Plasmon Enhanced Photoemission

Aleksandr Polyakov

Advanced Light Source Division

Fall 2012

Ph.D. Thesis

Plasmon Enhanced Photoemission

Aleksandr Polyakov
Ph.D. Thesis

Advanced Light Source Division
Lawrence Berkeley National Laboratory
University of California
Berkeley, California 94720

Fall 2012

This work was supported by the Director, Office of Science, Office of Basic Energy Sciences,
Materials Sciences and Engineering Division of the U.S. Department of Energy
under Contract no. DE-AC02-05CH11231.

Plasmon Enhanced Photoemission

by

Aleksandr Polyakov

Plasmon Enhanced Photoemission

by

Aleksandr N Polyakov

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

Doctor of Philosophy

in

Applied Science and Technology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Roger W Falcone, Chair
Dr. Howard A Padmore
Professor David Attwood
Professor Michael Crommie

Spring 2012

Plasmon Enhanced Photoemission

Aleksandr N Polyakov

Abstract

Plasmon Enhanced Photoemission

by

Aleksandr N Polyakov

Doctor of Philosophy in Applied Science and Technology

University of California, Berkeley

Professor Roger W Falcone, Chair

Next generation ultrabright light sources will operate at megahertz repetition rates with temporal resolution in the attosecond regime. For an X-Ray Free Electron Laser (FEL) to operate at such repetition rate requires a high quantum efficiency (QE) cathode to produce electron bunches of 300 pC per 1.5 μJ incident laser pulse. Semiconductor photocathodes have sufficient QE in the ultraviolet (UV) and the visible spectrum, however, they produce picosecond electron pulses due to the electron-phonon scattering. On the other hand, metals have two orders of magnitude less QE, but can produce femtosecond pulses, that are required to form the optimum electron distribution for high efficiency FEL operation. In this work, a novel metallic photocathode design is presented, where a set of nano-cavities is introduced on the metal surface to increase its QE to meet the FEL requirements, while maintaining the fast time response.

Photoemission can be broken up into three steps: (1) photon absorption, (2) electron transport to the surface, and (3) crossing the metal-vacuum barrier. The first two steps can be improved by making the metal completely absorbing and by localizing the fields closer to the metal surface, thereby reducing the electron travel distance. Both of these effects can be achieved by coupling the incident light to an electron density wave on the metal surface, represented by a quasi-particle, the Surface Plasmon Polariton (SPP).

The photoemission then becomes a process where the photon energy is transferred to an SPP and then to an electron. The dispersion relation for the SPP defines the region of energies where such process can occur. For example, for gold, the maximum SPP energy is 2.4 eV, however, the work function is 5.6 eV, therefore, only a fourth order photoemission process is possible. In such process, four photons excite four plasmons that together excite only one electron. The yield of such non-linear process depends strongly on the light intensity.

In this work, the structure consisted of rectangular nano-grooves (NGs) arranged in a subwavelength grating on a metal surface is presented that provides a dramatic increase in the metal's absorption, field localization, and field enhancement. When light is polarized perpendicular to the orientation of the grooves a standing SPP wave is excited along the vertical walls in the NGs, that act as Fabry-Perot resonators. By adjusting the geometry of

the NGs and the period of the subwavelength grating the resonance can be fine tuned to a desired position, for example, the laser fundamental wavelength, anywhere from the UV to the near infrared (NIR).

Two types of gratings are presented: (a) a gold grating with period of 600 nm, and (b) an aluminum-gold grating with a period of 100 nm; both with resonance at 720 nm. In each case, strong on-resonance absorption was observed, with over 98% for grating (b). Unlike the grating-coupled SPP waves, where the angle is well defined by the momentum matching condition, the resonant NGs allow coupling to the standing modes at a range of angles of incidence, referred to as the angular bandwidth. A new model for the on-resonance absorption based on the ensamble action of the NGs is presented that serves as the basis for a design of an NG grating with an ultrawide spectral as well as angular bandwidth. For sample (b), the angular bandwidth is 80 degrees, corresponding to an opening angle of 160 degrees.

The photoemission enhancement for such a grating was measured to be seven orders of magnitude for a four-photon photoemission. This is an incredible result demonstrating the power of the plasmonic grating presented, which is an efficient light trapper and field enhancer for a non-linear processes. These results demonstrate that the metal photocathode prepared with a NG grating on the metal surface will provide sufficient pulse charge driven by a 1 μJ $15 fs$ pulsed laser at 800 nm for the optimum FEL operation.

To the two most intelligent people I have ever met,
Irina Polyakova and Nikolay Polyakov,
my parents.

Contents

List of Figures

List of Abbreviations

ALD Atomic Layer Deposition 73

BS Beam Splitter 101

DI Deionized 59

EM Electro-Magnetic 8

FDTD Finite Difference Time Domain 45

FEL Free Electron Laser 1

FWHM Full Width Half Maximum 48

HF Hydrofluoric Acid 56

HMDS Hexamethyldisilazane 55

HSQ Hydrogen Silsesquioxane 64

IPA Isopropyl Alcohol 56

KOH Potassium Hydroxide 53

LBNL Lawrence Berkeley National Laboratory 53

MIBK Methyl Isobutyl Ketone 64

MIM Metal Insulator Metal 40

NG Nano Groove 39

NIR Near Infrared 2

PEEM Photoemission Electron Microscope 77

PMMA Poly(Methyl Methacrylate) 53

PNNL Pacific Northwest National Laboratory 77

Acknowledgments

Whether it was a chance or fate, I am grateful for the turbulent beginning of my PhD career that has inevitably led me to professor Roger Falcone, who agreed to take me on as a student. I want to thank Roger for being a great advisor to me, for supporting me throughout my graduate study and for introducing me to an amazing scientist Howard Padmore. Howard has been most supportive throughout my career working in his group, and it is a formidable challenge to summarize on a single page the amount of work and time he has put into creating the most nutritious environment for a scientist to grow. I can only say that I am very much indebted to Howard for his guidance throughout the past five years, and I can only hope I did not disappoint.

This dissertation may carry my name alone in the author list, however, many people have contributed countless hours to ensure the success of this project. First, I want to thank people from the Experimental Systems Group, who have helped me navigate the complex world of the ALS: Richard Celestre (for his wise advice in all things technical and beyond), Kyle Engelhorn (for all his help around the lab), Jun Feng (for his inspiring commitment to science), Mike Greaves (for his contagiously positive attitude as well as a massive amount of help around the lab), Kevin Thompson (for his dedication to the project), Theo Vecchione (for his mad dedication to doing good science), Dmitriy Voronov (for introducing me to nano fabrication techniques), Weishi Wan (for helping me make sense of the theory), and Max Zolotorev (for teaching me Maxwell's equations). And second, I want to thank the amazing team from the Molecular Foundry: Stefano Cabrini (who was once described as a white knight, which he truly is), Scott Dhuey (for all the help with the electron beam lithography), Bruce Harteneck (for teaching me how to use the wet bench), Xiaogan Liang (for teaching many nano fabrication techniques, and, most importantly, how to deliver a conference presentation), Deirdre Olynick (for making sense of etching), Peter James (Jim) Schuck (for teaching me plasmonics and being an amazing person as well as a scientist), Larry Southern (for being honest and fun), and Erin Wood (for all his patience, careful attention to detail, and for not taking things too seriously).

On the campus side of things, I want to thank professor David Attwood for explaining and helping me figure out the flow of the graduate school and what comes after that. Being a part of the Applied Science and Technology graduate group has been a magnificent experience for me, and this is primarily due to Patricia Berumen, our group coordinator. She has put endless hours into individually caring for every student and now that she is retiring, AS&T will never be the same.

Throughout these past years I have relied on my friends for keeping me sane: Nate Butler (for always being there), Paul Edwards (for his wise advice), Ian Henderson (for being an awesome and fearless dude), and Christoph Senft (for ensuring I do not lose track of what's important in life). I also want to thank my brother, Andrey Polyakov, for inspiring me to do science in the first place, my sister, Svetlana Polyakova, for her inspirational courage to do what she feels passionate about, and my parents, who are the most amazing people and it is impossible to adequately thank them.

Although I have now thanked most of the people for their contribution to this work and their companionship, it is no secret that life to me only makes sense when I am next to my wife, Quynh Lien Ngo, and my daughter Yuki. Thank you.

Enhanced Photocathodes for the FEL

1

1.1 Introduction

A photocathode is a device that produces electrons via the photoelectric effect: the photocathode is illuminated with a laser, and, provided that the incident photons have energy exceeding the cathode's work function, electrons are ejected from the surface [1, 2]. This simple principle is the basis for the generation of the electron beam that is used by a Free Electron Laser (FEL)[3–5], see figure 1.1. Once ejected, electrons form an Coulomb-expanding cloud that is accelerated close to the speed of light by a linear accelerator. The key to the FEL lasing lies in passing the electron beam through a set of undulators that cause electrons to undergo a transverse oscillatory motion, which produces a radiation that self-interacts with the electron beam causing it form into micro-bunches [6]. At this point, the electrons

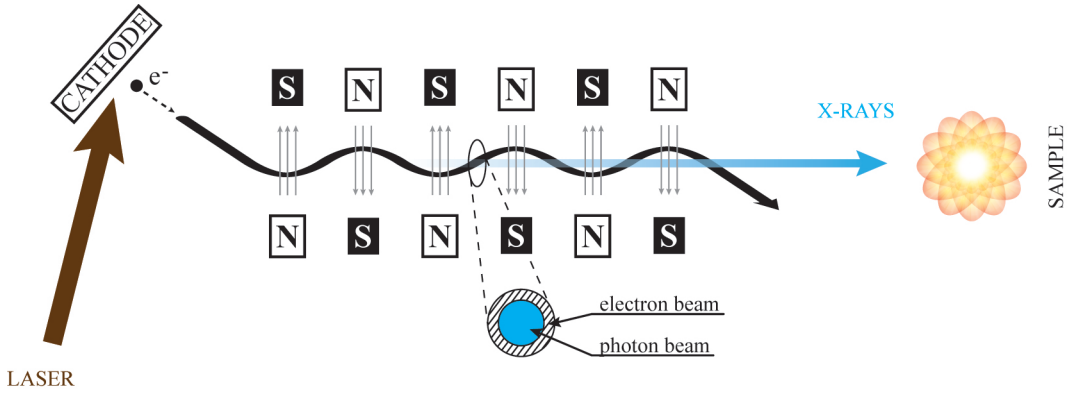

Figure 1.1: *Free Electron Laser is an ultra-bright light source that operates based on acceleration of electrons moving close to the speed of light. As the electrons pass through the undulators, they undergo an oscillatory motion in the transverse plane emitting radiation. This radiation interacts with the electron beam itself producing an in-phase radiation that scales as N_e^2, where N_e is the number of electrons.*

oscillate in phase resulting in a dramatic increase in the radiation intensity[1]

$$I \propto N_e^2.$$

For optimum operation, the radiation spot size produced by the electron beam must be *matched* to the size of the electron beam itself maximizing the self-interaction action (see figure 1.1 for an illustration of the photon and the electron beam overlap). The phase space size of the diffraction limited photon beam is $\lambda/4\pi$, so the phase space of the electron beam, called the *transverse emittance*, must be

$$\epsilon < \frac{\lambda}{4\pi}.$$

Another way to understand the importance of emittance is to examine the electron beam brightness, given by [7]

$$B \approx \frac{q_{pulse} f_{laser}}{\pi^2 \epsilon^2},$$

where q_{pulse} is the pulse charge and f_{laser} is the laser repetition rate. Maximizing brightness requires increasing the pulse charge via increasing the quantum efficiency (QE)–the number of electrons ejected per incident photon–and reducing the emittance. However, both of these quantities depend on the electron excess energy, $\hbar\omega - \phi_{eff}$, in the following way[2] [8–10]:

$$\begin{aligned} \epsilon &= \sigma_x \sqrt{\frac{\hbar\omega - \phi_{eff}}{3mc^2}}, \\ QE &\approx \eta_{tr} \frac{\left(\hbar\omega - \phi_{eff}\right)^2}{8\phi_{eff}\left(E_F + \phi_{eff}\right)}, \end{aligned}$$

where ϕ_{eff} is the metal's work function, E_F is the metals Fermi level, σ_x is the RMS beam size, and η_{tr} is the probability that a photon is absorbed and transported to the surface without scattering.

The minimum beam size σ_x is given by the space charge limit. When electrons are ejected from the cathode they encounter an radio frequency (RF) modulated field E_{rf} that is used to quickly pull them away from the metal surface. If, however, the amount of electrons ejected in one pulse is so large that they effectively screen the cathode from the RF field, those electrons that are ejected last will encounter a field that will push them back into the metal surface. This is the space charge limit, a point at which the photoemission saturates, which is given by [11]

$$Q_{max} = \epsilon_0 E_{rf} \times \sigma_x.$$

The only way then to reduce the emittance is to reduce the excess energy, which, in turn decreases metal's QE reducing the pulse charge. The two effects compete, and the balance is found for the excess energy of around 0.5 eV [9]. However, in this low energy regime metals

[1] In a synchrotron the electrons oscillate out of phase resulting in an emission $\propto N_e$.

[2] The detailed derivation for the QE is presented in later chapters.

are poor photoemitters: it takes over 10000 photons to eject one electron out [12]. Most of the excited electrons scatter with electrons in the bulk as they make their way to the surface. In this regime, each scattering is catastrophic (energy lost due to a single scattering event yields an electron with a below-work function total energy), which results in poor QE.

In semiconductors, once the electron is excited to the conduction band it does not scatter with other electrons, because it cannot loose too little energy, as there are no states in the gap. The primary energy-loss mechanism in this case is the electron-phonon scattering [13], which is a low energy loss event. As a consequence, an excited electron can stay a long time, on the picosecond scale, inside the bulk after excitation before exiting the material. This produces a long temporal spread from a semiconductor photocathode.

In a metal, due to the catastrophic nature of the electron-electron scattering, the temporal response is on the same time scale as the pulse length of the incident beam, because only those electron that travel directly to the surface without electron-electron scattering are able to exit the bulk material [12]. Due to this fast time response a metal photocathode could be operated in the *blow-out* regime, where the ejected electrons form a cloud with only linear internal fields [14]. The resulting electron bunch would have a shape of a pancake and immediately blow up under the internal Coulomb forces resulting in a uniform 3D ellipsoid with only linear space charge fields, which is an ideal distribution for a high-brightness FEL operation [15].

In this work, a new approach is presented for improving the QE of a bulk metal to satisfy the requirements of Next Generation Light Sources, while maintaining a metal's intrinsic fast response. This is done by introducing a set of nano-cavities on the metal surface, where Surface Plasmon Polariton (SPP) modes are excited enhancing the metal's performance as a photocathodes by orders of magnitude. The rest of this chapter describes the motivation for using the SPP coupling, with the detailed description of the phenomenon presented in chapters 2, 3, and 4. Chapter 5 describes the fabrication methods and chapters 6 and 7 present the experimental demonstration of the optical light trapping and the resulting increase in the photoemission from a plasmon enhanced photocathode.

1.2 Metal Quantum Efficiency

As explained earlier, for optimum FEL operation, the photon energy is chosen to be about 0.5 eV above the metal work function to balance minimizing the emittance and maximizing the QE. For metal films thicker than the optical skin depth–or for bulk metal–transmission is zero, in which case η_{tr} can be expressed as [8] (R is the reflectivity, λ_{opt} is metal's optical skin depth, and λ_{e-e} is the electron mean free path):

$$\eta_{tr} = \frac{1 - R(\omega)}{1 + \frac{\lambda_{opt}(\omega)}{\lambda_{e-e}(\omega)}}.$$

As will be shown later, it is possible to design a system where light couples to the SPP modes on the surface yielding a perfect absorber at a chosen wavelength, therefore it is

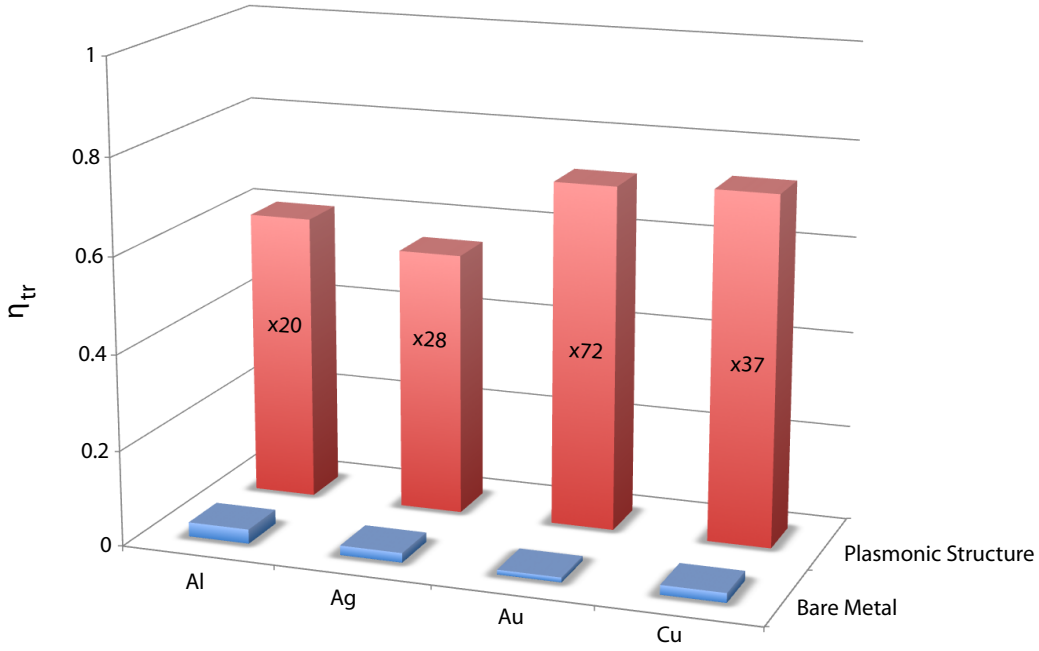

Figure 1.2: *Coupling to the SPP modes on the metal surface can be used to reduce the metal's reflectivity to zero and decrease the effective metal's skin depth. These effects lead to a dramatic increase in the metal's QE.*

Metal	ϕ (eV)	$E_{SP}^{max}(eV)$	λ_0 (nm)	n	k	R (%)	λ_{e-e} (nm)	λ_{opt} (nm)	λ_{SP} (nm)
Al	4.41	10.5	266	0.217	3.12	92.5	6.2	6.67	3.8
Ag	4.73	3.5	532	0.129	3.19	95.5	9.7	13.26	7.51
Au	4.8	2.45	800	0.181	5.12	97.4	11.6	12.45	4.7
Cu	4.3	2	800	0.25	5.03	96.3	11.6	12.67	4.88

Table 1.1: *Characteristics of typical materials used for photocathodes. The wavelengths of interest were chosen according to commonly available lasers in the regions where the corresponding material has a plasmonic response. The parameters λ_{e-e}, λ_{opt}, and λ_{SP} are the electron-electron scattering length, the penetration length of the normal-incidence field into the metal, and the spatial extension of the SPP evanescent fields into the metal respectively. The work function of gold varies from 4.7 eV to 5.6 eV depending of the preparation method [16, 17].*

possible to reduce the $R(\omega)$ factor to zero. However, it turns out that an SPP has shorter evanescent fields than optical skin depth, which can be expressed in terms of the material's dielectric function [18]:

$$\begin{aligned} \lambda_{opt} &= \frac{\lambda_0}{4\pi}\frac{1}{n'_{metal}}, \\ \lambda_{SP} &= \frac{\lambda_0}{2\pi}\left(\frac{\epsilon'_{metal}+\epsilon_{dielectric}}{[\epsilon'_{metal}]^2}\right). \end{aligned}$$

The only way to increase the QE without adversely affecting the emittance is to decrease the optical penetration depth and reduce the metal reflectivity. Coupling to the SPP modes on the metal surface is the ideal solution for this problem. Strong QE increase was demonstrated using this approach by Callcott *et al.* [19], where a prism was coated with a 20 nm thin aluminum film and the electrons were emitted from the backside. The dramatic increase in the QE is attributed to light coupling to the SPP modes on the aluminum–vacuum interface, where the reflectivity is extinguished, the field penetration depth is reduced, and the field strength is greatly enhanced. For comparison, a theoretical case study for four metals is presented in table 1.1 and the results are plotted in figure 1.2 (without accounting for the SPP field enhancement). For each material, coupling to the SPP modes leads to a dramatic increase in the QE.

1.3 Plasmon Enhancement

The plasmonic coupling on the metal surface can be achieved by a variety of coupling mechanisms as described in the next chapter. In this project, the SPP modes are excited in rectangular nano-grooves arranged in a subwavelength grating, see figure 4.1 on page 39 for a schematic. For an order-of-magnitude estimate of the photoemission from a metal with such plasmonic structure on the surface assume that all the photoelectrons come from the NGs and that the field enhancement inside the NGs can be modeled as an average value:

$$A^{NG}_{emit} = \left(\frac{w}{p}\right) A^{FS}_{emit},$$

where A_{NG} and A_{FS} is the effective emission area of the plasmonic and the bare-metal cathodes respectively, w is the NG width and p is the NG period. Similarly, the light intensity $I^{FS} = E_{ave}/(f_{laser}\tau_{pulse}A_{beam\ spot})$ (τ_{pulse} is the laser pulse length, $A_{beam\ spot}$ is the root mean square (RMS) beam size equivalent to σ_x in the previous section) at the cathode surface can be expressed in terms of the field enhancement factor FE as:

$$I^{NG} = FE \times I^{FS}.$$

The total emitted charge is then given by the Fowler-DuBridge photoemission model [20]:

$$Q_n = C_n\,(1-R)^n\, I^n A_{emit}, \tag{1.1}$$

where C_n represents all the photoemission factors unaffected by the plasmonic structure, except the field penetration depth, which is not accounted for explicitly in the context of this model.

For a higher order photoemission processes there is no clear definition of the quantum efficiency in terms of number of electrons produced as a function of the number of incident photons, because the total emitted charge increases non-linearly with the light intensity. Instead, it is useful to define relative quantum efficiency as a ratio between the energy required to produce the same emitted charge from the NG structured surface versus the flat metal surface:

$$QE_n^{rel} \equiv \frac{Q^{NG}}{E_{pulse}^{NG}} \times \frac{E_{pulse}^{FS}}{Q^{FS}} = \frac{E_{pulse}^{FS}}{E_{pulse}^{NG}}.$$

The total emitted charge for the two cases is given by combining the above equations:

$$Q_n = \begin{cases} C_n \left(1 - R^{FS}\right)^n \left(\frac{E_{pulse}^{FS}}{\tau_{pulse} A_{beam\ \ spot}}\right)^n A_{emit}^{FS} \\ C_n \left(1 - R^{NG}\right)^n \left(\frac{FE \times E_{pulse}^{NG}}{\tau_{pulse} A_{beam\ \ spot}}\right)^n A_{emit}^{NG} \end{cases}.$$

Setting the laser spot size equal, the relative quantum efficiency is given by:

$$QE_n^{rel} = FE \times \left(\frac{1 - R^{NG}}{1 - R^{FS}}\right) \left(\frac{w}{p}\right)^{\frac{1}{n}}. \tag{1.2}$$

A conservative estimate for a gold plasmonic cathode with the NG structure tuned to yield 10% reflectivity at 800 nm and FE of 50 would be a QE improvement over 300 for the multi-photon photoemission. Under low intensity illumination the multi-photon photoemission cross section is very small, rendering the overall yield negligible. However, for very short laser pulses (less than 50 fs), the peak laser intensity is high enough that multi-photon photoemission becomes the dominant process [20].

The plasmonic structures provide the right environment for enhancing multi-photon photoemission in the high laser intensity regime (short pulses) that is ideal for optimum FEL operation due to the ellipsoidal shape of the electron cloud formed at the cathode [14]. This is achieved by coupling to the SPP modes in the nanostructures that leads to reflectivity reduction close to zero, reduced field penetration, and strong field enhancement at the metal surface. In combination, these effects lead to a dramatic increase in the overall performance of a metal photocathode.

Theory and Applications of SPP Waves

2

2.1 Theory of Surface Plasmon Polaritons

The surface plasmon polariton (SPP) introduced in the previous chapter is a quasi particle representing an electron density wave on the metal-dielectric boundary [18]. A good model for the SPP wave is to treat it as a monochromatic wave ($E \propto \exp\left[i(\bar{k} \cdot \bar{r} - \omega t)\right]$) propagating along the metal-dielectric interface [21]. However, unlike the incident light, which is a transverse wave, the SPP wave has a longitudinal component. Therefore, from momentum conservation, the incident light must have a polarization component perpendicular to the surface as well as along the metal surface, that is, the incident light must be *p-polarized*. S-polarized light cannot couple to SPP modes, therefore in the discussion that follows only p-polarized light is considered.

The SPP wave solution comes naturally from Maxwell's equations at a boundary between the free electron metal and a dielectric, for simplicity, air. The SPP wave is itself made up from the "free charge" density fluctuations on the metal surface, and, therefore, for the

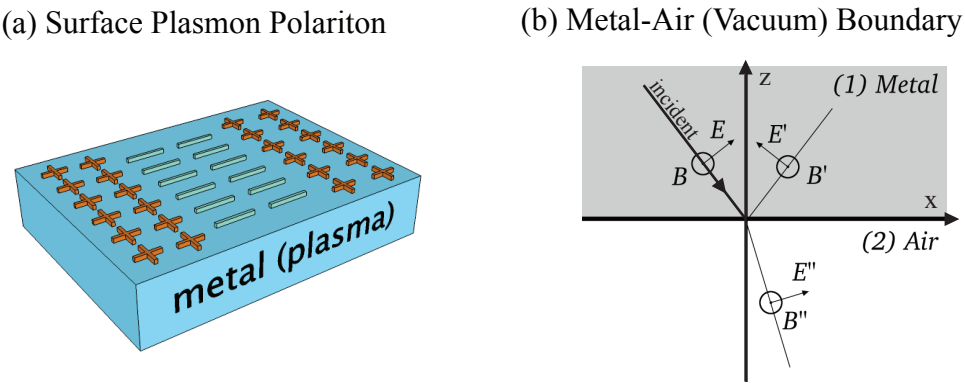

Figure 2.1: *The SPP wave is a quasi-particle representing electron density fluctuations (a). The solution for this mixed longitudinal and transverse wave comes directly from the Maxwell's equations by solving for the light propagation across the boundary (b).*

boundary conditions the surface charge is set to 0: $\boldsymbol{\nabla} \cdot \boldsymbol{D} = 0$ and $\boldsymbol{\nabla} \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}$. Since the SPP wave propagates *along* the surface at the interface, it can be decomposed into two waves: one in medium 1 (metal) and another in medium 2 (air). Then the SPP wave is treated as an electro-magnetic (EM) wave interference between the *reflected* and *refracted* radiation [18] (for notation, refer to the schematic in figure 2.1). The applicable boundary conditions are:

$$E_x^{(1)} = E_x^{(2)}, H_y^{(1)} = H_y^{(2)}, D_z^{(1)} = D_z^{(2)} \Rightarrow \epsilon_1 \epsilon_0 E_z^{(1)} = \epsilon_2 \epsilon_0 E_z^{(2)},$$

where $\epsilon_{1,2}$ is the relative dielectric constant.

From the boundary conditions the tangential component of the k-vector is continuous across the interface: $k_x^{(1)} = k_x^{(2)}$. The explicit formulas for the $\boldsymbol{H}$- and $\boldsymbol{E}$-fields are:

$$\begin{aligned}
\boldsymbol{H} &= \boldsymbol{H} \cdot \exp\left[i\left(\bar{k} \cdot \bar{r} - \omega t\right)\right] \\
&= (H_x \hat{x} + H_y \hat{y} + H_z \hat{z}) \cdot \exp\left[i\left(k_x x + k_y y + k_z z - \omega t\right)\right] \\
&= H_y \exp\left[i\left(k_x x + k_y y + k_z z - \omega t\right)\right] \hat{y}, \\
\boldsymbol{E} &= \boldsymbol{E} \cdot \exp\left[i\left(\bar{k} \cdot \bar{r} - \omega t\right)\right] \\
&= (E_x \hat{x} + E_z \hat{z}) \cdot \exp\left[i\left(k_x x + k_y y + k_z z - \omega t\right)\right].
\end{aligned}$$

Plugging these equations into Ampere's Law, $\boldsymbol{\nabla} \times \boldsymbol{H} = \frac{\partial \boldsymbol{D}}{\partial t}$, yields (superscripts $1, 2$ denote the quantities in the medium 1 and 2, see figure 2.1):

$$\boldsymbol{\nabla} \times \boldsymbol{H}^{(1,2)} = \frac{\partial \boldsymbol{D}^{(1,2)}}{\partial t}.$$

From $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$ the y-component of the wavevector is $k_y = 0$. Then, plugging the wave solutions into the above equation for medium 1 yields:

$$\begin{aligned}
-\frac{\partial}{\partial z} H_y^{(1)} \exp[...]\, \hat{x} + \frac{\partial}{\partial x} H_y^{(1)} \exp[...]\, \hat{z} &= -i\omega\epsilon_1\epsilon_0 \cdot \left(E_x^{(1)} \hat{x} + E_z^{(1)} \hat{z}\right) \exp[...] \\
-ik_z^{(1)} H_y^{(1)} &= -i\omega\epsilon_1\epsilon_0 E_x^{(1)} \\
ik_x^{(1)} H_y^{(1)} &= -i\omega\epsilon_1\epsilon_0 E_z^{(1)} \\
k_z^{(1)} H_y^{(1)} &= \omega\epsilon_1\epsilon_0 E_x^{(1)} \\
\frac{k_z^{(1)} H_y^{(1)}}{\epsilon_1\epsilon_0} &= \omega E_x^{(1)}.
\end{aligned}$$

The situation is the same for medium 2, except that the wave is propagating away from the surface, so there is a net negative sign on the k-vector. From the $\boldsymbol{H}$- and E_x-field continuity across the interface:

$$\begin{cases} H_y^{(1)} - H_y^{(2)} &= 0 \\ E_x^{(1)} - E_x^{(2)} &= 0 \end{cases} \Rightarrow \begin{cases} H_y^{(1)} - H_y^{(2)} &= 0 \\ \frac{k_z^{(1)}}{\epsilon_1} H_y^{(1)} + \frac{k_z^{(2)}}{\epsilon_2} H_y^{(2)} &= 0 \end{cases} \Rightarrow \begin{pmatrix} 1 & -1 \\ \frac{k_z^1}{\epsilon_1} & \frac{k_z^2}{\epsilon_2} \end{pmatrix} \cdot \begin{pmatrix} H_y^{(1)} \\ H_y^{(2)} \end{pmatrix} = 0.$$

For the non-trivial solution, the determinant of the coefficient matrix above must be zero, therefore:

$$\frac{k_z^{(1)}}{\epsilon_1} + \frac{k_z^{(2)}}{\epsilon_2} = 0.$$

From the definition of the monochromatic plane wave and the relations between the phase velocity in medium with index of refraction n ($v = \frac{c}{n}$, $n^2 = \epsilon$), the relation between k_x and k_z is found to be:

$$\begin{aligned}
\nabla^2 E - \frac{1}{v^2}\frac{\partial^2 E}{\partial t^2} &= 0 \\
-k^2 E + \frac{1}{v^2}\omega^2 E &= 0 \\
k_x^2 + k_z^2 &= \frac{\omega^2}{v^2} \\
&= \omega^2 \left(\frac{c}{n}\right)^{-2} \\
&= \left(\frac{\omega}{c}\right)^2 n^2, \\
k_x^2 + \left[k_z^{(1,2)}\right]^2 &= \epsilon_{1,2}\left(\frac{\omega}{c}\right)^2. \qquad (2.1)
\end{aligned}$$

Combining the above results yields:

$$\begin{cases} \frac{k_z^{(1)}}{\epsilon_1} + \frac{k_z^{(2)}}{\epsilon_2} & = 0 \\ k_x^2 + \left(k_z^{(1)}\right)^2 & = \epsilon_1 \cdot \frac{\omega^2}{c^2} \\ k_x^2 + \left(k_z^{(2)}\right)^2 & = \epsilon_2 \cdot \frac{\omega^2}{c^2} \end{cases} \Rightarrow \begin{cases} k_x^2 + \left(k_z^{(1)}\right)^2 & = \epsilon_1 \cdot \frac{\omega^2}{c^2} \\ k_x^2 + \left(-k_z^{(1)}\right)^2 \left(\frac{\epsilon_2}{\epsilon_1}\right)^2 & = \epsilon_2 \cdot \frac{\omega^2}{c^2} \end{cases},$$

$$\Rightarrow \begin{cases} k_x^2 + \left(\epsilon_1 \frac{\omega^2}{c^2} - k_x^2\right)\left(\frac{\epsilon_2}{\epsilon_1}\right)^2 & = \epsilon_2 \frac{\omega^2}{c^2} \\ k_z^2 + k_x^2 & = \epsilon_2 \left(\frac{\omega}{c}\right)^2 \end{cases}$$

$$\begin{aligned}
k_x^2\left[1 - \left(\frac{\epsilon_2}{\epsilon_1}\right)^2\right] &= \epsilon_2\frac{\omega^2}{c^2} - \epsilon_1\frac{\omega^2}{c^2}\left(\frac{\epsilon_2}{\epsilon_1}\right)^2 = \frac{\omega^2}{c^2}\left[\epsilon_2 - \epsilon_1\left(\frac{\epsilon_2}{\epsilon_1}\right)^2\right] \\
k_x^2 &= \frac{\omega^2}{c^2} \cdot \frac{\epsilon_2 - \epsilon_1\left(\frac{\epsilon_2}{\epsilon_1}\right)^2}{1 - \left(\frac{\epsilon_2}{\epsilon_1}\right)^2} \\
&= \frac{\omega^2}{c^2}\left[\frac{\epsilon_2\epsilon_1^2 - \epsilon_1\epsilon_2^2}{\epsilon_1^2 - \epsilon_2^2}\right] \\
&= \frac{\omega^2}{c^2}\left[\frac{\epsilon_1\epsilon_2\left(\epsilon_1 - \epsilon_2\right)}{\left(\epsilon_1 + \epsilon_2\right)\left(\epsilon_1 - \epsilon_2\right)}\right].
\end{aligned}$$

The above expression gives the dispersion relation for the SPPs:

$$k_x = \frac{\omega}{c}\sqrt{\frac{\epsilon_1\epsilon_2}{\epsilon_1+\epsilon_2}}. \tag{2.2}$$

In the above expression, $\epsilon_1' < 0$ in a metal and $\epsilon_2 > 0$ in air, therefore, for the SPP wave propagation: $\Re[k_x] = \frac{\omega}{c}\left(\frac{\epsilon_1'\epsilon_2}{\epsilon_1'+\epsilon_2}\right) > \frac{\omega}{c} \Rightarrow \epsilon_1' < -\epsilon_2$. In other words, if the real part of the dielectric constant of one material is more negative then the dielectric constant of the surrounding medium, an SPP wave can be propagate.

Free Electron Metal: The Drude Model

The dispersion relation given by equation 2.2 contains the optical properties of the metal in terms of the dielectric function. To determine the complete dispersion relation in terms of the fundamental constants alone, the metal can be approximated using Drude's model [18]. However, care must be taken in interpreting the results based on this model to ensure that metal under consideration is well approximated in the spectral region of interest. For example, for aluminum the model is applicable anywhere from the UV to the NIR, for gold and copper–due to the interband transitions–it is from ≈ 600 nm to the NIR. In the general case, the Drude model is

$$\epsilon(\omega) = 1 + \frac{Z \cdot n_a e^2}{\epsilon_0 m}\sum_j f_j\left(\omega_j^2 - \omega^2 - i\omega\gamma_j\right)^{-1},$$

where n_a is the atom density[1], $\sum f_j = Z$, ω_j is the resonance frequency, and γ_j is the damping factor. Treating the metal as a free electron gas assumes ω to be far above the highest resonance frequency

$$\epsilon_1(\omega) = 1 - \frac{n_a e^2}{\epsilon_0 m}\cdot\left(\frac{1}{\omega^2}\right) = 1 - \frac{\omega_p^2}{\omega^2},$$

where $\omega_p^2 = \frac{Z\cdot N e^2}{\epsilon_0 m}$ is the plasma frequency and the Z factor represents the average number of *conduction* electrons contributed by each atom. Plugging this result in equation 2.2 yields the dispersion relation expressed in terms of the fundamental constants alone (for the derivation below $\epsilon_2 = 1$):

[1]The atom density is given by $n_a = \frac{\rho N_A}{M_{At}}$, where ρ is the material density, N_A is the Avagadro's number, and M_{At} is the atomic weight of the material.

$$\begin{aligned}
k_x &= \frac{\omega}{c}\sqrt{\frac{\epsilon_1}{\epsilon_1+1}} \\
(k_x c)^2(1+\epsilon_1) &= \omega^2\epsilon_1 \\
(k_x c)^2\left(1+1-\frac{\omega_p^2}{\omega^2}\right) &= \omega^2\left(1-\frac{\omega_p^2}{\omega^2}\right) \\
(k_x c)^2\left(2\omega^2-\omega_p^2\right) &= \omega^4-\omega^2\omega_p^2 \\
\omega^4-\left[2(k_x c)^2+\omega_p^2\right]\cdot\omega^2+\omega_p^2(k_x c)^2 &= 0
\end{aligned}$$

Solving the above equation gives:

$$\begin{aligned}
\omega^2 &= \frac{1}{2}\left\{\left[2(k_x c)^2+\omega_p^2\right]\pm\sqrt{\left[2(k_x c)^2+\omega_p^2\right]^2-4\omega_p^2(k_x c)^2}\right\} \\
&= \frac{1}{2}\left\{2(k_x c)^2+\omega_p^2\pm\sqrt{4(k_x c)^4+4(k_x c)^2\omega_p^2+\omega_p^4-4\omega_p^2(k_x c)^2}\right\} \\
&= \frac{1}{2}\left\{2(k_x c)^2+\omega_p^2\pm\sqrt{4(k_x c)^4+\omega_p^4}\right\} \\
&= \frac{1}{2}\left\{2(k_x c)^2\pm 2(k_x c)^2\sqrt{1+\frac{1}{4}\left(\frac{\omega_p}{k_x c}\right)^4}+\omega_p^2\right\} \\
&= \frac{1}{2}\left\{2(k_x c)^2\cdot\left[1\pm\sqrt{1+\frac{1}{4}\left(\frac{\omega_p}{k_x c}\right)^4}\right]+\omega_p^2\right\}
\end{aligned}$$

The plus-minus sign in the above equation gives two solutions as shown in figure 2.2. Taking the positive sign of the square root corresponds to the top branch, where $k_x < \omega/c$. From equation 2.1, this condition corresponds to a real k_z, which, in turn, yields a propagating wave. This is the dispersion relation for light propagating in the metal.

