Simulation Framework for Electromagnetic Effects in Plasmonics, Filter Apertures, Wafer Scattering, Grating Mirrors, and Nano-Crystals

Daniel Peter Ceperley

Electrical Engineering and Computer Sciences
University of California at Berkeley

Acknowledgement

Simulation Framework for Electromagnetic Effects in Plasmonics, Filter Apertures, Wafer Scattering, Grating Mirrors, and Nano-Crystals

by

Daniel Peter Ceperley

B.S. (University of Virginia) 2002
M.S. (University of California, Berkeley) 2005

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering - Electrical Engineering and Computer Sciences
and the
Designated Emphasis in Nanoscale Science and Engineering

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Andrew R. Neureuther, Chair
Professor Jeffrey Bokor
Professor Xiang Zhang

Fall 2008

The dissertation of Daniel Peter Ceperley is approved:

Chair	Date

	Date

	Date

University of California, Berkeley

Fall 2008

Simulation Framework for Electromagnetic Effects in Plasmonics, Filter Apertures, Wafer Scattering, Grating Mirrors, and Nano-Crystals

Abstract

Simulation Framework for Electromagnetic Effects in Plasmonics, Filter Apertures,

Wafer Scattering, Grating Mirrors, and Nano-Crystals

by

Daniel Peter Ceperley

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Sciences

and the

Designated Emphasis in Nanoscale Science and Engineering

University of California, Berkeley

Professor Andrew R. Neureuther, Chair

This thesis presents a Finite-Difference Time-Domain simulation framework as well as both scientific observations and quantitative design data for emerging optical devices. These emerging applications required the development of simulation capabilities to carefully control numerical experimental conditions, isolate and quantifying specific scattering processes, and overcome memory and run-time limitations on large device structures. The framework consists of a new version 7 of TEMPEST and auxiliary tools implemented as Matlab scripts. In improving the geometry representation and absorbing boundary conditions in TEMPEST from v6 the accuracy

has been sustained and key improvements have yielded application specific speed and accuracy improvements. These extensions include pulsed methods, PML for plasmon termination, and plasmon and scattered field sources. The auxiliary tools include application specific methods such as signal flow graphs of plasmon couplers , Bloch mode expansions of sub-wavelength grating waves, and back-propagation methods to characterize edge scattering in diffraction masks.

Each application posed different numerical hurdles and physical questions for the simulation framework. The Terrestrial Planet Finder Coronagraph required accurate modeling of diffraction mask structures too large for solely FDTD analysis. This analysis was achieved through a combination of targeted TEMPEST simulations and full system simulator based on thin mask scalar diffraction models by Ball Aerospace for JPL. TEMPEST simulation showed that vertical sidewalls were the strongest scatterers, adding nearly 2λ of light per mask edge, which could be reduced by 20° undercuts.

TEMPEST assessment of coupling in rapid thermal annealing was complicated by extremely sub-wavelength features and fine meshes. Near 100% coupling and low variability was confirmed even in the presence of unidirectional dense metal gates.

Accurate analysis of surface plasmon coupling efficiency by small surface features required capabilities to isolate these features and cleanly illuminate them with plasmons and plane-waves. These features were shown to have coupling cross-sections up to and slightly exceeding their physical size. Long run-times for TEMPEST sim-

ulations of finite length gratings were overcome with a signal flow graph method. With these methods a plasmon coupler with over a 10λ 100% capture length was demonstrated.

Simulation of 3D nano-particle arrays utilized TEMPEST v7's pulsed methods to minimize the number of multi-day simulations. These simulations led to the discovery that interstitial plasmons were responsible for resonant absorption and transmission but not reflection.

Simulation of a sub-wavelength grating mirror using pulsed sources to map resonant spectra showed that neither coupled guided waves nor coupled isolated resonators accurately described the operation. However, a new model based on vertical propagation of lateral Bloch modes with zero phase progression efficiently characterized the device and provided principles for designing similar devices at other wavelengths.

Professor Andrew R. Neureuther
Dissertation Committee Chair

To my parents

Contents

List of Figures

List of Tables

Acknowledgments

I would like to start by thanking those who funded my stay in graduate school. The work presented in this thesis was funded by the Jet Propulsion Laboratory under the TPF Development of Technologies study contract, the Semiconductor Research Corporation's support of maskless lithography (task 460.013), and the IMPACT (Integrated Modeling, Process and Computation for Technology) program, a joint collaboration between the UC Discovery Program (award number Ele07-10283) and industry sponsors.

I searched for graduate schools without a specific goal in mind. I remember feeling bewildered and unprepared at the visit days. My peers had lists of professors to visit but all I had was an interest in academic research and a desire to live somewhere new. I was lost until the EECS department's visit day at U. C. Berkeley. I was sold on Berkeley on the spot. It was the people – they were happy, smart, and part of a great community. During my time at Berkeley many people made contributions to my happiness and success – thank you, all!

To my close friends, Erik Steltz, Sarah Krathwohl, Becca Krathwohl, Andrea Tao, Dave Nguyen, Ryan White, Joanna Lai, Rebecca Lynn, Dan Hazen, Elaine Cheong, Ashley Eden, Kurt Miller, Morgan Mager, Mervin Zhou, Babak Jamshidi, and Eric Tang, thank you for the the commiseration, the collaboration, the business ventures, the sports, the cinemas, and the adventures.

To my supervisors at summer internships, Mike Lieber (Ball Aerospace), Kun-

jithapatham Balasubramanian (JPL), Dan Hoppe (JPL), Tim Brunner (IBM), and Dave Medeiros (IBM), thank you for the guidance, the interesting work, the excuse to travel, and the preview of life after school.

To my peers in the Neureuther Group, Greg McIntyre, Yunfei Deng, Garth Robbins, Scott Hafeman, Mike Williamson, Lei Yuan, Wojtek Poppe, Chris Clifford, Marshal Miller, Eric Chin, Lynn Wang, Juliet Rubinstein, and Koji Kikuchi, thanks for the guidance, camaraderie, collaboration, commiseration, PhotoBears assistance, and golf. A special thanks to Michael Lam – you took me under your wing when I first arrived and transformed me from a scared outsider to a happy and productive member of the group.

To my adviser, Prof. Andy Neureuther, thank you for a spectacular education! You sparked both my interest in theory and practical applications. Your instruction extended well beyond the technical and into public speaking, career exploration, and leadership. A few weeks in the working world has shown me that the skills you've given me are invaluable. Thank you for directing me into interesting projects, providing stable funding, and rearranging your research schedule around my many summer internships.

Finally, thanks to my family for being there the whole way. From an early age I was bound and determined to go to grad school and it was your example and encouragement that set me on the path and made it happen. I remember computer and math lessons in early elementary school, coaching on my presentation skills for

innumerable projects in Ms. Klaus' third grade class, being kicked out the door to make friends and get exercise, and visits at all my different abodes. Thanks, from the bottom of my heart.

Chapter 1

Introduction

Many novel optical systems and devices are being explored for a broad variety of applications. These include telescopes for finding planets orbiting other stars, nano-particle based chemical sensors, rapid termal annealing systems, and optical microscope probles. These systems are extremely complex and require new physical insights and engineering of trade-offs in design.

This thesis adds understanding and quantitative data to several of these applications. This includes detailed analysis of edge scattering from diffraction masks used by a planet finding telescope, characterization of optical coupling through dense metal gate arrays in rapid thermal annealing, quantitative analysis of the effects of topography on surface wave coupling, the role of surface plasmons in silver nano-particle arrays, and investigation of the guided waves involved in a high reflectivity broad-band dielectric mirror.

This understanding and engineering characterization has been generated through rigorous electromagnetic simulation of the device operation. Generally these applications are so complex that simple intuition and algebraic models do not work. For example, many devices contain critical topographic elements that are on the order of the wavelength of light and/or use guided waves. These topographical elements are difficult to understand because the simple picture of light rays bouncing from surface to surface breaks down and instead light must be viewed as a wave that scatteres off corners and diffracts through openings in close proximity. Guiding of the waves often occurs and is also difficult to analyze because the wave's speed and pattern are often not algebraically tractable and conversion from guided to radiative waves is extremely difficult to predict.

Simulation of these electromagnetic effects faces challenges. Often devices are too large to fit into memory or too large to simulate in a resonable amount of time (a few days or less). Even when simulation is possible, understanding what is the dominant physical phenomenon within the device is difficult. Simulators automatically include all the many competing phenomena but reducing this down to a simple, intuitive understanding for device physics is difficult. Additionally, simulators contain approximations to physical reality and these approximations can cause simulation predictions to differ from reality and hide important phenomena if approximations are not properly understood and addressed.

Meeting these computational challenges has emerged as a central unifying theme

to this thesis. It is exemplified by the development of an extensive application specific framework to support the core simulator. This framework includes key components for accurately representing topographical elements, scattered field, guided wave, and plasmon sources, simulator accuracy and speedup balancing strategies, and post-processing to extract scattering conversion efficiencies and phases.

The dissertation flow begins with a chapter discussing TEMPEST, the core simulator used throughout this work. Many enhancements were made to TEMPEST and special emphasis is given to the scattered field and surface plasmon sources. These enhancements resulted in improved accuracy and speed in specific applications.

The first application is in Chapters 3 and 4 where the Terrestrial Planet Finder Coronagraph (TPF-C) is discussed. This optical system required rigorous computer simulation of light scattering off the corners of diffraction masks to determine if the telescope was designed well enough to see planets orbiting other stars. Computer simulation was required because scattering by sharp corners with different angles and coatings was not algebraically tractable. The diffractions masks were too large to simulate rigorously, so sharp corners were modeled with detailed simulations and the results were used to correct a fast, full system simulator. This project had to overcome large errors in wave speed caused by numerical approximations in the simulation method. This project utilized TEMPEST version 6 and motivated the development of the current version (v7).

The second application in Chapter 5 analyzed optical coupling in rapid thermal

annealing processes. Computer simulation was required because annealing light must pass between metal transistor gates that are orders of magnitude smaller than the wavelength, so complex wave and polarization phenomena are paramount.

The third application in Chapter 6 involved determining how efficiently small metal features could covert a light beam into a surface plasmon (a type of guided wave that runs across a metal surface). Rigorous simulation was required because of the complexity of converting between types of waves and the metal features performing the conversion were smaller than a wavelength. Multiple methods were used to demonstrate the physics and to overcome the long run-times associated with rigorous simulations of these structures.

The fourth application in Chapter 7 investigated the physical phenomena responsible for colors seen in arrays of silver nano-particles. This study required computer simulation because the nano-particles were much smaller than a wavelength and contained many sharp corners and small gaps. The simulations were very large and consequently pulsed methods were used to speed up the analysis.

The fifth and final application in Chapter 8 investigated the physical phenomena at work in the Sub-Wavelength Grating, a dielectric grating that forms a broadband, high reflectivity mirror. This application necessitated computer simulation because the features were on the order of the wavelength. While the structure is easily simulated the guided wave and CROW models do not explain the operation. However, using Bloch waves to find vertical propagation effects gave a more satisfactory physical

picture.

Chapter 2

TEMPEST Version 7

This chapter presents a simulation framework for quantifying optical scattering processes. The framework consists of a combination of the simulator TEMPEST and a collection of supporting Matlab scripts. TEMPEST[73] is a finite difference time domain (FDTD) simulator originally developed to analyze photomasks for the integrated circuit industry. TEMPEST performs the heavy number crunching and was written in C++ for speed. It directly solves Maxwell's equations. Consequently, it can be used in many applications. Matlab scripts perform pre- and post-processing functions that are often quite application specific. Pre-processing involves such things as as constructing bitmap layouts and sources, creating batches of input files, and computing mode profiles. Post-processing analyzes the resulting field patterns to compute reflection spectra, guided wave amplitudes, and resonance patterns.

FDTD as used in TEMPEST has been shown to perform well in accuracy and

efficiency with regard to other methods, such as RCWA (rigorous coupled wave analysis). For example, figure 2.1 shows a comparison of TEMPEST v7 and RCWA on reflected orders from a plasmonic grating with a variable depth. The test geometry is a gold grating ($n = 0.803 + j1.818$) with a 500 nm pitch and a 50% duty cycle illuminated at a wavelength of 500 nm and an angle $30°$ off normal incidence. Two grating orders are present: the 0th and the -1st. The solid line shows the RCWA predictions, taken from the literature [47], and the dashed lines show the TEMPEST predictions (the additional solid line is a perfect conductor model included in the literature). The two methods are in very close agreement, with less than a $0.01\lambda_0$ shift due to depth in the 0th order and a 2% over prediction by TEMPEST in reflectivity of the first order.

(a) Zeroth order.　　　　　　　　　　(b) Minus first order.
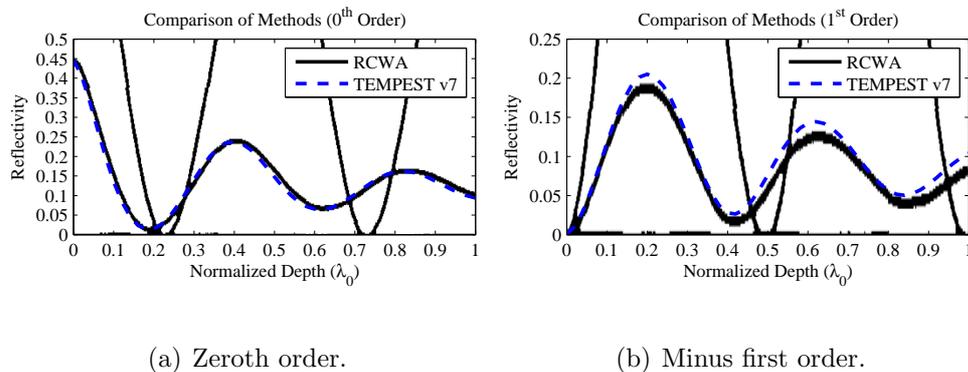
Figure 2.1: Comparison of RCWA and TEMPEST predicitions of reflectivity of the two visible orders from a gold grating.

In improving TEMPEST from v6 to v7 the accuracy has certainly not been degraded and perhaps improved. Past research has shown that TEMPEST v6 simulations matched RCWA predictions [45] [11] and experimental measurements [44] of

the sub-wavelength grating (see chapter 8). Figure 2.2 compares RCWA and TEM-PEST v6 predictions to TEMPEST v7's pulsed method's predictions of the reflection spectrum of this dielectric grating. The important broad reflection peak matches quite well. The simulation conditions of the RCWA and TEMPEST v6 results are not precisely known and give rise to the differences at short and long wavelengths. The differences in the location of the short wavelength peaks are due to differences in cell size (TEMPEST v7 used $10nm$ cells). The other differences are likely due to the thickness of the oxide layer separating the silicon substrate and the grating (TEMPEST v7 used a thickness of 790 nm).
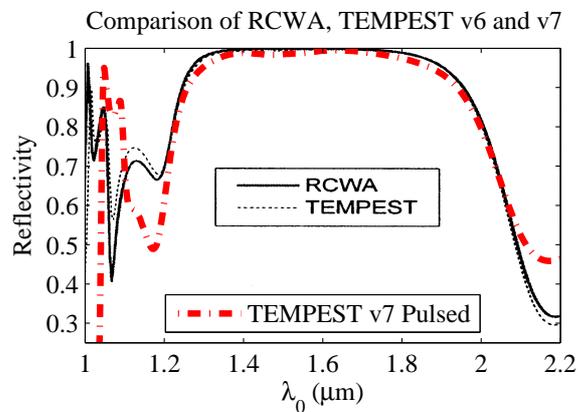
Figure 2.2: Comparison of reflection spectra computed with RCWA, TEMPEST v6, and TEMPEST v7.

Other improvements in TEMPEST v7 have yielded speed and accuracy improvements in particular applications. For example, improvements in the perfectly matched layers (PML) removed cross-talk between neighboring surface plasmon scattering features (see chapter 6) and the scattered field source enabled fast surface plasmon cou-

pling simulations without non-physical PML induced diffraction (see section 2.2.2). TEMPEST v7's pulsed methods enabled the measurement of an entire reflection spectrum from a single simulation of nano-octahedra, significantly reducing the number of long 3D simulations required (see chapter 7).

Some of these improvements came with a cost. At a number of points trade-offs were made in favor of precision, accuracy, and coding convenience at the price of larger memory requirements and longer run-times. For example, the floating-point precision was increased from single to double and the refractive index tables were made finer grain, doubling TEMPEST's memory requirements. Additionally, the C++ code was modularized for convenience at the price of speed. Increases in computer speed and memory are expected to diminish the cost of these trade-offs.

2.1   Internal Structure of TEMPEST

TEMPEST v7 was a complete rewrite of the source code motivated by the need to apply TEMPEST to a number of new applications. The FDTD method is quite general and many of the commonly used features were implemented in version 6. These features were directly transfered to version 7. The code was re-written to organize it and facilitate the addition of new features. It was organize into three distinct modules that are separated by standardized interfaces (figure 2.3). The numerical core handles the number crunching during the simulation. The parser uses an open source tool called ANTLR[52] to analyze input files and setup simulations. The object generators

are the glue between the parser and the numerical core. They are the routines that know how to create sources, plots, materials, etc. The parser calls them to create the required elements of the simulation and then plugs them into the numerical core.
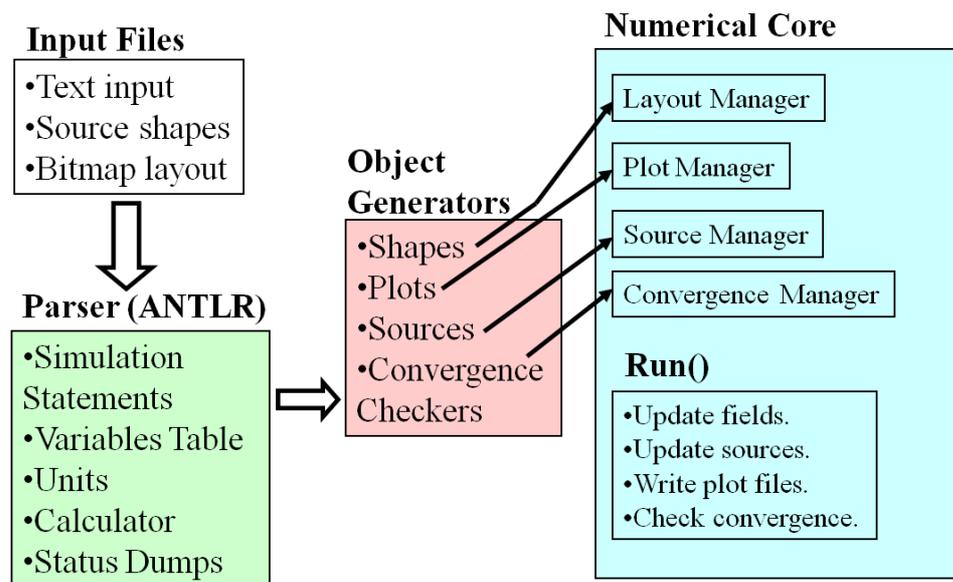
Figure 2.3: The major modules in TEMPEST v7's C++ code.

An example of a standardized interface is shown in the code snippet below. This is the interface for the sources, implemented as an abstract class in C++. This interface enables the numerical core to use different types of sources (e.g. sinusoidal or pulsed) without knowing the exact details of the source.

```
class Source
{
private:
bool active;

public:
void SetActive(bool a) { active = a; };
```

```
bool IsActive() const { return active; };

Source(){ active = false;};
virtual ~Source(){ active = false;};
virtual void Update(INT time_step) = 0;
virtual void Close() = 0;
virtual void Print(std::ostream &out) const = 0;
};
```

This interface contains three important functions: Update, Close, and Print. The Update function is called by the numerical core before each time step and it injects the source's current into the field matrices. The Close function shuts down the source at the end of the simulation. The Print function dumps the source's internal state to an output file to help the user debug the simulation.

A data management system was introduced to facilitate debugging simulations and code updates, save the simulation state for future reference, and to reduce numerical noise in simulation. This system includes commands in the input file to dump the internal state of the simulator into text files so that the user can see what sources, materials, layouts, and variables the simulator created during initialization. These same files contain the coefficients from the update equations used by TEMPEST during simulation. Access to these coefficients enables suppression of numerical dispersion in scattered field sources (section 2.2.2) and surface plasmon sources (section 2.2.1). The parser also supports variables and mathematical expressions in input files enabling the user to easily vary simulation parameters and reducing the chance of errors caused by forgetting to update related parameters. Additionally, the variables

have units enabling TEMPEST to automatically perform unit conversions, such as wavelength to frequency.

2.2   Guided Wave Features

Modeling guided wave devices required implementing new capabilities in TEMPEST v7 or extending capabilities found in version 6. Many of the new and extended capabilities were borrowed from the literature; however, two specialized sources were developed during the course of this research and are discussed in the following sections.

One of the most important features borrowed from the literature was an improved perfectly matched layers absorbing boundary condition. The PML in TEMPEST has three improvements: better absorbtion of evanescent fields, matching to lossy materials, and matching to inhomogeneous waveguiding structures. The improved performance around evanescent fields comes from a new formulation called Convolution PML developed by Roden and Gedney[55]. Their improved formulation also simplifies the implementation of PML matched to many different types of materials, notably the lossy materials found in surface plasmon simulations. Terminating waveguiding structures was enabled by modifying method of scaling the PML's absorbtion coefficients to preserve the impedance across interfaces between different media, as discussed by Taflove[61]. For example, surface plasmon simulations typically contain half-spaces of metal that run all the way to the edge of the simulation (see

figure 2.8(b)). Separate blocks of PML are created to match the metal and the free-space. If theses blocks are not constructed appropriately their interface will reflect surface plasmon waves. Figure 6.2 shows the fields from a surface plasmon simulation with properly constructed PML boundaries. As expected, the PML reflections are minimal, exhibited by the lack of standing waves near the boundaries.

Pulsed methods were critical for computing the optical response of nano-octahedra (chapter 7) and the sub-wavelength grating (chapter 8). Enabling pulsed simulations required new sources, new materials, and new plots. The new sources included impulse sources and modulated Gaussian sources. The former were plagued by large amounts of high frequency noise so the studies discussed later used the latter source. The material models implemented in TEMPEST v6 matched a lossy material's dielectric response at only one frequency; however, to accurately compute a broadband optical response the material's dielectric response must be accurate across the band. The literature contains many examples of broadband material models and the coefficients required to reproduce specific materials' dielectric responses. For this work a commonly used model of silver was employed[17]. The new plots facilitate saving the field distribution at every time step, instead of at only the end of the simulation.

2.2.1 Surface Plasmon Source

Sources in TEMPEST are phased arrays of line currents (essentially antennas) and the pattern determines the type of wave launched. Two surface plasmon sources were

developed in the course of this work: the longitudinal source and the transverse source (figures 2.4 and 2.5). The array of currents in the longitudinal source was placed parallel to the metal surface and phased to match the surface plasmon wavelength. The array of line currents in the transverse source was placed perpendicular to the surface and given the amplitude and phase profile of the surface plasmon mode. The longitudinal source was not used in the studies discussed later because it launched free-space radiation in addition to surface plasmons. This free-space radiation was caused by the finite length of the source, which introduced radiating sidelobes. This problem could be reduced by lengthening the source (and even adding a smooth taper) but this comes at the cost of more memory and longer run-times. The transverse source fared much better in this regard because the surface plasmon mode profile is localized to the metal surface and decays exponentially with a $1/e$ distance of less than one wavelength on silver surfaces in the visible spectrum. Therefore the source could be terminated where the plasmon tail was nearly zero. The transverse source was used in the studies discussed in chapter 6.

Illuminating tall structures with surface plasmons is difficult because surface plasmon tails decay exponentially and can be dominated by low levels of non-physical radiation introduced by source imperfections. The non-physical radiation can be reduced by impedance matching the source so that the portion in the metal and the portion in the air launch waves with matching amplitudes and phases. This section derives the radiation impedance equations used by TEMPEST v7 to relate the desired
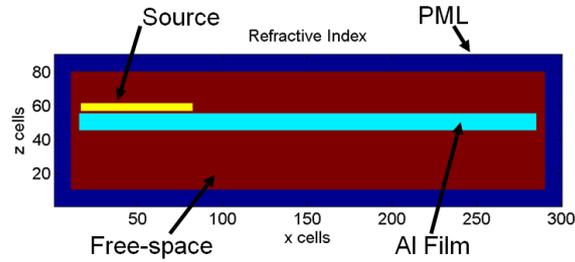
(a) Layout.
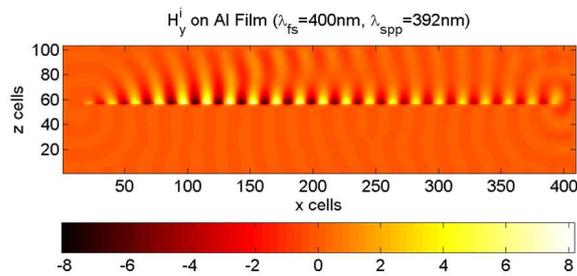
(b) Instantaneous H_y.

Figure 2.4: Layout and fields produced by a longitudinal surface plasmon source on a thin aluminum film ($\lambda_0 = 400$ nm, $n = 0.49 + j4.86$).

complex wave amplitude to the required complex source amplitude for the general case of complex wavenumbers (i.e. the decaying and propagating waves characteristic of surface plasmons). These equations come from the update equations used in TEMPEST and therefore automatically account for some numerical dispersion.

This section considers the case of the TM_{xy} fields (E_x, E_y, H_z) used in the surface plasmon work discussed later. For the non-magnetic materials considered in this work, TEMPEST uses the standard semi-implicit Yee update equation[76][61] for H_z
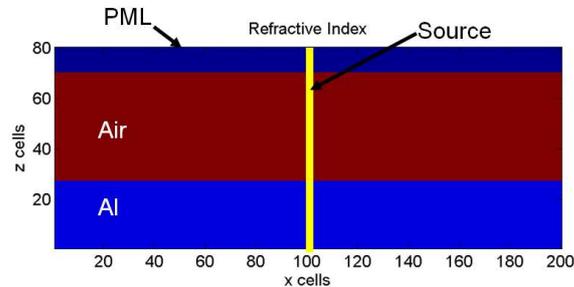
(a) Layout.
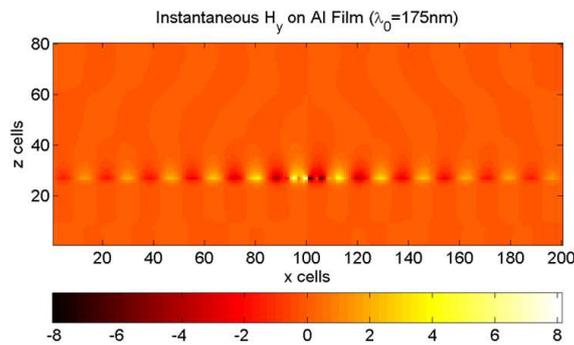
(b) Instantaneous H_y.

Figure 2.5: Layout and fields produced by a transverse surface plasmon source on a half-space of aluminum ($\lambda_0 = 175$ nm, $n = 0.0923 + j1.956$).

(2.1).

$$H_z^{t+1} = a_h H_z^t + b_h \left(E_x^{y-1/2} - E_x^{y+1/2} + E_y^{x+1/2} - E_y^{x-1/2} \right) + s_{hz}^{t+1/2} \delta(y - y_0), \quad (2.1)$$

where t is the time-step number and (x, y, z) is the location of H_z in cells. s_{hz} is the source, the source plane is located at $y = y_0$, and $\delta(y - y_0)$ is a Kronecker delta function that limits the source to an xz plane. a_h and b_h are material dependent coefficients derived from Faraday's Law. For the non-magnetic materials considered here $a_h = 1$ and $b_h = \frac{\Delta t}{\mu_0 \Delta x}$, where Δt is the time step size, Δx is the cell size, and

μ_0 is the permeability of free-space. TEMPEST uses a staggered grid in space and time causing the field components to be located half a cell and half a time step apart. This is indicated by the $(x \pm 1/2)$, $(y \pm 1/2)$, and $t + 1$ terms.

Next the update equation is converted to the frequency domain (2.4) with a sign convention of $e^{j(\vec{k}\cdot\vec{x}\Delta x - \omega t \Delta t)}$. For example, $H_z^{t+1} \rightarrow H_z e^{-j\omega(t+1)\Delta t}$ and $E_y^{x-1/2} \rightarrow E_y^{jk_x(x-1/2)\Delta x}$. Each of these corresponds to a single term in a discrete Fourier transform (DFT) when the wavenumber (or frequency) is real and a single term in a Z-transform when the wavenumber is complex. Special care must be taken of the E_x term because the wave above the source is propagating in the opposite direction as the wave below the source so both k_y and E_x change signs (this is the $N-1$ method):

$$E_x^{y+1/2}(t) \quad \rightarrow \quad E_x^{\uparrow}(\omega)e^{jk_y\Delta x/2} \tag{2.2}$$

$$E_x^{y-1/2}(t) \quad \rightarrow \quad E_x^{\downarrow}(\omega)e^{j(-k_y)(-\Delta x/2)} = -E_x^{\uparrow}(\omega)e^{jk_y\Delta x/2}, \tag{2.3}$$

where $E_x^{\uparrow}$ represents the wave traveling in the $+y$ direction and $E_x^{\downarrow}$ represents the wave traveling in the $-y$ direction. Making these substitutions, the update equation becomes:

$$\left(e^{-jw/2} - a_h e^{+jw/2}\right) H_z \quad = \quad b_h \left[2e^{+jK_y/2}E_x^{\uparrow}+\right.$$

$$\left.\left(e^{+jK_x/2} - e^{-jK_x/2}\right) E_y\right] + s_{hz}, \tag{2.4}$$

where $w = \omega\Delta t$ is the normalized angular frequency and $K_x = k_x\Delta x$ and $K_y = ky\Delta y$ are the normalized wavenumbers. The $e^{j(K_x x + K_y y - w(t+1/2))}$ common to all terms has been removed.

Defining two wave impedances (2.5) and (2.6), we arrive at an expression for the complex wave amplitude H_z in terms of the complex source amplitude s_{hz} (2.7).

$$Z_y = \frac{E_x}{H_z} \tag{2.5}$$

$$Z_x = \frac{E_y}{H_z} \tag{2.6}$$

$$\frac{s_{hz}}{H_z} = \left(e^{-jw/2} - a_h e^{+jw/2}\right) - b_h \left(2e^{jK_y/2}Z_y - j2sin(K_x/2)Z_x\right). \tag{2.7}$$

The wave impedances depend on material. Two materials appear in surface plasmon simulations, free-space and metal, and each use a different update equation.

In TEMPEST, the electric fields in free-space are modeled with the semi-implicit Yee update equation as well.

$$E_x^{t+1/2} = a_e E_x^{t-1/2} + b_e \left(H_z^{y+1/2} - H_z^{y-1/2}\right) \tag{2.8}$$

$$E_y^{t+1/2} = a_e E_y^{t-1/2} + b_e \left(H_z^{x-1/2} - H_z^{x+1/2}\right), \tag{2.9}$$

where a_e and b_e are coefficients that describe the topography and are derived from Ampere's Law. For free-space they are given by $a_e = 1$ and $b_h = \frac{\Delta t}{\epsilon_0 \Delta x}$, where ϵ_0 is the permittivity of free-space. Converting these expressions to the frequency domain and solving for the wave impedances yields:

$$Z_y = \frac{E_x}{H_z} = \frac{j2b_e sin(K_y/2)}{e^{-jw/2} - a_e e^{+jw/2}} \tag{2.10}$$

$$Z_x = \frac{E_y}{H_z} = \frac{-j2b_e sin(K_x/2)}{e^{-jw/2} - a_e e^{+jw/2}}. \tag{2.11}$$

The surface plasmon simulations discussed later use the frequency-dependent finite-difference time-domain method[42] for stability when simulating lossy metals.

The update equations include extra recursive terms for the material polarization:

$$E_x^{t+1/2} = a_e \left(E_x^{t-1/2} + P_x^{t-1/2} \right) + b_e \left(H_z^{y+1/2} - H_z^{y-1/2} \right)$$

$$P_x^{t+1/2} = c_e P_x^{t-1/2} + d_e E_x^{t+1/2} \qquad (2.12)$$

$$E_y^{t+1/2} = a_e \left(E_y^{t-1/2} + P_y^{t-1/2} \right) + b_e \left(H_z^{x-1/2} - H_z^{x+1/2} \right)$$

$$P_y^{t+1/2} = c_e P_y^{t-1/2} + d_e E_y^{t+1/2} \qquad (2.13)$$

The a_e, b_e, c_e, and d_e coefficients are complicated functions of the material's complex refractive index[72]. Converting these to the frequency domain and solving for the wave impedances we find:

$$Z_y = \frac{E_x}{H_z} = \frac{j2b_e sin(K_y/2)}{e^{-jw/2} - a_e e^{jw/2} \left(1 - \frac{d_e e^{-jw/2}}{e^{-jw/2} - c_e e^{+jw/2}} \right)} \qquad (2.14)$$

$$Z_x = \frac{E_y}{H_z} = \frac{-j2b_e sin(K_x/2)}{e^{-jw/2} - a_e e^{jw/2} \left(1 - \frac{d_e e^{-jw/2}}{e^{-jw/2} - c_e e^{+jw/2}} \right)} \qquad (2.15)$$

The source relation (2.7) and wave impedances (2.10),(2.11),(2.14),(2.15) are hard-coded into TEMPEST so that the user can launch waves with arbitrary amplitude and phase. These relations produce the desired complex wave amplitudes even when the wavenumbers are complex (decaying waves) and the medium is lossy.

The transverse surface plasmon source is actually a combination of two of these sources: one placed in free-space and one placed in the metal. The two sources work in concert to cleanly launch a surface plasmon wave. Figure 2.6 shows the amplitude (on a log scale) of the H_z component of a surface plasmon wave launched by an impedance matched source. The dark spot $2.8\mu m$ to the right of the source and 666 nm above

the source is caused by interference between the surface plasmon unintentionally introduced by the source. This dark spot represents the point above which the surface plasmon is weaker than the spurious radiation. An earlier, unmatched source that simply used a source amplitude equal to the surface plasmon mode profile instead of (2.7) exhibited an interference spot 184 nm closer to the surface and 228 nm closer to the source. This small improvement from impedance matching is important when studying tall structures. The free-space radiation is most likely caused by a number of factors, including numerical dispersion, the finite length of the source, and interactions with the upper PML. The numerical dispersion arises from the need to compute the surface plasmon mode shape beforehand. In this work the theoretical mode profile was used; however, this differs somewhat from the profile in TEMPEST due to the large numerical dispersion seen on lossy materials (see chapter 6).
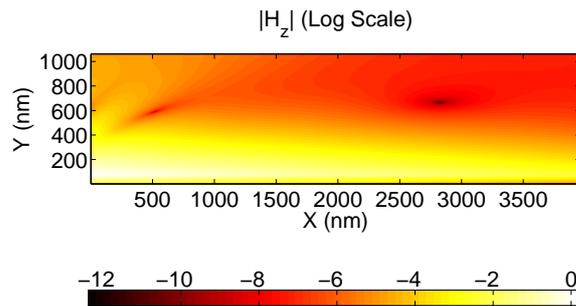
Figure 2.6: The amplitude of the H_z field component of the surface plasmon launched to the right by an impedance matched source ($\lambda_0 = 400$ nm). The data is shown on a log scale. The dark spot in the upper right is caused by unintentional free-space radiation interfering with the surface plasmon.

2.2.2 Scattered Field Source

Illuminating isolated objects with plane-waves is difficult because absorbing boundary conditions must be added to the sidewalls but these interact with the standard plane-wave source causing non-physical diffraction (figure 2.7). Very wide simulations with Gaussian beam illumination are used by other researchers[15] but results in large simulations and long run-times. The literature contains two types of localized sources designed to deal with this problem: total field/scattered field (TF/SF) and scattered field sources. TF/SF sources[61] introduce light on an artificial surface surrounding the object. Scattered field sources introduce the light as a polarization field in the object's volume. They were first developed to compute radar scattering cross-sections of aircraft[20][21]. TEMPEST v7 uses the scattered field source because it is straight-forward to implement for features near, or even embedded in, a substrate. This implementation of the scattered field source is derived directly from TEMPEST's update equations to avoid numerical dispersion effects.

Computations with a scattered field source are similar to using perturbation theory. Perturbation theory is a two step process. First, an incident wave is calculated along with the polarization it induces in the scattering object. Second, re-radiation from the induced polarization is computed. Perturbation theory is only approximate because it neglects self-interaction terms. TEMPEST is fully rigorous and the scattered field source includes these self-interaction terms.

Simulation with the scattered field source is also a two step process (figure 2.8).
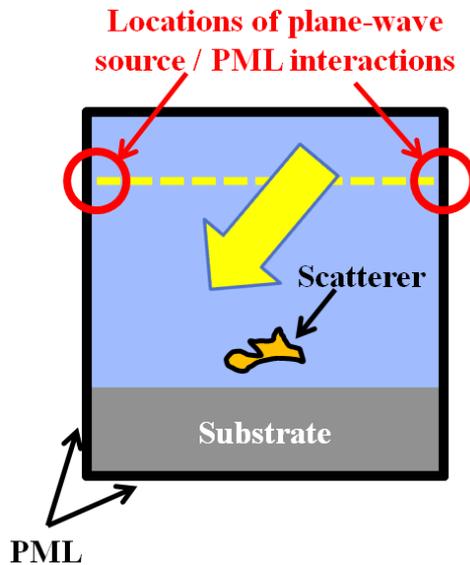
Figure 2.7: Simulation layout with plane-wave / PML interaction problems.

First, the incident wave is computed without the scattering object present and the polarization induced in the object is computed during post-processing. Second, the induced polarization and the object are loaded into a TEMPEST simulation and the re-radiation (including self-interaction terms) is computed. Typically the incident wave calculation is very quick. In simple cases in can be done analytically; however, it was performed with a fast 1D TEMPEST simulation in the work discussed later to avoid problems with numerical dispersion. The benefit of this two step process is that the absorbing boundaries on the sidewalls can be omitted during the first simulation, enabling clean plane-wave illumination (figure 2.8(a)). The absorbing boundary conditions can be inserted into the second simulation because the scattered field source is localized to the scattering object (figure 2.8(b)).

