optics. This amplitude grating creates a periodic intensity modulation on top of the wafer’s signal on the CCD camera. If the reflected signal of a sample dominates the amplitude modulation (for example we measure large-scale structures), we can still retrieve the phase image. In our case, the wafer has about 20 nm feature size (consider the Raleigh scattering intensity that scales with volume squared), therefore there is low signal contrast in the interferogram. In addition to amplitude modulation, the dust particle contamination of the grating is also notorious. The grating is in the intermediate image plane; thus, a low coherence light source cannot mitigate this issue. In the future, a dedicated dust cover should be used to protect the grating surface from attracting dust particles.
Fig. 6.7. Spectroscopic filtering using an SLM in the epi-wDPM system. (a) The interference beam orders immediately after the SLM without any filtering of the 1st-order beam. (b) The left lobe of the 1st-order diffraction order remains. (c) The right lobe of the 1st-order diffraction order remains.

After attempting many methods (almost every system modification or combination we can do in epi-wDPM), we found that the grating amplitude modulation issue can be solved with selective dark-field filtering. We can perform the dark-field filtering on the signal beam order (1st-diffraction order) with the SLM. This dark-field filtering is equivalent to spectroscopic filtering if done on the 1st-diffraction order, since this signal order’s colors are spread out in the Fourier plane on the SLM. According to the diffraction grating equation, the smaller the wavelength, the smaller the separation between the 1st- and the 0th-order beams. The white-light LED spectrum starts with a main peak at 450 nm, and then a plateau around 550 nm (refer to Fig. 4.8, this spectrum is probably due to the different color LEDs). Thus, on the SLM plane, the 1st-diffraction order will appear as two lobes. We use a lens to relay the SLM to the camera plane to measure the intensity distribution in Fig. 6.7(a). Notice that the 0th-order beam is filtered with a 200 µm pinhole to serve as the reference beam. In this image, we cleanly observe two lobes in the 1st-order. The left lobe corresponds to the 550 nm plateau region of the spectrum and the right lobe corresponds to the 450 nm peak region of the spectrum of the LED. We can either select the left lobe as in Fig. 6.7(b) or the right lobe as in Fig. 6.7(c). Actually, the selection of the spectrum depends on the specific application; we can actually perform spectroscopic epi-wDPM electronically.

When imaging the 22 nm node IDA, the best image contrast is achieved with a 450 nm spectrum band, realized with the spectroscopic filtering in Fig. 6.7(c). A retrieved phase image of this wafer’s underlying
pattern is presented in Fig. 6.8(a). (Fig. 5.1 and Fig. 5.2 provide this wafer’s structure and the 532 nm laser epi-wDPM image information.) The wafer’s underlying periodic structure is recovered in this image. To see the improvement of using this spectroscopic filtering, we also present in Fig. 6.8(b) a retrieved phase image of this wafer pattern without spectroscopic filtering. In this image, the phase is partially and poorly retrieved across the field of view. It is worth mentioning that the SLM spectroscopic filtering is identical to using a physical spectrum filter, which we have experimentally demonstrated and it matches the result in Fig. 6.8(a) when using a 450 nm band-pass filter.

![Image](image.png)

Fig. 6.8. 22 nm node wafer imaging with spectroscopic filtering epi-wDPM system. (a) The retrieved phase image with spectroscopic filtering. (b) The retrieved phase image without spectroscopic filtering.

After improving the signal contrast in the phase image, we perform defect inspection with transverse wafer scanning with a stepping size of 0.75 µm to match the period of the wafer’s structure. In Fig. 6.9(a), we show the stitched 2\textsuperscript{nd}-order difference phase image obtained from a 20 nm by 160 nm parallel bridge defect array. Unfortunately, the intentional defect at the center of the defect array is not detected in this phase image. For a sanity test, we measured the same defect array with bright-field imaging (remove the grating) in Fig. 6.9(b). The intentional defect now shows up, as indicated by the red arrow, indicating that there is actually a defect. In these two images, the wafer’s underlying structure still shows up even though
we have performed period stepping. This might be due to errors in stepping size or direction. With no success in the 22 nm node IDA wafer, we thought there might be a possibility of success in the 9 nm node IDA wafer. But the results were worse; we cannot even retrieve the underlying structure with good quality.

Fig. 6.9. 22 nm node IDA wafer defect inspection. (a) The stitched 2nd-order difference phase image with spectroscopic filtering in epi-wDPM. (b) The stitched 2nd-order difference amplitude image with bright-field imaging.

To date, we have not been able to come up with an experimentally verified reason for the failure of using the epi-wDPM system for the 22 nm and 9 nm node wafer inspection. We think that polarization is the most likely reason. The polarizer of the defect signal upon reflection is scrambled, but the SLM system can only collect a linear polarization. This results in most of the defect signal not being collected, which is not the case for the laser epi-DPM system. Another possible reason is the system’s mechanical stability. We see very frequent vibration of the wafer images on the CCD camera. This vibration transfers to imprecision of the system, and the key of wafer inspection is precision. The last possible reason is wafer contamination.
by dust particles. All of the wafers we inspected have been exposed in non-cleanroom labs for more than one year. There are many dust particles in each defect array, thus preventing us from improving the inspection system’s sensitivity. In the near future, we will first investigate the polarization issue, since this is the only issue we can possibly solve.

**6.6 Dark-field inspection**

Normally, dark-field imaging just blocks the unscattered light or the low spatial frequency with a disk to enhance the wafer pattern edge contrast. In the conventional dark-field imaging system, the light source, usually a laser, is focused onto the sample, which could damage the sample. It also has a small field of view. Instead of blocking the low-frequency light, it would be more beneficial if we block the light coming from the wafer underlying pattern. In this way, we can still preserve the bright-field while improving the defect signal contrast on the camera. Thus, this requires selective filtering in the Fourier plane. Using a filter mask in the Fourier plane of our inspection system is not practical due to the difficulties in alignment, and also the inability to adaptively adjust the pattern for the best performance. Thus, we propose to perform selective filtering using an SLM. In our epi-wDPM system, we already have an SLM used to filter one of the signal beams down to the reference. If we make a filter pattern on the signal beam order, then we can realize selective filtering. In this section, we first perform numerical simulations in order to find the optimum filter design for different wafer structures and defects. Then, we show preliminary dark-field inspection results by using an inverse physical pinhole (a small opaque disk). It should point out that the dark-field filtering of the underlying structure can be performed only when using the 0\(^{th}\)-order beam as the signal, but this is not what we do in the current epi-wDPM system.
6.6.1 Wafer simulation

Fig. 6.10. Dark-field inspection simulation in the 22 nm node IDA wafer for a 34 nm by 160 nm parallel bridge defect. (a) The Fourier spectrum of the defect-free array. (b) The Fourier spectrum of the parallel bridge defect array. (c) The Fourier spectrum difference which is the defect signal spectrum. The plots are in logarithm scale. (d) and (e) are the simulated 2nd-order difference image and tripole convolution image.

In this section, we show the simulation results for dark-field inspection. The simulations will help us to design our dark-filed SLM filter in the future. We first simulate the Fourier plane spectrum signal for a 22 nm node wafer under a numerical aperture NA = 0.9 objective. This wafer has a 0.8 µm x 0.8 µm unit cell arrayed in a rhombic pattern, see the SEM images in Fig. 5.1. The Fourier spectrum of a defect-free array is plotted on a log scale in Fig. 6.10(a). The Fourier spectrum of a 34 nm by 160 nm parallel bridge defect array is plotted in Fig. 6.10(b). The spectrum difference, i.e., the defect signal, is computed on a linear scale and then converted to a log scale and plotted in Fig. 6.10(c). As we can see from Fig. 6.10(a) and (b), most of the wafer underlying structure is in a few bright dot regions corresponding to the translational symmetries.
of the wafer structure whereas the defect signal covers a broad region about the center. Thus, we can design our filter to just block those dot regions and maintain the defect signal. Note that there will also be a strong spectrum spot near the origin which is due the direct reflection from the wafer and the laser speckle noise. Thus, the SLM filter needs to also block the central region spot.

To study the noise and underlying pattern’s effect on detection sensitivity, we also simulate the 2\textsuperscript{nd}-order difference images and tripole convolution image, i.e., 2DISC image. The images simulated assume no noise, but we define a parameter called peak-signal-to-noise-ratio (PSNR) to address the noise factor, which is defined as:

$$PSNR = 20 \times \log_{10} \left( \frac{|D_{\text{max}} - D_{\text{min}}|}{\sigma_n + \sigma_i} \right),$$

(6.1)

where $|D_{\text{max}} - D_{\text{min}}|$ is the defect contrast in the 4 µm by 4 µm defect region, $\sigma_i$ is the standard deviation of the image, and $\sigma_n$ is the standard deviation of the noise. $|D_{\text{max}} - D_{\text{min}}|$ and $\sigma_i$ can be extracted from the simulation but $\sigma_n$ can be obtained from experimental results. Note that our simulation does not include polarization effects, thus we lowered the experimental value of PSNR to obtain reasonable $\sigma_n$ values for the 2\textsuperscript{nd}-order difference image and tripole convolution image. For the 2\textsuperscript{nd}-order difference image in Fig. 6.10(d), the ratio of $\sigma_o/\sigma_i$ is 0.62 and the PSNR is 11 dB. And for the tripole convolution image in Fig. 6.10(e), the ratio of $\sigma_o/\sigma_i$ is 0.34 and the PSNR is 22 dB. As demonstrated experimentally, when PSNR > 15 dB, we have good detectability. Thus, the simulation shows that we can detect defects very well in the tripole convolution image. However, if we can also use our dark-field filter to block the wafer’s underlying structure, i.e., making $\sigma_i \approx 0$, then even the 2\textsuperscript{nd}-order difference image without image stitching or convolution may be good enough to detect defects in the 22 nm node wafer.
Fig. 6.11. Dark-field inspection simulation in the 9 nm node IDA wafer for a 15 nm by 90 nm parallel bridge defect. (a) The Fourier spectrum of the defect-free array. (b) The Fourier spectrum of the parallel bridge defect array. (c) The Fourier spectrum difference which is the defect signal spectrum. The plots are in logarithm scale. (d) and (e) are the simulated 2nd-order difference image and tripole convolution image.

As the wafer line structure shrinks down from the 22 nm to 9 nm node, $\sigma_i$ becomes an order of magnitude smaller. This is much smaller than $\sigma_n$, the noise coming from the microscope system, which then becomes the limiting factor. In Fig. 6.11, we show a similar simulation for the 9 nm node wafer with geometry defined in Fig. 6.1. Figures 6.11(a) and (b) are the Fourier spectra for the defect-free array pattern and the 15 nm by 90 nm bridge defect array pattern, respectively. From both images, we see that most of the underlying wafer structure’s periodic signal cannot be captured by the optical system due to the diffraction limit. Only small portions are left near the north and south poles. Figure 6.11(c) is the difference spectrum which is due to the defect. Figures 6.11(d) and (e) show the 2nd-order difference image and tripole convolution image. For the PSNR calculation, we use the $\sigma_n$ value from our previous 532 nm laser system.
As expected, the underlying structure and the defect signal become much weaker, i.e., $\sigma_i$ and $|P_{\max} - P_{\min}|$ are very small. Thus, the PSNR values are well below our detection limit. For the 2nd-order difference image, the ratio of $\sigma_n/\sigma_i$ is 46.5 and the PSNR is -12 dB, and for the tripole convolution image, the ratio of $\sigma_n/\sigma_i$ is 15.1 and the PSNR is only 4.7 dB. However, our 405 nm laser is 10x better in power stability compared with 532 nm laser. Thus, the $\sigma_n$ will be much smaller, and the PSNR can be above 15 dB for the tripole convolution image. This explains why we were able to detect defects experimentally in the 9 nm node wafer.

### 6.6.2 Inspection with an inverse-filter

For our preliminary experimental demonstration, we put an inverse pinhole filter in the Fourier plane of the image signal to filter out the laser speckle pattern. Here, we perform dark-field imaging to enhance the sensitivity of the 2DISC method for the 15 nm by 90 nm parallel bridge defect array. The inverse pinhole filter consists of a clear quartz slide with an opaque chrome region that has a 50 µm diameter. This filter is placed in a Fourier plane of the microscope and blocks the unscattered light, which thereby allows the camera’s full dynamic range to be used to measure the scattered signal. For simplicity, only intensity images are recorded and we use them to produce the 2DISC amplitude image. In this simplified configuration, no interferograms are formed and we do not recover phase images.

Figure 6.12(a) is the dark-field panoramic 2nd-order difference amplitude image, and Fig. 6.12(b) is the image after convolution with the tripole pattern, i.e., 2DISC image. In both images, we clearly see the intentional defect at the center. In addition, we also observed several other unintentional defects, indicating that this is a potentially more sensitive technique than the bright-field method. This inference will need to be verified with SEM in the future. For comparison, we also produced a bright-field image in Fig. 6.12(c) and (d), in which we could only clearly observe the center intentional defect and one unintentional defect indicated by the red arrow, but not the one indicated by the black arrow. To study the sensitivity
improvement in the 2DISC image, we define a quantity as: $\gamma = \frac{|D_{\text{max}} - D_{\text{min}}|_{\text{region1}}}{|D_{\text{max}} - D_{\text{min}}|_{\text{region2}}}$, which is the defect tripole pattern contrast ratio between the unintentional defect indicated by the black box and the central intentional defect. $D_{\text{max}}$ and $D_{\text{min}}$ are the maximum and minimum signal in a 4 µm by 4 µm rectangle around the defect. Region1 is marked by the black box and region2 is marked by the red box. The $\gamma$ value is found to be 0.81 for the dark field, whereas a value of 0.53 is found for the bright field. This means that in the dark-field image, the unintentional defect produces signal strength comparable to the intentional defect. However, in the bright-field image, the unintentional defect produces only half of the signal compared to the intentional defect. This indicates that we can potentially detect more defects in the dark-field image. Again, this needs to be verified by SEM in the future.