Taking the negative sign, however, results in the lower branch of figure 2.2, where $k_x > \omega/c$. In this case, from equation 2.1 it follows that k_z is imaginary. Noting that the electric field in the z-direction is proportional to $\exp[-ik_z z]$, an imaginary k_z yields an exponential decay function. This is the surface wave: it decays into the material and represents the SPP wave. The upper limit on the SPP energy is given by [18]:

$$E_{max}^{plasmon} = \frac{\hbar\omega_p}{\sqrt{2}} = \sqrt{2\pi a_0 N_A\left(\frac{Z\rho}{M_{At}}\right)}\cdot\left[\frac{(\hbar c)^2}{m_e c^2}\right],$$

where a_0 is Bohr's radius. Note that as $k_x \to \infty$ the maximum frequency goes to the limit $\omega \to \frac{\omega_p}{\sqrt{2}}$. For example, in aluminum each atom contributes more than one electron to

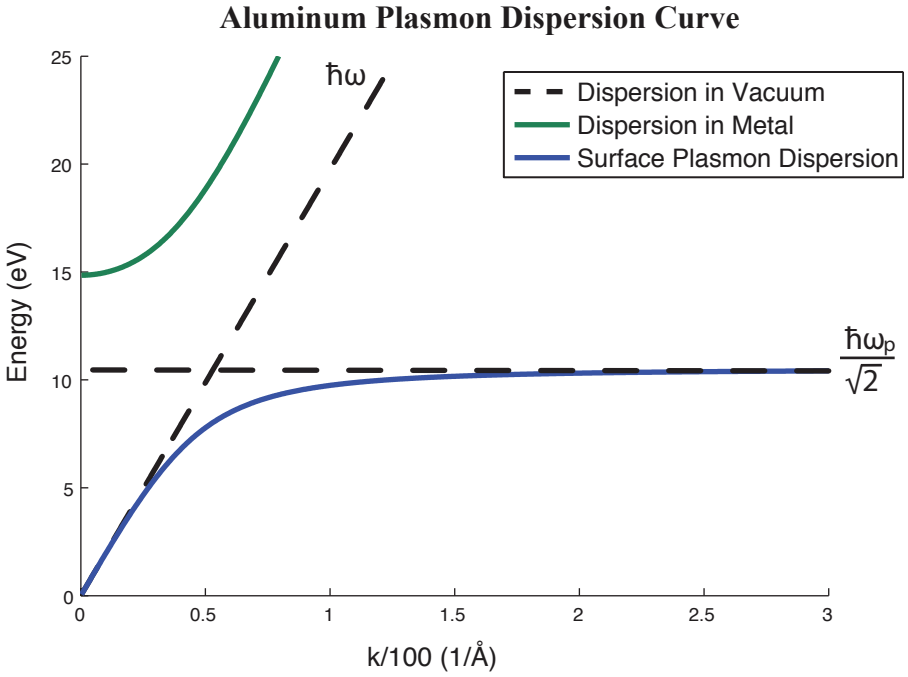

Figure 2.2: *The dispersion in vacuum (the light cone) is plotted with a dashed line. For the SPP wave, the dispersion curve lies to the right of the light cone, while the dispersion inside the metal is always to the left. The graphs shown are only valid in the regime where the metal is approximated well by the Drude model.*

the "sea of electrons" in the free electron gas model. The electron structure of aluminum is $[Ne]\,3s^2 3p^1$, so each atom can contribute up to $Z = 3$ electrons. Experimentally, the maximum SPP energy is ≈ 10.5 eV. Relating this to the expression for the ω_p yields the correct Z:

$$E_{max}^{plasmon} = \hbar\frac{\omega_p}{\sqrt{2}} = \frac{\hbar}{\sqrt{2}}\sqrt{\frac{Z \cdot n_a e^2}{m\epsilon_0}} \Rightarrow Z = \frac{2m\epsilon_0}{Ne^2}\left(\frac{E_{max}}{\hbar}\right)^2 \approx 2.66.$$

The final expression for the SPP dispersion is given below; plugging in the above value for Z yields the graph shown in figure 2.2:

$$\omega_{SP} = \frac{1}{\sqrt{2}}\left\{2\left(k_x c\right)^2 \cdot \left[1 - \sqrt{1 + \frac{1}{4}\left(\frac{\omega_p}{k_x c}\right)^4}\right] + \omega_p^2\right\}^{\frac{1}{2}}.$$

2.2 Applications

Exciting the SPP waves is used in a variety of applications [22–27] ranging from light filtering [28, 29] to harmonic generation [30–33]. Structures such as lamellar gratings, split-ring resonators, plasmonic coaxial guides and bowties among others have been investigated in great detail both theoretically [27, 34–37] and experimentally [38–43]. For a specific example, consider the setup, where the properties of a thin metal film are investigated by observing the reflectivity at a fixed wavelength from a sample as a function of the angle of incidence. This simple example illustrates many properties that will be later directly translated into more complex systems.

Coupling to SPP Modes via Prism Momentum

An important feature of the SPP dispersion curve is illustrated in figure 2.2: the dispersion curve always lies to the right of the light cone. This means that SPPs cannot couple directly to light, because there is a momentum mismatch between light in vacuum and the SPPs on the metal surface, which violates momentum conservation [18]. In this condition, SPPs are non-radiative. To compensate for the momentum mismatch, the light wave vector needs to be "stretched" to match the momentum of the SPP. This can be achieved by using a dielectric with a higher index of refraction than vacuum.

For example, the metal film can be deposited onto a prism (referred to as substrate below). Light inside the prism is of a larger k due to the higher index of refraction n_{gl} of the prism compared to vacuum (or air). The metal film is illuminated through the prism; this way the component of the momentum parallel to the metal surface can be adjusted by changing the angle of incidence. The shape of the substrate itself plays a very important role in determining the range of the angles of incidence which satisfy the light-SPP momentum matching condition. The k_0-vector of the incident light wave is always less then the corresponding (same energy) k_{SP}-vector of the SPP. When the EM-wave passes through the glass, the wave number is increased by a factor of n_{gl}:

$$\begin{aligned} f &= \frac{c}{\lambda_0} = \frac{c}{n} \cdot \frac{1}{\lambda_{gl}} \Rightarrow \lambda_{gl} = \frac{1}{n_{gl}} \lambda_0 \Rightarrow k_{gl} = \frac{1}{\lambda_{gl}} = n_{gl} \frac{1}{\lambda_0}, \\ k_{gl} &= n_{gl} k_0. \end{aligned}$$

With the larger k_{gl} vector, it is possible to match the momentum of light to the momentum of the SPP wave. Taking the projection of the k_{gl} vector onto the plane of the glass-aluminum interface yields:

$$k_{||} = k_{gl} \sin\theta_{gl} = k_0 \cdot n_{gl} \sin\theta_{gl}.$$

From Snell's law for the geometry (b) from figure 2.3, $n_{gl} \sin\theta_{gl} = n_0 \sin\theta_0$, where $n_0 = 1$ for air or vacuum. Therefore, $k_{||} = k_0 \sin\theta_0$. The situation is more complicated for the

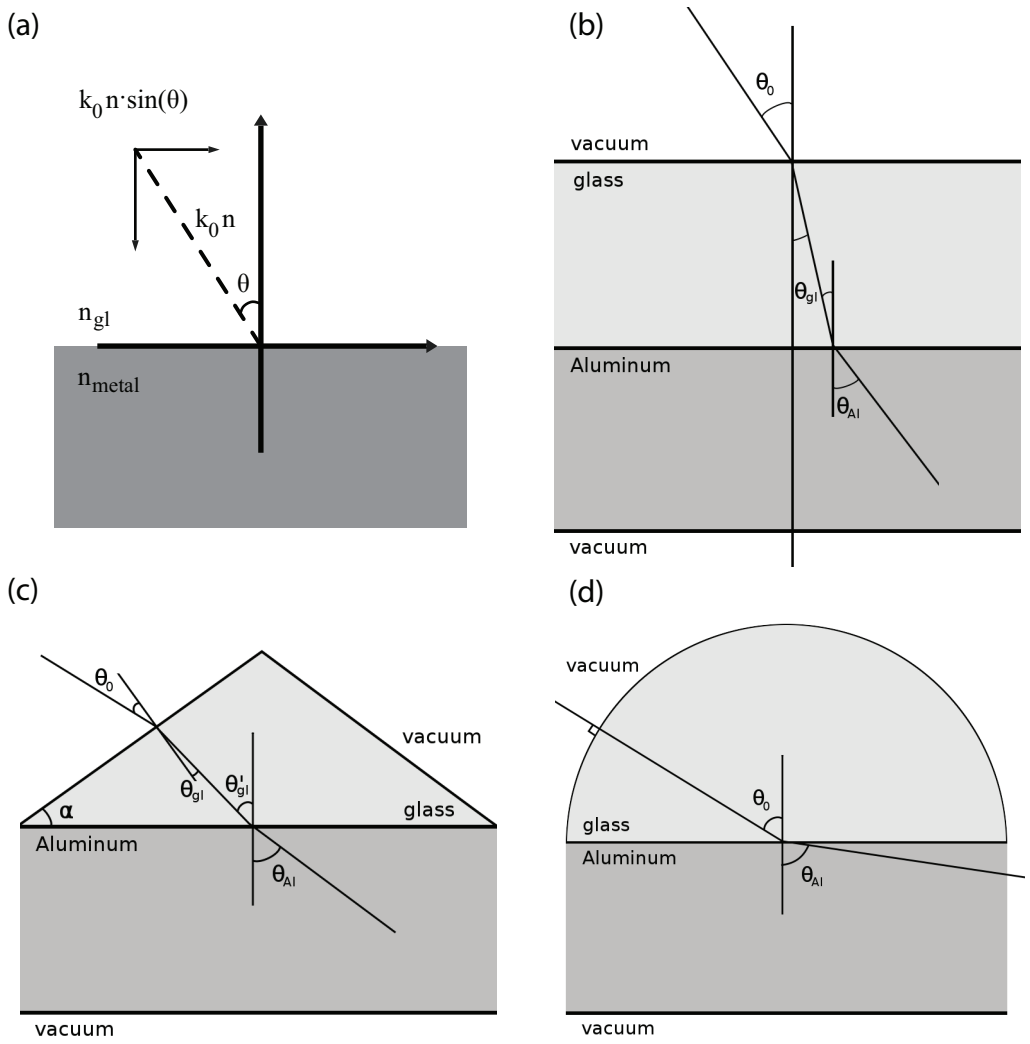

Figure 2.3: *The metal film is deposited on a prism substrate. Changing the angle of incidence of the light relative to the metal surface allows adjustment of the parallel component of the momentum (a) for coupling to the SPP waves. Three geometries are considered: (b) the planar geometry, (c) the right-angle prism, and (d) the half-cylinder prism. The geometry in (d) has the benefit of allowing a full range of angles of incidence at the glass-metal interface from 0 to 90 degrees. However, the curved surface of the prism focuses the beam, which must be compensated for in the experimental setup.*

second option, geometry (c). The angle of incidence in the glass medium at the glass-vacuum interface, θ_{gl} is expressed in terms of the prism angle α and the angle of incidence in the glass medium at the metal-glass interface, $\theta'_{gl} = \alpha + \theta_{gl}$, see figure 2.3(c) for reference.

$$\begin{aligned}
n_{gl}\sin\theta'_{gl} &= n_{gl}\sin\left(\theta_{gl}+\alpha\right) = n_{gl}Im\left\{\exp\left[i\left(\theta_{gl}+\alpha\right)\right]\right\} = n_{gl}Im\left\{\exp\left[i\theta_{gl}\right]\cdot\exp\left[i\alpha\right]\right\} \\
&= n_{gl}\left(\cos\theta_{gl}\sin\alpha+\sin\theta_{gl}\cos\alpha\right) = n_{gl}\cos\theta_{gl}\cdot\sin\alpha + n_{gl}\sin\theta_{gl}\cdot\cos\alpha \\
&= \sqrt{n_{gl}^2 - n_0^2\sin^2\theta_0}\cdot\sin\alpha + n_0\sin\theta_0\cdot\cos\alpha \\
&= n_0\sin\theta_0\left(\cos\alpha+\sqrt{\left(\frac{n_{gl}}{n_o\sin\theta_0}\right)^2-1}\right) \\
&= n_0\sin\theta_0\left(\cos\alpha+\sqrt{\left(\frac{n_{gl}}{n_{gl}\sin\theta_{gl}}\right)^2-1}\right) \\
&= n_0\sin\theta_0\left(\cos\alpha+\sqrt{\frac{\sin^2\theta_{gl}+\cos^2\theta_{gl}}{\sin^2\theta_{gl}}-1}\right) \\
&= n_0\sin\theta_0\left(\cos\alpha+\sqrt{\cot^2\theta_{gl}}\right) \\
&= n_0\sin\theta_0\left(\cos\alpha+\cot\theta_{gl}\right) \\
\Rightarrow\ k_{||} &= k_0\left(\cos\alpha+\cot\theta_{gl}\right)\cdot\sin\theta_0
\end{aligned}$$

In this geometry, adjusting the angles θ_0 results in stretching of the parallel component of the momentum. However, the accessible angles are limited by Snell's law: figure 2.4 shows the range of accessible angles of incidence at the metal-glass interface for all possible angles of incidence at the glass-vacuum interface. As this angle increases past 60 degrees, the internal angle of incidence at the metal-glass interface saturates at values below 43 degrees.

The full range of angles from 0 to 90 degrees is accessible only in the geometry shown in figure 2.4(d), where $\theta_{gl} = \theta_0$. The prism shaped as a half-cylinder yields $k_{||} = n_{gl}\sin\theta_0$. However, due to the curved surface of the prism, the incoming beam of finite diameter will focus requiring a compensation prism to keep the beam parallel at the metal-glass interface.

Surface Plasmon Resonance

The localization of the SPP wave close to the metal-dielectric boundary results in a strong dependance of the SPP dispersion on the boundary conditions. For example, a thin dielectric coating with high index of refraction produces a noticeable shift in term of the angle of incidence required to excite the plasmon wave. This effect is used directly in surface plasmon resonance (SPR) applications [44], for example, for real-time analysis of biospecific interactions. In the study that follows, the sample is prepared consisting of a thin film of aluminum

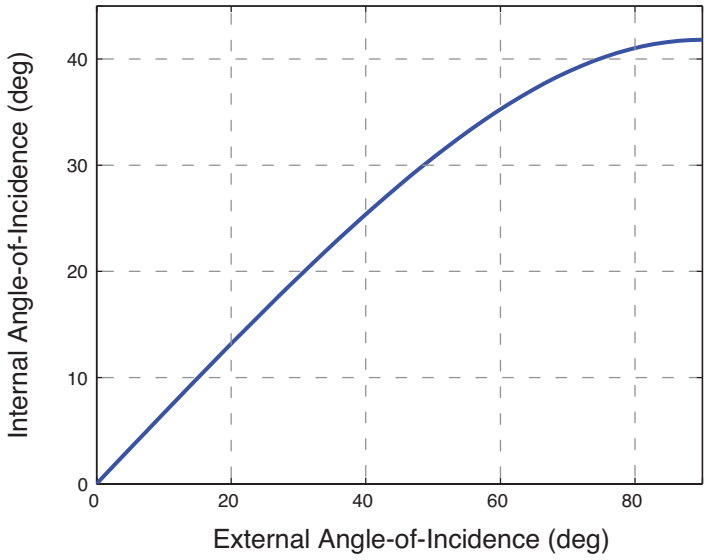

Figure 2.4: *The angles accessible in the right-angle prism geometry from figure 2.3(c) for* $n_0 = 1, n_{gl} = 1.4$. *As the angle of incidence increases at the glass-vacuum interface, so does the angle of incidence at the metal-glass interface, that produces the required change in the parallel component of the momentum of light. However, this internal angle of incidence saturates at below 43 degrees.*

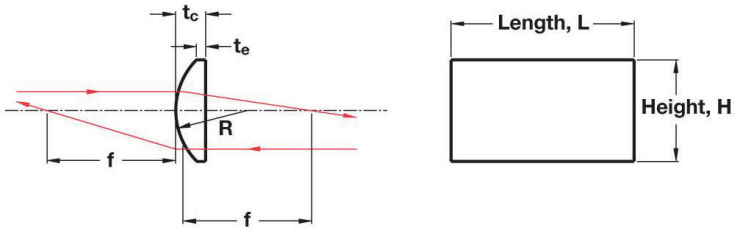

Figure 2.5: *Sample substrate geometry (image taken from the Thorlabs catalogue, part #LJ4918; all dimensions in mm):* $f = 10.0$, $H = 9.0$, $L = 14.0$, $R = 4.6$, $t_c = 5.6$, $t_e = 2.0$.

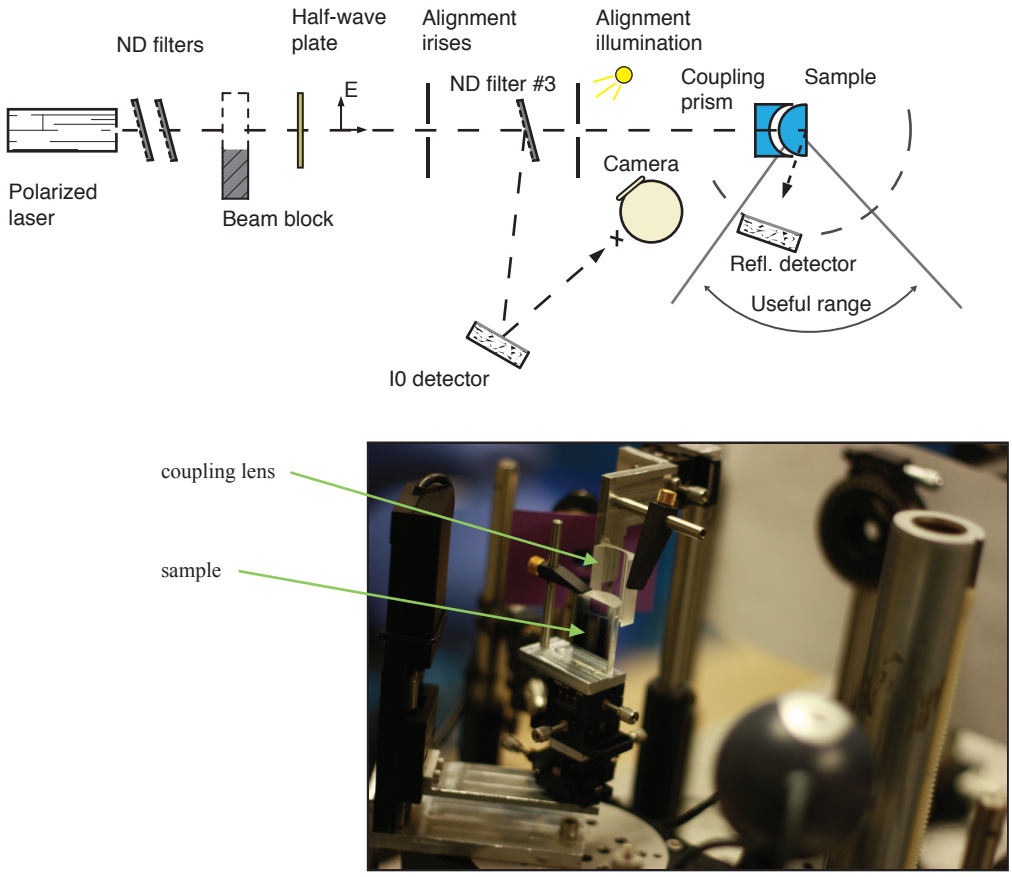

Figure 2.6: *The light passes through a set of conditioning elements providing 10 nm bandwidth light with a defined polarization. The sample is rotated starting from 0 up to 90 degrees with the detector collecting the reflected signal in a $\theta - 2\theta$ arrangement.*

on a semi-cylinder prism with an unknown thickness of aluminum oxide (alumina). SPR analysis is used to determine the optical properties and the thickness of the alumina.

Theta-2Theta Setup

To measure the 0-order reflectivity, a theta-2theta system was assembled. In such a system, the sample is rotated by an angle θ with respect to the incident beam, and the detector is then placed at an angle 2θ to record the reflectivity signal. The setup layout is shown in figure 2.6. The intensity of a linearly polarized laser[2] was adjusted by a set of three neutral density (ND) filters, with the polarization controlled via a $\lambda/2$-plate. The first two filters appear immediately after the laser itself, while the third ND filter is placed in between the

[2]The two polarized lasers used were: 594 nm @ 2.0 mW, and 543 nm @ 5 mW.

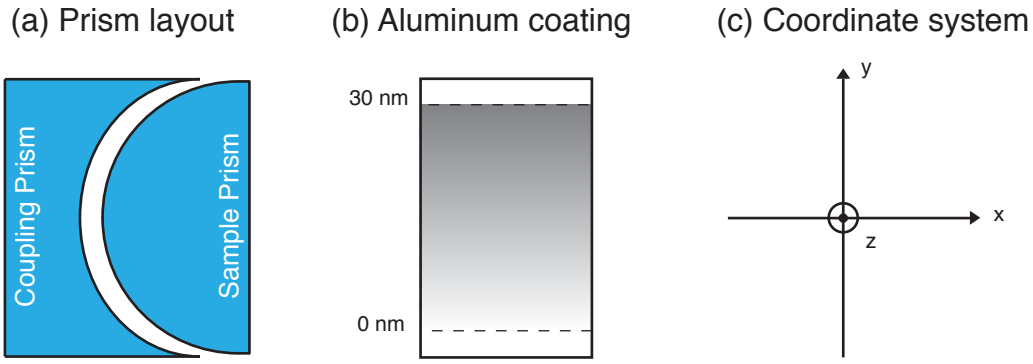

Figure 2.7: *The sample was prepared by magnetron sputtering a varying thickness of aluminum film in a wedge shape onto the half-cylinder prism. The sample prism is shown together with the coupling prism required to produce a parallel laser beam at the metal-prism interface. The thickness of the aluminum film was varied from 0 nm to 30 nm.*

two pinholes (used for alignment). The reference laser power (I_0) was measured from the scattered beam off the ND filter #3. For noise measurements a beam block was placed after the two ND filters and right before the $\lambda/2$-plate.

To ensure consistent alignment across measurements, the second pinhole was observed by the camera as indicated in figure 2.6. The sample was placed at the end of the beam path, preceded by the coupling prism that compensates the focusing effect of the sample prism's shape. Due to the dimensions of the coupling prism, the useful range of the setup was limited to the angles as indicated in the detector's path (shown by the dashed semi-circle).

Sample Geometry

The sample prism was made from UV-grade fused silica, it is shown schematically together with the coupling prism in figure 2.7. The coupling prism was made with dimensions exactly matching the sample prism to ensure a parallel beam at the aluminum-prism interface. The aluminum coating was varied in thickness from 0 nm to 30 nm in a wedge shape, deposited by magnetron sputtering, leaving 2 mm clearance for reference measurements.

The rotation axis of the sample's rotation stage was the only fixed reference point in the setup, with everything else aligned in reference to it. First, the coupling prism was aligned to the sample by observing the gap between the two prisms while rotating the sample. Second, using the two pinholes, the laser was aligned to pass through the middle of both prisms by observing the scattered reflections. Third, using the pinholes and the ND filter #3, the beam was adjusted to be 1 mm in diameter at the aluminum film.

Using Snell's law, the index of refraction is obtained as a function of the transmission cut-off angle (due to the total internal reflection):

$$n_{gl} = \frac{1}{sin\,(\theta_{max})}.$$

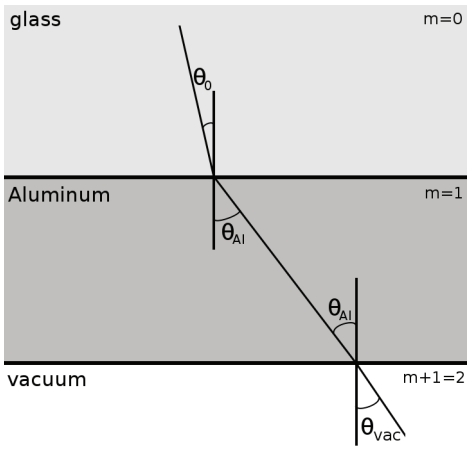

Figure 2.8: *The section of the sample under consideration is modeled well by solving the Maxwell's equations for only the three-layer system consisting of: glass, aluminum film, and vacuum.*

Meanwhile, the optical constants for aluminum and the aluminum oxide used were $n_{Al} = 1.16 + i \cdot 7.19$ and $n_{Al_2O_3} = 1.77$ respectively for 594 nm light and $n_{Al} = 0.93 + i \cdot 6.595$, $n_{Al_2O_3} = 1.77$ for 543 nm [45].

Reflectivity of the Multi-Layer System

Using the matrix formulation for solving the Maxwell's equations for a multi-layer system as shown in figure 2.8, the electric field can be found everywhere inside the metal film [46]. For simplicity, consider the system composed of an aluminum film on the semi-infinite glass substrate.

The relationships between the reflected and the incident waves at the incident (E_0^+, E_0^-) and metal-vacuum boundaries (E_{m+1}^+, E_{m+1}^-), is given by:

$$\begin{pmatrix} E_0^+ \\ E_0^- \end{pmatrix} = \frac{C_1 C_2 ... C_{m+1}}{t_1 t_2 ... t_{m+1}} \begin{pmatrix} E_{m+1}^+ \\ E_{m+1}^- \end{pmatrix}, \tag{2.3}$$

$$C_j = \begin{pmatrix} \exp\left[-i\delta_{j-1}\right] & r_j \cdot \exp\left[-i\delta_{j-1}\right] \\ r_j \cdot \exp\left[i\delta_{j-1}\right] & \exp\left[i\delta_{j-1}\right] \end{pmatrix}, \tag{2.4}$$

where the r_j, t_j are the Fresnel coefficients given in terms of the complex index of refraction n_j. The expression for δ_j involves $\cos\theta_j$ and the thickness of the j^{th} layer ($j \in [1, m] : \delta_j = \frac{2\pi}{\lambda_0} n_j \cos\theta_j h_j, \delta_0 = 0, \cos\theta_j = \frac{1}{n_j}\sqrt{n_j^2 - n_0^2 \sin^2\theta_0}$):

$$\begin{aligned} r_j &= \tfrac{n_{j-1}\cos\theta_j - n_j \cos\theta_{j-1}}{n_{j-1}\cos\theta_j + n_j \cos\theta_{j-1}}, \\ t_j &= \tfrac{2 n_{j-1}\cos\theta_{j-1}}{n_{j-1}\cos\theta_j + n_j \cos\theta_{j-1}}, \end{aligned} \tag{2.5}$$

For later convenience define matrices D_j:

$$\begin{aligned} D_j &= C_{j+1}C_{j+2}...C_{m+1} = \begin{pmatrix} a_j & b_j \\ c_j & d_j \end{pmatrix}, \\ D_{m+1} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\ D_0 &= C_1C_2...C_{m+1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \end{aligned}$$

In the air (on the back side) there is only the outgoing wave, therefore $E^-_{m+1} = 0$, which yields the expression for the total reflection coefficient to be:

$$R = |r|^2 = \left|\frac{E^-_0}{E^+_0}\right|^2 = \left|\frac{c}{a}\right|^2. \tag{2.6}$$

Measuring Alumina Thickness via SPR

Treating the aluminum film thickness, the aluminum oxide thickness, and the index of refraction of the aluminum oxide as fit parameters, equation 2.6 can be used to precisely measure the optical properties and the thickness of the oxide layer. The measurement was carried out for two wavelengths, and for a range of aluminum film thicknesses. The results are summarized in figure 2.9. Each measurement was repeated five times for statistical error analysis, with the standard deviation given by:

$$s = \sqrt{\frac{1}{N}\sum\left([x_i - x_{bg}] - \bar{x}\right)^2}.$$

Good agreement is seen between the theory and the experimental data, with minor deviation in the absolute reflectivity at the resonance point due to the finite roughness of the metal film. This demonstrates the power of using the SPR to precisely measure the properties of thin films.

The presented model treats the SPP wave as the interference between the reflected and refracted waves at the metal-dielectric boundary. This model is an accurate description of the surface plasmon coupling at optical frequencies from the UV to the NIR. The SPR measurements of the optical properties as well as the thickness of the of a dielectric are possible due to the localization of the fields close to the boundary, rendering the coupling condition strongly dependent on the exact composition of the dielectric layer at the boundary. The same phenomenon is used later in this work for tuning the SPP resonance in metal nanostructures.

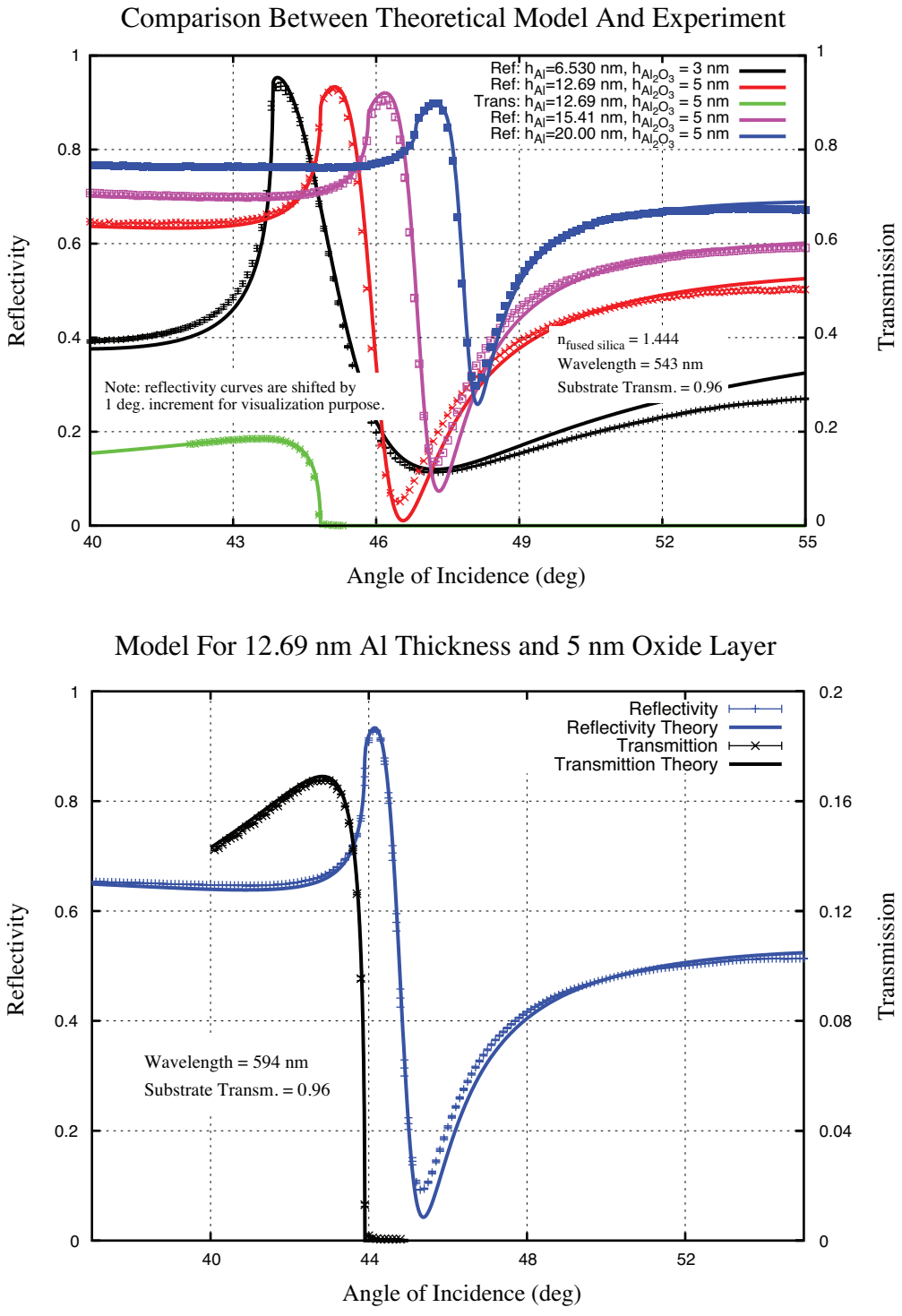

Figure 2.9: *Due to the localized nature of the SPP wave at the metal-air boundary, precise measurement of the optical properties of thin films is possible.*

Photoemission Enhancement in SPP Waves

3

3.1 S-Polarization

SPR provides the optimum environment for enhanced photoemission due to the increased absorption and strong field intensity localized close to the metal-vacuum interface. This field enhancement due to the excited SPP coupling is measured in reference to the plain metal surface. In s-polarization no SPP modes are excited, and in this case, the response is identical to that of a plain surface. Consider the situation shown in figure 2.1: s-polarization is equivalent to swapping the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields. The same boundary conditions apply in such a case ($E_x = 0$, $E_z = 0$, $H_y = 0$, $\mu_1 \approx \mu_2$):

$$\begin{cases} E_x^{(1)} = E_x^{(2)}, & E_y^{(1)} = E_y^{(2)} \\ H_x^{(1)} = H_x^{(2)}, & H_y^{(1)} = H_y^{(2)} \\ \epsilon_1 E_z^{(1)} = \epsilon_2 E_z^{(2)}, & \mu_1 H_z^{(1)} = \mu_2 H_z^{(2)} \end{cases} \Rightarrow \begin{cases} E_y^{(1)} = E_y^{(2)} \\ H_x^{(1)} = H_x^{(2)} \\ H_z^{(1)} = H_z^{(2)} \end{cases}$$

.

Expressing the $\boldsymbol{H}$ and $\boldsymbol{E}$ fileds in terms of a monochromatic plane wave solution gives:

$$\begin{aligned} \boldsymbol{H} &= \boldsymbol{H} \cdot \exp\left[i\left(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t\right)\right], \\ \boldsymbol{E} &= \boldsymbol{E} \cdot \exp\left[i\left(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t\right)\right]. \end{aligned}$$

Applying Maxwell's equations as before yields:

$$
\begin{aligned}
\boldsymbol{\nabla}\cdot\boldsymbol{E} &= 0 \Rightarrow k_y = 0, \\
\boldsymbol{\nabla}\cdot\boldsymbol{H} &= 0 \Rightarrow H_x i k_x + H_z i k_z = 0, \\
\boldsymbol{\nabla}\times\boldsymbol{E} &= -\frac{\partial \overline{D}}{\partial t} \Rightarrow \begin{cases} \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} &= -\mu \frac{\partial H_x}{\partial t} \\ \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} &= -\mu \frac{\partial H_y}{\partial t} \\ \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} &= -\mu \frac{\partial H_z}{\partial t} \end{cases}, \\
&\Rightarrow \begin{cases} -ik_z E_y &= \mu i\omega H_x \\ ik_x E_y &= \mu i\omega H_z \end{cases}, \\
&\Rightarrow \begin{cases} \frac{k_z}{\mu} E_y &= -\omega H_x \\ \frac{k_x}{\mu} E_y &= \omega H_z \end{cases}.
\end{aligned}
$$

The above equations have similar form for medium 1 and 2 with $k_z^{(2)} = -k_z^{(1)}$ for the propagation in the opposite direction. Applying the boundary conditions yields a system of linear equations:

$$
\begin{aligned}
\begin{cases} H_x^{(1)} - H_x^{(2)} &= 0 \\ E_y^{(1)} - E_y^{(2)} &= 0 \end{cases} &\Rightarrow \begin{cases} \frac{k_z^{(1)}}{\mu} E_y^{(1)} + \frac{k_z^{(2)}}{\mu} E_y^{(2)} &= 0 \\ E_y^{(1)} - E_y^{(2)} &= 0 \end{cases} \\
\begin{pmatrix} \frac{k_z^{(1)}}{\mu} & \frac{k_z^{(2)}}{\mu} \\ 1 & -1 \end{pmatrix} \begin{pmatrix} E_y^{(1)} \\ E_y^{(2)} \end{pmatrix} = 0 &\Rightarrow \begin{vmatrix} \frac{k_z^{(1)}}{\mu} & \frac{k_z^{(2)}}{\mu} \\ 1 & -1 \end{vmatrix} = 0, \\
k_z^{(1)} + k_z^{(2)} &= 0.
\end{aligned}
$$

Plugging in the above result into equation 2.1 yields:

$$
k_x^2 + \left[k_z^{(1,2)}\right]^2 = \frac{\epsilon_{1,2}}{\epsilon_0}\left(\frac{\omega}{c}\right)^2 \Rightarrow \begin{cases} k_x^2 &+ \left[k_z^{(1)}\right]^2 = \frac{\epsilon_1}{\epsilon_0}\left(\frac{\omega}{c}\right)^2 \\ k_x^2 &+ \left[-k_z^{(1)}\right]^2 = \frac{\epsilon_2}{\epsilon_0}\left(\frac{\omega}{c}\right)^2 \end{cases} \Rightarrow \epsilon_1 = \epsilon_2.
$$

The above result is invalid when $\epsilon_1 \neq \epsilon_2$, which is the case for the metal-dielectric interface. Therefore, in s-polarization no SPP wave is excited and the response is that of a plain metal surface. This way, the s-polarization response can be used as a reference for measuring the field enhancement in the SPP-coupled systems.