(a) Incident wave simulation.

(b) Scattered field simulation.
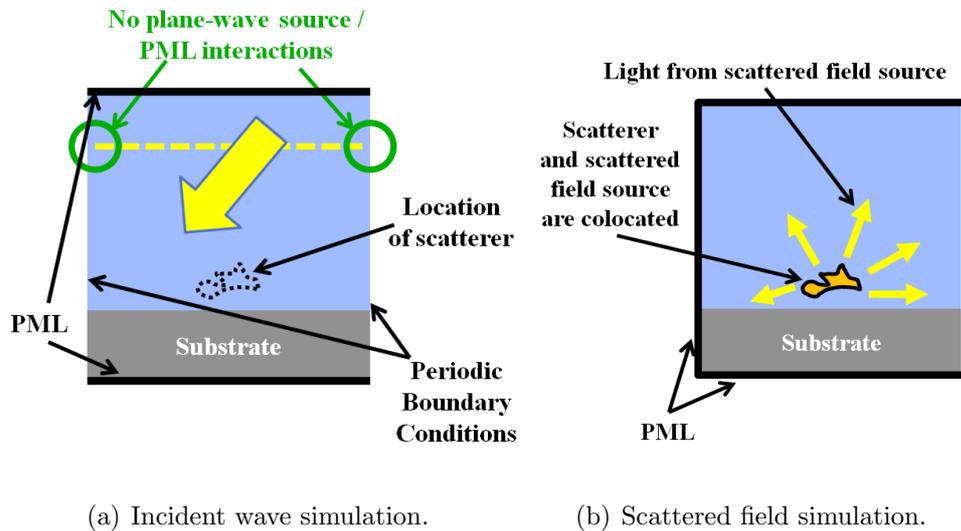
Figure 2.8: Layouts of the two simulations used in the scattered field source process.

This implementation of the scattered field source used in the surface plasmon studies avoids numerical dispersion issues by computing the induced polarization with the coefficients used in TEMPEST's update equations, much like the surface plasmon source. The scattered field source equations are derived directly from the update equations and rely on the linearity of Maxwell's equations. For example, the steady-state version of the E_x update equation (2.8) in the total simulation has the form:

$$E_x e^{-jw/2} = a_e(x,y)e^{+jw/2}E_x + b_e(x,y)\left(H_z^{y+1/2} - H_z^{y-1/2}\right) + s_{PW}(x)\delta_{y_{src}}, \quad (2.16)$$

where $a_e(x,y)$ and $b_e(x,y)$ are the coefficents that describe the topography and $s_{PW}(x)\delta_{y_{src}}$ is the plane-wave source. The source is uncovered by first splitting the total fields into incident and scattered fields: $E_x = E_x^{inc} + E_x^{scat}$ and $H_z = H_z^{inc} + H_z^{scat}$.

Substituting these into (2.16) and collecting terms:

$$\left[E_x^{inc}\left(e^{-jw/2} - a_e(x,y)e^{+jw/2}\right) \quad - \quad b_e(x,y)\left(H_z^{inc}|^{y+1/2} - H_z^{inc}|^{y-1/2}\right)\right.$$

$$- \quad s_{PW}(x)\delta_{y_{src}}\left] + \right.$$

$$\left[E_x^{scat}\left(e^{-jw/2} - a_e(x,y)e^{+jw/2}\right) \quad - \quad b_e(x,y)\left(H_z^{scat}|^{y+1/2} - H_z^{scat}|^{y-1/2}\right)\right]$$

$$= \quad 0. \tag{2.17}$$

Finally, the incident topography is defined as the total topography without the scattering object. It is described by different update coefficients $a_e^{inc}(x,y)$ and $b_e^{inc}(x,y)$. The total and incident topographies are linked by:

$$a_e(x,y) \quad = \quad a_e^{inc}(x,y) + \Delta a_e(x,y) \tag{2.18}$$

$$b_e(x,y) \quad = \quad b_e^{inc}(x,y) + \Delta b_e(x,y), \tag{2.19}$$

where $\Delta a_e(x,y)$ and $\Delta b_e(x,y)$ are the missing scattering object.

Substituting this topography decomposition into the first line of (2.17) and grouping terms we get the update equations for the incident and scattered simulations:

$$\left[E_x^{inc}\left(e^{-jw/2} - a_e^{inc}(x,y)e^{+jw/2}\right) \quad - \quad b_e^{inc}(x,y)\left(H_z^{inc}|^{y+1/2} - H_z^{inc}|^{y-1/2}\right)\right.$$

$$- \quad s_{PW}(x)\delta_{y_{src}}\left] + \right.$$

$$\left[E_x^{scat}\left(e^{-jw/2} - a_e(x,y)e^{+jw/2}\right) \quad - \quad b_e(x,y)\left(H_z^{scat}|^{y+1/2} - H_z^{scat}|^{y-1/2}\right)\right.$$

$$\left. - \left(\Delta a_e(x,y)e^{+jw/2}E_x^{inc} \quad + \quad \Delta b_e(x,y)\left(H_z^{inc}|^{y+1/2} - H_z^{inc}|^{y-1/2}\right)\right)\right]$$

$$= \quad 0. \tag{2.20}$$

The terms in square brackets are the update equations for the first and second simulations respectively.

The scattered field source is the last term in the left side:

$$s_{SF}(x,y) = \Delta a_e(x,y)e^{+jw/2}E_x^{inc} + \Delta b_e(x,y)\left(H_z^{inc}|^{y+1/2} - H_z^{inc}|^{y-1/2}\right). \qquad (2.21)$$

This source uses the steady-state fields from the first simulation, E_x^{inc} and H_z^{inc}, to drive polarization in the second simulation. It is localized to the scattered object because $\Delta a_e(x,y)$ and $\Delta b_e(x,y)$ are localized to the source. A second, almost identical equation is used for the scattered field source in the E_y field component. This source is used in only $\vec{E}$ because no magnetic materials were used in these studies.

The scattered field source was validated against more standard TEMPEST simulations employing the surface plasmon source. The test geometry was a rounded silver bar 360 nm across with a corner rounding radius of 90 nm and suspended 170 nm off a smooth silver surface. This geometry was illuminated with a surface plasmon wave from the left at a wavelength of 700 nm. To compute the standard TEMPEST result, two simulations were run. The first simulation was used to compute only the incident surface plasmon wave and contained only the smooth silver surface and a surface plasmon source (on the left edge). The second simulation included the suspended bar and computed the total fields. The scattered fields were computed by taking the difference of the two simulations, shown in figure 2.9(a). The scattered field source simulation used a scattered field source constructed from the fields in the incident wave simulation, so any non-physical scattering from the surface plasmon

source affected the standard and the scattered field source simulations equally. The fields from the scattered field source simulation are shown in figure 2.9(b). The scattered fields from the standard and scattered field simulations differed in magnitude by only 0.3%. Including phase increased the differences to 3.6%; however, most of this error likely comes from small phase problems in the C++ code for the source that were subsequently fixed.
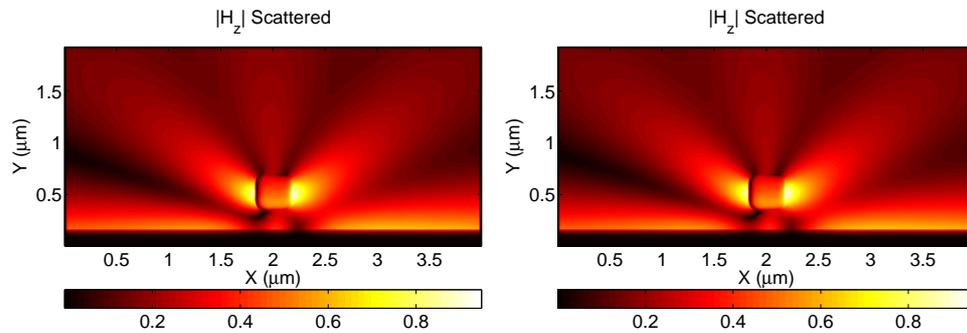
(a) Difference of simulations (standard method).

(b) Scattered field source.

Figure 2.9: Amplitude of the H_z field scattered by a rounded bar computed in two different ways to verify the scattered field source ($\lambda_0 = 700$ nm).

2.3   Conclusion

The TEMPEST v7 system contains many extensions for modeling guided wave systems, including surface plasmon sources and scattered field sources. The new data management features enable these sources to access the coefficients used internally in the update equations and thus avoid noise from numerical dispersion. The following

chapters discuss quantitative anaylses of five optical applications. The first application, vector edge scattering in the Terrestrial Planet Finder Coronagraph (chapters 3 and 4), motivated the development of TEMPEST v7 and the surface plasmon coupling project. The surface plasmon coupling study used the two new sources and the improved PML (chapter 6). The pulsed methods enabled the mapping of resonances in the sub-wavelength grating (chapter 8) and the computation of reflection spectra from arrays of silver nano-octahedra (chapter 7). An increase in the floating point precision enabled the extremely fine computational grids required to analyze optical coupling from the Laser Spike Annealing system (chapter 5).

Chapter 3

Terrestrial Planet Finder

Coronagraph: Methods

This chapter presents methods for computing vector effects in pupil masks designed for the Terrestrial Planet Finder Coronagraph (TPF-C), quantifying their severity, and computing corrections for the Integrated Telescope Model. The TPF-C is designed to image Earth-like planets orbiting other stars[70]; however, electromagnetic scattering from edges and corners its pupil masks can generate enough stray starlight to hide the planet signal. TEMPEST was used to rigorously calculate thick mask transmission patterns for a variety of mask structures. These transmission patterns were compared to the intended thin mask, scalar transmission patterns in the near-field to extract the vector edge effects. Post-processing scripts were developed to accurately perform this comparison and distill the vector edge effects into equivalent

thin mask edge biases. These calculations were complicated by numerical dispersion.

The following chapter discusses the library of mask geometries that were analyzed with this methodology and assess their impact on the planet finding capabilities of the full system.

3.1 Technical Challenges and Team Structure

The chief technical problem faced by the TPF-C is overcoming the extreme contrast between star light and the planet light. In the visible region of the spectrum in which the TPF-C will operate (approximately $\lambda = 500 - 800nm$) terrestrial planets are predicted to be 10 billion times dimmer than their parent stars. A number of different mechanisms have been proposed to adequately suppress the start-light, chief among them pupil remapping[18], focal plane masks[33], and pupil plane masks[66]. This work focused on the latter.

Accurate modeling of pupil mask performance to the 10^{-10} level was extremely challenging. The task was split into three concurrent projects, each addressing major difficulties:

1. Pupil mask pattern design (Princeton University),

2. Rigorous electromagnetic edge modeling (U. C. Berkeley), and

3. Rapid full telescope modeling (Ball Aerospace).

Pupil plane masks act as spatial frequency filters that alter the point spread function (PSF) of the telescope to create dark planet search regions in the habitable orbits around the target star (Fig. 3.1). Mask patterns are created with an optimization program that can engineer the tradeoffs between throughput, angular size of the planet search region, and inner working distance. Many different mask patterns have been created to strike different balances in these tradeoffs[66][67][30].
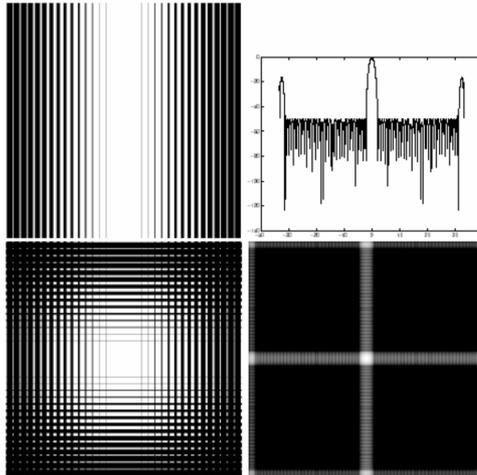
Figure 3.1: The barcode pupil mask, one of the mask patterns designed for the TPF-C. Clockwise from upper left: single mask pattern (white denotes opening and black denotes material), cutline through point spread function (PSF) of the single mask, PSF of dual mask configuration (dark quadrants provide necessary levels of starlight suppression), dual mask pattern. Image courtesy of Princeton University.

The chief shortcoming of the the mask pattern design software is that it does not accurately model edge scattering in thick mask structures. Edge scattering is negligible in many optical models; however, the extreme contrast ratio necessary for the TPF-C make accurate modeling of these effects necessary. TEMPEST was used to rigorously model transmission through thick mask structures and quantify the edge

scattering. FDTD simulations of the full $10cm$ diameter pupil masks are beyond the capabilities of current computers. Therefore, TEMPEST was used to compute scattering from individual edges and generate corrections that were stitched onto the edges of the mask pattern in the full telescope model.

The Integrated Telescope Model[40] (ITM) was used to assess the TPF-C's planet finding capabilities. It is a full-system simulator of the TPF-C written in Matlab[1] that models not only the telescope's optical systems (fig. 3.2) but its thermal and mechanical properties as well. It was used to test different pupil mask patterns and the impact of vector edge effects on these patterns. The optical code uses thin mask and scalar diffraction models for speed.

The vector edge effect modeling proceeded in two steps. First, the edge effects were quantified using TEMPEST. Second, the edge effects were distilled into corrections for the ITM.

3.2 Characterizing Vector Effects in the Near-Field with TEMPEST

Vector edge effects are the additional scattering processes caused by thick mask structures that are not captured by thin mask scalar diffraction models. These effects include reflection off thick mask sidewalls, corner scattering, and polarization and ma-

---

[1]MATLAB is a trademark of The Mathworks
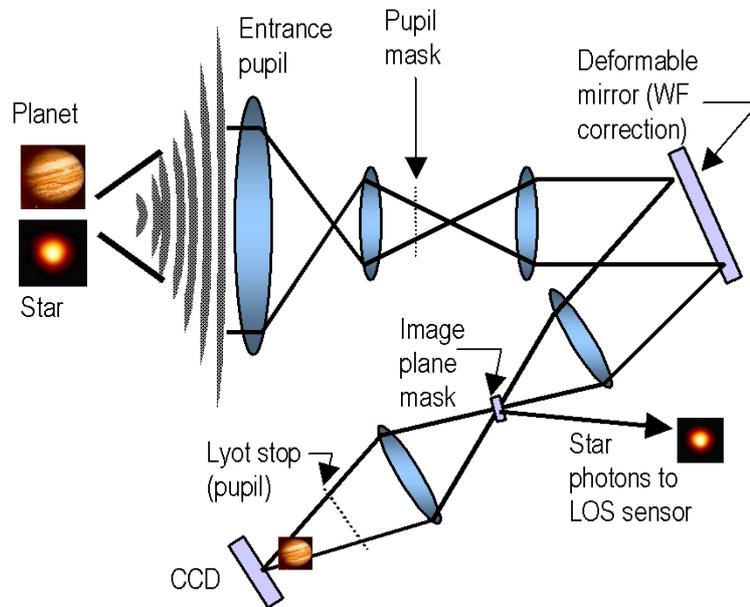
Figure 3.2: Simplified schematic of the optical portion of the Integrated Telescope Model of the Terrestrial Planet Finder Coronagraph. Image courtesy of Ball Aerospace.

terial dependent scattering. To quantify vector edge effects one must know both the thick mask transmission pattern and how that differs from the thin mask transmission pattern. The difference between the two transmission patterns can be quite large, as shown in figure 3.3, and these differences are the stray light scattered by the vector edge effects. Vector edge effects are quite important in photolithography[1][64][46] where mask openings are small and mask patterns are extremely complicated.

In this work, TEMPEST was used to compute thick mask transmission patterns. TEMPEST layout capabilities are quite flexible and enable modeling of many mask topographies. The background was modeled as free-space and the mask material was defined by rectangles. Undercut was modeled with free-space-filled triangles. More

(a) Layout.   (b) Thin mask. (c) Thick mask, (d) Thick mask,

TE.   TM.
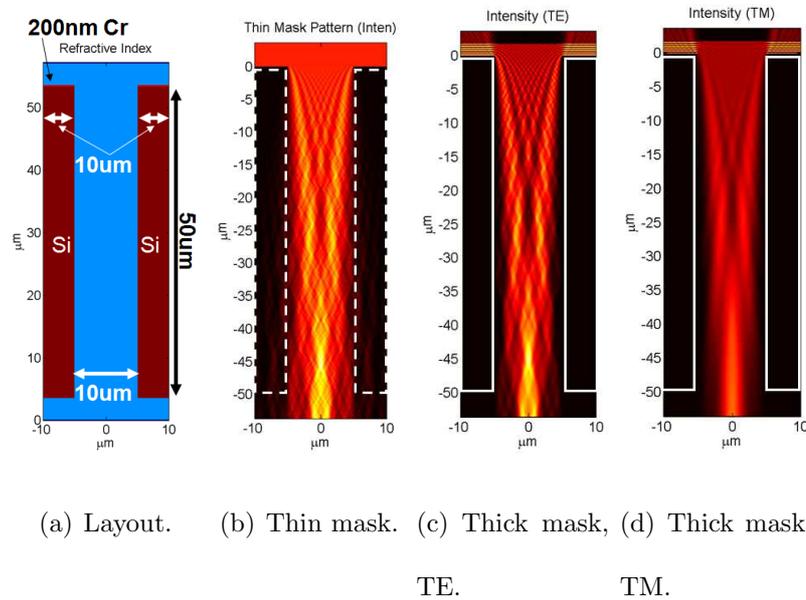
Figure 3.3: A comparison of vector thick mask (figures 3.3(c) and 3.3(d)) and scalar thin mask (fig. 3.3(b)) diffraction patterns. These images show cross-sections through a mask opening. Notice that the thin-mask fields diffract into the thick mask regions (outlined by rectangles), a process that is not physically allowed. The thick mask diffraction patterns are dark near the sidewalls. This difference around the sidewalls is the stray-light light from vector edge effects.

complex geometries were created by adding different geometrical primitives (such as thin rectangles for metal top-coats and cylinders for manufacturing roughness on the sidewalls) and through bitmap layouts (used for 3D manufacturing roughness). Light was introduced by polarized plane-wave sources placed above the mask. In this work, H_y was tracked for TM fields and E_y was tracked for TE fields (where y is the direction out of the page). An example of light moving through a typical simulation of an opening is shown in figure 3.4. As the light propagated downwards it underwent many scattering phenomena which can be seen in the field patterns, such as reflection off the top of the mask, scattering off the corners, reflection off the sidewalls, and
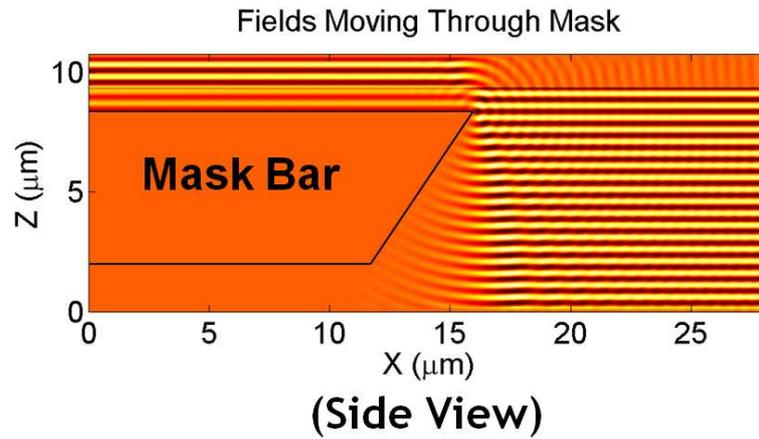
Figure 3.4: The instantaneous field from a TEMPEST simulation of a typical mask opening.

diffraction from the top and bottom apertures. For 2D simulations, the steady-state fields throughout the entire structure were saved at the completion of the simulation for later analysis. 3D simulations were too large to save all the data. Instead, selected slices (cut-planes) were stored for post-processing.

After computing a thick mask transmission pattern in TEMPEST, the analysis shifted into Matlab, which was used to compute the differences between thick and thin mask transmission patterns. The reference plane for comparing the two transmission patterns is arbitrary, but was chosen as the top of the mask to coincide with the mask plane used by the ITM. Consequently, the thick mask transmission pattern had to be referred to the top of the mask. Since the thick mask transmission pattern is the desired output of the thin mask scalar diffraction model used by the ITM, the scalar diffraction model was used to move the thick mask transmission pattern

to the reference plane. The actual movement was performed by a four step process termed back-propagation (fig. 3.5). First, the thick mask transmission pattern was decomposed into a spectrum of plane-waves with a Fast Fourier Transform, FFT, (3.1).

$$F_{bot}(k_x) = \sum_{x=0}^{P} F_{bot}(x)e^{-jk_x x},$$ (3.1)

where $F_{bot}(x)$ is the spatial field pattern at the output plane beneath the thick mask, $F_{bot}(k_x)$ is the spectrum of plane-waves beneath the mask (as a function of the wavenumber in the horizontal direction, k_x), and P is the pitch (in units of cells). Sums are used rather than integrals throughout this work because TEMPEST data is discrete in both real- and Fourier-space. Second, the evanescent orders were removed with a low pass filter to leave only the propagating orders (3.2).

$$
\begin{aligned}
F_{bot}^{prop}(k_x) &= W(k_x)F_{bot}(k_x) \\
W(k_x) &= \begin{cases} 1 & \text{if } |k_x| \leq k_0 \text{ (i.e. propagating orders)} \\ 0 & \text{if } |k_x| > k_0 \text{ (i.e. evanescent orders)} \end{cases}
\end{aligned}, \quad (3.2)
$$

where $F_{bot}^{prop}(k_x)$ is the spectrum of propagating plane-waves beneath the mask, $W(k_x)$ is the low pass filter (window) used to remove the evanescent orders, and k_0 is the wavenumber of free-space. Third, the spectrum of plane-waves was time reversed to move them to the reference plane (3.3).

$$F_{ref}(k_x) = F_{bot}^{prop}(k_x)e^{-j\sqrt{k_0^2-k_x^2}(z_{ref}-z_{bot})},$$ (3.3)

where $F_{ref}(k_x)$ is the spectrum of plane-waves at the reference plane, $k_z = \sqrt{k_0^2 - k_x^2}$ is the wavenumber in the vertical direction, z_{ref} and z_{bot} are the locations of the

reference and output planes respectively, and the minus sign in front of the square root reflects the fact that the waves are time reversed (a sign convention of $e^{-j\omega t}$ is used in this work). Finally, the spectrum of plane-waves was converted into a spatial field pattern at the reference plane using an inverse FFT (3.4).

$$F_{ref}(x) = \frac{1}{P} \sum_{k_x=-k_0}^{k_0} F_{ref}(k_x)e^{jk_x x}, \tag{3.4}$$

where $F_{ref}(x)$ is the spatial field pattern at the reference plane.
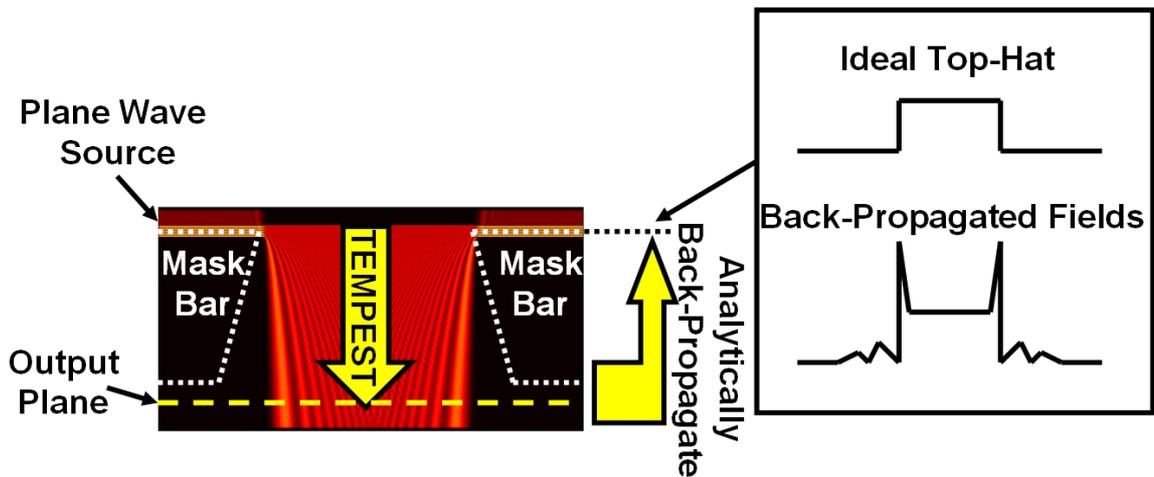
Figure 3.5: Cross-sectional view of the intensity of light diffracting through a pupil mask opening in a TEMPEST simulation (left). A thin-mask transmission pattern, called the "ideal top-hat" (upper right), and a schematic of a thick mask transmission pattern after back-propagation to the reference plane (lower right).

Stray-light from the vector edge effects was computed by subtracting the thin-mask transmission pattern from the thick mask transmission pattern at the reference plane, $F_{ref}(x)$. The thin mask model assumes a binary on-off transmission pattern for the mask, i.e. 100% transmission through openings and zero transmission through

mask material. A stray-light pattern (a.k.a. difference fields) is shown in figure 3.6. In all but the narrowest openings these corrections were strongly localized to the edges because scattering off the top corners and the upper sidewalls was the strongest source of stray-light.

There are two critical assumptions underlying this modeling process. First, vector effects at one edge are assumed to be independent of neighboring edges, so vector edge effects computed in a single TEMPEST simulation of one opening can be used as the basis of corrections for most edges in an entire mask pattern, independent of the opening size. This independence was demonstrated previously [6], but with a prior version of the back-propagation method that neglected phase. This assumption was verified with the final version of the back-propagation method and the edge-to-edge cross-talk was found to be very small. For example, simulations of $32\mu m$ wide, $64\mu m$ wide, and $128\mu m$ wide openings in a 35° undercut Aluminum mask (500nm wavelength, $5.67\mu m$ thick, and 50% duty cycle) show that the total stray-light generated by vector edge effects changed only 3% as the opening size increased. This approximation breaks down for narrow openings ($10\lambda_0$ and narrower) and high aspect ratio openings (e.g. 5:1 height to width) as documented in section 4.3.

Second, TEMPEST simulations of straight edges were assumed to accurately model vector effects on gently curved edges. Many edges in pupil mask patterns are gently curved, with radii of curvature much greater than a wavelength, so it is unlikely that the vector edge effects would differ greatly from those found on straight

edges. However, the curvature changes the local definition of polarization. To account for this the polarized corrections were simply rotated to match the local polarization convention before being applied to mask patterns in the ITM. Sharp 3D corners are much more complicated and require geometry specific simulations, as discussed in section 4.6.

3.3 Equivalent Thin-Mask Model Corrections

To be useful in the ITM, the difference fields must be turned into corrections for thin-mask patterns. The difference fields shown in figure 3.6 cannot be directly added to the ITM because because of a large mismatch in the computational grid densities. The ITM uses $100\mu m$ cells to discretize a $10cm$ mask whereas an entire TEMPEST simulation is on the order of $64\mu m$ across. So the difference fields must be first reduced to a useable form. To be accurate, this data reduction must only preserve the average value of the difference field because the f-number of the TPF-C is so large that only the on-axis order from TEMPEST reaches the image plane. The higher angle orders are scattered out of the system.

Therefore, the difference fields are reduced with the Limited Parameter Model for Low Angular Acceptance Systems. This model converts each spike in the difference fields into a simple box-car (figure 3.7). The height of the box-car is set to one, commensurate with the height of the binary transmission model. The width of the box-car is adjusted to match the average field in one spike of the difference fields (3.5)

and the phase is adjusted to match the phase of the same average field (3.6).

$$W = \frac{1}{2}\left|\frac{1}{N}\sum f(x)\right| \tag{3.5}$$

$$\phi = phase\left(\frac{1}{N}\sum f(x)\right), \tag{3.6}$$

where W is the width of the box-car, ϕ is the phase of the box-car, $f(x)$ represents the complex difference field, N is the total number of cells in one period of the TEMPEST simulation, and the $1/2$ is to account for two edges. Intuitively this box-car is a complex bias that is applied to every edge in a thin-mask pattern. The bias represents the stray-light scattered by vector effects. The phase of the complex bias respresents the phase of the scattered light relative to the un-scattered light. This phase can make the opening appear narrower or wider than the physical width. Since only the on-axis data (DC term) reaches the final image plane the aspect ratio of the box-car can be adjusted as long as the area is preserved.

3.4   Numerical Issues

The strongest numerical issue, numerical dispersion, cropped up in the back-propagation method. Numerical dispersion refers to a numerical shortening of the wavelength inherent to FDTD simulations. It arises from the difference approximation to the derivative operators in Maxwell's equations and slightly changes the dispersion relation of a plane-wave propagating in a uniform medium. Even small errors in the dispersion relation are a large concern in the back-propagation algorithm

because the thick-mask transmission pattern must be moved hundreds of wavelengths from the output plane beneath the mask to the reference plane at the top of the mask. If the wavenumbers used for this movement are even slightly incorrect large phase errors accrue and result in non-physical difference fields across the entire mask opening rather than localized to the corners.

A two-pronged attack was used to combat numerical dispersion. The easiest fix for numerical dispersion was to increase the density of the computational grid. FDTD is second order accurate in cell size, so the errors rapidly diminsh. However, memory requirements and run-times rapidly increase. This work employed a cell density of $53\text{cells}/\lambda_0$ which resulted in a wavelength error of 0.04%. This was not sufficient for accurate back-propagation and a second strategy was employed to further reduce the error.

The second strategy was to use a second simulation to measure the numerical wavelength and use the numerical value in the back-propagation process. In this simulation a single on-axis plane-wave was run through free-space and the wavelength was extracted with a least-squares fit to the phase progression in the direction of propagation. The measured wavelength was used to calculate the numerically correct free-space wavenumber used in the back-propagation process (3.3). This resulted in an order of magnitude reduction in the wavelength error.
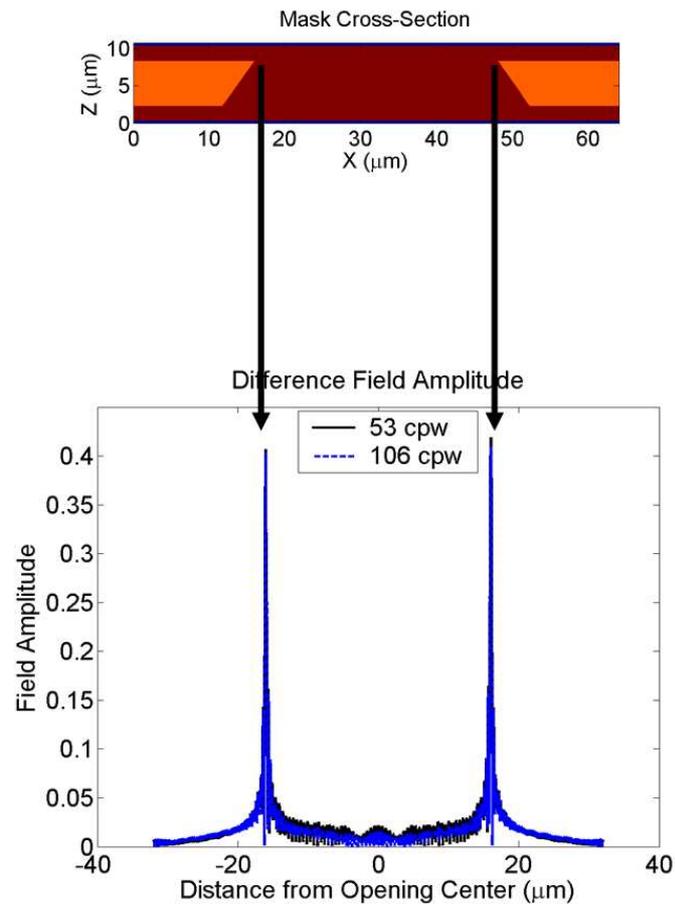
Figure 3.6: The layout of a typical two-dimensional TEMPEST simulation and a set of difference fields resulting from this simulation. In the layout the light orange color represents the mask absorber, dark red represents free-space, and the blue bands at the top and bottom were PML. The mask structure was a $6\mu m$ thick Aluminum mask with $35.3°$ undercut sidewalls, a $64\mu m$ period, and a 50% duty cycle illuminated at 500nm wavelength and TE polarization. The two difference field curves overlaid on top of one another came from two simulations of the same structure with different computational grid densities (53 and 106 cells per wavelength). This demonstrates that the spikes in the difference field were not numerical artifacts.
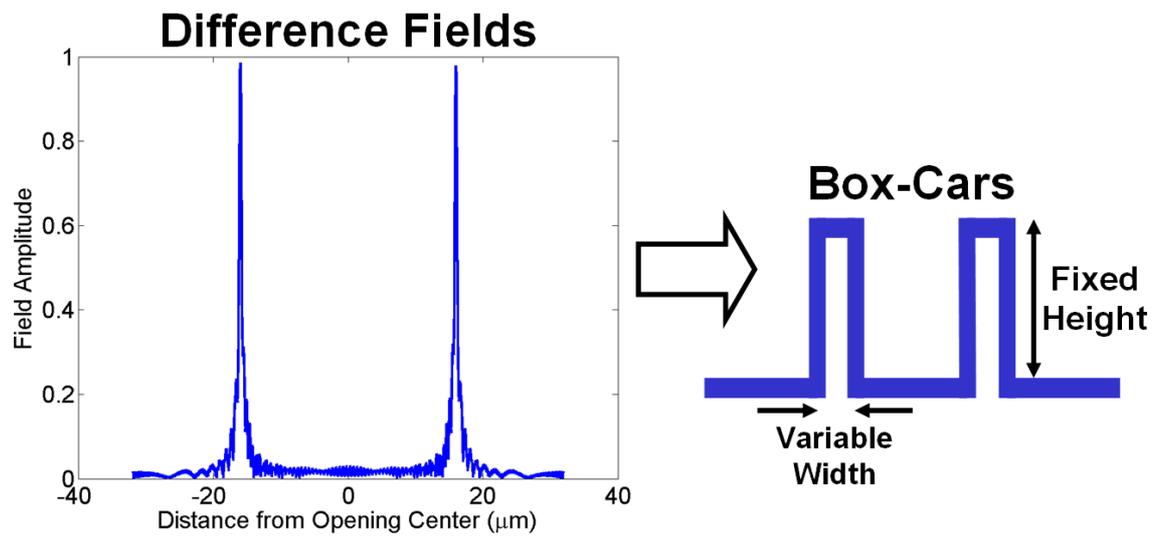
Figure 3.7: A typical set of difference fields (left) and the resulting box-cars (right) use to correct thin-mask patterns in the ITM.

Chapter 4

Terrestrial Planet Finder

Coronagraph: Results

The methodology presented in the previous chapter was used to quantify vector edge effects on numerous mask structures to uncover problems with current mask designs and guide future mask design efforts. This chapter presents the resulting library of vector edge effects and a discussion of their impact on the telescope's planet finding capabilties. The dominant source of stray starlight was diffraction confinement from vertical sidewalls. Narrow openings were also a cause for concern. Initial investigations into 3D mask features shows that TEMPEST is capable of analyzing them with hero runs. Ultimately, vector edge effects may limit the TPF-C from achieving the last order of magnitude of starlight suppression required for terrestrial planet finding.

4.1 Silicon Masks: Effects of Undercut Angle, Polarization, Wavelength, and Chrome Top-Coats

Silicon was chosen as a mask material because of the highly developed processing techniques from the integrated circuit industry. Thick material is desirable for mechanical stability when the telescope is launched into space; however, thick mask sidewalls were predicted to be a strong source of stray-light. Thick sidewalls keep light from spreading, whereas in the thin mask model light is free to diffract into the shadow region beneath the mask. The first round of masks manufactured for the High Contrast Imaging Testbed (HCIT) [65] were very thick. They were designed with vertical sidewalls and a thickness of $50\mu m$, but after manufacturing were found to be $100\mu m$ thick [3].

Polarization was a potential problem because the different boundary conditions from Maxwell's equations could lead to different edge scattering profiles.

Wavelength can change vector effects in two ways: through material properties and diffraction spreading. The rate of diffraction spreading depends on the width of the aperture in wavelengths. As the wavelength increases, the width in wavelengths decreases, leading to more rapid spreading and thus stronger interactions with sidewalls. Material properties refers to the refractive index, which in some materials is a strong function of wavelength. However, in the visible region Silicon has a nearly constant refractive index.

Chrome was investigated as a means to prohibit light leakage directly through the mask material. Silicon becomes transparent near the infrared because the imaginary part of the refractive index decreases to nearly zero. Chrome is opaque; however, its addition could alter edge scattering. TEMPEST is not accurate enough to measure leakage directly through the mask material so these simulations focused on changes in edge scattering. Simple thin-film interference theory predicted that $200nm$ of chrome was enough to reduce leakage to below the 10^{-10} level.

Vector effects on the original mask design were assessed with TEMPEST simulations of $50\mu m$ thick masks with a $96\mu m$ period and a 50% duty cycle. Two sidewall undercuts were studied: vertical ($0°$) and $20°$. Two wavelengths were tested, $630nm$ and $785nm$, which correspond to two lasers available in the HCIT. Both polarization were tested, TE ($\vec{E}$ parallel to the mask edges) and TM ($\vec{H}$ parallel to the mask edges). All the mask designs were tested with and without a $200nm$ chrome capping layer (figure 4.1).