Fig. 6.12. Dark-field 2DISC amplitude image for detecting a parallel bridge defect. (a) The dark-field panoramic 2nd-order difference amplitude image. (b) Image (a) after convolution with the tripole pattern. (c) The bright panoramic 2nd-order difference amplitude image. (d) Image (c) after convolution with the tripole pattern. In the image, the defect is marked by a red rectangle, and in each image two unintentional defects are marked by a red arrow and a black arrow, respectively.
6.7 Summary
We have presented our progress on inspecting defects on a 9 nm node dense wafer. Several approaches have been proposed and implemented to successfully detect the defects, including using the 2DISC method with the 405 nm laser epi-DPM system, 3D wafer scanning with white light. Currently, we are working on implementing the epi-wDPM system for 9 nm node wafer inspection, and also adding selective filtering to it. However, there are several technical difficulties that we need to solve to really break the detection limit with this system.
CHAPTER 7: SUMMARY ON WAFER INSPECTION

The wafer defect inspection research is all driven by the semiconductor industry. In the past three years, we have been developing wafer metrology tools by specializing newly developed general optical imaging methods to wafer inspection.

In the previous chapters, we went through our wafer defect inspection work. We first discussed the current challenges and future trends in the wafer metrology industry. Then, we introduced the optical wafer defect inspection tools that we have developed, using laser and white-light interferometry, wafer scanning, and image post-processing. After the metrology tools development, we showed the results of their implementation for 22 nm node and 9 nm node patterned wafer defect inspection. Using the 532 nm laser based epi-DPM system, we detected defects down to a size of 20 nm by 100 nm in a 22 nm node wafer. For the 9 nm node wafer inspection, we used the 405 nm laser-based epi-DPM system and the white-light bright-field system. Combined with image post-processing, those two systems were able to detect 15 nm by 90 nm size defects in a densely patterned 9 nm node wafer.

To further improve the sensitivity of the system for 9 nm node wafer inspection, we have been working on developing an epi-wDPM white-light interferometry system. Now, the system is able to retrieve phase and amplitude images from the wafer. Also, an automated 3D scanning system integrated with image processing has been also built. This allows us to perform real-time wafer defect inspection while we scan the wafer. At the end, selective signal filtering was tested on the system with a projector-based SLM. In addition to building the white-light wafer inspection tool, we constructed a table-top cleanroom to host the system for the 9 nm node wafer inspection system.

There are still several issues we need to solve to really break the detection limit in the 9 nm node wafer using epi-wDPM. One of the issues relates to the polarization. The SLM filter needs to use two linear
polarizers which are preventing us from collecting the scattered defect signal with random polarization distributions. The second issue is the LED power. In order to perform white-light interferometry and selective filtering, we need to ensure enough power on the camera to capture low-noise images (mostly due to light shot noise that cannot be improved on the camera with a larger dynamic range). However, the interferometric system involves a lot of spatial filtering and transmission through optical surfaces (mostly due to the grating and polarizers) where a significant amount of power is lost. The third issue is the wafer contamination by dust particles from air. As the system was built in a lab and not a cleanroom environment, dust particles accumulate on the wafer, which finally deteriorates the sensitivity of the inspection. The last issue is the system’s mechanical stability. The method we use to mount the wafer in epi-wDPM is not very stable (especially the tip/tilt stage); thus, building vibrations can be easily coupled into the mounts. Note that the wafer inspection system is on the 4th floor of the Beckman Institute, which is only several yards away from the construction of the new ECE building. Mechanical vibrational noise needs to be very low for wafer inspection, since we are looking at nanoscale defects which require high precision. Similarly, in superresolution imaging (essentially super-precision localization), to achieve better than 10 nm resolution, a ground floor lab is very necessary. We are currently working to address those issues to optimize the detection system.

With the already-built 3D scanning system epi-wDPM, we can also work on reconstructing the 3D wafer structure, i.e., solving the scattering inverse problem in the wafer. If the scattering from the wafer is weak compared to the incident wave, the 1st-order Born approximation [151] and the Rytov approximation [152] can be used for the reconstruction. However, the Rytov approximation easily breaks down compared with the 1st-order Born approximation, since it assumes the total field correction only takes account of the phase due to the scatterer and the amplitude remains a constant [153]. For the wafer structure, the 1st-order Born approximation may not be valid since the perturbation due to the wafer’s underlying structure can be
relatively strong. In this case, an iterative procedure using the *distorted Born approximation*, which is equivalent to the Newton iteration method, can be used [154, 155]. For the current node process wafer, the pattern size is very small and the patterned wafer approximates an optically flat surface. Thus, the scattering is relatively weak and the 1st-order Born approximation may still work. With the 3D reconstruction of the wafer, buried defects can be revealed, which will be especially useful for high aspect ratio (HAR) structures such as through silicon via (TSV) and 3D layer structures.

Another future research direction is on developing methods to improve the image resolution and quality using digital image processing techniques. Since 2000, many digital image processing methods have been developed, and one of the most promising techniques is using sparse representations [31, 156]. For example, a sparse representation using the wavelet transform or the recently proposed nonsubsampled contourlet transform is widely used for denoising of images in the literature [157]. Thus, we can also develop a sparse representation based technique to process the images from our wafer inspection system. The processing will be physically inspired, using a physical model to be developed for 3D wafer reconstruction. Ultimately, we expect to reconstruct the sub-wavelength features and achieve super-resolution reconstruction of the wafer structure.

There are many other problems to solve in wafer defect inspection. One of the problems is on pattern referencing. For wafer defect inspection, we need a defect-free pattern to serve as a reference for detecting defects in the test pattern. When the wafer has self-repeating features, we can use a cross-correlation technique to eliminate the need of a reference pattern. This technique is described in the Appendix A.4. However, for non-periodic patterns, we will have to use a reference pattern which means we need to do measurements on both the reference and the test wafers, and then use the cross-correlation technique to find the defects. This means our system needs to be perfectly calibrated and the translation stages require precision and repeatability over large translations. With our newly built 3D translation stage, we can
perform experiments on pattern cross-referencing. Other problems include analyzing image stitching errors with the 2DISC method, and finding and sizing dust particles on the wafer.
CHAPTER 8: TOMOGRAPHIC IMAGING WITH LOW-COHERENCE LIGHT

Optical diffraction tomography (ODT) was first introduced by Wolf in 1969 [36]. In recent years, advances in CCD/CMOS cameras and computers have significantly advanced the research in digital holographic microscopy (DHM) [9, 13-15], a technique to measure the complex field emitting from the object, giving birth to a new research field called quantitative phase imaging (QPI). Combining ODT with QPI allows for revealing the 3D structure of the object. Over the past several years, this technique has gained popularity for 3D cell and tissue imaging for early stage disease detection [82-85]. In ODT, multiple far-field scattering plane measurements are performed which require well-defined plane waves. Thus, lasers are typically used. However, the sensitivity and resolution of the technique are limited by the laser speckle artifact. With the advent of optical coherence tomography (OCT) in the early 1990s [88], low-coherence light (LCI) gained popularity for tomographic imaging. LCI sources are laser speckle free; thus, offering high image contrast. In addition, due to their short temporal coherence lengths, strong depth-sectioning is achieved, making it ideal for tomographic imaging. In this chapter, we introduce the principle for tomographic imaging with LCI.

8.1 Low-coherence interferometry

LCI uses a broadband light source, such as a light emitting diode (LED), superluminescent diode (SLD), white-light lamp, or short-pulse laser, to form an interference signal. Low-coherence light is typically defined as light with a short temporal coherence length, i.e. on the order of tens of microns or even less. Due to this short coherence length, LCI methods provides better depth sectioning. Here, we show the theoretical basis for LCI.
We start the discussion by considering a typical Michelson interferometer with a broadband light source as shown in Fig. 8.1. The input beam is split into two arms as $U_1$ and $U_2$, where $U_2$ is delayed through a scanning mirror. On the detector plane, the two light fields are combined; thus, the instantaneous field is,

$$U(t) = U_1(t) + U_2(t + \tau),$$

where $\tau = 2\Delta d / c$ is the delay imposed by the mobile mirror and $\Delta d$ is the mirror translation distance. At the detector, the intensity of this total field is measured, thus,

$$I(\tau) = \langle |U(t)|^2 \rangle = I_1 + I_2 + \langle U_1(t) \cdot U_2^* (t + \tau) \rangle + \langle U_1^* (t) \cdot U_2 (t + \tau) \rangle.$$  \hspace{1cm} (8.2)

In this equation, the brackets indicate temporal averaging, $\langle f(t) \rangle \approx \int f(t) \, dt$, where the range of the integral is over the integration time of the detector. With the temporal cross-correlation between the reference and sample field defined as, $\Gamma_{12}(\tau) = \langle U_1(t) \cdot U_2^* (t + \tau) \rangle$, Eq. (8.2) is simplified to,

$$I(\tau) = I_1 + I_2 + 2 \text{Re} \left[ \Gamma_{12}(\tau) \right].$$  \hspace{1cm} (8.3)
Equation (8.3) implies that one can experimentally measure the real part of the temporal cross-correlation function \( \text{Re}[\Gamma_{12}(\tau)] \) by varying the temporal shift or the shift in the reference arm mirror. For a perfectly balanced interferometer, \( U_2(\omega) = U_1(\omega) \), the measurement reduces to,

\[
I(\tau) = 2I_i + 2\text{Re}[\Gamma(\tau)],
\]

where \( \Gamma(\tau) \) is the autocorrelation function of the field generated by the broadband source, which is a measurable quantity. By subtracting the DC component, \( 2I_i \), the measurement reduces to the real part of the autocorrelation function. The complex analytic signal associated with this measured signal can be recovered using the Hilbert transformation, which yields the imaginary part of the autocorrelation,

\[
\text{Im}\{\Gamma(\tau)\} = -\frac{1}{\pi} P\int\frac{\text{Re}[\Gamma(\tau')]}{\tau - \tau'} d\tau',
\]

where \( P \) indicates a principal value integral. The Wiener-Kintchin theorem shows that the temporal cross-correlation relates to the cross-spectral density \( W_{12}(\omega) \) through a Fourier transform,

\[
W_{12}(\omega) = \int \Gamma_{12}(\tau) e^{i\omega \tau} d\tau,
\]

and Eq. (8.6) can be expressed in terms of the temporal Fourier transform of the two balanced fields,

\[
W_{12}(\omega) = U_1(\omega) \cdot U_1^*(\omega).
\]

Thus, the cross-spectral density reduces to the spectrum of the input field, \( S(\omega) \). Therefore, this interferometric measurement recovers the full complex analytic signal associated with the spectrum.

Now, assuming that the scattered field is a modification of the reference field by the transfer function, \( h(\omega) \), thus, \( U_2(\omega) = U_1(\omega)h(\omega) \). Then the cross-spectral density becomes \( W_{12}(\omega) = S(\omega)h(\omega) \) and the temporal cross-correlation, \( \Gamma_{12}(\tau) \), can be expressed as a Fourier transform of \( W_{12} \).
\[ \Gamma_{12}(\tau) = \int W_{12}(\omega) \cdot e^{i\omega \tau} d\omega = \Gamma(\tau) \otimes h(\tau). \quad (8.8) \]

In the case of a \( \delta \) response function, \( h(\tau) = \delta(\tau - \tau_o) \), i.e. a mirror, the cross-correlation in Eq. (8.8) reduces to,

\[ \Gamma_{12}(\tau) = \Gamma(\tau - \tau_o). \quad (8.9) \]

where \( \tau_o = 2d_o/c \) is the time delay due to the mirror translation. Therefore, by scanning the reference mirror, the position of the sample mirror is measured with an accuracy given the width of the cross-correlation function. In other words, the depth resolution is determined by the envelope of \( \Gamma \), or the coherence length, \( l_c = c \tau_o/n \), where \( c \) is the speed of light, \( n \) the refractive index of the surrounding medium. 

Also, \( \tau_c \propto 1/\Delta\omega \) is the coherence time of the broadband source, where \( \Delta\omega \) is the bandwidth of the source. Furthermore, when the reference mirror and the sample mirror are more than a coherence length away from each other in path length, then the signal will be extinct. This effect is referred to as coherence gating.

OCT, a technique based on this LCI principle, is capable of performing depth-resolved imaging [88], which we will discuss in more detail in Section 8.2. Also, combining LCI with angle-resolved scattering, a technique called angle-resolved low-coherence interferometry (aLCI) was also developed and used for solving the inverse scattering problem to determine the scattering geometries. One of the main applications of aLCI is in measuring the size distribution of cellular structures in tissue in situ for cancer detection [158].

### 8.2 Optical coherence tomography (OCT)

OCT is a non-invasive 3D imaging technique that is similar to ultrasound imaging which uses the time of flight principle. However, light has a significantly higher speed than ultrasound; thus, it is not feasible to measure the time delay directly. Instead, OCT performs interferometric measurements with a low-coherence light source, i.e., utilizing the coherence gating effect of LCI. The first experimental demonstration of OCT in 1991 showed 3D imaging of the retina [88], and still, OCT performs an
irreplaceable role in ophthalmology such as retinal disease detection and glaucoma diagnostics [159-162].

OCT is also found in many other clinical applications, such as cardiology [163-166], oncology [167-169], dermatology [170-173], gastroenterology [174-178], and dentistry [179-181], where an imaging method that shows deeper into the tissue is needed.

Fig. 8.2. A typical OCT system based on a 2 by 2 fiber interferometer.

Based on LCI, a typical experimental setup for OCT uses a broadband light source as the illumination and 2 by 2 optical fiber interferometer system with the four fiber ends located at the source, detector, reference mirror, and the sample, respectively. Figure 8.2 shows a schematic design of a conventional OCT system. By scanning the delay of the reference mirror, the specimen’s axial dimension information is obtained on the detector through the LCI principle in Eq. (8.9). In other words, if the time delay of the signals from the reference arm and the sample arm are within the coherence time, the detector detects the interference signal. For example, considering a broadband source with a Gaussian spectrum profile, the coherent length is \( \sim 0.44\bar{\lambda}^2/\Delta \lambda \), where \( \bar{\lambda} \) is the mean wavelength and \( \Delta \lambda \) is the spectrum width [89]. For a broadband light source with \( \bar{\lambda} \) equal to 1064 nm and \( \Delta \lambda \) equal to 200 nm, the axial resolution is 2.5 \( \mu \text{m} \).