3.2 Field Intensity Profile

For the following discussion the p-polarization is assumed, in this case the SPP wave is excited in the film (deposited onto a glass substrate) at the metal-vacuum interface. In a

bulk metal, the field exponentially decays into the metal at the characteristic length, called the metal skin depth. In a thin film, however, the situation is more complicated. From equations 2.4, the field amplitude at the j^{th} interface is given by:

$$\begin{aligned} E_j^+ &= t_1 t_2 ... t_j \cdot \frac{a_j}{a} E_0^+, \\ E_j^- &= t_1 t_2 ... t_j \cdot \frac{c_j}{a} E_0^+. \end{aligned}$$

Expressing the amplitude $E(z)$ as a superposition of the forward and backward propagating waves gives:

$$\begin{aligned} E^+(z) &= E_j^+ \exp\left[iK_j \Delta z\right], \\ E^-(z) &= E_j^- \exp\left[-iK_j \Delta z\right], \end{aligned}$$

where Δz is the distance into the j^{th} layer and the K_j is the z-component of the wave vector and is given by

$$K_j = \frac{2\pi}{\lambda_0} n_j \cos\theta_j = \frac{\delta_j}{h_j}.$$

Combining the above results yields the normalized field intensity $F(z)$:

$$\begin{aligned} E_x(z) &= \left[E^+(z) + E^-(z)\right] \cdot \cos\theta_j \\ E_z(z) &= \left[E^+(z) - E^-(z)\right] \cdot \sin\theta_j \\ F_x(z) &= \frac{|E_x(z)|^2}{\left|E_0^+\right|^2} \\ F_z(z) &= \frac{|E_z(z)|^2}{\left|E_0^+\right|^2} \\ F(z) &= F_x(z) + F_z(z) \\ &= \frac{1}{\left|E_{op}^+\right|^2} \left(|E_x(z)|^2 + |E_z(z)|^2\right) \end{aligned}$$

From Snell's law $\sin\theta_j = \sin\theta_0 (n_0/n_j)$, which yields:

$$
\begin{aligned}
F(z) &= \frac{1}{|E_{op}^{+}|^2}[|(E_p^{+}(z)+E_p^{-}(z))\cdot\frac{1}{n_j}\sqrt{n_j^2-n_0^2\sin^2\theta_0}|^2+|((E_p^{+}(z)-E_p^{-}(z))\cdot\frac{n_0}{n_j}\sin\theta_0|^2] \\
F_j(z) &= \frac{1}{\left|E_{op}^{+}\right|^2}\left[\left|(E_j^{+}e^{i\Delta z\frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0}}+E_j^{-}e^{-i\Delta z\frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0}})\cdot\frac{1}{n_j}\sqrt{n_j^2-n_0^2\sin^2\theta_0}\right|^2\right. \\
&\quad \left.+\left|(E_j^{+}e^{i\Delta z\frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0}}-E_j^{-}e^{-i\Delta z\frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0}})\cdot\frac{n_0}{n_j}\sin\theta_0\right|^2\right] \\
&= \frac{1}{|E_{op}^{+}|^2}\left[|(t_1t_2...t_j\frac{a_j}{a}E_0^{+}e^{i\Delta z\frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0}}\right. \\
&\quad +t_1t_2...t_j\frac{c_j}{a}E_0^{+}e^{-i\Delta z\frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0}})\cdot\frac{1}{n_j}\sqrt{n_j^2-n_0^2\sin^2\theta_0}|^2 \\
&\quad \left.+|(t_1t_2...t_j\frac{a_j}{a}E_0^{+}e^{i\Delta z\frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0}}-t_1t_2...t_j\frac{c_j}{a}E_0^{+}e^{-i\Delta z\frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0}})\cdot\frac{n_0}{n_j}\sin\theta_0|^2\right] \\
&= \left|\frac{t_1t_2...t_j}{a}\right|^2\left[\left|\left(a_je^{i\Delta z\frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0}}+c_je^{-i\Delta z\frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0}}\right)\cdot\frac{1}{n_j}\sqrt{n_j^2-n_0^2\sin^2\theta_0}\right|^2+\right. \\
&\quad \left.+\left|\left(a_je^{i\Delta z\frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0}}-c_je^{-i\Delta z\frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0}}\right)\cdot\frac{n_0}{n_j}\sin\theta_0\right|^2\right]
\end{aligned}
$$

There are three variables expressed in terms of the C- and D-matrices (see equation 2.4): $a = D_{11}^0$, $a_j = D_{11}^j$, $c_j = D_{21}^j$. It is convenient to convert the square-root expressions to expressions involving K_j:

$$K_j = \frac{2\pi}{\lambda}n_j\cos\theta_j = \frac{2\pi}{\lambda}n_j\cdot\frac{1}{n_j}\sqrt{n_j^2-n_0^2\sin^2\theta_0} = \frac{2\pi}{\lambda}\sqrt{n_j^2-n_0^2\sin^2\theta_0},$$

$$\frac{1}{n_j}\sqrt{n_j^2-n_0^2\sin^2\theta_0} = K_j\cdot\frac{\lambda}{2\pi n_j}.$$

Define roi to correspond to the metal layer index in the multi-layer stack: see figure 2.8; then $t_{roi}^j \equiv t_1t_2...t_j$, and plugging the above result into the expression for $F(z)$ yields:

$$
\begin{aligned}
F_j(z) &= \left|\frac{t_{roi}^j}{D_{11}^0}\right|^2\left\{\left|\left(D_{11}^je^{iK_j\Delta z}+D_{21}^je^{-iK_j\Delta z}\right)\cdot\frac{\lambda}{2\pi}\frac{K_j}{n_j}\right|^2+\right. \\
&\quad \left.\left|\left(D_{11}^je^{iK_j\Delta z}-D_{21}^je^{-iK_j\Delta z}\right)\cdot\frac{n_0}{n_j}\sin\theta_0\right|^2\right\}
\end{aligned}
$$

In the above expression, Δz refers to the distance into the layer roi. Making a replacement $\Delta z \to z$ yields:

$$
\begin{aligned}
F_j(z)\cdot\left(\frac{1}{\left|\frac{t^j_{roi}}{D^0_{11}}\right|^2}\right) &= \left|\frac{\lambda}{2\pi}\cdot\frac{K_j}{n_j}\right|^2\left(a_j^* e^{-iK_j^* z}+c_j^* e^{iK_j^* z}\right)\left(a_j e^{iK_j z}+c_j e^{-iK_j z}\right)\\
&\quad+\left|\frac{n_0}{n_j}\sin\theta_0\right|^2\left(a_j^* e^{-iK_j^* z}-c_j^* e^{iK_j^* z}\right)\left(a_j e^{iK_j z}-c_j e^{-iK_j z}\right)\\
&= \left|\frac{\lambda}{2\pi}\cdot\frac{K_j}{n_j}\right|^2\cdot\left(|a_j|^2\cdot\exp\left[-i\left(K_J^*-K_J\right)z\right]+|c_j|^2\cdot\exp\left[i\left(K_j^*-K_j\right)z\right]\right.\\
&\quad\left.+\,a_j^* c_j\cdot\exp\left[-i\left(K_j^*+K_j\right)z\right]+c_j^* a_j\cdot\exp\left[i\left(K_j^*+K_j\right)z\right]\right)\\
&+ \left|\frac{n_0}{n_j}\sin\theta_0\right|^2\cdot\left(|a_j|^2\cdot\exp\left[-i\left(K_J^*-K_J\right)z\right]+|c_j|^2\cdot\exp\left[i\left(K_J^*-K_J\right)z\right]\right.\\
&\quad\left.-\,a_j^* c_j\cdot\exp\left[-i\left(K_j^*+K_j\right)z\right]-c_j^* a_j\cdot\exp\left[i\left(K_j^*+K_j\right)z\right]\right)
\end{aligned}
$$

The above expression is simplified by noting that

$$\left(a_j^* c_j\cdot\exp\left[-i\left(K_j^*+K_j\right)z\right]\right)^* = c_j^* a_j\cdot\exp\left[i\left(K_j^*+K_j\right)z\right].$$

Any complex number represented in the form $z=a+ib$ has a complex conjugate $z^*=a-ib$, which gives $z+z^*=2a=2Re\left[z\right]$. Applying this to the above expression gives:

$$
\begin{aligned}
a_j^* c_j\cdot\exp\left[-i\left(K_j^*+K_j\right)z\right] &+ c_j^* a_j\cdot\exp\left[i\left(K_j^*+K_j\right)z\right]\\
&= a_j^* c_j\cdot\exp\left[-i\left(K_j^*+K_j\right)z\right]+\left(a_j^* c_j\cdot\exp\left[-i\left(K_j^*+K_j\right)z\right]\right)^*,\\
&= 2\cdot Re\left(a_j^* c_j\cdot\exp\left[-i\left(K_j^*+K_j\right)z\right]\right),\\
&= 2\cdot Re\left(a_j^* c_j\cdot\exp\left[-2i\cdot Re\left(K_j\right)\cdot z\right]\right).
\end{aligned}
$$

Similarly, note that $z^*-z=-2i\cdot Im\left[z\right]\Rightarrow K_j^*-K_j=-2i\cdot Im\left(K_j\right)$. Putting the above results together yields:

$$\begin{aligned}
F_j(z)\cdot\left(\frac{1}{\gamma}\right) &= \alpha\cdot\left\{|a_j|^2\cdot\exp\left[-i\left(-2i\cdot Im\left[K_j\right]\right)z\right]+|c_j|^2\cdot\exp\left[i\left(-2i\cdot Im\left[K_j\right]\right)z\right]\right.\\
&\quad \left.+2\cdot Re\left(a_j^*c_j\cdot\exp\left[-2i\cdot Re\left(K_j\right)\cdot z\right]\right)\right\}\\
&+ \beta\cdot\left\{|a_j|^2\cdot\exp\left[-i\left(-2i\cdot Im\left[K_j\right]\right)z\right]+|c_j|^2\cdot\exp\left[i\left(-2i\cdot Im\left[K_j\right]\right)z\right]\right.\\
&\quad \left.-\left[2\cdot Re\left(a_j^*c_j\cdot\exp\left[-2i\cdot Re\left(K_j\right)\cdot z\right]\right)\right]\right\}\\
&= \alpha\cdot\left\{|a_j|^2\cdot\exp\left[-2\cdot Im\left(K_j\right)z\right]+|c_j|^2\cdot\exp\left[2\cdot Im\left(K_j\right)z\right]\right.\\
&\quad \left.+2\cdot Re\left(a_j^*c_j\cdot\exp\left[-2i\cdot Re\left(K_j\right)\cdot z\right]\right)\right\}\\
&+ \beta\cdot\left\{|a_j|^2\cdot\exp\left[-2\cdot Im\left(K_j\right)z\right]+|c_j|^2\cdot\exp\left[2\cdot Im\left(K_j\right)z\right]\right.\\
&\quad \left.-\left[2\cdot Re\left(a_j^*c_j\cdot\exp\left[-2i\cdot Re\left(K_j\right)\cdot z\right]\right)\right]\right\}\\
&= (\alpha+\beta)\cdot\left\{|a_j|^2\cdot\exp\left[-2\cdot Im\left(K_j\right)z\right]+|c_j|^2\cdot\exp\left[2\cdot Im\left(K_j\right)z\right]\right\}\\
&\quad +(\alpha-\beta)\cdot\left\{2\cdot Re\left(a_j^*c_j\cdot\exp\left[-2i\cdot Re\left(K_j\right)\cdot z\right]\right)\right\}.
\end{aligned}$$

The above formula is convenient to use in the computer code, but for an analytical expression a further simplification can be introduced:

$$\begin{aligned}
F_j(z)\cdot\left(\frac{1}{\gamma}\right) &= (\alpha+\beta)\cdot\left\{|a_j|^2\cdot\exp\left[-2\cdot Im\left(K_j\right)z\right]+|c_j|^2\cdot\exp\left[2\cdot Im\left(K_j\right)z\right]\right\}\\
&\quad +(\alpha-\beta)\cdot\left\{2\cdot Re\left(a_j^*c_j\cdot\left\{\cos\left[2\cdot Re\left(K_j\right)\cdot z\right]-i\cdot\sin\left[2Re\left(K_j\right)\cdot z\right]\right\}\right)\right\}\\
&= (\alpha+\beta)\cdot\left\{|a_j|^2\cdot\exp\left[-2\cdot Im\left(K_j\right)z\right]+|c_j|^2\cdot\exp\left[2\cdot Im\left(K_j\right)z\right]\right\}\\
&\quad +2(\alpha-\beta)\cdot Re\left[\left\{Re\left[a_j^*c_j\right]+i\cdot Im\left[a_j^*c_j\right]\right\}\cdot\right.\\
&\quad \left.\left\{\cos\left[2\cdot Re\left(K_j\right)\cdot z\right]-i\cdot\sin\left[2Re\left(K_j\right)\cdot z\right]\right\}\right]\\
&= (\alpha+\beta)\cdot\left\{|a_j|^2\cdot\exp\left[-2\cdot Im\left(K_j\right)z\right]+|c_j|^2\cdot\exp\left[2\cdot Im\left(K_j\right)z\right]\right\}\\
&\quad +2(\alpha-\beta)\cdot\left\{Re\left[a_j^*c_j\right]\cos\left[2\cdot Re\left(K_j\right)\cdot z\right]+Im\left[a_j^*c_j\right]\sin\left[2Re\left(K_j\right)\cdot z\right]\right\},
\end{aligned}$$

where the constants are defined as

$$\begin{aligned}
\alpha &= \left|\sqrt{1-\left(\tfrac{n_0}{n_j}\sin\theta_0\right)^2}\right|^2,\\
\beta &= \left|\tfrac{n_0}{n_j}\sin\theta_0\right|^2,\\
\gamma &= \left|\tfrac{t^j_{roi}}{D^0_{11}}\right|^2.
\end{aligned}$$

Further simplifying the above expression yields:

$$\begin{aligned}
F_j(z) = \;& \gamma\cdot\left[(\alpha+\beta)\cdot\left\{|a_j|^2\cdot\exp\left[-2\cdot Im\left(K_j\right)z\right]+|c_j|^2\cdot\exp\left[2\cdot Im\left(K_j\right)z\right]\right\}\right.\\
& \left.+2(\alpha-\beta)\left\{Re\left[a_j^*c_j\right]\cos\left[2\cdot Re\left(K_j\right)\cdot z\right]+Im\left[a_j^*c_j\right]\sin\left[2Re\left(K_j\right)\cdot z\right]\right\}\right]
\end{aligned}\tag{3.1}$$

3.3 SPP Coupling

The above equation for the field intensity has two components: one for the exponential decay inside the material, and another for the exponential growth. Consider the specific case of the aluminum film on a glass substrate: the region of interest is the aluminum film of thickness d, therefore $z \in (0, d)$, and $D^1 = C_2$:

$$D^1 = C_2 = \begin{pmatrix} e^{-i\delta_1} & r_2 e^{-i\delta_1} \\ r_2 e^{i\delta_1} & e^{i\delta_1} \end{pmatrix} \Rightarrow D^1_{21} = r_2 e^{i\delta_1} = \frac{n_1 \cos\theta_2 - n_2 \cos\theta_1}{n_1 \cos\theta_2 + n_2 \cos\theta_1} e^{i\frac{2\pi}{\lambda} n_1 \cos\theta_1 d}.$$

Expanding the terms $|c_1|^2 = |D^1_{21}|^2$ and $|a_1|^2 = |D^1_{11}|^2$ yield:

$$\begin{aligned}
|c_1|^2 &= \left|\tfrac{n_1\cos\theta_2 - n_2\cos\theta_1}{n_1\cos\theta_2 + n_2\cos\theta_1}\right|^2 \cdot \exp\left[-i\tfrac{2\pi}{\lambda} n_1^* (\cos\theta_1)^* d\right] \cdot \exp\left[i\tfrac{2\pi}{\lambda} n_1 \cos\theta_1 d\right] \\
&= \left|\tfrac{n_1\cos\theta_2 - n_2\cos\theta_1}{n_1\cos\theta_2 + n_2\cos\theta_1}\right|^2 \cdot \exp\left[-i\tfrac{2\pi d}{\lambda} \left\{(n_1\cos\theta_1)^* - (n_1\cos\theta_1)\right\}\right] \\
&= \left|\tfrac{n_1\cos\theta_2 - n_2\cos\theta_1}{n_1\cos\theta_2 + n_2\cos\theta_1}\right|^2 \cdot \exp\left[-i\tfrac{2\pi d}{\lambda} \left\{\left(\sqrt{n_1^2 - n_0^2\sin^2\theta_0}\right)^* - \sqrt{n_1^2 - n_0^2\sin^2\theta_0}\right\}\right] \\
&= \left|\tfrac{n_1\cos\theta_2 - n_2\cos\theta_1}{n_1\cos\theta_2 + n_2\cos\theta_1}\right|^2 \cdot \exp\left[-i\tfrac{2\pi d}{\lambda} \left\{-2i \cdot Im\left(\sqrt{n_1^2 - n_0^2\sin^2\theta_0}\right)\right\}\right] \\
&= \left|\tfrac{n_1\cos\theta_2 - n_2\cos\theta_1}{n_1\cos\theta_2 + n_2\cos\theta_1}\right|^2 \cdot \exp\left[-\tfrac{4\pi d}{\lambda} \cdot Im\left(\sqrt{n_1^2 - n_0^2\sin^2\theta_0}\right)\right] \qquad (3.2)
\end{aligned}$$

$$\begin{aligned}
|a_1|^2 &= \exp\left[i\tfrac{2\pi}{\lambda} n_1^* (\cos\theta_1)^* d\right] \cdot \exp\left[-i\tfrac{2\pi}{\lambda} n_1 \cos\theta_1 d\right] \\
&= \exp\left[i\tfrac{2\pi d}{\lambda} \left\{(n_1\cos\theta_1)^* - (n_1\cos\theta_1)\right\}\right] \\
&= \exp\left[i\tfrac{2\pi d}{\lambda} \left\{-2i \cdot Im\left(\sqrt{n_1^2 - n_0^2\sin^2\theta_0}\right)\right\}\right] \\
&= \exp\left[\tfrac{4\pi d}{\lambda} Im\left(\sqrt{n_1^2 - n_0^2\sin^2\theta_0}\right)\right]
\end{aligned}$$

Similarly for $a = D^0_{11}$:

$$D_0 = C_1 C_2 = \begin{pmatrix} 1 & r_1 \\ r_1 & 1 \end{pmatrix} \begin{pmatrix} e^{-i\delta_1} & r_2 e^{-i\delta_1} \\ r_2 e^{i\delta_1} & e^{i\delta_1} \end{pmatrix} = \begin{pmatrix} e^{-i\delta_1} + r_1 r_2 e^{i\delta_1} & r_2 e^{-i\delta_1} + r_1 e^{i\delta_1} \\ r_1 e^{-i\delta_1} + r_2 e^{i\delta_1} & r_1 r_2 e^{-i\delta_1} + e^{i\delta_1} \end{pmatrix}.$$

$$
\begin{aligned}
\Rightarrow |a|^2 &= \left|e^{-i\delta_1} + r_1 r_2 e^{i\delta_1}\right|^2 \\
&= \left(\exp\left[i\delta_1^*\right] + \left(r_1 r_2\right)^* \exp\left[-i\delta_1^*\right]\right)\left(\exp\left[-i\delta_1\right] + \left(r_1 r_2\right)\exp\left[i\delta_1\right]\right) \\
&= \exp\left[i\delta_1^*\right]\exp\left[-i\delta_1\right] + \left(r_1 r_2\right)^*\exp\left[-i\delta_1^*\right]\left(r_1 r_2\right)\exp\left[i\delta_1\right] \\
&\quad + \left(r_1 r_2\right)^*\exp\left[-i\delta_1^*\right]\exp\left[-i\delta_1\right] + \left(r_1 r_2\right)\exp\left[i\delta_1\right]\exp\left[i\delta_1^*\right] \\
&= \left(1 + \left|r_1 r_2\right|^2\right)\exp\left[i\left(\delta_1^* - \delta_1\right)\right] + \left(r_1 r_2\right)^*\exp\left[-i\left(\delta_1^* + \delta_1\right)\right] + \left(r_1 r_2\right)\exp\left[i\left(\delta_1^* + \delta_1\right)\right] \\
&= \left(1 + \left|r_1 r_2\right|^2\right)\exp\left[2\cdot Im\left(\delta_1\right)\right] + \left(r_1 r_2\right)^*\exp\left[-2i\cdot Re\left(\delta_1\right)\right] + \left(r_1 r_2\right)\exp\left[2i\cdot Re\left(\delta_1\right)\right] \\
&= \left(1 + \left|r_1 r_2\right|^2\right)\exp\left[2\cdot Im\left(\delta_1\right)\right] \\
&\quad + \left\{\left(r_1 r_2\right)\exp\left[2i\cdot Re\left(\delta_1\right)\right]\right\}^* + \left\{\left(r_1 r_2\right)\exp\left[2i\cdot Re\left(\delta_1\right)\right]\right\} \\
&= \left(1 + \left|r_1 r_2\right|^2\right)\exp\left[2\cdot Im\left(\delta_1\right)\right] + 2\cdot Re\left\{\left(r_1 r_2\right)\exp\left[2i\cdot Re\left(\delta_1\right)\right]\right\} \\
&= \left(1 + \left|r_1 r_2\right|^2\right)\exp\left[2\cdot Im\left(\delta_1\right)\right] \\
&\quad + 2\cdot Re\left\{\left(Re\left[r_1 r_2\right] + i\cdot Im\left[r_1 r_2\right]\right)\left(\cos\left[2Re\left(\delta_1\right)\right] + i\cdot\sin\left[2Re\left(\delta_1\right)\right]\right)\right\} \\
&= \left(1 + \left|r_1 r_2\right|^2\right)\exp\left[2\cdot Im\left(\delta_1\right)\right] \\
&\quad + 2\cdot\left\{Re\left[r_1 r_2\right]\cos\left[2Re\left(\delta_1\right)\right] - Im\left[r_1 r_2\right]\sin\left[2Re\left(\delta_1\right)\right]\right\}
\end{aligned}
$$

From the definition $\delta_1 = \frac{2\pi d}{\lambda}\sqrt{n_{Al} - n_{gl}\sin^2\theta_0}$ it follows that $Im\left[\delta_1\right] \propto d$. For large d the contribution from the sin and cos become negligible; therefore, $|a|^2 \propto e^d$. Similarly, from $\gamma \propto \frac{1}{|a|^2}$ it follows that $\lim_{d\to\infty}(\gamma) = 0$. In general, $|c_j|^2 \propto e^{-d}$, $|a_j|^2 \propto e^d$, and $\gamma \propto e^{-d}$. Pulling out the exponentially increasing and decreasing parts from equation 3.1 yields the prefactors $\gamma|a_j|^2 \propto const$ and $\gamma|c_j|^2 \propto e^{-2d}$, therefore an increase in the metal thickness decreases the amplitude of the exponentially increasing component of the field intensity, until it vanishes completely for large d. The exponentially growing contribution is $\propto \exp\left[C_1 z - C_2 d\right]$ (C_1 and C_2 are numerical factors) with the maximum value at position $z = d$. This result confirms that the SPP modes is excited at the metal-vacuum interface, not the metal-prism interface. The field intensity strength $F(z)$ is a function of the thickness of the aluminum film. By adjusting the thickness of the metal layer the field strength can be maximized.

The exponentially increasing and decreasing components from equation 3.1, $F_1(z) \propto f_{dec}(z) + f_{inc}(z)$, can be rewritten for the glass-aluminum-vacuum system (as in figure 2.8) by plugging in the expressions for $|a_1|^2$ and $|c_1|^2$ from the equation 3.2. Note that the oscillatory parts in the equation 3.1 have the period of oscillation comparable to $\lambda \gg d$, therefore the oscillatory contribution can be ignored. The exponentially increasing and decreasing components are given by (note $K_1 = \frac{2\pi}{\lambda}\sqrt{n_1^2 - n_0^2\sin^2\theta_0}$):

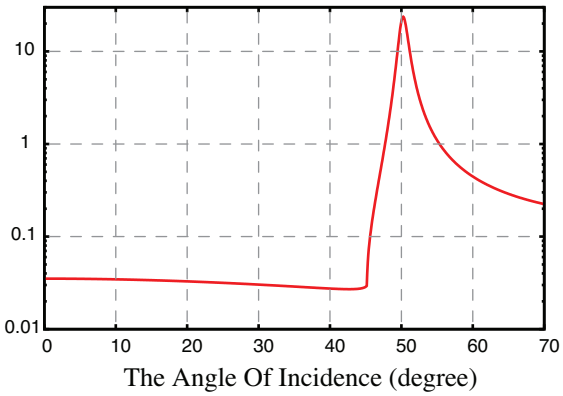

Figure 3.1: *The ratio between the exponentially increasing and decreasing components of the field intensity profile inside the metal film. The sharp increase at* $\approx 50^o$ *corresponds to the resonance condition due to the SPP coupling.*

$$
\begin{aligned}
f_{dec} &= \exp\left[\frac{4\pi d}{\lambda} Im\left(\sqrt{n_1^2 - n_0^2 \sin^2\theta_0}\right)\right] \cdot \exp\left[-\frac{4\pi z}{\lambda} \cdot Im\left(\sqrt{n_1^2 - n_0^2 \sin^2\theta_0}\right)\right] \\
&= \exp\left[\frac{4\pi\,(d-z)}{\lambda} Im\left(\sqrt{n_1^2 - n_0^2 \sin^2\theta_0}\right)\right], \\
f_{inc} &= \left|\frac{n_1\cos\theta_2 - n_2\cos\theta_1}{n_1\cos\theta_2 + n_2\cos\theta_1}\right|^2 \times \exp\left[-\frac{4\pi d}{\lambda} Im\left(\sqrt{n_1^2 - n_0^2 \sin^2\theta_0}\right)\right] \\
&\quad \cdot \exp\left[\frac{4\pi z}{\lambda} Im\left(\sqrt{n_1^2 - n_0^2 \sin^2\theta_0}\right)\right] \\
&= \left|\frac{n_1\cos\theta_2 - n_2\cos\theta_1}{n_1\cos\theta_2 + n_2\cos\theta_1}\right|^2 \times \exp\left[-\frac{4\pi\,(d-z)}{\lambda} Im\left(\sqrt{n_1^2 - n_0^2 \sin^2\theta_0}\right)\right],
\end{aligned}
$$

Off resonance, the f_{dec} component dominates and the field intensity profile in the metal film is similar to that in the bulk metal. However, on resonance, the f_{inc} component dominates, and the field intensity profile has the exponentially growing shape, exactly opposite to that of the bulk material. Taking the ratio between the two components yields:

$$
A_{ratio} = \frac{\left|\frac{n_1\cos\theta_2 - n_2\cos\theta_1}{n_1\cos\theta_2 + n_2\cos\theta_1}\right|^2}{\exp\left[\frac{4\pi d}{\lambda} Im\left(\sqrt{n_1^2 - n_0^2 \sin^2\theta_0}\right)\right]}.
$$

When the above ratio is greater than unity the exponential increase dominates, otherwise the field intensity has the exponential decay profile. The graph of the magnitude of this

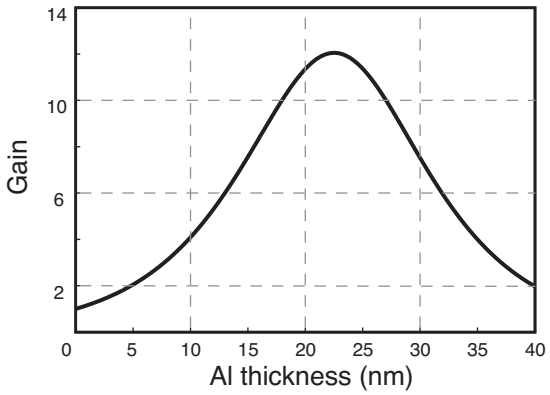

Figure 3.2: *The field intensity gain in the aluminum film was numerically calculated for the glass-aluminum-vacuum system. For each thickness, the angle of incidence was adjusted to ensure the SPP coupling condition is met. The peak occurs for the aluminum film thickness $d = 22.52$ nm.*

ratio as a function of the angle of incidence is shown in figure 3.1. The ratio becomes larger than unity for $\theta_0 \in (47.7^o, 55.3^o)$, with the maximum value of 23.7 at $\theta_0 = 50.3^o$, which corresponds to optimum coupling between the incident radiation and the SPP wave at the aluminum-vacuum interface.

As described above, the field intensity at the aluminum-vacuum interface is a function of the angle of incidence as well as the aluminum film thickness. The angle of incidence is chosen to satisfy the SPP coupling condition for a given thickness. To find the optimum layer thickness that yields the largest increase, it is useful to define the field gain in the aluminum layer as follows:

$$\begin{aligned} g \equiv & \quad \frac{F(h_j)}{F(0)} \\ \approx & \quad \frac{|a_j|^2 \cdot \exp[-2 \cdot Im(K_1) h_j] + |c_j|^2 \cdot \exp[2 \cdot Im(K_1) h_j]}{|a_j|^2 + |c_j|^2} \\ \approx & \quad \frac{\left(1 + |r_1 r_2|^2\right) \cdot \exp[2 \cdot Im(\delta_1)] \cdot \exp[-2 \cdot Im(\delta_1)] + |r_2|^2 \cdot \exp[-2 \cdot Im(\delta_1)] \cdot \exp[2 \cdot Im(\delta_1)]}{|a_j|^2 + |c_j|^2} \\ \approx & \quad \frac{1 + |r_1 r_2|^2 + |r_2|^2}{\left(1 + |r_1 r_2|^2\right) \cdot \exp[2 \cdot Im(\delta_1)] + |r_2|^2 \cdot \exp[-2 \cdot Im(\delta_1)]} \end{aligned}$$

The gain curve for the glass-aluminum model is shown in figure 3.2. The maximum gain in the field intensity is obtained for an aluminum thickness of 22.5 nm. For the case study of the glass-aluminum-vacuum system[1], the field intensity is increased by a factor of 12. The optical properties of this system are summarized in figure 3.3, where the reflection and transmission are plotted in (a), and the field profile inside the metal film is shown in (b). Note that off-resonance, at 0 and 45 degrees, the field profile is similar in shape to that in the bulk material. On-resonance, however, field profile flips in orientation and the exponential growth is observed, corresponding to the SPP wave excited at the metal-vacuum interface.

[1]Glass refers to the magnesium fluoride prism.

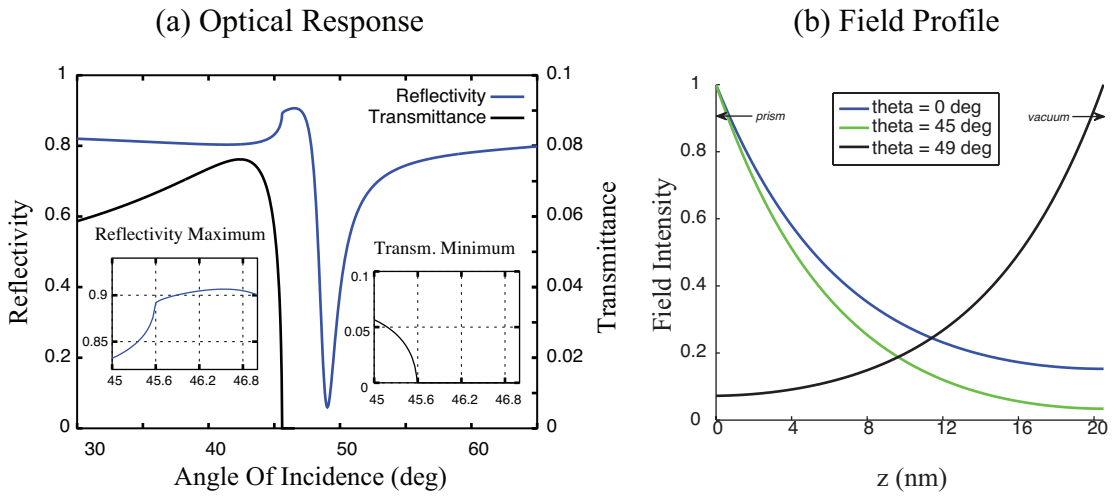

Figure 3.3: *The transmittance coefficient and the field intensity profile. The graph is plotted for the system* $MgF_2\,(n_0 = 1.40)\,|\,Al\,(n_1 = 0.22 + i3.16)$*, illuminated by* $265\,nm$ *light; the aluminum coating is* $23.5\,nm$ *thick. The critical angle for the total internal reflection from the system is around* 45.6 *incidence angle at the* $MgF_2 - Al$ *interface. Up to this angle the reflectance increases corresponding to the drop in transmittance. Right at the critical angle, the transmission abruptly stops corresponding to a discontinuity in the reflectance curve.*

3.4 The Photoemission Three Step Model

Step One: Photon Absorption and Electron Excitation

Assuming that the number of photons $n\,(z)$ at a distance z within the metal film is proportional to the field intensity $F\,(z)$ [8], it must therefore have the same functional form:

$$n\,(z) = A \cdot \exp\left[-\frac{z}{\lambda_{opt}}\right] + B \cdot \exp\left[\frac{z}{\lambda_{opt}}\right] + C \cdot \cos\left[\frac{z}{\lambda_{osc}}\right] + D \cdot \sin\left[\frac{z}{\lambda_{osc}}\right],$$

where the constants are

$$\begin{array}{lll} A = \gamma\,(\alpha+\beta)\cdot|a_j|^2 & C = 2\gamma\,(\alpha-\beta)\,Re\left[a_j^* c_j\right] & \lambda_{opt} = \frac{1}{2\cdot Im[K_j]} \\ B = \gamma\,(\alpha+\beta)\cdot|c_j|^2\ , & D = 2\gamma\,(\alpha-\beta)\,Im\left[a_j^* c_j\right]\ , & \lambda_{osc} = \frac{1}{2\cdot Re[K_j]}\ . \end{array}$$

Substituting these definitions into the expression for $n\,(z)$ yields (d is the thickness of the aluminum film):

$$\begin{aligned}
\int_0^d n(z)\,dz &= \int_0^d \left[A\cdot e^{-\frac{z}{\lambda_{opt}}} + B\cdot e^{\frac{z}{\lambda_{opt}}} + C\cdot\cos\left(\frac{z}{\lambda_{osc}}\right) + D\cdot\sin\left(\frac{z}{\lambda_{osc}}\right)\right]dz \\
&= -\lambda_{opt}\left[A\cdot\left(e^{-\frac{d}{\lambda_{opt}}}-1\right) - B\cdot\left(e^{\frac{d}{\lambda_{opt}}}-1\right)\right] \\
&\quad +\lambda_{osc}\left[C\cdot\sin\left(\frac{d}{\lambda_{osc}}\right) - D\cdot\left\{\cos\left(\frac{d}{\lambda_{osc}}\right)-1\right\}\right] \\
&= \lambda_{opt}\left(B\cdot\exp\left[\frac{d}{\lambda_{opt}}\right] - A\cdot\exp\left[-\frac{d}{\lambda_{opt}}\right]\right) \\
&\quad +\lambda_{osc}\left(C\cdot\sin\left[\frac{d}{\lambda_{osc}}\right] - D\cdot\cos\left[\frac{d}{\lambda_{osc}}\right]\right) + \lambda_{osc}D + \lambda_{opt}(A-B)
\end{aligned}$$

At normal incidence $\theta_j = 0$, therefore $Im[K_j] = \frac{2\pi}{\lambda_0} Im[n_j\cos\theta_j] = 0$ and $\lambda_{opt} = \frac{\lambda_0}{4\pi Im[n_j]}$, which is the metal optical skin depth. The photon absorption probability is defined as:

$$\begin{aligned}
P(z) &\equiv \frac{n(z)}{\int_0^d n(z)dz} \\
&= \frac{A\cdot\exp\left[-\frac{z}{\lambda_{opt}}\right]+B\cdot\exp\left[\frac{z}{\lambda_{opt}}\right]+C\cdot\cos\left[\frac{z}{\lambda_{osc}}\right]+D\cdot\sin\left[\frac{z}{\lambda_{osc}}\right]}{\lambda_{opt}\left(B\cdot\exp\left[\frac{d}{\lambda_{opt}}\right]-A\cdot\exp\left[-\frac{d}{\lambda_{opt}}\right]\right)+\lambda_{osc}\left(C\cdot\sin\left[\frac{d}{\lambda_{osc}}\right]-D\cdot\cos\left[\frac{d}{\lambda_{osc}}\right]\right)+\lambda_{osc}D+\lambda_{opt}(A-B)}
\end{aligned} \tag{3.3}$$

Then the number of states at energy E_0 is given by $N(E_0)$ and the occupancy of those states is given by the Fermi-Dirac function $f_{FD}(E_0)$. The probability that an electron is promoted from the *occupied* state to an *unoccupied* state is then given by:

$$P_{Exc}(E) = N(E)\cdot f_{FD}(E)\times N(E+\hbar\omega)\cdot[1-f_{FD}(E+\hbar\omega)].$$

Step Two: Transport to the Surface

Assuming any electron-electron scattering is catastrophic implies that after the first collision, an electron does not have sufficient energy remaining to overcome the potential barrier to exit from the metal surface. This collision probability is an exponential function with the characteristic decay length corresponding to the electron mean free path λ_{e-e}. The probability that an electron excited a distance $(d-z)$ away from the metal-vacuum interface will reach the surface without scattering is given by:

$$F_{e-e} \equiv \int_0^d P(z)\cdot e^{-\frac{d-z}{\lambda_{e-e}}}dz.$$

Plugging the expression for $P(z)$ into the above expression yields:

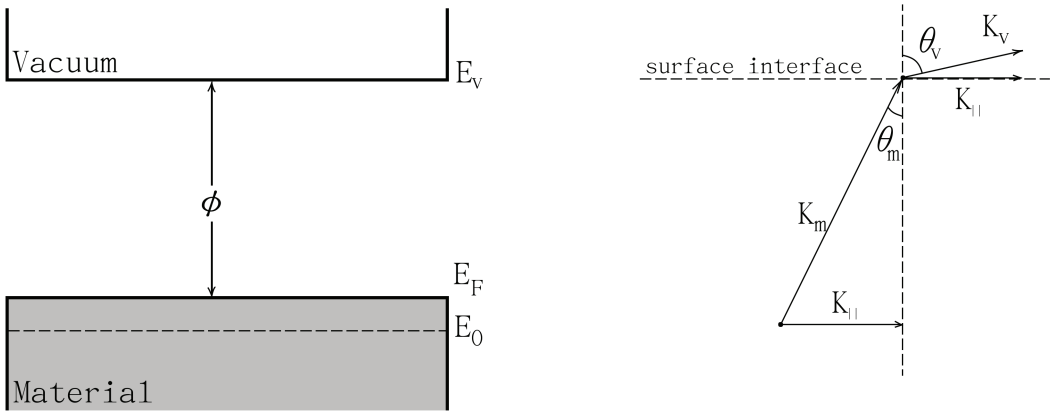

Figure 3.4: *Momentum conservation at the metal-vacuum interface determines which direction electrons must travel within the metal in order to have sufficient energy in the longitudinal direction to exit from the surface. These electron trajectories are limited to be within the cone as shown.*

$$F_{e-e} = \int_0^d \frac{\left(A\cdot\exp\left[-\frac{z}{\lambda_{opt}}\right]+B\cdot\exp\left[\frac{z}{\lambda_{opt}}\right]+C\cdot\cos\left[\frac{z}{\lambda_{osc}}\right]+D\cdot\sin\left[\frac{z}{\lambda_{osc}}\right]\right)\cdot\exp\left[-\frac{d-z}{\lambda_{e-e}}\right]\cdot dz}{\lambda_{opt}\left(B\cdot\exp\left[\frac{d}{\lambda_{opt}}\right]-A\cdot\exp\left[-\frac{d}{\lambda_{opt}}\right]\right)+\lambda_{osc}\left(C\cdot\sin\left[\frac{d}{\lambda_{osc}}\right]-D\cdot\cos\left[\frac{d}{\lambda_{osc}}\right]\right)+\lambda_{osc}D+\lambda_{opt}(A-B)},$$

where the electron-electron scattering length is given by [8]:

$$\overline{\lambda}_{e-e} = \frac{2\lambda_m E_m^{\frac{3}{2}}}{\hbar\omega\sqrt{\phi_{eff}}}\cdot\frac{1}{1+\sqrt{\frac{\phi_{eff}}{\hbar\omega}}},$$

with $\lambda_m = \lambda_{e-e}(E_m = 8.6\,eV) = 22\,\mathring{A}$. For aluminum $\phi_{eff} \approx 4.08\ eV$ at $\lambda = 239\ nm$, $\hbar\omega = 5.188\ eV \Rightarrow \bar{\lambda}_{e-e} = 56\,\mathring{A}$. The exponent in the electron-electron scattering probability has the form of $e^{-\frac{d-z}{\lambda_{e-e}}}$, where $(d-z)$ is the distance from the location of the photon absorption (photoelectron creation) to the metal-vacuum interface.

Step Three: Escape Probability

The parallel component of the wave vector of the electron is conserved on transmission through the interface, therefore, the momentum in the longitudinal direction must be sufficient to allow the electron to exit from the metal surface [47]. This condition is illustrated in figure 3.4, where E_0 – the initial energy of the electron in the conduction band, E_F – the Fermi energy, E_v – the vacuum energy level, ϕ – the work function of the metal, θ_m – the incidence angle inside the material, θ_v – the incidence angle in vacuum, K_v and K_m are the electron wave vectors inside the vacuum and metal respectively, and $K_{||}$ is the parallel component of the wave vector. On photon absorption, the electron picks up the photon energy to get $E_0 + \hbar\omega$. Once the electron passes over the barrier, its remaining energy is

$E_0 + \hbar\omega - E_F - \phi$, with the parallel component of K given by:

$$K_{||} = \frac{p_{||}}{\hbar} = \frac{\sin\theta_m}{\hbar}\sqrt{2m \cdot (E_0 + \hbar\omega)} = \frac{\sin\theta_v}{\hbar}\sqrt{2m \cdot (E_0 + \hbar\omega - E_F - \phi)}.$$

The maximum angle for θ_v is $\frac{\pi}{2}$, therefore all electron trajectories must fall within a cone with the opening angle θ_{max} given by:

$$\begin{aligned} \sin\theta_{max} &= \sqrt{\tfrac{E_0+\hbar\omega-E_F-\phi}{E_0+\hbar\omega}}, \\ \cos\theta_{max} &= \sqrt{\tfrac{E_0+\hbar\omega-E_0-\hbar\omega+E_F+\phi}{E_0+\hbar\omega}} \\ &= \sqrt{\tfrac{E_F+\phi}{E_0+\hbar\omega}}. \end{aligned} \tag{3.4}$$

For any angle $\theta_m > \theta_{max}$ the electron will not pass through the surface barrier. For the low energy emitted electrons θ_{max} is close to $\pi/2$. The electron escape probability is then given by:

$$P_{Esc} = \frac{1}{4\pi}\int\limits_{\cos(\theta_{max})}^{1} d(\cos\theta)\int\limits_{0}^{2\pi} d\Phi.$$

The $\frac{1}{4\pi}$ is the normalization prefactor. Note that the angle θ in the expression above is the spherical coordinate angle, *not* the angle of incidence.

3.5 Quantum Efficiency

The quantum efficiency is defined as the number of emitted electrons per incident photon. It is expressed in terms of the probabilities derived above as:

$$QE(\omega) = \chi_{ph} \times \frac{\int P_{Exc}F_{e-e}P_{Esc}dE}{\int P_{Exc}dE},$$

where $\chi_{ph} \equiv 1 - R(\omega,\theta) - T(\omega,\theta)$ is the number of absorbed photons, which is equal to the number of photoelectrons, assuming that each absorbed photon yields a photoelectron; P_{Exc} is the electron excitation probability; and P_{Esc} is the probability an electron will travel to the metal surface without scattering.