The edge biases from each simulation are shown in table 4.1. Diffraction confine-

λ	Polar.	Vertical Sidewalls		$20°$ Sidewalls	
		No Top Coat	Top Coat	No Top Coat	Top Coat
630 nm	TE	$2.76\lambda, -135°$	$2.77\lambda, -135°$	$0.239\lambda, -85°$	$0.278\lambda, -84°$
	TM	$2.49\lambda, -146°$	$2.50\lambda, -146°$	$0.028\lambda, -157°$	$0.029\lambda, 141°$
785 nm	TE	$2.49\lambda, -134°$	$2.49\lambda, -134°$	$0.23\lambda, -87°$	$0.26\lambda, -84°$
	TM	$2.18\lambda, -146°$	$2.18\lambda, -147°$	$0.023\lambda, 135°$	$0.031\lambda, 104°$

Table 4.1: Box-cars for a single edge of a $50\mu m$ thick silicon mask with a period of $96\mu m$ and a 50% duty cycle. The simulation layout is shown in figure 4.1.
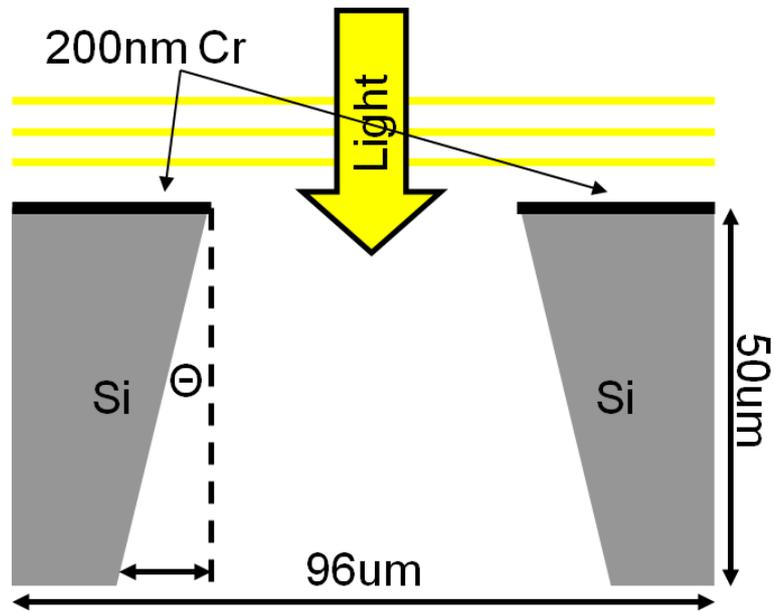
Figure 4.1: Layout of silicon mask simulations. θ refers to the mask undercut angle. Box-cars from these simulations are shown in table 4.1.

ment by the vertical mask sidewalls was the largest source of stray-light. Column two of the table shows that this vector edge effect accounted for over 2λ wide box-cars per edge on a mask with vertical sidewalls. This scattering was reduced by an order of magnitude by introducing a 20° undercut angle (column four). Polarization was only a $\lambda/4$ effect, as can be seen by comparing rows three to four and five to six. This effect is weak because most of the light that hit the sidewalls entered the opening at nearly gazing incidence. These large angles lead to nearly 100% reflection for both polarizations. Wavelength was also a $\lambda/4$ effect, which can be seen by comparing row three to five and four to six. The impact of wavelength was weak because the refractive index of silicon is nearly constant over the spectrum of interest and the

mask opening was very wide in terms of wavelength (leading to negligible changes in diffraction spreading). Thin metal top coats did not appreciably alter the edge scattering, accounting for only a few hundredths of a wavelength change (compare columns three to four and five to six).

4.2   Surface Plasmon Radiation

Surface plasmon polaritons[54] (SPP's) are light waves confined to the surface of a metal that appear under only the TM polarization (for a more detailed explanation see chapter 6). They are a concern on TPF-C diffraction masks because they are strongly polarization dependent and they can cause the bottom corners of the mask to glow.

Surface plasmons are launched at the top corners of the mask. As incident light enters the mask, the top corners scatter some of it into surface plasmons. These plasmons run down the sidewalls and re-radiate at the bottom corners, causing the corners to glow even though they are in the shadow regions.

The plasmon scattering path is complicated and has dependencies on geometry and material. This particular study focused on the material and sidewall angle dependencies of plasmons. A two-step process was developed to isolate the light that traveled through the surface plasmon path.

Each step corresponds to a separate TEMPEST simulation. The first simulation was used to measure the surface plasmon excitation strength. It had the same layout

as the simulations used in the standard back-propagation analysis – a plane-wave illuminated an opening in a thick mask (figure 4.2(a)). However, the simulation was stopped after the surface plasmon first reached the bottom corners of the mask. This was to avoid forming a surface plasmon standing wave pattern, which introduces an additional dependency on the length of the sidewall. The second simulation was used to measure re-radiation at the bottom corner of the mask. It modeled the same geometry but used a pair of surface plasmon sources placed midway down the sidewall to launch surface plasmons at the bottom corner of the mask (figure 4.2(b)).

The surface plasmon excitation strength was measured as the peak of the magnetic field in the middle of the sidewall of the first simulation. The surface plasmon measurement point and the surface plasmon sources were co-located to avoid complications from propagation losses. The re-radiated light was caught 2 μm beneath the mask in the second simulation.
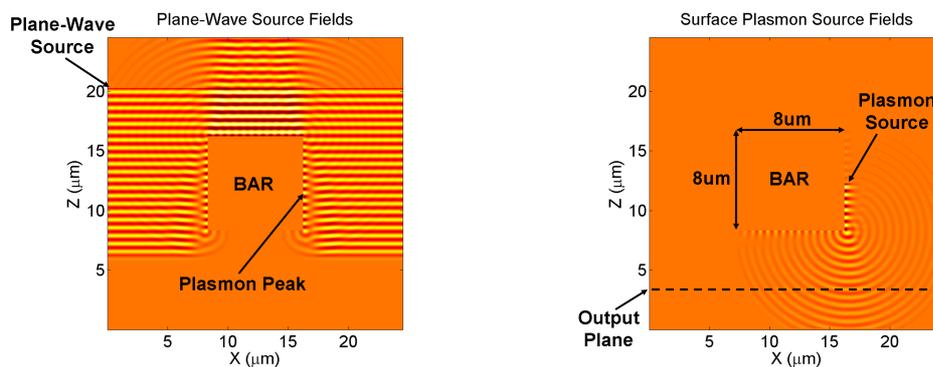
(a) Plasmon excitation simulation.   (b) Plasmon re-radiation simulation.

Figure 4.2: Instantaneous H_z fields from simulations used to quantify the surface plasmon scattering path in TPF-C pupil masks (see table 4.2).

The re-radiated light was converted into a box-car in a similar manner to the standard back-propagation process (3.5). However, modifications were made to account for the surface plasmon excitation efficiency, normalize the surface plasmon source, and account for the fact that only one edge was excited in the second simulation (4.1).

$$W = \frac{\left|H_z^I\right|}{\left|H_z^{II}\right|} \left|\frac{1}{N} \sum H_z^{rad}(x)\right|, \tag{4.1}$$

where $\left|H_z^I\right|$ was the surface plasmon excitation efficiency measured in the first simulation, $\left|H_z^{II}\right|$ was the normalization for the surface plasmon source in the second simulation, N was the pitch (in units of cells), and $H_z^{rad}(x)$ was the light radiated from the bottom corner. The phase of the box-car was not calculated because it depends on the thickness of the mask.

These studies modeled $16\mu m$ openings in $8\mu m$ thick metal masks. Sidewall angles of $0°$, $10°$, and $20°$ were simulated. Silver and Aluminum masks were tested at 670nm as illustrative examples of materials exhibiting strong surface plasmons. Chrome masks were tested at 500nm to illustrate weak surface plasmons.

Table 4.2 shows the box-car widths for the surface plasmon path on these masks. The surface plasmon path on Silver and Aluminum masks with vertical sidewalls was very large, adding one wavelength wide box-cars to each edge of the opening. However, undercutting the mask by $20°$ reduced the light by a factor of three to five. The Chrome mask showed much weaker surface plasmons, with box-car widths less than a fifth of a wavelength for vertical sidewalls.

Ultimately, the surface plasmon path was not important on practical masks for

Undercut Angle	0°	10°	20°
Ag (670 nm)	0.94λ	0.54λ	0.38λ
Al (670 nm)	1.03λ	0.64λ	0.17λ
Cr (500 nm)	0.14λ	—	0.07λ

Table 4.2: Box-car widths for the surface plasmon path on one sidewall.

two reasons. First, surface plasmons decay as they propagate and re-radiation will be much dimmer on the $100\mu m$ thick HCIT masks than on these $8\mu m$ simulated masks. Table 4.3 shows the travel distances of surface plasmons for the materials and wavelengths[58] simulated in table 4.2. All of the travel distances are less than the manufactured mask thickness. Second, surface plasmons do not exist on Silicon, the

Material (Wavelength)	Travel Distance
Ag (670nm)	55.5μm
Al (670nm)	36.3μm
Cr (500nm)	4.7μm

Table 4.3: Surface plasmon travel distances on three metals at two wavelengths. The travel distance is defined as the distance a surface plasmon propagates before its amplitude falls to $1/e$ of its starting value.

material used for the TPF-C masks.

4.3 Sub-20λ Openings

Narrow mask openings are critical to obtaining the last few orders of magnitude in contrast required to image planets; however, narrow openings are heavily impacted by vector effects. Stray-light from vector edge effects can account for a very large

percentage of the light moving through an opening. For example, a $1\lambda/edge$ effect on a $48\mu m$ wide opening introduces only 4% stray-light whereas the same effect on a $10\mu m$ wide opening introduces 20% stray-light. Additionally, cross-talk between edges may be important when openings become narrow. For example, figure 4.3(a) shows the sharp corner spikes characteristic of minimal cross-talk in a wide opening. In contrast, figure 4.3(b) shows difference fields from narrow openings. The corner spikes in these difference fields are not well defined and may contain side-to-side cross-talk.
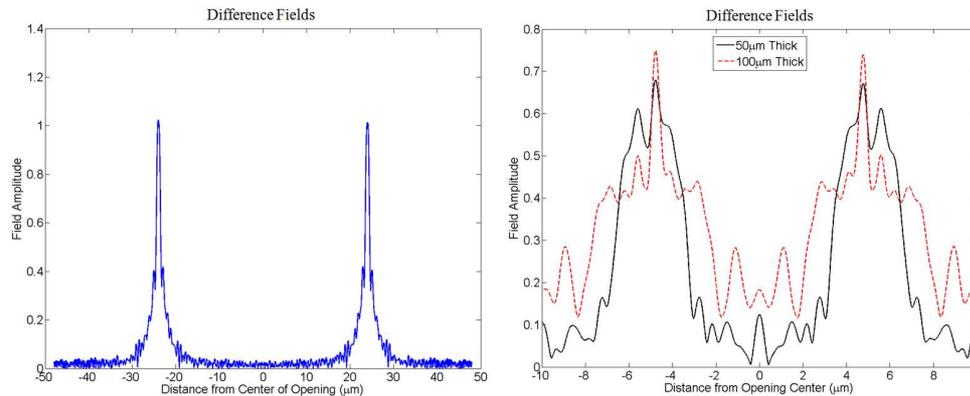
(a) $48\mu m$ wide opening.   (b) $10\mu m$ wide openings.

Figure 4.3: Difference fields from a wide opening (left) and two narrow openings (right) with vertical sidewalls. The wide opening mask and one of the narrow opening masks were $50\mu m$ thick. The other narrow opening mask was $100\mu m$ thick. The TE polarization is shown.

On-axis illumination of three mask geometries was used to study vector effects in narrow openings: a wide $48\mu m$ opening in a $50\mu m$ thick mask, a narrow $10\mu m$ opening in a $50\mu m$ thick mask, and a $10\mu m$ opening in a $100\mu m$ thick mask. All masks

Opening	Thickness	TE	TM
$48\mu m$	$50\mu m$	$2.77\lambda, -135.2°$	$2.50\lambda, -146.2°$
$10\mu m$	$50\mu m$	$2.83\lambda, -129.4°$	$2.66\lambda, -143.1°$
$10\mu m$	$100\mu m$	$3.31\lambda, -132.5°$	$3.47\lambda, -139.5°$

Table 4.4: Box-cars from narrow openings.

were made of silicon, had a 50% duty cycle, and had a $200nm$ chrome top-coat. A wavelength of $630nm$ was used.

Table 4.4 shows the box-cars for the three mask geometries. The width of the box-cars did not change much in the $50\mu m$ thick mask when the opening was narrowed from $48\mu m$ to $10\mu m$; however, the shape of the corner spikes changed significantly (figure 4.3) and the percentage of stray-light increased from 6% to 28%. Increasing the thickness of the mask increased the sidewall interaction length and thus the box-car widths. In this case, doubling the thickness of the mask increased the box-car widths by 17% to 30%, depending on polarization. The electromagnetic width of the opening tended to follow the resolution equation for contact printing:

$$w = 2\sqrt{\lambda t_{mask}}, \tag{4.2}$$

where w was the effective width of the opening, λ was the free-space wavelength, and t_{mask} was the distance from the mask.

4.4 Tilt

Mask tilt may be introduced either intentionally to reduce reflections from the top of the mask back into the optical path or unintentionally due to alignment errors. Mask tilt affects transmission through three mechanisms:

1. Projection shortening, which is a reduction in the light flux through the opening. This effect is modeled in scalar thin-mask theory and is small, accounting for only 0.005% reduction in opening size with a 1.8° tilt (0.008λ reduction on a $10\mu m$ wide opening at $\lambda = 630nm$).

2. Geometrical obstruction by thick mask sidewalls, is a much larger effect, accounting for a 2.5λ opening reduction on a $50\mu m$ thick mask and 5λ opening reduction on a $100\mu m$ thick mask (again with a $10\mu m$ opening at $\lambda = 630nm$).

3. Increased vector effects due to strong illumination of one sidewall. This is especially important in narrow openings.

Mask tilt was simulated in TEMPEST by tilting the incident wavefront rather than the mask. Tilt angles were set by the periodicity requirement for the boundary conditions in TEMPEST v6. The back-propagation algorithm properly accounts for the change in wavenumber caused by off-axis illumination. Both wide and narrow openings were investigated. All masks were made of silicon, had vertical sidewalls and 50% duty cycles, did not have chrome top-coats, and were simulated at a wavelength of $630nm$.

	TE		TM	
Tilt	Left	Right	Left	Right
0°	$2.76\lambda, -135.4°$	$2.76\lambda, -135.4°$	$2.49\lambda, -146.2°$	$2.49\lambda, -146.2°$
0.376°	$2.62\lambda, -132.0°$	$2.96\lambda, -139.5°$	$2.34\lambda, -142.2°$	$2.66\lambda, -150.1°$
1.128°	$2.32\lambda, -124.3°$	$3.40\lambda, -148.6°$	$2.09\lambda, -134.1°$	$3.02\lambda, -158.8°$

Table 4.5: Complex boxcar biases for $48\mu m$ wide openings in a $50\mu m$ thick Si mask at $\lambda = 630nm$. The un-tilted mask numbers have been included for reference.

Table 4.5 shows the box-cars through tilt on a wide $48\mu m$ opening. Phase differences between the left and right box-cars due to tilt in the incident wavefront were removed. The remaining phase differences are due to differences in the vector edge effects excited on each corner. On wide openings, increasing the tilt from 0° to 1.1° increased the box-car width by 20% on the brightly illuminated side (right). On the weakly illuminated side (left) the box-cars decreased by 16% because the tilt moved the edge farther into shadow.

Table 4.6 shows the box-cars for narrow openings, where the vector effects were much larger. For example increasing the tilt from 0° to 1.8° increased the box-car widths by 50% to over 100% depending on thickness. Additionally, the dark (left) sidewalls on the thickest masks actually got brighter when the tilt was increased, probably due to cross-talk from the bright edge.

For small mask tilts, boxcar widths on the highly illuminated sidewalls followed these rules of thumb:

- $\lambda/2$ per degree tilt with wide openings in $50\mu m$ thick masks,

- 1λ per degree tilt with narrow openings in $50\mu m$ thick masks, and

Thick-		TE		TM	
ness	Tilt	Left	Right	Left	Right
50μm	0°	$2.83\lambda, -129.4°$	$2.83\lambda, -129.4°$	$2.66\lambda, -143.1°$	$2.66\lambda, -143.1°$
	1.8°	$2.04\lambda, -136.7°$	$4.43\lambda, -166.9°$	$1.88\lambda, -138.7°$	$3.84\lambda, -173.6°$
100μm	0°	$3.31\lambda, -132.5°$	$3.31\lambda, -132.5°$	$3.47\lambda, -139.5°$	$3.47\lambda, -139.5°$
	1.8°	$4.16\lambda, -117.2°$	$7.43\lambda, -158.7°$	$3.86\lambda, -129.1°$	$7.00\lambda, -167.4°$

Table 4.6: Complex boxcar biases for $10\mu m$ wide openings in $50\mu m$ and $100\mu m$ thick Si masks at $\lambda = 630nm$. The un-tilted mask numbers have been included for reference.

- 2λ per degree tilt with narrow openings in $100\mu m$ thick masks.

4.5   Summary of 2D Effects

The effects investigated with 2D simulations are summarized in figure 4.4. This table highlights the worst structures (narrow openings in a tilted $100\mu m$ thick mask with vertical sidewalls) to the best structures (wide openings in an un-tilted $50\mu m$ thick mask with 20° undercut sidewalls). It also lists the severity of the individual phenomena, with undercut angle being the most important.

4.6   3D Manufacturing Effects

Up to this point, only smooth 1D mask patterns have been discussed; however, most, if not all, masks have roughness and 2D or 3D features. These features come in the form of sharp 2D corners intentionally designed into the mask pattern and roughness accidentally introduced during the manufacturing process. Many of these

Mask Structure:	Severity:
Vertical sidewalls, Off-axis illumination, 10um Opening, 100um Thick	7 λ
Vertical sidewalls, Off-axis illumination, 10um Opening, 50um Thick	4 λ
Vertical sidewalls, On-axis illumination, 10um Opening, 100um Thick	3-4 λ
Vertical sidewalls, On-axis illumination, 10um Opening, 50um Thick	3 λ
Vertical sidewalls, On-axis illumination, 48um Opening, 50um Thick	3 λ
Vertical sidewalls, On-axis illumination, 48um Opening, 50um Thick	2-3 λ
20° undercut sidewalls, On-axis illumination, 48um Opening, 50um Thick	λ / 4
Effects on 50um thick Si masks with 48um openings at 630nm:	Worst case:
Undercut angle	3 λ
Small mask tilt	λ / 2 per degree tilt
Polarization	λ / 4
Wavelength (630nm to 785nm)	λ / 4
200nm Cr capping layer	λ / 100

Figure 4.4: Summary of the mask structures investigated with 2D TEMPEST simulations and the physical phenomena ranked from strongest vector effect to weakest.

2D features require 3D TEMPEST simulations, which tax current computers. The following sections discuss example simulations of manufacturing roughness and 2D mask features.

4.6.1 Sawtooth Edges

Sharp 2D corners appear in mask patterns either intentionally or from linear approximations to smooth curves (figure 4.5(a)). Sharp corners introduce localized vector effects, rather than vector effects distributed along an entire 1D edge. Additionally, sharp features can introduce cross-polarization (change in polarization) not present in locally straight edges. TEMPEST can model scattering from 2D corners and quantify the cross-polarization.
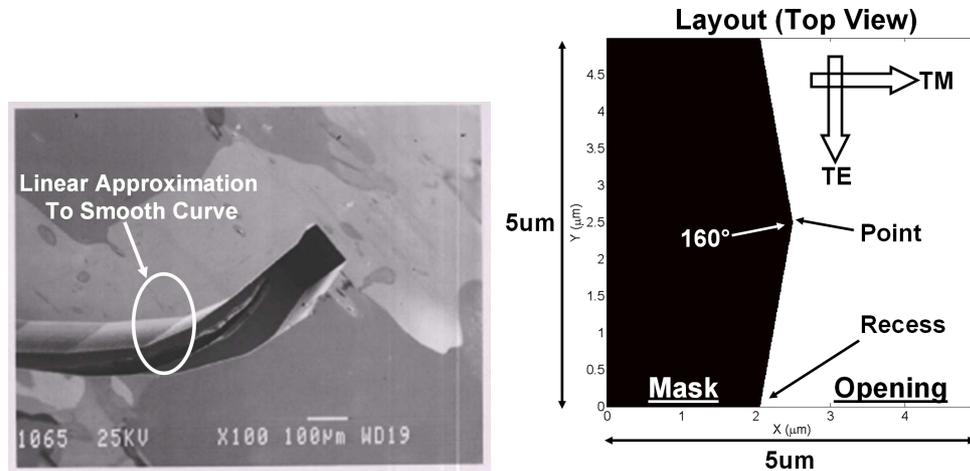
(a) SEM of linear approximation to smooth curves (courtesy of Jim Beall, NIST).

(b) Top view of sawtooth simulation layout.

Figure 4.5: Sharp corners in 2D mask patterns.

Figure 4.5(b) shows a top view of the layout of a simulation modeling a sawtooth edge, which has two types of 2D corners: protrusions and indentations. By simply switching the post-processing applied to the fields, a single simulation of this layout can be used to assess vector effects on both intentional 2D corners (e.g. those designed into the mask pattern) and unintentional corners (e.g. those from linear approximations to smooth curves). To model the former, the thin-mask design was given the saw-tooth shape shown in figure 4.5(b). To model the later, the thin-mask design was given a 1D pattern, which was a straight sidewall located midway between the maximum intrusion and protrusion.

To correct the ITM for 2D localized vector effects a 2D box-car model was required.

		TE Output	TM Output
TE Input	Point	$2.18\lambda^2, -121.3°$	$0.0014\lambda^2, -131.1°$
	Recess	$1.65\lambda^2, -113.0°$	$0.0011\lambda^2, 66.1°$
TM Input	Point	$8.7 \times 10^{-4}\lambda^2, -134.5°$	$1.99\lambda^2, -140.1°$
	Recess	$1.4 \times 10^{-5}\lambda^2, 49.6°$	$1.18\lambda^2, -139.3°$

Table 4.7: 2D box-cars from protrusions and recesses on an unintended sawtooth edge.

This 2D box-car was a simple extension of the 1D case: the height was set equal to the amplitude of the incident light and the area and phase are scaled to preserve the on-axis transmission coefficient.

3D simulations require very large amounts of memory and runtime, even when run in parallel on a network of workstations. Consequently, the saw-tooth simulations were designed to be small, occupying a volume only $5\mu m \times 5\mu m$ across containing a $10\mu m$ thick Si mask with vertical sidewalls. The opening varied from $2.5\mu m$ at the narrowest to $3\mu m$ at the widest. The wavelength was $630nm$. These simulations took up to 60 hours to complete on 8 workstations in parallel (using the Millennium cluster at U. C. Berkeley[10]).

The vector effects are summarized in tables 4.7 and 4.8. The upper left and lower right quadrants of each table account for the direct vector effects and the other two quadrants account for cross-polarization. The box-cars show that these protrusions and recesses account for approximately $2\lambda^2$ boxcars. Unintentional protrusions generate more stray-light than unintentional recesses (table 4.7) whereas the opposite occurs for intentionally designed protrusions and recesses (table 4.8).

		TE Output	TM Output
TE Input	Point	$1.95\lambda^2, -106.8°$	$0.0014\lambda^2, -131.1°$
	Recess	$1.98\lambda^2, -129.8°$	$0.0011\lambda^2, 66.1°$
TM Input	Point	$8.7 \times 10^{-4}\lambda^2, -134.5°$	$1.59\lambda^2, -126.9°$
	Recess	$1.4 \times 10^{-5}\lambda^2, 49.6°$	$1.70\lambda^2, -153.0°$

Table 4.8: 2D box-cars from protrusions and recesses in intended sawtooth edge.

One explanation for this imbalance is that thin-mask models do not account for diffraction confinement, so a large portion of the box-car width on a 1D edge comes from the shadow created by the thick mask sidewalls. However, near a sharp protrusion light can more deeply penetrate into the mask material thus shrinking this shadow and reducing box-car size. Conversely, recesses reinforce the diffraction confinement and force light even farther away from the sidewall, increasing the width of the shadow and thus increasing box-car size. In the unintentional case, recesses are more effective at pushing light out to the intended location of the sidewall than are protrusions so the recesses generate smaller box-cars. In the intentional case, recesses reinforce diffraction confinement making the recess extra dark and increasing box-car size. For protrusions, penetration into the mask material probably offsets diffraction confinement reducing box-car size.

The cross-polarization was extremely small for these mask designs (upper right and lower left quadrants of the two tables). This can be explained by the shallow angle of the saw-tooth and by the fact that the grazing incidence reflection coefficients are independent of polarization.

4.6.2   2D RIE Corrugation

Reactive ion etching[9] during the manufacturing process introduced significant roughness on the bottom and sidewalls of the HCIT masks (Fig. 4.6). This roughness comes in two forms:

1. Horizontal corrugation, caused by alternating etching and passivation cycles. This was modeled with a modification to the 2D simulations discussed previously.

2. Vertical erosion, possibly due to ion bombardment. This required full 3D simulations and special Matlab scripts to grow 3D mask patterns.
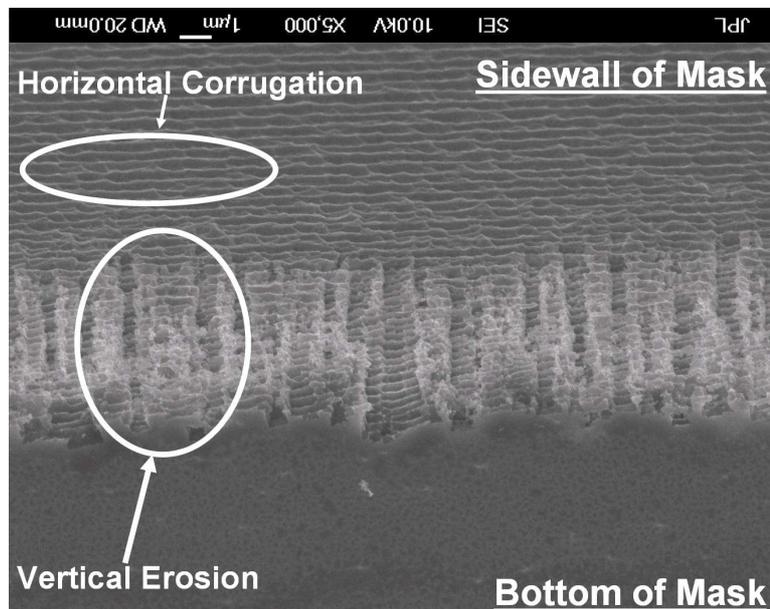
Figure 4.6: SEM of mask sidewall showing manufacturing roughness. Note: the image has been inverted to be consistent with the orientation used throughout this paper.

In a joint study with Marshal Miller, horizontal corrugation was added to 2D smooth-walled simulations by introducing a periodic array of cylinders along the sidewalls. Sharp roughness, like that seen in the SEM image (Fig. 4.7), was modeled with air-filled cylinders periodically cut into the sidewalls whereas thicker, rounded corrugation was modeled by alternating air-filled and Silicon-filled cylinders to remove and add mask material respectively. The dashed line indicates the intended location of the smooth sidewall.

(a) Sharp roughness.   (b) Rounded roughness.
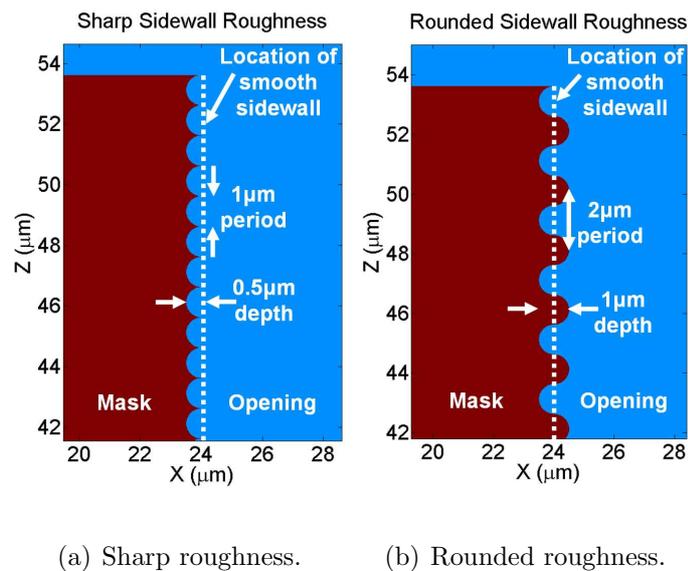
Figure 4.7: Layout of 2D roughness simulations (side-view).

Table 4.9 summarizes the vector effects from simulations of $48\mu m$ wide openings in $50\mu m$ thick Silicon masks. It compares a smooth sidewall to two rough sidewalls. The box-cars show that light penetrated through the thin, sharp corrugations and reduced the box-cars in comparison to the smooth sidewall case; however, the thick,

Roughness	TE	TM
None	$2.76\lambda, -135.4°$	$2.49\lambda, -146.2°$
Sharp	$2.38\lambda, -132.3°$	$2.39\lambda, -134.0°$
Rounded	$3.27\lambda, -145.9°$	$2.98\lambda, -150.3°$

Table 4.9: Box-cars from simulations of horizontal RIE corrugations. Sidewall geometry shown in figure 4.7.

rounded corrugations were thick enough to reduce this penetration and thus restrict the aperture by an additional $\lambda/2$ per edge beyond the smooth sidewall.

4.6.3   3D RIE Erosion

Masks modeling combined vertical erosion and horizontal corrugation were grown in Matlab and loaded into TEMPEST as a refractive index bitmap. The masks were grown through a combination of extrusion to define the vertical openings and convolution and thresholding to define the horizontal corrugations. This method can handle both vertical sidewalls with sloping erosion and bottom awnings with vertical erosion.

A simulation of a $2.5\mu m$ wide opening in a $10\mu m$ thick Si mask with a $200nm$ Aluminum top-coat ($n_{Al} = 1.3587 + j7.591$ at a wavelength of $630nm$) was conducted as a demonstration of TEMPEST's 3D capabilities. An "awning" structure, which resticts the aperture, was added at the bottom of the opening as a worst case scenario. This "awning" protruded $625nm$ from each sidewall and had a $10°$ slope. A series of vertical grooves and holes was etched through this awning and horizontal corrugation,

with a period of $1\mu m$, was added along each sidewall (Figures 4.8 and 4.9).
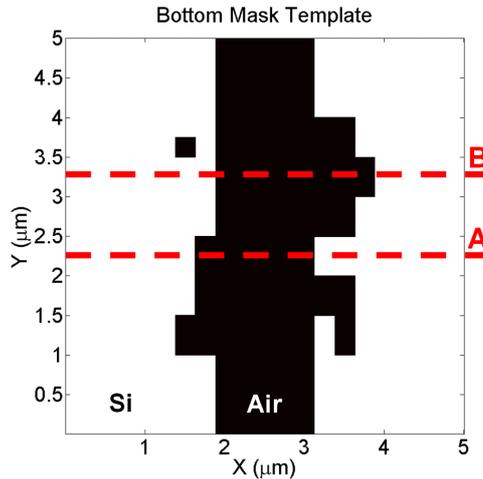
(a) Top-view.

Figure 4.8: Layout of the 3D roughness simulation modeling an opening with an intended width of $2.5\mu m$. This images shows manufacturing roughness in the form of vertical holes and grooves etched into the sidewalls.

The vector effects were very strong in this narrow opening. The opening was only four wavelengths wide and acted as a waveguide. The waveguiding and roughness caused a 70° phase difference in transmission between the scalar model's predictions and TEMPEST's thick-mask results across the entire opening. Waveguiding effects also caused a strong polarization imbalance with stronger TM transmission than TE transmission. Additionally, the near-fields show a wider opening under the TM polarization than the TE polarization (figure 4.10). The awning structure further restricted the bottom aperture, narrowing the opening to a physical width of $1.25\mu m$. Cross-polarization was also evident and was strongest at the corners in the vertical grooves and "awning" protrusion (Fig. 4.11).
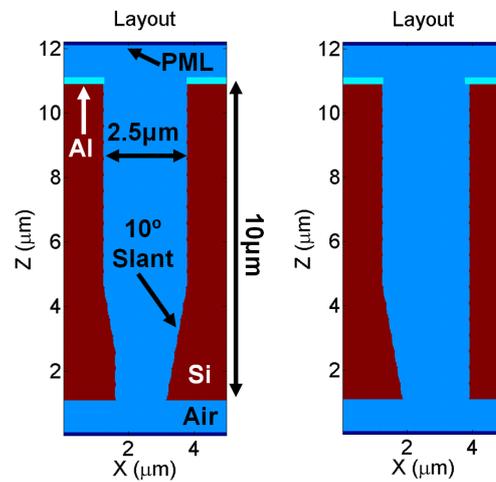
(a) Side-view (cut-line A).

(b) Side-view (cut-line B).

Figure 4.9: Layout of the 3D roughness simulation modeling an opening with an intended width of $2.5\mu m$. These images show manufacturing roughness in the form of an "awning" structure that narrowed the bottom of the opening and horizontal corrugation. These simulations were performed at a wavelength of $630nm$.

Due to the strong waveguiding effects, the reference plane was moved to the bottom of the mask. Different vector effects were isolated by comparing the rough 3D mask's near-fields to three different thin mask models:

1. a thin-mask model with a width equal to the opening's designed width $(2.5\mu m)$,

2. a thin-mask model with a width equal to the aperture set by the "awning" structure $(1.28\mu m)$, and

3. a smooth walled thick-mask model (2D TEMPEST simulation) with a width equal to the designed width $(2.5\mu m)$. Comparison with this model enabled measurement of contributions of the 3D roughness effects above and beyond
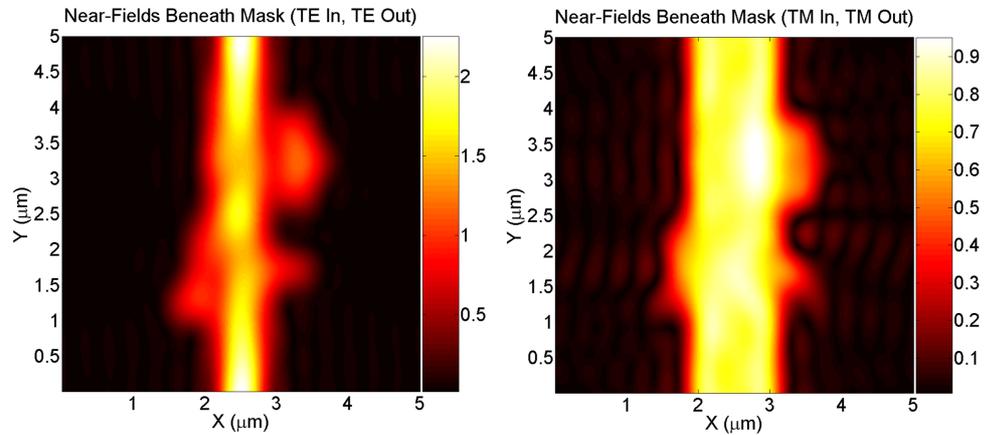
(a) TE near-fields.   (b) TM near-fields.

Figure 4.10: Direct near-fields at the bottom of the 3D manufacturing roughness mask. Non-propagating fields have been removed.

the 2D diffraction confinement.

The opening was so narrow that the edges were not independent. However, the box-car analysis was carried out for comparison with previous results. The box-cars are shown in table 4.10. The numbers reported are the average box-car along one sidewall. Surprisingly, comparison against the smooth walled thick-mask reference produced the smallest difference fields of the three models even though the thick-mask model's opening was much wider than the narrow opening defined by the "awning." This indicates that waveguiding effects were stronger than aperture restriction affects caused by the "awning." The smooth thick mask model had the following box-cars when compared to a thin mask model: TE $\rightarrow$ $(1.10\lambda, -121.2°)$, TM $\rightarrow$ $(0.79\lambda, -167.9°)$. Of the two thin-mask models, the $1.28\mu m$ wide opening best matched the 3D simulation results indicating that diffraction confinement by the "awning" was the second
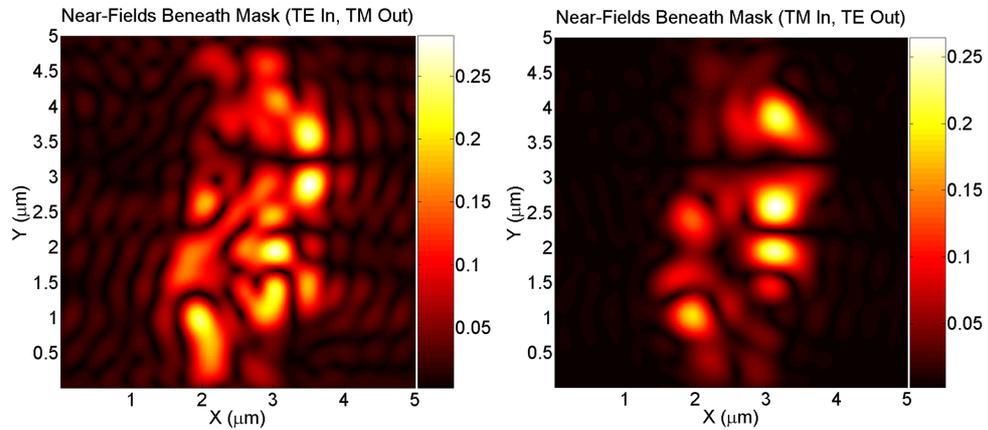
(a) TE input, TM output.

(b) TM input, TE output.

Figure 4.11: Cross-polarized near-fields at the bottom of the mask. Non-propagating fields have been removed.

Reference (Opening Width)	TE	TM
Scalar $(2.5\mu m)$	$1.49\lambda, -154.7°$	$1.16\lambda, -167.3°$
Scalar $(1.28\mu m)$	$1.34\lambda, -107.2°$	$0.78\lambda, -127.0°$
Vector $(2.5\mu m)$	$0.84\lambda, -162.4°$	$0.37\lambda, -151.4°$

Table 4.10: Average boxcars from simulations of a narrow opening with 3D roughness.

strongest vector effect.

4.7 Impact of Vector Effects on Planet Finding

The assessment of the vector effects' impact on the final image was carried out by Michael Lieber of Ball Aerospace [39] and this section contains a discussion of his methodology and conclusions.