Due to the usage of fiber, this type of OCT images a single point at a time. Therefore, in order to obtain a 3D image, a 3D scan (combining A-scan and B-scan modes) needs to be performed. Most of the early works on OCT was based on the time-domain measurements, which rely on the axial scanning of the
reference mobile mirror. Further development of broadband sources and detectors has brought a few new OCT techniques, including spectral-domain OCT and swept-source OCT.

8.3 Interferometric synthetic aperture microscopy (ISAM)
In a time-domain OCT system, the focal depth is fixed and the depth scanning is done by scanning the reference arm. In this case, the image quality deteriorates away from the focal plane, thus, reducing the lateral resolution. Therefore, a large depth of focus, achieved by a low NA objective, is important to achieve a high-quality OCT image. However, a low NA objective provides poor lateral resolution, usually on the order of 10 µm. In 2006, Ralston et al. developed an interferometric synthetic aperture microscopy (ISAM) to solve the inverse scattering problem in OCT, which was used to computationally achieve depth independent images with high resolution in both lateral and axial dimensions [182]. ISAM, based on diffraction tomography theory, solves the inverse scattering problem with the Green’s function approach to establish a relation between the measured signal and the sample’s scattering potential. ISAM requires resampling of the raw data, a process analogous to the synthetic aperture radar (SAR) imaging, to achieve high-resolution 3D reconstruction. Since ISAM reconstruction is done computationally, it can be used for high-speed, high-3D resolution, in vivo noninvasive imaging of cellular morphology. Recently, the same group has developed fast computation methods for tomographic imaging of scattering tissues [183, 184].

8.4 Summary
In this chapter, we have introduced tomographic imaging using low-coherence light. We first provided the theory basics for tomographic imaging using LCI. Then, we discussed OCT, which is a tomographic imaging method based on LCI. At the end, we introduced ISAM which is an interferometric tomographic imaging method based on solving the inverse scattering problem.


CHAPTER 9: WHITE-LIGHT DIFFRACTION TOMOGRAPHY

9.1 Introduction

ODT uses the measured complex field to infer the structure of the object, i.e., solving the inverse scattering problem. In solving the inverse problem, Rytov and the 1st-order Born approximations are typically used.

In the Rytov approximation, the field perturbed by the object is assumed to be the product of the unperturbed field and a phase term $e^{i\phi(r)}$. Thus $U(r) = U_i(r) e^{i\phi(r)}$, where $U_i(r)$ is the incident field which is also the unperturbed system solution. Since the perturbation is only in the phase, the Rytov approximation is a better choice for reconstructing objects that are smooth and thick with respect to the wavelength of the light [152]. This limits it to low-resolution imaging. The 1st-order Born approximation, on the other hand, assumes the field inside the object is $U(r) = U_i(r) + U_s(r) \approx U_i(r)$. This field interacts with the object to act as a secondary source, which allows for direct solving of the scattered field. The Born approximation works better for reconstructing fine and thin structures (see, e.g., p. 485 in Ref. [153]). Cells, which are of interest to us, are thin and have fine structures, making the 1st-order Born approximation a better choice.

A schematic description of the 1st-order Born approximation is depicted in Fig. 9.1. The figure shows a plane wave’s wavefront that is perturbed by the object. This wavefront can be measured with a DHM system and we use SLIM as the approach. In developing white-light diffraction tomography (WDT) theory, we first solve for the forward scattered field $U_i$ in the wavevector space instead of using the traditional Green’s function and Weyl’s formula approach [36]. Then we extend it to white-light illumination for WDT theory. WDT is an imaging approach and it differs from laser-based projection tomography, which measures in the far zone at different scattering planes. The WDT experiments were carried with the recently developed SLIM system. In WDT, the depth dimension is reconstructed by scanning the object through the focus and acquiring a stack of phase-resolved images. Using WDT, we calculated the 3D PSF of the SLIM
system and achieved 350 nm transverse resolution and 900 nm axial resolution. With the 3D PSF, we reconstructed the 3D structures of live, unlabeled, red blood cells. The results agree well with confocal and scanning electron microscopy images.

Fig. 9.1. A schematic description of the 1st-order Born approximation.

9.2 Theory formulation

9.2.1 SLIM

In SLIM, we measure the intensity distribution due to the interference of the reference and the scattered field. The scattered field carries the phase information $\Delta \phi(r)$. Thus, we can write the intensity distribution as [185, 186]:

$$ I(r, \tau) = I_r + I_s(r) + 2|\Gamma_{12}(r; \tau)| \cos[\omega_0 \tau + \Delta \phi(r)], $$  \hspace{1cm} (9.1)

where $I_r$ and $I_s(r)$ denote the reference field and scattered field intensity, $\omega_0$ is the mean frequency of the light source, $\tau$ is the delay between the reference and scattered field, and $\Gamma_{12}(r; \tau)$ is temporal cross-correlation function which relates to the scattered and reference field as
\[
\Gamma_{12}(r, \tau) = \langle U_s(r, \tau) U_r^*(r, \tau + \tau) \rangle. \tag{9.2}
\]

We can also write \(\Gamma_{12}(r, \tau)\) as

\[
\Gamma_{12}(r, \tau) = |\Gamma_{12}(r, \tau)| e^{i \phi(r, \tau)}, \tag{9.3}
\]

where \(\phi(r, \tau)\) is the phase associated with the image field. In SLIM, using the four phase shifted interferograms we can measure \(\Delta \phi(r)\) as

\[
\Delta \phi(r) = \arg \left[ \frac{I(r; \tau_3) - I(r; \tau_1)}{I(r; \tau_0) - I(r; \tau_2)} \right]. \tag{9.4}
\]

Also, it is found that \(\Gamma_{12}(\tau)\) is actually evaluated near the origin or \(\tau = 0\) [185]. Thus, measuring the phase directly solves \(\Gamma_{12}(r, \tau)\). However, we should distinguish the scattered field phase \(\Delta \phi(r)\) with the image field phase \(\phi(r)\). They are related with the relation [15]

\[
\phi(r, \tau) = \tan^{-1} \left[ \frac{\beta(x, y) \sin (\Delta \phi(x, y))}{1 + \beta(x, y) \cos (\Delta \phi(x, y))} \right], \tag{9.5}
\]

where \(\beta(x, y) = |U_s/U_r|\). The generalized Wiener-Khintchine theorem [187, 188] allows us to relate \(\Gamma_{12}(r, \tau)\) to the cross-spectral density through Fourier transform of

\[
\Gamma_{12}(r, \tau) = \int_0^\infty W_{12}(r, \omega) e^{i \omega \tau} d\omega, \tag{9.6}
\]

where

\[
W_{12}(r, \omega) = \langle U_s(r, \omega) U_r^*(r, \omega) \rangle. \tag{9.7}
\]

During the phase shifting measurement, \(\Gamma_{12}(r, \tau)\) is evaluated at \(\tau = 0\), which allows us to write
\[ \Gamma_{12}(\mathbf{r}, \tau = 0) = \int_{0}^{\infty} \left( U_{s}(\mathbf{r}, \omega) U^{*}_{s}(\mathbf{r}, \omega) \right) d\omega. \tag{9.8} \]

We see from Eq. (9.8) that the unknown quantity is the scattered field, \( U_{s} \). In Section 9.3, we first derive an analytical solution for \( U_{s} \) and then establish the relationship between the measurable quantity, \( \Gamma_{12} \), and the 3D object structure of interest.

### 9.2.2 White-light diffraction tomography (WDT)

The inhomogeneous Helmholtz equation, which describes the field \( U \) in a medium with an index distribution of \( n(\mathbf{r}) \) is:

\[ \nabla^{2} U(\mathbf{r}, \omega) + n^{2}(\mathbf{r}, \omega) \beta_{0}^{2}(\omega) U(\mathbf{r}, \omega) = 0, \tag{9.9} \]

where \( \beta_{0}(\omega) = \omega / c \) is the wavenumber in vacuum. We re-arrange Eq. (9.9) as:

\[ \nabla^{2} U(\mathbf{r}, \omega) + \beta^{2}(\omega) U(\mathbf{r}, \omega) = -\chi(\mathbf{r}, \omega) \beta_{0}^{2}(\omega) U(\mathbf{r}, \omega), \tag{9.10} \]

where \( \beta(\omega) = \bar{n} \beta_{0}(\omega) \) (we assume non-dispersive objects), \( \bar{n} = \left\langle n(\mathbf{r}) \right\rangle_{r} \) is the spatially averaged refractive index, and \( \chi(\mathbf{r}, \omega) = n^{2}(\mathbf{r}, \omega) - \bar{n}^{2}(\omega) \) is the scattering potential. The total field \( U(\mathbf{r}, \omega) \) can be written as \( U(\mathbf{r}, \omega) = U_{i}(\mathbf{r}, \omega) + U_{s}(\mathbf{r}, \omega) \), where \( U_{i}(\mathbf{r}, \omega) \) is the incident wave. So \( U_{i}(\mathbf{r}, \omega) \) satisfies the homogeneous wave equation,

\[ \nabla^{2} U_{i}(\mathbf{r}, \omega) + \beta^{2}(\omega) U_{i}(\mathbf{r}, \omega) = 0. \tag{9.11} \]

Equation (9.11) has a plane-wave solution \( U_{i}(\mathbf{r}, \omega) = A(\omega) e^{i\beta(\omega)z} \) where \( A(\omega) \) is the spectral amplitude of the incident field. Subtracting Eq. (9.10) from Eq. (9.11) gives

\[ \nabla^{2} U_{s}(\mathbf{r}, \omega) + \beta^{2}(\omega) U_{s}(\mathbf{r}, \omega) = -\chi(\mathbf{r}) \beta_{0}^{2}(\omega) U(\mathbf{r}, \omega). \tag{9.12} \]
From Eq. (6.12) we can clearly see that the driving term of the right-hand side represents the interaction of the object scattering potential $\chi$ with the total field $U$. Under the first Born approximation, i.e., $|U_s(r,\omega)| \ll |U_i(r,\omega)|$, we can approximate $U(r,\omega)$ on the right-hand side as $U_i(r,\omega)$, allowing Eq. (9.12) to be re-written as,

$$\nabla^2 U_s(r,\omega) + \beta^2(\omega) U_s(r,\omega) = -\chi(r,\omega) \beta_0^2(\omega) A(\omega) e^{i\beta(\omega)z}.$$  

(9.13)

Instead of employing the traditional Green’s function approach and the angular spectrum representation (Weyl’s formula), we solve for the scattered field directly in the wavevector space, using the 3D Fourier transformation. We first perform a 3D Fourier transform of Eq. (9.13), which yields

$$[\beta^2(\omega) - k^2] U_s(k,\omega) = -\beta_0^2(\omega) A(\omega) \chi[k_\perp, k_z - \beta(\omega)].$$  

(9.14)

In Eq. (9.14), we used the shift theorem of Fourier transforms, namely, $\chi(r,\omega) e^{i\beta(\omega)z} \rightarrow \chi[k_\perp, k_z - \beta(\omega)]$, where the arrow indicates Fourier transformation. Note that, throughout the dissertation, we use the same symbol for a function and its Fourier transform but carry all the arguments explicitly, which clearly identifies the domain in which the function operates. For example, $\chi(k)$ is the Fourier transform of $\chi(r)$.

From Eq. (9.14), the scattered field $U_s$ is obtained immediately,

$$U_s(k,\omega) = -\beta_0^2(\omega) A(\omega) \frac{\chi[k_\perp, k_z - \beta(\omega)]}{\beta^2(\omega) - k^2 - k_\perp^2}.$$  

(9.15)

Next, we derive an expression for the field as a function of axial distance $z$, i.e., we arrange the terms such that a 1D inverse Fourier transform with respect to $k_z$ can be easily performed. Toward this end, we define a $k_\perp$-dependent propagation constant, $q = \sqrt{\beta^2(\omega) - k_\perp^2}$, and re-write Eq. (9.15) as:

$$U_s(k,\omega) = -\beta_0^2(\omega) A(\omega) \chi[k_\perp, k_z - \beta(\omega)] \frac{1}{2q} \left( \frac{1}{k_z - q} - \frac{1}{k_z + q} \right).$$  

(9.16)
Since our imaging experiment only measures the transmitted field or the forward scattered field, we can ignore the backscattered field, \( \frac{1}{k_z + q} \). We perform an inverse Fourier transform on Eq. (9.16) with respect to \( k_z \) in order to obtain the forward scattered field as a function of transverse wavevector, \( k_\perp \), axial distance, \( z \), and angular frequency, \( \omega \)

\[
U_s (k_\perp, z; \omega) = -\frac{\beta^2(\omega) A(\omega)}{2q} \mathcal{X}(k_\perp, z) e^{i\beta(\omega)z} \bigotimes e^{iqz}. \tag{9.17}
\]

In Eq. (9.17), \( \bigotimes \) indicates convolution along the \( z \)-dimension, \( f(z) \bigotimes e^{iqz} = \int_{-\infty}^{\infty} f(z') e^{iq(z-z')} dz' \). It can be easily seen that the convolution of a function with a complex exponential yields the Fourier transform of that function multiplied the complex exponential, which yields the simple result

\[
\mathcal{X}(k_\perp, z; \omega) e^{i\beta(\omega)z} \bigotimes e^{iqz} = e^{iqz} \mathcal{X}(k_\perp, q - \beta(\omega)).
\]

In order to insert the result of Eq. (9.17) into Eq. (9.8), we need to Fourier transform \( \Gamma_{12}(r, \tau) \) with respect to the transverse position vector, \( r_\perp = (x, y) \). Since \( U_r (r, \omega) \) is a plane wave propagating in the \( z \)-direction, Eq. (9.7) in \( k_\perp \)-space is

\[
W_{12} (k_\perp, z, \omega) = \langle U_s (k_\perp, z, \omega) U_r^* (z, \omega) \rangle. \tag{9.18}
\]