Normalization is done by integrating over all excited electrons from the minimum energy of $E_F - \hbar\omega$, to the maximum energy, the Fermi energy E_F. For an electron to escape, the minimum energy must be above $E_F - \hbar\omega + \phi$. Furthermore, for $\theta_{max} \approx \frac{\pi}{2}$ F_{e-e} is effectively independent of the electron's trajectory, and so F_{e-e} can be taken outside the integral. In a free-electron-like metals, for electrons emitted from a small energy range close to the Fermi level, a flat band electron density of states is a good approximation. In such case, the electron excitation probability is approximately constant [8] and also comes outside the integral (θ is the angle of incidence):

$$QE\left(\omega,\theta\right)=\left[1-R\left(\omega,\theta\right)-T\left(\omega,\theta\right)\right]\times\frac{F_{e-e}\left(\omega\right)\cdot P_{Exc}}{P_{Exc}}\times\frac{\int_{E_F-\hbar\omega+\phi}^{E_F}\left[\int_{\cos\theta_{max}}^{1}d\left(\cos\theta'\right)\int_{0}^{2\pi}d\Phi\right]\cdot dE}{4\pi\int_{E_F-\hbar\omega}^{E_F}dE}$$

Plugging in the expression for $\cos\left(\theta_{max}\right)$ from equation 3.4 yields (E_0 is the electron energy):

$$QE\left(\omega,\theta\right)\;=\;\left[1-R\left(\omega,\theta\right)-T\left(\omega,\theta\right)\right]\cdot F_{e-e}\times\frac{\int_{E_F-\hbar\omega+\phi}^{E_F}\int_{\sqrt{\frac{E_F+\phi}{E+\hbar\omega}}}^{1}\int_{0}^{2\pi}d\Phi\cdot d\left(\cos\theta'\right)\cdot dE}{2\int_{E_F-\hbar\omega}^{E_F}dE}$$

$$\begin{aligned}\frac{\int_{E_F-\hbar\omega+\phi}^{E_F}\int_{\sqrt{\frac{E_F+\phi}{E+\hbar\omega}}}^{1}\int_{0}^{2\pi}d\Phi\cdot d\left(\cos\theta'\right)\cdot dE}{\int_{E_F-\hbar\omega}^{E_F}\int_{-1}^{1}\int_{0}^{2\pi}d\Phi\cdot d\left(\cos\theta'\right)\cdot dE} &= \frac{\int_{E_F-\hbar\omega+\phi}^{E_F}\left[1-\sqrt{\frac{E_F+\phi}{E+\hbar\omega}}\right]dE}{2\int_{E_F-\hbar\omega}^{E_F}dE}\\ &= \left[1-\sqrt{\frac{E_F+\phi}{E_F+\hbar\omega}}\right]^2\\ &= \frac{\hbar\omega+2E_F+\phi-2\sqrt{\left(E_F+\hbar\omega\right)\cdot\left(E_F+\phi\right)}}{2\hbar\omega}\\ &= \frac{E_F+\hbar\omega}{2\hbar\omega}\left[1+\frac{E_F+\phi}{E_F+\hbar\omega}-2\sqrt{\frac{E_F+\phi}{E_F+\hbar\omega}}\right].\end{aligned}$$

The above expression can also be Taylor expanded around $\hbar\omega-\phi\approx 0$, which yields the quadratic dependance of the QE on the electron excess energy:

$$QE\left(\omega\right)\propto\left(\hbar\omega-\phi\right)^2.$$

In an experiment where the field is applied to the cathode to pull the emitted electrons away, the photocurrent increases as a function of the strength of the applied field, which is known as the Schottky effect [8]. Defining $\phi_{eff}\equiv\phi-\phi_{Schottky}$, where $\phi_{Schottky}=3.7947\cdot 10^{-5}\sqrt{E\left(V/m\right)}\,eV$ is the barrier energy shift due to the Schottky effect, yields the general expression for the quantum efficiency:

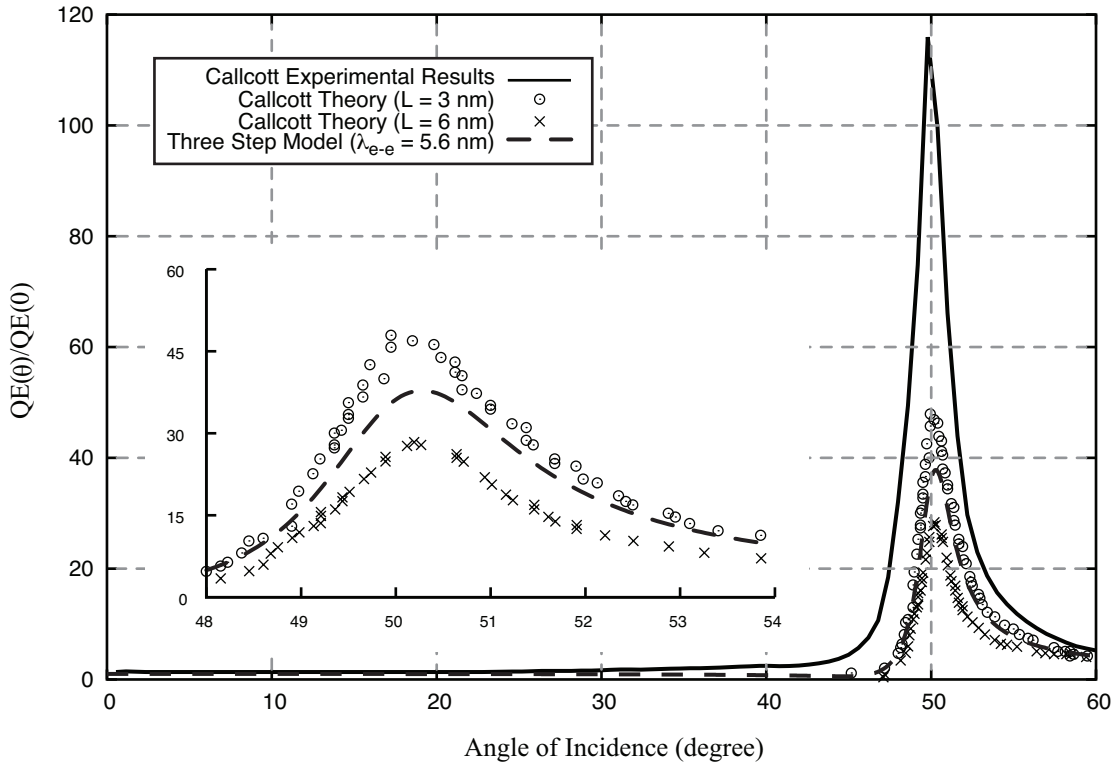

Figure 3.5: *Quantum efficiency of an aluminum film on a glass substrate. The experimental data is compared to the theoretical calculations by Callcott* et al. *as well as the three step model developed here shown by the the dashed curve. All three theoretical models are in close agreement, however, they all underestimate the experimental measurement of the photoemission enhancement by more than a factor of 2.*

$$\begin{aligned} QE\left(\omega,\theta\right) = \;& F_{e-e}\left(\omega,\theta\right)\cdot\left[1-R(\omega,\theta)-T\left(\omega,\theta\right)\right] \\ & \times\tfrac{E_F+\hbar\omega}{2\hbar\omega}\cdot\left[1+\tfrac{E_F+\phi_{eff}}{E_F+\hbar\omega}-2\sqrt{\tfrac{E_F+\phi_{eff}}{E_F+\hbar\omega}}\right]. \end{aligned} \tag{3.5}$$

The example of the glass-aluminum-vacuum has been experimentally investigated by Callcott *et al.* [19]. In that experiment, the aluminum thickness was chosen to be $d = 23.5$ nm to produce the maximum on-resonance field intensity gain. This is close to the optimum thickness found in figure 3.2. Then the sample was illuminated with a 266 nm laser for a range of angles and the photocurrent was measured in reference to the normal-incident light. At $\approx 50^o$ angle of incidence, the resonance condition was met corresponding to the large increase in the measured quantum yield. The experimental data is plotted in figure 3.5 along with the theoretical calculations presented by Callcott (L is equivalent to the electron mean free path). The three step model with the field intensity calculated above is shown by

the dashed curve – it is in close agreement with the published theoretical calculations [19, 48]. However, both theoretical models underestimate the quantum yield by a factor of 2.

The surface plasmon resonance is an effective way to increase the photoemission efficiency by a large factor due to the increased absorption, field enhancement, and field localization close to the metal-vacuum interface. The theory derived here is a useful tool for analyzing the photoemission enhancement due to the SPP coupling on a thin metal film. In the following chapters the ideas developed here are applied to nano-cavities on the metal surface to design a substrate for enhanced multi-photon photoemission.

SPP Resonance in Metallic Nano-cavities

4

In the previous chapter it was shown that exciting an SPP wave on the metal film leads to an increase in the electric field intensity at the metal-vacuum interface. However, the SPP dispersion relation never intersects the light line, and, as a consequence, the SPPs are non-radiative: a momentum-matching mechanism is needed to couple light to the SPP on the surface. For the enhanced photoemission from an aluminum film, a prism was used [19], as described in detail in the previous two chapters. For applications where intensities are high, however, a bulk material is needed, in which case another mechanism has to be used. SPP coupling can also be done via addition (or subtraction) of the grating momentum [18], where the grating momentum $2\pi/g$, where g is the grating period, is added or subtracted from the light momentum to match to the SPP k-vector.

If the coupling is perfect, all the incoming light will be converted into an SPP wave providing the most intense fields. An idea was proposed theoretically by Le Perchec *et al.* [27] where a set of rectangular grooves in a form of a subwavelength grating is introduced on the metal surface (see figure 4.1) that provides for complete light absorption on-resonance. The dimensions of these nano-grooves (NGs) are on the order of the metal skin depth, 10 to 20 nm in width and 40 to 60 nm deep. When discussing the NG grating below, the polarization is defined as follows (irrespective of the angle of incidence): s-polarization

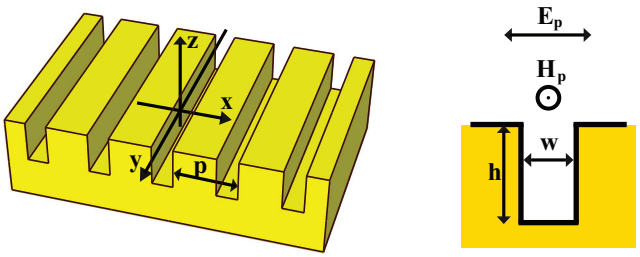

Figure 4.1: *Schematic of the plasmonic nano-grooves on a metal surface. The SPPs propagate down the NG along the z-axis, then reflect off the bottom and travel back to the mouth. On resonance, all the incident energy is trapped within the grooves.*

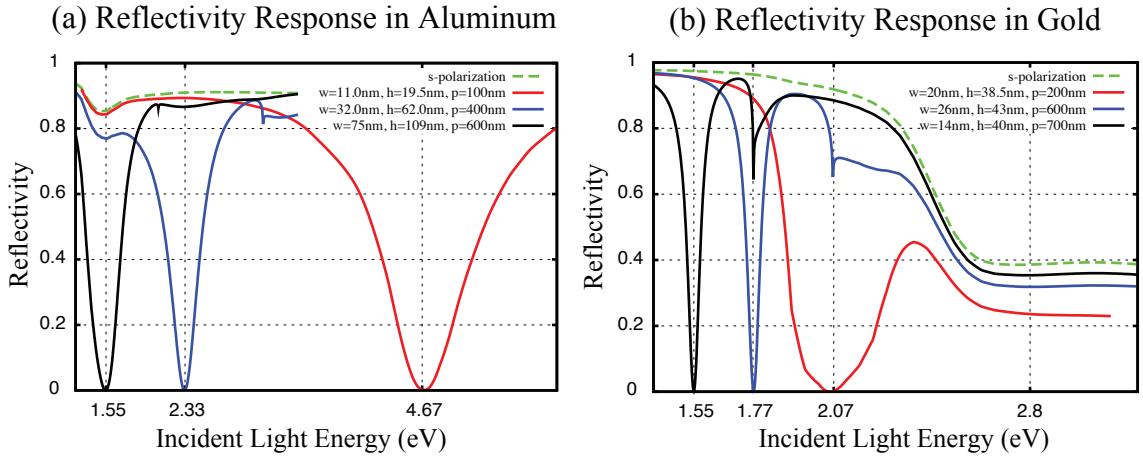

Figure 4.2: *The NGs on the metal surface serve both as the resonant nano-cavities and as the coupling mechanism. Adjusting the dimensions of the NGs allows designing the subwavelength grating for resonance anywhere from the UV to the NIR. Two case studies for aluminum (a) and gold (b) are shown. Reprinted with permission from Polyakov* et al. *[40] ©2011, American Institute of Physics.*

means light is polarized along the orientation of the NGs, and p-polarization means light is polarized perpendicular to the NGs. For all the discussion that follows polarization is always defined in reference to the orientation fo the NGs.

Due to the sharp edge of the NGs containing all Fourier frequencies, the NGs serve as a coupling medium, where light, incident onto the grating, couples to an SPP mode inside the grating. This mode is excited in the form of a wave travelling in the metal-insulator-metal (MIM) structure closed on the bottom end. This way, the NG also serves as a Fabry-Perot resonator further enhancing the electric field. In the design stage, the NG dimensions are chosen to produce the complete on-resonance absorption anywhere from the UV to the NIR, see figure 4.2 for the reflectivity spectrum samples for aluminum and gold. The reference s-polarization reflectivity response is plotted by a dashed line.

In the following section a dispersion relation is derived for the SPP wave in the MIM waveguide, that will be used to identify the metals, and the appropriate geometry, for enhancing photoemission from a bulk metal.

4.1 Metal-Insulator-Metal Dispersion Relation

The width of the NGs is on the order of the optical metal skin depth, which is much less then the wavelength, so the waveguide solution shows no optical propagating modes. Unlike the SPP wave at the metal surface derived in section 2.1, in the MIM waveguide, the solution

involves the coupled SPP waves at the two interface of the MIM waveguide [49]. Maxwell's equations take on the following form [21] (for convenience, the Gaussian system of units is used in the following derivation):

$$\begin{aligned}
\boldsymbol{\nabla}\cdot\boldsymbol{D} &= 0\\
\boldsymbol{\nabla}\cdot\boldsymbol{H} &= 0\\
\boldsymbol{\nabla}\times\boldsymbol{E} &= -\frac{1}{c}\frac{\partial\boldsymbol{H}}{\partial t}\\
\boldsymbol{\nabla}\times\boldsymbol{H} &= \frac{1}{c}\frac{\partial\boldsymbol{D}}{\partial t}
\end{aligned}$$

where $\boldsymbol{D}=\epsilon\boldsymbol{E}$. In solving the 2D system (as shown in figure 4.1) the $y-$component of the electric field can be ignored: $E_y=0$. Furthermore, the SPP wave is propagating in the $z-$direction with the amplitude that changes in the $x-$direction given by $F(x)$ and decays in the $z-$direction with no $y-$dependence. Therefore, the solution has the form:

$$F\left(x\right)e^{-i(\omega t-kz)}.$$

From the Maxwell's equations above:

$$\begin{aligned}
\boldsymbol{\nabla}\cdot\boldsymbol{D} &= 0=\epsilon_i\left(\frac{\partial E_x}{\partial x}+\frac{\partial E_y}{\partial y}+\frac{\partial E_z}{\partial z}\right)\\
E_z &= -\int\frac{dE_x}{dx}dz=-\frac{d}{dx}\int F\left(x\right)e^{-i(\omega t-kz)}dz\\
&= -\frac{d}{dx}\left[\frac{1}{ik}F\left(x\right)e^{-i(\omega t-kz)}\right]=\frac{i}{k}\frac{dE_x}{dx}\\
\boldsymbol{\nabla}\times\boldsymbol{E} &= \begin{vmatrix} i & j & k\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ E_x & 0 & E_z\end{vmatrix}=\begin{cases}\frac{\partial E_z}{\partial y} & =\frac{\omega}{c}H_x\\ \frac{\partial Ex}{\partial z}-\frac{\partial Ez}{\partial x} & =\frac{\omega}{c}H_y\\ -\frac{\partial Ex}{\partial y} & =\frac{\omega}{c}H_z\end{cases}\Rightarrow\begin{cases}H_x & =0\\ H_z & =0\end{cases},\\
\boldsymbol{\nabla}\times\boldsymbol{H} &= \begin{vmatrix} i & j & k\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ 0 & H_y & 0\end{vmatrix}=\begin{cases}-\frac{\partial Hy}{\partial z} & =\frac{-i\omega\epsilon}{c}E_x\\ 0 & =\frac{-i\omega\epsilon}{c}E_y\\ \frac{\partial Hy}{\partial x} & =\frac{-i\omega\epsilon}{c}E_z\end{cases}\Rightarrow -ikH_y=\frac{-i\omega\epsilon}{c}E_x
\end{aligned}$$

From the equations above, the relationship between the y-component of the **H** field and the x-component of the **E**-field is given by:

$$H_y=\frac{\omega\epsilon}{kc}E_x \tag{4.1}$$

Define $\kappa \equiv \sqrt{k^2 - \epsilon\frac{\omega^2}{c^2}}$. The SPP wave has both longitudinal and the transverse components. Rearranging the terms in the above result yields a wave equation in the x-direction[1]:

$$\begin{aligned}
\tfrac{dE_x}{dz} - \tfrac{dE_z}{dx} &= \tfrac{\omega}{c}H_y = \tfrac{\omega^2\epsilon}{kc^2}E_x, \\
ikE_x - \tfrac{i\epsilon}{k}\tfrac{\omega^2}{c^2}E_x &= \tfrac{dEz}{dx}, \\
\left(k^2 - \epsilon\tfrac{\omega^2}{c^2}\right)E_x &= \tfrac{k}{i}\tfrac{dEz}{dx}, \\
\tfrac{k}{i}\tfrac{d}{dx}\left[\tfrac{i}{k}\tfrac{dE_x}{dx}\right] - \kappa^2 E_x &= 0, \\
\tfrac{d^2E_x}{dx^2} - \kappa^2 E_x &= 0.
\end{aligned} \tag{4.2}$$

The MIM waveguide can be divided into three regions: (1) metal on the left, (2) vacuum, and (3) metal on the right. The general boundary conditions are [50]:

$$\begin{aligned}
\epsilon_i E_i^{\perp} &= \epsilon_j E_j^{\perp} \\
B_i^{\perp} &= B_j^{\perp} \\
\boldsymbol{E_i}^{\parallel} &= \boldsymbol{E_j}^{\parallel} \\
\frac{1}{\mu_i}\boldsymbol{B_i}^{\parallel} &= \frac{1}{\mu_j}\boldsymbol{B_j}^{\parallel}
\end{aligned}$$

To eliminate the time dependance, equation 4.2 can be rewritten in terms of $F(x)$ instead: $F_x'' - \kappa^2 F_x = 0$ and $F_z = (i/k)\,F_x'$. The solution for this equation can be written as a combination of two exponentials. Imposing the condition that the waves must go to zero at infinity inside the metal (regions 1 and 3) yields:

$$\begin{array}{cccccccccc}
 & & x<0: & & & 0\le x<d: & & & d\le x: \\
{}^1F_x & = & A_1 e^{\kappa' x} & {}^2F_x & = & C_1 e^{\kappa x} + C_2 e^{-\kappa x} & {}^3F_x & = & A_2 e^{-\kappa' x} \\
{}^1F_z & = & \frac{i\kappa'}{k}A_1 e^{\kappa' x} & {}^2F_z & = & \frac{i\kappa}{k}\left(C_1 e^{\kappa x} - C_2 e^{-\kappa x}\right) & {}^3F_z & = & -\frac{i\kappa'}{k}A_2 e^{-\kappa' x}
\end{array}$$

The conditions at each boundary are:

$$\begin{array}{cccccccc}
(i) & {}^1F_z(x=0) & = & {}^2F_z(x=0) & (iii) & \epsilon^1\cdot F_x(x=0) & = & \epsilon_0^2\cdot F_x(x=0) \\
(ii) & {}^2F_z(x=d) & = & {}^3F_z(x=d) & (iv) & \epsilon_0^2\cdot F_z(x=d) & = & \epsilon^3\cdot F_z(x=d)
\end{array}.$$

$$\begin{aligned}
(i) \quad & \frac{A_1\kappa'}{\kappa} = C_1 - C_2 \Rightarrow \frac{\kappa'}{\kappa} = \frac{C_1 - C_2}{A_1} \\
(ii) \quad & \frac{i\kappa}{k}\left(C_1 e^{\kappa d} - C_2 e^{-\kappa d}\right) = -\frac{i}{k}A_2\kappa' e^{-\kappa' d} \Rightarrow -\frac{\kappa'}{\kappa} = \frac{C_1 e^{\kappa d} - C_2 e^{-\kappa d}}{A_2 e^{-\kappa' d}} \\
(iii) \quad & \frac{\epsilon}{\epsilon_0} = \frac{C_1 + C_2}{A_1} \\
(iv) \quad & \frac{\epsilon}{\epsilon_0} = \frac{C_1 e^{\kappa d} + C_2 e^{-\kappa d}}{A_2 e^{-\kappa' d}}
\end{aligned}$$

[1]The x-direction wave is propagating perpendicular to the metal-insulator interface.

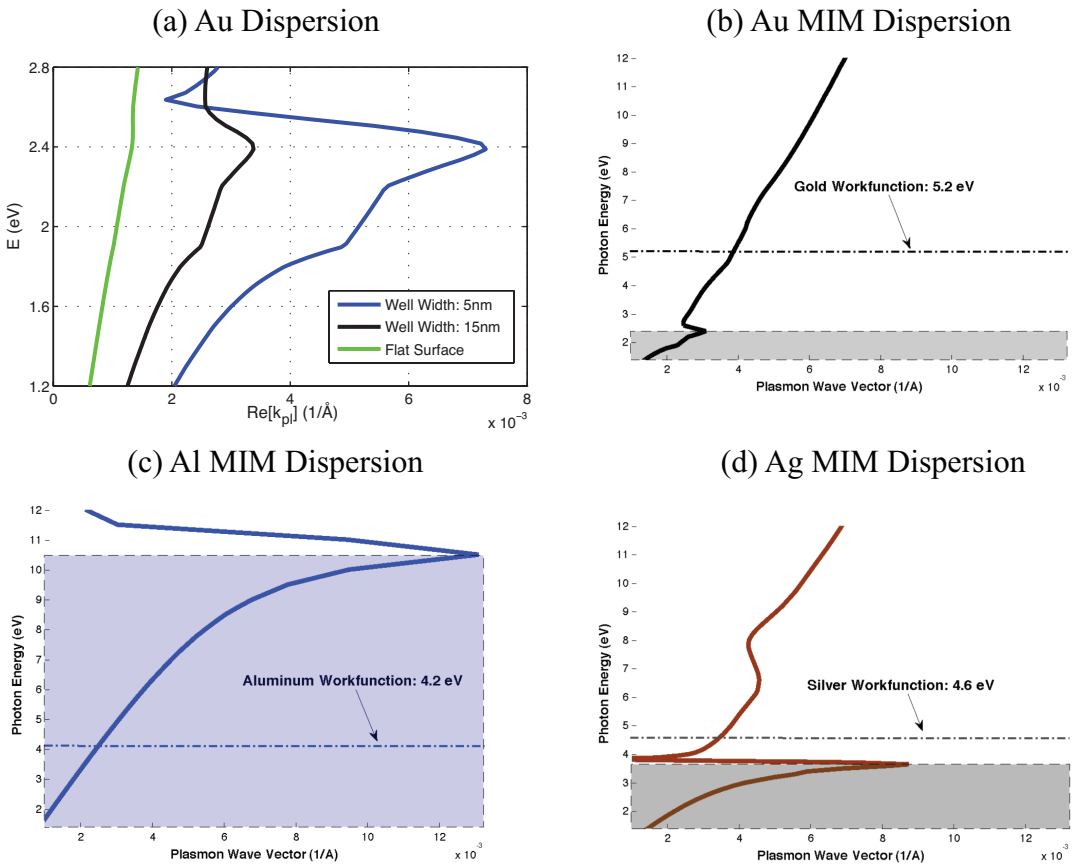

Figure 4.3: *The SPP mode in the MIM waveguide squeezes the SPP wavelength to a smaller volume, pushing the dispersion curve further towards the right. The SPP cut-off energy, however, remains constant (a). Comparing this cut-off energy to the metal's work function determines the resonance wavelength needed to excite an electron with sufficient energy to overcome the surface barrier for a one photon excitation. For gold (b) and silver (d), this results in multi-photon photoemission, while in aluminum (c), a single SPP can carry sufficient energy. Reprinted with permission from Polyakov* et al. *[40] ©2011, American Institute of Physics.*

Combining (i) and (iii) yields:

$$
\begin{aligned}
C_1 &= \frac{A_1}{2}\left(\frac{\epsilon}{\epsilon_0}+\frac{\kappa'}{\kappa}\right) \\
C_2 &= \frac{A_1}{2}\left(\frac{\epsilon}{\epsilon_0}-\frac{\kappa'}{\kappa}\right) \\
\frac{C_1}{C_2} &= \frac{\kappa\epsilon+\kappa'\epsilon_0}{\kappa\epsilon-\kappa'\epsilon_0}.
\end{aligned}
$$

Combining (ii) and (iv) and plugging the result into the above equations for C_1 and C_2 yields:

$$\begin{aligned}
-\frac{\kappa'}{\kappa} &= \frac{\epsilon}{\epsilon_0}\frac{C_1e^{\kappa d}-C_2e^{-\kappa d}}{C_1e^{\kappa d}+C_2e^{-\kappa d}} = \frac{\epsilon}{\epsilon_0}\frac{C_1-C_2e^{-2\kappa d}}{C_1+C_2e^{-2\kappa d}},\\
-\frac{\kappa'\epsilon_0}{\kappa\epsilon} &= \frac{\kappa\epsilon+\kappa'\epsilon_0-\kappa\epsilon e^{-2\kappa d}+\kappa'\epsilon_0e^{-2\kappa d}}{\kappa\epsilon+\kappa'\epsilon_0+\kappa\epsilon e^{-2\kappa d}-\kappa'\epsilon_0e^{-2\kappa d}} = \frac{\kappa\epsilon\left(1-e^{-2\kappa d}\right)+\kappa'\epsilon_0\left(1+e^{-2\kappa d}\right)}{\kappa\epsilon\left(1+e^{-2\kappa d}\right)+\kappa'\epsilon_0\left(1-e^{-2\kappa d}\right)}.
\end{aligned}$$

$$\begin{aligned}
-\kappa'\epsilon_0\kappa\epsilon\left(1+e^{-2\kappa d}\right)-\left(\kappa'\epsilon_0\right)^2\left(1-e^{-2\kappa d}\right) &= \left(\kappa\epsilon\right)^2\left(1-e^{-2\kappa d}\right)+\kappa'\epsilon_0\kappa\epsilon\left(1+e^{-2\kappa d}\right),\\
\left(\kappa\epsilon\right)^2+2\kappa'\epsilon_0\kappa\epsilon+\left(\kappa'\epsilon_0\right)^2 &= \left(\kappa\epsilon\right)^2e^{-2\kappa d}-2\kappa'\epsilon_0\kappa\epsilon e^{-2\kappa d}+\left(\kappa'\epsilon_0\right)^2e^{-2\kappa d},\\
\left(\kappa'\epsilon_0+\kappa'\epsilon\right)^2 &= \left(\kappa\epsilon-\kappa'\epsilon_0\right)^2e^{-2\kappa d},\\
\frac{\kappa\epsilon+\kappa'\epsilon_0}{\kappa\epsilon-\kappa'\epsilon_0} &= \pm e^{-\kappa d}.
\end{aligned}$$

The resulting relation between the C_1 and C_2 coefficients is given by:

$$\frac{C_1}{C_2} = \pm e^{-\kappa d}.$$

Combining (ii) and (iv) and plugging in the above result yields:

$$\begin{aligned}
-\frac{\kappa'\epsilon_0}{\kappa\epsilon} &= \frac{C_1e^{\kappa d}-C_2e^{-\kappa d}}{C_1e^{\kappa d}+C_2e^{-\kappa d}} = \frac{1-\left(C_2/C_1\right)e^{-2\kappa d}}{1+\left(C_2/C_1\right)e^{-2\kappa d}} = \frac{1\pm e^{-\kappa d}}{1\mp e^{-\kappa d}},\\
\frac{\kappa'\epsilon_0}{\kappa\epsilon} &= \tanh\left(-\frac{\kappa d}{2}\right),\\
\frac{\kappa'\epsilon_0}{\kappa\epsilon} &= \coth\left(-\frac{\kappa d}{2}\right).
\end{aligned}$$

Inserting the definitions for $\kappa' = \sqrt{k^2-\epsilon_{metal}\frac{\omega^2}{c^2}}$ and $\kappa = \sqrt{k^2-\epsilon_{insulator}\frac{\omega^2}{c^2}}$ yields the final dispersion relation in the MIM waveguide:

$$\frac{\epsilon_{insulator}\sqrt{k^2-\epsilon_{metal}\frac{\omega^2}{c^2}}}{\epsilon_{metal}\sqrt{k^2-\epsilon_{insulator}\frac{\omega^2}{c^2}}} = \begin{cases}\tanh\left(-\frac{d}{2}\sqrt{k^2-\epsilon_{insulator}\frac{\omega^2}{c^2}}\right)\\ \coth\left(-\frac{d}{2}\sqrt{k^2-\epsilon_{insulator}\frac{\omega^2}{c^2}}\right).\end{cases} \tag{4.3}$$

The equation above cannot be solved analytically and must instead be solved numerically– see appendix A for an example *MatLab* code. The dispersion relation is plotted for different MIM widths in comparison to the SPP dispersion on the flat surface in 4.3(a). The advantage gained by using the MIM waveguide is the extreme stretch of the dispersion relation, effectively *squeezing* light into a smaller volume. However, the energy cut-off remains constant,

which sets the maximum energy an SPP can carry, and, therefore, the maximum energy that can be transferred to an electron. This parameter will determine the resonance wavelength needed to photoemit from a given metal. Three case studies are shown figure 4.3: (b) for gold, the maximum SPP energy is ≈ 2.2 eV, while the work function is 5.2 eV, therefore, a three-photon process is required to photoemit; similarly in (d) two photons are required to give enough energy for a single electron to overcome the metal's work function. Aluminum is a special metal in this regard (c): the work function is much lower then the cut-off energy, and, therefore, a single-photon photoemission is possible.

4.2 Model for On-Resonance Impedance Matching

As illustrated in figure 4.2, the dimensions of the NGs can be optimized to produce a strong resonance at any target wavelength from the UV to the NIR yielding complete absorption. This effect was investigated in a Finite Difference Time Domain (FDTD) simulation. To develop a model for the light interaction with the NGs, the time evolution of the incident and reflected light, as well as the formation of the SPP wave, was observed as shown in figure 4.4. For this study the NG was chosen to be 1400 nm deep to observe the decay of the SPP wave packet. The details for the FDTD simulation are described in appendix B.

When considering light coupling structures such as the NG grating presented here (see an example for NGs 15 nm wide, 30 nm deep, at 200 nm period in figure 4.5), a fundamental question becomes whether these units can be treated independently or must be treated in terms of a collective excitation of the system. The assumption that the performance of these types of systems can be modeled on the basis of a single unit has also been followed in a number of other studies [49, 51, 52]. However, it is well known that collective effects can play an important role in determining the behavior of some types of plasmonic systems [25, 53–57]. The analysis of a single resonant NG narrower then the metal skin depth can be connected to the collective excitation of the subwavelength grating in a simple relationship between the NG width w, height h, and period p required to achieve complete light absorption at the metal surface.

To give an insight into the fundamental nature of this phenomenon, a surface impedance-based analytical model for complete omni-directional light absorption in the NGs is presented below. First, the general features of the light-SPP interaction are investigated to define the region of validity for an analytical theory. When light interacts with a sharp metal edge, as in the case of a grating structure, SPPs are generated along the top surface as well as inside the grooves (for an experimental time resolved study of this effect see the work by Kubo *et al.* [58]). To determine which SPPs contribute to the on-resonance absorption, consider two gratings $NG1$ and $NG2$ optimized for complete absorption at different wavelengths. These give the reflectivity responses as shown by the dashed lines in figure 4.6. When two such gratings are interleaved, a composite structure is formed with the unit cell as shown in the inset of figure 4.6. The design of this system is such that the SPPs generated by $NG1$ that travel on the top surface are blocked by $NG2$, yet the reflectivity response of the composite

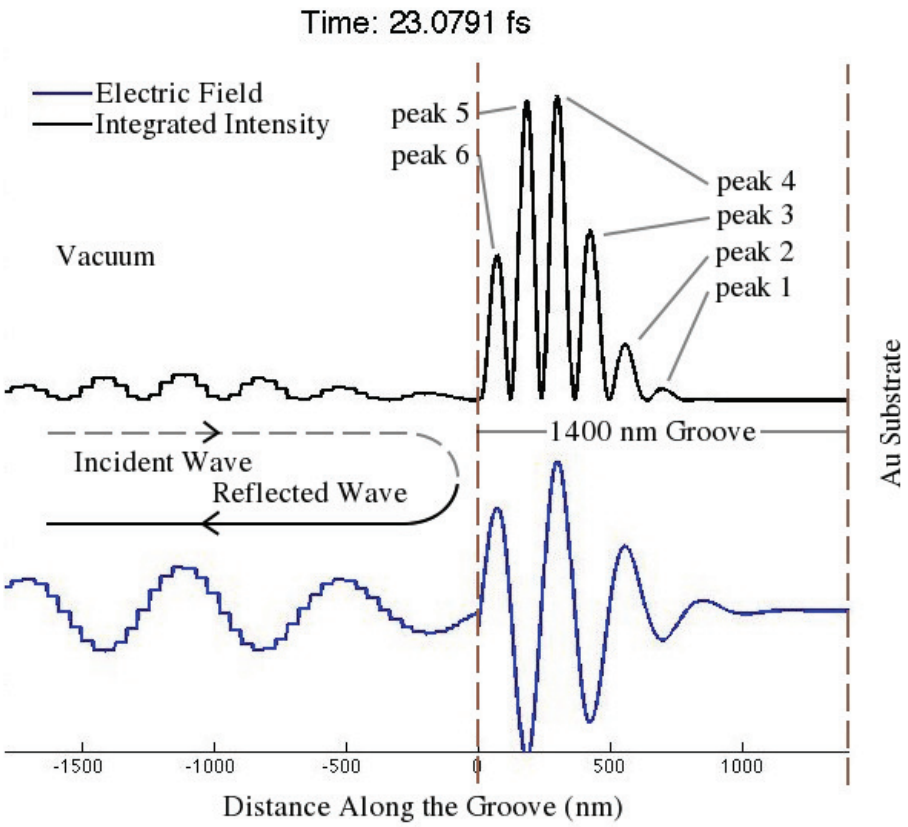

Figure 4.4: *The FDTD simulation is useful to simulate the fields' time evolution (top). In this case study, the gold NG grating was designed to have an extreme depth of 1400 nm to observe the SPP propagation inside the NG. The incoming and reflected waves were observed to the right of the metal boundary, while the SPP wave packet evolution is seen inside the NG. Such simulation is useful when developing a physical model for the light interaction.*

grating is a linear combination of the reflectivity responses of $NG1$ and $NG2$ (see the solid curve in figure 4.6) similar to the independent transmission of the photon sorters [59]. This independent behavior of each grating, when put in combination, shows that on-resonance, the top-surface SPPs do not contribute to the reflectivity response of the subwavelength grating.

However, in a dense grating the neighboring NGs can interact via the evanescent fields of the SPPs travelling down the walls of the NGs. Figure 4.7 illustrates the reflectivity effect of this interaction for small (a) and large (b) grating period where reflectivity is plotted for $\lambda_0 = 800$ nm normal incidence light when w is held constant and h is varied from 0 to 1.5 μm. From the set of reflectivity dips the SPP wavelength inside the groove can be calculated using the Fabry-Perot analysis: $2k_{pl}h_m + \phi = 2\pi m$, where $k_{pl} = k'_{pl} + ik''_{pl} = \frac{2\pi}{\lambda_{pl}}$ is the plasmon wavevector, m is the index of each dip and h_m is the corresponding depth. In

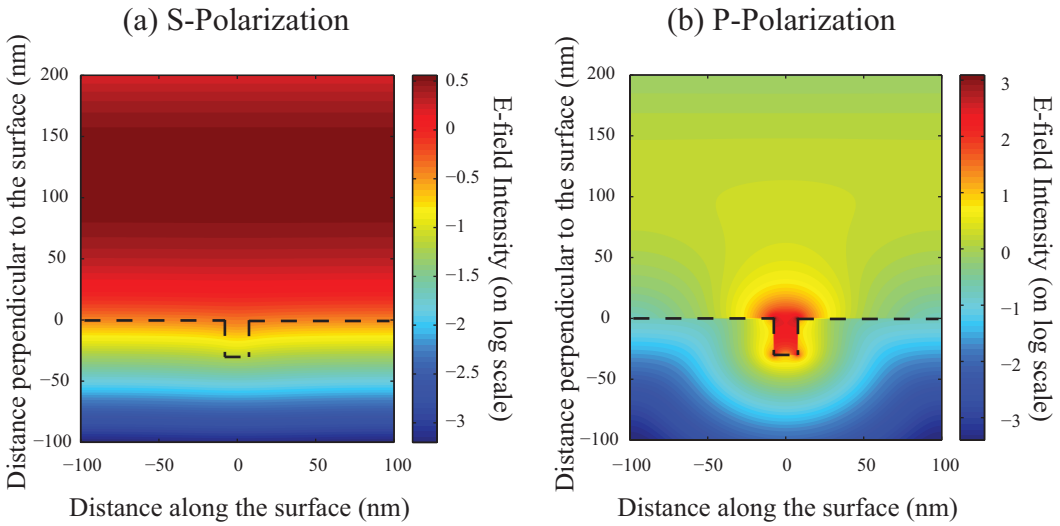

Figure 4.5: *When the NGs are illuminated at normal incidence with the light polarized parallel to the NGs (s-polarization), no SPP modes are excited and bulk-like response is observed: the field decays exponentially into the metal (a). However, for light polarized perpendicular to the NGs, the SPP mode is excited with the field profile as shown in (b). In this case, the field enhancement is beyond 500 creating an optimum environment for the non-linear process enhancement, such as the multi-photon photoemission.*

this way, the dispersion relation is numerically calculated for a given NG width and varying NG period: the result is shown in the inset of figure 4.7b. The plasmon wavelength varies from the incident light wavelength (for p close to w) to the SPP dispersion in the metal-insulator-metal (MIM) waveguide [21] as p approaches the wavelength of light λ_0. For 10 nm wide NGs in gold, the SPP evanescent wave decay length in the metal is given by [18] $1/|\sqrt{\epsilon_{metal}[2\pi/\lambda_0]^2 - [k_{pl,z}]^2}| \approx 30$ nm. The overlap between the evanescent fields from the neighboring NGs can be neglected for $p > 100$ nm, but is significant for $p = 50$ nm. For such a dense grating achieving complete absorption requires micron-deep NG channels due to the long SPP wavelength. The case of interest, where the reflectivity extinction occurs for $h < 70$ nm yielding strong field enhancement [51] and field localization close to the metal surface, is shown in figure 4.7b corresponding to a sparse grating. The distance corresponding to the overlap of the evanescent fields sets lower limit on the region of validity for the analytical model presented here.