Vector effects were included in the ITM by adding complex box-cars to the thin-

mask model along every mask edge (figure 4.12). Extra wide box-cars used to obtain
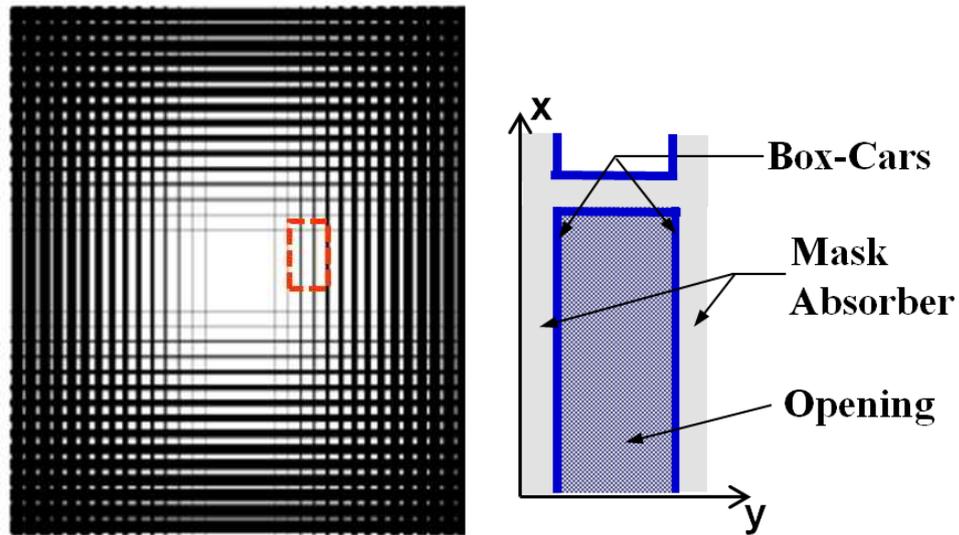
(a) Checkboard mask pattern (top-view).   (b) Box-car stitching.

Figure 4.12: To put vector effects into the ITM, the original thin-mask pattern was altered by adding complex box-cars at every edge in the mask. The image on the right shows a magnified piece of the pattern on the left to show the location of the box-cars at the edges. Image reproduced from [39].

a conservative upper bound on the vector effects. Specificall, a $3\mu m$ wide box-car with $-123°$ phase was used for the TE polarization and a $2\mu m$ wide box-car with $-157°$ phase was used for the TM polarization at a wavelength of $500nm$. A complication arose from a mismatch in the definition of polarization. TEMPEST defines polarization relative to each edge, with $\vec{E}$ oriented parallel to the edge for TE and $\vec{E}$ oriented perpendicular to the edge for TM. The ITM uses a global definition of polarization, i.e. $\vec{E}$ pointed along a coordinate axis. Consequently, a super-position

of TE and TM box-cars were used along curved edges to force the vector effects into compliance with the ITM's definition of polarization (figure 4.13). 3D vector edge
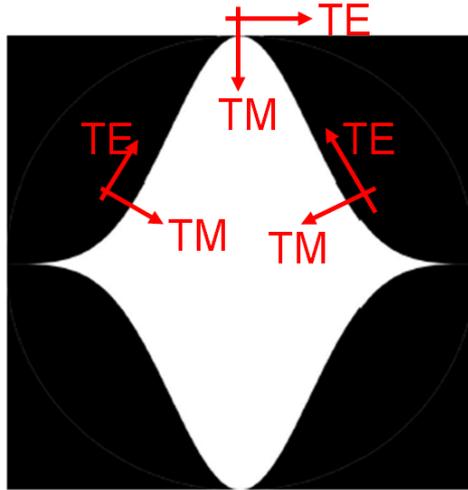
Figure 4.13: The polarization convention for the vector edge effects (defined in TEM-PEST) had to be rotated to conform to the local edge curvature in the ITM. This image shows examples of the local polarization definition at three points on an edge in the SKV mask pattern. Image courtesy of Michael Lieber.

effects from sharp corners were omitted in these investigations.

The ITM analysis showed that the impact of vector edge effects on the telescope's planet finding capabilities depended on mask design but did not depend on edge perimeter. A cutline through the planet finding regions of the PSF from two mask patterns is shown in figure 4.14. Stray-light from vector effects on the Spergel-Kasdin-Vanderbei (SKV) mask[30] (figure 4.13) exceeded the 10^{-10} intensity threshold in the planet search region, whereas stray-light from the checkerboard mask[66] (figure 4.12(a)) did not. However, the impacts may be larger because these numbers represent the intensity of only the vector effects and not the coherent interaction of

the vector effects with the thin-mask PSF. Additionally, these results used the TE polarization only.

It was surprising that the vector effects were stronger on the SKV mask than the checkerboard mask. The original hypothesis was that vector effects come from edges, so the masks with longer edge perimeters would have worse vector effects. However, the intuition from the ITM analysis is that vector effects on the bar-code mask produce near-fields that look very similar to the original mask design and thus coherently reinforce the desired point spread function (PSF) in the far-field. Conversely, the vector effects on the SKV mask do not resemble the original mask design and therefore do not reinforce the desired PSF leading to worse stray light.

A final study explored the effects of mask size. The ITM showed that as mask size decreased the impact of vector effects worsens. For example, when the mask diameter was decreased from $10cm$ to $0.5cm$ the intensity of the vector effects in the image increased two to three orders of magnitude, exceeding the planet finding threshold for both mask patterns. Performance degrades as mask size shrinks because the vector effects account for a larger percentage of the light moving through each opening when the opening size is decreased.

The understanding that narrower openings degrade mask performance led to a rethinking of the design and a proposal to separate the TPF-C into two space-craft: a telescope and a mask [34]. These two space-craft would require very careful orbital alignment, but the mask could be made so large that vector effects could be neglected.

4.8   Conclusion

A combined TEMPEST/ITM assessment of vector effects from thick mask structures showed that edge scattering may keep the TPF-C from finding terrestrial planets. However, vector edge effects can be mitigated through careful selection of mask pattern and sidewall profile.

In these two chapters, an improved back-propagation method based on precise measurements of numerical propagation parameters was introduced that allowed the retrieval of phase as well as amplitude information about stray-light sources. A thin-mask equivalent model (box-cars) was developed to quantify edge effect strength in the near-field and stitch the edge effects into the integrated telescope model for full telescope analysis. Using these methods a number of physical phenomena were investigated and characterized through box-car widths and phases.

On smooth-walled masks, undercut angle was found to be the dominant source of stray-light, accounting for over 2λ of electromagnetic bias per edge with vertical sidewalls on $50\mu m$ thick masks. Undercutting the sidewalls by $20°$ reduced this effect by an order of magnitude. Narrow openings were also problematic and vector effects were found to change transmission by 28% in $10\mu m$ wide openings in a $50\mu m$ thick silicon mask.

3D simulations of sharp 2D corners and 3D manufacturing effects were within TEMPEST's capabilties when run on a network of workstations, but took up to 60 hours each. These simulations demonstrated that vector effects accounted for $2\lambda^2$

boxcars on shallow 2D sawtooth edges.

Manufacturing roughness observed on HCIT masks was very complicated and was broken down into two categories for modeling purposes: horizontal corrugations from the Bosch process and vertical erosion through "awning" structures from ion bombardment. The severity of vector effects from the horizontal corrugation was found to depend on the thickness of the corrugation, with $1\mu m$ rounded corrugation producing half a wavelength more box-car than a smooth sidewall and thin, sharp corrugations producing up to $\lambda/3$ less box-car than a smooth sidewall. 3D studies of "awnings" and vertical erosion in narrow openings demonstrated strong waveguiding, cross-polarization at sharp corners, and aperture restriction.
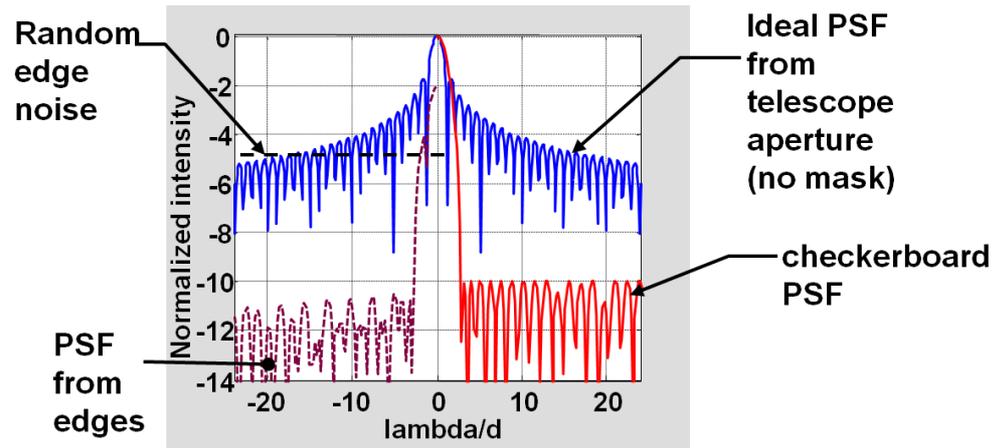
(a) Checkerboard PSF.
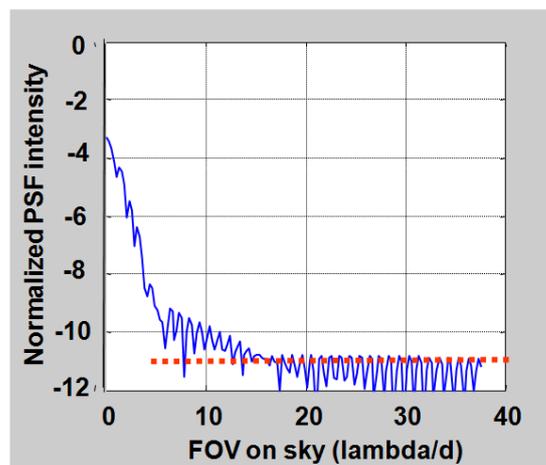
(b) SKV PSF.

Figure 4.14: Cutlines through the dark planet search regions showing vector effects in the final PSF for two different mask patterns. Reproduced from [39].

Chapter 5

Rapid Thermal Annealing

TEMPEST simulation of metal gate line/space patterns show that pattern orientation, topography, and density dependencies strongly affect optical heating power delivered to the substrate during the Laser Spike Annealing process. Advanced rapid thermal annealing (RTA) processes enable improved device performance; however, they require precise control of heat delivery and distribution. Metal gate structures placed on the wafer before the RTA step act like wire grid polarizers with strong reflection coefficients that can vary depending on local pattern geometry. TEMPEST enables investigations into geometry dependencies in gate reflectivity and near-field intensity distributions but requires extremely fine meshing due to a large length scale mismatch between the $10\mu m$ wavelength and the $50nm$ gates.

TEMPEST simulations of $50nm$ Tungsten gates show that optical power delivered to the substrate can vary between 4% and 100% depending on gate orientation.

Thick silicon nitride caps catch additional light, increasing the minimum power delivered to 53%. Near-field intensity distributions show hotspots at metal corners and waveguiding in high index poly-silicon pedestals.

5.1 Introduction

Careful control over the RTA process can improve device performance by increasing conductivity by a factor of two. This increase comes from boosting doping concentrations above the solid solubility limit but requires very rapid and precise heating [57]. Researchers at IBM demonstrated that these faster heating processes can cause large circuit performance variations. They tied these variations to small differences in pattern density that change the local reflectivity and thus optical heating[2].

The studies presented here used TEMPEST simulations to quantify reflectivity dependencies on the geometry of tungsten gate structures. They focused on one particular RTA system, Laser Spike Annealing from Ultratech, which uses a $10\mu m$ polarized laser source optimized to deliver light near silicon's Brewster angle to maximize power coupling into the substrate. Ultratech optimized their Laser Spike Annealing system by polarizing the laser illumination in the TM mode and changing the angle of illumination to Brewster's angle in order to maximize optical power delivered into the substrate.

5.2 Description of Physical Characterization Studies

TEMPEST was used to investigate transmission and reflection from the three gate stacks shown in figure 5.1 which commonly appear in memory and logic layouts. These are simple line/space patterns that were modeled with fast 2D simulations. The
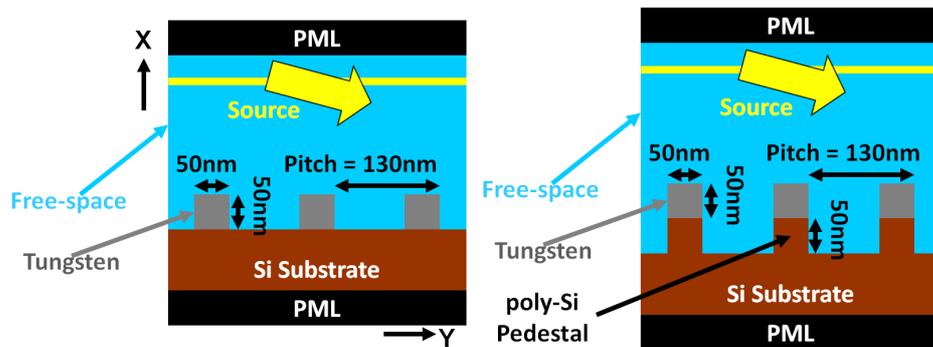
(a) Gate directly on wafer.

(b) Gate on pedestal.
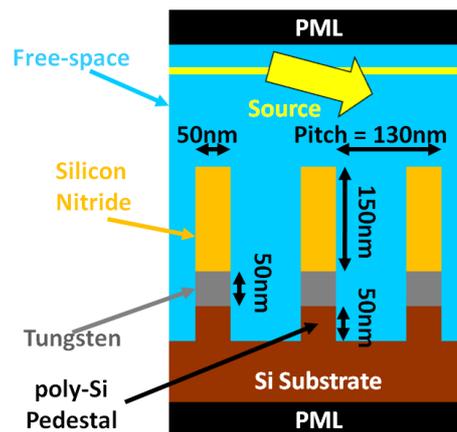
(c) Gate with capping layer.

Figure 5.1: A cross-sectional view of the layouts of the geometries explored in this study.

silicon substrate ($n = 3.5$), poly-silicon pedestals ($n = 3.5$), and free-space cladding ($n = 1$) were simulated as loss-less materials with the standard Yee model[76]. The tungsten ($n = 10.1 + j46.4$) and silicon nitride ($n = 1.184 + j1.604$) were simulated with the frequency-dependent FDTD method[42]. Cell sizes of $10nm$ were used to resolve the small $50nm$ gate structures. Illumination was provided by plane-wave sources placed above the gate stack. Floquet boundary conditions were used to enable off-axis illumination at arbitrary angles.

Transmission and reflection were measured as a percentage of the incident power. For example, to compute the transmission coefficient, first the power flow into the wafer was computed with the Poynting vector (5.1):

$$S_{trans}(y) = \hat{x} \cdot \vec{S}_{trans}(y) = 0.5 \cdot Real\left\{\vec{E}^{trans}(y) \times \left(\vec{H}^{trans}(y)\right)^*\right\}, \qquad (5.1)$$

where $\hat{x}$ is the unit vector normal to the surface, $\vec{S}_{trans}$ is the Poynting vector from the transmitted wave, and the asterix denotes complex conjugation. Bright spots in the near-field caused small fluctuations in S_{trans}, which were removed by averaging in the y direction (parallel to the wafer surface). The four field components are located at slightly different locations due to the staggered grid and, in theory, must be interpolated onto a common grid. However, the cell size was so small that the interpolations produced negligle improvements and were omitted. Second, the transmission coefficient was computed as the ratio of the transmitted power flow to the

incident power flow (5.2).

$$T = \frac{S_{trans}}{S_{inc}}, \tag{5.2}$$

where T is the transmission coefficient, S_{trans} is the transmitted power flow, and S_{inc} is the incident power flow (measured from a separate 1D simulation). Reflection was computed in the same manner, but with data from a cutline located above the gate stack rather than in the substrate. Absorbtion due to resistive heating in the gate stack was computed through conservation of power.

The simulations were constructed to explore the effects of illumination conditions and gate topography. The following sections start with a discussion of the illumination studies followed by the topography studies and a discussion of near-field intensity patterns around the gate structures.

5.3   Impact of Illumination Direction

Reflectivity was expected to depend strongly on illumination direction. Changing the illumination direction changes the polarization of the light relative to the gates and polarization strongly affects the reflectivity of wire grids. Illumination direction is described by two angles: inclination angle θ (the angle of incidence relative to the surface normal) and the azimuthal angle ϕ (the angle in the plane of the wafer surface). Changes in the inclination angle were used to assess the impact of small misalignments in the LSA equipment. Changes in the azimuthal angle were used to

assess both variations in reflectivity from gate arrays aligned in different directions in the die and rotational misalignments of the wafer in the LSA system.

To study the impact of small variations in the inclination angle around Brewster's angle (75° for silicon) θ was varied between 70° and 80° and ϕ was set to zero (for reference $\theta = 0°$ corresponds to normal incidence). Illumination came from a TM plane-wave source for the p-polarization ($\vec{H}$ in the plane of the wafer) and a TE plane-wave source for the s-polarization ($\vec{E}$ in the plane of the wafer).

This study looked at the optical properties of simple tungsten gates placed directly on the wafer surface (figure 5.1(a)). Table 5.1 shows variations in the reflectivity for angles between 70° and 80° for the two polarizations. Reflectivity is a strong function of polarization, with nearly 100% reflection under s-polarized light and nearly zero reflection under p-polarized light. Reflectivity depends only weakly on angle of incidence, so small errors in θ in the LSA system will not greatly affect optical coupling.

Inclination Angle θ	P-Polarization	S-Polarization
70°	1.2%	98%
72.5°	0.21%	99%
75°	0.16%	99%
77.5°	1.6%	99%
80°	5.4%	100%

Table 5.1: Reflectivity of simple tungsten gates (layout shown in figure 5.1(a)) as a function of small variations in the inclination angle around Brewster's angle.

The second set of illumination direction simulations looked at the impact of az-

imuthal angle on reflection with p-polarized light. Variations in azimuthal angle are used to study dependencies on the orientation of gate arrays. In the LSA system the source direction is fixed and the gate orientation varies depending on local pattern orientation and rotational wafer misalignment. However, in TEMPEST, the modeling is greatly simplified by fixing pattern orientation and rotating the illumination.

Rotating the illumination out of plane complicates the plane-wave source and requires an extra post-processing before computing the Poynting vectors shown previously. Out of plane illumination requires placing sources in two different field components and ensuring that these sources are in phase and have the correct amplitude ratio. This is a difficult proposition because the staggered grid locates field components at different locations and the radiation impedance varies for each field component and direction. However, Maxwell's equations are linear in the source terms so a methodology was developed in which each source was run in a separate simulation and the results were combined through superposition during post-processing. A benefit of this method is that one pair of simulations can be used to compute transmission from all polarizations of the incident wave. The two sources cannot be run in the same simulation when out of plane illumination is used because the polarizations couple at material boundaries[19].

The complex weightings used in the superposition sum were computed from the incident waves. Two 1D empty simulations (free-space only) were run, one excited with an H_z source (termed TM here) and one with an E_z source (termed TE here).

A matrix equation of the form $Ax = b$ was formed from the electric fields parallel to the wafer surface, where A contained the complex electric field values from the two simulations (5.3):

$$A = \begin{bmatrix} E_y^{TE} & E_y^{TM} \\ E_z^{TE} & E_z^{TM} \end{bmatrix},$$

(5.3)

x contained the superposition weightings (5.4):

$$x = \begin{bmatrix} \alpha_{TE} \\ \alpha_{TM} \end{bmatrix},$$

(5.4)

and b contained the unit vector in the wafer plane pointing in the direction of $\vec{E}$ for the p-polarized incident wave (5.5):

$$b = \begin{bmatrix} cos(\phi) \\ sin(\phi) \end{bmatrix}.$$

(5.5)

The matrix equation was solved for x to obtain the weightings. Finally, the two fields were combined through superposition (5.6):

$$\vec{E}_{total} = \alpha_{TE}\vec{E}_{TE} + \alpha_{TM}\vec{E}_{TM}.$$

(5.6)

These weightings were used to computed $\vec{H}_{total}$ as well. Then $\vec{E}_{total}$ and $\vec{H}_{total}$ were used in (5.1) to find the transmission and reflection coefficients as before. There was one special case: when $\phi = 0$ the illumination was in plane and the A matrix was singular. However, there was a solution to the problem – the p-polarization simply corresponded to the TM simulation and the TE simulation was neglected.

In the azimuthal studies, the inclination angle was fixed at $\theta = 75°$ and the azimuthal angle was varied between $\phi = 0°$ (illumination perpendicular to the gates) and $90°$ illumination parallel to the gates (figure 5.2). Table 5.2 shows the reflection,
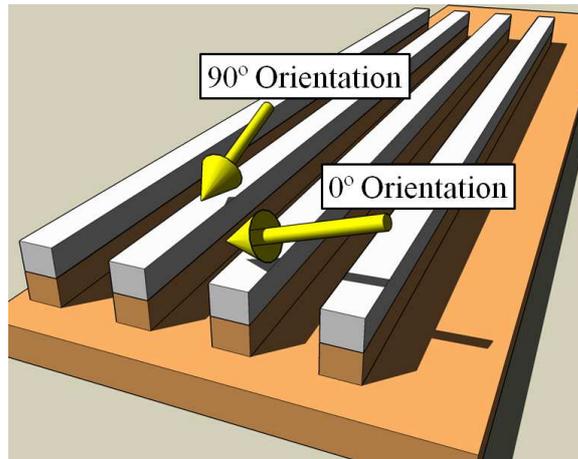
Figure 5.2: Layout of line/space pattern used in Laser Spike Annealing Studies. The arrows show the illumination directions for the two orthogonal orientations.

transmission, and absorbtion due to resistive heating in the tungsten gates for different azimuthal angles. The data shows that reflectivity is a strong function of orientation, with both a good and bad orientation. The good orientation, $\phi = 0°$, showed nearly 100% transmission so all of the optical power intended for substrate heating was actually delivered to the substrate. The bad orientation, $\phi = 90°$, showed 80% reflection, so only one fifth of the optical power went into heating the substrate. The bottom line is that a die with different grating orientations will show strong variations in heating under the LSA process. Reflection does not change rapidly with ϕ near $0°$ or $90°$ so LSA is insensitive to small orientation errors. Resistive heating in the

tungsten is significant in the bad orientation but is expected to contribute to substrate heating due to the large coefficients of thermal diffusion in tungsten and silicon.

Azimuthal Angle ϕ	Reflection	Transmission	Resistive Heating
0°	0.16%	99%	1.0%
2.5°	0.17%	99%	1.0%
5°	0.19%	99%	1.0%
15°	0.48%	98%	1.3%
30°	2.7%	95%	2.4%
45°	11%	84%	4.8%
60°	31%	60%	8.6%
75°	63%	24%	13%
85°	78%	6.4%	15%
87.5°	80%	4.5%	16%
90°	80%	3.8%	16%

Table 5.2: Variations in reflection, transmission, and resistive heating in the gate as a function of azimuthal angle ϕ for a simple tungsten gate (layout shown in figure 5.1(a)).

5.4 Impact of Gate Stack Topography

The effects of gate stack topography were investigated by repeating the azimuthal angle study with two more complicated topographies: pedestal-tungsten (figure 5.1(b)) and pedestal-tungsten-cap (figure 5.1(c)). The reflectivities are shown in table 5.3, with the reflectivities from the simple tungsten gate studies (table 5.2) repeated for convenience. The data shows that adding a $50nm$ poly-silicon pedestal beneath the gate did not greatly affect the reflectivity. However, adding a silicon nitride cap drastically reduced the reflectivity in the 90° (bad) orientation from 80% to 47% without

Azimuthal Angle ϕ	Gate on Wafer	Gate on Pedestal	Capped
0°	0.16%	0.71%	5.6%
2.5°	0.17%	0.72%	5.6%
5°	0.19%	0.74%	5.6%
15°	0.48%	1.0%	5.6%
30°	2.7%	3.2%	6.7%
45°	11%	11%	11%
60°	31%	31%	22%
75°	63%	63%	38%
85°	78%	78%	46%
87.5°	80%	80%	46%
90°	80%	80%	47%

Table 5.3: Reflectivity vs. azimuthal angle for the three gate topographies shown in figure 5.1.

greatly increasing reflectivity. Therefore the silicon nitride capping layer may be useful for improving heating in the bad orientation and reducing variability from pattern orientation.

Table 5.4 shows the reflection, transmission, and resistive heating in the capped topography. The last row shows that the decrease in reflectivity in the bad orientation did not come from direct transmission into the substrate but rather resistive heating in the silicon nitride cap. However, this heating is expected to rapidly diffuse into the substrate and contribute to substrate heating.

5.5 Fill Factor Dependence

The final geometry dependence investigated was the effects of pitch and fill factor. The azimuthal angle study was repeated with the tungsten on pedestal topography

Azimuthal Angle ϕ	Reflection	Transmission	Resistive Heating
0°	5.6%	80%	15%
2.5°	5.6%	80%	15%
5°	5.6%	80%	15%
15°	5.7%	79%	16%
30°	6.7%	74%	19%
45°	11%	63%	26%
60°	22%	43%	35%
75°	38%	16%	46%
85°	46%	4.2%	50%
87.5°	46%	2.9%	51%
90°	47%	2.4%	51%

Table 5.4: Reflection, transmission, and resistive heating in the silicon nitride capped gate topography (layout shown in figure 5.1(c)).

(figure 5.1(b)) with two pitches (130nm and 200nm) and two fill factors (42% and 28%). The results in table 5.5 show that fill factor is a stronger source of variation in transmission than pitch, with a 33% reduction in fill factor leading to a 160% increase in transmission in the 90° (bad) orientation. Fill factor and pitch did not noticeably affect transmission in the 0° orientation.

5.6   Near-Fields

TEMPEST enables visualization of the near-fields around the gate topography, which reveals interesting hotspots and waveguiding effects. Fig. 5.3(a) shows a cut-line of the electric field intensity at the top surface of the three different gate stack topographies illuminated in the 0° (good) azimuthal orientation with a 70° angle of inclination. In all three cases, hotspots were seen on the down-stream corners of

Pitch	130nm Pitch	200nm Pitch	200nm Pitch
Fill Factor	42% Filled	43% Filled	28% Filled
0°	98%	98%	99%
2.5°	98%	98%	99%
5°	98%	98%	98%
15°	98%	98%	98%
30°	94%	95%	94%
45°	84%	84%	84%
60°	60%	60%	61%
75°	24%	25%	28%
85°	6.4%	7.2%	12%
87.5°	4.4%	5.3%	11%
90°	3.8%	4.6%	10%

Table 5.5: Transmission vs. azimuthal angle ϕ for different pitches and fill factors. The data comes from simulations of the tungsten on pedestal gate topography (figure 5.1(b)).

the gate stack. Fig. 5.3(b) shows shows the same cutline but illuminated in the 90° (bad) azimuthal orientation. In this case hotspots appeared on both the upstream and downstream corners and were symmetric. The hotspots around the metal corners were significantly stronger than those around the dielectric, silicon nitride, corners. Additionally, the hotspots in the bad orientation were more than 2x brighter than the hotspots in the good orientation, except for the capped geometry where the difference was only 1.6x.

Looking at the near-field intensities immediately beneath the gate topography, the hotspot in the 0° orientation, shown in figure 5.4(a), moved to the upstream corners. Again, the hotspot at the metal corner (in the tungsten on wafer topography) is brighter than the hotspots on the dielectric corners (poly-silicon in this case). Also,

(a) 0° orientation. (b) 90° orientation.
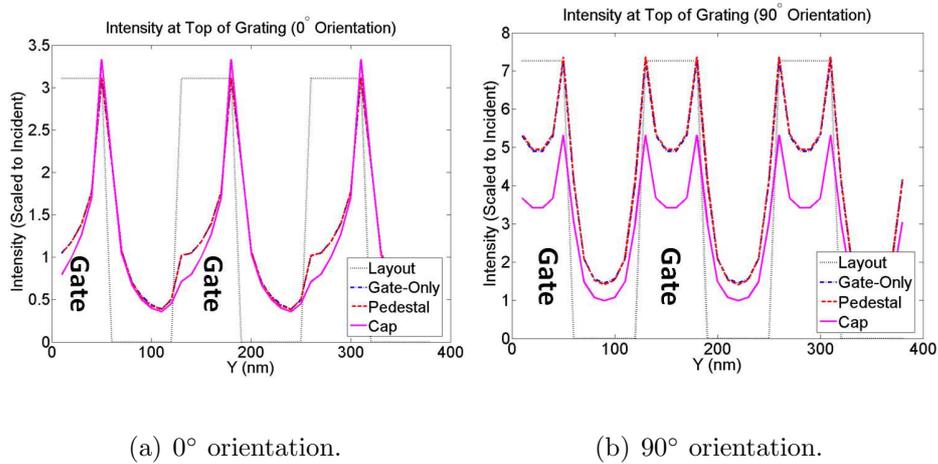
Figure 5.3: A cut-line showing the near-field electric field intensity at the top of the gate topographies shown in figure 5.1.

the tungsten gate created a shadow directly beneath it in the substrate.

The 90° orientation showed very interesting wave-guiding effects (figure 5.4(b)). The tungsten on wafer topography shows strong field localization between the gates, caused by confinement between the highly conducting tungsten sidewalls. Hotspots were also visible at the bottom corners. The two geometries with poly-silicon pedestals showed field localization in the opposite place – directly beneath the gates. This was due to wave-guiding by the high-index pedestals. Again, the capping layer geometry showed weaker near-fields than the other two geometries due to heating in the capping layer. Additionally, the brightness of the hotspots beneath the topography switched compared to the hotspots at the top. The bottom hotspots in the 0° (good) orientation appeared two orders of magnitude brighter than those in the 90° (bad) orientation.

The fields used to form the 90° orientation images shown above were interpolated

(a) 0° orientation. (b) 90° orientation.
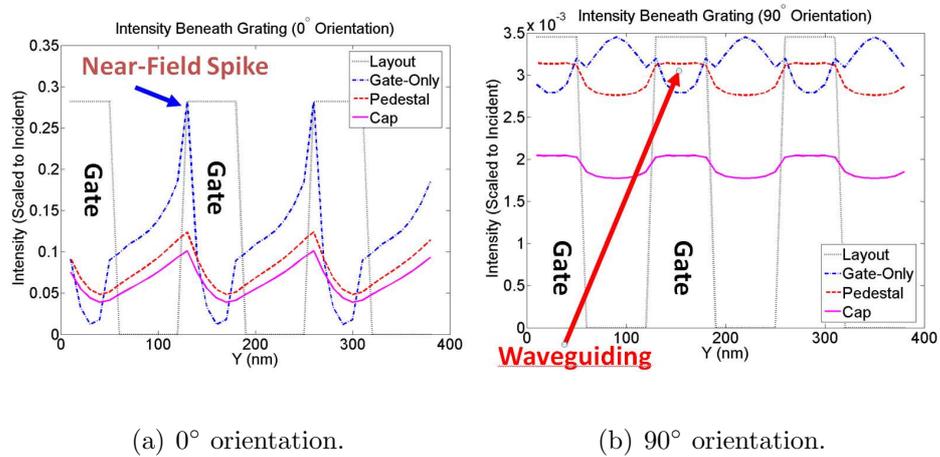
Figure 5.4: A cut-line showing the near-field electric field intensity just beneath the bottom of the gate patterns.

onto a common grid to avoid asymmetries. The asymmetries ere a numerical artifact of the staggered grid. This is in contrast to the reflection and transmission measurements described in section 5.2 which were performed without interpolation because the data was taken far away from hotspots.

5.7   Viability of 2D Pattern Assessment

TEMPEST simulations have proven to be useful in understanding optical coupling in 1D line/space gate arrays; however, realistic logic layouts contain 2D patterns. The 1D patterns assessed in this study require only 2D simulations, which are relatively quick, taking only a few minutes. Moving to 3D simulations is possible, but it is a time consuming proposition.

FDTD run-time depends on the number of cells and the number of time-steps. A

representative 2D simulation from these studies had $21,000$ cells and was run through $566,400$ time steps to achieve convergence. These simulations were usually run on a 3 GHz Intel Xeon processor with 16 GB of RAM and took between 19 minutes 39 seconds and 33 minutes and 60 seconds depending on the system usage level from other users. This works out to a time dependence of between $0.153\mu s$/cell/time step and $0.219\mu s$/cell/time step.

A 2D pattern representative of a small element of logic circuitry includes shapes that are roughly $500nm$ to $1\mu m$ on a side. For accurate measurements these require a $2\mu m$ thick simulation, leading to simulation volumes on the order of $0.5\mu m^3$ to $2\mu m^3$. With a typical cell size of $\Delta x = 10nm$ this leads to $500,000$ to $2,000,000$ cells. Running these simulations through the same $566,400$ time steps results in simulation times of 16 to 63 hours. Thus 2D patterns on wafers are possible but require hero runs.

5.8   Conclusion

TEMPEST is a useful tool for assessing optical coupling through metal gate topographies. It enables quantitative analysis of reflection, transmission, and resistive heating through line/space patterns and a qualitative assessment of the near-field intensity patterns around the gate structures.

TEMPEST simulations showed that optical coupling with p-polarized light depends strongly on grating orientation, with the good orientation (light delivered per-

pendicular to the gates) seeing nearly 100% transmission and the bad orientation (light delivered parallel to the gates) seeing 80% reflection. Thick silicon nitride capping layers may provide a pathway to reducing this variation because it reduced the reflection in the bad orientation to less than 50% with little change to transmission in the good orientation.

The near-field intensity patterns show strong localization at the metal gate corners and waveguiding effects in the poly-silicon pedestals.

Chapter 6

Optimizing Surface Plasmon Generation

Surface plasmons were first discovered over a century ago by R. Wood [74] and their use was explored for a number of applications [43] [22] [54]. Surface plasmon grating couplers are emerging as a key component in a number of devices, including near-field optical probes [5], surface enhanced Raman sensors [68], and solar cells [15]. These optical components are difficult to design because even a simple plane-wave is scattered into a multitude of plane-waves at a corner and each of these new plane-waves is scattered into an additional multitude of plane-waves at the next corner. Surface plasmons devices have the additional complexity of waves bound to the surface that can take multiple paths around the object and interfere. These complications make algebraic analysis intractable and most researchers rely on numerical simulation

[35][41]. These studies used a combination of FDTD and signal flow graph methods to investigate the contributions of individual elements and optimize their collective behavior to maximize grating output.

6.1   Surface Plasmons

A surface plasmon is the two wave solution of Maxwell's equations at a planar interface (similar to Brewster's angle). The dispersion relation is shown in (6.1)[54].

$$k_y = k_0\sqrt{\frac{\epsilon_m}{1+\epsilon_m}} \tag{6.1}$$

$$k_x^{metal} = \sqrt{k_0^2\epsilon_m - k_y^2} \tag{6.2}$$

$$k_x^{air} = \sqrt{k_0^2 - k_y^2} \tag{6.3}$$

where k_y is the wavenumber parallel to the interface, k_x is the wavenumber perpendicular to the interface, k_0 is the wavenumber of free-space, and ϵ_m is the complex dielectric function of the metal. (6.2) and (6.3) are simply the dispersion relations in the metal and air respectively and are used to compute the surface plasmon mode profile used by the surface plasmon source (see section 2.2.1).

In general, the surface plasmon wavenumbers are complex, representing both propagation and attenuation. Here the wavenumber is defined as $k = \beta + j\alpha$, where β is the propagation constant, and α is the attenuation constant. Surface plasmon propagation distance depends on material and wavelength. Figure 6.1 shows the travel distances for plasmons on silver, gold, and aluminum. The travel distance is largest

in the infrared. At 700 nm (used in these studies) the plasmon travel distance on silver is $68\mu m$.
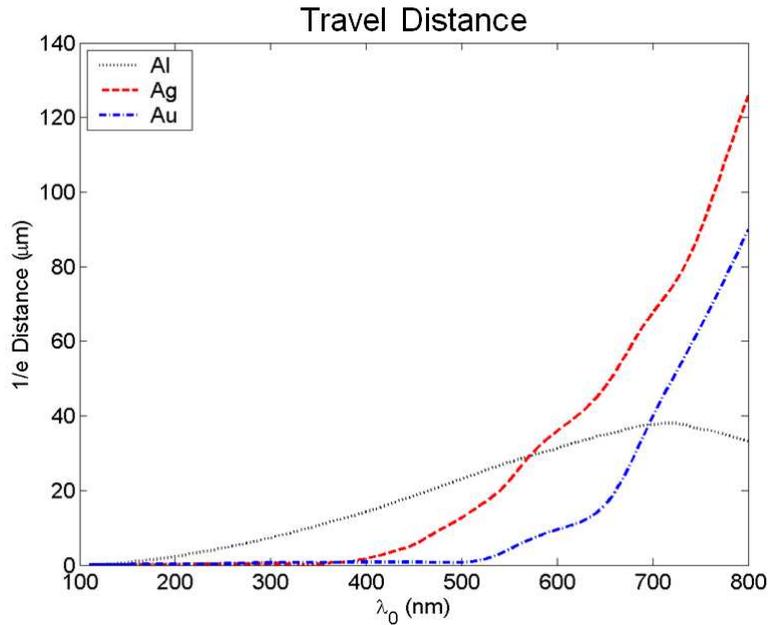
Figure 6.1: Surface plasmon travel distances as a function of wavelength at an planar interface between free-space and silver, gold, or aluminum.

The principle source of numerical error in surface plasmon work is numerical dispersion. The FDTD method is second order accurate in wavelength but only first order accurate in decay length, so increasing cell density only slowly improves errors in propagation distance at the cost of added memory and run-time. For example, table 6.1 shows the errors in the surface plasmon attenuation and propagation constants as a function of cell density. These values were measured in TEMPEST by first launching a surface plasmon across a smooth silver surface at $\lambda_0 = 400$ nm ($n = 0.173 + j1.95$) with the surface plasmon source. Then the propagation constant

was measured by performing a least-squares fit of the phase of H_z along the surface to a line and the attenuation constant was measured by performing a least-squares fit of the log of the amplitude of H_z along the surface to a line. Both of these operations were performed with Matlab's polyfit function.

Cells per λ_0	Relative Error	
	α	β
30	53%	10%
60	39%	5.9%
120	13%	2.1%
180	7.3%	1.4%

Table 6.1: The errors in surface plasmon wavenumbers caused by numerical dispersion at a wavelength of 400 nm. α is the attenuation constant and β is the propagation constant.