Using the solution of \( U_s \) from Eq. (9.17) and \( U_r (z, \omega) = A(\omega) e^{i\beta(\omega)z} \), we have

\[
W_{12} (k_\perp, z, \omega) = -\frac{\beta^2(\omega) S(\omega) e^{i[\beta(\omega)z]} }{2q} \mathcal{X}[k_\perp, q - \beta(\omega)], \tag{9.19}
\]

where \( S(\omega) = |A(\omega)|^2 \) is the power spectrum of the illumination field. Using Eq. (9.8) we obtain the temporal cross correlation at zero-delay as a function of the frequency integral.
\[
\Gamma_{12}(k_\perp, z, 0) = \frac{c}{2\pi^3} \frac{\beta^2 S(\omega) e^{ik_\parallel q} \chi_k[\chi, q - \beta(\omega)]}{2q} d\omega.
\] (9.20)

With the relation \( \beta(\omega) = \bar{n}^2 \omega = \bar{n} \omega / c \), we can change the integral from \( d\omega \) to \( d\beta \), \( S(\omega) \) to \( S(\beta / \bar{n}) \), such that Eq. (9.20) becomes

\[
\Gamma_{12}(k_\perp, z, 0) = -\frac{c}{2\pi^3} \int_0^\infty \frac{\beta^2 S(\beta / \bar{n}) e^{ik_\parallel q} \chi_k[\chi, q - \beta]}{2q} d\beta.
\] (9.21)

Experimentally, we measure \( S(\lambda) \). So to obtain the spectrum distribution for \( S(\beta / \bar{n}) \) from \( S(\lambda) \) we need to consider the Jacobian transformations, \( S(\beta / \bar{n}) \leftrightarrow \bar{n} S(\beta) / c \) and \( S(\beta) \leftrightarrow \lambda^2 S(\lambda) / 2\pi \). In order to evaluate the integral in Eq. (6.21), we define a new variable \( Q = q - \beta = \sqrt{\beta^2 - k_\perp^2} - \beta \), from which we have \( \beta = -\frac{Q^2 + k_\perp^2}{2Q} \) and \( q = \left(-\frac{Q}{2} + \frac{k_\perp^2}{2Q}\right) \). Substituting \( d\beta \) for \( dQ \), we need to consider the Jacobian

\[
\frac{d\beta}{dQ} = \left(-\frac{1}{2} + \frac{k_\perp^2}{2Q^2}\right),
\]

then Eq. (9.21) becomes

\[
\Gamma_{12}(k_\perp, z, 0) = \frac{1}{8\pi^3} \int_0^\infty \left(\frac{Q^2 + k_\perp^2}{Q^3}\right)^2 S\left(-\frac{Q^2 + k_\perp^2}{2Q}\right) \chi_k[\chi, Q] e^{iQ} dQ
\]

\[
= \frac{1}{8\pi^3} FT_Q^{-1}\left[\left(\frac{Q^2 + k_\perp^2}{Q^3}\right)^2 S\left(-\frac{Q^2 + k_\perp^2}{2Q}\right)\right] \odot z \chi(k_\perp, -z).
\] (9.22)

In Eq. (9.22), \( \Sigma \) is the function that incorporates all the details of the instrument response. Note that the \( k_\perp \) coverage of \( \Sigma \) is limited to a maximum value \( k_\perp^{\text{max}} = \beta_0 NA \), where \( NA \) is the numerical aperture of the objective. By measuring \( z \)-stacks in SLIM, we can reconstruct the object’s 3D distribution through deconvolution of Eq. (9.22). Alternatively, we can write Eq. (9.22) in the spatial frequency domain as a product, namely,
\[ \Gamma_{12}(k_x, Q, 0) = \Sigma(k_x, Q) \chi(k_z, Q), \quad (9.23) \]

where

\[ \Sigma(k_x, Q) = \frac{1}{8n^2} \frac{(Q^2 + k_\perp^2)^2}{Q^3} S \left( -\frac{Q^2 + k_\perp^2}{2Q} \right). \quad (9.24) \]

**9.3 3D PSF**

Equation (9.24) gives the Fourier transform of the point spread function (PSF) of the imaging system, which we call the coherent transfer function of the system. From this equation, we see that the 3D PSF relates to the spectrum of the white-light source. Therefore, by measuring the light source spectrum and filtering it according to the NA of the objective, we can obtain the coherent transfer function. The spectrum of the light source \( S(\lambda) \) is measured as a function of wavelength \( \lambda \) using a fiber optic spectrometer, see Fig. 9.2. However, in our coherent transfer function, the 3D frequency space grid is set to be \((k_x, k_y, Q)\), i.e., the \( z \) dimension is resampled by \( Q \). Thus, we need to perform a Jacobian transformation as introduced in Section 9.2 to obtain the spectrum in new variable \(-\left(Q^2 + k_\perp^2\right)/2Q\). After this transformation, the spectrum \( S \) is interpreted in the 3D space \((k_x, k_y, Q)\). Finally, we obtain the coherent transfer function in 3D. The function \( \Sigma(k_x, k_y, k_z) \) for our imaging system is illustrated in Fig. 9.3. As expected, the width of the \( k_z \) coverage increases with \( k_x \), which indicates that the sectioning is stronger for finer structures or, equivalently, higher scattering angles. The structure of the 3D is recovered through a sparse deconvolution algorithm [30, 33], see more details in our recent work in reference [34].
Fig. 9.2. Spectrum of the white-light source for SLIM, measured over the range of 340 nm to 1028 nm. The y-axis is in arbitrary units.

Fig. 9.3. The 3D coherent transfer function. (a) 3D rendering of the instrument transfer function, using the proposed WDT calculation. (b) Cross-section of the transfer function at $k_y = 0$ plane.
9.4 3D red blood cell imaging

We validated our WDT method using a polystyrene microbead sample, see the supplementary information in reference [34]. Then, we applied WDT to measure spiculated red blood cells, which are known as echinocytes. This morphological abnormality is well documented and can be an indication of transitory stress (e.g., osmotic stress) or a sign of a serious disease [95, 189]. We used this interesting 3D morphology as a test sample and employed a scanning electron microscopy (SEM) image and a confocal fluorescence microscopy image [190] as the control imaging methods (Fig. 9.4(a)). The sample was prepared as a blood smear on a glass slide, and phase images were measured using spatial light interference microscopy (SLIM). Unlike SEM and confocal microscopy, which need sample preparations, such as metal deposition and fluorescence labeling (Calcein and DiI in Fig. 9.4(a)), WDT is label-free and works without sample preparation. Furthermore, the irradiance at the sample plane in WDT is 6-7 orders of magnitude lower than in confocal microscopy [15], which provides a less harmful environment for the sample. The axial data (z-stack) was acquired with a step of 250 nm and a precision of 10 nm ensured by the piezoelectric nosepiece. With the Σ(x, y, z) function computed for a 40x/0.75NA objective, we performed the 3D deconvolution based on the sparsity constraint [30, 33]. Figure 9.4(b) shows a SLIM projection image of an echinocyte and its corresponding deconvolved image, along with a SEM image of a similar echinocyte. Sharper surface structures of this red blood cell are observed in the deconvolved image, as expected. Figure 9.4(c) shows the 3D rendering of the raw z-stack images as well as its corresponding deconvolution, as indicated (see also Supplementary Movie 1 in reference [191]). Again, the protrusions in the cell membrane show more details in the deconvolved image. This result is more clearly evidenced by investigating one slice from the tomogram, Fig. 9.4(c), which shows one slice where the empty space between spicules is revealed in greater detail after deconvolution.
Fig. 9.4. A spiculated red blood cell imaged using a 40x/0.75NA objective. The reconstruction uses a z-stack of 100 images each with 128 by 128 pixels, which takes about 5 minutes for sparse deconvolution. (a) Measured z-slice of a spiculated red blood cell and the corresponding deconvolution. An SEM image and a confocal of similar cells are shown for comparison. (b) 3D rendering of the raw data and the corresponding 3D deconvolution (Media 1 in reference [35]). (c) A measured z-slice and the corresponding deconvolution result, which show the empty space between spicules on the red blood cell. The scale bars represent 2 μm in space.
9.5 Summary

In conclusion, we have shown that by using spatially coherent and temporally incoherent light, 3D structure information can be retrieved unambiguously simply by scanning the focus through the object of the microscope. This type of reconstruction requires a new theoretical description of the interaction between a weakly scattering object and broadband light, which we presented here for the first time. In essence, our theory generalizes Wolf’s diffraction tomography to white light and correctly predicts that the sectioning capability is the result of the combined effect of coherence and high numerical aperture gating. As a result, we can calculate the imaging system’s response (point spread function) and quantify the resolution, which turns out to be 350 nm transversally and 890 nm axially. Solving scattering problems using a quantitative phase microscope, i.e., measuring at the image plane instead of the far-zone, is a powerful new concept that allows us to acquire light scattering data from completely transparent objects with high sensitivity and dynamic range. In practical terms, WDT renders 3D images of unlabeled cells in a traditional environment of an inverted microscope, allowing for cell imaging for an extended period of time. Therefore, we anticipate that WDT will become a standard imaging modality in cell biology, complementing established technologies such as confocal microscopy.
CHAPTER 10: INVERSE SCATTERING SOLUTIONS USING LOW-COHERENCE LIGHT

10.1 Introduction

Noninvasive measurement of three-dimensional (3D) refractive index distributions of biological tissues and cells are desired for early disease diagnoses and cancer detection. X-ray diffraction emerged in the 1940s was first used for analyzing crystal structures [76], and became widely used for measuring intracellular structures [53, 77]. In the 1960s, x-ray computed tomography (CT), based on absorption along the projection path, was developed [78]. However, it is well known that x-ray CT suffers from phase problem, as addressed in a review by Wolf [79]. The iterative phase retrieval method using the modulus of the Fourier transform (F.T) has been proposed, however, the uniqueness of solution was not guaranteed [80, 81]. Inspired by x-ray CT, in 1969 Wolf published the seminal paper on optical diffraction tomography (ODT) [36]. With recent advances in image sensors and computers, ODT has been successfully used for 3D reconstruction of transparent objects. Usually in an ODT experiment, one needs to either scan the incident beam or rotate the object to different angles and measure the scattered field in the far zone [82-85].

OCT, which is essentially spatially-resolved low-coherence interferometry (LCI), was demonstrated first in 1991 as a label-free 3D imaging method for deep tissue imaging [88]. OCT is an amplitude technique, which uses the narrow temporal cross-correlation function between the sample and reference fields to achieve axial sectioning, even with close-to-zero numerical aperture (NA) optics. Izatt et al. used the time derivatives of the phase (frequency shifts) to generate Doppler imaging [192]. The combination of OCT and ODT was not been demonstrated until much later [89, 90, 182]. Ralston et al. showed that using the phase information in OCT measurements allows for solving the scattering inverse problem, which in turn can be used for reconstructing the 3D distribution of the tissue scattering potential with spatially
invariant resolution [90, 182]. Villiger and Lasser presented image formation in OCT using a coherent transfer function approach [193].

Very recently, ODT was implemented with white-light illumination in high NA imaging, as discussed in Chapter 9. This approach, called white-light diffraction tomography (WDT), shares the coherence gating principle of OCT, while measuring the signal in transmission [27]. WDT can be realized in any quantitative phase imaging (QPI) system, such as spatial light interference microscopy (SLIM) or diffraction phase microscopy (DPM) [15, 18, 24, 30]. Unlike OCT, the numerical aperture of the objective is crucial to achieve sectioning in WDT. With this method, 3D imaging of unlabeled live cells with submicron resolution in all three directions has been demonstrated. Furthermore, WDT presents an efficient and accurate way of solving the wave equation in the wavevector space, rather than using the traditional Green’s function approach. In this chapter, we apply this wavevector space method to solve inverse scattering under the Fresnel approximation, in both the time-domain and the Fourier-domain. This inverse scattering solution can then be used for 3D reconstruction from OCT and LCI measurements, which typically deploy low NA systems suitable for the Fresnel approximation. Furthermore, the solution is physically interpreted on the Ewald sphere. We also provide analytical solutions for the time-domain reconstructions, which can be used for 3D imaging by scanning the optical path-length delay between the reference and sample field.

10.2 Inverse scattering solutions

For this formulation, we will start with the results from Section 9.2.2 (see more details in reference [34]) where the forward and backward scattered field can be written as:

\[
U_s(k, z; \omega) = \mathcal{F}_\omega \left( \frac{\beta_0^2(\omega)A(\omega)}{2q} \int \chi(k, z)e^{i\beta(\omega)z} \right) \otimes e^{izq}.
\]
[\chi(k_\perp, z)e^{i\beta(z)z}] \ast e^{iqz} = e^{iqz}\chi(k_\perp, q - \beta(\omega)].$$ Therefore, we can immediately evaluate Eq. (10.1) as

$$U_f(k_\perp, z, \omega) = -\frac{\beta_0^2(\omega)A(\omega)e^{iqz}}{2q} \chi(k_\perp, q - \beta(\omega)), \quad (10.2a)$$

$$U_b(k_\perp, z, \omega) = \frac{\beta_0^2(\omega)A(\omega)e^{-iqz}}{2q} \chi(k_\perp, -q - \beta(\omega)), \quad (10.2b)$$

where $U_f$ and $U_b$ are the forward and backward scattered field, respectively. Note that, although we assume a dispersionless medium, the scattered fields have strong dependence on optical frequency. In most interferometric microscopy experiments, the measured signal is the cross-correlation between the scattered and the reference field. Under the 1st-order Born approximation, where the reference field equals the incident field, the cross-spectral density as a function of $k_\perp$ and $z$ can be written as [191]

$$W(k_\perp, z, \omega) = U_f(k_\perp, z, \omega)U_r^*(\omega, z_R), \quad (10.3)$$

where $U_r(z_R, \omega) = A(\omega)e^{i\beta(z_R)z_R}$ and $z_R$ is the propagation distance of the reference beam with respect to $z = 0$. The coordinate system is fixed once the scattering potential is specifically defined. Thus, the general solutions to the inverse scattering are

$$\chi(k_\perp, q - \beta) = -\frac{2qW_f(k_\perp, z, \beta)}{\beta_0^2S(\beta)} e^{-i(q - \beta)z_R}, \quad (10.4a)$$

$$\chi(k_\perp, -q - \beta) = \frac{2qW_b(k_\perp, z, \beta)}{\beta_0^2S(\beta)} e^{i(q + \beta)z_R}, \quad (10.4b)$$

where $W_f$ and $W_b$ are the forward and backward measured scattered and $S(\beta) = |A(\beta)|^2$ is the power spectrum of the incident field. Note that in experiments we usually measure $S(\lambda)$ [34]. Therefore, in order
to obtain $S(\beta)$, we need to consider a Jacobian transformation, i.e., $S(\beta) = -\lambda^2 S(\lambda)/(2\pi \kappa)$. If $W_f(\lambda)$ and $W_b(\lambda)$ are measured, as in Fourier-domain OCT, the Jacobian is also necessary in order to transform $W$ to the $\beta$ domain.