As the period is further increased, for wavelengths of light close to the NG period, the SPPs generated at the metal surface in between the NGs can outcouple to photons due to the addition of grating momentum. To quantify this effect, the concept of angular bandwidth is used, which is defined as the angle of incidence where reflectivity reaches 0.5 starting from 0 at normal incidence (figure 4.8 shows reflectivity calculated for a range of angles of incidence). Achieving complete absorption for NG periods approaching the wavelength of

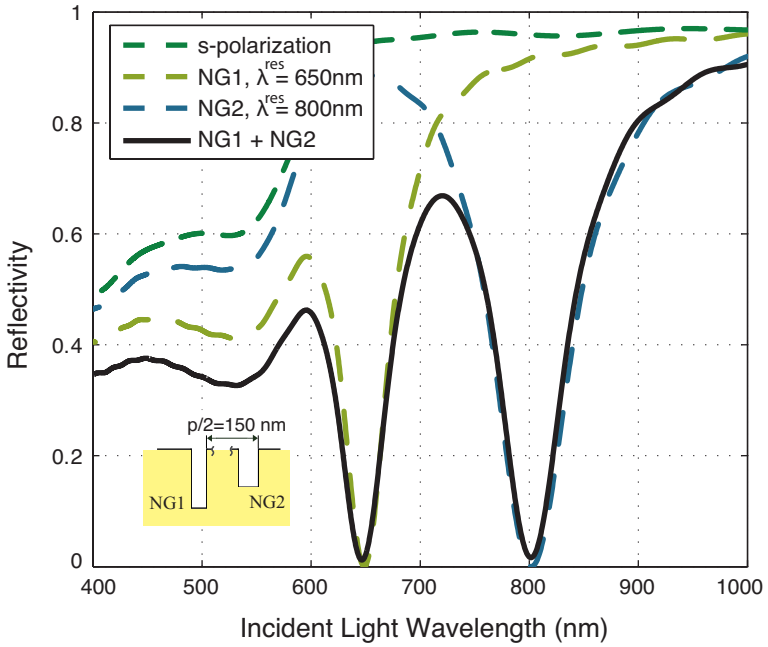

Figure 4.6: *Reflectivity response for a combined double grating where one NG set is inserted in between another one. The dimensions for the sets NG1 and NG2 are $w_1 = 19$ nm, $h_1 = 45$ nm, $w_2 = 10.8$ nm, and $h_2 = 56.1$ nm. Period for both gratings is $p = 300$ nm. The independent behavior of the two gratings when combined shows that, on resonance, the SPPs traveling on the flat surface do not contribute to the reflectivity response of the NG grating.*

light yields extremely narrow angular bandwidth and the reflectivity response becomes that of a regular grating: a sharp resonance that is highly angle dependent. Therefore, the upper limit on the region of validity for the analytical model is approximately the wavelength of light λ_0.

Achieving the resonance condition requires a set of NG period, width, and height, which, in the range of validity $w << p < \lambda_0$, can always be found by numerical optimization to produce on-resonance reflectivity consistently below 0.1%. For the following discussion, the reflectivity is measured when the geometry of individual NGs is optimized to produce such resonance for a given wavelength of light. The resulting NG dimensions are plotted in figure 4.9 along with the characteristic spectral bandwidth measured by the full width half maximum (FWHM) of the resulting reflectivity dip (for reflectivity spectrum examples see figure 4.6). In the context of our model a subwavelength grating can be optimized for both large spectral and angular bandwidth yielding omni-directional (see the bottom left inset of figure 4.8) complete light absorption at the metal surface.

The model for the on-resonance behavior of such subwavelength gratings consists of calculating the surface impedance due to the ensemble of the NGs, each of which are treated as independent radiation sources. For p-polarization, where the E-field of the incoming light

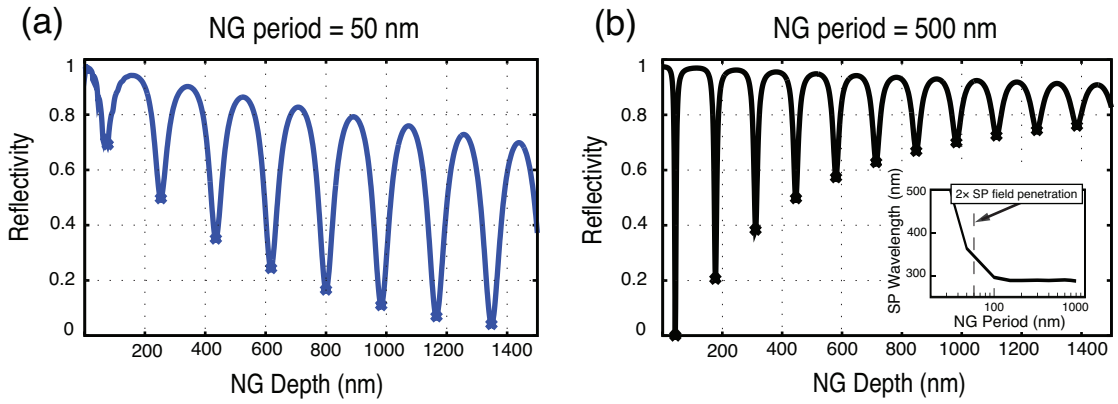

Figure 4.7: *Effect of the NG period on the reflectivity response in gold. The NGs constitute a set of resonant cavities: (a) SPP field penetration into the metal is the dominant factor for a dense grating, where complete absorption is achieved only for extreme h; (b) in a sparse grating NGs act as resonant cavities useful for field enhancement and field localization close to the metal surface. The inset shows the change in the real part of the plasmon wavelength due to the change in the NG period.*

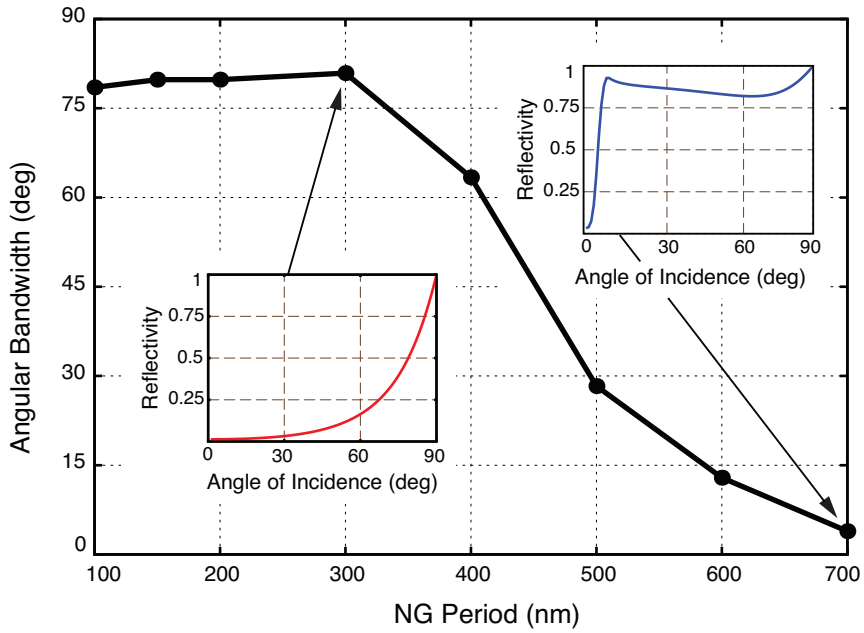

Figure 4.8: *Angular bandwidth is affected by the NG period. As the period is increased approaching the wavelength of incident light the angular bandwidth sharply drops and the reflectivity response acquires more of a grating-like character: a sharp resonance that is highly sensitive to the angle of incidence and wavelength.*

is polarized perpendicular to the orientation of the NGs (as shown in figure 4.1) the H-field is approximately constant across the flat surface and the mouth of the NG [27]. From the boundary conditions, the electric field at the metal surface away from the NGs is close to zero, which is equivalent to an out-of-phase sum of the incoming and reflected waves. In this case, the reflected H-field is in phase with the incoming H-field, and hence the amplitude is doubled: $H_{surface} \approx 2H_0$. It is important to note that no reflected wave exists in the far field, so this approximation is only valid at the metal boundary. On resonance, all the incoming electric field at the metal interface is trapped within the NG [27], and is further amplified due to the NG cavity [51]: $E_{surface} = E_0 \times \frac{w}{p}Q$, where Q is the NG quality factor that will be defined below.

The reflectivity can be expressed in terms of the surface impedance [56], which captures the effect of the ensemble of the NGs, $Z_{surf} = \frac{E_{surface}}{H_{surface}}$, where Z_0 is the impedance of free space:

$$Z_{surf} = \frac{E_0}{H_0} \times \frac{w}{p}\frac{Q}{2} \tag{4.4}$$

$$R = rr*,\ r = \frac{\eta - 1}{\eta + 1} \tag{4.5}$$

Substitute the relationship $E_0/H_0 = i \times k_{pl}c/\omega = i \times k_{pl}/k_0^R$ from the solutions to Maxwell's equations for a transverse magnetic monochromatic plane wave [21] ($H_x = H_z = E_y = 0$), where $k_0^R = \omega/c$ is the on-resonance wave vector of the incident light. The final expression for the surface impedance of an ensemble of the NGs becomes:

$$Z_{surf} = i \times \frac{k_{pl}}{k_0^R} \cdot \frac{w}{2p} \cdot Q \tag{4.6}$$

Next, consider the fields inside the individual grooves. On resonance, the mode in the NG forms a standing wave, where H- and E-fields travel 90^o out of phase, and have the same mouth-to-bottom amplitude variation. In this way, the Q-factor can be defined as the ratio between the H-field at the bottom to H-field at the mouth of the NG: $H_{bottom} = Q \times H_{top}$, $H_{NG} = 2H_0 \cos(k_{pl}z)$. On resonance, the plasmon wavevector can be rewritten using $\Delta k_{pl} \equiv k'_{pl} - k_{pl}^R$ to be $k_{pl} = k_{pl}^R + \Delta k_{pl} + ik''_{pl}$, where k_{pl}^R is the SPP wavevector on resonance. Taylor expanding the cosine term yields the final expression for Q:

$$Q = \frac{2}{\pi} \times \frac{1}{\frac{\Delta k_{pl}}{k_{pl}^R} + i\frac{k''_{pl}}{k_{pl}^R}} \tag{4.7}$$

The Δk_{pl} term can be rewritten in terms of the derivative $dk_{pl}/d\omega$:

$$\frac{\Delta k_{pl}}{k_{pl}^R} = \frac{dk_{pl}}{k_{pl}^R}\frac{\Delta\omega}{d\omega}\frac{\omega_R}{\omega_R} = \frac{\Delta\omega}{\omega_R}\frac{\omega_R}{k_{pl}^R} \cdot \frac{dk_{pl}}{d\omega} \tag{4.8}$$

Putting together equation 4.7 for an individual NG for $k_{pl} \rightarrow k_{pl}^R$ with the effect of an ensemble of NGs given by equation 4.6 yields:

$$Z_{surf} = i \times \frac{\alpha/\omega}{\Delta\omega/\omega + i \cdot \beta/\omega}, \begin{cases} \alpha/\omega = \frac{1}{Z_0}\frac{\omega c}{\pi p}\frac{k_{pl}}{\omega}^2 \frac{d\omega}{dk_{pl}} \\ \beta/\omega = \frac{k''_{pl}}{\omega}\frac{d\omega}{dk_{pl}} \end{cases} \tag{4.9}$$

Equation 4.5 can now be rewritten in terms of energy lost due to radiation α and energy dissipation in the metal β:

$$R = \frac{\Delta\omega^2 + (\alpha - \beta)^2}{\Delta\omega^2 + (\alpha + \beta)^2} \tag{4.10}$$

On resonance, when $\Delta\omega = 0$, the optimum geometry of the NG yields a balance between the radiative loss and the energy dissipation in the metal: $\alpha = \beta$. The period-to-width relationship then is given by:

$$\frac{p}{w} = \frac{1}{\pi}\frac{[k'_{pl}]^2}{k_0^R k''_{pl}} \tag{4.11}$$

A case study was performed for copper and gold: a comparison between the analytical model given by equation 4.11 and the FDTD calculations for 800 nm normal incident light is shown in figure 4.9. In the region of validity, the p/w ratio from the analytical model is in excellent agreement with the FDTD simulation for both materials (see figure 4.9a). The small disagreement in the absolute value for the p/w ratio–as seen in figure 4.9b–is attributed to the effects of the SPPs propagating on the surface in between the NGs, which is beyond the present model. This effect is more prominent in the case of gold where the flat surface SPP dispersion relation lies closer to the light line, rendering top-surface SPP coupling more efficient than in the case of copper.

An analogy to an electronic circuit can be made using the idea of surface impedance. For example, for copper, using equations 4.6 and 4.7, throughout the region of validity the surface impedance $Z_{surf}^{Cu} = Z_0 \pm 5\%$. Tuning the NG geometry and period can be understood as an impedance matching phenomenon. Most importantly, for practical applications the subwavelength grating can be designed to be impedance matched for both maximum angular and spectral bandwidth (see figures 4.8 and 4.9). The period-to-width relationship given by equation 4.11 is a powerful tool in designing subwavelength plasmonic gratings for omni-directional complete light absorption at the metal surface.

This way, the NGs on the metal surface provide a powerful tool for designing the optical response with controllable angular bandwidth, resonance position, and absorption strength. The interaction model presented here is the basis for choosing the metal most suitable as a plasmon enhanced photocathode presented in the following chapters. Comparing the dispersion relation of each metal to the metal's work function determines the order of the multi-photon process required, which is directly related to the photoemission efficiency.

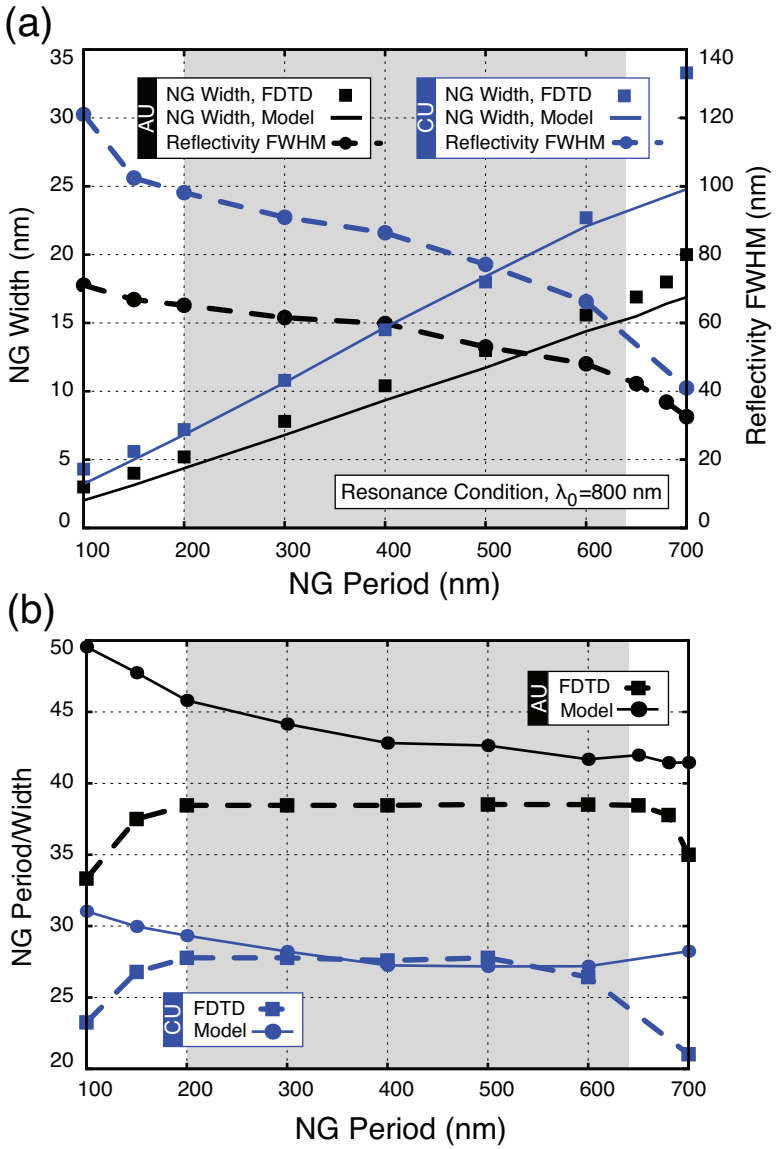

Figure 4.9: *Resonant behavior of the nano-groove. Producing a minimum reflectivity on-resonance at constant incident wavelength for a range of NG periods requires adjusting the NG dimensions. In the region of validity (shown in gray background), the analytic model presented here agrees well with the numerical FDTD calculations on the absolute scale as shown in (a). On the relative scale, the NG period-to-width ratio–shown in (b)–is flat in the validity region, which is closely approximated by our model: the NG width increases linearly with the period accompanied by bandwidth (FWHM) narrowing as shown for Au (black) and Cu (blue).*

5 Nanofabrication

5.1 Silicon Template

Producing rectangular grooves on the metal surface with dimensions as shown in figure 4.9 is a major fabrication challenge [60]. For best light trapping performance, the flat surface in between the grooves must be as smooth as possible to reduce the SPP coupling via surface roughness. This requirement means that etching of the gold surface itself is not possible, due to surface roughness of the gold film. The first step is to prepare an ultra flat gold surface. One way to do this is to evaporate gold onto the atomically smooth surface of a prime silicon wafer: the roughness of the gold film at the gold-silicon interface will be minimal.

This led to the fabrication approach, where the silicon was patterned, instead of gold, forming a template with an inverse geometry. Then, gold was evaporated onto the template and later debonded. Two template fabrication approaches were used. The first one involves etching a template from the silicon wafer, and in the second approach, the template is prepared by the electron beam lithography[1]. The advantage of the first method is the ability to produce sub 10 nm features with the atomically smooth walls at the cost of the fabrication difficulty. The second approach is substantially easier with greater control and repeatability, however, the minimum size of features is limited to above 10 nm.

The steps involved in preparing the silicon template are illustrated in figure 5.1. The silicon wafer used for this process is cut in $< 110 >$ orientation so that (110) direction is perpendicular to the wafer surface and the (111) direction is along the surface. In this case the advantage is taken of the anisotropic etch of silicon [62], where silicon in the (111) direction etches hundreds of times slower than in the (110) direction in the potassium hydroxide (KOH) solution. The initial pattern is formed in the Poly(methyl methacrylate) (PMMA) electron beam resist, which is then transferred onto the silicon wafer via the Cr lift-off process and the reactive ion etch (RIE). After a short exposure to KOH a pattern is etched in the silicon, that is used as a template: gold is evaporated onto the silicon template and later debonded to produce the final device.

[1]This work was done as a part of a user project at the Molecular Foundry, Lawrence Berkeley National Laboratory (LBNL). The electron beam lithography is described in greater detail in ref [61].

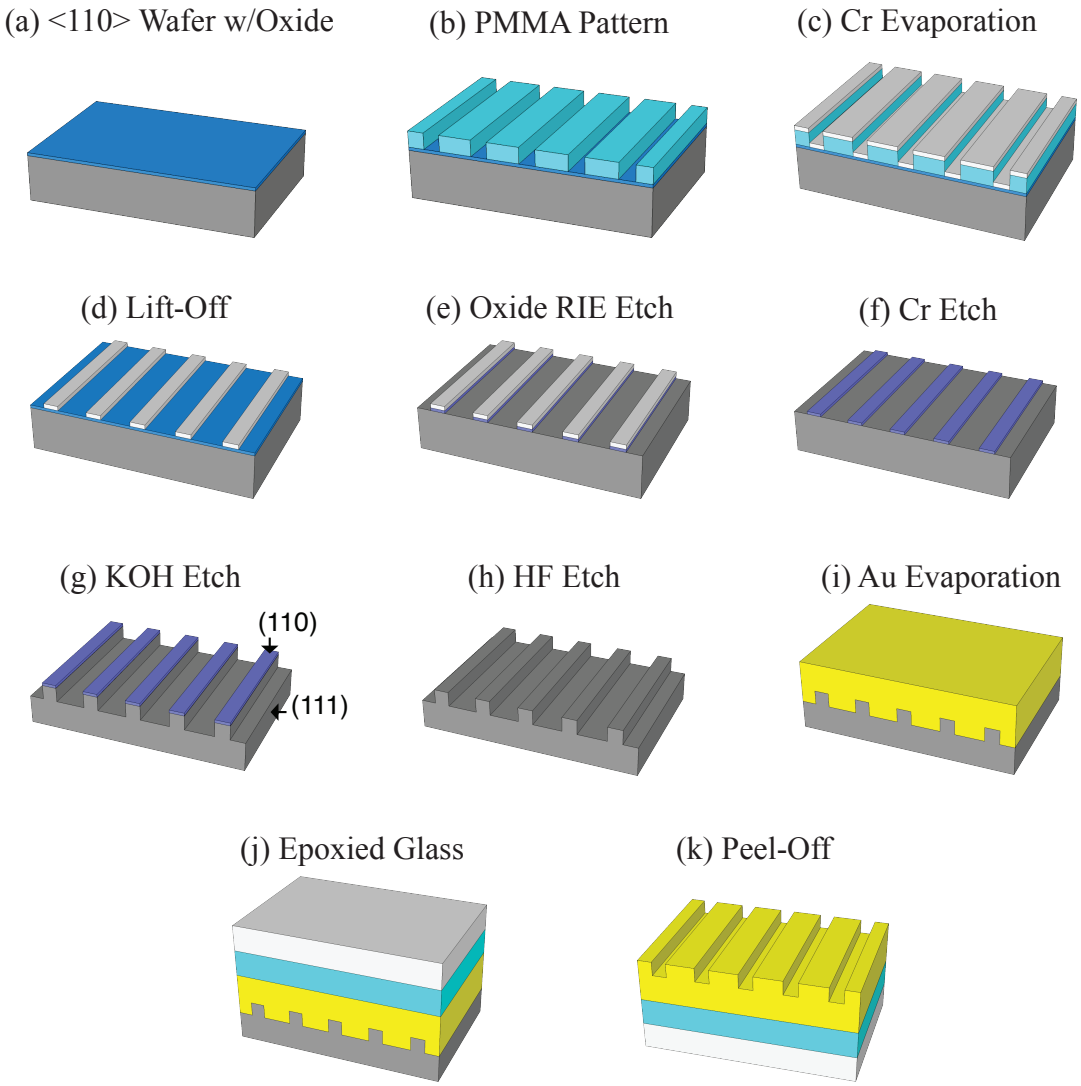

Figure 5.1: *The outline of the fabrication process based on preparing a template in the silicon wafer for producing rectangular grooves in the gold film.*

5.2 Wafer Preparation

The most important part in nanofabrication is keeping the samples clean, and this starts with the prime wafer, right out of the fresh new box. The following process is a version of the cleaning process developed by the Radio Corporation of America (RCA); throughout the following discussion it is simply referred to as the RCA clean. It removes organic contaminants and also forms ≈ 1 nm oxide on a pure piece of silicon. It is used for both cleaning the wafer, as well as forming a thin oxide layer.

Recipe 1 ***RCA Clean***
Use a solution of $NH_4OH : H_2O_2 : DI$ water in 1 : 1 : 5 ratio at 75^o for 15 minutes, follow by a 5 minute DI water bath, then dry the sample with nitrogen.

Immediately before spinning on the resist onto a wafer, bake it for 10 minutes at 80^o and then place the wafer into an ozone chamber for 5 minutes. This process also removes organic contaminants as well as mildly roughens the surface, which should enhance the adhesion of the resist.

5.3 PMMA Pattern

The initial line pattern is formed in the electron beam PMMA resist that will be translated by a set of steps into a silicon dioxide mask on the wafer for the later KOH etch. The width of these initial lines is 15 nm, therefore PMMA is a good choice, since it is a positive tone resist (the exposed areas are dissolved during the development) requiring short write times. The procedure for spinning on the PMMA resist is outlined below:

Recipe 2 ***Resist Spin-On***

1. *RCA clean the wafer per recipe 1*
2. *Preheat the wafer at 100^o for 1 minute*
3. *Put 10 drops of hexamethyldisilazane (HMDS) onto the wafer and spin at 2000 rpm*
4. *Bake the wafer at 100^o for 1 minute*
5. *Spin on PMMA at 3000 rpm*
6. *Bake the wafer at 180^o for 5 minutes*

After the electron beam exposure the PMMA is developed using the following recipe:

Recipe 3 ***PMMA Developing***

1. *Cool down the sonic bath using ice packs for 1.5 hours (to 5^o C)*

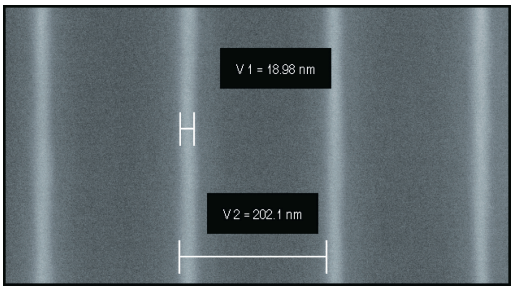

Figure 5.2: *The SEM image of the PMMA pattern after developing. The lines are not very sharp due to low resolution of the PMMA resist. However, this is not essential at this stage of the process.*

2. *Prepare an isopropyl alcohol (IPA) : DI water solution in 7 : 3 ratio and place it into the sonic bath*
3. *With sonification ON, develop the wafer for 100 seconds*
4. *Immediately dry the wafer with nitrogen*

The resulting lines in the PMMA resist are not very sharp due to the low resolution of the resist. However, this is not a problem, because the later processing steps will reduce the roughness and the final template will have extremely sharp features. A scanning electron microscope (SEM) image of the lines produced this way is shown in figure 5.2.

5.4 Deposition and Lift-Off

Notice in figure 5.1(c) that a thin layer of Cr does not form a continues layer across the trenches in the PMMA. Instead, the bottom of the trench is filled in a separate layer from the top. Dissolving the PMMA removes the top layer, while leaving the bottom layer intact. This process is called lift-off.

The first step is to run a quick *descum* process on the RIE machine to remove any remaining undeveloped or underdeveloped PMMA from the trenches as in figure 5.1(b). The parameters for the descum are as follows:

Recipe 4 ***De-Scum RIE Etch***
Flow O_2 at 30 sccm at 20 W forward power for 10 seconds.

This gentle de-scum also opens up the trenches by a few nanometers, which can be later compensated for by a short hydrofluoric acid (HF) dip later. Then Cr is evaporated onto the wafer using the electron beam evaporation at a 3 Å/s rate. For the PMMA prepared per recipe 2, 10 nm of Cr layer works well for the lift-off process. After that, dissolving the PMMA is done by submerging the wafer in a Dichloromethane solution as follows:

(a) Lift-Off Failure (b) Proper Cr Mask

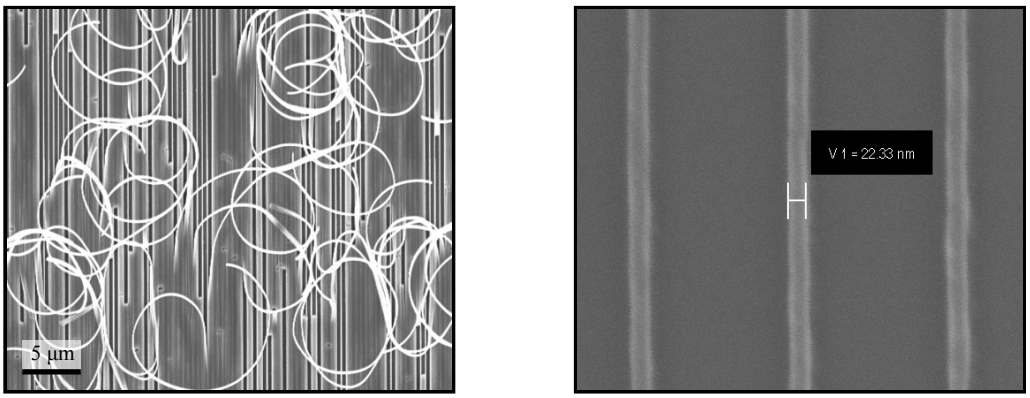

Figure 5.3: *A lift-off process can fail if either the Cr layer is too thick or the lift-off is done without sonification (a). Normally, the Cr layer will result in a clean mask (b) with minor defects spread across the sample. However, the number of defects should be minimal, no more than a few spots per square millimeter.*

Recipe 5 ***Lift-Off via PMMA***

1. *Submerge the wafer face down, with one end propped up, in Dichloromethane for 6-10 minutes. Watch for the Cr to flake off. The next steps must be done without letting the wafer surface dry*
2. *Wash off the extra Cr by spraying the wafer with acetone*
3. *Place the wafer into acetone in a sonic bath for 2 minutes*
4. *Clean the wafer with IPA and dry with nitrogen*

If the lift-off process fails, and Cr fails to detach with the dissolved PMMA, the sample will resemble the SEM image in figure 5.3(a). There is no way to rescue such a sample and the process must be restarted from the beginning. However, if the lift-off is successful, then the Cr layer will produce a clean mask as seen in figure 5.3(b). Some defects may be present throughout the sample: it is common to see a couple spots where Cr failed to detach per square millimeter.

5.5 RIE Etch

After the lift-off, Cr forms the mask for etching the silicon dioxide layer[2] in the RIE. This step of the process is the most reliable. The process outlined below is used to fully etch 50

[2]The oxide was thermally grown by the wafer supplier.

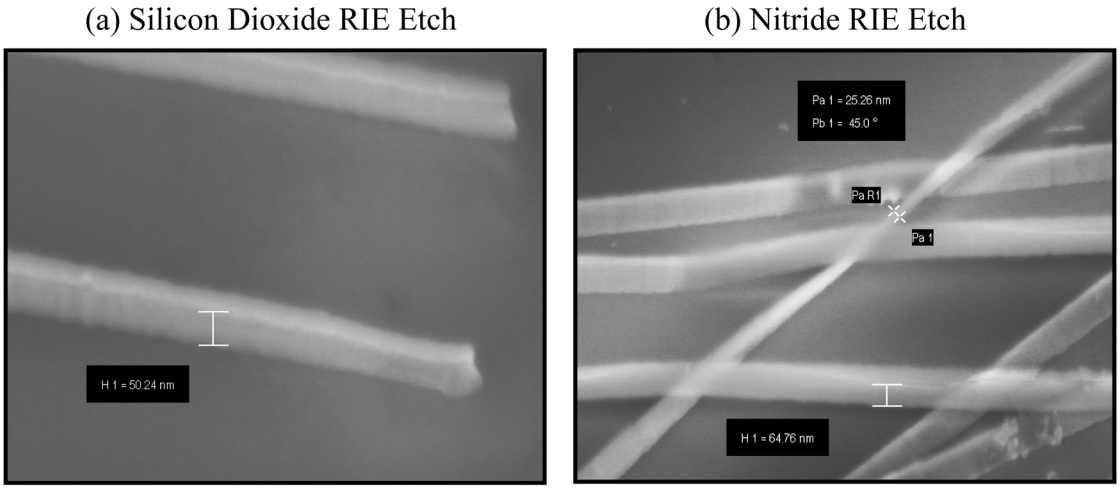

Figure 5.4: *Due to the excellent adhesion between the thermally grown oxide layer and the silicon wafer, nanometer-size high aspect ratio silicon dioxide structures can be prepared on the silicon wafer (a). This is not the case for silicon nitride, which adheres poorly to silicon, and will tend to detach once the dimensions are below 20 nm (b).*

nm of thermally grown silicon dioxide[3]:

Recipe 6 ***50 nm Si_2O_3 RIE Etch***

time	=	3 : 30	*ramp rate*	=	5	*table temp.*	=	$20°C$
set pressure	=	10.0 $mTorr$	*power*	=	100 W	CHF_3	=	25 $sccm$
strike pressure	=	10.0 $mTorr$	*cap* 1	=	50 %			
DC bias	=	50 V	*cap* 2	=	50 %			

The SEM image of the resulting structure is shown in figure 5.4; see the corresponding schematic in figure 5.1(f). The silicon dioxide was chosen as the mask for the KOH etch instead of the silicon nitride due to the better adhesion of the silicon dioxide. Once the dimensions of the structures reach values below 20 nm, the nitride starts to detach from the underlying silicon wafer as seen in figure 5.4(b). Silicon dioxide, on the other hand, adheres strongly allowing formation of much smaller structures, below 10 nm.

Before continuing to the next step, the remaining Cr must be removed from the sample, because leaving it on the sample interferes with the KOH etching completely ruining the sample. Cr is removed by placing the sample into a Cr etch for 10 minutes.

[3]All the numbers refer to the recipe used for the *Oxford Instruments Plasmalab* RIE etcher

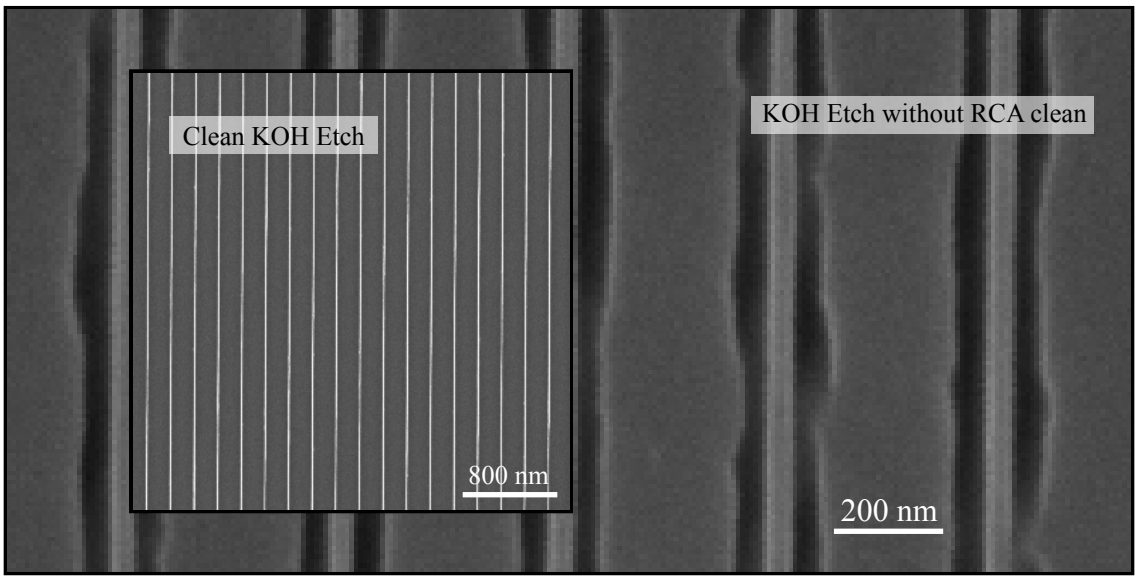

Figure 5.5: *The KOH etch can fail due to contamination. An RCA clean immediately before the KOH etch ensures a clean etch profile as seen in the inset figure.*

5.6 KOH Etch

This is the main focus of the fabrication process. Silicon crystal etches 100-200 times faster in the (100) direction than it does in the (110) direction. This is the basis for the anisotropic etch of silicon [62–68]. Mechanically, the etching procedure is straightforward:

Recipe 7 ***KOH Etch***

1. *RCA clean (all steps must be done in immediate succession)*
2. *HF dip*
3. *KOH etch*
4. *HF etch*
5. *DI water rinse*

All the steps above must be done in the immediate succession, without letting the sample surface dry.

RCA Clean

The first step is to clean the silicon piece in RCA (see recipe 1) for more than 5 minutes. This step is very important, and must be done every time before the KOH etch to ensure the

surface is clean. For example, a sample that was RCA cleaned and then sat in a plastic box in the clean room for two hours did not etch properly, as seen in figure 5.5. The figure inset shows the properly cleaned sample that was also KOH etched as part of the same batch. The difference is drastic, and is due to the RCA clean immediately prior to KOH etching.

HF-KOH-HF

The KOH etch is composed of three parts: (i) HF dip, (ii) KOH etch, and (iii) HF etch. The HF and KOH solutions are prepared as follows:

Recipe 8 ***HF***
130 mL DI water + 12 mL HF solution

Recipe 9 ***KOH***
110 mL DI water + 90 g KOH pellets,
set spinner to 120 rpm at 60 C

15 s HF dip: the purpose is two-fold. First, the lines formed in the oxide mask, as described above, are too thick and need to be shrunk. Roughly, 15 s of HF removes 5 nm of silicon dioxide, reducing the linewidth from 20 nm to 12 nm. Second, the HF removes the residual native oxide and hydrogen passivates the surface for the KOH etch.

KOH etch: the silicon piece must be immediately placed into the KOH solution after the HF dip. Prepare the KOH solution according to recipe 9, and submerge the sample with the pattern facing towards the center of the beaker. A spinner helps remove the bubbles–the byproducts of the KOH reaction–from the silicon surface, which yields a clean and smooth feature profile.

HF etch: the last HF etch also serves a two-fold purpose. First, the wet etch is hard to control and get the time to be exactly to-the-second. The second HF etch immediately stops the KOH reaction helping to prevent overetching. And second, the remaining oxide mask is removed by submerging the sample for about 1 minute in the HF solution ensuring that no silicon dioxide remains on the surface.

The process described above has been used to produce a Si template with fins of an extreme width:height ratio above 1:16 for widths below 10 nm [60]. The SEM images for such a template are shown in figure 5.6(a). Great care was taken to ensure fins would not collapse during the KOH etch, because if the oxide mask is misaligned with the crystal planes, the etching will follow the crystal orientation resulting in lines collapsing. To avoid this problem, the wafer was carefully oriented before the initial PMMA pattern was exposed. Furthermore, the KOH etch rate was measured (see figure 5.6) to ensure the fins would not be overetched.

(a) The Profile of the Etch Si Fins (b) The SI Etch Rate

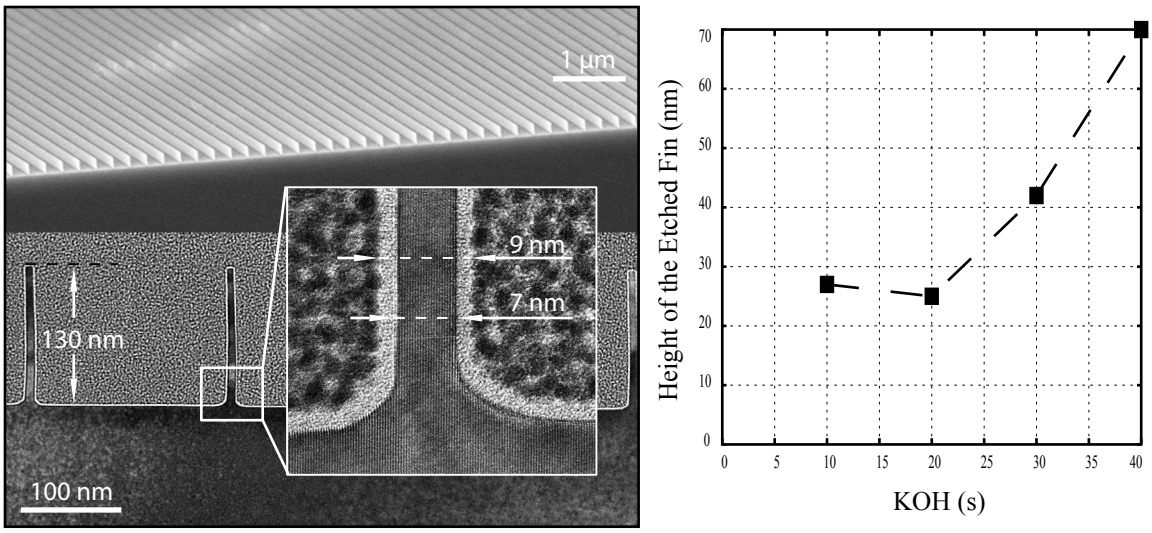

Figure 5.6: *Extreme aspect ratio fins of width:height above 1:16 have been produced. The profile of the etched Si fins using the anisotropic KOH etch is shown in (a). The Si etch rate measurement is plotted in (b). Reprinted with permission from Polyakov* et al. *[60] ©2011, American Vacuum Society.*

The impressive results achieved using the etch method presented here demonstrate the power of using the anisotropic etch of silicon to produce a template with fins of extreme dimensions. Note the vertical lines visible in the transmission electron microscope (TEM) image of the Si fin in figure 5.6, these are the Si lattice planes. They can be traced all the way from the base of the groove to the top; this illustrates the atomically smooth nature of the Si fins formed by this process.