The data shows that over 100 cells per wavelength are required to reduce the errors in the attenuation constant to less than 10%. By comparison, photomask simulations typically use 30 cells per wavelength. The studies discussed here utilized two means of reducing numerical dispersion. First, a cell density of 175 cells per wavelength was used to reduce errors while keeping the run-time to only a few hours. Second, signal flow graphs used the surface plasmon propagation and attenuation constants measured from simulation rather than their theoretical values.

6.2 Calculating Surface Plasmon Scattering Parameters

Surface plasmon coupling through grating structures is a combination of the effects of individual elements and their coherent interactions. These studies broke these two processes apart and first quantified the scattering properties of individual grating elements. Second, these studies quantified the efficiency of finite length gratings and demonstrated the link between the elements scattering properties and the total grating efficiency.

Scattering by an individual grating element is a complex process; however, it can be characterized by three scattering parameters: the collection efficiency, the plasmon transmission, and the plasmon reflection. Separate simulations were required to assess collection efficiency and plasmon scattering. These simulations employed a cell size of 4 nm to resolve sharp corners and the skin depth. They contained a $4\mu m \times 2\mu m$ area and required 2 to 6 hours to complete on a 3 GHz Intel Xeon processor with 16 GB of RAM.

Surface plasmon transmission and reflection was assessed by illuminating the topography with a surface plasmon wave emanating from a surface plasmon source placed to one side. An example of this is shown in figure 6.2(a), which contains the steady-state H_z field scattered by a $200nm \times 200nm$ silver bar suspended 200 nm above a silver substrate ($n = 0.1419 + j4.5242$ [58]) at a wavelength of 700 nm. The

(a) Suspended bar.

(b) Triangle, SPP from left.

(c) Triangle, SPP from right.
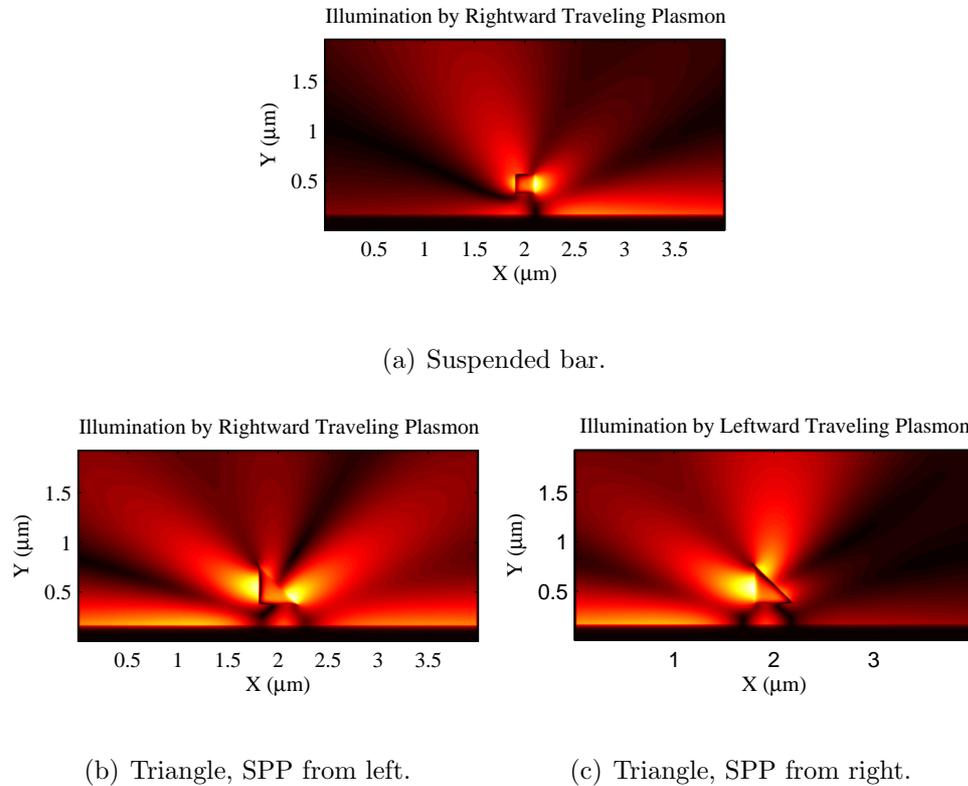
Figure 6.2: Amplitude of the H_z (out of the page) field component after a surface plasmon has scattered off square (a) and triangular (b),(c) bars.

detailed geometry is shown in the center of figure 6.3. This field was computed by placing a surface plasmon source to the left of the topography. The reflected and transmitted surface plasmon power flowing parallel to the interface was extracted during a post-processing step in Matlab. To accurately measure the surface plasmon power in the presence of low-angle radiation and reactive near-fields a mode overlap method was first applied to the transverse electric and magnetic fields:

$$H_z^{SPP} = \frac{\int H_z^{FDTD} \psi_{hz}^* dy}{\int \psi_{hz} \psi_{hz}^* dy},$$

(6.4)

where ψ_{hz} is the surface plasmon mode profile in the H_z field component. For greater
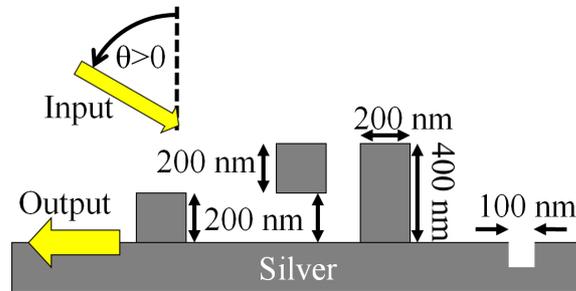
Figure 6.3: Four example element topographies. Positive angles correspond to plane-waves moving to the right.

accuracy the E_y field component was also treated in a similar manner. The mode overlap method was applied throughout an 800nm wide region $3.2\mu m$ from the center of the topography. The numbers shown have been corrected for a reference point at the center of the element. Additionally, re-radiated power was measured by integrating the power flow across a rectangular surface around the element (after removing the scattered plasmons). Absorption in the element was computed through conservation of power.

The mode overlap was validated by looking at the residual field at the silver surface after the detected surface plasmons were removed. The test geometry was a silver bar $100nm \times 100nm$ suspended 100 nm above a silver surface. It was illuminated with a plane-wave at normal incidence with a wavelength of 400 nm. The mode overlap method was used to detect the surface plasmons on each side of the bar in regions 400 nm to 900 nm away from the bar's center. The resulting complex amplitudes were used to reconstruct the surface plasmons in the detection region. These plasmons were subtracted from the original fields, resulting in the fields shown in figure 6.4.

After plasmon removal the amplitude of the H_z field component at the silver surface in the detection region was more than 7× lower demonstrating that the mode overlap method accurately measured surface plasmon amplitude and phase.

Scattered |H_z| with Surface Plasmons Removed

Surface plasmons removed
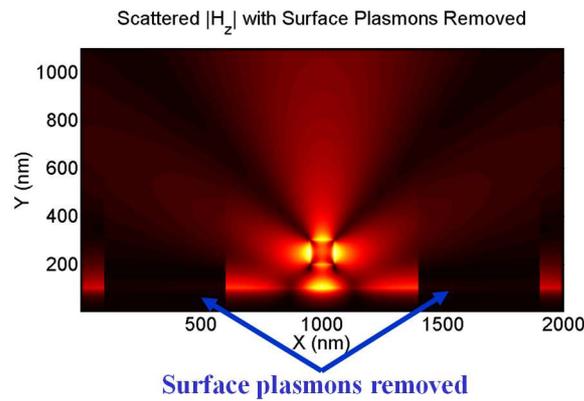
Figure 6.4: Magnitude of the scattered H_z field component after the detected surface plasmons were removed.

Table 6.2 shows the reflection and transmission coefficients from four example grating elements: the suspended bar discussed above, the same bar placed in direct contact with the surface (a ridge), a 100 nm square cross-section trench, and a 100 nm wide by 280 nm deep resonant trench. The second and third columns show the power in each scattering path. The signal flow graphs discussed later use complex scattering parameters to coherently compose waves from neighboring elements. These S parameters (columns 4 and 5) are the square-root of the power numbers, with a phase taken from the H_z field component at the metal/dielectric interface.

These coefficients show a broad range of values, with reflection ranging from moderate to near zero and transmission ranging from 12% to nearly 100%. Losses due to

Topography	R (power)	T (power)	R (S param.)	T (S param.)
200 nm square bar suspended 200 nm above surface	3.0%	67%	$0.17, -69°$	$0.82, 31°$
200 nm square ridge	34%	12%	$0.58, -52°$	$0.34, 25°$
100 nm wide x 100 nm deep trench	9.5%	53%	$0.31, -152°$	$0.73, -11°$
100 nm wide x 280 nm deep trench	0.039%	99%	$0.020, 104°$	$0.995, -3°$

Table 6.2: The complex surface plasmon reflection and transmission coefficients (columns 4 and 5) and the corresponding powers as a percentage of the incident wave (columns 2 and 3).

re-radiation and resistive heating are implicitly included in these coefficients. Conservation of power shows that these losses range from 50% down to near zero. The suspended bar is highly transmitting as the surface plasmon is able to pass through the gap. High transmission is important in a grating so that the plasmon launched by one element is able travel past neighboring elements. The ridge is a poor grating elements as it is highly reflecting and functions like a wall. Narrow trenches act like stub waveguides in series with the smooth surface and transmission is a smooth function of depth. Transmission is maximized when the trench is a half-wavelength deep, at its 560 nm propagation wavelength [12][13].

The phase of the complex S parameters indicate a slowing down or speeding up of the wave as it interacts with the topography (these studies used a sign convention of $e^{-j\omega t}$). The ridges transmission coefficient has a positive phase due to the additional path length incurred by going up and over the ridge. The positive phase of the

suspended bars transmission coefficient comes from the increase in wave-number as the plasmon moves beneath the bar. The negative phases of the reflection coefficients are due to the reflection occurring at the front face, 100 nm from the reference point at the center of the element.

The conversion efficiency between an incident plane-wave and a surface plasmon is characterized by a complex surface plasmon coupling coefficient, c, which relates the field amplitude of the plane-wave to the field amplitude of the plasmon. The plasmon is measured with the mode overlap method described above and is corrected for a reference point at the center of the element. This scattering coefficient is related to the coupling cross-section, σ, by:

$$c(\theta) = \sqrt{\sigma(\theta)}e^{j\phi(\theta)} = \sqrt{\frac{P_{SPP}(\theta)}{S_{inc}}}e^{j\phi(\theta)}, \tag{6.5}$$

where P_{SPP} is the surface plasmon power, S_{inc} is the power density in the incident wave (in the direction of propagation), θ is the direction of propagation, and ϕ is the surface plasmon phase. Coupling changes with angle and plasmon direction. Separate coupling coefficients were computed for leftward and rightward traveling surface plasmons.

Coupling cross-sections can range from nearly zero to slightly larger than the physical width of the topography. Figure 6.5 shows coupling cross-sections from four topographies, represented as antenna reception patterns. The suspended bar has coupling cross-section between 1% and 108% of the bar's physical width. This is not a theoretical limit but was roughly the maximum coupling we encountered for
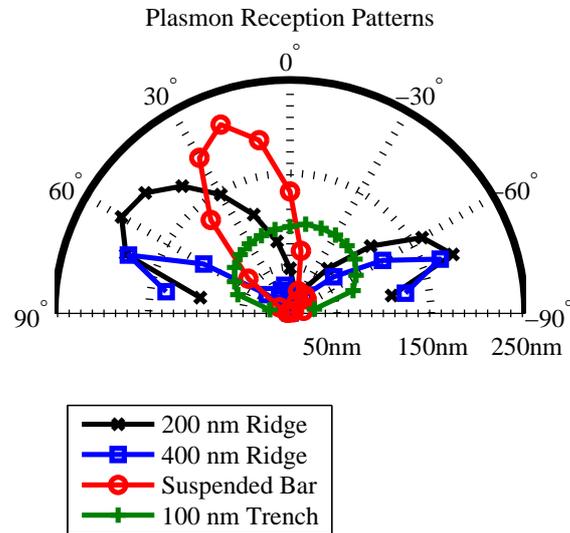
Figure 6.5: The plasmon coupling cross-sections for the leftward traveling plasmon represented as antenna patterns.

small perturbations of the topographies shown in figure 6.3. Ridges exhibited strong collection near grazing incidence and suspended bars showed strong coupling closer to normal incidence. Additionally, the suspended bar and ridge showed maximal coupling in the back-fire direction.

Due to reciprocity, strong plasmon coupling at one angle is equivalent to strong plasmon reradiation in the opposite direction. For example, the two re-radiation lobes seen in figure 6.2(a) correspond to the antenna lobes in figure 6.5 (the translation from amplitude to power greatly reduces one lobe in the antenna pattern). In these studies, half-wavelength and smaller topographies were found to have antenna patterns dominated by one or two lobes. Larger elements are required for three or more lobes.

Trenches exhibit a wide range of plasmon reception patterns. A 100nm deep $\times$ 100nm wide trench is an example of a strong, isotropic coupler (figure 6.5). A wide, shallow trench 200 nm deep and 900 nm wide is an example of a strong, directive coupler, with maximal coupling in the forward-fire direction. A half-wavelength deep trench, 100 nm wide and 280 nm deep, couples in the forward fire direction, but has a maximum coupling cross-section of only 1% of its physical width.

Due to reciprocity a symmetric element that is a strong receiver has poor transmission, an example of which is the half wavelength deep trench. An asymmetrical element topography may be able to break the direct relationship between reception and re-radiation by overlapping a null in the elements antenna pattern with a grating order. Asymmetric elements also enable a tradeoff between re-radiation, reflection, and resistive heating, as shown in table 6.3. However, surface plasmon transmission through asymmetric elements remains left/right symmetric due to reciprocity.

Plasmon Source Location	Transmission	Reflection	Re-Radiation	Resistive Heating
Left	35%	21%	43%	1%
Right	34%	3%	11%	52%

Table 6.3: Transmission, reflection, re-radiation, and resistive heating from the triangular bar when illuminated by a plasmon from the left and right.

6.3 Modeling Finite Length Gratings with Signal Flow Graphs

Finite length arrays can be modeled with signal flow graphs built from two port network models of each element (figure 6.6). These signal flow graphs track the complex surface plasmon amplitudes entering and leaving each side of every grating element. The smooth surfaces between elements are also modeled as simple two port networks. This leads to a set of linear coupling equations (6.6). In the notation used here, the n variable denotes the element number, b is the complex plasmon amplitude leaving an element, a is the complex plasmon amplitude entering an element, and L and R denote the left and right side of the element respectively.

$$
\begin{aligned}
b_{nL} &= r_L a_{nL} + t a_{nR} + c_L(\theta)\sqrt{S_{inc}}e^{jk_0 sin(\theta)Pn} \\
b_{nR} &= r_R a_{nR} + t a_{nL} + c_R(\theta)\sqrt{S_{inc}}e^{jk_0 sin(\theta)Pn},
\end{aligned}
\tag{6.6}
$$

where $c_R(\theta)$ and $c_L(\theta)$ are the leftward and rightward coupling coefficients, r_L and r_R are the leftward and rightward reflection coefficients, t is the transmission coefficient, $k_0 sin(\theta)$ is the wave-number of the incident wave parallel to the interface, and P is the pitch. The equations for plasmon propagation along the smooth surface between elements model a simple transmission process (6.7) that effectively couples the ports of neighboring devices, e.g. $b_{(n+1)L} \rightarrow a_{nR}$.

$$
a_{nR} = b_{(n+1)L}e^{jk_{spp}P}
$$

$$a_{(n+1)L} \;\; = \;\; b_{nR}e^{jk_{spp}P}, \tag{6.7}$$

where k_{spp} is the complex surface plasmon wavenumber parallel to the surface. These equations are implemented as an $Ax = b$ matrix equation in Matlab, which allows for rapid investigation of dependencies on pitch, angle of incidence, and number of grating elements. The b matrix contains the plane-wave source terms (e.g. $c_L(\theta)\sqrt{S_{inc}}e^{jk_0 sin(\theta)Pn}$), the A matrix contains the remaining terms, and the x matrix contains the unknown complex wave amplitudes (e.g. b_{nL} and $a_{(n+1)R}$). The ports at the ends of the grating are assumed to be matched, resulting in no reflections or inbound surface plasmons ($a_{1L} = 0$ and $a_{NR} = 0$).

The main assumption in this model is that the only interactions between neighboring elements occur through surface plasmons. This assumption was validated through FDTD simulation of full grating structures. Figure 6.7 shows the surface plasmon power output through pitch from a six element grating composed of suspended bars. The output was computed with signal flow graphs and spot checked with full TEMPEST simulations. The agreement between the models is quite good, with only an 11% over prediction of the output by the signal flow graph and a 9 nm shift in optimal pitch. The agreement between the models is quite good, with only an 11% over prediction of the output by the signal flow graph and a 9 nm shift in optimal pitch.

Typically, the optimal pitch is chosen with the Bragg condition, which uses the grating's periodicity to make up the differences in momentum between the incident
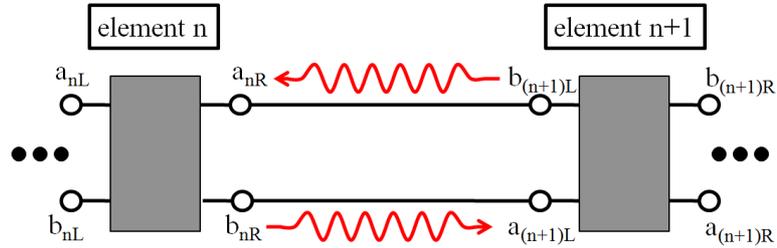
(a) Two-port model.
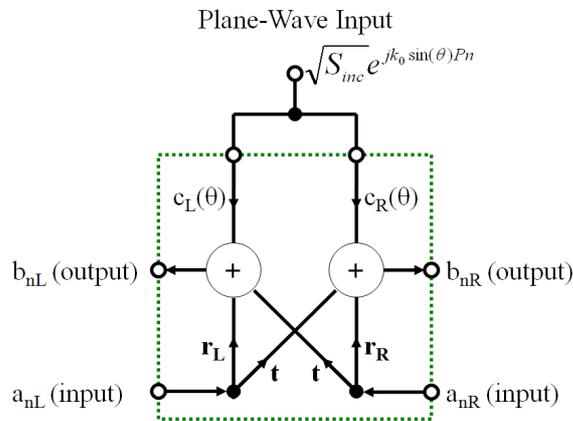
(b) Internals of the nth element.

Figure 6.6: Schematic of the signal flow model of a surface plasmon grating coupler.

plane-wave and the surface plasmon:

$$\frac{m}{\lambda_{SPP}} = \frac{sin\left(\theta\right)}{\lambda_0} + \frac{1}{P}, \tag{6.8}$$

where λ_{SPP} is the surface plasmon wavelength, λ_0 is the free-space wavelength, θ is the angle of incidence, m is the band number, and P is the pitch. However, the Bragg condition does not accurately predict the optimal pitch due to additional phase picked up as the surface plasmon interacts with each grating element (represented as
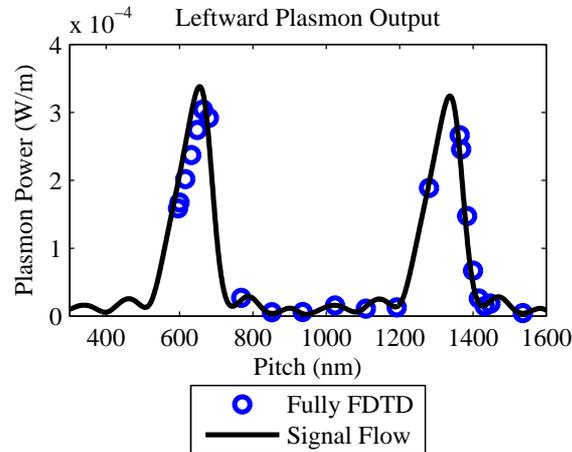
Figure 6.7: The plasmon power output exiting to the left from a six element grating composed of suspended bars illuminated at normal incidence. The smooth curve shows the predictions of signal flow graphs and the circles show the results of FDTD simulation of the full grating.

phases in the reflection and transmission coefficients). For example, under normally incident illumination the Bragg condition predicts an optimal pitch of 683 nm (the surface plasmon wavelength). The actual optimal pitches for six element gratings composed of suspended bars and 100nm wide by 200nm deep trenches are 655 nm and 734 nm respectively. The suspended bar has a shorter pitch than expect probably due to the increase in wave-vector as the plasmon sneaks beneath the bar. The trench has a longer pitch than expected probably due to the decrease in wave-vector as the plasmon jumps over the top of the trench.

The total grating plasmon output is determined by two factors: the collection efficiency of each element and the collective interactions of all the elements. These two factors are represented by the surface plasmon cross-section, σ, and the array

factor. If these gratings were classical antennas which collected the incident light and fed it down a waveguide, the total field collected would simply be a direction multiplication of the element factor and the array factor. However, the plasmon case has the additional complication that surface plasmons generated by one element must pass beneath the neighboring elements, and in so doing their transmission is reduced below unity. As was shown a large reception cross-section results in a small transmission coefficient. Therefore, we must examine the balance between reception and transmission.

Figure 6.8 illustrates how reception and transmission interplay in producing the grating output normalized to the incident power density versus the number of elements. The half wavelength trench has near 100% transmission (Table 6.2). Con-
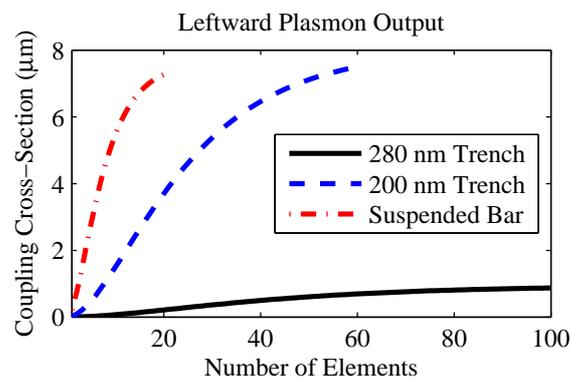
Figure 6.8: Normalized plasmon coupling cross-sections of finite length gratings composed of different numbers of elements.

sequently, many elements can make a contribution and the grating output does not saturate until the grating contains over 100 elements, where it becomes limited by

plasmon propagation loss. This is simply the absorption limit and is a strong function of material properties. However, such a high transmission leads to a very weak element coupling cross-section and the total grating output is weak, saturating at a cross-section of 750 nm. Switching to elements with larger cross-sections and lower transmissions, we find that the total grating output is an order of magnitude better, with a cross-section of $7.5\mu m$ (over ten wavelengths). Additionally, the output saturates at 60 elements for the non-resonant trench and 20 elements for the suspended bar. For many applications a smaller number of elements are preferred to reduce sensitivities to pitch, wavelength, and manufacturing uniformity.

6.4   Conclusion

Plasmon generation and scattering can be treated with finite difference time domain with 4 nm discretization of $20\lambda^2$ domains and desktop run-times of a few hours. A two-step methodology using a combination of FDTD and signal flow graphs was found to accurately predict finite length grating performance. Small surface features were found to show a wide range of characteristics, from near 100% transmission, to cross-sections exceeding their physical widths, to directive coupling patterns. Complex transmission coefficients were found to change the optimal pitch from the Bragg condition, and maximal grating output was found to come from a balance of element collection efficiency and transmission.

Chapter 7

Octahedral Nano-Crystal Arrays

TEMPEST simulations of arrays of silver nano-octahedra show that surface plasmon resonances are responsible for resonant absorbtion, transmission, and field enhancement. Colloidally assembled arrays of nano-crystals constructed through bottom-up synthesis methods form highly tunable optical materials in the visible spectrum[63]. Experimentally collected reflection spectra showed a broad, strong reflection peak and the working hypothesis was that surface plasmon resonances drove this peak. TEMPEST is capable of computing reflection spectra and near-field patterns throughout arrays of octahedra.

Examination of the time response of a reflected pulse, the near-field pattern, and the reflection spectrum with TEMPEST shows that slow reflection processes due to surface plasmon resonances give rise to a long wavelength dip in the reflection spectrum instead of the broad peak.

7.1 Introduction

There has been a lot of scientific research aimed at understanding optical processes in periodic arrays. The interesting questions are what range of phenomena are possible and how can they be controlled to engineer optical properties.

One exciting application for surface plasmonics is surface enhanced raman spectroscopy (SERS) for chemical detection. SERS requires very large field enhancements to be practical, with a goal of 1,000x. Many researchers have investigated surface plasmonic devices for SERS because surface plasmons may provide the necessary large field enhancements. For example, researchers at Stanford and the Molecular Foundry are looking into bow-tie antennas[59]. The Nordlander and Halas groups at Rice University work with artificial plasmonic molecules[69]. The Schatz group at Northwestern University has conducted experimental and FDTD investigations into oddly shaped plasmonic particles[56]. Researchers at U. C. Berkeley are investigating a plasmonic dimple lense[68].

Another application is materials with tunable reflection responses. A lot of research has gone into understanding dielectric photonic crystals. When used as a reflecting surface, photonic crystals have very sharp reflection bands generated by guided resonances[14]. In constrast, plasmonic crystals have broader reflection bands, as demonstrated with the particle arrays discussed in this chapter.

These studies considered the case of hexagonally packed nano-octahedral particles, $148nm$ on a side, suspended in water with variable interparticle distances and number

of layers. These particles were developed by the Yang group in the School of Chemistry at U. C. Berkeley[63]. Studies of mono-layer films in a Langmuir-Blogett setup showed that packing density controlled the dominant color of reflected light. Later studies showed that collidal solutions of these particles self-assemble into hexagonal arrays (shown in figure 7.1) and also demonstrate tunable optical responses. The synthesis method used to create octahedra can also be used to create cubes and cuboctahedra. The focus of the simulation studies discussed here was to understand the optical processes driving the reflection spectrum and uncover the role of surface plasmons in arrays of octahedra.
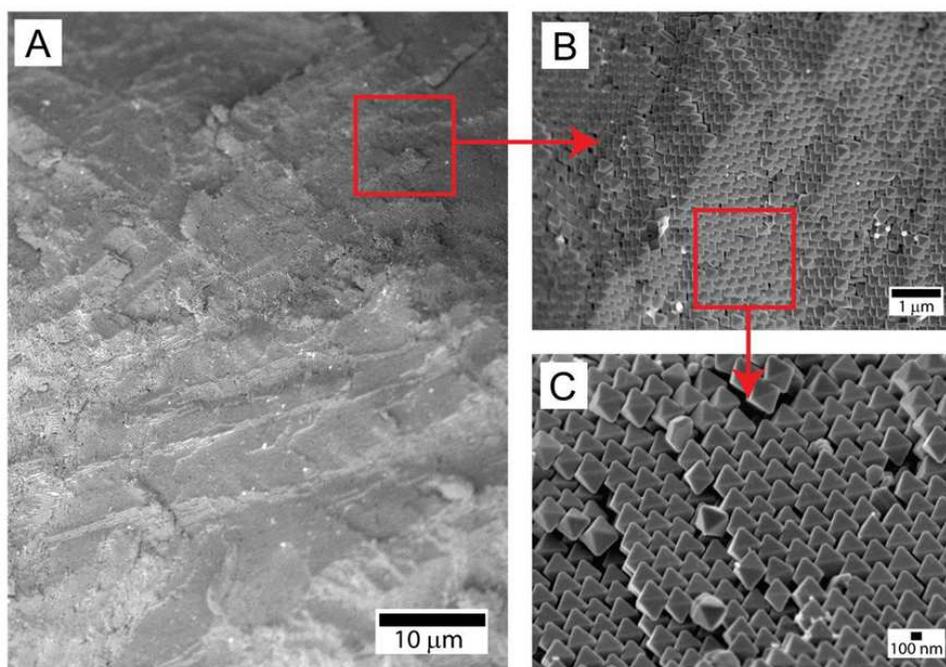
Figure 7.1: SEM images of self-assembled arrays of silver nano-octahedra. Each particle is a single crystal. Image courtesy of A. Tao.

The complicated octahedra arrays were difficult to load into TEMPEST. They

wree first grown in Matlab and saved as a bitmap layout that was later loaded into TEMPEST (a layout is shown in figure 7.2). The octahedra pack in hexagonal arrays; however, TEMPEST uses a rectangular simulation domain. To overcome this difficulty, each period of the TEMPEST simulation contained pieces of nine separate octahedra.

The 3D simulations were long, taking 18 to 40 hours to simulate 6 million cells on a 3 GHz Intel Xeon processor with 16 GB of RAM. The run-time was dominated by the small time-step ($1.93 \times 10^{-4} fs$) required for plasmonic simulations. The long run-times were mitigated by using pulsed methods to compute an entire reflection spectrum in a single simulation. The near-fields of the peaks and valleys in these spectra were spot checked with single frequency simulations. The pulsed simulations necessitated a new broad-band material model for silver. The model added to TEMPEST v7 was a Drude model and implemented with an auxiliary differential equation formulation borrowed from the literature[17]. The background medium, water, was modeled as loss-less with a frequency independent refractive index of 1.2 using the standard Yee update equations.

Reflection spectra were computed by illuminating the array with a pulsed plane-wave and fast Fourier transforming the reflected signal. The incident pulse had a modulated Gaussian temporal profile with a center frequency of 8.3×10^{14} Hz (361 nm wavelength) and a full-width half-max (in field amplitude) bandwidth of 89% of the center frequency. Periodic boundary conditions were used in the plane of the

supercrystal to model an infinitely periodic array. Complex frequency shifted perfectly matched layers were used to terminate the computational grid without reflection in the third dimension. The cell size was chosen as $2nm$ to resolve sharp corners and reduce stair-casing effects.
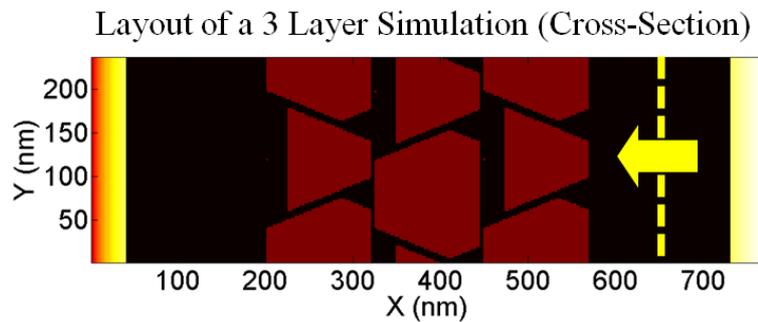
Figure 7.2: A cross-section through the layout of a simulation of three layers of octahedra. The dashed yellow line indicates the source plane. The bright bands on the left and right of the image are PML.

7.2 Simulation Matches Experiment

Simulated reflection spectra matched experimentally obtained reflection spectra when the particles were close packed, but not when the inter-particle spacing was large. Figure 7.3 shows reflection spectra obtained from experiment and from TEM-PEST. In the close packed case (the bottom curves) the reflection spectra exhibited a broad reflection peak bounded by a dip on the long wavelength side (indicated with dashed arrows in the figure). The differences between simulation and experiment arose from inaccuracies in the refractive indices of silver[12] and water, differences in

the illumination angles, and inaccuracies in particles size. TEMPEST used normal incidence illumination while the experiment was setup with 45° illumination. This difference in illumination angle is not as severe as might be expected because the structure is only weakly angle dependent (see section 7.5). Inaccuracies in particle size arose from difficulties in experimentally determining particle size. The simulated particles were $148nm$ on a side and the experimental particles were roughly $140nm$ on a side. Additionally, the inter-particle spacing was difficult to determine experimentally. In TEMPEST, the close packed particles were simulated with a spacing of $6nm$.

Increasing the distance between particles shifted the reflection peak. The experimental curves show the main peak split and eventually the long wavelength peak red-shifted (shown with arrows in the center of figure 7.3). The long wavelength peak likely arose from a photonic crystal-like mode wherein the light penetrated deeply into the array and was confined by the particle facets. Surface plasmon resonances localized to each particle may have influenced this reflection process; however, plasmonic coupling between neighboring particles was negligible due to the large separation distances.

The simulated curves showed strong blue-shifts as inter-particle spacing increased (shown with arrows in the right-most figure) and the intensity of the main reflection peak diminished. Additionally, a large reflection peak blue-shifted into the visible spectrum from the near infrared. This peak wa most likely due to coherent Bragg
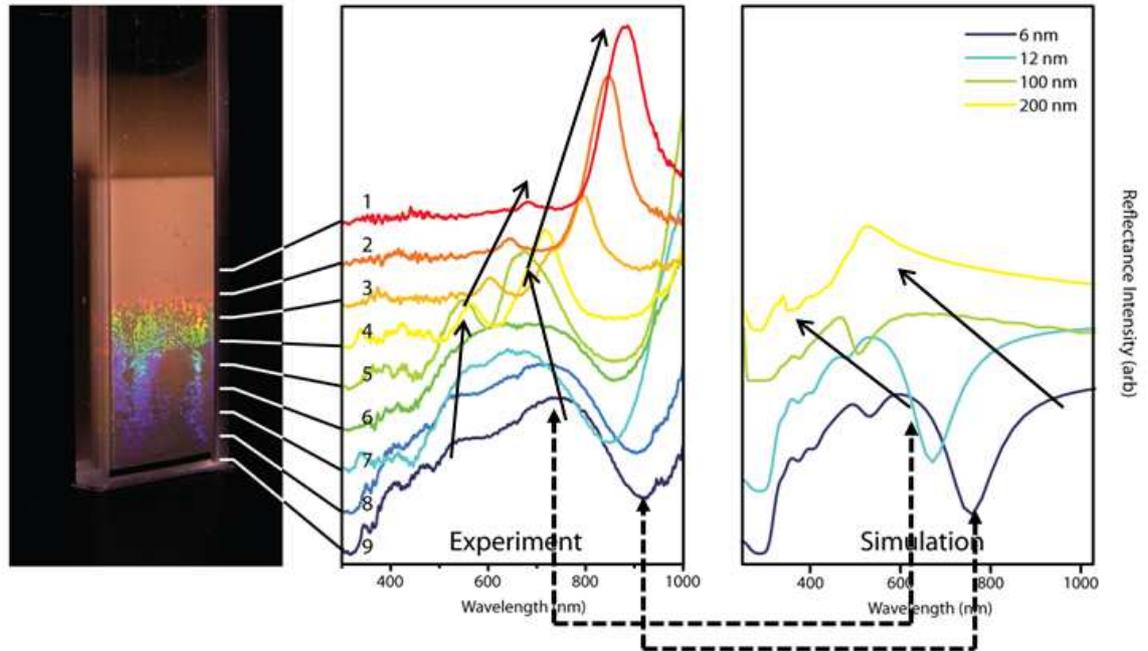
Figure 7.3: (Left) Digital images of the reflection from a many-layered sample of octahedra nano-particles illuminated with white light. The particles have been allowed to settle in solution in a glass cuvette. The particles at the bottom are most densely packed (due to gravity) and the particles at the top are loosely packed. (Center) Experimentally obtained reflection spectra from the sample shown on the left. The bottom curve correspond to close packed particles at the bottom of the cuvette and the top curve corresponds to loosely packed particles at the top of the cuvette. (Right) Reflection spectra computed by TEMPEST for a single layer of octahedra.

scattering off the array of particles and not waveguiding.

The discrepancies between simulation and experiment at large separations came primarily from light scattered from succeeding layers of octahedra. The experimental setup contained many layers whereas the simulations contained only a mono-layer. However, experiment and simulation match in the close packed case because the top layer of crystals shielded the deeper layers. Figure 7.4 shows reflection spectra from one, two, and three layers of close packed octahedra ($6nm$ inter-particle spacing).

The first layer set the size and shape of the curve and the second layer introduces only minor red-shifts in the reflection spectrum. The third layer produces almost no shift at all and suceeding layers are expected to have no effect.
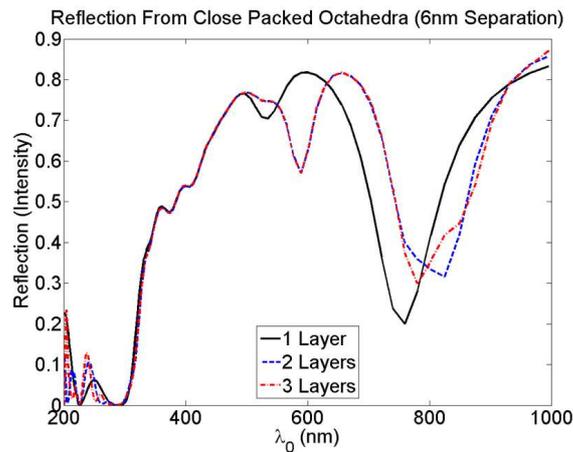
Figure 7.4: Reflection from arrays of octahedra with different numbers of layers.

7.3   Surface Plasmons Drive Resonant Absorbtion and Transmission

Arrays of nano-octahedra were expected to exhibit strong surface plasmons in the visible spectrum because the particles are made of silver. Long lifetime plasmon resonances can be seen in a particle array's temporal response to pulsed excitation using a method originally developed for mapping photonic crystal resonances[14]. Figure 7.5(a) shows the simulated light decaying out of a single layer of a close packed

array of octahedra in response to pulsed plane-wave illumination. Two features are evident in this response: a fast reflection and a slow reflection. The fast reflection looks similar to the incident pulse and came primarily from reflection off the top facets of the particles. The slow reflection came from light that resonated in the interstitial spaces before leaking back into the reflection. The rapid decay of the slow reflection (on the order of $5fs$) indicates that this is a localized process, not a long-range guided wave process characteristic of photonic crystals (which s decay times an order of magnitude longer).
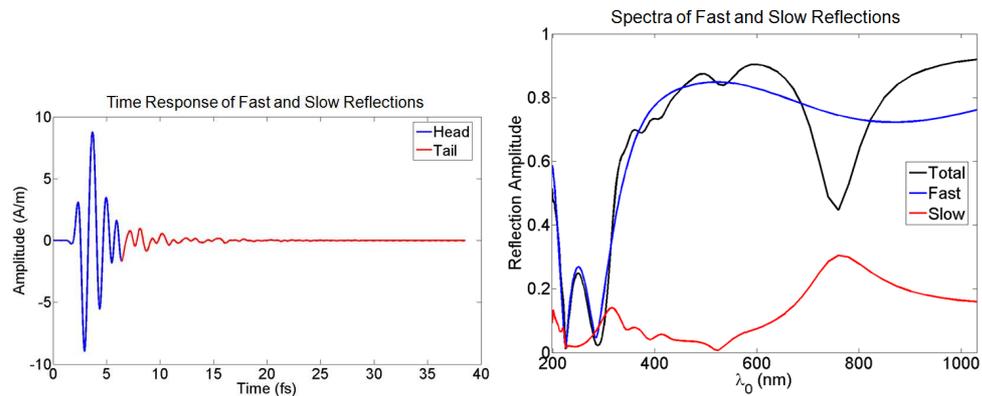
(a) Temporal response.　　(b) Reflection spectrum response.