There are several significant implications of Eq. (10.4). First, it establishes the 3D reconstruction model for ODT. As shown in the scattering potential $\chi$, the object sectioning is determined by $q - \beta$ in transmission and $-q - \beta$ in reflection. In other words, the 3D reconstruction can be achieved by scanning the incident plane wave angle, $q = \sqrt{\beta^2 - k_{\perp}^2}$, or scanning the frequency, $\beta$. Notice that, to numerically reconstruct the 3D object, a resampling of the axial frequency $k_z$ as $k_{z} = q - \beta = \sqrt{\beta^2 - k_{\perp}^2} - \beta$ is necessary. As demonstrated previously, regularization or sparse deconvolution methods can be used to obtain uniform object reconstruction [34, 90, 194]. In particular, Eq. (10.4b) can be used for high-resolution Fourier-domain OCT, without the far-field approximation used in ISAM [83]. Finally, Eq. (10.4) also describes the scattering measurement in aLCI for determining depth-resolved angular scattering from tissues, without the need for the Mie theory assumption (note that $k_{\perp}$ is proportional to the sine of the scattering angle) [195, 196].

10.3 The Fresnel regime solutions

In order to integrate the forward and backward scattered signal in Eq. (10.4) with respect to frequency $\beta$, and obtain the time-domain solution, we use the Fresnel approximation. For an imaging system under this approximation, using the Taylor expansion, we have

$$q = \sqrt{\beta^2 - k_{\perp}^2} \approx \beta \left(1 - \frac{k_{\perp}^2}{2\beta^2}\right), \quad \text{thus,} \quad \beta^2 / q \approx \beta \left(1 + \frac{k_{\perp}^2}{2\beta^2}\right).$$  \hspace{1cm} (10.5)

Therefore, the solutions to 3D object reconstruction using the forward and backward scattering become
\[ \chi(k, k^2/2\beta) \approx -2\pi^2 W(k, z, \beta) e^{-i\beta(z-2z_k)} e^{i k^2/2\beta}, \]  
\[ \chi(k, k^2/2\beta - 2\beta) \approx 2\pi^2 W(k, z, \beta) e^{-i\beta(z+2z_k)} e^{-i k^2/2\beta}. \]  

In the following, we provide a physical interpretation for Eq. (10.6). In elastic scattering, we have the dispersion relation \(|k| = |k| = \beta\), where \(k = k_x + k_y\) and \(k\), are the scattering and incident wavevector, respectively. Figure 10.1(a) shows a cross-section of the Ewald sphere of radius \(\beta\), which explains the sectioning effect. In the right triangle ABC, the lengths of the segments BC, AC and BD are related to the wavevectors as BC = |p|, AC = 2\beta, and BD = |k|, where p = k - k, is the momentum transfer wavevector. Using the similarities of triangles ABC and BCD, BD^2 = AD \cdot DC, meaning that 
\[ k^2 = |\beta + k_x||\beta - k_x| \approx 2\beta|\beta - k_x|, \]  
which implies 
\[ |\beta - k_x| = \frac{k^2}{2\beta}. \]  

|\beta - k_x| is the axial projection of p, \(p_x = k^2/2\beta\), whose spread determines the sectioning in z. Note that, this projection appears explicitly in Eq. (10.6a) through the scattering potential, \(\chi(k, k^2/2\beta)\). This is expected, as measuring the scattered field at a scattering wavevector provides a single spatial frequency of the object. Considering the wavevector geometry described in Fig. 10.1(b), we find that in backward scattering, the momentum transfer wavevector has a projection of \(p_x = -2\beta + k^2/2\beta\), which also appears in Eq. (10.6b). From the comparison of the two wavevector geometries, we see that backscattering has a better sectioning effect due to the longer axial projection coming from 2\beta.
10.4 Time-domain solutions

Next, we derive the time-domain tomographic reconstruction model in both transmission and reflection geometry. At $z = z_R$, the forward scattering interference signal $W_f$, described in Eq. (10.6a), becomes

$$W_f(k_z, z, \beta) \approx -\frac{1}{2\pi^2} \beta S(\beta) \chi(k_z, \frac{k_z^2}{2\beta}) e^{- \frac{k_z^2}{2\beta}}.$$  \hspace{1cm} (10.8)

In order to obtain the time-domain solution, we perform the inverse FT over $\Omega = k_z^2 / 2\beta$, which ensures proper resampling of $k_z$. As $\beta = k_z^2 / 2\Omega$, the Jacobian transformation relates the power spectrum $ S(\beta) $ and $ S(\Omega) $,

$$S(\beta) = S(\Omega) \frac{d\Omega}{d\beta} = -S(\Omega) \frac{k_z^2}{2\beta^2} = -\frac{2\Omega^2 S(\Omega)}{k_z^2}.$$  \hspace{1cm} (10.9)

Note that $\Omega$ has a unit of $m^{-1}$ and its conjugate variable is $c\tau$, with $c$ the speed of light in the medium and $\tau$ the time delay between the reference and object field. Using the new variable, Eq. (10.8) becomes

$$W_f(k_z, z, \Omega) \approx -\frac{1}{2\pi^2} \Omega S(\Omega) \chi(k_z, \Omega) e^{-i\Omega \tau}.$$  \hspace{1cm} (10.10)
Taking the inverse FT of Eq. (10.10) with respect to \( \Omega \), and using the FT property:

\[ \Omega S(\Omega) \leftrightarrow -i \partial \Gamma(c \tau) / \partial \tau, \]

we obtain

\[ \Gamma_0(k_z, c \tau) = -\frac{i}{2 \pi^2} \frac{\partial \Gamma_0(k_z, c \tau)}{\partial (c \tau)} \Theta(c \tau) \chi(k_z, c \tau - z), \quad (10.11) \]

where \( \Gamma_0 \) is the temporal auto-correlation function of the source field, defined as

\[ \Gamma_0(k_z, c \tau) = FT^{-1} \{ S(\Omega) \} = \Gamma_0(k_z, c \tau) e^{i \Delta \tau}. \quad (10.12) \]

For simplicity, here we assume a symmetric optical spectrum. Under the slowly varying envelope approximation for \( \Gamma_0 \),

\[ \frac{\partial \Gamma_0(k_z, c \tau)}{\partial (c \tau)} = i \Omega \Gamma_0(k_z, c \tau) + \frac{\partial}{\partial (c \tau)} \left[ \Gamma_0(k_z, c \tau) \right] e^{i \Delta \tau} \approx i \Omega \Gamma_0(k_z, c \tau). \quad (10.13) \]

With this approximation, the temporal cross-correlation function, simplifies to

\[ \Gamma_i(k_z, z, c \tau) \approx \frac{i}{2 \pi^2} \Omega \Gamma_0(k_z, c \tau) \Theta(c \tau) \chi(k_z, c \tau - z). \quad (10.14) \]

Equation (10.14) provides the reconstruction model for a transmission measurement with LCI. This is analogous to a transmission OCT. This equation indicates that the axial reconstruction can be obtained by simply scanning the delay \( c \tau \) or by translating the object in \( z \). As a result of the convolution operation, the axial resolution is given by the width of \( \Gamma_0 \), i.e., the temporal coherence length of the field. If the source \( \Gamma_0 \) is measured, deconvolution or regularization methods can be also used to improve the axial reconstruction of \( \chi \). In backscattering, for \( z = z_b \), Eq. (10.6b) becomes

\[ W_b(k_z, z, \beta) = \frac{1}{2 \pi^2} \beta S(\beta) \chi \left( k_z^2 \frac{k_z^2}{2 \beta} - 2 \beta \right) e^{\left( \frac{k_z^2}{2 \beta} - 2 \beta \right) z_b}. \quad (10.15) \]
In order to obtain the time-domain reconstruction model, let us define \( Q = k_{\perp}^2 / 2\beta - 2\beta \). Thus, 
\[
\beta = \left( -Q + \sqrt{Q^2 + 4k_{\perp}^2} \right) / 4 \text{ (the other solution for } \beta \text{ is negative, therefore unphysical). Again, considering the Jacobian transformation,}
\]
\[
S(\beta) = S(Q) \frac{dQ}{d\beta} = S(Q) \frac{4\sqrt{Q^2 + 4k_{\perp}^2}}{Q - \sqrt{Q^2 + 4k_{\perp}^2}},
\]
Eq. (10.16) becomes
\[
W_\omega(k_{\perp}, z, \beta) \approx -\frac{2}{n^2} \sqrt{Q^2 + 4k_{\perp}^2} S(Q) \chi(k_{\perp}, Q)e^{i\omega z}.
\]
When \( k_{\perp} \ll |Q| \), under the Fresnel approximation, we can use the Taylor expansion by keeping the linear and quadratic terms, \( \sqrt{Q^2 + 4k_{\perp}^2} \approx Q + 2k_{\perp}^2 / Q \). Further, under the Fraunhofer approximation, only the linear term is kept, \( Q + 2k_{\perp}^2 / Q \approx Q \). Thus, Eq. (20) becomes
\[
W_\omega(k_{\perp}, z, \beta) \approx \frac{2}{n^2} Q S(Q) \chi(k_{\perp}, Q)e^{i\omega z}.
\]
This equation resembles Eq. (10.10), and can be used to obtain the temporal correlation function described in Eq. (10.14) as well. However, the meaning of the variable \( Q \) which determines the sectioning is different than \( \Omega \). \( Q \) introduces a better sectioning effect due to the extra \(-2\beta\) factor [182]. From Eq. (10.17), we also see that the light scattering associated with the scattering vector \( k_{\perp} \) affects the reconstruction in time-domain OCT.

### 10.5 Summary

In summary, we presented a new way of solving the inverse scattering problem in the wavevector space, without the far-field approximations. This method provides a theoretical foundation for achieving high-resolution Fourier-domain 3D reconstructions. We applied our method for systems under the Fresnel
approximation, and formulated both the Fourier-domain and time-domain 3D reconstruction models in transmission and reflection. The models are expected to improve experimental results by achieving spatially invariant high-resolution reconstruction with low-coherence light. Aside from solving inverse scattering problems, we envision solving wave equations in the wavevector space, i.e., bypassing the Green’s function, to be widely used in the future for many other optical problems.

As quantitative phase imaging and light scattering are now merged into a single discipline, we anticipate that better understanding of the light-tissue interaction will emerge. In particular, due to lack of better tools for studying tissue scattering, researchers have assumed for a long time a size distribution of spheres, scattering independently, and, thus, obeying Mie theory. Today, QPI of thin tissues is much more informative, revealing strong spatial correlations in unlabeled tissues and connecting these to the scattering of the bulk [197].
CHAPTER 11: OUTLOOK ON OPTICAL TOMOGRAPHIC IMAGING

From Chapter 8 to Chapter 10, we introduced the fundamentals of low-coherence light imaging and tomography, developed an inverse scattering theory to describe WDT, and applied it to imaging systems under the Fresnel approximation. Our wavevector space field decomposition method developed by us can be applied to many other optical problems. The first problem is to solve refocusing in QPI. During live imaging, the cells or tissues may move out of focus due to the environmental changes. Thus, it is necessary to develop an algorithm to correct it. It is found that the theory we developed for WDT can be modified for this particular problem. The theoretical description on how to solve the refocusing problem is presented in Appendix B.1. We carried out some experimental work on this and tried to refocus SLIM tissue images. However, we found that the measured SLIM tissue images have much larger depth of field than expected in theory. We are trying to identify the issues and then implement our equations to bring the SLIM images back in focus. One of the issues might be the spatial coherence. Usually, as we close down the condenser aperture to make the light more coherent, the depth of field appears to be smaller.

The second problem to solve is the optical trapping force. We have been working on using the wavevector space method to calculate the force, without any approximations, and eventually apply it to the study of cell mass transportation. For this calculation, we represent the spherical wave as well as the aperture of the lens in the wavevector space. By using the Fourier transform properties, the optical trapping force can be formulated in a very simple way. So far, we have been able to write a complete form for the trapping force. The derivation is included in Appendix B.2. The future work will be on further simplifying the analytical solution of the trapping force; and using the formula for calculating the force of the cells in live quantitative phase images. We believe that the calculated force will help us to better understand the mass transportation of the cells.
The third problem that we are trying to solve is the diffraction integral, i.e. calculating the near-field and field diffraction after an aperture in an accurate way using the wavevector space method. One of the most important theories on diffraction is the Fresnel-Kirchhoff theory. This theory imposes boundary conditions for both the field strength and its normal derivative on the aperture plane, which indicates the field should be zero everywhere behind the aperture. However, this contradicts the physics. Thus, the Fresnel-Kirchhoff diffraction integral formula cannot predict an accurate field distribution near the aperture, and it only works in the far-field [198, 199]. However, by properly choosing the Green’s function, the contradiction can be removed, as this is done in the Rayleigh-Sommerfeld (RS) diffraction integral formula. Most recently, Colin Sheppard et al. [200, 201] have done a wavevector space calculation of the RS diffraction integral formula, directly starting from the Green’s function. However, the work we plan to do is solving the diffraction integral in the wavevector space starting from the very beginning, i.e., wave equation. We have some preliminary results on this. One of my concern in the calculation is that we do not have any obliquity factors, like they do in the Rayleigh-Sommerfeld or Fresnel-Kirchhoff diffraction integral.

The fourth problem that we will work on is deriving the lens equation using our diffraction tomography result. This calculation is expected to accurately simulate the focusing of the field after the lens, which is also expected to automatically include the dispersion and aberration effects of the lens. The initial calculations are summarized in the Appendix B.3.