5.7 Gold Evaporation

In this work, gold was the material of choice, which was evaporated onto the template. However, due to the high aspect ratio of the fins on the silicon template, gold evaporation requires extra care. If evaporated onto a stationary sample, gold tends to form a large grain at the tip of the fin, see figure 5.7(a). This grain shadows part of the fin producing large gaps along the fin all the way up to the surface of the evaporated gold film (see figure 5.7).

To reduce this effect, the sample was set at 45^o tilt with respect to the evaporation source in the electron beam evaporation setup. Gold was then evaporated at the 2 $\mathring{A}/s$ rate for the first 50 nm and then at 4 $\mathring{A}/s$ for the remaining 250 nm. Throughout the evaporation the sample was continuously rotated from -45^o to 45^o tilt. This dramatically decreased the

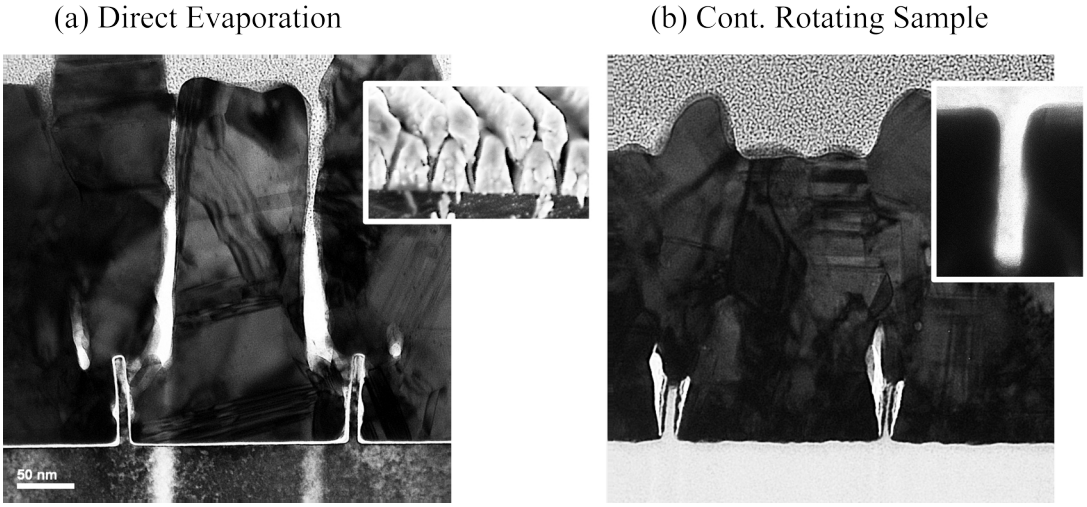

Figure 5.7: *Due to the high aspect ratio of the fins, metal evaporation requires continues rotation of the sample. If evaporated onto a stationary sample, gold forms a large grain at the tip of the fin that grows shadowing the fin itself (a). The inset in (a) shows the SEM image illustrating the effect of the grain forming at the tip of a fin. The gaps surrounding the fin completely change the geometry of the final grooves. One way to reduce this effect, is to continuously rotate the sample, this exacerbates the gaps at the base of the fin, but the metal layer above the fin forms without any gaps (b). The gaps at the base can be later closed up by slightly heating up the sample–see the inset in (b) showing the resulting groove profile in the metal surface.*

gaps forming to the top of the metal layer as seen in figure 5.7(b). However, at the base of the fin the gaps became larger. They, in turn, were closed up by heating the sample (with the evaporated gold film) to 100 C for 3 hours. The resulting shape of the groove is shown in the inset of figure 5.7(b).

5.8 The Peel-Off Process

With the gold evaporated onto the template, the metal grooves are formed. The remaining steps involve separating the silicon template from the gold film to reveal the final device. To do this an advantage is taken of the poor adhesion of gold to the silicon dioxide. A glass support plate is epoxied to the gold film and then the silicon piece is mechanically separated from the gold film in a peel-off motion.

The key to a successful peel-off is to prepare an ultra-clean glass support plate, because any contamination on the surface will weaken the epoxy bond between the glass and the gold layer. The cleaning procedure for the glass plate is described below. It is helpful to ensure that the gold evaporation is finished a minute or two before the final cleaning step. The idea is to take the glass plate from the ozone cleaning chamber and immediately epoxy

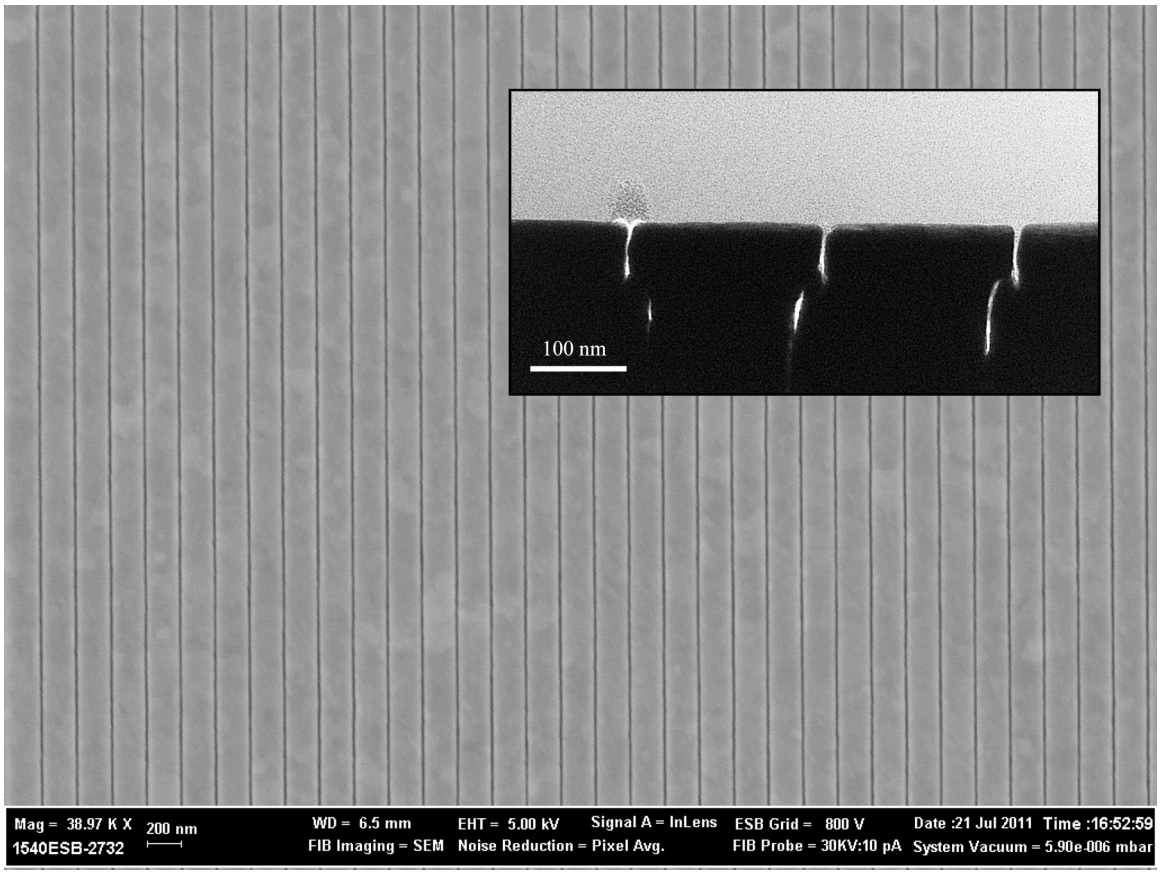

Figure 5.8: *The SEM image of the final device showing the deformation-free rectangular grooves formed on the gold surface. The TEM cross section showing minor gaps formed during the evaporation is shown in the figure inset.*

the sample to it.

Recipe 10 ***Glass Plate Cleaning***

1. *3 minute acetone sonification*
2. *IPA clean*
3. *10 minute RCA clean*
4. *IPA clean*
5. *10 minute UV ozone cleaning*

The sample is epoxied by placing the minimal amount of epoxy onto the gold film and pressing it onto the glass plate. In this step, a small portion of the silicon piece is left hanging

off the edge of the glass plate. Once the epoxy is cured, the silicon template is peeled-off from the gold film by prying onto this hanging-off portion of the silicon piece[4]. Figure 5.8 shows the top-down view of the final device as well as the cross section image. Notice that the peel-off process does not produce any deformation in the shape of the grooves over a large area.

5.9 Sample Contamination and Cleaning

Once the sample is prepared it must be kept in a nitrogen atmosphere to avoid forming a carbon contamination layer that will fill up the grooves changing the response of the subwavelength grating. If, however, the contamination layer does form, it can be cleaned off with the following procedure:

Recipe 11 ***Sample Cleaning***

1. *3 minute acetone sonification*
2. *IPA clean*
3. *10 minute ozone cleaning*
4. *3 minute acetone sonification*
5. *IPA clean*

However, it is important to note that acetone weakens the epoxy, which holds the gold film on the glass support plate. So this cleaning technique can be used no more than three-four times before the sample begins to form holes in the metal film and large-scale defects begin to form on the metal surface.

5.10 HSQ Template

The fabrication of the NGs on the metal surface has been described in the previous section where the fabrication method involved etching the template in the silicon wafer itself. Taking advantage of the anisotropic KOH etch of silicon, atomically smooth fins were prepared with extreme dimensions: 9 nm wide and over 130 nm tall. However, the challenge of preparing a Si template lies in the KOH etch, which is difficult to control. For this reason, an alternative template fabrication method was explored.

Another way to prepare a template, is to use the high resolution hydrogen silsesquioxane (HSQ) electron beam resist[5], which allows producing lines around 10 nm wide using single-pass electron beam writing. The fabrication steps are illustrated in figure 5.9: the template

[4]The orientation of the grooves with respect to the peel-off direction does not affect the result.

[5]The HSQ is prepared in low concentration diluted by the methyl isobutyl ketone (MIBK).

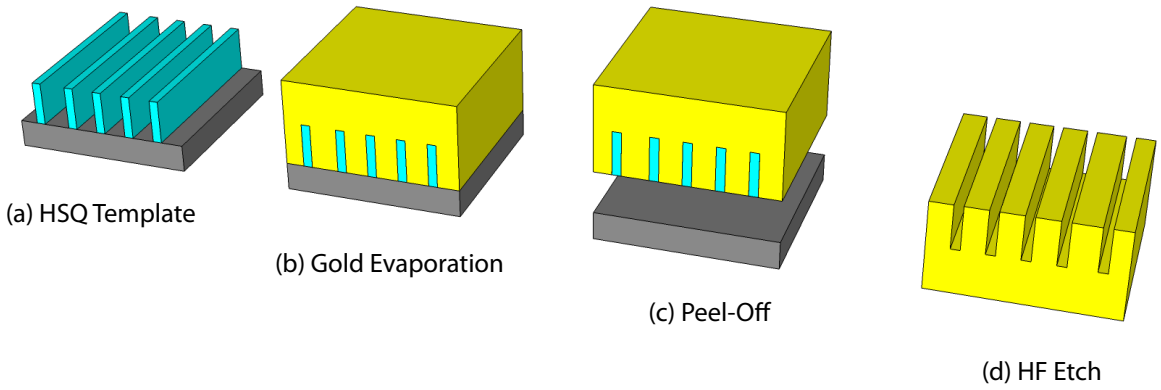

Figure 5.9: *The HSQ template fabrication method is based on preparing a line pattern using electron beam lithography on a plain Si wafer (a). Then metal is evaporated onto the template (b). Afterwards, the gold film is epoxied to a support plate and the Si wafer is peeled off (c). The resulting gold film contains grooves filled with HSQ, which can be removed in the final HF etch (d).*

is prepared by electron beam lithography, then the metal is evaporated onto the template, after that the silicon wafer is separated from the metal film and the remaining HSQ is etched from the inside of the grooves. Every step of this process is well controlled, which makes this process much more reliable than the Si etch, and substantially easier to follow. First, HSQ is spun on onto the prime Si wafer using the following recipe:

Recipe 12 ***HSQ Spin-On***

1. *10 minute RCA clean (see recipe 1)*
2. *10 minute pre-heat at* 180^o
3. *10 minute ozone clean*
4. *2% HSQ spin-on at 1000 rpm (yielding a 60 nm thick layer)*
5. *1 minute HSQ cure at* 100^o

Then, after the HSQ is developed, the metal is evaporated onto the template as described in section 5.7 and the underlying Si wafer is peeled-off following the directions in section 5.8. The result is a gold film epoxied onto a glass plate with a pattern of nano-grooves formed on the gold surface that are filled with the exposed HSQ[6]. The final step is to etch out the HSQ using the HF solution as given by recipe 8. This is done by placing a drop of HF from a plastic syringe onto the metal surface. The goal is to cover the pattern with the HF solution, while minimizing exposure of the underlying epoxy to the HF.

[6]Exposed HSQ is very similar to the silicon dioxide in terms of its optical properties.

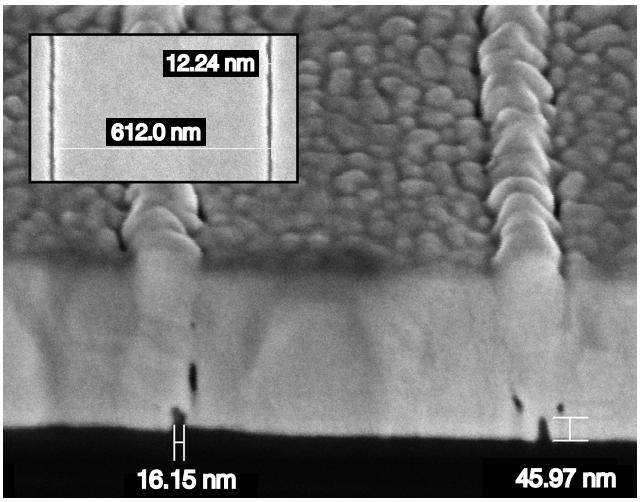

Figure 5.10: *Cross-section of the HSQ-template based device. The dimensions are close to the design values with two types of deviations: nano-voids due to shadowing during the metal evaporation, and a variation of line widths in the range between 12 nm and 16 nm. Reprinted with permission from Polyakov* et al. *[60] ©2011, American Vacuum Society.*

The result is a reliable fabrication process with good control over the final geometry of the grooves on the metal surface. An example cross section of a device prepared in this way is shown in figure 5.10. The final groove dimensions are within 2 nm from the designed values; an even better match in dimensions can be achieved by slightly varying the exposure during the electron beam lithography step, producing a few patterns with varying exposure doses.

5.11 Fabrication Process Comparison

The Si template fabrication technique described in this section has been used to produce sub 10 nm wide and over 100 nm deep grooves on the gold surface. The prime advantage of using the Si template is the ability to fabricate atomically smooth Si fins at extreme dimensions, which corresponds to an improved overall quality of the final sample.

An alternative fabrication process that is based on producing a template from the HSQ resist on the silicon wafer has an advantage of greater reliability in producing a sample with dimensions close to the designed values. Although HSQ-based fabrication cannot achieve the high edge quality of the Si fins, especially in terms of the edge sharpness and flatness, it is sufficient for producing a light trapping grating for the visible spectrum. However, producing an NG grating with resonance in the NIR spectrum requires narrow grooves, which cannot be made using the HSQ resist [60].

Both fabrication methods are well suited for producing grooves on the metal surface,

and must be chosen according to the application requirements. For the designs where the groove width exceeds 10 nm and the groove flatness is not critical the HSQ-based fabrication method is a better match. For other high-demand cases, the Si template-based approach is better suited.

Optical Properties of Plasmonic Nano Grooves

6

Light trapping in the visible spectrum via a subwavelength grating composed of rectangular nano-grooves (NGs) has been proposed theoretically by Le Perchec *et al.* [27]. The schematic for such a grating is shown in figure 6.2(a): on resonance, light polarized perpendicular to the orientation of the NGs (denoted p-polarization) excites a standing wave plasmon mode in a fashion such that the SPP wave travels up and down the NG producing the electric field pattern as illustrated in figure 6.2(b). This SPP mode has strongest electric field at the mouth of the groove and strongest magnetic field at the bottom. The field enhancement achieved at the mouth of the groove is advantageous for non-linear optical applications such as harmonic generation and multi-photon photoemission [69]. In this way, the NGs serve both as an SPP coupling mechanism, as well as the resonant cavities with strong field enhancement [51] for absorption at the resonant wavelength λ_{res}.

6.1 Imaging Setup

Reflectivity measurements were done by imaging the device illuminated by tunable monochromatic light with a camera and recording the image normalized to the s–polarized light, which cannot couple to the SPPs inside the grooves and therefore behaves as a plain gold surface [40]. The ratio between the s– and p–polarized light was measured experimental and compared to an FDTD theoretical model based on the measure optical constants [45]. For these measurements a halogen light source was used in combination with a monochromator (see figure 6.1) as a tunable randomly polarized 5 nm bandwidth light source. A Glan-Laser Calcite polarizer was used to select a desired polarization with respect to the orientation of the grooves on the sample surface. Finally, a broadband beam splitter was placed in front of the sample to allow normal-incidence reflectivity measurements using a camera to record a magnified image of the sample for each wavelength.

The difference in the illumination intensity between the s- and p-polarized light was removed by recording the signal R_{NG} normalized to the plain gold surface R_{FS} in the following way:

$$K = \frac{R_p^{NG}}{R_p^{FS}} \times \frac{R_s^{FS}}{R_s^{NG}}, \tag{6.1}$$

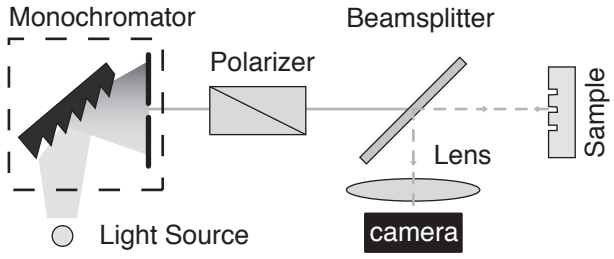

Figure 6.1: *The imaging setup used for the reflectivity measurements consists of a polarized light illuminating the device and the resulting image recorded by a camera. Reprinted with permission from Polyakov* et al. *[60] ©2011, American Vacuum Society.*

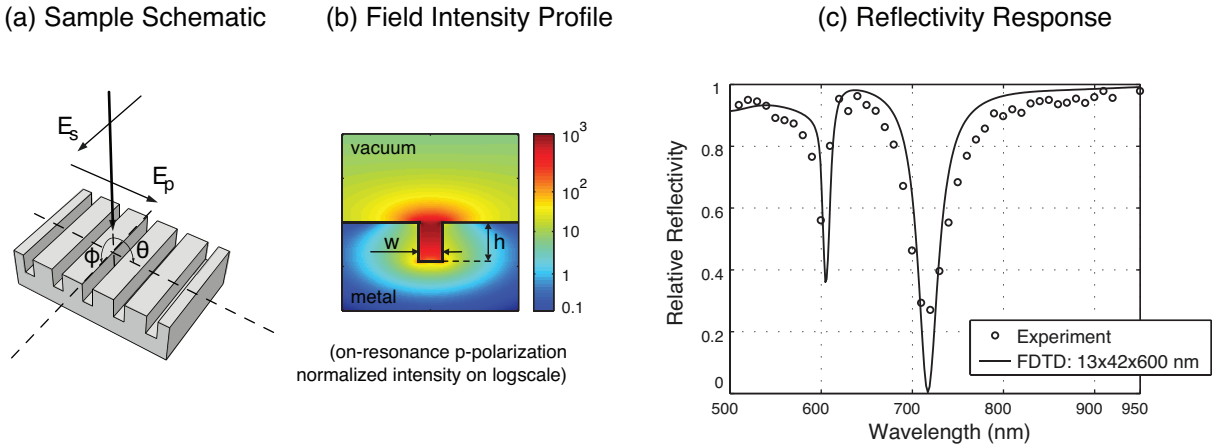

Figure 6.2: *The nanostructure is composed of a set of rectangular grooves (a), which provide for efficient coupling of p-polarized light to the standing SPP modes inside these resonant cavities yielding the field profile as illustrated in (b). This effect is readily observed by a direct normal-incidence reflectivity measurement for a gold grating (c). On resonance, the absorption is increased by 16 times relatively to the flat gold. The theoretical fit based on the FDTD numerical model shows good agreement with the experimental data.*

where K is the measured relative reflectivity.

Figure 6.2(c) shows the reflectivity spectrum of a plasmonic subwavelength grating, measured at normal incidence; the black solid curve shows the corresponding numerical FDTD calculation. The grating was designed by numerically optimizing the period p, width w and depth h of the NGs to produce a resonance in the UV-NIR region. In this case, for a grating made on a gold surface: $\lambda_{res} = 720$ nm, $w = 13$ nm, $h = 42$ nm, $p = 600$ nm. Two reflectivity dips are observed in the reflectivity spectrum shown in figure 6.2(c). The broader dip, at 720 nm, is due to the resonant SPP modes excited within the NGs, which yields a

16 times absorption increase compared to the plain gold surface. In contrast, the classical grating-coupled SPPs are excited on the top surface [18], in between the NGs. The narrow reflectivity dip at 600 nm corresponds to this classical grating-coupled SPP mode.

6.2 Maximizing Spectral Bandwidth

One of the advantages of the NG-coupled SPP modes is the increase in the spectral bandwidth. In section 4.2 it was shown that the triplet (p, w, h) is not unique, and the period can take any value between about 100 nm up to the resonance wavelength. Changing p, however, requires adjusting w and h to maintain the resonance condition at λ_{res}. The relationship between the period and the NG width is given by (k_0^R is the resonant wave vector and the $k'_{pl} + i \cdot k''_{pl}$ is the SPP wave vector):

$$\frac{p}{w} = \frac{1}{\pi} \frac{[k'_{pl}]^2}{k_0^R k''_{pl}}.$$

The case where $p \approx \lambda_{res}$ is the classical grating case, while the NG-coupled grating has $p < \lambda_{res}$. Choosing $p \ll \lambda_{res}$ yields a plasmonic grating with a larger spectral bandwidth. For a case study we prepared an aluminum grating with NG dimensions of w = 15 nm, h = 60 nm, and p = 100 nm resonant at λ_{res} = 720 nm. The fabrication process has been described in the previous chapter–see figure 6.3(a) for reference; it consists of four steps: (i) preparation of a template on a silicon wafer with an HSQ electron beam resist pattern in the form of lines, (ii) evaporation of a thin gold layer ($\approx$ 4 nm thick) to reduce the adhesion of aluminum to the silicon wafer, (ii) evaporation of the target metal, aluminum in this case, onto the template, and (iv) separation of the silicon wafer from the metal-HSQ structure by epoxying a glass plate onto the metal side and peeling off the silicon wafer to reveal the final device–see figure 6.3(b) for the schematic of the final device as used for the FDTD model. To keep the NG dimensions below 20 nm in width the amount of the evaporated gold was minimized to a thickness where gold would not form a continuous film. Consequently, upon separation, the exposed aluminum formed a stable oxide layer as indicated in figure 6.3(d). The TEM cross section image of the final grating is shown in figure 6.3(d).

The presence of the HSQ filler in the groove slightly expands the spectral bandwidth, but is a minor effect. Good agreement is seen between the FDTD model and the experimental data: this device, fabricated with $p \ll \lambda_{res}$, is over 98% absorbing at the target resonance wavelength and has an extreme spectral bandwidth, greater than 350 nm, as shown in figure 6.3(c). The deviation in the absolute absorption between the FDTD model and the experiment is attributed to the edge rounding of the actual NG geometry (as seen on the TEM image) compared to the perfectly square shape used in the FDTD model.

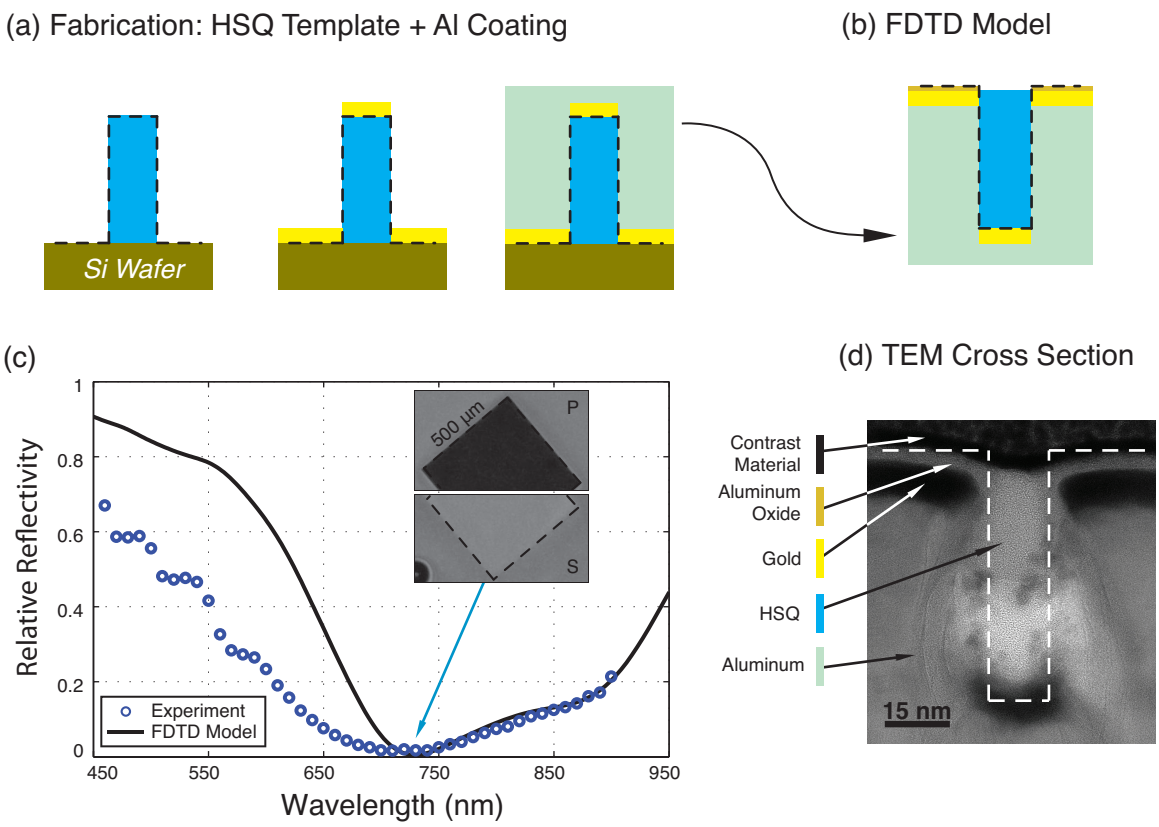

Figure 6.3: *Fabrication and reflectivity response of a grating designed for the maximum spectral bandwidth. The fabrication process starts with electron beam patterned HSQ resist producing a template for the final device; a thin layer of gold is then deposited to decrease the adhesion between the target metal and the silicon wafer; finally, the target metal (aluminum) is deposited onto the template. Separating the resulting metal structure from the template reveals the final grating on the metal surface. The reflectivity response was measured by recording the intensity change between the s- and p-polarized light (with respect to the NG orientation) at normal incidence (see figure 6.2 for the definition of the polarization convention used); the superimposed optical image of the pattern on the metal substrate is shown in the inset of (c), where the bottom was recorded at s-polarization showing no plasmon coupling, and the top was recorded for p-polarized light showing the dramatic decrease in the reflected intensity due to the plasmon coupling inside the NGs. The cross section of this grating was measured with the TEM (d), showing the layered structure. The corresponding schematic used for the FDTD model is shown in (b).*

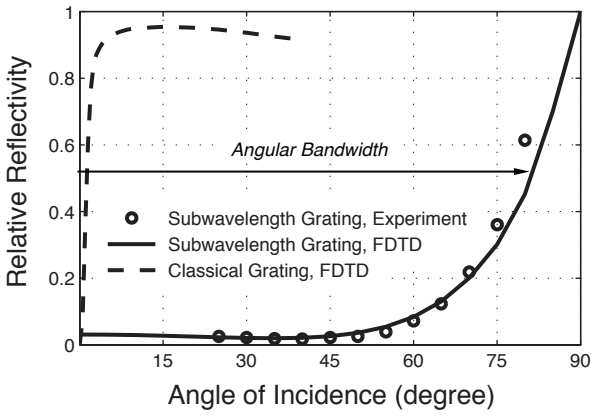

Figure 6.4: *The aluminum grating was designed for maximum angular bandwidth, the angle at which the on-resonance reflectivity increases to 50%. The experimental measurement is shown with circles and the corresponding FDTD model with a solid line. The dashed line shows the response of a classical grating coupled system at the same resonance wavelength. The advantage of the subwavelength grating is the ability to design a light trapping substrate with an angular bandwidth from a few degrees up to 80 degrees, which corresponds to a full opening angle of 160 degrees.*

6.3 Maximizing Angular Bandwidth

For photovoltaic applications it is desirable to produce large spectral bandwidth light trapping substrates [70], while for other applications, such as sensors [44], narrower bandwidth is preferred. Both types of response can be achieved by adjusting the p/λ_{res} ratio. However, another important aspect of light trapping is the angular bandwidth–the angle of incidence at which the on-resonance reflectivity increases to 50%. The aluminum grating presented above is a demonstration of a device where the angular bandwidth is maximized, due to $p \ll \lambda_{res}$. The on-resonance relative reflectivity dependance on the angle of incidence for the aluminum grating measured at 735 nm is shown in figure 6.4 by circles with the corresponding FDTD model plotted as a solid line. For comparison, the response of the classical grating coupled system is shown by a dashed line exhibiting a characteristic angular bandwidth of ≈ 3 degrees. Reducing the period with respect to the resonance wavelength increases the angular bandwidth. However, due to the the evanescent wave of the SPP mode inside the NG extending into the metal, the period of the grating must be chosen to avoid an overlap between the SPP modes from the neighboring NGs. In varying the period of the grating from the resonance wavelength λ_{res} down to about 100 nm, the angular bandwidth can be designed from a few degrees up to 80 degrees corresponding to a 160 degree opening angle.

6.4 Resonance Tuning

The aluminum grating presented above was designed for maximum absorption at the target wavelength with maximum spectral and angular bandwidths. However, an NG grating can also be designed with a narrow spectral bandwidth such as the gold gratings shown in figure 6.2(c). Such gratings, composed of narrow NGs on the scale of the metal skin depth are extremely sensitive to the presence of any dielectric on the metal surface. A small amount– less than a nanometer thick layer–of a dielectric on the surface shifts the resonance position by more than 10 nm. The dielectric layer on the metal surface changes the boundary condition for the SPP dispersion [71], shifting the resonance wavelength for the SPP standing wave excited in the NG. This effect was successfully used in the fabrication process to test for the presence of surface contaminants on the gold grating discussed above. The fabrication of the gold grating is identical to the aluminum grating described above, with the additional step of submerging the structure in HF to remove the HSQ from inside the NGs. A properly cleaned structure–rinsed in acetone and IPA–had a resonance at $\lambda_{res} = 705$ nm. If, however, a thin residue layer formed during the final cleaning, the resonance position would shift by 15 nm to $\lambda_{res} = 720$ nm. Due to the nanometer dimensions of the NGs, even a few atomic layers of a dielectric result in an appreciable shift in the resonance wavelength.

This sensitivity of the plasmonic grating to the presence of a dielectric coating can also be effectively used to tune the grating in post-fabrication. A well-controlled thickness of the dielectric coating can be conformally deposited onto the structure by atomic layer deposition (ALD) growth, yielding a high-quality conformal coating [72–74]. For this study we employed a room temperature alumina (Al_2O_3) plasma activated ALD deposition (25 C to avoid deterioration of the NG shape on the gold surface) using the Oxford Instruments FlexAL reactor. The process consisted of the following steps: (1) 30 ms exposure to trimethylaluminium (TMA) dose at 15 mTorr without any oxygen flow; (2) followed by 1 s TMA purge and (3) oxygen inflow at 60 sccm for another 1.5 s; (4) with the final 2 s 400 W oxygen plasma oxidation at 15 mTorr with oxygen flow maintained at 60 sccm. Throughout the entire process argon gas flow was set at 100 sccm.

The TEM cross section view of the completely filled NG is shown in figure 6.5(b). The top-surface alumina coating has uniform density and appears of consistent shade, while the alumina filling of the NG varies greatly in shade corresponding to the density variation of the alumina filling: as the dimensions of the cavity approach values less than 10 nm in width, the alumina growth is no longer uniform and instead forms a less dense material. A model based on this consists of approximating the non-uniformity of the alumina growth inside the NGs as an exponential decay in the index of refraction of the alumina filling starting from the nominal refractive index of $n_{ox} = 1.6$ to the reduced refractive index at the center of the NG n_c, which is treated as a fit parameter. Additionally, the growth inside the NGs is modeled to be twice that on the surface due to the less dense nature of the alumina filling inside the NG. The mathematical model for the alumina index of refraction is then given by (d is the distance away from the center of the groove, and t is the exponential decay length;

figure 6.5(b) shows the schematic used for the FDTD model):

$$n(d) = (n_{ox} - n_c)\exp[-d/t] + n_c.$$

Two gold gratings of closely matched dimensions were prepared to measure the dielectric-based tuning: sample A with $w = 13$ nm, $h = 42$ nm, $p = 600$ nm, and sample B with $w = 15$ nm, $h = 42$ nm, $p = 600$ nm. For sample A the alumina was grown via 2, 4, or 8 ALD cycles per session, with sample taken to air after each session to measure the resonance wavelength. Blue circles in figure 6.5(a) correspond to the measured resonance position; the solid line is the FDTD model fit based on the dielectric constant as described above (the steps in the theory curve are due to the finite size of the mesh used). When the sample was taken out into air, a contamination layer would form that resulted in the overall alumina growth to be less dense yielding $n_c = 1.2$ and $t = 2.5$ nm. The growth rate on the top surface was measured to be 1.1 Å/cycle based on the TEM picture shown in figure 6.5(b). The growth inside the NG is modeled to be twice that amount (2.2 Å/cycle) as discussed above in good agreement with the TEM cross section showing the light region of the alumina in the center of the NG.

For sample B, the alumina was grown to the target thickness all in one session starting from a completely cleaned surface. After measuring the resonance position at the target alumina thickness, the sample was cleaned using the combination of the HF, acetone, and IPA, and the resonance position of the cleaned sample was consistently confirmed to be $\lambda_0^{res} = 705 \pm 2$ nm before growing the new target thickness of the alumina layer. This way, a total target thickness of alumina was deposited onto the clean structure without any contamination. This process was repeated nine times and the result is plotted in figure 6.5(a) (black squares) with the corresponding FDTD model. Alumina grown this way yielded a more dense coating with the FDTD model fit parameters of $n_c = 1.4$ and $t = 4$ nm. Unlike sample A, sample B was not exposed to air in between the successive layers, therefore the alumina growth rate is slightly higher, 1.25 Å/cycle on the top surface, or 2.5 Å/cycle inside the NG, which is consistent with growth rates reported for alumina ALD at room temperate under similar conditions [74–76].

Both samples exhibit the same tuning behavior, in good agreement with the FDTD: in the region where the NG is not completely filled, sub nanometer size increase in the thickness of the dielectric layer partially filling the NG produces a large shift in the resonance position. In this regime, the resonance can be tuned anywhere up to the NIR region. Once the NG is filled, the tuning saturates. The thickness of the alumina layer at this point is half the NG width, this critical thickness is marked with vertical dashed lines in figure 6.5(a). Any continued alumina growth on the surface is equivalent to small decrease in the periodicity of the subwavelength grating yielding only minor further shift in the resonance wavelength.

The flexibility offered by the dielectric-based tuning is useful beyond post-fabrication resonance adjustment. Preparing a plasmonic grating resonant far into the NIR is a major fabrication challenge [60], which is solved by designing a grating for resonance towards the red part of the visible spectrum and later converting it into the NIR grating by controllable

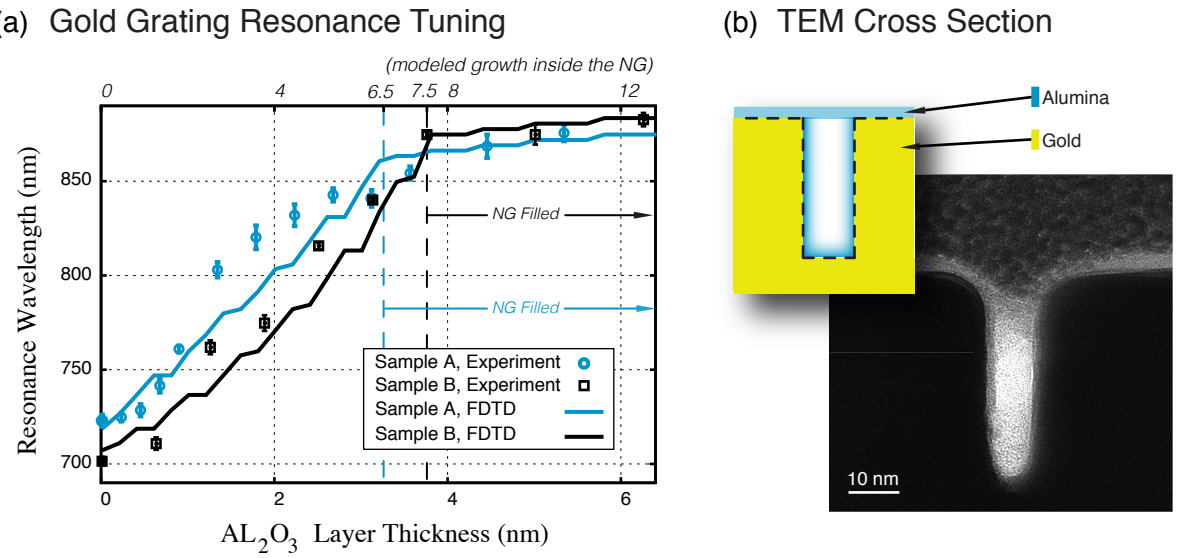

Figure 6.5: *Tuning the resonance wavelength of a plasmonic gold grating by ALD conformal growth of nanometer thick layers of alumina. Two gold gratings, samples A and B, were prepared, both with cross sections as shown in (b). For sample A, the alumina growth was incremental, with the sample exposed to air in between each successive session. For sample B, the alumina was grown in one session from 0 nm to the target thickness, then alumina completely removed returning the resonance to the original wavelength. And the entire process repeated. Both samples are tuned from from the resonance position up to ≈ 860 nm by ALD alumina growth due to the sensitivity of the SPP modes inside the NGs to the sub nanometer thick layers of a dielectric. This effect saturates once the NG is completely filled, this critical thickness is marked by dashed vertical lines in (a).*

growth of a thin dielectric coating. This technique is completely reversible: using HF the alumina is readily removed from the gold grating and the initial response is restored. This effect opens the possibility of using such gratings as a sensor for probing sub-nanometer thick layers of a dielectric.

The design presented here and the post-fabrication tuning of a subwavelength plasmonic metallic grating allow for complete control over the resonance wavelength, the absorption strength, the spectral and angular bandwidths. The aluminum grating presented was designed for an extreme spectral bandwidth of above 350 nm, and an angular bandwidth of 80 degrees with 98% absorption at the resonance wavelength of 720 nm. The ultimate control over the resonance position was then demonstrated in the study of the gold gratings that were tuned from the resonance wavelength of 720 nm up to 860 nm by ALD growth of a few nanometer thick layers of alumina. This coating also serves as a protective layer for the metal structure, allowing fabrication of highly controlled resonant light trapping metal surface with resistance to environmental contaminants. Both types exhibit optical characteristics in good

agreement with the FDTD model presented.