Figure 7.5: Reflection response of a single layer of close-packed octahedra. Processes with two different speeds are visible: a fast response (blue) and a slow response (red).

The fast and slow reflections are active in different parts of the reflection spectrum. Figure 7.5(b) shows the spectrum of the fast and slow reflections compared to the total reflection spectrum. The fast reflection spectrum exhibits a large plateau, which forms the basis of the broad reflection peak. The short wavelength end of the plateau is

caused by silver's plasma frequency. Below this frequency silver becomes transparent and does not support surface plasmons.

The slow reflection shows a single peak at $760nm$. This light destructively interfered with the fast reflection to generate the long wavelength dip in reflection. This slow process also contributed to transmission and resistive heating. Figure 7.6 shows the reflection, transmission, and resistive heating in a close packed ($6nm$ spacing) array of octahedra. The dominant process ws resistive heating, followed by relatively weak transmission.
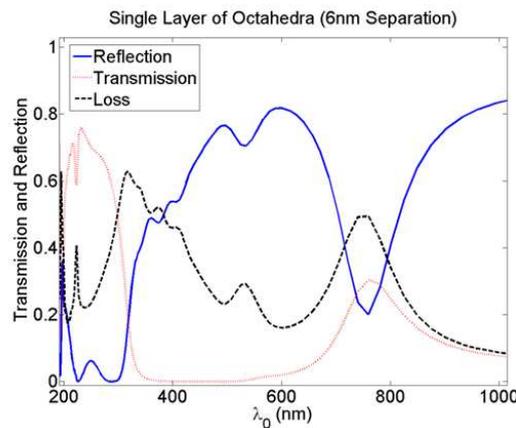
Figure 7.6: Reflection, transmission, and loss (resistive heating) in a close packed array of octahedra. The vertical axis has units of amplitude, not power.

Images of the near-fields show that plasmon resonances in the interstitial spaces were active in the reflection dip. Figure 7.7 shows the near-fields at the top surface and middle cutplane of an array of close packed octahedra in the reflection dip ($\lambda = 760nm$). The near-fields at the top surface show strong field enhancement around the sharp corners of the particles caused by a lightning rod effect. The near-fields

in the middle of the array show strong localization in the interstital gaps and strong coupling between neighboring crystal facets caused by surface plasmon modes. Light was most strongly localized between neighboring facets in a parallel plate waveguide mode, rather than in the horn antenna-like voids formed near the top and bottom surfaces.
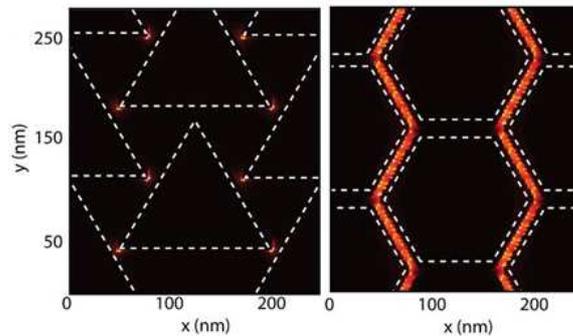
Figure 7.7: Cross-sectional images of the near-fields at the reflection dip ($\lambda = 760nm$) in an array of close packed octahedra. The image on the left shows the fields on the top surface of the array. The image on the right show the fields in the middle of the array. The dashed lines indicate the boundaries of the octahedra.

The interstitial gaps (parallel plate waveguides) can be thought of as weakly coupled resonators. TEMPEST shows that reflection peaks correspond to different coupling symmetries between neighboring. For example, figure 7.8 shows the intensity enhancement in the gap between two layers of octahedra caused by capacitive coupling. The two layers are staggered, causing the strange pattern. All of this light traveled through the interstitial gaps in the top layer and shows the phase relations between neighboring gaps. At long wavelengths (figure 7.8(b)) there are no nodes visible, indicating that each gap was in phase with its nearest neighbors. At short

wavelengths (figure 7.8(a)) nodes were visible, indicating that the gaps were out of phase with one another.
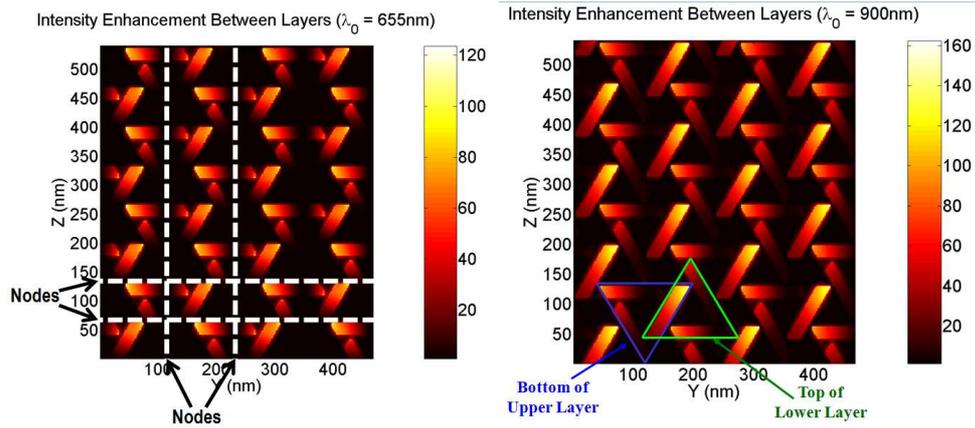
(a) Short wavelength, $\lambda_0 = 655nm$.   (b) Long wavelength, $\lambda_0 = 900nm$.

Figure 7.8: Intensity enhancement between layers in a two-layer array of octahedra spaced at 6 nm. The bright patches correspond to the capacitive coupling between the triangular bottom faces of the top layer and the triangular top faces of the bottom layer. The two layers are offset from one another.

7.4   Field Enhancement

Field enhancement is important for SERS applications. The original hypothesis was that field enhancement would be strongest in the reflection peaks; however, the field enhancement was maximum in the reflection dip due to the interstitial surface plasmon resonances. Table 7.1 shows the maximum near-field enhancement in close packed arrays of octahedra ($6nm$ inter-particle spacing) computed with single frequency TEMPEST simultions. The field enhancement for a single layer of particles

was $2.35\times$ higher in the reflection dip than in the peak.

	Wavelength	Amplitude Enhancement
Single Layer	$600nm$ (peak)	17
	$760nm$ (dip)	40
Double Layer	$655nm$ (peak)	15
	$825nm$ (dip)	28

Table 7.1: The maximum field enhancements throughout an array of close packed octahedra.

The field enhancement numbers shown in table 7.1 are only rough estimates. FDTD simulations face difficulties measuring field enhancement around sharp corners due to stair-casing and the different locations of field components in the staggered grid. The numbers quoted here were computed without interpolating onto a common grid. Performing this interpolation reduces these estimates by roughly a factor of 1.7x.

7.5 Frequency Selective Surfaces

Investigation of dependencies on the angle of incidence and polarization show that both were weak. Periodic structures with these same weak dependencies have been created for radar applications. They are called frequency selective surfaces and they are utilized to hide antenna on stealth aircraft[48].

As shown in section 7.3 the interstitial resonances were localized and therefore their dependence on angle of incidence as expected to be weak. This was tested by

computing the reflection 36° off-axis in the y direction (axes shown in figure 7.8). The results, shown in figure 7.9, demonstrate that this change in incidence angle did not change the shape of the reflection spectrum, shift the main reflection peak or dip, nor did it significantly change the total reflected intensity. It did introduce a small splitting in the polarizations; however, this structure is only weakly polarization dependent due to its approximately hexagonal symmetry.
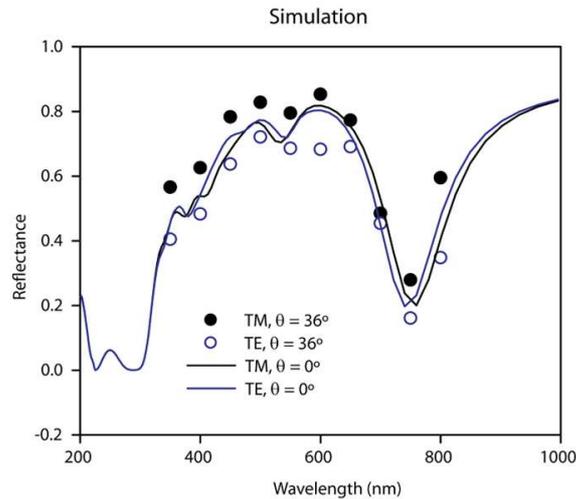
Figure 7.9: Reflection spectra on-axis and 36° off-axis from a single layer of octahedra under both polarizations.

7.6   Conclusion

FDTD simulation is a useful tool for uncovering the role of surface plasmon in periodic arrays of silver nanoparticles. Pulsed methods were used to minimize the number of 18 to 40 hours simulations. The run-time was dominated by the small

time-step. Pulsed simulations required the implementation of a broad-band model of silver in TEMPEST v7. These studies required the addition of a broad-band silver model to TEMPEST. TEMPEST simulations correspond well with experimental results and overturned the hypothesis that interstitial surface plasmon resonances drive broad reflection peaks. It was discovered that these resonances drive resistive heating, transmission, and field enhancements of 40x. Additionally, the angle of incidence and polarization dependencies were found to be weak due to the localized nature of the resonances driving the reflection process.

Chapter 8

The Sub-Wavelength Grating

The Sub-Wavelength Grating (SWG), figure 8.1, is a simple 1D dielectric grating that forms a surprisingly high reflectivity mirror over a broad bandwidth for the TM polarization. The reflection spectrum (figure 8.2) is centered around the $1.5\mu m$ wavelength, important for fiber optic communications. The device is lightweight and forms the basis of a tunable vertical cavity surface emitting laser (VCSEL)[24][25][31]. The device has been demonstrated in simulation and experiment [45] [11]; however, uncertainty remains about whether the device is best intuitively understood as a coupled wave system or a series of weakly coupled resonators.

Highly resonant systems like the SWG are often described in terms of either a coupled wave or coupled resonator model. These models are typically used to describe waves traveling horizontally along the structure. In the limit of small gaps, the coupled wave model treats the SWG as a solid silicon waveguide with waves traveling
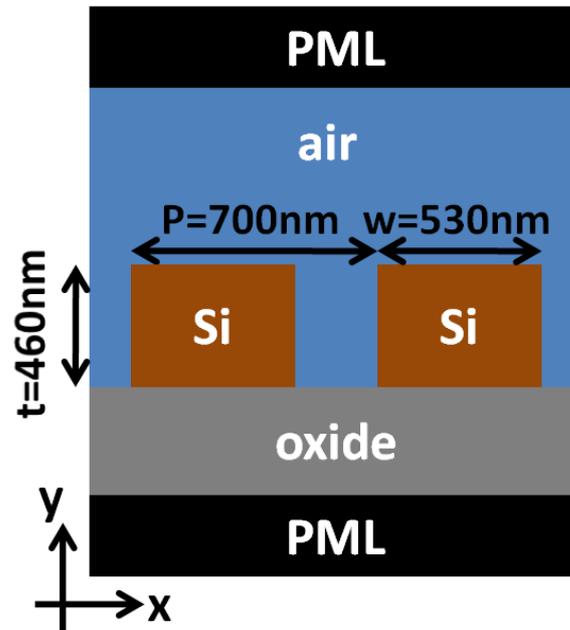
Figure 8.1: Layout of the sub-wavelength grating. The structure is composed of three materials which were modeled with frequency independent refractive indices: air ($n_{air} = 1$), SiO_2 ($n_{oxide} = 1.47$), and silicon ($n_{Si} = 3.48$). The grating is infinite but only two periods are shown.

horizontally. The small, periodic breaks couple and reflect the forward and backward traveling waves. This model has difficulty with the strong coupling caused by large gaps in the SWG. In the coupled resonator optical waveguide (CROW) [75] model each bar is treated as an individual resonator that localizes light and very weakly couples to its neighbors. This model has difficulty with the sub-wavelength gaps in the SWG, which make coupling strong enough to shift the resonances. This chapter shows that these two resonance pictures do not accurately describe the reflection band. A model based on approximating the SWG as a layered and representing the fields as vertically traveling Bloch waves is developed. This model is shown to accurately

reproduce the high reflection band.

8.1 Computing the Reflection Spectrum

Due to TEMPEST's origins in photolithography, previous studies of the SWG computed reflection spectra point by point with continuous-wave simulations[45]. However, this process is prone to missing narrow features in the reflection spectrum. TEMPEST v7's pulsed methods compute the entire reflection spectrum in one simulation and catch all the narrow features. For example, the SWG's reflection spectrum computed with both single frequency and pulsed methods is shown in figure 8.2. If too few single frequency simulations are used the pair of spikes at short wavelengths can be missed. However, the pulsed simulations can take longer to complete. In this example, the pulsed simulation took 5 minutes and 22 seconds to complete whereas each single frequency simulation completed in 7 seconds (on a 3 GHz Intel Xeon processor with 16 GB of RAM). The simulations must iterate through many time steps in order to achieve the spectral resolution necessary to resolve the narrow peaks and to allow the waves to converge in this highly resonant system. In this example, and all the work discussed below, the cell size was $\Delta x = 10$ nm and the time step was $\Delta t = 1.93 \times 10^{-3}$ fs.

The pulsed simulations illuminated the grating with a plane-wave source placed above the structure that emitted a pulse with a modulated Gaussian temporal profile. This pulse had a center frequency of 2.10×10^{14} Hz (corresponding to a wavelength of
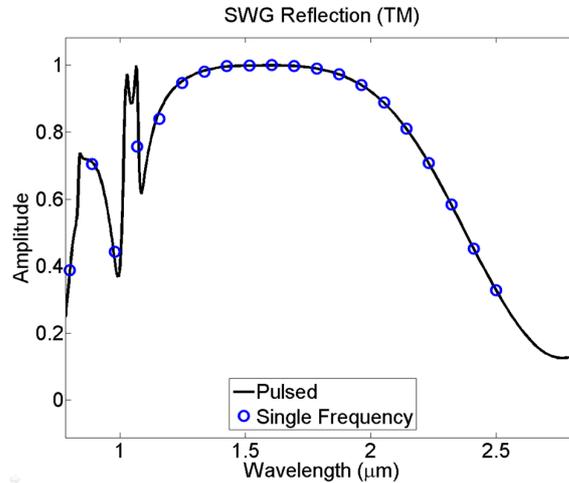
Figure 8.2: The SWG's reflection spectrum calculated with single frequency simulations (circles) and a single pulsed simulation (solid line).

$1.43\mu m$) and a full-width half-max of 1.81×10^{14} Hz. Monitor points placed along the top surface of the grating recorded the time response of the H_z field component. To compute the reflection spectrum, the temporal responses were first averaged across the top of the grating to remove all waves except the propagating, on-axis wave:

$$H_z^{On-Axis}(n) = \frac{1}{M} \sum_{m=0}^{M} H_z(x_m, n), \tag{8.1}$$

where n is the time step number, m is the monitor point number, M is the total number of monitor poitns, $H_z(x_m, t)$ is the signal recorded at monitor point m, and $H_z^{On-Axis}(t)$ is the on-axis signal. Second, the incident wave (computed in a separate simulation) was subtracted off to yield only the back-scattered signal, $H_z^R(t) = H_z^{On-Axis}(t) - H_z^{Incident}(t)$. Third, the result was discrete Fourier transformed, with Matlab's fast Fourier transform (FFT) routine, to get the spectrum of the back-scattered response. The simulation was run through $200,000$ time steps so

the high-Q resonances had time to decay and the FFT was not corrupted by trunca-
tion error. The temporal signal was additionally zero-padded to 500, 000 points for
frequency resolution (an operation handled by the FFT routine).

$$
\bar{H}_z^R(n) \;=\; \begin{cases} H_z^R(n), & 0 \leq n \leq 199,999. \\[2mm] 0, & 200,000 \leq n \leq 499,999. \end{cases}
$$

$$
s_R(\omega) \;=\; \frac{1}{500,000} \sum_{n=0}^{499,999} \bar{H}_z^m(n) e^{-j\omega n \Delta t}, \tag{8.2}
$$

where $\bar{H}_z^R(n)$ was the zero-padded back-scattered signal and $s_R(\omega)$ is the spectrum of
this signal. Finally, the back-scattered spectrum was divided by the incident signal's
spectrum to yield the reflection spectrum:

$$
R(\omega) = \frac{s_R(\omega)}{s_{Incident}(\omega)}, \tag{8.3}
$$

where $R(\omega)$ was the reflection spectrum (shown in figure 8.2) and $s_{Incident}(\omega)$ was the
incident wave's spectrum.

8.2   Resonances in the SWG

The coupled wave and CROW models are designed to predict resonances. Reso-
nances are source-free solutions to Maxwell's equations and respresent long lifetime
paths through the structure. Since these resonances decay slowly they can be mapped
with TEMPEST v7's pulsed methods to determine if they influence the broad re-
flection band. The resonances in the SWG were mapped with TEMPEST and the

resonances in the high reflection band were found to possess the wrong symmetry to couple to radiation at normal incidence. Therefore resonances do not drive the high reflection band. Consequently, the CROW and coupled wave models do not explain the large reflection band.

However, because of the large amount of interest in resonant optical systems, the accuracy of the CROW and coupled wave models' predictions of the resonances in the SWG was investigated. These models' accuracies were determined by comparing the spectral locations and near-field patterns of the SWG resonances with the coupled wave and CROW predictions. CROW was found to accurately predict the spectral locations of the resonances with light centered in the bars but did not account for resonances with light centered in the bars or explain why some SWG resonances are coupled in phase to their nearest neighbors and some are out of phase. Coupled wave theory accounted for the location of light concentration, the near field patterns of the resonances, and the nearest neighbor coupling. However, the coupled wave theory used here over predicted the resonances' wavelengths and predicted resonances centered in the gaps that did not show up in TEMPEST.

Previous work showed that TEMPEST simulations match both experimental measurements [44] and RCWA predictions [45]. TEMPEST v7 was used to map the resonance spectrum with a method very similar to the method for computing reflection spectra discussed above. Instead of a plane-wave source, a point source introduced a pulse of light into the structure and monitor points placed throughout the structure

recorded the temporal response (the "pinging" method discussed by G. Burr[62]).
Floquet boundary conditions were used to control the phase between nearest neighbors. A discrete Fourier transform of the temporal responses yielded the resonance spectrum, which contained peaks where the long-lived resonances lie. These simulations had the same modulated Gaussian pulse, cell size, and time step size as the reflection spectrum simulations. The source point was placed in the upper right corner of the bar to avoid nulls in the low-order resonances. Multiple monitor points were used to avoid nulls in the near-field resonance patterns as well. Each monitor point had a different sensitivity to each resonance, which was determined by the resonance's near-field pattern. To maximize the liklihood of finding every resonance, the spectrum of the signal recorded at each monitor point was separately computed and then all the spectra were combined with a maximum function. The spectra were computed with a Fast Fourier Transform (FFT) zero-padded to $500,000$ points for frequency resolution and windowed with a Hann function to remove noise and sidelobes. (8.4) gives the equations used to compute a resonance spectrum.

$$
\bar{H}_z^m(n) = \begin{cases} H_z^m(n), & 0 \le n \le 199,999. \\ 0, & 200,000 \le n \le 499,999. \end{cases}
$$

$$
w(n) = \begin{cases} 0.5\left(1 - cos\left(\frac{2\pi n}{200,000}\right)\right), & 0 \le n \le 199,999. \\ 0, & 200,000 \le n \le 499,999. \end{cases}
$$

$$
s(\omega) = \max_m \left(\left| \sum_{n=0}^{499,999} \bar{H}_z^m(n) w(n) e^{-j\omega n \Delta t} \right| \right), \tag{8.4}
$$

where $H_z^m(n)$ is the temporal response recorded at monitor point m, $\bar{H}_x^m(n)$ is the

zero-padded temporal response (this is automatically performed in Matlab's FFT routine), $w(n)$ is the Hann window function (implemented in Matlab's hann routine), Δt is the time-step size, and $s(\omega)$ is the resonance spectrum. For on-axis illumination, periodic boundary conditions are used on the left and right boundaries.

A similar procedure was used to compute the near-field patterns for each resonance. The difference is that every point inside the bar was monitored. Each point's signal was separately FFT-ed and the terms corresponding to a resonance were selected to form the resonance pattern.

$$\bar{H}_z(x,y,n) = \begin{cases} H_z(x,y,n), & 0 \leq n \leq 199,999. \\ \\ 0, & 200,000 \leq n \leq 499,999. \end{cases}$$

$$w(n) = \begin{cases} 0.5\left(1 - cos\left(\frac{2\pi n}{200,000}\right)\right), & 0 \leq n \leq 199,999. \\ \\ 0, & 200,000 \leq n \leq 499,999. \end{cases} \tag{8.5}$$

$$s(x,y,\omega) = \sum_{n=0}^{499,999} \bar{H}_z(x,y,n)w(n)e^{-j\omega n\Delta t}, \tag{8.6}$$

where $H_z(x,y,n)$ is the temporal response at every point in the bar, $\bar{H}_x(x,y,n)$ is the zero-padded temporal response, $w(n)$ is the Hann window function, Δt is the time-step size, and $s(x,y,\omega)$ is the resonance spectrum. The resonance pattern,e.g. $|s(x,y,\omega_i)|$ for the ith resonance, was obtained by selecting a frequency term corresponding to a resonance.

Figure 8.3 shows the resonance spectrum of the SWG, $|s(\omega)|$, when all neighbors were in phase and 180° out of phase. In the in phase cse, two resonances lie in the middle of reflection band and two resonances lie just below it. Figures 8.4(a) and (a)

show the near-field patterns of these two resonances. The vertical nulls in the middle of the bar show that these patterns have odd symmetry and since the pitch is less than the wavelength they cannot couple to on-axis plane-waves. The fundamental resonance (one bright point in the center of the bar) does not appear. Figures 8.4(c) and (d) show the near-field patterns for the two resonances just below the reflection band. The resonance at $1.0812\mu m$ couples to on-axis plane-waves because it has even symmetry. The resonance at $1.0384\mu m$ has odd symmetry; however, its first grating orders are just below cutoff in the oxide and so this resonance affects the reflection spectrum.
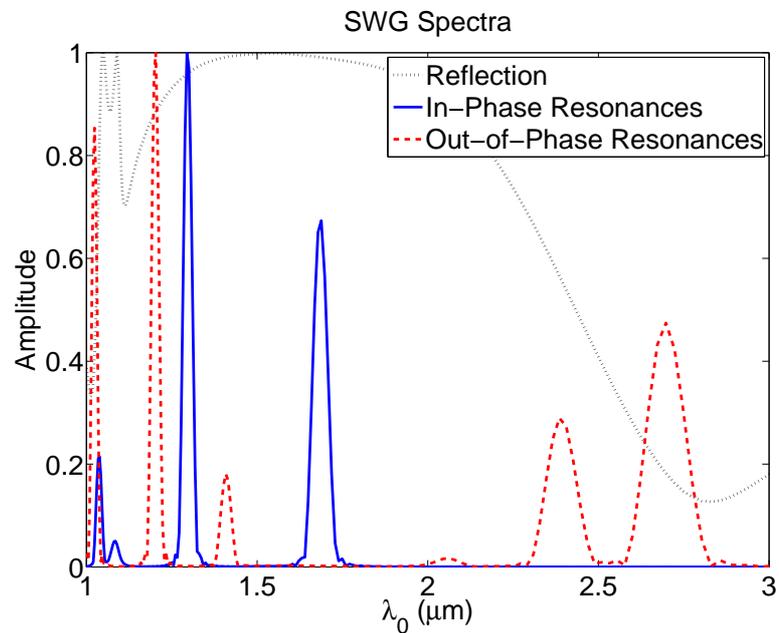
Figure 8.3: SWG resonance spectrum showing resonances with the nearest neighbors in-phase and 180° out-of-phase. The SWG reflection spectrum is shown for reference. The corresponding near-field patterns are shown in Figures 8.4 and 8.5.

(a) $\lambda_0 = 1.69 \mu m$.

(b) $\lambda_0 = 1.30 \mu m$.

(c) $\lambda_0 = 1.08 \mu m$.

(d) $\lambda_0 = 1.04 \mu m$.

Figure 8.4: The near-field patterns of the SWG resonances when nearest neighbors are in-phase. The images show the amplitude of the H_z field component (out of the page). White indicates bright and black indicates dim. The pattern shown in (c) is a super-position of multiple resonances. The corresponding resonance spectrum is shown in figure 8.3.

The first orders turn on when the wavelength in the oxide equals the pitch of the structure:

$$\lambda_0 = N_{oxide} \times P, \tag{8.7}$$

where λ_0 is the free-space wavelength, $N_{oxide} = 1.47$ is the refractive index of the oxide, and $P = 700nm$ is the pitch. In the SWG the first orders turn on at $\lambda_0 = 1.029\mu m$. This number may be slightly altered by numerical dispersion.

The 180° out of phase case shows more resonances than the in phase cse. Figure 8.5 shows the corresponding near-field patterns. These modes cannot couple to normal incidence radiation due to symmetry and the fact that the pitch is less than the wavelength, but they are useful for determining the applicability of the CROW and coupled wave models. The lowest order resonance appears at $2.70\mu m$ (Figure 8.5(a)) and a similar resonance, but with the light centered over the gap rather than the bar, appears at a shorter wavelength (Figure 8.5(b)).

(a) $\lambda_0 = 2.70\mu m$.

(b) $\lambda_0 = 2.39\mu m$.

(c) $\lambda_0 = 1.41\mu m$.

(d) $\lambda_0 = 1.20\mu m$.

(e) $\lambda_0 = 1.02\mu m$.
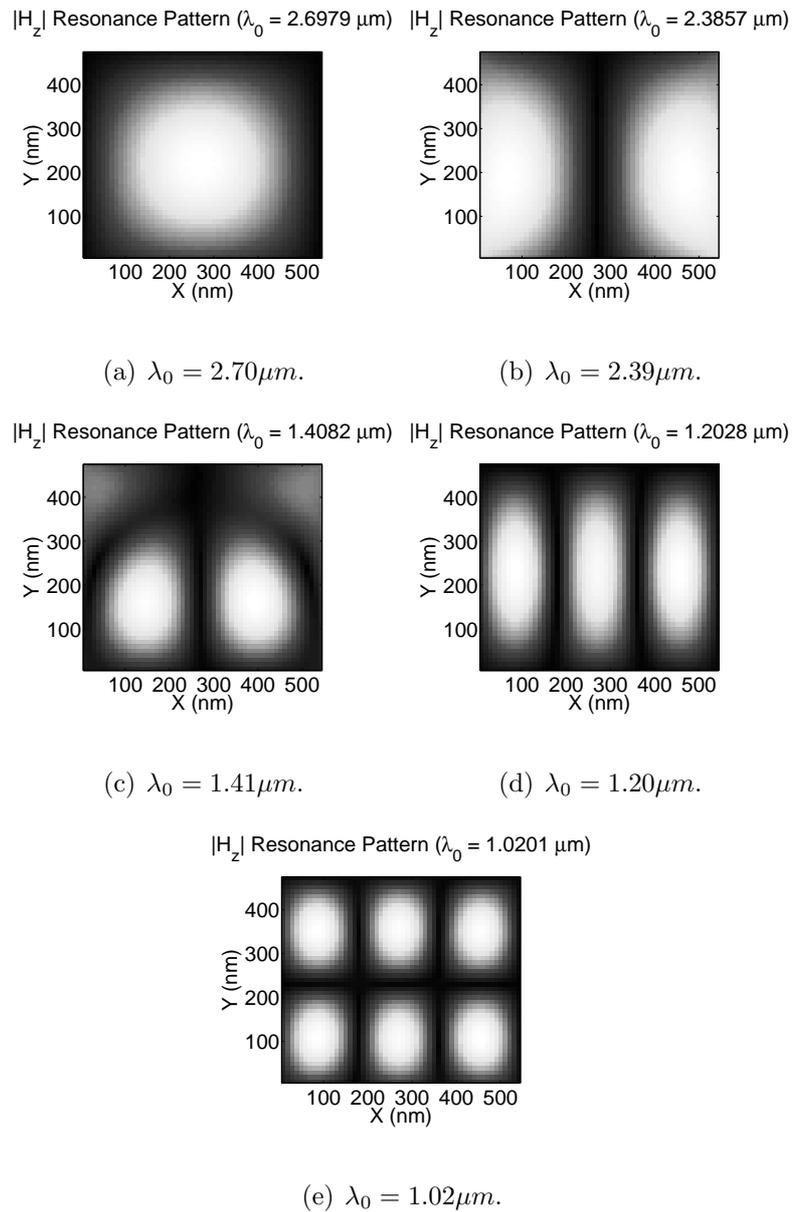
Figure 8.5: The near-field patterns of the SWG resonances when nearest neighbors are 180° out-of-phase. The pattern shown in (c) is a super-position of multiple resonances. The corresponding resonance spectrum is shown in figure 8.3.

8.2.1 The CROW Picture

The large refractive index contrast in the SWG and the broad width of the reflection peak led to the hypothesis that the bars act as low-Q isolated resonators, with weak coupling between nearest neighbors, like the resonators used to design CROW's. The resonances of an isolated bar were measured with TEMPEST and found to correspond to some of the resonances in the SWG – specifically the ones with light centered in the bars. However, the SWG exhibits resonances with light centered between the bars which indicates a high degree of coupling that is not handled by the CROW model. These types of resonances are predicted by the coupled wave model, which is discussed in the next section. Additionally, CROW model predicts a resonance corresponding to the fundamental resonance of the isolated bar coupled in phase to its nearest neighbors, but the SWG does not exhibit this resonance.

The CROW model is based on the resonances of an isolated bar. These resonances have patterns with light concentrated inside the bar. The model predicts the resonances seen in the SWG will fall at roughly the same spectral location as the resonances of the isolated bar. Weak coupling between the bars splits the resonances into pairs, with in phase coupling between nearest neighbors appearing at a slightly longer wavelength and 180° out of phase coupling appearing at a slightly shorter wavelengths.

The resonances of the isolated bar were mapped with the methods discussed above. The simulation contained a single bar from the SWG (with a width of $530nm$ and

a height of $460nm$) sitting on top of a semi-infinite oxide layer and bounded by air above. PML absorbed the outbound waves along the x and y boundaries.

Figure 8.6 shows a comparison of the SWG's in phase and out of phase resonance spectra with the resonance spectrum from an isolated bar. The SWG resonances exhibit higher Q's than the isolated bar's resonances (indicated by the resonances' linewidths). Figure 8.6 shows the near-field patterns of the isolated bar's resonances. These patterns show that isolated bar's resonances corresponded to four resonances in the SWG. The overlapping resonances were:

1. the isolated bar's fundamental resonance at $2.65\mu m$ and the SWG's longest wavelength out of phase resonance at $2.70\mu m$ (Figures 8.7(a) and 8.5(a)),

2. the isolated bar's second resonance at $1.74\mu m$ and the SWG's first in phase resonance at $1.70\mu m$ (Figures 8.7(b) and 8.4(a)),

3. the isolated bar's third resonance at $1.24\mu m$ and the SWG's fourth out of phase resonance at $1.20\mu m$ (Figures 8.7(c) and 8.5(d)), and

4. the isolated bar's fourth resonance at $1.05\mu m$ and the SWG's fifth out of phase resonance at $1.02\mu m$ (Figures 8.7(d) and 8.5(e)).
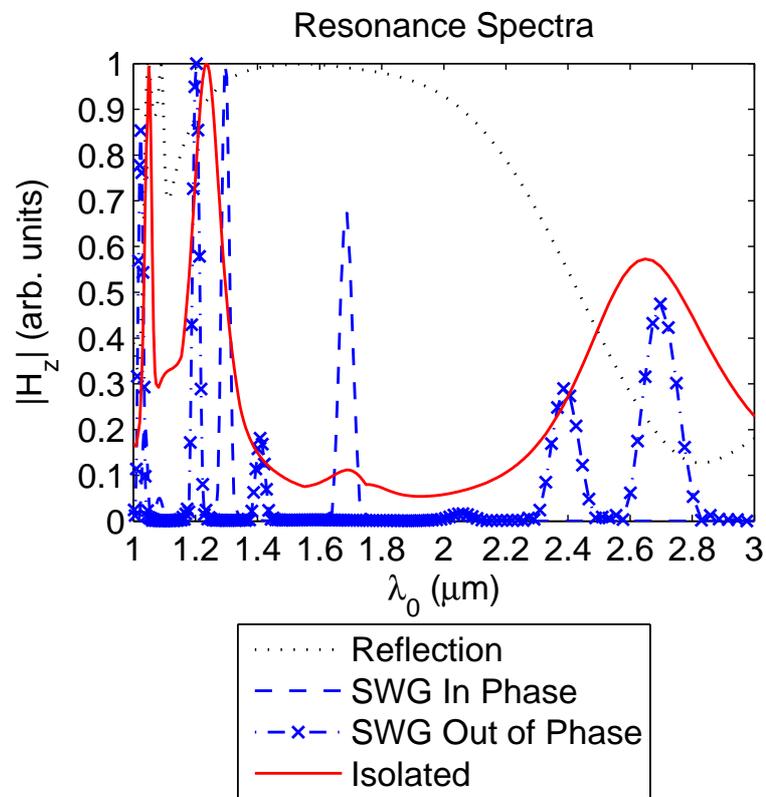
Figure 8.6: The SWG resonance spectrum with the nearest neighbors in-phase, the resonance spectrum of an isolated bar, and the SWG reflection spectrum. The corresponding near-field patterns are shown in figure 8.7.

(a) $\lambda_0 = 2.65\mu m$.

(b) $\lambda_0 = 1.74\mu m$.

(c) $\lambda_0 = 1.24\mu m$.

(d) $\lambda_0 = 1.05\mu m$.

Figure 8.7: The near-field patterns of the isolated bar's resonances. The pattern shown in (c) is a super-position of the 3×1 and 1×2 resonances. The corresponding resonance spectrum is shown in figure 8.6.

Even though the isolated bar has many resonances in common with the SWG the CROW picture does not explain why some of these resonances match the SWG's in phase resonances and some match the out of phase resonances. The SWG shows additional resonances that do not appear in the isolated bar; most notably the second out of phase resonance at $2.39\mu m$ (Figure 8.5(b)). This resonance has light centered in the gap between bars, a distribution not predicted by the CROW model. This distribution is contained in the coupled wave picture [32]. Additionally, the in-phase coupling of the fundamental resonance predicted by CROW does not appear in the SWG resonances, but this missing resonance is explained by cutoff in the coupled wave picture.

8.2.2 The Coupled Wave Picture

Previous work has shown that guided waves in photonic crystals give rise to narrow features in the reflection spectrum[14], like the narrow peaks in the short wavelengths of the SWG's reflection spectrum (Figure 8.2). As shown in the previous section the SWG resonances have a higher Q than predicted by the isolated resonator model, which is more in line with the guided waves seen in photonic crystals. Additionally, the SWG contains resonances with light centered in the gaps – a distribution predicted by coupled wave theory but not CROW.

In the coupled wave model the SWG is viewed as a slab waveguide with waves running to the left and right (figure 8.8(a)). These waves are coupled by the breaks

in the guide. This model predicts resonances that are distinguished by the number of nulls in the x and y directions and by the location of the light (concentrated in either the bars or the gaps). The resonances come in pairs spectrally centered around the crossing points (band gaps) at $\beta_x = 0$ (nearest neighbors in phase) and $\beta_x = \pi/P$ (nearest neighbors out of phase). In a resonance pair, the longer wavelength resonance concentrates light in the bars and the shorter wavelength resonance concentrates light in the gaps. The number of nulls in the y direction is determined by the corresponding wave in the slab waveguide. Wave 1 has no nulls in the y direction, wave 2 has one null, etc. As the wavelength decreases the resonances show more nulls in the x direction. The first time a wave's dispersion relation hits the $\beta_x = 0$ or π/P points the resonances exhibit only one null per period. The second time, two nulls, and so on.