Besides the above problems, we can also use our diffraction tomography theory for better understanding of 3D imaging resolution in the Ewald sphere, which relates to the momentum transfer between the scattering and excitation. Recently, time-reversal and refocusing in deep tissue imaging has been a hot topic. By treating the tissue as a stack of thin slices, thus, allowing for the 1st-order Born approximation, we can apply our scattered field solutions to calculate the diffusion parameters. This may be used to prove or disprove time-reversal based on the fundamental principle of the uncertainty relation.
CHAPTER 12: CONCLUSION

In conclusion, in this dissertation we have discussed developing two major applications of interferometric light microscopy, one in defect metrology and one in 3D biomedical imaging. Chapter 1 provided the background and motivation for the dissertation. Chapter 2 reviewed quantitative phase imaging. Chapter 3 started the first application of interferometric light imaging, and we did a survey on the semiconductor defect metrology field. In Chapter 4, we discussed out development of semiconductor wafer inspection tools based on interferometric light imaging for inspecting patterned silicon wafers. In Chapters 5 and Chapter 6, we presented the progress we have made so far on implementing the inspection tools for 22 nm and 9 nm node wafer defect inspection. In Chapter 7, we concluded the wafer defect inspection work and proposed the future work. The second application of the interferometric light imaging we proposed is on 3D tomographic object reconstruction. Chapter 8 served as the foundation for low-coherence tomographic imaging. Chapter 9 provided the scattering inverse problem in white-light quantitative phase imaging and demonstrated 3D cellular structure reconstruction. Chapter 10 generalized inverse scattering and also applied our solutions to imaging systems under the Fresnel approximation. In Chapter 11, we concluded the 3D object reconstruction work and laid out the future work.
A.1 System noise characterization

Fig. A.1. Noise histogram with and without diffuser for 532 nm laser epi-DPM. (a)(c) The spatial noise histogram without and with diffuser, respectively. (b)(d) The temporal noise histogram without and with diffuser, respectively.

Our wafer inspection system is a highly sensitive technique since it uses epi-DPM to achieve common-path interferometry. By inserting a rotating diffuser in the input port of our system, we can further improve the...
system sensitivity. In order to quantitatively study the noise in our system we use the following noise characterization procedure: we first measure a plain part of the wafer (no pattern in the region) and collect 256 frames at one field of view and calculate an average phase and use it as a calibration phase image. Then we collect another 256 frames at another field of view still in a plain part of the wafer and subtract the calibration phase image from each frame. With the calibrated images, we calculate the standard deviation at each pixel and obtained a histogram for the temporal noise. We use all the pixel values of all the frames to obtain the spatial noise histogram. In Fig. A.1, we show the noise histogram with and without a diffuser for the 532 nm laser epi-DPM. Figure A.1(a) and (b) are the spatial noise and temporal histogram without a diffuser, respectively. The standard deviation for spatial noise is 9.16 nm and the temporal noise is 1.45 nm. With the diffuser in place, the spatial noise is improved from 9.16 nm to 5.18 nm, and the temporal noise is improved from 1.45 nm to 0.85 nm, which demonstrate that our system indeed has nanometer sensitivity. There are about 2x improvements for both spatial and temporal noise. Thus, using a rotating diffuser is very necessary to mitigate the residual noise.

A.2 Mathematical description of 2DISC method

A.2.1 2nd-order difference image

The mathematical description for the $n^{th}$ amplitude or phase frame can be written as

$$
F_n(x, y, t_n) = \left\{ L(x - n \cdot ds, y) \otimes P(x, y) \right. \\
+ N_s(x, y) + N_i(x, y) + N_i(x, y, t_n) \left\} W(x, y) S(x, y),
$$

(A.1)

where

$$
W(x, y) = \text{rect} \left( \frac{x}{F_x} \right) \text{rect} \left( \frac{y}{F_y} \right).
$$

(A.2)
\( L(x, y) \) is the image due to the wafer’s underlying structure and any defects, \( ds \) is the translation step size between adjacent frames, \( P(x, y) \) is the PSF of the imaging system, \( N_s(x, y) \) is the additive time-invariant spatial noise, \( N_i(x, y) \) is the additive time-invariant noise from system imperfections, \( N_v(x, y, t_n) \) is the additive time-variant spatial noise, \( W(x, y) \) is the window function of the camera that has field of view \( F_x \) in \( x \) and \( F_y \) in \( y \), and \( S(x, y) \) is the multiplicative time-invariant noise, i.e. the system response function that accounts for spatial variations in illumination intensity, illumination path length, or camera response.

The symbol \( \otimes \) denotes two-dimensional convolution. Due to sample translation, the imaged structure is shifted by \( n \cdot ds \) in \( x \). To completely remove the additive time-invariant spatial noise and system imperfection noise, we calculate the 2\(^{nd}\)-order difference image frame. The \( n \)\(^{th}\) frame is defined as:

\[
F_{n+1}(x, y, t_{n+1}) - 2 \times F_n(x, y, t_n) + F_{n-1}(x, y, t_{n-1}).
\]

Using Eq. (A.1), the 2\(^{nd}\)-order difference image frame \( n \) is written as:

\[
H_n(x, y, t_n) = \left\{ \left( L(x - (n+1) \cdot ds, y) - 2L(x - n \cdot ds, y) + L(x - (n-1) \cdot ds, y) \right) \otimes P(x, y) \\
+ N_s(x, y, t_{n+1}) - 2N_s(x, y, t_n) + N_s(x, y, t_{n-1}) \right\} \times W(x, y) S(x, y).
\] (A.3)

Note that in addition to removing these noise sources, the 2\(^{nd}\)-order difference operation also high-pass filters the image. Thus, it suppresses the low-spatial frequency components of noise in favor of the high-frequency ones, which are precisely the ones coming from the defect or from the wafer’s underlying structure.

### A.2.2 Image stitching

The additive time-variant spatial noise has random spatial variations at each image capture time \( t_n \). Thus, the best way to remove temporal noise is through averaging. Since we have a sequence of 2\(^{nd}\)-order difference images, we can stitch the sequence together to produce a panoramic average image. However, due to the small translation between images, there are large overlaps between frames. Thus, during the
image stitching, the overlapping area is averaged and we can drop the time-variant spatial noise terms and obtain

\[ G(x, y) = \left( (L(x - ds, y) - 2L(x, y) + L(x + ds, y)) \ast P(x, y) \right) W'(x, y) S'(y), \quad (A.4) \]

where

\[ W'(x, y) = \text{rect} \left( \frac{x}{F_x + N \cdot ds} \right) \text{rect} \left( \frac{y}{F_y} \right), \quad (A.5) \]

\( G(x, y) \) is the 2\textsuperscript{nd}-order difference panoramic image of the wafer, \( N \) is the total number of steps and \( S'(y) \) is the average of \( S(x, y) \) over \( x \). Note that the FOV in the panoramic image described by \( W'(x, y) \) is increased by \( N \cdot ds \) in \( x \). Further, the multiplicative noise has been greatly reduced since the system response \( S(x, y) \) has been averaged over \( x \). By using the translation invariant property of convolution [202], and applying a few approximations (a rigorous proof is given later in Appendix A.2.4), we can equivalently write Eq. (A.4) as:

\[ G(x, y) = \hat{L}(x, y) \ast K(x, y), \quad (A.6) \]

where

\[ K(x, y) = P(x - ds, y) - 2P(x, y) + P(x + ds, y), \quad (A.7) \]

and \( L(x, y) \) is the stitched wafer structure after some slight cropping at each end in \( x \).

\subsection{A.2.3 Defect extraction}

After the 2\textsuperscript{nd}-order difference and image stitching steps, the defect will appear as a tripole pattern of \( +1 -2 +1 \). To separate the defect and the patterned wafer’s underlying structure, we let

\[ \hat{L}(x, y) = U(x, y) + V(x, y), \quad (A.8) \]
where \( U(x, y) \) is the underlying structure and \( V(x, y) \) is the defect such as a small dot or an extra line.

From Eqs. (A.3) and (A.4), we see that the defect is convolved with a +1 -2 +1 tripole pattern of the system’s PSF \( P(x, y) \). To detect this signature in the panoramic image, we convolve the 2\textsuperscript{nd}-order difference panoramic image \( G(x, y) \) with a matched tripole pattern. This results in

\[
Q(x, y) = (U(x, y) + V(x, y)) \otimes P(x, y) \otimes T(x, y) \otimes d(x, y), \tag{A.9}
\]

where

\[
d(x, y) = \delta(x + 2ds, y) - 4\delta(x + ds, y) + 6\delta(x, y) - 4\delta(x - ds, y) + \delta(x - 2ds, y), \tag{A.10}
\]

and \( d(x, y) \) is the new defect pattern, which now has an extent of \( 4ds \) along the \( x \) direction. Now we can see that the test function \( T(x, y) \) needs to be specifically designed such that it prefers the defect pattern \( V(x, y) \) but rejects the underlying pattern \( U(x, y) \).

**A.2.4 A mathematical proof**

The sifting property of Dirac delta function in 2D is described as [202]

\[
f(x, y) \otimes \delta(x - a, y - b) = f(x - a, y - b). \tag{A.11}
\]

Using the sifting property, Eq. (A.4) can be written as

\[
G(x, y) = \left[ (L(x, y) \otimes (\delta(x - ds, y) - 2\delta(x, y) + \delta(x + ds, y))) \otimes P(x, y) \right] W'(x, y)S'(y). \tag{A.12}
\]

Applying the associativity property of convolution, we obtain

\[
G(x, y) = \left[ L(x, y) \otimes (P(x, y) \otimes (\delta(x - ds, y) - 2\delta(x, y) + \delta(x + ds, y))) \right] W'(x, y)S'(y). \tag{A.13}
\]

Use the sifting property again, and Eq. (A.13) becomes

\[
G(x, y) = (L(x, y) \otimes K(x, y)) W'(x, y)S'(y), \tag{A.14}
\]

where
\[ K(x, y) = P(x - ds, y) - 2P(x, y) + P(x + ds, y). \] (A.15)

Now we assume the multiplicative noise term is approximately constant: \( S'(y) = S_0 \) and we normalize the image such that \( S_0 = 1 \). It should be pointed out that averaging at the beginning and at the end of the panoramic image uses a fewer number of frames than at the middle of the panoramic image. In the \( x \) direction, the number of averaging frames increases linearly from 1 to \( [F_x / ds] \) with a slope of 1, stays constant at \( [F_x / ds] \), and decreases back to 1 linearly with a slope of -1. Thus, to maintain a reasonably low background noise level across the image, we crop the stitched image at both ends in \( x \) such that the number of averages is at least \( [F_x / ds] / 2 \). Let \( W_c(x, y) \) be the crop function

\[ W_c(x, y) = \text{rect}\left( \frac{x}{N \cdot ds} \right) \text{rect}\left( \frac{y}{F_y} \right). \] (A.16)

After cropping, we obtain a panoramic image having an approximate area of \( N \cdot ds \) by \( F_y \)

\[ G(x, y) = (L(x, y) \otimes K(x, y)) W_c(x, y). \] (A.17)

Consider \( L(x, y) = L(x, y) W_c(x, y) \). Since \( K(x, y) \) is non-zero only in a very small region (approximate width of \( dx \sim 2(ds + \Delta \rho) \) and height of \( dy \sim \Delta \rho \), where \( \Delta \rho \) is the diffraction spot size) compared to the FOV defined by \( W_c(x, y) \), we find that

\[ L(x, y) \otimes K(x, y) = (L(x, y) \otimes K(x, y)) W_c(x, y), \] (A.18)

except for the strips of width \( dx \) or of height \( dy \) at the perimeter of the image. Thus, we can simply write Eq. (A.17) as

\[ G(x, y) = L(x, y) \otimes K(x, y). \] (A.19)
A.3 Optimal parameters selection

In this section, we discuss parameter selections and present a systematic study of the PSNR for each part of the 2DISC method. First, we explain the choice of translation direction (horizontal, i.e. parallel to the underlying line structure). Since adjacent images will be subtracted after mechanical translation, we would like the lines to partially overlap so that there is partial cancellation of the background signal. Since the width (22 nm) of the lines is much smaller than their length (120 nm or 260 nm), the precision needed for alignment is much tighter for vertical translation. The individual lines are actually blurred by the PSF of the system into overlapping ellipses so some cancellation does still occur for translations in either direction. We pick horizontal translation since it has looser alignment tolerances.

Next, we discuss the optimal shape for the matched test function $T(x, y)$. An isolated defect appears like the PSF of the microscope system. Thus, to maximize the defect signal, the profile of the test function should match the PSF. The PSF for the amplitude image in an ideal system is expected to be the jinc function, $J_1(z)/z$, where $z = 3.83 r_0$, $r = \sqrt{x^2 + y^2}$, and $r_0 = 720$ nm is the diffraction spot size. The lineshape of the PSF for the phase image is more complex and still an open question [203]. We will fit both PSFs with Gaussian functions. The Gaussian function has a similar lineshape as the jinc function near the origin. From another perspective, the convolution with the defect pattern as described by Eq. (A.9) is analogous to a Canny edge detector [204] for detecting the defect pattern. Further, the Gaussian function was also previously successful in modeling the amplitude PSF [74]. Let

$$T(x, y) = e^{-\frac{(x^2+y^2)}{2w^2}},$$

(A.20)

where $w$ is the Gaussian width, i.e. the standard deviation. Let $w_0 = 2r_0/3.83 = 375$ nm. For the Gaussian to match the jinc function, we need to set $w = w_0$. Thus, $w_0$ represents the equivalent width of the PSF when using a Gaussian function.
Fig. A.2. Test function lineshape and width study for 2DISC (a) amplitude and (b) phase images. The best PSNR in both cases occurs for the Gaussian test function when the width is around 360 nm.