Starting from the model for the NG on-resonance behavior presented in section 4.2 a grating can be designed with resonance anywhere from the UV to the NIR and have designer spectral and angular bandwidths. The fine-tuning of the resonance position is done later in post-fabrication by ALD growth on nanometer thick layers of a dielectric. The optical properties of the NG grating are fully customizable allowing the preparation of a robust light absorbing substrate for maximizing field enhancement, useful for non-linear processes. This enhancement is the subject of the next chapter, where experimental measurements are presented demonstrating the increase in the non-linear photoemission due to the SPP coupling in the NG grating.

Enhanced Multi-Photon Photoemission

7

When a metal is illuminated, electrons will come out from the surface provided that the photon energy exceeds the metal's work function. The ejected electrons can be focused using a set of accelerating magnetic or electric lenses to produce an image at a fluorescent screen. This image can then be captured by a camera to show optically the electron emission pattern at the metal surface. This is the principle behind the Photoemission Electron Microscope (PEEM) that was used to investigate the photoemission from the plasmonic grating.

The range of energies where the SPP modes can be excited on the metal surface were defined in chapter 4, see figure 4.3 on page 43. For gold, the maximum accessible energy in the SPP dispersion curve is $\approx 2.4\ eV$ corresponding to $\lambda = 515$ nm. Any wavelength shorter than this will not excite surface plasmons on the gold surface [18]. For this reason, a longer wavelength, the fundamental of the Ti:sapphire laser was chosen at 800 nm (1.55 eV). However, gold's work function is 5.6 eV [17], so only fourth order photoemission is possible from gold at this wavelength. That is, four photons must combine together to eject on electron. The probability of such process at low intensities is extremely low. As the intensity is increased, the emitted charge increase non-linearly [20], which makes this setup an excellent choice for investigating the photoemission enhancement due to the plasmonic coupling in the NGs that provides reflectivity reduction and strong field enhancement[1].

7.1 PEEM Measurement

The PEEM measurements were carried out at Pacific Northwest National Laboratory (PNNL) using two light sources: the Ti:Sapphire laser with parameters given in table 7.1 and an Energetiq EQ1500 plasma driven UV source. The laser was aligned at 75 degrees angle of incidence with polarization controlled by a half wave plate. First, the sample was illuminated with both light sources showing the outline of the pattern on the metal substrate as in figure 7.1(a). Then, the UV light source was turned off, leaving only the laser illuminating the sample in p-polarization (E-vector oriented perpendicular to the grooves). In this case, photocurrent was observed only at the pattern and not at the flat surface. In p-polarization,

[1] See chapter 1 for the full description of the photocurrent dependance on the light intensity.

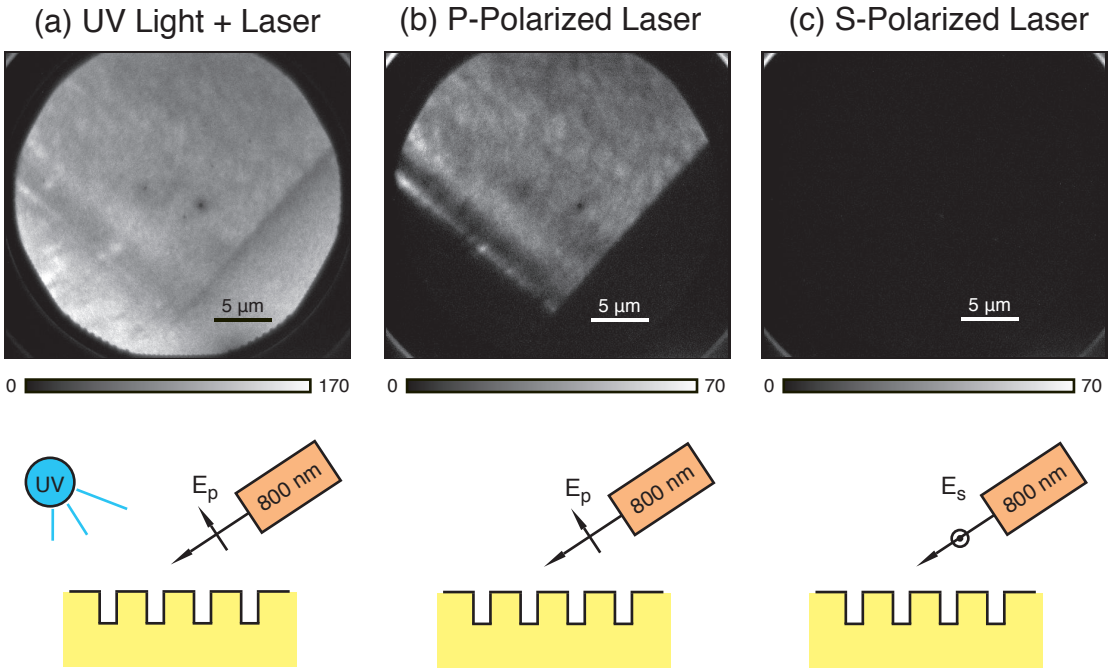

Figure 7.1: *The photocurrent emitted from the NG pattern was imaging in PEEM. Under UV illumination in combination with the laser (a) the photoemission is seen from the entire surface including the flat metal substrate, however, once the UV lamp is turned off, the photocurrent is seen only from the pattern (b). The contrast between the dark and the bright areas visually represents the photoemission enhancement due to the SPP coupling for p-polarized light. For s-polarization (c), the photoemission from the pattern is identical to that from the substrate, which is below the noise level and thus appears equally dark. This experiment demonstrates the dramatic enhancement in photoemission due to the SPP coupling in the NGs.*

light couples to the SPP modes inside the NGs, resulting in an increase in the optical absorption as well as field enhancement, which all lead to a dramatic increase in the amount of the emitted charge, the photocurrent. This enhancement is illustrated as the contrast between the dark background and the light pattern in figure 7.1(b). Due to the low probability of the fourth-order photoemission, the signal from the flat surface was below the noise limit, however, the large enhancement provided by the NGs made the pattern appear very bright.

In s-polarization, no SPP coupling occurs, and the response in the pattern is identical to that of the flat surface 7.1(c). In this case, the signal from both the pattern and the substrate were below the noise limit. This experiment demonstrates the dramatic improvement in the multi-photon photoemission due to the NG grating on the metal surface.

	λ	f_{laser}	τ_{pulse} (meas. #1)	τ_{pulse} (meas. #2)	d_{spot}	nJ/pulse
PNNL	800 nm	90 MHz	60 fs	–	100 μm	3.3
LBNL	805 nm	62 MHz	43 fs	60 fs	160 μm	0.2-3.2

Table 7.1: *The laser parameters for the PNNL and LBNL experiments. The LBNL experiment consisted of two measurements with the pulse length stretched by 40% from 43 fs to 60 fs to measure the intensity dependance curve.*

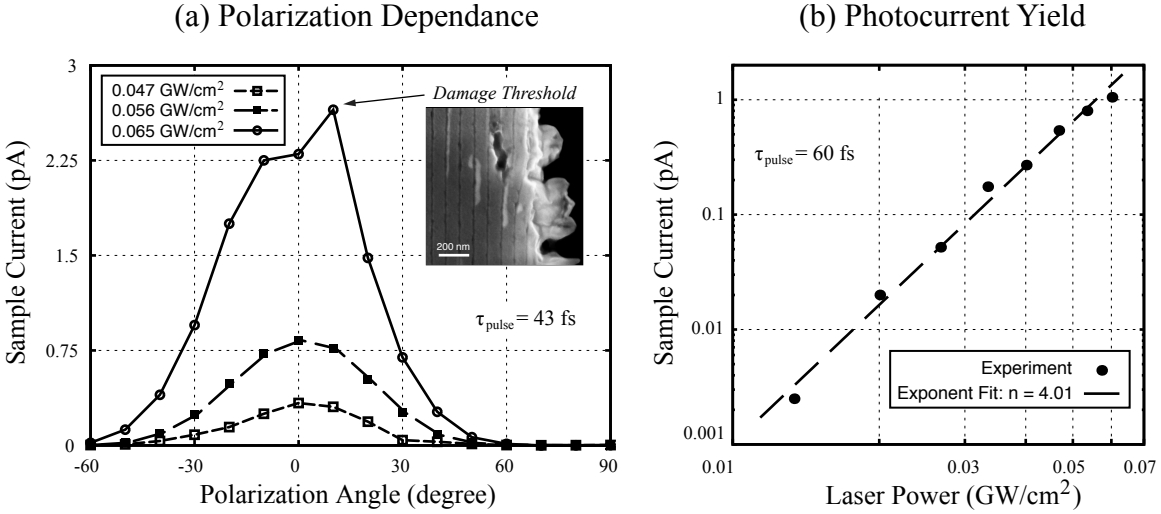

Figure 7.2: *The photocurrent as a function of the polarization angle is shown in (a), where p-polarization is set at 0 degrees and s-polarization at 90. Numerically, the photoemission enhancement due to the SPP coupling is the ratio between the maximum and minimum measured photocurrent for a given light intensity at the cathode. In (b), the pulse length was stretched by 40% to reduce the peak light intensity allowing resolution of the order of the multi-photon photoemission: n = 4.01. However, the four-photon photoemission process on the flat surface is highly improbable, which results in the s-polarization photocurrent values to be below the noise level. However, the dramatic enhancement due to the excitation of the SPP modes in the NGs improved the photoemission by seven orders of magnitude, sufficient for a direct photocurrent measurement.*

7.2 Photoemission and Field Enhancement Measurement

To quantify the photoemission enhancement a separate experiment was done at LBNL, where the photoemission was measured by focusing a Ti:Sapphire laser onto a 160 μm spot on the metal surface (the laser parameters for this experiment are listed in table 7.1). In the first part of this experiment, the sample was placed in a vacuum chamber with a laser illuminating the sample surface at normal incidence through a viewport. The laser pulse length was 43

fs with the polarization controlled via a half wave plate. The results are shown in figure 7.2(a), where the photocurrent is plotted as a function of the polarization angle with 0 set to p-polarization. At 90 degrees (s-polarization) the SPP coupling is forbidden resulting in the sample response identical to that of the flat surface, consistent with the PEEM measurements presented in the previous section. However, as the polarization is rotate towards 0 degrees (p-polarization), the incident light begins to excite the SPP modes that yield a large increase in the photoemission leading to a measurable photocurrent. The measurement was repeated three times at different light intensities at the cathode given by ($A_{spot} = \frac{\pi}{4} d_{spot}^2$):

$$I_{cathode} = \frac{P_{ave}}{f_{laser} \tau_{pulse} A_{spot}} \tag{7.1}$$

At the critical light intensity of $\approx 0.06\ GW/cm^2$, the epoxy that holds the metal film to the support plate began to burst due to the low heat conduction. However, the NGs themselves were not damaged: the NGs were imaged in the SEM confirming their structural integrity (see the inset in figure 7.2). The highest photocurrent measured at the pattern was 2.45 pA at $0.065\ GW/cm^2$. In s-polarization, due to the high sample reflectivity, a much higher laser intensity could be used without damaging the sample, yet the photocurrent was still below the noise level. The maximum estimate on the photocurrent from the flat surface is the noise level of 2.5 fA at $0.65\ GW/cm^2$.

Using equation 1.1 on page 5 the photoemission enhancement can be defined at the photocurrent i per intensity to the fourth power I^4 for the four-photon photoemission process:

$$\eta \equiv \frac{i_{NG}/I_{NG}^4}{i_{FS}/I_{FS}^4} = \frac{2.45\ pA\ /\ (0.065\ GW/cm^2)^4}{2.5\ fA\ /\ (0.65\ GW/cm^2)^4} = 10^7.$$

The plasmonic subwavelength grating gains *seven orders of magnitude photoemission enhancement* due to the reflectivity reduction and field enhancement in the NGs on the metal surface. To verify the order of the photoemission process, a separate measurement was done, where the pulse length was stretched by 40% to allow more data points below the damage threshold. The straight line fit is shown on the the log-log plot in figure 7.2(b), which yield an exponent of $n = 4.01$, confirming the fourth order photoemission.

The field intensity given by equation 7.1 represents the average light intensity over the entire beam spot. However, in the NGs, the field at the top is enhanced yielding an increase in the intensity. The photocurrent in this case is found by rewriting equation 1.1 as

$$i_n = C_n \left(1 - R\right)^n I^n F E^n\ A_{spot},$$

where FE is the field enhancement factor, which is 1 for the plain surface. The minimal estimate on the field enhancement is then given by

$$\begin{aligned} \frac{i_{FS}\ /\ I_{FS}}{i_{NG}\ /\ I_{NG}} &= \left(\tfrac{1-R_{FS}}{1-R_{NG}}\right)^4 \tfrac{1}{FE^4}, \\ FE &= \eta^{1/4}\ \left(\tfrac{1-R_{FS}}{1-R_{NG}}\right). \end{aligned}$$

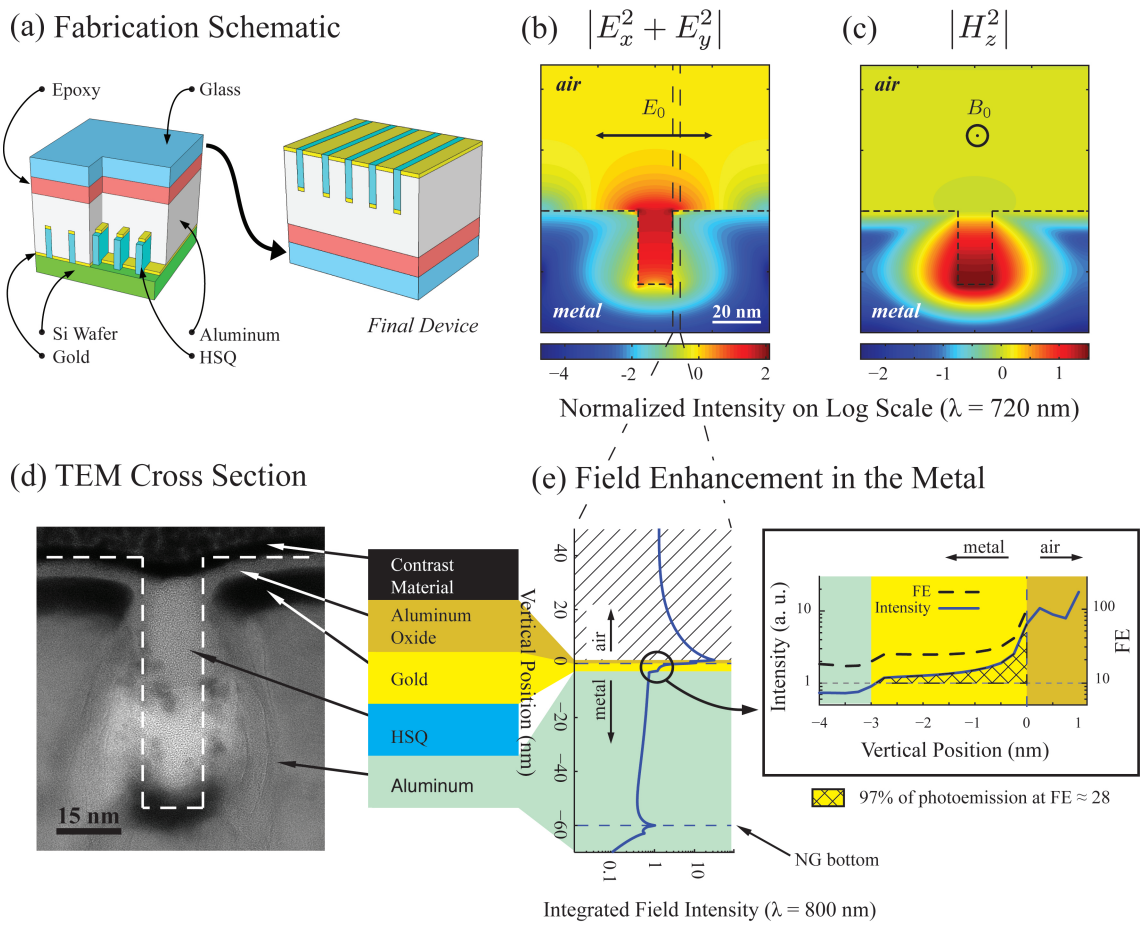

Figure 7.3: *The fabrication of NGs on the metal surface is illustrated in (a). In p-polarization a standing SPP wave is excited in the NGs, with the field profiles as shown ion (b) and (c), the TEM picture of the final device is shown in (d). The field intensity at 800 nm, is shown in (e), illustrating the field enhancement inside the metal is about 28 compared to a much larger field enhancement in the HSQ region.*

The sample reflectivity at 800 nm on the plain surface was $R_{FS} = 87\%$, while on the pattern in p-polarization $R_{NG} = 10\%$. Plugging these values into the above expression gives $FE = 8$. For the theoretical comparison the FDTD model was used to calculate the field enhancement. Figure 7.3(b) illustrates the field intensity profile at 800 nm inside the NGs. The highest field intensity occurs in the HSQ region, where there is no metal. Therefore, the field that participates in the photoemission is concentrated in the narrow region around the metal edge is indicated by the dashed lines. The SPP field drops sharply in both the x and y directions, with 97% of all photoelectrons coming from the area within 2 nm away from the edge and only 3 nm down into the NG. The field intensity and the field enhancement in this region are illustrated in figure 7.3(e). In this region FE varies slowly and is ≈ 28.

The experimental measurement of the FE is close to the FDTD model. A much larger enhancement could be achieved for a similar geometry by choosing a larger period. In this case, however, the small period of 100 nm was chosen for its benefits of a wide angular bandwidth, that is required for the PEEM measurements due to the fixed 75 degree angle of incidence of the laser beam.

The experimental results presented in this chapter confirm that placing a subwavelength array of narrow grooves on the metal surface can be used to dramatically increase the photoemission. The seven orders of magnitude enhancement of the fourth order photoemission demonstrates the dramatic improvement in light trapping and field enhancement on the metal surface. This field enhancement factor can be controlled to fit the application from low FE for a short NG period (as in this case to widen the angular bandwidth) to FE $>$ 300 for a longer NG period. With the benefits of designer angular and spectral bandwidth, the NG grating provides the greatest flexibility of light trapping on the metal surface for enhancing non-linear processes such as the multi-photon photoemission.

Plasmon Enhanced Photocathode

8

A novel approach for increasing metal's quantum efficiency has ben demonstrated experimentally in the previous chapter. By introducing a set of narrow grooves on the metal surface arranged in a subwavelength grating, the four-photon photoemission has been increased by *seven orders of magnitude*. Although under low intensities multi-photon photoemission is negligible, in a regime where the light intensity at the cathode is above 10 GW/cm^2 the multi-photon photoemission is the dominant process [20]. Such high intensities are accessible via femtosecond pulses, that lead to the development of electron beams that have a natural linear space-charge fields. This results in a correlation between the electron energy and position in the electron bunch allowing beam propagation without increasing the emmitance. Such behavior is critical for optimum X-ray FEL operation[14, 15], and could be useful in time dependent electron diffraction. In combination, the plasmonic grating on the metal surface provides the right environment for dramatic photoemission enhancement from a metal photocathode.

In the experimental demonstration of the photoemission enhancement presented in the previous chapter, the maximum photocurrent was limited by the poor heat conduction of the epoxy that held the metal film attached to the glass support plate. At the critical intensity of 0.06 GW/cm^2 the epoxy began to melt forming hot bubbles that burst rapturing the metal film. Figure 7.2(a) shows the photocurrent as a function of the laser polarization. As the polarization was rotated towards the p-polarization most of the energy was coupled into the NGs and absorbed by the metal film, heating up the underlying epoxy. The damage is illustrated by the SEM image in figure 7.2(a): NGs appear to suffer no thermal damage themselves. Therefore, the only limitation in terms of the incident laser power is the epoxy used in the fabrication process. This limitation can be readily overcome by using a heat-conductive epoxy that, once cured, does not contain any polymer component.

The four-photon photoemission enhancement can be extended to higher laser intensities. In this experiment, for $I_{ave} = 0.065\ GW/cm^2$ intensity at the cathode the pulse charge measured was $Q_{ave} = 3.63 \times 10^{-20}$ C (roughly 1 electron per ten pulses). In the four-photon photoemission the pulse charge increases as the fourth power of the laser intensity at the cathode. Therefore, to produce 0.3 nC per pulse, as required for the FEL operation [77], the

intensity at the cathode has to be [78]:

$$I_{FEL} = I_{ave} \left(\frac{0.3 \times 10^{-9}}{3.63 \times 10^{-20}} \right)^{\frac{1}{4}} = 20 \ GW/cm^2.$$

Such intensity is readily achieved by a commercially available Ti:sapphire lasers with an average power of 1 W at 1 MHz repetition rate for 30 fs pulses focused to the same spot size of 160 μm diameter. Such ultra-short laser pulse has a finite bandwidth due to the uncertainty principle. This is called the time-bandwidth product given by ($\Delta\lambda$, Δf, and Δt refer to the FWHM measurements):

$$\Delta f \Delta t = 0.44 \Rightarrow \Delta\lambda\Delta t = \lambda_c^2/c \tag{8.1}$$

A 10 fs pulse would have a 210 nm bandwidth, while a 30 fs pulse would have a 70 nm bandwidth. The large spectral bandwidth grating presented in the previous chapter can accomodate the shortest pulses.

For high brightness electron beams the electron-electron interaction plays an important role: the emission of electrons creates a screening effect, where the electrons towards the end of the pulse see the field of the electrons that were emitted before them. This is called the space charge limit. The critical charge that can be emitted is determined by the RF field that is used to pull out the charge, which is given by the Child-Langmuir law [11]:

$$Q_{max} = E_{gun}^{rf} \ A_{spot} \epsilon_0.$$

Although the emission comes from the NGs, which occupy only a part of the beam spot, due to the electron-electron interaction, the emitted electrons will rapidly expand to form a cloud that covers the entire area, which is then space-charge limited. The beam spot can be expressed as a function of the emitted charge as follows:

$$A_{spot} = \frac{Q_{max}}{E_{gun}^{rf} \epsilon_0}.$$

The above expression can then be used to express the maximum laser intensity at the cathode as a function of the emitted charge at the space charge limit:

$$I_{max} = \frac{P_{laser}}{\tau_{pulse} f_{laser} A_{spot}} = \frac{P_{laser} \ E_{gun}^{rf} \epsilon_0}{\tau_{pulse} f_{laser} Q_{max}}.$$

Based on the experimental measurements presented in the previous chapter, for Q_{ave} emitted charge at I_{ave} laser intensity at the cathode, the total emitted charge is expressed as a function of the laser intensity by

$$Q_{max} = Q_{ave} \left(\frac{I_{max}}{I_{ave}} \right)^4.$$

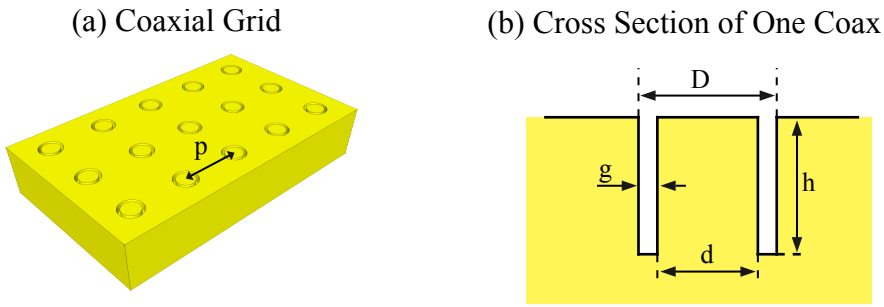

Figure 8.1: *A coaxial set of holes on a grid is a 3D extension of the plasmonic grating. In this case, the gap between the core and the surrounding metal is equivalent to the NG width. Adjusting the gap to period ratio would yield high field enhancement and control over the spectral bandwidth as well as angular bandwidth.*

Combining the above expressions and solving for the Q_{max} yields:

$$Q_{max} = Q_{ave}^{1/5} \left(\frac{P_{laser}\ E_{gun}^{rf}\ \epsilon_0}{\tau_{pulse}\ f_{laser}\ I_{ave}} \right)^{4/5} . \tag{8.2}$$

For a high-power laser as mentioned above, in a 120 MV/m RF field in the gun the maximum pulse charge within the space charge limit is 0.2 nC from a 500 μm diameter spot illuminated with a 17.7 GW/cm^2 laser intensity. The intensity calculated for an average field at the metal surface is amplified due to the field enhancement in the NGs. However, this field enhancement can be controlled to be within a factor of 100, which puts the maximum field intensity at below the damage threshold [20].

This pulse charge of 0.2 nC is very close to the FEL requirement [77] of 0.3 nC. In fact the same laser with a pulse compressed down to 15 fs would yield a 0.35 nC of charge per pulse. For comparison, upconverting the 1 W laser to the UV light with 10% efficiency would yield only[1] 4.3 pC at 1 MHz repetition rate. By taking advantage of the high intensities provided by the ultra-short laser pulses, a plasmon enhanced photocathode yields a sufficient pulse charge for FEL operation at a MHz repetition rate.

The plasmonic grating can be extended to 3D, where the resonant cavity would take the coaxial shape as shown in figure 8.1. The coaxial shape is analogous to the infinite grooves. The gap between the core and the surrounding metal would be equivalent to the width of the NG. Adjusting this gap with respect to the diameter of the coax would correspond to adjusting the NG width to period ratio, which can be effectively used to alter the spectral bandwidth, angular bandwidth, and field enhancement. The strength of the field enhancement in all such structures can then be directly measured via two photon photo luminescence (TPPL) microscopy. A coaxial structure would have an additional benefit of

[1] Metal's quantum efficiency is assumed to be $\approx 0.01\%$.

polarization indifference: any polarization, including circular polarized light, would couple to the resonant SPP mode. Furthermore, the 3D structure can be used to launch the photoemitted electrons from a grid, thereby reducing their stochastic interaction to produce a cold electron beam.

In summary, introducing plasmonic nano-cavities in the shape of rectangular grooves arranged in a subwavelength grating on the metal surface leads to a dramatic increase of the multi-photon photoemission. Due to the light trapping effect of the NGs, complete reflectivity extinction is achievable on the metal surface. Furthermore, such a grating has the benefit of flexibility in terms of the resonance position, angular bandwidth, and spectral bandwidth, with a possibility of fine tuning the resonance position in a reversible manner. The experimental results demonstrate the seven orders of magnitude enhancement in the photoemission from the metal photocathode in the regime that allows optimum performance of the FEL.

MIM Dispersion: MatLab Code

A

```
% This function will attempt to find the solutions to the dispersion
% relation for a Metal-Insulator-Metal Plasmon Waveguide
%
% If
%   a) ei is the epsilon of the insulator,
%   b) em is the epsilon of the metal,
%   c) w is the width of the groove,
%   d) k = k' + ik" is the plasmon wave vector,
%   e) k0 = 2*pi/lambda0, where l0 is the incident light wavelength,
% then the equations read:
%
% i) ei*sqrt(k^2 - em*k0^2)/(em*sqrt(k^2 - ei*k0^2)) = ...
%       tanh((w/-2) * sqrt(k^2 - ei*k0^2))
% ii) ei*sqrt(k^2 - em*k0^2)/(em*sqrt(k^2 - ei*k0^2)) = ...
%       coth((w/-2) * sqrt(k^2 - ei*k0^2))
%
% in looking for the solution, the code will sweep the decay length, which
% corresponds to L = 1/(2*k"); and the plasmon wavelength, given by
% k = 2*pi/lambda.
%
% Assume that the plasmon wavelength is smaller then the incident
% light (it, certainly, cannot be larger). Then sweep lambda from 0 to
% lambda0. On the other hand, the decay length can be quite long, but,
% shorter then the surface plasmon polariton case. So, set the limits
% at 1nm to 10*lambda.

function [E0 kpl Li] = dispersion(mat, wRange, eRange)
    ePlot = figure; kPlot = figure; dPlot = figure;

    colors = ['b' 'k' 'g' 'r'];

    for i = 1:length(wRange)
        [E, kPl_r, kPl_i, decay] = calcDispersion(mat, wRange(i), eRange);

        figure(ePlot); plot(kPl_r, E, sprintf('-%so', colors(i))); hold all;
        figure(kPlot); plot(kPl_r, kPl_i, sprintf('-%so', colors(i))); hold all;
        figure(dPlot); plot(E, decay, sprintf('-%so', colors(i))); hold all;

        % save the output parameters:
        E0 = E; kpl = kPl_r; Li = decay;
    end
end


function [E, kPl_r, kPl_i, decay] = calcDispersion(mat, w, eRange)
    nm = 1e-9; hc = 1240;
```

```
% Construct the plasmon wave vector:
kPl = @(lambda, L)(complex(2*pi/(lambda*nm), 1/(2*L*nm)));

% Consider the tanh solution, let's define the left-hand-side:
lhs = @(lambda, L, props)( ...
    props.ei*sqrt(kPl(lambda, L)^2 - props.em*props.k0^2) / ...
    (props.em*sqrt(kPl(lambda, L)^2 - props.ei*props.k0^2)) );
% Now, define the right-hand-side:
rhs = @(lambda, L, props)( ...
    tanh((props.w/-2) * sqrt(kPl(lambda, L)^2 - props.ei*props.k0^2)) );
% Calculate the difference:
calcDiff = @(lambda, L, props)(abs(lhs(lambda, L, props) - rhs(lambda, L, props)));
% Calculate the energy of the photon
phE = @(lambda)(hc/lambda);

% read in the optical constants data for the metal:
fid = fopen(sprintf('/Users/apolyakov/LBL/Archive/Theory/Materials/nk.dir/%s.nk', mat), 'r');
props.mats = procData(textscan(fid, '%f %f %f', 'CommentStyle', ';'), eRange, phE);
fclose(fid);

% setup the properties structure:
props.w = w*nm; props.N = 50;
for i = 1:length(props.mats{1})
    % setup the properties of MIM structure
    props.disp.Eph(i) = hc/(0.1 * props.mats{1}(i));
    props.disp.lambda0(i) = props.mats{1}(i)/10;
    props.k0 = kPl(props.disp.lambda0(i), Inf);
    props.ei = 1;
    props.em = complex(props.mats{2}(i)^2 - props.mats{3}(i)^2, ...
        2*props.mats{2}(i)*props.mats{3}(i));

    % search for minimum:
    props.diffMin = 1e10;
    for lambda = 1 : props.disp.lambda0(i)/props.N : ...
            props.disp.lambda0(i)
        for L = 1:10*lambda/props.N:10*lambda
            % calculate the RHS - LHS:
            diff = calcDiff(lambda, L, props);
            if props.diffMin > diff
                props.diffMin = diff;
                props.lambdaMin = lambda;
                props.LMin = L;
            end
        end
    end

    % now optimize:
    min = fminsearch(@(x) calcDiff(x(1), x(2), props), [props.lambdaMin, props.LMin]);

    % now record the minimum for this energy:
    props.disp.decay(i) = min(2);
    props.disp.kPl(i) = 0.2*pi/min(1);
    props.disp.kPl_imag(i) = 0.1/(2*min(2));

    % for a selected photon energy display the fit results
    %fprintf(1, 'E = %.2f: lambda = %.0f nm; L = %.0f nm\n', ...
    %        phE(props.mats{1}(i)/10), min(1), min(2));

    % display the progress:
    fprintf(1, '%.1f%% done...\n', 100*i/length(props.mats{1}));
end

E = props.disp.Eph; decay = props.disp.decay;
```

```
    kPl_r = props.disp.kPl; kPl_i = props.disp.kPl_imag;
end

function mats = procData(data, eRange, phE)
    % init the data counter:
    j = 0;

    for i = 1:length(data{1})
        % Limit the energy region of interest:
        if phE(data{1}(i)/10) > eRange(2) || phE(data{1}(i)/10) < eRange(1)
            continue;
        else
            j = j + 1;
        end

        % copy the data
        mats{1}(j) = data{1}(i);
        mats{2}(j) = data{2}(i);
        mats{3}(j) = data{3}(i);
    end
end
```

Introduction to FDTD: Lumerical

B

The FDTD simulation used for this project was developed using the *Lumerical* software package[79]. The description of how to use this software is presented here in a tutorial-like fashion using the example of NGs on the gold surface. The NG grating dimensions are: 200 nm period, 15 nm width, and 30 nm depth.

B.1 Simulation Setup

Begin by opening a new simulation $File \rightarrow New \rightarrow 2D\ Simulation \rightarrow From\ default\ project$, this will open a layout as shown in figure B.1. The simulation consists of (1) drawing the structure, (2) setting up the simulation region, (3) inserting the light source, (4) placing monitors to record the data, and (5) analyzing the simulation results. The default dimensions are microns, which is what will be used in specifying dimensions later on. However, it is convenient to set the drawing grid to nanometers for ease of alignment. To do this, open the drawing grid setup window: $View \rightarrow Drawing\ grid \rightarrow Edit\ drawing\ grid$ and set a1 and a1 lattice vector to 0.001 microns (1 nm).

B.2 Sample Geometry and Simulation Setup

Substrate

From the action panel, choose $Structures \rightarrow Rectangle$. Notice the new element has appeared in the objects panel, right-click the substrate object and select *Edit object* to open the object properties panel. In the name field enter "substrate"; in the *Geometry* tab set $x = 0$, $x\ span = 1$, $y\ min = -1$, and $y\ max = 0$. This will set the dimensions of the substrate large enough to ensure that the simulation region does not end up on the substrate boundary.

Next, in the *Material* tab, set *material* to $Au\ (Gold) - Johnson\ and\ Christy$ and hit OK. Open $Simulation \rightarrow Material\ explorer$ and set the *simulation bandwidth* to the desired wavelength range. Then, hit *Fit and Plot* to see how the FDTD will approximate the optical constants at each wavelength in comparison to the empirical data. Adjust the

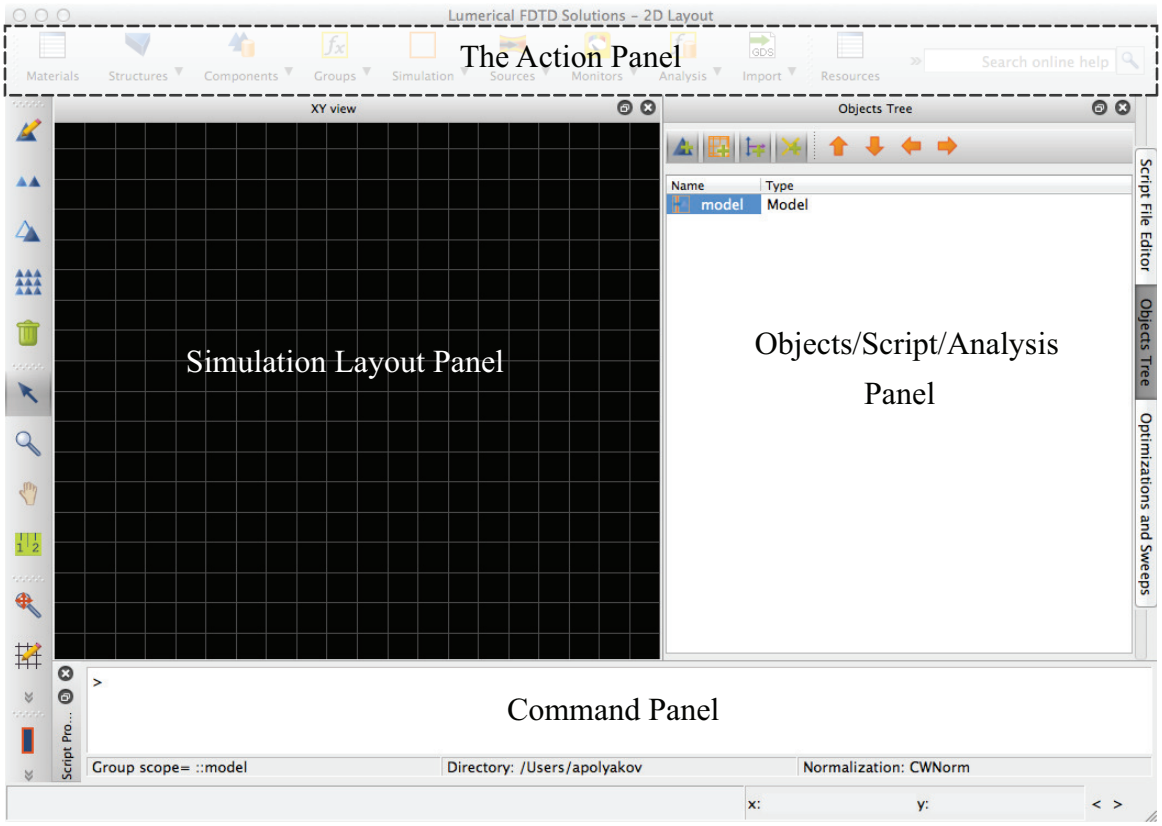

Figure B.1: *The new project window. Note the location of the action panel, the objects/script panel, and the command panel referred to in the text.*

max coefficients parameter until the fit is sufficiently close (click "Save" then "Yes" if prompted to save the new optical constants fit).

Nano-Groove

In a similar fashion, place another rectangle and open its properties panel. Set the parameters as follows: set *name* to "NG"; under *Geometry* tab: $x = 0$, $x\ span = 0.015$, $y\ min = -0.03$, $y\ max = 0$ (all dimensions are in microns); under the *Material* tab set *material* to *etch*, this will ensure–in case objects happen to overlap, as they do in this case–that the region where the etch object is located will have the index of refraction of vacuum.

To save computer resources, it is advantageous to define a fine-mesh region around the NG and use a coarse mesh elsewhere. To do this, select *Simulation* $\rightarrow$ *Mesh* from the action panel (see figure B.1). A new object will appear similar to the rectangles before in the objects panel, open the object parameters panel for the mesh. Set *name* to "NG Mesh" and both *override x mesh* and *override y mesh* to 0.0005 to force the mesh to be 0.5 nm

around the NG. Under the *Geometry* tab, set $x = 0$, $x\ span = 0.019$, $y\ min = -0.032$, $y\ max = 0.002$. This will set a fine mesh around the NG with 2 nm clearance.

Simulation Region

Select *Simulation* $\rightarrow$ *Region* to open an FDTD simulation region: a corresponding object will appear in the objects panel named "FDTD", open the parameters panel for it. Under the *Geometry* tab set $x = 0$, $x\ span = 0.2$, $y\ min = -0.3$, and $y\ max = 1$. Under the *Mesh settings* tab set[1] *mesh type* to *auto non − uniform*, *mesh accuracy* to 2, and *mesh refinement* to *conformal variant* 2. In general, to improve simulation accuracy the *mesh accuracy* value can be increased, this, however, has little effect for the geometry considered here. In the *Boundary conditions* tab set *allow symmetry on all boundaries* to on, $x\ min\ bc$ to *periodic*, which will also set the $x\ max\ bc$ to the same value.

B.3 Light Source and Monitors

Plane Wave Light Source

Select *Sources* $\rightarrow$ *Plane wave* to create a plane wave light source object in the object panel. Open the parameters window for it. Set *name* to "linSource"; under the *General* tab set *direction* to *Backward* and *polarization* to TE. Under the *Geometry* tab set $x = 0$, $x\ span = 1$, and $y = 0.8$. In the $Frequency/Wavelength$ tab set $set\ frequency/wavelength$ to on and then *wavelength start* to 0.4 and *wavelength stop* to 1.0. This will set the light source to be polarized in the direction perpendicular to the NGs with a bandwidth of 600 nm from 400 nm to 1 micron injected in the proper orientation.

Cavity Q

It is always useful to place a time monitor near the structure to check that the fields decay to 0 before the simulations stops. Use $Monitors \rightarrow Field\ time$ to insert a monitor. Then select it from the object panel and open the preferences window. Set *name* to "q"; under the *Geometry* tab set *monitortype* to *Point*, $x = 0$, and $y = -0.025$. This monitor will record all the data as listed under the *Data to record* tab, and can also be used to measure the Q of the cavity.