The coupled wave dispersion relation shows where the resonances are predicted to lie. This dispersion relation is based on the slab waveguide's dispersion relation. In this work, the dispersion relation for the slab waveguide was numerically computed by looking for guided solutions to Maxwell's equations with a 2π round-trip phase. The numerical method utilized a combination of Snell's Law, Fresnel reflection coefficients, and the transmission line method. For example, the round-trip phase for the TM case was computed with:

$$\theta_{air}(\theta_{Si}) = arccos\left(\frac{n_{Si}}{n_{air}}cos(\theta_{Si})\right) \tag{8.8}$$

$$\theta_{ox}(\theta_{Si}) = arccos\left(\frac{n_{Si}}{n_{ox}}cos(\theta_{Si})\right) \tag{8.9}$$

$$\phi_{air}(\theta_{Si}) = angle\left(\frac{-n_{Si}sin(\theta_{air}) + n_{air}sin(\theta_{Si})}{n_{air}sin(\theta_{Si}) + n_{Si}sin(\theta_{air})}\right) \quad (8.10)$$

$$\phi_{ox}(\theta_{Si}) = angle\left(\frac{n_{ox}sin(\theta_{Si}) - n_{Si}sin(\theta_{ox})}{n_{ox}sin(\theta_{Si}) + n_{Si}sin(\theta_{ox})}\right) \quad (8.11)$$

$$\beta_y(\theta_{Si}) = \frac{2\pi n_{Si}}{\lambda_0}sin(\theta_{Si}) \quad (8.12)$$

$$\beta_x(\theta_{Si}) = \frac{2\pi n_{Si}}{\lambda_0}cos(\theta_{Si}) \quad (8.13)$$

$$\phi_{rt}(\theta_{Si}) = \phi_{air}(\theta_{Si}) + \phi_{ox}(\theta_{Si}) + 2\beta_y(\theta_{Si})t. \quad (8.14)$$

θ_{Si} is the propagation angle in the Silicon guide. To find guided modes (which propagate in the Silicon but not in the air or oxide), this angle was scanned from $0°$ to $arccos\left(\frac{n_{ox}}{n_{Si}}\right)$ (radiation cutoff angle in the oxide). (8.8) and (8.9) are Snell's Law, which yield the corresponding angles in the air and oxide, $\theta_{air}(\theta_{Si})$ and $\theta_{ox}(\theta_{Si})$. These angles are imaginary for guided modes. n_{Si}, n_{ox}, and n_{air} are the refractive indices of the silicon, oxide, and air respectively. (8.10) and (8.11) are the phase of the Fresnel reflection coefficients at the top and bottom surfaces of the guide. The phase is caused by the reactive reflection in the the evanescent tails in the air and oxide. t is the thickness of the guide. (8.12) is the propagation constant in the guide transverse to the direction of propagation and it was used to compute the phase incurred as the guided wave moved from the top of the guide to the bottom and back again. (8.14) is the total round-trip phase. This was monitored as θ_{Si} was changed and whenever it reached a multiple of 2π a mode was recorded with propagation constant $\beta_x(\theta_{Si}^m)$ (from (8.13)). This process was repeated for different wavelengths, λ_0, to form the

dispersion relation:

$$\beta_x^{Slab}(\lambda_0) = \sum_m \beta_x(\lambda_0, \theta_{Si})\delta(\theta_{Si} - \theta_{Si}^m), \tag{8.15}$$

where m is the mode number and δ is a Dirac delta.

The dispersion relation for these modes is shown in figure 8.8(b). There are three TM modes in the spectrum of interest. These modes are characterized by their number of nulls in the y direction. For example, mode 1 has no nulls in the H_z pattern and mode 2 has one null.
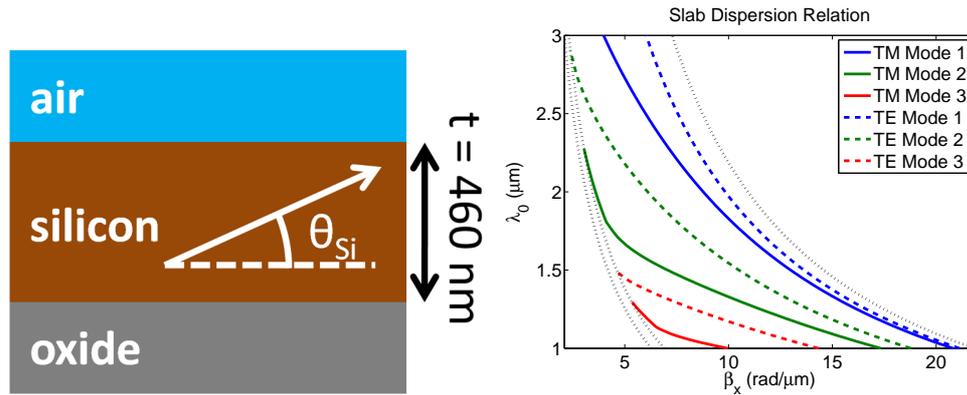
(a) Slab guide layout. (b) Dispersion relation.

Figure 8.8: Layout of the silicon slab waveguide and its dispersion relation. The three dotted lines in the dispersion relation are the light cones. The left-most corresponds to free-space, the middle corresponds to the oxide, and the right-most corresponds to silicon.

The simplest guided wave model of the SWG is the empty lattice approximation, which is simply a periodic repetition of the slab's dispersion relation (8.15) and its

mirror image.

$$\lambda_0^{EL}(\beta_x) = \sum_{n=-\infty}^{+\infty} \left(\lambda_0^{Slab}\left(\beta_x + \frac{2\pi n}{P}\right) + \lambda_0^{Slab}\left(-\beta_x + \frac{2\pi n}{P}\right) \right), \qquad (8.16)$$

where $\lambda_0^{EL}(\beta_x)$ is the empty lattice approximation to the SWG's dispersion relation (8.15) and P is the SWG's pitch. The resulting dispersion relation is shown in figure 8.9. This model gives the approximate locations of the mode crossing points (shown in Table 8.1). If the model is accurate, the resonance pairs will straddle these points. This model neglects coupling and consequently gives no estimate of the degree of splitting. Additionally, this model does not take into account blue-shifting due to the removal of silicon to form the gaps.
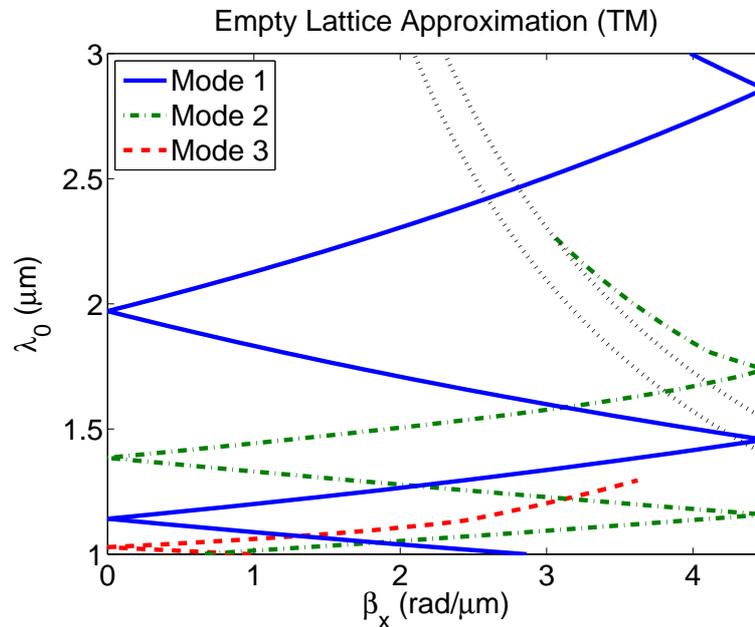

Figure 8.9: Approximate SWG dispersion relation formed with the empty lattice approximation. The dotted curves in the upper right are light cones. The left-most one corresponds to free-space and the right-most one corresponds to the oxide.

Slab Mode	Coupling	Wavelength	
1	out of phase	$2.86\mu m$	$1.46\mu m$
	in phase	$1.97\mu m$	$1.14\mu m$
2	out of phase	$1.74\mu m$	$1.16\mu m$
	in phase	$1.39\mu m$	$1.03\mu m$
3	out of phase	—	
	in phase	$1.03\mu m$	

Table 8.1: Locations of the slab mode crossing points ($\beta_x = 0$ and π/P) as predicted by the empty lattice approximation.

The empty lattice approximation correctly predicts the spectral order of the resonances and their near-field patterns. For example, the empty lattice dispersion relation predicts the longest wavelength resonances (the crossing appearing in the upper right of Figure 8.9) should correspond to mode 1 in the slab (no nulls in the y direction) and nearest neighbors should be out of phase. Additionally, this is the first time mode 1's dispersion relation hit a folding point and consequently no nulls in the x direction should appear. The SWG resonance spectra measured by TEMPEST confirms that the longest wavelength pair of resonances indeed has out of phase nearest neighbors, Figure 8.3. The near-field patterns computed by TEPEST, Figures 8.5(a) and 8.5(b), show no nulls in the y direction, indicating this resonance arose from slab mode 1. The longer wavelength resonance also shows only nulls in the x direction at the edges of the bar (one null per period) and its light is centered in the bar, two more predictions of coupled wave theory confirmed. The shorter wavelength resonance is centered over the gap and again shows only one null per period, as predicted by coupled wave theory. The empty lattice approximation predicts this pair of resonances

to be centered at $2.86\mu m$ (Table 8.1); however, the resonances actually lie at $2.70\mu m$ and $2.39\mu m$. This blue-shifting was caused by the removal of silicon to form the gaps. The empty lattice approximation's predictions would be improved by including coupling.

The second longest wavelength pair of resonances predicted by the empty lattice approximation is again mode 1 but with nearest neighbors in phase (upper left side of the dispersion relation in Figure 8.9). The SWG resonance spectrum measured by TEMPEST confirms that nearest neighbors are in phase (Figure 8.3). The near-field pattern (Figure 8.4(a)) again shows no nulls in the y direction, indicative of slab mode 1. This time there are two nulls per period in the x direction, indicating this is the second time slab mode 1's dispersion relation has encountered a folding point. Coupled wave theory predicts a corresponding resonance at shorter wavelengths with light centered over the gap, but this resonance does not appear in the TEMPEST measurements. Two possible explanations for its disappearance are that it shifted too close to another resonance (e.g. Figure 8.4(b)) and was not detected or the gaps between the bars are so large that the light is not confined. Again, a blue-shift is evident between the empty lattice approximation of the center of the resonance pair ($1.97\mu m$) and the location of the resonance ($1.69\mu m$).

All of the resonances measured by TEMPEST correspond to empty lattice crossings. The third, out of phase, crossing for mode 1 shows up as an out of phase resonance at $1.20\mu m$ (Figure 8.5(d)). Slab mode 2 resonances (characterized by one

null in the y direction) appear at $1.41 \mu m$, $1.30 \mu m$, and $1.02 \mu m$ (Figures 8.5(c), 8.4(b), and 8.5(e) respectively). A slab mode 3 resonance (characterized by two nulls in the y direction) appears at $1.04 \mu m$ (Figure 8.4(d)). All of these resonances have light centered in the bars and are missing their counterparts with light centered in the gaps. Additionally, they show blue-shifts relative to the empty lattice approximations.

The coupled wave model accurately predicts the properties of the SWG's resonances, including the spectral order of the resonances, coupling to nearest neighbors, number of nulls, and where the light is concentrated. However, this model predicts some resonances with light centered in the gaps that were not found in TEMPEST. Additionally, the empty lattice approximation neglected coupling and so predicted the resonances to lie at longer wavelengths than TEMPEST measured.

8.3   SWG Topography as a Layered Medium

The previous section explored the resonances in the SWG and found that the resonances do not affect the reflection spectrum due to symmetry. Resonances are source-free solutions to Maxwell's equations; however, reflections are a forced (i.e. particular) solution to Maxwell's equations. This section relaxes the source-free constraint and models the silicon bars as a segment of an infinitely tall layered medium. The layered medium contains alternating vertical strips of silicon and air. In general waves travel through the structure in the x and y directions. To find the self-consistent wave vectors, a Bloch formulation was applied in the x direction and constrained to

zero phase progression to match the normal incidence source condition. It is shown that two modes with the correct symmetry drive the high reflectivity band.

Using a modal picture originally developed for microwave waveguide analysis [71] the SWG structure was broken down into three regions (see figure 8.10): the cladding (air), a layered medium (the grating region), and the substrate (oxide). Initially each region was considered by itself and assumed to be infinite in extent. The characteristic modes of each region are computed to determine how light travels through the regions. Finally, the actual structure is analyzed by coupling the modes at the boundaries between regions with boundary conditions.
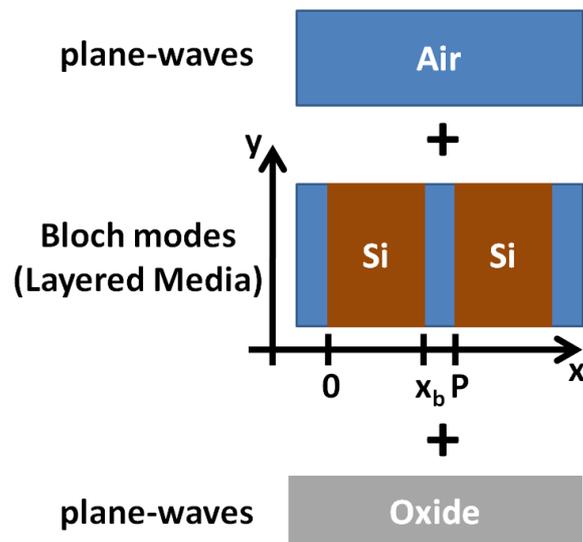

Figure 8.10: The three regions that define the SWG in the layered medium picture: the air above the grating, the oxide below the grating, and layered medium representing the grating. Each region is characterized by the set of waves listed to the left.

The grating is represented by an infinite layered medium, which is turned $90°$

from its typical orientation. Light propagates in a set of Bloch modes which move downwards, parallel to the silicon/air interfaces, rather than perpendicular to them. These modes consist of a standing wave pattern formed from plane-waves traveling at oblique angles through the layers and automatically take into account coupling between neighboring bars. Two plane-waves exist in each medium (air and silicon) – one traveling to the left and one traveling to the right. The basic Bloch mode pattern consists of a periodic piece, called h_z and e_y here, and a horizontal phase progression $(e^{j\kappa x})$ which combine to give the total mode pattern, H_z and E_y. The equation of the basic mode pattern appears in (8.17):

$$
\begin{aligned}
H_z(x, k_y, \lambda_0, \kappa) &= h_z(x, k_y(\lambda_0), \lambda_0)e^{j\kappa x} \\[2mm]
E_y(x, k_y, \lambda_0, \kappa) &= e_y(x, k_y(\lambda_0), \lambda_0)e^{j\kappa x} \\[2mm]
h_z(x, k_y, \lambda_0) &=
\begin{cases}
Ae^{jk_x^{Si}(k_y(\lambda_0),\lambda_0)x} + Be^{-jk_x^{Si}(k_y(\lambda_0),\lambda_0)x}, \\[2mm]
\qquad\qquad 0 \le x < x_b. \\[2mm]
Ce^{jk_x^{air}(k_y(\lambda_0),\lambda_0)x} + De^{-jk_x^{air}(k_y(\lambda_0),\lambda_0)x}, \\[2mm]
\qquad\qquad x_b \le x < P.
\end{cases} \\[2mm]
e_y(x, k_y, \lambda_0) &=
\begin{cases}
\frac{k_x^{Si}(k_y(\lambda_0),\lambda_0)}{\omega\epsilon_{Si}}\left(Ae^{jk_x^{Si}(k_y(\lambda_0),\lambda_0)x} - Be^{-jk_x^{Si}(k_y(\lambda_0),\lambda_0)x}\right), \\[2mm]
\qquad\qquad 0 \le x < x_b. \\[2mm]
\frac{k_x^{air}(k_y(\lambda_0),\lambda_0)}{\omega\epsilon_{air}}\left(Ce^{jk_x^{air}(k_y(\lambda_0),\lambda_0)x} - De^{-jk_x^{air}(k_y(\lambda_0),\lambda_0)x}\right), \\[2mm]
\qquad\qquad x_b \le x < P.
\end{cases}
\end{aligned} \tag{8.17}
$$

$$
k_x^{air}(k_y, \lambda_0) = \sqrt{\left(\frac{2\pi}{\lambda_0}\right)^2 - k_y^2(\lambda_0)} \tag{8.18}
$$

$$k_x^{Si}(k_y, \lambda_0) = \sqrt{\left(\frac{2\pi n_{Si}}{\lambda_0}\right)^2 - k_y^2(\lambda_0)}, \tag{8.19}$$

where λ_0 is the free-space wavelength, k_y is the Bloch mode's propagation constant (wavenumber parallel to the air/silicon interfaces), and $\epsilon_{Si} = n_{Si}^2 \epsilon_0$ and $\epsilon_{air} = n_{air}^2 \epsilon_0 = \epsilon_0$ are the permittivities of silicon and air respectively. There are three wave numbers in the horizontal direction: k_x^{Si}, k_x^{air}, and κ. The first two are local wavenumbers for the plane-waves in each medium that form the Bloch mode. κ is the global horizontal wavenumber that describes the phase progression of the mode in the horizontal direction. In general κ is arbitrary; however, this study considers only the on-axis case so $\kappa = 0$. The mode profiles depends on both the wavelength λ_0 and the propagation constant k_y. The propagation constant itself depends on wavelength and to find the allowed values of k_y requires solving for the dispersion relation. Determining the mode profile also involves finding the amplitudes of the four plane-waves that comprise the Bloch mode, A, B, C, and D.

To find the dispersion relation, boundary conditions were imposed along the vertical air/silicon interfaces. They enforce the continuity of E_x and H_z at $x = x_b$ and P (coordinates shown in figure 8.10).

$$H_z(x_b^-, k_y(\lambda_0), \kappa) = H_z(x_b^+, k_y(\lambda_0), \kappa)$$

$$E_x(x_b^-, k_y(\lambda_0), \kappa) = E_x(x_b^+, k_y(\lambda_0), \kappa)$$

$$H_z(P^-, k_y(\lambda_0), \kappa) = H_z(P^+, k_y(\lambda_0), \kappa) = H_z(0^+, k_y(\lambda_0), \kappa)e^{j\kappa P} \tag{8.20}$$

$$E_x(P^-, k_y(\lambda_0), \kappa) = E_x(P^+, k_y(\lambda_0), \kappa) = E_x(0^+, k_y(\lambda_0), \kappa)e^{j\kappa P}. \tag{8.21}$$

Since $\kappa = 0$ the last two boundary conditions reduce to periodic boundary conditions.

The boundary conditions result in four equations in four unknown wave amplitudes – A, B, C, and D – which were solved as a matrix equation. Finding the non-zero solutions (which are the Bloch modes) required that the determinant be zero, which led to an equation with two parameters: λ_0 and $k_y(\lambda 0)$. This equation was numerically solved (with Matlab's fminsearch method) by scanning k_y at a fixed λ_0 until a zero determinant was found. This yielded the dispersion relation shown in figure 8.11(a). This method occasionally found spurious solutions which were ruled out by inspection.
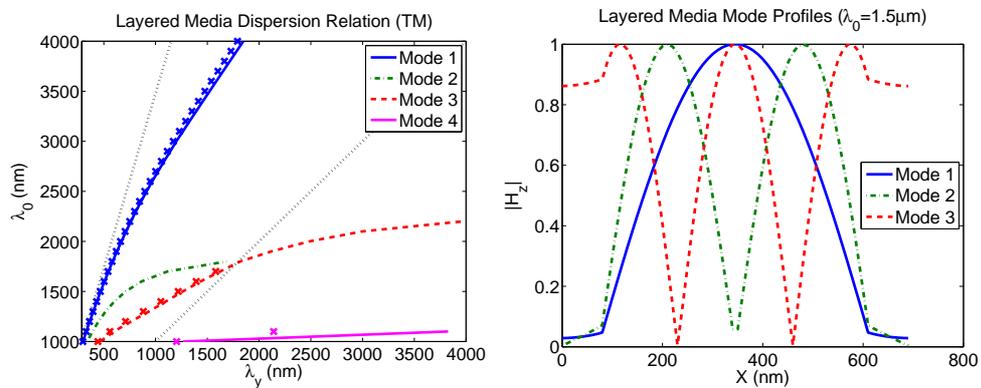
(a) Dispersion relation.   (b) Mode profiles.

Figure 8.11: Dispersion diagram of the layered media model of the SWG and the profiles ($|H_z|$) of the associated modes. In the dispersion diagram the lines denote the theoretical calculations and the x's denote TEMPEST measurements. The dotted lines are the silicon light cone (left) and the air light cone (right). The mode profiles are shown with the silicon centered (running from $x = 80$ nm to 610 nm.

After solving for the dispersion relation, the mode profiles were obtained from the matrix equation. The amplitudes of the plane-waves comprising the mode profiles,

A, B, C, and D, were computed by solving for the eigen-vectors that corresponded to a zero eigen-value (using Matlab's eig function). Figure 8.11(b) shows profiles of the three modes present at $1.5\mu m$. These modes differ in the number of peaks they have. Each modes' peaks are 180° out of phase with their neighbors.

Mode profiles were also computed with TEMPEST (using a method nearly identical to that used to uncover resonance patterns (8.6)). This method produced nearly identical results; however, the profiles computed from TEMPEST contained small phase errors near the silicon/air interfaces that made the modes non-orthogonal (orthogonality is used below). Theory predicts that since the layered medium is loss-less the modes should be completely real. For this reason, the mode profiles computed from the matrix equation were used for the later calculations.

Since this study considered the case of normal incidence, only even Bloch modes were excited in the SWG. The mode profiles in Figure 8.11(b) show that mode two had odd symmetry and therefore was not active. Modes one and three had even symmetry and the dispersion relation in Figure 8.11(a) shows they were both present in the high reflectivity spectrum. For this reason, the discussion below considers only modes one and three.

The theoretical dispersion relation was confirmed by TEMPEST simulations of semi-infinite layered medium (layout shown in Figure 8.12(a)). In these simulations, a single frequency plane-wave in air illuminated the top of the layered medium. The layered medium terminated in PML to remove all reflections.

The Bloch modes can be seen in the near-field patterns from these simulations. At $3.0\mu m$, shown in Figure 8.12(c), each vertical bar contained one bright region, indicating that mode one was excited and no higher order modes were active. At $1.5\mu m$, shown in Figure 8.12(c), the fields were more complicated. Each bar contained three bright spots in the horizontal direction. The central spot was brightest because it was a combination of modes one and three. The two spots to either side came from mode three. The periodicity in the vertical direction comes from modes one and three beating against one another.

The modes' wavelengths were verified by computing the wavelength spectrum from near-fields at each wavelength. This was accomplished by Fourier transforming a vertical cut-line of the fields taken within a silicon bar (figure 8.12(c)):

$$s(\lambda_y, \lambda_0) = \sum_{y=0}^{499,999} \bar{H}_z(y, \lambda_0) w(y) e^{-j\frac{2\pi}{\lambda_y}y}, \tag{8.22}$$

where $\bar{H}_z(y, \lambda_0)$ was the zero-padded H_z data, $w(y)$ was the Hann window (8.5), and $s(\lambda_y)$ was the spatial spectrum. An example of the spatial spectrum at $1.5\mu m$ is shown in Figure 8.12(b). The number of spectral peaks indicated that two modes were present at this wavelength. The location of each peak yielded the modes' propagation wavelengths. The dispersion relation measured in TEMPEST is overlaid with the theoretical dispersion relation in figure 8.11(a). The TEMPEST results are shown by x's and closely match the theoretical dispersion relation.
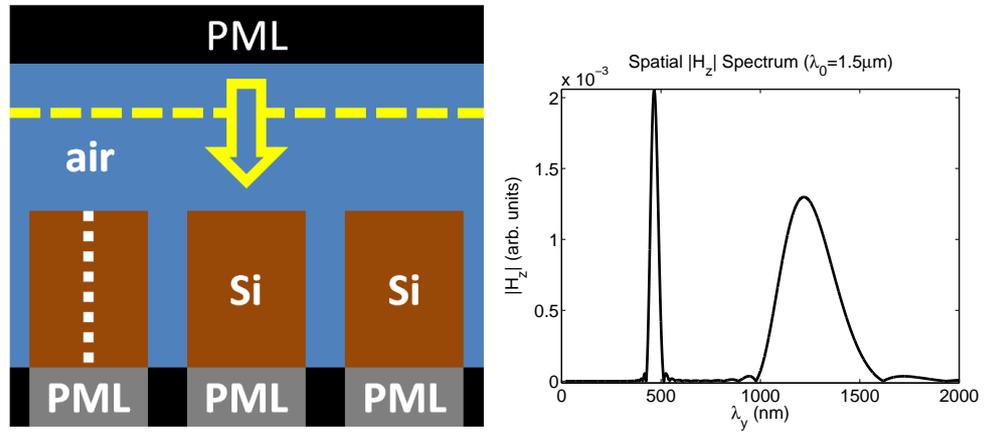
(a) Layout.

(b) Spatial spectrum along the cutline.
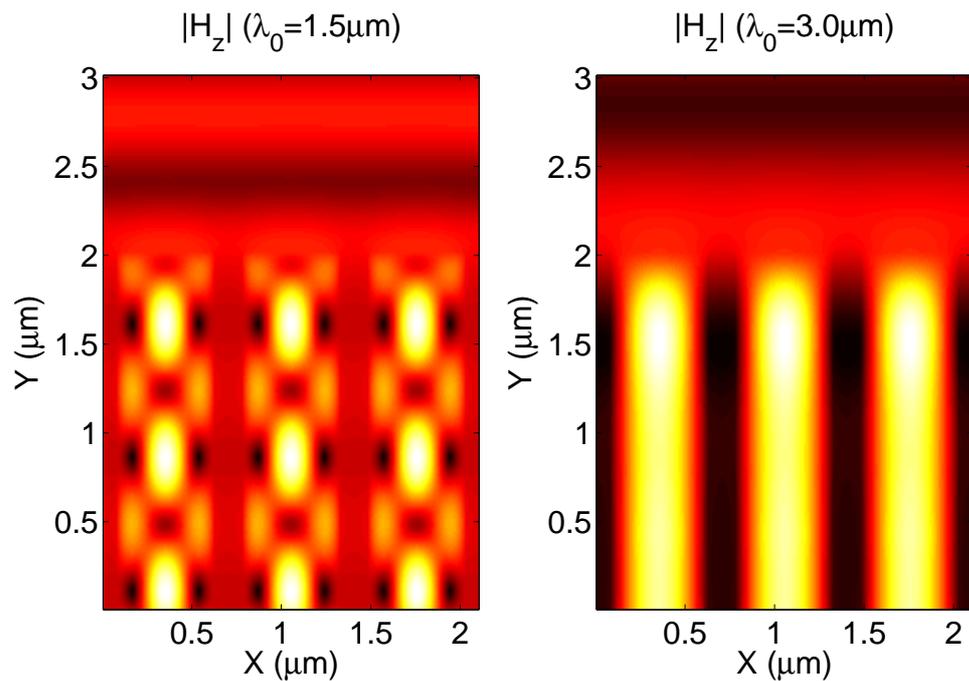
(c) Near-fields at $\lambda_0 = 1.5 \mu m$.

(d) Near-fields at $\lambda_0 = 3.0 \mu m$.

Figure 8.12: The layout of the TEMPEST simulations used to confirm the layered media dispersion relation, an example of the spatial spectrum computed from the near-fields, and two examples of the near-fields. Three periods are shown, but only one was simulated. The vertical dashed line in the layout shows the location of the cutline used to form the spatial spectrum.

8.4 Measuring Bloch Modes in the SWG

This section discusses the role of the Bloch waves in the SWG reflection spectrum. Two waves with the correct symmetry are active in the spectrum of interest: modes one and three from figure 8.11. The broad reflection peak was found to be a complex combination of both waves.

The theory described above shows that there are three Bloch waves active in an infinitely tall layered medium at wavelengths in the high reflection band. However, the actual grating has a finite height and the theory says nothing how the Bloch waves couple to plane-waves in the air above the grating and the oxide below the grating. This coupling process is quite complicated. It can be computed with a mode matching method based on the Bloch waves [71]; however, the approach taken here was to simulate the grating with TEMPEST and measure the amplitudes of the excited Bloch waves during post-processing.

Doing this required the near-field patterns throughout the grating structure under normal incidence and each wavelength. Single frequency TEMPEST simulations with plane-wave sources were used to compute these patterns. These were the same simulations used to compute the point by point reflection spectrum shown in Figure 8.2.

Since the layered medium was lossless its modes were orthogonal. Orthogonality was used to extract the power flow in each mode in the SWG structure during post-processing. However, oppositely direct waves of the same mode were not orthogonal. So two equations, based on the mode amplitude in the E_y and H_z fields were simul-

taneously solved to get the amplitudes of the upward and downward traveling waves (essentially an impedance condition). The process began by considering the fields inside the SWG (at one wavelength) as a super-position of Bloch modes:

$$
\begin{aligned}
E_y(x, \lambda_0) &= \sum_m \left[\left(a_m^\uparrow(\lambda_0) + a_m^\downarrow(\lambda_0) \right) \psi_e^m(x, \lambda_0) \right] \\
H_z(x, \lambda_0) &= \sum_m \left[\left(-a_m^\uparrow(\lambda_0) + a_m^\downarrow(\lambda_0) \right) \psi_h^m(x, \lambda_0) \right],
\end{aligned}
\tag{8.23}
$$

where $E_y(x, \lambda_0)$ and $H_z(x, \lambda_0)$ were the fields in the SWG, m was the mode number, $\psi_e^m(x, \lambda_0)$ was the E_y mode profile, $\psi_h^n(x, \lambda_0)$ was the H_z mode profile, $a_m^\uparrow(\lambda_0)$ was the complex amplitude of the upward traveling wave, and $a_m^\downarrow(\lambda_0)$ was the complex coefficient of the downward traveling wave. The mode profiles were constructed from the downward traveling wave which dictated the signs used in the $H(z)$ equation.

The modes were normalized according to (8.24).

$$
\sum_{x=0}^{P} \psi_e^m(x, \lambda_0) \psi_h^n(x, \lambda_0) dx = \delta_{m,n},
\tag{8.24}
$$

where m and n were the mode numbers and $\delta_{m,n}$ was a Kronecker delta function.

The complex mode amplitudes were extracted by inverting a simple matrix equation based on the super-position of modes (8.23):

$$
\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}
\begin{pmatrix} a_m^\uparrow(\lambda_0) \\ a_m^\downarrow(\lambda_0) \end{pmatrix}
=
\begin{pmatrix} \sum_{x=0}^{P} E_y(x, \lambda_0) \psi_h^m(x, \lambda_0) \\ \sum_{x=0}^{P} H_z(x, \lambda_0) \psi_e^m(x, \lambda_0) \end{pmatrix}.
\tag{8.25}
$$

E_y and H_z were the field patterns computed by TEMPEST. This equation was used for each of the two active modes (modes one and three) at each frequency in the spectrum of interest.

These wave amplitudes indicate which modes were active in the SWG at each wavelength and the direction of power flow. Figure 8.13 shows the normalized power flowing in each mode in the SWG. The normalized power represents the average power density across a period of the structure. The modal power was normalized to the incident power density according to:

$$p_m^\uparrow(\lambda_0) = \frac{1}{2} \frac{|a_m^\uparrow(\lambda_0)|^2}{PS_{inc}}, \tag{8.26}$$

where $p_m^\uparrow(\lambda_0)$ is the normalized power flowing upward in mode number m, $a_m^\uparrow(\lambda_0)$ is the mode amplitude extracted with (8.25), P is the pitch, and S_{inc} is the power density of the incident wave. An identical equation was used for the downward waves.

The circle and x markers show the power in the two downward traveling waves and the non-marked lines show the power in the upward traveling waves. At long wavelengths, only mode one was active and it predominantly carried power downwards, leading to low reflectivity. Around $2.0\mu m$ mode three became active and mode one carried less downward power, increasing reflectivity. The modal power peaked at $1.55\mu m$, where mode three carried more power than mode one. At this peak, the modal power density exceeded the incident power density, indicating a high-Q system. Additionally, the upward and downward waves carried equal amounts of power indicating that all the power entering the top of the SWG was ultimately returned to the top of the grating (more about this below). Around $1.3\mu m$ mode one was again dominant and the reflectivity begain to decrease. Near $1.0\mu m$ the modal powers be-
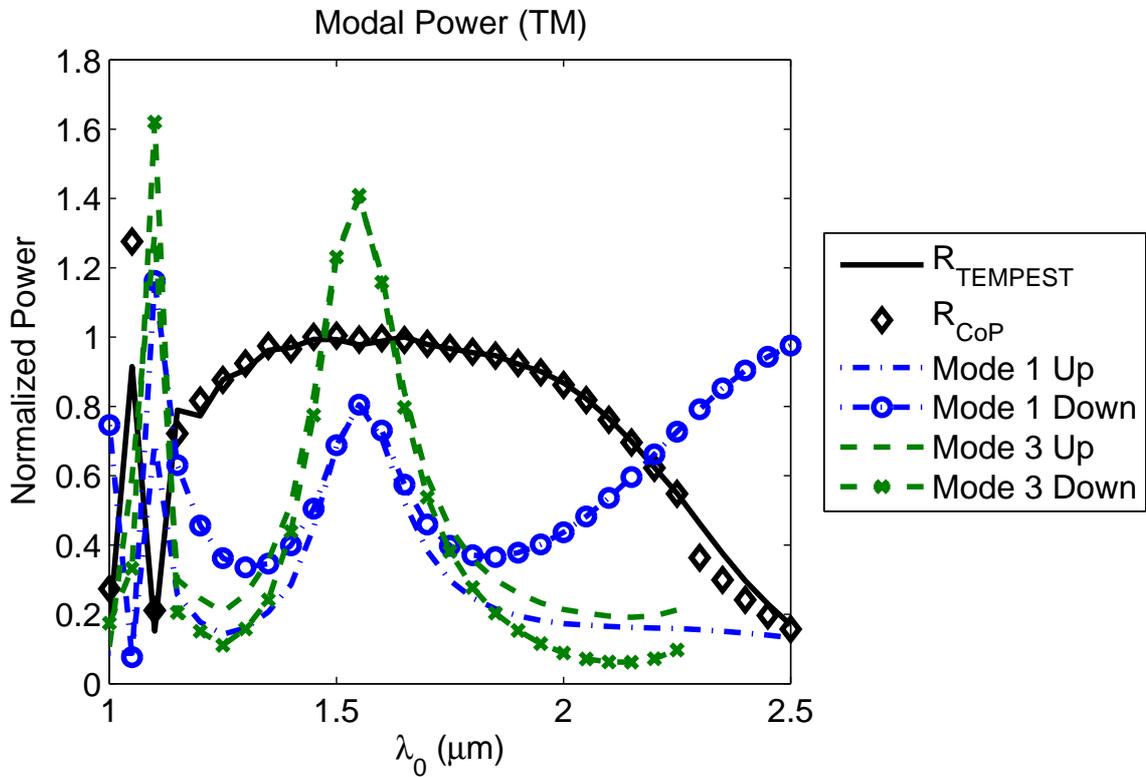
Figure 8.13: Power flowing in each mode in the SWG. CoP stands for conservation of power.

come erratic because mode four was neglected from this analysis due to complications near cutoff.

Conservation of power shows that the modal powers extracted during post-processing account for all of the power flowing vertically through the SWG. This is shown in Figure 8.13, which contains a comparison of the reflectivity measured directly from the reflected plane-wave and the reflectivity computed from the modal powers. The solid line shows the reflectivity measured directly from the reflected plane-wave in the TEMPEST simulations. The diamond markers show the reflectivity computed from

conservation of power applied to the modal powers at the top surface:

$$R_{CoP}(\lambda_0) = 1 + \sum \left(p_m^\uparrow(\lambda_0) - p_m^\downarrow(\lambda_0) \right), \tag{8.27}$$

where CoP stands for conservation of power and the one represents the incident wave. Conservation of power shows that 100% reflectivity requires the upward waves carry the same amout of power as the downward waves. Figure 8.13 shows this condition was satisfied at $1.5\mu m$. The conservation of power measurement is quite accurate, differing from the direct reflectivity measurement by only 1% in the high reflection band. The accuracy of this method indicates that the Bloch mode picture captures the physics behind the SWG. The conservation of power calculations display errors at wavelengths longer than $2.25\mu m$ because the third mode was removed from the calculations due to complications near cut-off. The calculations also display errors at wavelengths shorter than $1.2\mu m$ because the fourth mode was removed due to the same complications from cut-off.

The changes in modal power are not due to changes in coupling between the incident wave and the Bloch modes at the top surface. This coupling was characterized in TEMPEST by measuring the power coupled through the top surface of the semi-infinite layered medium used to validate the Bloch dispersion relation (layout shown in Figure 8.12(a)). Figure 8.14 shows the total power coupled through the top surface normalized to the incident wave (labeled transmission), the total power not coupled (labeled reflection), and the power in modes one and three. The total power coupled into the medium was roughly 80% and is nearly independent of wavelength. Small

glitches in total coupling are visible where modes three and four reach cutoff (roughly $2.4\mu m$ and $1.1\mu m$ respectively). Above $2.4\mu m$ only mode one was active and all the power coupled into this mode. At $2.4\mu m$ mode three turned on and began to carry some power. The power in mode three peaked at $1.5\mu m$ and the power in mode one reached a minimum at the same wavelength.
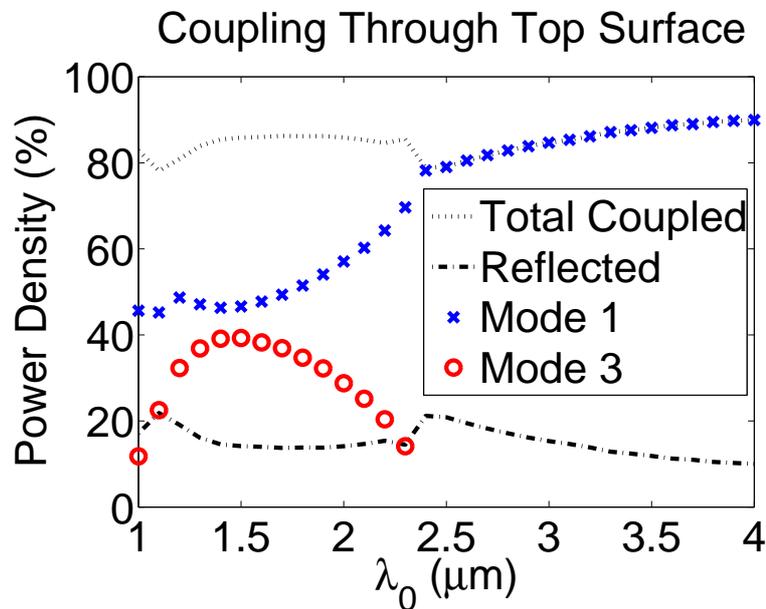

Figure 8.14: Coupling through the top surface of the SWG in terms of power normalized to the incident wave. Total coupling is labeled transmission, power not coupled is labeled reflection, and the x's denote power coupled into modes one and three.