Figure A.2 shows an optimization study for the width of the Gaussian function (see blue curves). From this study, we find the best PSNR can be achieved for both phase and amplitude 2DISC images for $w$ around 360 nm. Convolution serves as a low pass filter whose bandwidth is determined by $w$. When $w$ is large, i.e., comparable to the step size, the defect signal is filtered out and thus the PSNR is reduced. When $w$ is small, i.e. significantly less than $w_0$, the underlying features of the wafer structure are barely filtered and so there is little benefit from the convolution step. The Lorentzian function has a similar profile to a Gaussian but decays much more slowly, which may perform more filtering and possibly improve the PSNR. Thus, a test function with Lorentzian lineshape is also investigated (see red curves). The Lorentzian function is defined by

$$T(x, y) = \frac{1}{\pi} \frac{\Gamma/2}{\left(x^2 + y^2\right) + \left(\Gamma/2\right)^2}.$$  \hfill (A.21)

For the Lorentzian test function, the best PSNR value and the optimal choice of $\Gamma$ for amplitude images are comparable to the Gaussian test function case; however, for phase images, the optimal choice for $\Gamma$ occurs
at a significantly different value. Since the Gaussian test function works well in both cases with the same parameter, we chose it as the test function and \( w \) to be 360 nm.

Fig. A.3. PSNR versus order of the central difference for: a difference only method (D), a difference plus image stitching method (DIS), and the full DISC method for (a) amplitude and (b) phase.

Fig. A.4. 2DISC method step size study for (a) amplitude and (b) phase images. A step size of 0.75 \( \mu m \) or 1.0 \( \mu m \) yields the best performance in both cases.

Now, we study the improvements in PSNR as we increase the order of the difference method from \( 0^{th} \)-order to \( 4^{th} \)-order. As we change the difference order, the defect pattern \( d(x, y) \) will change accordingly. Thus, for the convolution step, we need to convolve the different orders with their corresponding patterns, e.g. dipole +1 -1 for \( 1^{st} \)-order, tripole +1 -2 +1 for \( 2^{nd} \)-order, etc. We also study the PSNR for three methods:
a difference only method (D), a difference plus image stitching method (DIS), and the full DISC method. Figure A.3 shows these results for the amplitude and phase images. From this figure, we see that the raw single frame image (0\textsuperscript{th}-order, D) has the worst PSNR and that increasing the order improves the PSNR until it saturates around 2\textsuperscript{nd}-order for each method. Going from 0\textsuperscript{th}-order to 2\textsuperscript{nd}-order difference is adequate to remove the additive time-invariant spatial noise such as speckle and system imperfections and high-pass filter the image to remove the low-spatial frequency components of other noise sources. Beyond 2\textsuperscript{nd}-order, there is no noticeable performance improvement. This is because high-frequency noise is also accentuated as much as the signals from the defect and the wafer pattern. From the figure, we also see that the image stitching step produces a few dB of improvement over the difference only method. This indicates that there is some additive time-variant noise due to instabilities in the system or some multiplicative noise due to non-uniformities in the system response. However, these noise sources are not as strong as other noise sources such as speckle. This is due to the excellent stability of the epi-DPM system and the relatively uniform system response. The final trend to discern from the figure is that the convolution step is critical. It significantly reduces the high-frequency components from the background pattern while preserving the strength of the defect signal. The PSNR increases 7-10 dB and as a consequence, the defect detectability is greatly enhanced. Given the conclusions of the optimal parameters selection, we select the 2DISC method with a Gaussian test function with $w = 360 \text{ nm}$.

Finally, we discuss the selection of the translation step size. To produce a strong tripole pattern that is spatially separated, the step size needs to be larger than or at least comparable to the diffraction spot size of 720 nm. However, the step size should be relatively small so that noise that is invariant to small shifts can be canceled. The step size also must be a multiple of the minimum step size of the scanning stage, which was 0.25 $\mu$m for our microscope. Given these constraints, we chose step size of 0.75 $\mu$m. The optimal step size may also depend on the wafer’s underlying structure. To verify these hypotheses, a sequence of 360
interferogram frames with adjacent frame horizontal translation steps of 0.25 µm is collected. Thus, the step size can effectively be varied by forming subsequences of images separated by a fixed number of steps. Figure A.4 shows the general trends for several different choices of the step sizes. For small steps (0.25 and 0.50 µm), the PSNR is small because the contrast of the defect’s tripole pattern is low since the positive and negative signals overlap spatially. For large steps (1.25, 1.5, and 1.75 µm), the PSNR is small because the shift-invariant noise cancellation is less effective for large steps. From this step size study, we conclude that the optimal step size is in the range of 0.75 to 1.0 µm and that the performance is not critically dependent on the exact choice of step size.

A.4 Wafer referencing

In practical wafer inspection systems, we need a defect-free pattern to detect defects in the test pattern. This pattern is used as a reference for generating a difference image to extract the defect from the test pattern. Here, we propose a method using cross-correlation to address this problem. This method uses pattern self-referencing which does not need a reference pattern, but it requires the pattern itself to be periodic. For non-periodic patterns, we can use our cross-correlation method provided we have access to the signal from a reference pattern.

For this wafer referencing study, we will use the 22 nm node wafer as described in Chapter 5. For the stitched 2nd-order difference images, there is undesired underlying structure from the wafer pattern. In order to remove the underlying structure, a reference pattern is needed. Since the IDA wafer pattern is periodic, we can use it as a self-referenced pattern. The underlying structure can be canceled, leaving only the defect tripole pattern in the final stitched 2nd-order difference images. Here, we show the steps of how to perform the self-referencing in the IDA wafer. We start with the scan amplitude images, Fig. A.5(a)-(c) show an example set of three adjacent frames with separation of 0.75 µm. The defect is marked by a red rectangular box. For self-referencing, we need to achieve 1.6 µm frame separations. This can be done simply by
digitally shifting Figs. A.5(a) and (b) by 0.85 µm to the right and left, respectively. However, due to the translation stage stepping error, the actual shifting amount can be different. Thus, we perform cross-correlation between Figs. A.5(a) and (b), and Figs. A.5(b) and (c) to find the exact shifting amount. With cross-correlation method, we found that Fig. A.5(a) needs 0.833 µm digital shifting to the right and Fig. A.5(c) needs 0.786 µm digital shifting to the left. After the digital shifting, we produce a 2nd-order difference image which is shown in Fig. A.5(d) where the defect starts to become noticeable. We repeat this digital-shifting for all the scan frames, and produce a sequence of 2nd-order difference images. Then, we stitch all the images together to produce a panoramic image as shown in Fig. A.5(e). Now, we see that the defect signal to underlying structure noise ratio is very high, leading to clear detection of the defect in the 70 µm by 27 µm field of view. A tripole matched convolution can be performed on Fig. A.5(e) to further enhance the defect signal. The image shows detection of a parallel bridge defect with 20 nm by 100 nm size which is verified by SEM. This self-referencing method can also be applied to the scan phase images.

![Illustration of pattern self-referencing](image)

**Fig. A.5.** Illustration of pattern self-referencing. (a)-(c) is an example set of three adjacent amplitude images with 0.75 µm separation, (d) is the 2nd-order difference after digital shifting with realization of 1.6 µm frame separation, and (e) is the panoramic 2nd-order difference amplitude image after stitching. This detection image has a 20 nm by 100 nm parallel bridge defect which is verified by SEM.
APPENDIX B: WAVEVECTOR SPACE CALCULATIONS

B.1 Refocusing

Here we derive the equation for refocusing using WDT theory. Using this equation, we expect to bring out of focus image frames back in focus during live imaging. For a slice at \( z \), the measured cross-correlation signal can be written as

\[
\Gamma_{12}(\mathbf{k}_z, z; 0) = \sum (\mathbf{k}_z, -z) \otimes \chi(\mathbf{k}_z, -z).
\]  

(B.1)

If the slice is defocused by \( \Delta z \), then we have

\[
\Gamma_{12}(\mathbf{k}_z, z + \Delta z; 0) = \sum (\mathbf{k}_z, -(z + \Delta z)) \otimes \chi(\mathbf{k}_z, -z).
\]  

(B.2)

Assuming the object is thin and centered at \( z = z_0 \), we have

\[
\chi(\mathbf{k}_z, -z) = \chi(\mathbf{k}_z) \delta(z - z_0).
\]  

(B.3)

Thus, Eq. (B.2) can be written as:

\[
\Gamma_{12}(\mathbf{k}_z, z_0 + \Delta z; 0) = \sum (\mathbf{k}_z, -(z + \Delta z)) \otimes (\chi(\mathbf{k}_z) \delta(z - z_0)) \\
= \chi(\mathbf{k}_z) \sum (\mathbf{k}_z, -(z + \Delta z)) \otimes \delta(z - z_0) .
\]  

(B.4)

For simplification, let \( z_0 \to 0 \) and \( \Delta z \to z \), we have

\[
\Gamma_{12}(\mathbf{k}_z, z; 0) = \chi(\mathbf{k}_z) \sum (\mathbf{k}_z, -z).
\]  

(B.5)

At \( z = 0 \)

\[
\Gamma_{12}(\mathbf{k}_z, 0; 0) = \chi(\mathbf{k}_z) \sum (\mathbf{k}_z, 0).
\]  

(B.6)

Thus, for focus correction, we use Eq. (B.7)
Thus, we need to find an optimum $z$ to make $\Gamma_{12}$ sharp. Recall:

$$\Gamma_{12}(r, r) = |\Gamma_{12}(r, r)| e^{i\phi(r, r)}. \quad (B.8)$$

B.2 Optical trapping

Start from the Helmholtz equation

$$\nabla^2 U_s(r) + \beta^2 U_s(r) = -\chi(r) \beta^2 U(r), \quad (B.9)$$

where $\beta(\omega) = \bar{n} \beta_0(\omega)$, with $\bar{n}$ the spatial average of the refractive index associated with the object,

$$\bar{n} = \langle n(r) \rangle_r, \quad \chi(r) = n^2(r) - \bar{n}^2. \quad \text{Under the 1st-order Born approximation, we can approximate } U(r, \omega) \text{ on the right-hand side as } U_i(r, \omega), \text{ allowing Eq. (B.9) to be re-written as}$$

$$\nabla^2 U_s(r) + \beta^2 U_s(r) = -\chi(r-r_0) \beta_0^2 U_i(r). \quad (B.10)$$

Performing a Fourier transform over both sides yields

$$\left[ \beta^2 - k^2 \right] U_s(k, \omega) = -\beta_0^2 \chi(k) e^{ik_0} \otimes U_i(k). \quad (B.11)$$

The scattering field can be solved as

$$U_s(k) = -\beta_0^2 \chi(k) e^{ik_0} \otimes U_i(k) \frac{1}{\beta^2 - k^2}. \quad (B.12)$$

Assume $U_i(k)$ is a spherical wave cropped by the numerical aperture function $A(k_\perp)$, thus

$$U_i(k) = \frac{A(k_\perp)}{k^2 - \beta^2}. \quad (B.13)$$

Then we have

$$\Gamma_{12}(k_\perp, z; 0) = \Gamma_{12}(k_\perp, z = 0; 0) \frac{\Sigma(k_\perp, -z)}{\Sigma(k_\perp, 0)}. \quad (B.7)$$
\[ U_s(k) = \beta_0^2 \left( \chi(k) e^{ik \cdot r_0} \otimes \left( A(k_\perp) \frac{1}{k^2 - \beta^2} \right) \right) \frac{1}{k^2 - \beta^2}. \]  

(B.14)

Taking an inverse Fourier transform to bring this back to \( r \) space,

\[ U_s(r) = \beta_0^2 \left( \chi(r - r_0) F.T^{-1} \left( A(k_\perp) \frac{1}{k^2 - \beta^2} \right) \right) \otimes F.T^{-1} \left( \frac{1}{k^2 - \beta^2} \right) \]

\[ = \beta_0^2 \left( \chi(r - r_0) A(r_\perp) \otimes F.T^{-1} \left( \frac{1}{k^2 - \beta^2} \right) \right) \otimes F.T^{-1} \left( \frac{1}{k^2 - \beta^2} \right) \]

\[ = \beta_0^2 \chi'(r - r_0) \otimes \frac{e^{-i \beta r}}{r} \]

\[ = \beta_0^2 \chi'(r - r_0) \otimes \frac{e^{-i \beta r}}{r} \]

(B.15)

To calculate the trapping force in the \( z \)-direction, we need to calculate the momentum defined by

\[ \langle k_z \rangle = \int_{-\infty}^{\infty} k_z |U_s(k)|^2 \, dk \]

\[ = \mathcal{F} \left\{ k_z |U_s(k)|^2 \right\}_{r=0} \]

\[ = \left\{ \frac{\partial}{\partial z} \mathcal{F}^{-1} \left| \left\{ U_s(k) \right\} \right| \right\}_{r=0} \]

\[ = -i \left\{ \frac{\partial}{\partial z} \left[ U_s(-r) \otimes U_s(-r) \right] \right\}_{r=0} \]

where we can write

\[ U_s(r) \otimes U_s(r) = \beta_0^2 \left( \chi'(r - r_0) \otimes \frac{e^{-i \beta r}}{r} \right) \otimes \left( \chi'(r - r_0) \otimes \frac{e^{-i \beta r}}{r} \right) \]

\[ = \beta_0^2 \left( \chi'(r - r_0) \otimes \chi'(r - r_0) \right) \otimes \left( \frac{e^{-i \beta r}}{r} \otimes \frac{e^{-i \beta r}}{r} \right) \]

(B.17)

\[ \langle k_z \rangle = -i \left\{ \frac{\partial}{\partial z} \left[ U_s(-r) \otimes U_s(-r) \right] \right\}_{r=0} \]

\[ = -i \beta_0^2 \frac{\partial}{\partial z} \left\{ \left( \chi'(r - r_0) \otimes \chi'(r - r_0) \right) \otimes \left( \frac{e^{-i \beta r}}{r} \otimes \frac{e^{-i \beta r}}{r} \right) \right\}_{r=0} \]

(B.18)

Making a Taylor expansion for \( \chi'(r - r_0) \)
\[ \chi'(\mathbf{r} - \mathbf{r}_0) \approx \chi'(\mathbf{r}) + x_0 \frac{\partial \chi'(\mathbf{r}')}{\partial x} \bigg|_{r'=r} + y_0 \frac{\partial \chi'(\mathbf{r}')}{\partial y} \bigg|_{r'=r} + z_0 \frac{\partial \chi'(\mathbf{r}')}{\partial z} \bigg|_{r'=r}. \] (B.19)

We will work on further simplification in the future.