Reflectivity

To record the reflectivity, place a $Monitors \rightarrow Frequency - domain$ monitor. In the preferences, set *name* to "R", set *override global monitor setting* to on, and $frequency\ points$

[1]The exact settings in this tab are not essential, as the fine mesh defines the NG, which is the prime point of light-matter interaction.

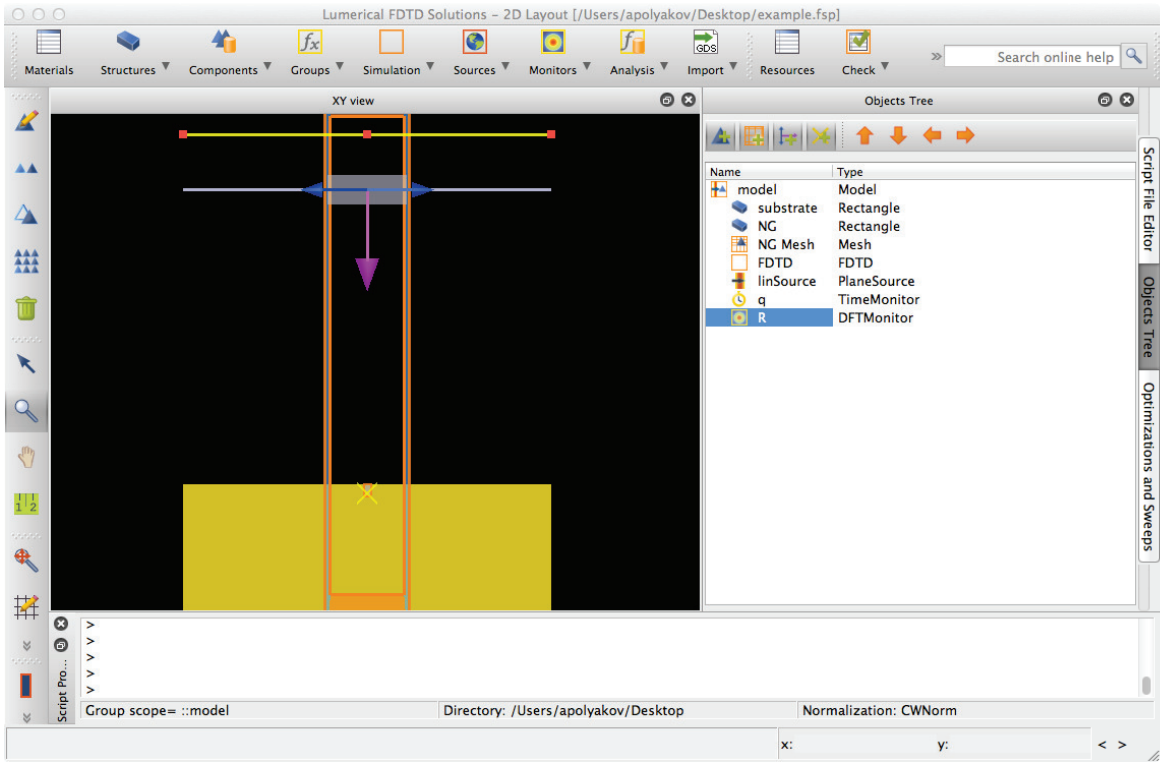

Figure B.2: *The properly setup simulation for the NG on the gold surface.*

to 600, which will produce wavelength resolution of 1 nm per point; under *Geometry* tab, set *monitor type* to *Linear X*, $x = 0$, *x span* $= 1$ (the part of the monitor outside the FDTD region is ignored automatically), and $y = 0.95$. This monitor is placed above the light source, so it only records the normalized power flow *reflected* off the sample. It is important to have a time field monitor somewhere in the FDTD region, to check that all the fields decay when the simulation stops, otherwise the fields recorded by the field profile monitor will be inaccurate. The *Lumerical* window should now look similar to figure B.2.

Structure Automation

When preparing an FDTD simulation it is worthwhile to invest time in setting up the structure to automatically adjust the codependent components. For example, the "NG Mesh" must follow the dimensions of the NG. This is done by inserting a short script into the *model* object in the object panel. Open the parameters window and select the *Setup* tab. In there, select the *Variables* sub-tab, where the parameters for the system can be inserted. Add parameters by hitting the *Add* button and then adjusting the values specifying each parameter. Add three parameters: width (*name*="width", *type* = *Length*, *value* = 0.015, *units* = *micron*), height (*name*="height", *type* = *Length*, *value* = 0.03,

units = *micron*), period (*name*="period", *type* = *Length*, *value* = 0.2, *units* = *micron*), polarization (*name*="polarization", *type* = *Number*, *value* = 0: 0 is p-polarization and 1 is s-polarization).

Switch to the *Script* sub-tab and insert the following text:

```
# set the NG dimensions
select("NG");
set("x span", width);
set("y min", -height);
set("y max", 0);

# configure the simulation region and the b.c.:
select("FDTD");
set("allow symmetry on all boundaries", 1);
set("x span", period);
# polarization: 0 is p-polarized, 1 is s-polarized
if (polarization == 0) {
    set("x min bc", 5);
    set("x max bc", 5);
} else {
    set("x min bc", 4);
    set("x max bc", 4);
}

# set the light source polarization:
select("linSource");
if (polarization == 0) {
    set("polarization", 2);
} else {
    set("polarization", 1);
}

# adjust the fine mesh to match the NG dimensions:
select("NG Mesh");
set("x span", width + 2*2e-9);
set("y min", -height-2e-9);
set("y max", 2e-9);
```

Now changing the parameters in the "model" object will automatically adjust the parameters of all the necessary objects in the model.

B.4 Data Analysis

The most important part of using the FDTD model is to find the optimum geometry to produce the resonance at the target wavelength. The following script can be effectively used to do that by repeatedly measuring the reflectivity for a different triplet of (w, h, p) until the desired reflectivity minimum is found.

```
##### Choose NG Dimensions (in nm): #####
w = 15; h = 30; p = 200;
########################################

switchtolayout; nm = 1e-9;

# set the dimension of the NG:
```

```
select("::model");
set("width", w*nm);
set("height", h*nm);
set("period", p*nm);

# run the s-polarization:
set("polarization", 1);
run(3);

# record the reflectivity:
refl = transmission("R");

# run the same model in p-polarization:
switchtolayout;
select("::model");
set("polarization", 0);
run(3);

# get the frequency from the time monitor:
x = c/getdata("R", "f");

# get the fields from the time monitor:
refl = transmission("R") / refl;

# convert reflectivity to transmission:
nm=1e-9; y = 1-refl;

## find the peaks:
idx = findpeaks(y,2);
?"peaks: " + num2str(x(idx(1))/nm) + " " + num2str(x(idx(2))/nm);
refl_l = x(idx(1))/nm;
refl_min = 1-y(idx(1));

# output the minimum found
?"(p,w,h)=("+num2str(w)+","+num2str(h)+","+num2str(p)+ ")=>R(@"+num2str(refl_l)+")="+num2str(refl_min);

# also plot the result
plot(wlen, refl, "Wavelength (nm)", "Relative Reflectivity");
```

Once the dimensions are optimized to produce a sufficient reflectivity minimum, the code shown below can be used to calculate the cavity Q, defined as

$$Q \equiv \frac{Resonance\ Frequency}{Full\ Width\ Half\ Maximum}. \tag{B.1}$$

```
# get the frequency from the time monitor:
x = fftw(getdata("q","t"),2)/(2*pi);

# get the fields from the time monitor:
y = abs(fft(getdata("q","Ex"),2))^2 +
    abs(fft(getdata("q","Ey"),2))^2;# +
#    abs(fft(getdata("q","Ez"),2))^2;

# Fourier transform
k=fftk(x,3);
y_k=czt(y,x,k);

# re-sample data at 10x
x_hi=linspace(min(x),max(x),20000);
y_hi=abs(1/length(k)*czt(y_k,-k,x_hi));
#plot(x_hi,abs(y_hi),"f","intensity","p="+num2str(p/nm));
```

```
## find the peaks:
idx = findpeaks(y_hi);

# setup the indices
lidx = 1; ridx = 1; found = 0;

# first, find the left index of half-max:
for(i=idx; i>0; i=i-1) {
    if (found) {1;}
    else {
        if (2*y_hi(i) < y_hi(idx)) {
            found = 1;
            lidx = i;
        }
    }
}

# second, find the right index of half-max:
found = 0;
for(i=idx; i<size(y_hi); i=i+1) {
    if (found) {1;}
    else {
        if (2*y_hi(i) < y_hi(idx)) {
            found = 1;
            ridx = i;
        }
    }
}

res_FWHM = x_hi(ridx) - x_hi(lidx);
res_Q = x_hi(idx)/res_FWHM;
res_l = 1e9*c/x_hi(idx);

?"The cavity Q is " + num2str(res_Q) + " at " + num2str(res_l) + " nm";
```

B.5 On-Resonance Behavior

Time Resolved Fields

A script from the previous section displays the resonance wavelength for a given dimensions of the NG. To investigate the time-dependent fields in the structure, first, open the parameters window for the "linSource" object from the objects panel. From the *Frequency/Wavelength* tab set *wavelength start* and *wavelength stop* to the same value of 0.6, corresponding to the resonance wavelength of 598 nm from the example above.

Next, insert another field time monitor by going to *Monitors* $\rightarrow$ *Field time*. Open it's parameter window and set the following: *name*="tFields", *monitor type* = *Linear Y*, $y\ min = -0.3$, $y\ max = 0.7$.

The incoming light is polarized perpendicular to the NGs, that is in the x-direction, therefore, only the x component of the E-field will be varying by much. Entering the following commands in the command panel will produce a *MatLab* compatible file with all the relevant field data saved in a matrix form:

```
> y = getdata("tFields", "y");
> t = getdata("tFields", "t");
```

```
> E = getdata("tFields", "Ex");
> matlabsave("fields.mat", y, t, E);
```

The following *MatLab* script will produce and AVI movie file from the saved fields data:

```
function tResFields(path,  yLim)
    load(path);

    % prepare the movie file:
    aviobj = avifile('fields.avi', 'compression', 'None');

    % prepare the graphing window to show progress:
    pWin = figure; set(gca, 'nextplot', 'replacechildren');

    % get the dimensions of the E-field array:
    [~, yLen, tLen] = size(E);

    for i = 1:tLen
        plot(y' * 1e9, E(:,:,i), '-b', 'LineWidth', 2);
        xlabel('Distance (nm)'); ylabel('E-field (normalized to unity)');
        title(sprintf('Time: %.4f fs', t(i)*1e15));
        axis([y(1)*1e9 y(yLen)*1e9 yLim(1) yLim(2)]);
        aviobj = addframe(aviobj, getframe(pWin));
    end

    close(pWin); aviobj = close(aviobj);
end
```

For the example presented here, invoke it with the following parameters:

```
>> tResFields('fields.mat', [-12 12])
```

Changing the position of the field time monitor "tFields" will record the time evolution of the fields at a different position. In the same way, the geometry can be changed to a point to observe the time variation of the fileds at a single point, as was done above for the cavity Q calculations.

Field Enhancement

The fields recorded by the $Frequency-domain\ field\ profile$ represent the amplitude of the steady-state solution for the field oscillating with the corresponding time dependance: $i\omega t$ at a given angular frequency ω. Internally, the fields are normalized to unity, which makes it easy to read off the field enhancement at the NGs from the field profile monitor.

Two more object need to be added to the simulation to obtain smooth fields. First, go to $Simulation \rightarrow Mesh$ to add another mesh. Change its parameters as follows: *name*="fieldMesh"; under *General* tab, *override x mesh* and *override y mesh* both set to 0.0005; under *Geometry* tab, $x = 0$, $x\ span = 0.2$, $y\ min = -0.05$, and $y\ max = 0.15$. Second, add $Monitors \rightarrow Frequency-domain\ field\ profile$ monitor and set its *name* to "FE"; then set the geometry settings for it to be identical to the "fieldMesh".

After running the simulation issue the following commands in the command panel:

```
> x = getdata("FE", "x");
> y = getdata("FE", "y");
> Ex = getdata("FE", "Ex");
```

```
> Ey = getdata("FE", "Ey");
> Hz = getdata("FE", "Hz");
> matlabsave("fields.mat", x, y, Ex, Ey, Hz);
```

The commands above will save the fields data into a *MatLab* ready matrix. The resulting fields are readily plotted on the logscale by issuing the following commands in *MatLab*:

```
>> load('fields.mat');
>> z0 = 376.73;
>> EFI = Ex.*conj(Ex) + Ey.*conj(Ey);
>> MFI = z0^2 * (Hz.*conj(Hz));
>> contourf(x*1e9, y*1e9, log10(EFI'), 200);
>> contourf(x*1e9, y*1e9, log10(MFI'), 200);
```

The magnetic field needs to be multiplied by the free space impedance, because *Lumerical* normalizes to unity only the electric field. Figure 4.5 was produced in this way.

Calculating the p-polarization field enhancement factor inside the metal in reference to the s-polarization fields is a bit more complicated, because the intensity has to be integrated over the edge of the NG. To do this, first, rerun the simulation with light in s-polarization, then save the fields as follows (the x- and y-components of the E- and H-field are taken from above):

```
> Ez = getdata("FE", "Ez"); > Hx = getdata("FE", "Hx"); > Hy =
getdata("FE", "Hy"); > matlabsave("fields.mat", x, y, Ex, Ey, Ez,
Hx, Hy, Hz);
```

Then the average FE factor is calculated by the following *MatLab* script:

```
load('fields.mat');

% load the p-polarization field intensity
EFI = Ex.*conj(Ex) + Ey.*conj(Ey);

% load the s-polarization field intensity
EFIs = Ez.*conj(Ez);

% the starting x-position of the NG metal edge
xs = 198; % 4 points per nm

% integrate the field intensity over the first 2 nanometers
dx = 2*4;
for i = 1:length(y); yt(i) = 0; for j = xs:xs+dx; yt(i) = yt(i) + EFI(j,i); end; end
for i = 1:length(y); yts(i) = 0; for j = xs:xs+dx; yts(i) = yts(i) + EFIs(j,i); end; end

% take the ratio to get the field enhancement over the NG (including the gold bottom)
ymin = 38; ymin = 274; ymax = 290; yNG = y(ymin:ymax);
fe_ary = yt(ymin:ymax) ./ yts(ymin:ymax);

figure; plot(yNG, fe_ary);

% calculate the average FE:
fe = trapz(fe_ary) / length(fe_ary)
```

The scripts provided here are meant to illustrate how to reproduce the theoretical calculations used in this work, they are by no means exhaustive of the capacities of the *Lumerical* software.

Optical Reflectivity Setup | C

C.1 Imaging Mode

The optical setup for measuring the relative reflectivity of the NG pattern has been described in figure 6.1. A more detailed description of this setup is presented here. Figure C.1 shows the photograph of the actual setup together with the corresponding schematic. The underlying principle is to record an image of the pattern under narrow bandwidth normal-incidence illumination. To achieve this, a broadband light source is used in combination with a monochromator producing a 5 nm bandwidth light (this part of the setup is not shown in the photograph). An off-axis poraboloid (OAP) aluminum mirror is then used to collimate the beam before the polarizer. The polarizer is mounted on the rotation stage, which completes the light conditioning. After the polarizer the light has a defined polarization, wavelength, and divergence.

The normal-incidence image is then recorded via a beam splitter as shown in the schematic. An example of a set of recorded photographs is shown at the bottom of figure C.1. At every wavelength, two photographs are recorded: one in p-polarization (E-vector perpendicular to the grooves), and one in s-polarization (E-vector parallel to the grooves). Before the reflectivity scan is performed, two areas are identified on the image where the signal from the pattern and the signal from the flat surface will be collected from, as indicated on the photographs in figure C.1. From each area, and average signal is calculated and is later used to measure the relative reflectivity using equation 6.1.

C.2 Integrated Reflectivity

The main advantage of the imaging setup described above is the minimal numerical aperture (NA). This is essential for resolving sharp features in the spectral scan. However, the NG grating as described in this work has an angular bandwidth of over 15 degrees. Such spectral features can be readily resolved with a higher NA. Therefore, a setup was designed, where a microscope objective lens with NA of 0.14 was used to measure an integrated signal from the grating. The schematic and the photograph of the actual setup are shown in figure C.2. The idea is to illuminate the sample with broadband light, then use an iris to select a 100 micron spot and record the signal with a spectrometer in one shot. Such an approach is much faster (less than 1 minute) then an image scan (over three hours) and has a greater

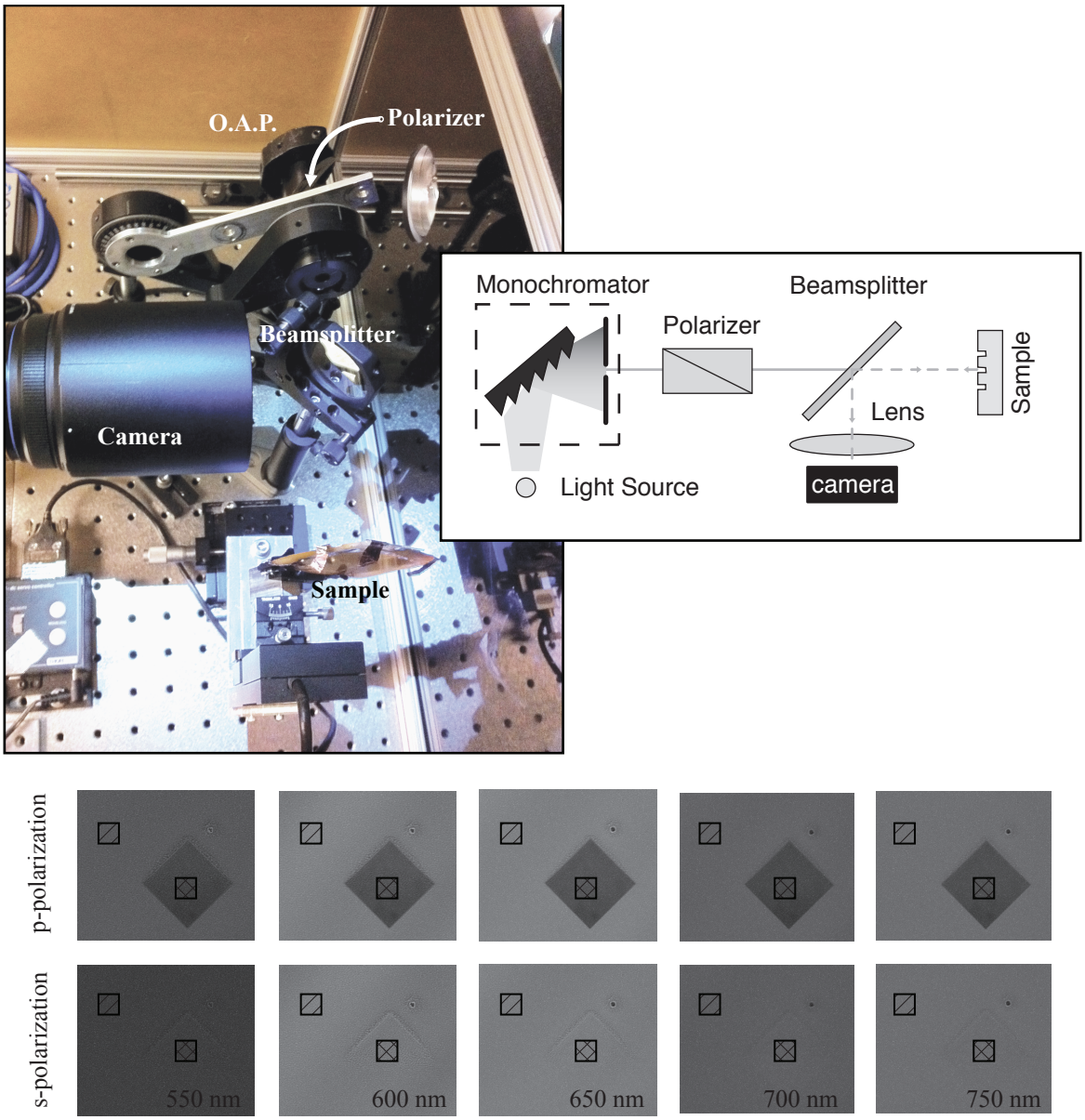

Figure C.1: *In the imaging mode, a photograph of the pattern is recorded in s- and p-polarization at normal incidence. Then, on each photograph, two areas are identified: one on the pattern, and one on the flat surface. The integrated intensity from these areas is then used to calculated the relative reflectivity.*

(a) Optical Reflectivity Setup

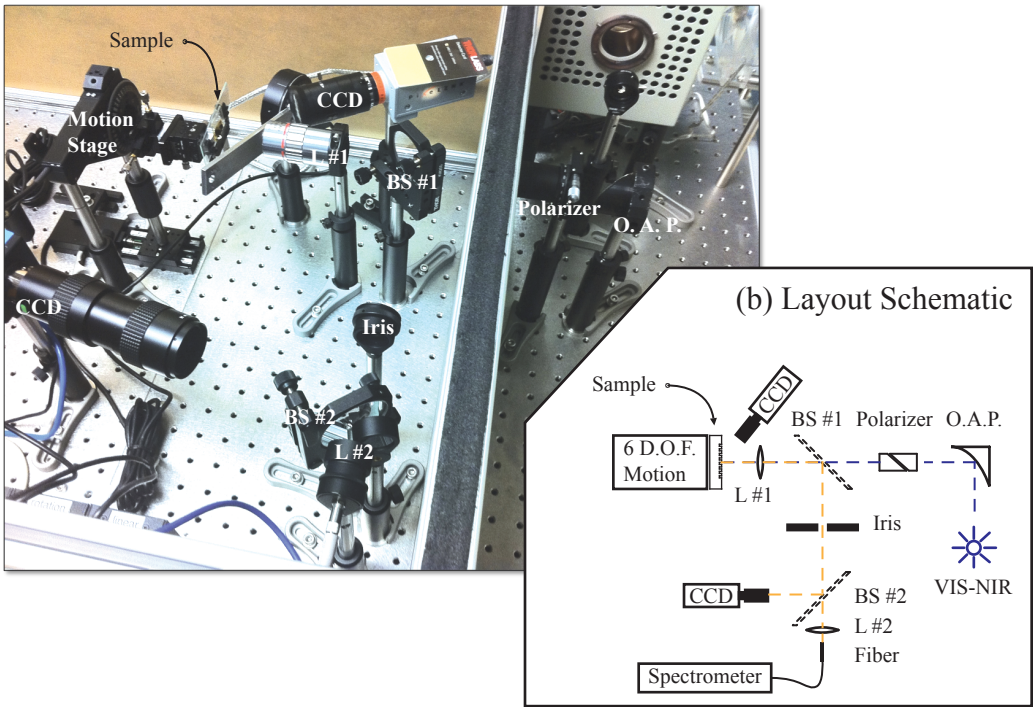

Figure C.2: *The optical setup combines the microscope objective for one-shot full-spectrum scan, as well as the imaging mode for precise alignment.*

resolution of 2 $\mathring{A}$ using only 17% of light intensity. The only disadvantage is higher NA. The setup consists of the following elements:

O.A.P. The off-axis paraboloid (O.A.P.) is used to parallelize the light from the UV-NIR light source (*Newport* part #50338AL).

Polarizer The polarizer selects out only one polarization; after the polarizer, the illumination is collimated with a defined polarization (Double Glan-Taylor Polarizer, *Thorlabs* part #DGL10).

BS #1 The first beam splitter (BS) is used to achieve a normal incidence illumination (*Thorlabs* part #BSW27).

L #1 The first lens (L) is the microscope objective lens with NA = 0.14; it is used for imaging the sample to the iris. The CCD next to this lens is used for a coarse alignment of the illumination spot to the sample pattern.

Motion The 6 degrees-of-freedom (D.O.F.) motion stage is used to achieve a complete control of the sample alignment and position relative to the incident illumination.

Iris The iris is placed in the image plane of the objective lens (L #1) to pick out a 100 μm diameter spot to record the reflectivity from. All the light that passes through this iris is assumed to come from the pattern and no further spatial resolution is required. This way, the minimum diameter of the iris determines the spatial resolution of this optical setup.

BS #2 The second beam splitter is used to focus the CCD camera onto the iris for maximum resolution alignment.

L #2 The second lens is an aspherical broadband lens of 4 mm focal length, it is used to couple the reflected beam into the fiber.

Fiber The extended throughput, 600 μm diameter fiber is used to pass as much of the reflected light signal to the spectrometer as possible.

Spectrometer This is the device that records the relative intensity per wavelength. The exposure used was around 4-6 seconds, however, the sample has to be rotated in between the readings, brining the total measurement time to about 1 minute (*Ocean Optics* part #USB4000).

The broadband light source is collimated via the OAP mirror prior for producing a collimated beam to pass through a statically mounted polarizer. Instead of rotating the polarizer, in this setup, the sample is mounted on a stage the allows sample rotation, tilt, and X-Y-Z movement. For normal-incidence measurement, a beam splitter is used as shown in the schematic (see figure C.2). To collect the reflected light from the sample, a microscope objective is used, which defines the NA of the system. To achieve a magnification of 10X, a working distance below 35 mm has to be used, which results in the necessity to pass the light through an objective lens to illuminate the sample. A longer working distance would allow placing a beamsplitter between the lens and the sample reducing the system NA to the same small values as in the imaging setup described in the previous section.

The reflected signal is collected by the objective lens, which produces an image at the iris (see the schematic in figure C.2). Closing the iris allows picking out the reflected signal from a defined region on the sample. The minimal iris diameter is 700 microns corresponding to a 70 micron resolution with a 10X optical magnification. Past the iris, only the signal from the selected region of interest (ROI) is allowed to propagate. This signal is split into two by the second beam splitter to allow observing the iris with a camera for alignment and to later record the signal by the spectrometer. The signal is coupled by a short focal distance lens into a fiber, which is attached to a spectrometer producing the full reflectivity spectrum in one shot. The reflection is recorded as an integrated signal from the ROI, which is then normalized to the reflectivity of the sample by repeating the measurement on the flat surface.

Bibliography

1. Millikan, R. A. A Direct Determination of "h." *Phys. Rev.* **4,** 73–75 (1914).
2. Millikan, R. A. A direct photoelectric determination of Planck's h. *Phys. Rev.* **7,** 355–388 (1916).
3. Emma, P., Akre, R., Arthur, J., Bionta, R. & Bostedt, C. First lasing and operation of an ångstrom-wavelength free-electron laser. *Nature Photonics* (2010).
4. Pelligrini, C. & Reiche, S. *Encyclopedia of Applied High Energy and Particle Physics* (2009).
5. Huang, Z. Review of x-ray free-electron laser theory. *Physical Review Special Topics-Accelerators and Beams* (2007).
6. Attwood, D. *Soft X-Rays and Extreme Ultraviolet Radiation* (Cambridge Univ Pr, 2007).
7. Roberson, C. Free-electron laser beam quality. *Quantum Electronics, IEEE Journal of* **21,** 860–866 (1985).
8. Dowell, D. H. & Schmerge, J. F. Quantum efficiency and thermal emittance of metal photocathodes. *Phys. Rev. Special Topics - Accelerators and Beams* **12** (2009).
9. Dowell, D. H. *et al.* Cathode R&D for future light sources. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **622,** 685–697 (2010).
10. Fowler, R. The Analysis of Photoelectric Sensitivity Curves for Clean Metals at Various Temperatures. *Phys. Rev.* **38,** 45–56 (1931).
11. Valfells, Á., Feldman, D. W., Virgo, M., O'Shea, P. G. & Lau, Y. Y. Effects of pulse-length and emitter area on virtual cathode formation in electron guns. *Physics of Plasmas* **9,** 2377 (2002).
12. Furman, M. A. & Chou, W. Beam Dynamics Newsletter. *International Committee for Future Accelerators* 1–209 (2008).
13. Ferrini, G., Michelato, P. & Parmigiani, F. A Monte Carlo simulation of low energy photoelectron scattering in Cs2Te. *Solid State Communications* **106,** 21–26 (1998).

14. Musumeci, P., Moody, J., England, R., Rosenzweig, J. & Tran, T. Experimental Generation and Characterization of Uniformly Filled Ellipsoidal Electron-Beam Distributions. *Phys. Rev. Lett.* **100** (2008).

15. Luiten, O., van der Geer, S., de Loos, M., Kiewiet, F. & van der Wiel, M. How to Realize Uniform Three-Dimensional Ellipsoidal Electron Bunches. *Phys. Rev. Lett.* **93** (2004).

16. Rentenberger, S., Vollmer, A., Zojer, E., Schennach, R. & Koch, N. UVozone treated Au for air-stable, low hole injection barrier electrodes in organic electronics. *Journal of Applied Physics* **100,** 053701 (2006).

17. Muggli, P. & Brogle, R. Two-color photoemission produced by femtosecond laser pulses on copper. *Journal of Optical Society of America B* (1995).

18. Raether, H. *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer-Verlag, 1988).

19. Callcott, T. A. Volume and surface photoemission processes from plasmon resonance fields. *Phys. Rev. B* (1975).

20. Musumeci, P. *et al.* Multiphoton Photoemission from a Copper Cathode Illuminated by Ultrashort Laser Pulses in an rf Photoinjector. *Phys. Rev. Lett.* **104** (2010).

21. Economou, E. N. Surface Plasmons in Thin Films. *Phys. Rev.* **182,** 539–554 (1969).

22. Wood, R. W. On a remarkable case of uneven distribution of light in a diffraction grating spectrum. (1902).

23. Wood, R. W. Optical properties of homogeneous and granular films of sodium and potassium. *Phil. Mag.* **38,** 98–112 (1919).

24. Mills, D. L. Enhanced electric fields near gratings: comments on enhanced Raman scattering from surfaces. *Phys. Rev. B* (1982).

25. López-Ríos, T., Mendoza, D., García-Vidal, F., Sánchez-Dehesa, J. & Pannetier, B. Surface Shape Resonances in Lamellar Metallic Gratings. *Phys. Rev. Lett.* **81,** 665–668 (1998).

26. Markel, V. A. *et al.* Near-field optical spectroscopy of individual surface-plasmon modes in colloid clusters. *Phys. Rev. B* **59,** 10903–10909 (1999).

27. Le Perchec, J., Quémerais, P., Barbara, A. & López-Ríos, T. Why Metallic Surfaces with Grooves a Few Nanometers Deep and Wide May Strongly Absorb Visible Light. *Phys. Rev. Lett.* **100** (2008).

28. Barnes, W. L., Dereux, A. & Ebbesen, T. W. Surface plasmon subwavelength optics. *Nature* **424,** 824–830 (2003).

29. García de Abajo, F. J. Colloquium: Light scattering by particle and hole arrays. *Reviews of Modern Physics* **79,** 1267–1290 (2007).

30. Beversluis, M. R., Bouhelier, A. & Novotny, L. Continuum generation from single gold nanostructures through near-field mediated intraband transitions. *Phys. Rev. B* **68,** 115433 (2003).

31. Castro-Lopez, M., Brinks, D., Sapienza, R. & van Hulst, N. F. Aluminum for Nonlinear Plasmonics: Resonance-Driven Polarized Luminescence of Al, Ag, and Au Nanoantennas. *Nano Lett.* **11,** 4674–4678 (2011).

32. Stockman, M., Bergman, D. J., Anceau, C., Brasselet, S. & Zyss, J. Enhanced second-harmonic generation by metal surfaces with nanoscale roughness: Nanoscale dephasing, depolarization, and correlations. *Phys. Rev. Lett.* **92,** 057402 (2004).

33. Fort, E. & Gresillon, S. Surface enhanced fluorescence. *Journal of Physics D: Applied Physics* **41,** 013001 (2008).

34. Popov, E., Neviere, M., Enoch, S. & Reinisch, R. Theory of light transmission through subwavelength periodic hole arrays. *Phys. Rev. B* **62,** 16100–16108 (2000).

35. Pendry, J. B., Holden, A. J., Robbins, D. J. & Stewart, W. J. Magnetism from Conductors and Enhanced Nonlinear Phenomena. *IEEE Trans. Microw. Theory Tech.* **47,** 2075–2084 (1999).

36. Moreau, A., Granet, G., Baida, F. I. & Van Labeke, D. Light transmission by subwavelength square coaxial aperture arrays in metallic films. *Opt Express* **11,** 1131 (2003).

37. Grober, R. D., Schoelkopf, R. J. & Prober, D. E. Optical antenna: Towards a unity efficiency near-field optical probe. *App. Phys. Lett.* **70,** 1354 (1997).

38. Bozhevolnyi, S. I., Volkov, V. S., Devaux, E. & Ebbesen, T. W. Channel plasmon-polariton guiding by subwavelength metal grooves. *Phys. Rev. Lett.* **95,** 046802 (2005).

39. Dennis, M. R., Zheludev, N. I. & García de Abajo, F. J. The plasmon Talbot effect. *Opt. Express* **15,** 9692–9700 (2007).

40. Polyakov, A. *et al.* Plasmonic light trapping in nanostructured metal surfaces. *App. Phys. Lett.* **98,** 203104 (2011).

41. Smith, D. R., Padilla, W. J., Vier, D. C., Nemat-Nasser, S. C. & Schultz, S. Composite Medium with Simultaneously Negative Permeability and Permittivity. *Phys. Rev. Lett.* **84,** 4184–4187 (2000).

42. Fan, W., Zhang, S., Minhas, B., Malloy, K. J. & Brueck, S. R. J. Enhanced Infrared Transmission through Subwavelength Coaxial Metallic Arrays. *Phys. Rev. Lett.* **94,** 033902 (2005).

43. Fromm, D. P, Sundaramurthy, A., Schuck, P. J., Kino, G. & Moerner, W. E. Gap-Dependent Optical Coupling of Single "Bowtie" Nanoantennas Resonant in the Visible. *Nano Letters* **4,** 957–961 (2004).

44. Homola, J., Yee, S. S. & Gauglitz, G. Surface plasmon resonance sensors: review. *Sens. Actuator B-Chem.* **54,** 3–15 (1999).

45. Palik, E. D. *Handbook of the Optical Constants of Solids* (Academic Press, New York, 1985).

46. Ohta, K. Matrix formalism for calculation of electric field intensity of light in stratified multilayered films. *Applied Optics* (1990).

47. Hüfner, S. *Photoelectron spectroscopy* (Springer Verlag, 2003).

48. Pepper, S. V. Optical analysis of photoemission. *JOSA* (1970).

49. Gordon, R. Light in a subwavelength slit in a metal: Propagation and reflection. *Phys. Rev. B* **73** (2006).

50. Griffiths, D. J. *Introduction to electrodynamics* (Benjamin-Cummings Pub Co, 1999).

51. Zhang, S., Liu, H. & Mu, G. Electromagnetic enhancement by a single nano-groove in metallic substrate. *J. Opt. Soc. Am. A* **27,** 1555–1560 (2010).

52. Lalanne, P., Hugonin, J. P. & Rodier, J Theory of Surface Plasmon Generation at Nanoslit Apertures. *Phys. Rev. Lett.* **95** (2005).

53. Temnov, V. V., Woggon, U., Dintinger, J., Devaux, E. & Ebbesen, T. W. Surface plasmon interferometry: measuring group velocity of surface plasmons. *Optics Letters* **32,** 1235–1237 (2007).

54. Lalanne, P. & Hugonin, J. P. Interaction between optical nano-objects at metallo-dielectric interfaces. *Nature Physics* **2,** 551–556 (2006).

55. Schouten, H. *et al.* Plasmon-Assisted Two-Slit Transmission: Young's Experiment Revisited. *Phys. Rev. Lett.* **94** (2005).

56. Teperik, T. V. *et al.* Omnidirectional absorption in nanostructured metal surfaces. *Nature Photonics* **2,** 299–301 (2008).

57. Pendry, J. B. Mimicking Surface Plasmons with Structured Surfaces. *Science* **305,** 847–848 (2004).

58. Kubo, A. *et al.* Femtosecond Imaging of Surface Plasmon Dynamics in a Nanostructured Silver Film. *Nano Letters* **5,** 1123–1127 (2005).

59. Ebbesen, T. W., Lezec, H. J., Ghaemi, H. F., Thio, T. & Wolff, P. A. Extraordinary optical transmission through sub-wavelength hole arrays. *Nature* **391,** 667–669 (1998).

60. Polyakov, A. *et al.* Light trapping in plasmonic nanocavities on metal surfaces. *J. Vac. Sci. Technol. B* **29,** 06FF01 (2011).

61. Peroz, C. *et al. Nanotechnology* **21,** 445301 (2010).

62. Waggener, H. A., Krageness, R. C. & Tyler, A. L. *Two-way etch* (Electronics, 1967).

63. Josse, M. & Kendall, D. L. Rectangular-Profile Diffraction Grating From Single-Crystal Silicon. *Applied Optics* **19,** 72–76 (1980).

64. Kendall, D. L. Vertical Etching of Silicon at Very High Aspect Ratios. *Annual Review of Materials Science* **9,** 373–403 (1979).

65. Nagel, D. J., Bean, K. E. & Watts, R. K. Spectroscopic Applications of Structures Produced by Orientation-Dependent Etching. *Nuclear Instruments & Methods* **172,** 321–326 (1980).

66. Voronov, D. L. *et al.* A 10,000 groove/mm multilayer coated grating for EUV spectroscopy. *Optics Express* **19,** 6328–6333 (2011).

67. Liang, X. & Chou, S. Y. Nanogap Detector Inside Nanofluidic Channel for Fast Real-Time Label-Free DNA Analysis. *Nano Letters* **8,** 1472–1476 (2008).

68. Chou, S. Y. & Smith, H. I. Xray lithography for sub100nmchannellength transistors using masks fabricated with conventional photolithography, anisotropic etching, and oblique shadowing. *J. Vac. Sci. Technol. B* (1985).

69. Polyakov, A. *et al.* Photocathode performance improvement by plasmonic light trapping in nanostructured metal surfaces. *Proc. SPIE - Int. Soc. Opt. Eng. (USA)* **8094,** 809407 (2011).

70. Atwater, H. A. & Polman, A. Plasmonics for improved photovoltaic devices. *Nat. Mater.* **9,** 205–213 (2010).

71. Dionne, J. A., Sweatlock, L. A., Atwater, H. A. & Polman, A. Planar metal plasmon waveguides: frequency-dependent dispersion, propagation, localization, and loss beyond the free electron model. *Phys. Rev. B* **72,** 075405 (2005).

72. Leskela, M. & Ritala, M. Atomic layer deposition (ALD): from precursors to thin film structures. *Thin Solid Films* **409,** 138–146 (2002).

73. Groner, M. D., Fabreguette, F. H., Elam, J. W. & George, S. M. Low-Temperature Al2O3 Atomic Layer Deposition. *Chem. Mater.* **16,** 639–645 (2004).

74. Profijt, H. B., Potts, S. E., van de Sanden, M. C. M. & Kessels, W. M. M. Plasma-Assisted Atomic Layer Deposition: Basics, Opportunities, and Challenges. *J. Vac. Sci. Technol. A* **29,** 050801 (2011).

75. Ott, A. W., Klaus, J. W., Johnson, J. M. & George, S. M. Al2O3 thin film growth on Si (100) using binary reaction sequencechemistry. *Thin Solid Films* **292,** 135–144 (1997).

76. Kim, S. K., Lee, S. W., Hwang, C. S. & Min, Y. S. Low Temperature (¡100 C) Deposition of Aluminum Oxide Thin Films by ALD with O3 as Oxidant. *Journal of The Electrochemical Society* (2006).

77. Corlett, J. N. A next generation light source facility at lbnl. *Particle Accelrator Conference Pac'11, New York, NY, March 28-April 1* (2011).

78. Ferrini, G., Banfi, F., Giannetti, C. & Parmigiani, F. Non-linear electron photoemission from metals with ultrashort pulses. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **601,** 123–131 (2009).

79. *Lumerical Solutions, Inc.* `http://www.lumerical.com`.