The Bloch modes do not act independently in the SWG, but are coupled to one another at the top and bottom interface. Therefore, it is unclear in Figure 8.14 whether an upward traveling waves are fed by direct reflection or mode conversion. Coupling at the top interface was assessed with the same semi-infinite layered medium TEMPEST simulations (Figure 8.12(a)); however, this time they were excited with

single mode sources placed in the layered medium. These sources launched a single mode and the reflected modes were extracted with the orthogonal mode expansion described above. Separate simulations were run for each mode and each wavelength.

Figure 8.15(a) shows the power radiated into the air and coupled into each downward mode when the top surface was illuminated with mode one traveling upwards. At long wavelengths over 80% of the power radiated into the air. The radiation rapidly decreased when mode three became active, but reflected power does not exceed radiated power. Most of the non-radiated power was coupled back into mode one, with only 10% of the power converted into mode three. Figure 8.15(b) shows
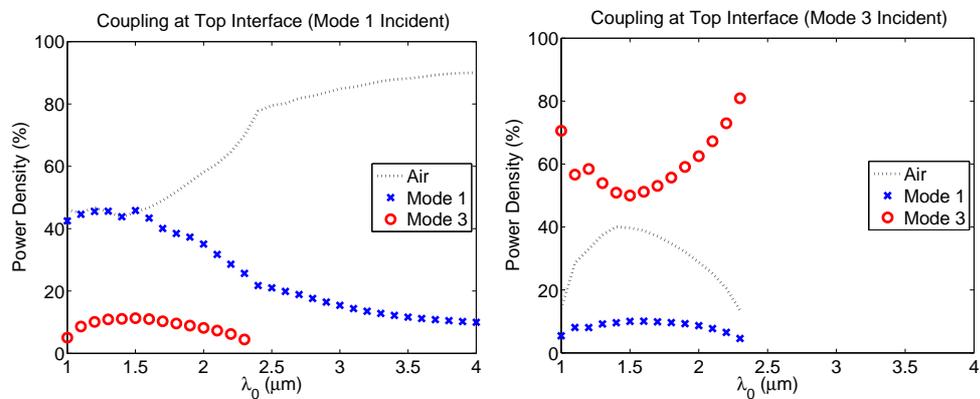
(a) Mode 1 incident.   (b) Mode 3 incident.

Figure 8.15: Coupling between modes at an interface between the layered medium and air, representative of the top surface of the SWG. The simulation layout is shown in Figure 8.12(a).

the same analysis but with mode three incident on the top surface. In this case, radiation was weaker than reflection. Again, mode conversion was weak. Reciprocity dictates that power conversion between modes one and three be symmetric, which

was confirmed by these plots. Additionally, conversion between radiation in air and layered medium modes should be reciprocal. Comparison of Figures 8.14 and 8.15 shows that reciprocity was satisfied.

Coupling at the bottom interface was quantified in a similar manner, but in this case the layered medium terminated in an oxide layer rather than air (Figure 8.16(a)). Figure 8.16(b) shows the coupling when mode one was incident on the bottom interface. Transmission into the oxide was much stronger than transmission into the air, probably because the higher refractive index of the oxide (1.47) more closely matched the silicon's index (3.48). Mode conversion was a few percent stronger as well; however, direct reflection back into mode one was stronger than mode conversion. Figure 8.16(c) shows the same analysis when mode three was incident. Transmission into the oxide was actually weaker than into the air. Reciprocity again kept mode conversion symmetric between modes one and three.

(a) Simulation layout.
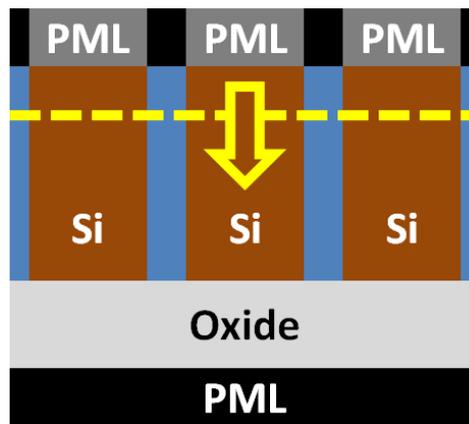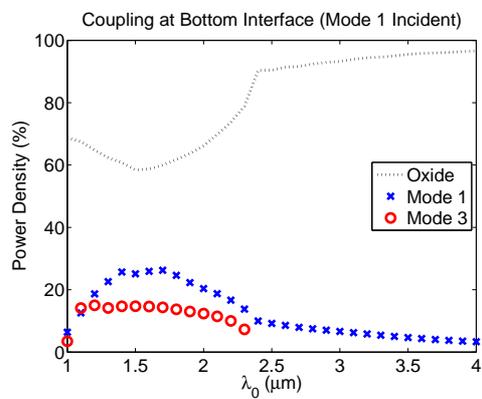

(b) Mode 1 incident.
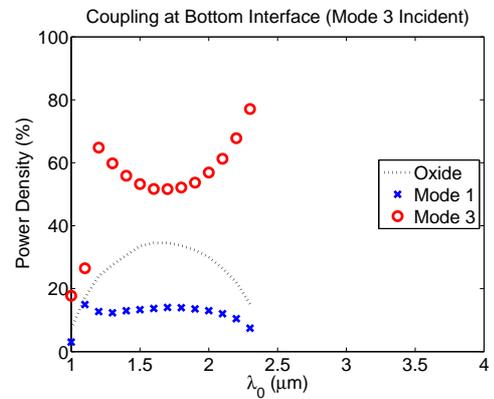
(c) Mode 3 incident.

Figure 8.16: Coupling between modes at an interface between the layered medium and oxide, representative of the bottom surface of the SWG.

8.5 Effects of Thickness

The effects of thickness was investigated by previous researchers [45] and for thicknesses between 300 nm and 700 nm they found two reflection bands, shown in figure 8.17. These bands joined when the grating was 460 nm thick to form the broad band used by the SWG. Altering the thickness changes only the distance each Bloch mode must travel and not the Bloch mode profiles or the coupling processes at the top and bottom interfaces. Therefore, the two reflection bands must arise from an interference pattern between the upward and downward waves.
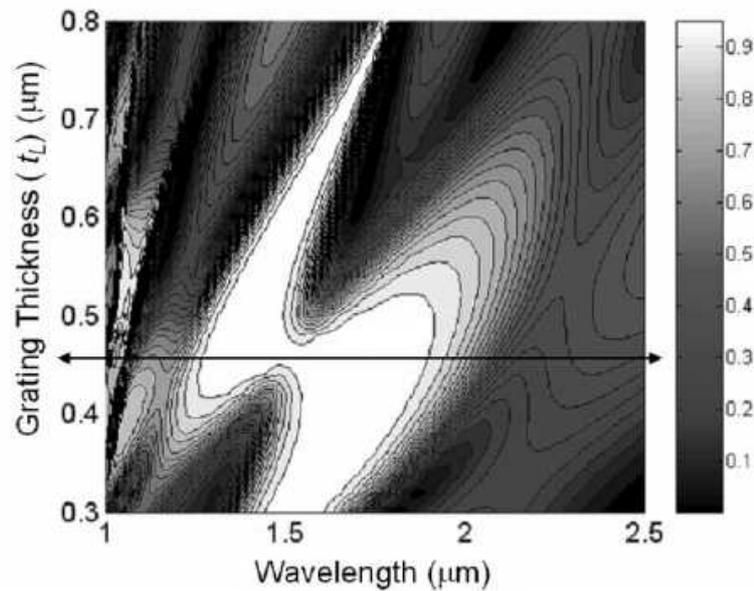

Figure 8.17: Dependence of the SWG's reflectivity on silicon thickness. Reproduced from [45].

Figures 8.18 and 8.19 show the modal powers in 300 nm and 620 nm thick gratings. These figures show that modal powers peak on either side of the high reflection bands.

These peaks correlate the transitions in the reflectivity from low to high or high to low. Comparing the two figures shows the that the peaks shift to longer wavelengths as the SWG is made thicker. Intuitively, this makes sense because thicker structures resonator at longer wavelengths. Additionally, the reflectivity from conservation of power shows that the Bloch model accurately calculates the reflectivity regardless of thickness.

Figure 8.18: Power flowing in each mode in a version of the SWG that is only 300 nm thick. CoP stands for conservation of power.

The two reflection bands probably arise when the structure is effectively a half-wavelength thick. If the SWG were a simple thin-film, reflectivity would be maximized at a wavelength where the film was half a wavelength thick. The SWG picture is complicated by reactive near-fields at the top and bottom surfaces that change the
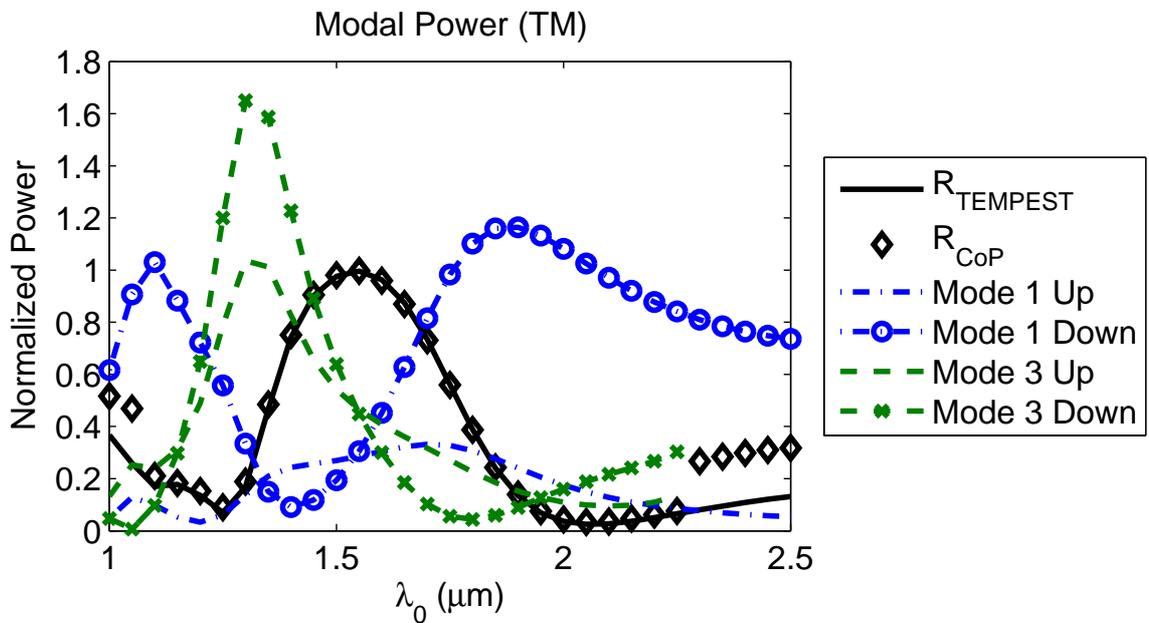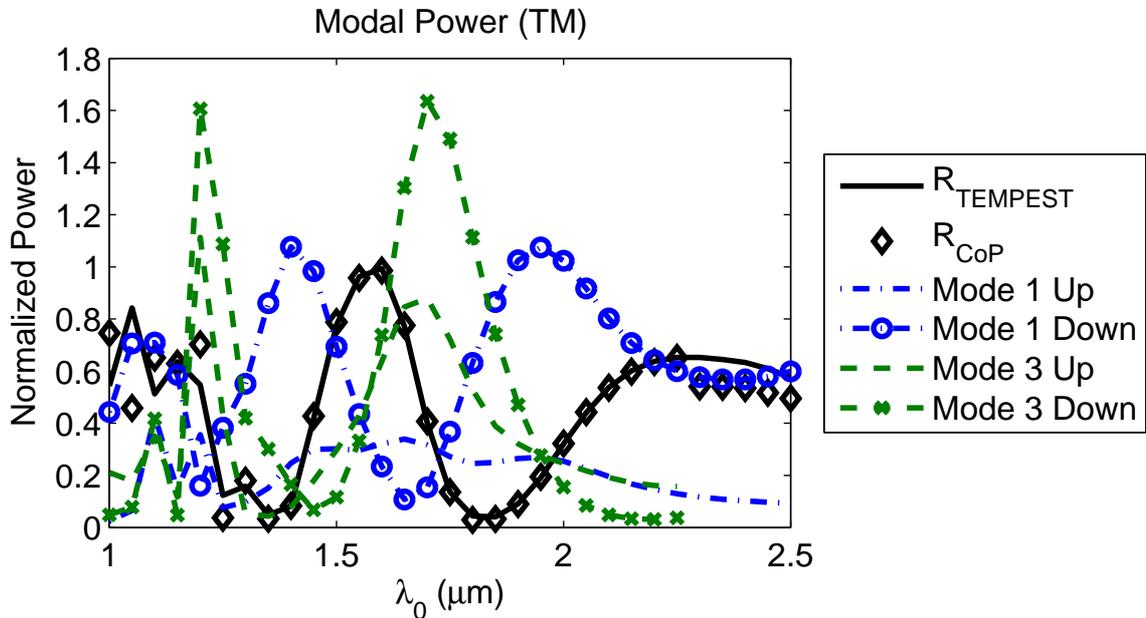
Figure 8.19: Power flowing in each mode in a version of the SWG that is 620 nm thick. CoP stands for conservation of power.

phase length of the structure and mode conversion. Due to conservation of power, 100% reflection occurs when the upward and downward modal powers are equal. This is the signature of effectively half-wavelength cases and can be seen in high reflectivity spectra in Figures 8.13, 8.18, and 8.19.

However, effectively half-wavelength cases are not enough for an extremely wide reflection band. To get the additional width, the wavelengths of the two modes must be an odd integer number of wavelengths apart, causing the modes to be 180° out of phase after propagating from the top surface to the bottom. In this case, even though individually the modes each radiate into the oxide, they cancel each others radiation and result in high reflectivity. This situation happens in the high reflectivity band of

the SWG. Figure 8.20 shows the Bloch mode dispersion relation with additional lines drawn in showing mode one's wavelength scaled by integer multiples. At $\lambda_0 = 1.7\mu m$ mode three's wavelength is three times mode one's wavelength and so mode one incurs an extra half wavelength when moving between the top and bottom interfaces.

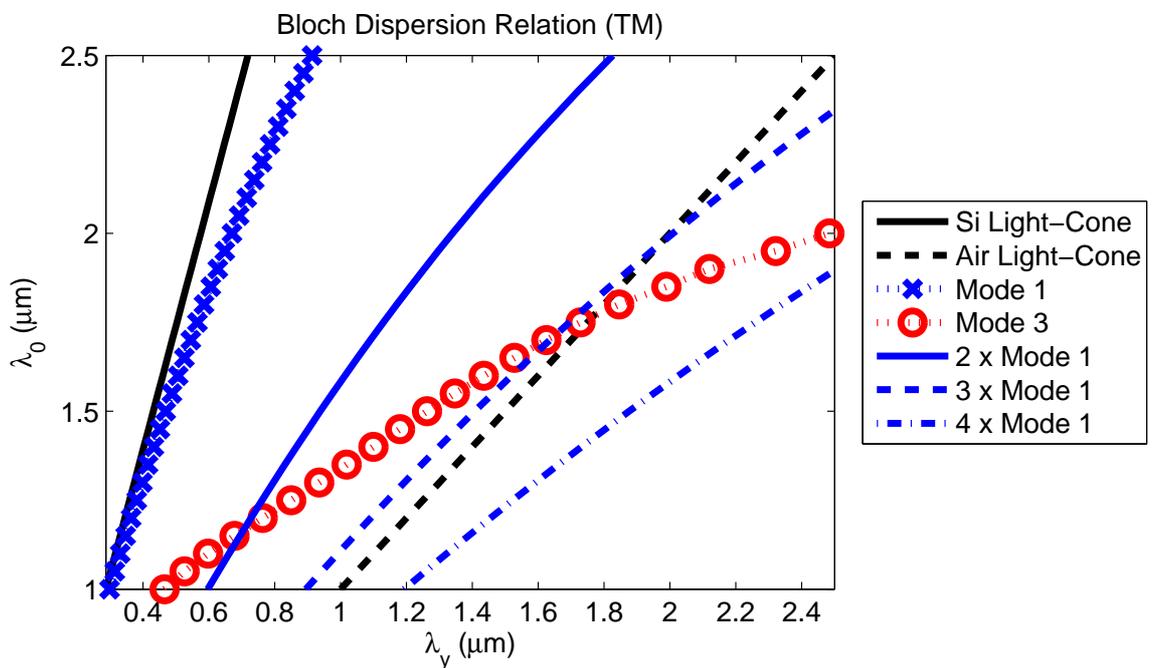Figure 8.20: The Bloch wave dispersion diagram. Extra curves were added to show where mode three's wavelength is an integer multiple of mode one's.

8.6 Conclusion

The reflection process in the SWG is extremely complicated due to mode conversion and reactive near-fields at the top and bottom surfaces. Discovery of the broad reflection band by previous researchers was fortuitous.

TEMPEST v7's pulsed methods enabled mapping the resonances in the SWG and required only a few minutes per simulation. The SWG was shown to contain two resonances in the broad reflection band; however, these resonances did not have the correct symmetry to couple to on-axis radiation. The resonance locations were compared to predictions from two simple models, which were based on horizontally traveling waves. The isolated resonance model accurately predict the spectral locations of some of the resonances. The coupled wave picture accurately predicted the near-field patterns, nearest neighbor coupling, and spectral order of the resonances.

A pictures based on Bloch modes in a layered medium was shown to accurately predict the broad reflection band. This picture uses vertically traveling waves and automatically handles coupling between neighbors. The Bloch mode dispersion relation and near-field patterns were confirmed with TEMPEST. An orthogonal mode expansion was used to measure the modal power in the SWG. Strong modal power was shown to correspond to transitions between low and high reflectivity in gratings of different thicknesses. The broad reflection band in the SWG corresponded to colocation of the power peaks in modes one and three.

Since the SWG is a sub-wavelength structure and radiates only on-axis planewaves, it is likely that another sub-wavelength structure could be designed with a similar broad reflection band. It is suggested that future design efforts look into broadening the high reflectivity band by modifying the topography of the top and bottom surfaces to change the modal reflection and conversion.

Chapter 9

Conclusions

This thesis presented finite difference time domain simulation studies of emerging optical applications. These studies gave physical insight and quantitative engineering data for each of these applications. The breadth of these applications also stimulated the development of a simulation framework. The framework consists of numerical tools designed to enable detailed control over numerical experiments, provide observation of physical phenomena, and enable simulation of large devices. Some numerical tools that were useful to multiple applications were built into the core simulator, TEMPEST version 7. However, each application also required additional specialized tools, which were added to a Matlab framework surrounding TEMPEST.

Even with fully rigorous simulation, developing physical understanding can be very difficult. This is because devices can be so complex that intuitive models do not exist as a starting point and because a multitude of scattered waves can obscure the

processes dominating device operation. This physical phenomena visibility problem was addressed with auxiliary tools.

Even though TEMPEST formed the core of the simulation framework many of the studies benefited from linking other classical analysis methods with the FDTD method. In some cases, the analysis methods added physical insight, such as found with the surface plasmon signal flow graphs and the Bloch mode expansions. In other cases, these analysis methods enabled simulation of electromagnetically large structures, such as the combination of TEMPEST edge models and the full system simulator created by Ball Aerospace for analyzing diffraction masks in the TPF-C.

These studies also highlighted short comings in the FDTD method. The primary shortcoming was memory, which limited the maximum topography size. However, surface plasmon and rapid thermal annealing simulations were limited by computer speed. The long run-times were caused by the fine spatial meshes required to resolve the skin depth and extremely sub-wavelength gates. The fine mesh in turn led to small time steps, required for stability.

This thesis presents a broad range of research, from basic scientific understanding of novel devices to guiding engineering decisions. The core FDTD methods were useful across the entire range and the flexibility of the framework enabled rapid development of auxiliary tools to support each specific study. The science oriented studies, especially the nano-particle work, used TEMPEST (without much of the framework) to observe near-field patterns that were not available in experimental setups. The

engineering studies, e.g. the TPF-C and surface plasmon work, heavily used the framework to quantify specific scattering processes and assess design trade-offs.

In the TPF-C studies, a combination of TEMPEST, analytical back-propagation, and Fourier propagators enabled analysis of the performance of large diffraction masks inside a telescope. Detailed analysis of edge scattering determined that scattering from vertical mask sidewalls was the largest source of straylight, contributing over 2λ per edge. Undercutting by $20°$ reduced this by an order of magnitude. Plasmons on thick metal masks were not an issue due to propagation losses.

TEMPEST simulations of optical coupling in LSA required extremely fine meshes to resolve small gate topographies. These fine meshes necessitated double precision field matrices and long run-times. Three-dimensional simulations of 2D layouts is possible, but will require 1-2 days per simulation with current desktop computer technology. This study confirmed that the LSA system maintained nearly 100% coupling and low variability due to topography even when unidirectional dense metal gates were present on the substrate.

Scattered field sources and PML matched to plasmonic interfaces enabled studies of surface plasmon coupling elements. These elements showed maximum coupling cross-sections slightly exceeding their physical width and strong directivity. A combination of S parameter characterization with TEMPEST and signal flow graphs enabled rapid investigation into pitch and number of element dependencies in finite length gratings. Maximizing plasmon output required a trade-off between elemen-

tal coupling efficiency and transmission. A 100% coupling length of over 10λ was demonstrated.

Multi-day run-times of 3D simulations of nano-octahedra were mitigated by using TEMPEST v7's pulsed methods to compute an entire reflection spectrum per simulation. These spectra were screened for interesting features and single frequency simulations were used to spot check the near-fields. This study led to the discovery that plasmons in the interstitial gaps drive resonant absorption and transmission, instead of reflection.

Pulsed methods enabled mapping the resonance spectrum of the SWG. These maps showed that the resonances have the wrong symmetry to affect the reflection spectrum. Additionally, the coupled wave and CROW models of isolated bars accurately predicted some of the SWG's resonances. The Bloch wave picture of two concurrent vertically traveling waves whose transmissions cancel gave the most physical insight. Since sub-wavelength gratings produce only one plane-wave it is likely highly reflective SWG structures could be produced in the future at other wavelengths.

Bibliography

[1] K. Adam and A. Neureuther. Simplified models for edge transitions in rigorous mask modeling. In *Proc. SPIE*, volume 4346, 2001.

[2] I. Ahsan, N. Zamdmer, O. Glushchenkov, R. Logan, E. J. Nowak, H. Kimura, J. Zimmerman, G. Berg, J. Herman, E. Maciejewski, A. Chan, A. Azuma, S. Deshpande, B. Dirahoui, G. Freeman, A. Gabor, M. Gribelyuk, S. Huang, M. Kumar, K. MIyamoto, D. Mocuta, A. Mahorowala, E. Leobandung, H. Utomo, and B. Walsh. Rta-driven intra-die variations in stage delay, and parametric sensitivities for 65nm technology. In *IEEE Symposium on VLSI Technology Digest of Technical Papers*, 2006.

[3] K. Balasubramanian, P. Echternach, M. Dickie, R. Muller, V. White, D. Hoppe, S. Shaklan, R. Belikov, N. J. Kasdin, R. Vanderbei, D. P. Ceperley, and A. Neureuther. Fabrication and characteristics of free-standing shaped pupil masks for tpf-coronagraph. In *Space Telescopes and Instrumentation I: Optical, Infrared, and Millimeter*, volume 6265. SPIE, June 2006.

[4] S. Boutami, B. Bakir, J.-L. Leclercq, X. Letartre, P. Rojo-Romeo, M. Garrigues, and P. Viktorovitch. Highly selective and compact tunable moems photonic crystal fabry-perot filter. *Optics Express*, 14(8):3129–3137, April 2006.

[5] Carlson. Harvard medical school. private communications, 2004.

[6] D. Ceperley. Vector scattering analysis of tpf coronagraph pupil masks. Master's thesis, U. C. Berkeley, May 2005.

[7] D. Ceperley, A. Neureuther, M. Lieber, and J. Kasdin. Characterizing edge-generated stray light sources for tpf coronagraph pupil masks. In *Techniques and Instrumentation for Detection of Exoplanets II*, volume 5905, pages 162–172. SPIE, Aug. 2005.

[8] D. Ceperley, A. Neureuther, M. Lieber, J. Kasdin, and T. Shih. Vector scattering analysis of tpf coronagraph pupil masks. In *Stray Light in Optical Systems: Analysis, Measurement, and Suppression*, volume 5526B. SPIE, Aug. 2004.

[9] C. Chang. Etching submicrometer trenches by using the bosch process and its application to the fabrication of antireflection structures. *Journal of Micromechanics and Microengineering*, 15:580–585, 2005.

[10] U.C. Berkeley Clustered Computing. The millennium cluster. http://www.millennium.berkeley.edu, May 2006.

[11] Y. Deng. *Modeling Innovations in EUV and Nanoimprint Lithography.* PhD thesis, University of California, Berkeley, 2005.

[12] J. A. Dionne, L. A. Sweatlock, H. A. Atwater, and A. Polman. Planar metal plasmon waveguides: frequency-dependent dispersion, propagation, localization, and loss beyond the free electron model. *Physical Review B*, 72:075405, 2005.

[13] N. Economou. Surface plasmons in thin films. *Physical Review*, 182(2):539, 1969.

[14] S. Fan and J. Joannopoulos. Analysis of guided resonances in photonic crystal slabs. *Physical Review B*, 65:235112, 2002.

[15] V. Ferry. California institute of technology. private communications, 2008.

[16] A. Goncharov, M. Owner-Petersen, and D. Puryayev. Apodization effect in a compact two-mirror system with a spherical primary mirror. *Opt. Eng.*, 41(12):3111, 2002.

[17] S. Gray and T. Kupka. Propagation of light in metallic nanowire arrays: Finite-difference time-domain studies of silver cylinders. *Physical Review B*, 68:045415, 2003.

[18] O. Guyon. Phase-induced amplitude apodization of telescope pupils for extrasolar planet imaging. *Astronomy and Astrophysics*, 404:379–387, 2003.

[19] R. F. Harrington. *Time-Harmonic Electromagnetic Fields.* Wiley-IEEE Press, 2 edition, 2001.

[20] R. Holland. Threde: A free-field emp coupling and scattering code. *IEEE Trans on Nuclear Science*, 24(6):2416–2421, Dec. 1977.

[21] R. Holland, L. Simpson, and K. Kunz. Finite-difference analysis of emp coupling to lossy dielectric structures. *IEEE Trans on Electromagnetic Compatibility*, 22(3):203–209, Aug. 1980.

[22] J. Homola, S. S. Yee, and G. Gauglitz. Surface plasmon resonance sensors: review. *Sensors and Actuators B*, 54:3–15, 1999.

[23] D. Hoppe and T. Cwik. Full-wave effects in coronagraph masks. *TPF Science, Technology, and Design Expo*, Oct. 2002. http://planetquest1.jpl.nasa.gov/tpf/expo03/dsp_tpfAgendaDetail.cfm?id=a186.

[24] M. Huang, Y. Zhou, and C. Chang-Hasnain. A surface-emitting laser incorporating a high-index-contrast subwavelength grating. *Nature Photonics*, 1:119–122, Feb. 2007.

[25] M. Huang, Y. Zhou, and C. Chang-Hasnain. A nanoelectromechanical tunable laser. *Nature Photonics*, 2:180–184, March 2008.

[26] R. A. Hurd. The edge condition in electromagnetics. *IEEE Trans. Antennas and Propagation*, pages 70–73, Jan. 1976.

[27] J. Kasdin, R. Belikov, J. Beall, R. Vanderbei, M. Littman, M. Carr, and A. Give'on. Shaped pupil coronagraphs for planet finding: optimization, man-

ufacturing, and experimental results. In *Techniques and Instrumentation for Detection of Exoplanets II*, volume 5905, pages 128–136. SPIE, Aug. 2005.

[28] J. Kasdin, R. Vanderbei, D. Spergel, and M. Littman. Extrasolar planet finding via optimal apodized-pupil and shaped-pupil coronagraphs. *The Astrophysical Journal*, 582:1147–1161, 2003.

[29] N. J. Kasdin, R. Vanderbei, J. Beall, R. Belikov, M. Littman, and M. Carr. Shaped pupil coronagraphs for planet finding: optimization, manufacturing, and experimental results. In *Optical Engineering and Instrumentation: Techniques and Instrumentation for Detection of Exoplanets II*, volume 5905, Aug. 2005.

[30] N. J. Kasdin, R. Vanderbei, D. Spergel, and M. Littman. Optimal shaped pupils for extrasolar planet detection. In *High Contrast Imaging for Exo-planet Detection*, volume 4860. SPIE, Aug. 2002.

[31] J. Kim, L. Chrostowski, E. Bisaillon, and D. Plant. Dbr, sub-wavelength grating, and photonic crystal slab fabry-perot cavity design using phase analysis by fdtd. *Optics Express*, 15(16):10330–10339, Aug. 2007.

[32] C. Kittel. *Introduction to Solid State Physics*, page 162. John Wiley & Sons, Inc., 6 edition, 1986.

[33] M. Kuchner and W. Traub. A coronagraph with a band-limited mask for finding terrestrial planets. *Astrophysical Journal*, 570:900, 2002.

[34] J. Leitner. Formation flying system design for a planet-finding telescope-occulter system. In *UV/Optical/IR Space Telescopes: Innovative Technologies and Concepts III*, volume 6687. SPIE, 2007.

[35] G. Leveque and O. J. F. Martin. Numerical study and optimization of a diffraction grating for surface plasmon excitation. In *Proc. SPIE*, volume 5927, 2005.

[36] M. Lieber, S. Kilston, J Kasdin, R. Vanderbei, and M. Littman. Evolving exosolar planet detection methods with lab experiments and integrated modeling: I. modeling. In *Techniques and Instrumentation for Detection of Exoplanets*, volume 5170. SPIE, Aug. 2003.

[37] M. Lieber, A. Neureuther, D. Ceperley, J. Kasdin, D. Hoppe, and A. Eisenman. Evaluating the end-to-end performance of tpf-c with vector propagation models: Part i. pupil mask effects. In *Techniques and Instrumentation for Detection of Exoplanets II*, volume 5905, pages 173–184. SPIE, Aug. 2005.

[38] M. Lieber, A. Neureuther, D. Ceperley, J. Kasdin, and N. Ter-Gabrielyan. Vector wavefront propagation modeling for the tpf coronagraph. In *SPIE Astron. Telescopes*, June 2004.

[39] M. Lieber, A. Neureuther, D. Ceperley, N. J. Kasdin, D. Hoppe, and A. Eisenman. Evaluating the end-to-end performance of tpf-c with vector propagation models and wavefront control. In *Optical Engineering and Instrumentation:*

Techniques and Instrumentation for Detection of Exoplanets II, volume 5905. SPIE, Aug. 2005.

[40] M. Lieber, C. Noecker, and S. Kilston. Integrated system modeling for evaluating the coronagraph approach to planet detection. In *High Contrast Imaging for Exo-planet Detection*, volume 4860. SPIE, Aug. 2002.

[41] J. Lu, C. Petre, E. Yablonovitch, and J. Conway. Numerical optimization of a grating coupler for the efficient excitation of surface plasmons at an ag-sio$_2$ interface. *Journal of the Optical Society of America B*, 24(9):2268–2272, Sept. 2007.

[42] R. Luebbers, F. Hunsberger, K. S. Kunz, R. B. Standler, and M. Schneider. A frequency-dependent finite-difference time-domain formulation for dispersive materials. *IEEE Trans. Electromagnetic Comp.*, 32(3):222, 1990.

[43] S. A. Maier and H. A. Atwater. Plasmonics: Localization and guiding of electromagnetic energy in metal/dielectric structures. *Journal of Applied Physics*, 98:011101, 2005.

[44] C. Mateus, M. Huang, L. Chen, and C. Chang-Hasnain. Broad-band mirror (1.12 - 1.62μm) using a subwavelength grating. *IEEE Photonics Technology Letters*, 16(7):1676–1678, July 2004.

[45] C. Mateus, M. Huang, Y. Deng, A. Neureuther, and C. Chang-Hasnain. Ultra-

broadband mirror using low-index cladded subwavelength grating. *IEEE Photonics Technology Letters*, 16(2):518–520, Feb. 2004.

[46] M. Miller, A. Neureuther, D. P. Ceperley, J. Rubinstein, and K. Kikuchi. Characterization and monitoring of photomask edge effects. In *Proc. SPIE*, volume 6730, 2007.

[47] M. G. Moharam and T. K. Gaylord. Rigorous coupled-wave analysis of metallic surface-relief gratings. *Journal of the Optical Society of America A*, 3(11):1780–1787, Nov. 1986.

[48] B. A. Munk. *Frequency Selective Surfaces: Theory and Design*. Wiley-Interscience, 2000.

[49] A. Neureuther, M. Lieber, S. Kilston, and Y. Deng. Vector wavefront simulation for tpf coronagraph performance evaluation. Proposal response to JPL RFP for Development Technologies for the Terrestrial Planet Finder Mission, June 2003.

[50] A. R. Neureuther and D. P. Ceperley. Modeling and simulation for nanometrics. *J. Vac. Sci. Tech. B*, 23(6):2578, 2005.

[51] C. Noecker, B. Woodruff, and C. Burrows. Optical system engineering for a terrestrial planet finder coronagraph. In *High-Contrast Imaging for Exo-planet Detection*, volume 4860. SPIE, Aug. 2002.

[52] T. Parr. Another tool for language recognition v2. Technical report, University of San Francisco, 2008. http://www.antlr.org/.

[53] T. Pistor. *Electromagnetic Simulation and Modeling with Applications in Lithography.* PhD thesis, University of California, Berkeley, 2001.

[54] H. Raether. Surface plasma oscillations and their applications. In G. Hass, M. Francombe, and R. Hoffman, editors, *Physics of Thin Films*, volume 9, pages 145–261. Academic Press, 1977.

[55] J. A. Roden and S. D. Gedney. Convolution pml (cpml): An efficient fdtd implementation of the cfs-pml for arbitrary media. *Microwave and Optical Technology Letters*, 27(5):334–339, 2000.

[56] L. Sherry, R. Jin, C. Mirkin, G. Schatz, and R. Van Duyne. Localized surface plasmon resonance spectroscopy of single silver triangular nanoprisms. *Nano Letters*, 6(9):2060–2065, 2006.

[57] A. Shima, Y. Wang, D. Upadhyaya, L. Feng, S. Talwar, and A. Hiraiwa. Dopant profile engineering of cmos devices formed by non-melt laser spike annealing. In *2005 Symposium on VLSI Technology*, pages 144–145. IEEE, June 2005.

[58] B. W. Smith. Optical properties of thin films for duv and vuv microlithography. Technical report, Rochester Institute of Technology, 2008.

http://www.rit.edu/kgcoe/microsystems/lithography/thinfilms/thinfilms/thinfilms.html.

[59] A. Sundaramurthy, K. Crozier, G. Kino, D. Fromm, P. Schuck, and W. Moerner. Field enhancement and gap-dependent resonance in a system of two opposing tip-to-tip au nanotriangles. *Physical Review B*, 72:165409, 2005.

[60] A. Taflove and S. Hagness. *Computational Electrodynamics: The Finite Difference Time Domain Method*. Artech House, 2 edition, 2000.

[61] A. Taflove and S. Hagness. *Computational Electrodynamics: The Finite Difference Time Domain Method*. Artech House, 3 edition, 2005.

[62] A. Taflove and S. Hagness. *Computational Electrodynamics: The Finite Difference Time Domain Method*, pages 774–790. Artech House, 3 edition, 2005.

[63] A. Tao, P. Sinsermsuksakul, and P. Yang. Tunable plasmonic superlattices of silver nanocrystals. *Nature Nanotechnology*, 2:435, 2007.

[64] J. Tirapu-Azpiroz, P. Burchard, and E. Yablonovitch. Boundary layer model to account for thick mask effects in photolithography. In *Proc. SPIE*, volume 5040, 2003.

[65] J. Trauger, C. Burrows, B. Gordon, J. Green, A. Lowman, D. Moody, A. Niessner, F. Shi, and D. Wilson. Coronagraph contrast demonstration with the high-

contrast imaging testbed. In *Optical, Infrared, and Millimeter Space Telescopes*, volume 5487. SPIE, Oct. 2004.

[66] R. Vanderbei, N. J. Kasdin, and D. Spergel. Checkerboard-mask coronagraphs for high-contrast imaging. *Astrophysical Journal*, 615(1):555–561, 2004.

[67] R. Vanderbei, D. Spergel, and N. J. Kasdin. Spiderweb masks for high contrast imaging. *Astrophysical Journal*, 590:593–603, 2003.

[68] S. Vedantam, H. Lee, J. Tang, J. Conway, M. Staffaroni, J. Lu, and E. Yablonovitch. Nanoscale fabrication of a plasmonic dimple lens for nano-focusing of light. In *Proc. SPIE*, volume 6641, Sept. 2007.

[69] H. Wang, D. Brandl, P. Nordlander, and N. Halas. Plasmonic nanostructures: Artificial molecules. *Acc. Chem. Res.*, 40(1):53–62, 2006.

[70] Terrestrial Planet Finder Website. http://planetquest.jpl.nasa.gov/tpf/ tpf_index.html, Aug. 2005.

[71] A. Wexler. Solution of waveguide discontinuities. *IEEE Trans. Microwave Theory and Techniques*, MTT-15(9):508–517, Sept. 1967.

[72] A. Wong. *Rigorous Three-Dimensional Time-Domain Finite-Difference Electro-magnetic Simulation*. PhD thesis, University of California, Berkeley, 1994.

[73] A. K. Wong and A. R. Neureuther. Rigorous three-dimensional time-domain

finite-difference electromagnetic simulation for photolithographic applications. *IEEE Trans. Electron Devices*, 41(6):419–431, 1995.

[74] R. W. Wood. On a remarkable case of uneven distribution of light in a diffraction grating spectrum. *Proceedings of the Physical Society of London*, 18:269–275, 1902.

[75] A. Yariv, Y. Xu, R. Lee, and Z. Scherer. Coupled-resonator optical waveguide: a proposal and analysis. *Optics Letters*, 24(11):711–713, June 1999.

[76] K. S. Yee. Numerical solution of initial boundary-value problems involving maxwell's equations in isotropic media. *IEEE Trans. Ant. Prop.*, AP-14:302–307, May 1966.