### B.3 Lens equation

For a ball lens, the scattering potential is:

\[ \chi(r) = \Pi \left( \frac{r}{2a} \right) \left( n_1^2 - n_0^2 \right). \] (B.20)

Here we consider the following lenses:

(a) Plano-convex lens

\[ \chi(r) = \left( \Pi \left( \frac{r}{2a} \right) \otimes \delta(z + d) \right) \cdot H(z) \cdot \left( n_1^2 - n_0^2 \right). \] (B.21)

(b) Bi-convex lens
\[ \chi(r) = \left( \prod \left( \frac{r}{2a} \right) \otimes \delta(z - d) \right) \cdot H(-z) + \left( \prod \left( \frac{r}{2a} \right) \otimes \delta(z + d) \right) \cdot H(z) \cdot \left( n_1^2 - n_0^2 \right). \]  \hspace{1cm} (B.22)

(c) Concave-convex lens

\[ \chi(r) = \left( \prod \left( \frac{r}{2a} \right) \otimes \delta(z - d) \right) - \left( \prod \left( \frac{r}{2a} \right) \otimes \delta(z + d) \right) \cdot H(z) \cdot \left( n_1^2 - n_0^2 \right). \]  \hspace{1cm} (B.23)

For the plano-convex lens, its Fourier transform is

\[ \chi_{pc}(k) = \left( \Pi(k) e^{-ik \cdot d} \right) \otimes H(k) \cdot \left( n_1^2 - n_0^2 \right), \]  \hspace{1cm} (B.24)

where (from Chapter 2 in reference [13])

\[ \Pi(k) = \frac{\sin(ka) - ka \cos(ka)}{k^3}, \]  \hspace{1cm} (B.25)

\[ H(k) = \pi \delta(k_z) - \frac{i}{k_z}. \]  \hspace{1cm} (B.26)

So we can write Eq. (B.24) as

\[ \chi_{pc}(k) = \left( n_1^2 - n_0^2 \right) \cdot \left( \frac{\sin(ka) - ka \cos(ka)}{k^3} e^{i k \cdot d} \right) \otimes \left( \pi \delta(k_z) - \frac{i}{k_z} \right) \]

\[ = \left( n_1^2 - n_0^2 \right) \cdot \left( \pi \cdot \frac{\sin(ka) - ka \cos(ka)}{k^3} e^{i k \cdot d} - i \frac{\sin(ka) - ka \cos(ka)}{k^3} e^{i k \cdot d} \otimes \left( \frac{1}{k_z} \right) \right). \]  \hspace{1cm} (B.27)

The bi-convex is just adding another flipped plano-convex lens, so

\[ \chi(k) = \chi_{pc}(k) + \chi_{pc}(-k). \]  \hspace{1cm} (B.28)
From the diffraction tomography theory [191], the forward scattering field is

$$U_s(k_\perp,z,\omega) = -\frac{\beta_0^2(\omega)A(\omega)e^{iqz}}{2q}X[k_\perp,q-\beta(\omega)], \quad \text{(B.29)}$$

where $q = \sqrt{\beta^2(\omega) - k_\perp^2}$. The reference field is a plane wave

$$U_r(z,\omega) = A(\omega)e^{i\beta(\omega)z}. \quad \text{(B.30)}$$

The intensity is

$$I = |U|^2 = |U_r + U_s|^2. \quad \text{(B.31)}$$

Here we show some simulation results for a ball lens by assuming wavelength equal to 0.532 µm, medium refractive index $n_0 = 1$, and the ball is glass with $n_1 = 1.5$. The intensity distributions for different ball radius are calculated.

(a) $R = 2560\lambda = 1.36$ mm.

(b) $R = 1280\lambda = 0.68$ mm.
The simulated intensity distributions show many foci after the ball lens, and the focal length increase with R, but the value is too large compared with predicted value [205]:

(c) $R = 640\lambda = 0.34\text{ mm.}$
Our assumption of the 1st-Born approximation may not be valid here. In the future we plan to calculate the plano-convex lens focusing, since it is more suitable under the Born approximation. In fact, using the Born approximation to calculate the light diffraction of a plano-convex lens was done previously [206, 207].

\[ F = \frac{n_i D}{4(n_i - 1)}. \]  

(B.32)
APPENDIX C: CO2 GAS SENSING

C.1 Introduction

Compared to traditional solid-state lasers, fiber lasers can have a more compact design, higher beam quality, and a narrower spectral linewidth [208]. Over the past years, fiber lasers at different wavelengths have been developed to cover the spectrum from visible to near infrared (NIR) [209]. Recently, NIR fiber lasers have been used for light detection and ranging (LIDAR) and gas sensing [210-213]. In particular, measurement of the CO$_2$ concentration has become important for monitoring atmospheric climate changes and spoilage in agri-food storage systems [213, 214]. Here, we are interested in developing a low-cost system for distributed monitoring of the CO$_2$ level in large volume corn storage bins to minimize post-harvest food loss.

Previous experiments demonstrated CO$_2$ sensing using 1573 nm and 2051 nm laser sources [211, 213]. CO$_2$ also has strong and sharp absorption peaks near 2005 nm. The absorption cross section at 2005 nm is approximately 100x stronger than at 1573 nm and 5x stronger than at 2051 nm. We built an all fiber based thulium holmium (Tm-Ho) co-doped 2005 nm fiber laser for CO$_2$ sensing as shown in Fig. C.1(a). The fiber laser is pumped by a commercially available 808 nm fiber coupled diode laser and has a threshold pump power of 60 mW. The output power is 2.5 mW at 90 mW pump which is adequate for CO$_2$ sensing. The fiber Bragg grating (FBG) cavity mirrors are temperature modulated to achieve laser wavelength modulation around the 2005 nm absorption peak. We measure the signal transmittance through a 15 cm long 2% CO$_2$ test cell and a 100% CO$_2$ reference cell, both of which have anti-reflective (AR) coated angled facets to minimize etalon effects. Results show that we can quantify the CO$_2$ concentration in the 2% cell with decent accuracy. Water vapor adds a small slope to the absorption spectrum and so with calibration,
we expect this fiber laser sensing system to be able to measure both the CO\textsubscript{2} concentration and the relative humidity at very low concentration with great accuracy.

**C.2 Tm-Ho co-doped fiber laser**

The experimental setup for measuring the light-light (L-L) curves of the 2005 nm Tm-Ho fiber laser is illustrated in Fig. C.1(a). A 30 cm long Tm-Ho co-doped fiber is end pumped by a single mode (SM) fiber coupled 808 nm diode laser. The Tm-Ho fiber laser cavity is closed by two FBGs with center wavelength of 2004.5 nm and reflectivity of 96\% and 53\%, respectively. The full width at half maximum (FWHM) values of the FBGs are 0.2 nm and 3.5 nm, respectively. After the fiber, the beam is collimated by a CaF\textsubscript{2} lens. A long pass filter blocks the pump power. The 2005 nm signal is focused onto an extended wavelength InGaAs detector through another CaF\textsubscript{2} lens to measure the signal power. The measured signal power without the 50 m of SMF28 fiber is plotted in Fig. C.1(b). From this plot, we find the laser has a clear threshold at around 60 mW pump power. The signal power is about 2.5 mW at 90 mW pump.

![Diagram](image)

*Fig. C.1. (a) Experimental setup for the Tm-Ho co-doped 2005 nm fiber laser sensor system; (b) measured laser signal versus pump power without the 50 m of higher-order mode filtering SMF28 fiber.*

148
Each fiber component is single mode at 1.55 µm, but may not be at 2 µm. Two lasing modes were observed, at 2004.4 nm and 2005.9 nm at room temperature, in the laser output spectrum using Fourier transform infrared spectroscopy (FTIR). Since the 2005.9 nm mode is outside the reflection bandwidth of the 0.2 nm FBG, we believe that this corresponds to a higher-order transverse mode. The mode mismatch at Tm-Ho active fiber and passive fiber splice interfaces can excite a higher-order mode that can then satisfy the Bragg condition at a slightly different wavelength than the fundamental mode. Thus, to produce a single wavelength output signal, we added 50 m of SMF-28 fiber to filter out the higher-order mode since that mode has a relatively higher loss. Successful filtering was confirmed by FTIR. Although the fiber reduced the overall signal power by 5.6 dB, the noise in the output power was greatly reduced.

C.3 CO₂ sensing

![Graphs](image)

Fig. C.2. (a) Transmission spectrum for the 100% and 2% CO₂ cells as the FBG temperature is scanned to sweep the laser across an absorption line. The oscillatory ripples are due to etalon effects. (b) Logarithmic ratio plot and best fit line that is constrained to pass through the origin. The unknown concentration is determined as the slope of the line, which in this case is 2.22%.

To demonstrate CO₂ sensing, we used a 15 cm long test cell filled with 2% CO₂ and a reference cell with 100% CO₂ so that the CO₂ concentration can be determined without prior knowledge of the CO₂ absorption
spectrum or the operating wavelength of the fiber laser. The 2005 nm signal is split and goes to three measurement arms as shown in Fig. C.2(a). The test arm has the 15 cm long 2% CO2 cell, the reference arm has the 100% CO2 cell, and the input arm has no cell and is used to measure the input laser power. We modulated the temperature of both FBGs and the Tm-Ho active fiber between 45 C and 65 C, and measured the spectrum of the laser at 45 C, 55 C and 65 C using FTIR. The FTIR spectrum showed peak wavelength modulation from 2004.7 nm to 2004.95 nm (0.013 nm/C). As can be seen in Fig. C.2, this temperature range spans one of the CO2 absorption peaks. For each temperature, we record the transmitted power for each arm and normalize by the input power to obtain the two transmission ratios: for the 2% and 100% cells. The temperature scan is repeated several times to verify that the system is stable. The spectra are averaged to reduce noise. The two averaged ratio spectra are normalized to unity at 45 C. Thus, they represent

\[
 r_{\text{sens}} = \frac{P_{\text{sens}}}{P_{\text{in}}} = \exp \left[ -C_{\text{sens}} L \left( \alpha (\lambda_T) - \alpha (\lambda_{\text{45C}}) \right) \right], \tag{C.1}
\]

and

\[
 r_{\text{ref}} = \frac{P_{\text{ref}}}{P_{\text{in}}} = \exp \left[ -C_{\text{ref}} L \left( \alpha (\lambda_T) - \alpha (\lambda_{\text{45C}}) \right) \right], \tag{C.2}
\]

where \( C_{\text{sens}} \) is the unknown concentration to be determined, \( C_{\text{ref}} = 1 \) for the 100% CO2 cell, \( L = 15 \) cm, \( \alpha \) is the absorption per length and \( \lambda_T \) is the laser wavelength at temperature T. The measurement results for the ratios are plotted in Fig. C.2(a). From this figure, the CO2 absorption peak is clearly seen for both cells at about 56 C. The noise in the spectrum for the 2% cell is more noticeable because the absorption is less pronounced and thus the oscillatory noise from small etalon effects in the system is more visible. In Fig. C.2(b), we plot \( \ln(r_{\text{sens}}) \) versus \( \ln(r_{\text{ref}}) \) since according to Eqs. (C.1) and (C.2) should result in a line with zero intercept and slope given by \( C_{\text{sens}} \). For improved accuracy, we limit the data to 52.5 C to 57.5 C so that
the values of $\ln(r_{sens})$ and $\ln(r_{ref})$ are large and thus less influenced by noise. We determine the concentration in the test cell to be $C_{sens} = 2.22\%$ and the 95% confidence interval for the measurement as 2.16-2.28%. These results are in good agreement with the manufacturer’s specification tolerances for the 2% cell. The adjusted $R^2$ for the linear fit was 0.94. For future experiments, an additional detection arm will be added with a reference $H_2O$ cell. Also, the test arm will be placed inside the storage bin using a long fiber tether. With 4 arms, we expect to not only be able to accurately determine the $CO_2$ concentration but also monitor the relative humidity in the corn bin.

C.4 Summary

A low-cost single wavelength 2005 nm Tm-Ho fiber laser has been developed. With this fiber laser source, we built a $CO_2$ concentration measurement system. By modulating the laser wavelength across a $CO_2$ absorption peak, we can easily detect and accurately quantify $CO_2$ concentrations down to 2% in a 15 cm long cell. In the future, we expect to measure $CO_2$ concentration and relative humidity simultaneously to better monitor and prevent food spoilage.
APPENDIX D: FUNDING SUPPORT ACKNOWLEDGMENTS

D.1 Semiconductor wafer defect inspection
This research was mainly funded by the Semiconductor Research Corporation (contract P13117) from 2011-2014. The work used instrumentation being developed with support from the National Science Foundation (NSF grant CBET-1040462 MRI) and matching funds from the University of Illinois. The authors express their gratitude to SEMATECH for providing the large-area nano-patterned wafer. From 2013-2014, this research also received support from the Beckman Foundation through a Beckman Graduate Fellowship award to R.Z.

D.2 Inverse scattering and 3D reconstruction
This research was supported in part by the National Science Foundation: grants CBET 08-46660 CAREER, CBET-1040462 MRI, and the Science and Technology Center for Emergent Behaviors of Integrated Cellular Systems (EBICS, CBET-0939511). The authors thank Rashid Bashir and Kidong Park for providing HT29 cells, Ido Golding and Michael Bednarz for providing E. coli cells, and Scott Robinson for assistance with SEM imaging of RBCs. The authors also acknowledge Jonathon Howard, Khaled Khairy, and Ji Jinn Foo for providing the confocal image of RBC. In addition, R. Z was supported by the Beckman foundation through a Beckman Graduate Fellowship.

D.3 CO2 gas sensing
This research was supported in part by the National Science Foundation (Award ID: ECCS-0901388) and by a grant from Archer Daniels Midland company.
REFERENCES

1. C. Dobell, and A. v. Leeuwenhoek, *Antony van Leeuwenhoek and his "Little animals"; being some account of the father of protozoology & bacteriology and his multifarious discoveries in these disciplines* (Russell & Russell, 1958).


