# Metal Optics Based nanoLEDs: In Search of a Fast, Efficient, Nanoscale Light Emitter

*Michael Eggleston*

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2015-122
http://www.eecs.berkeley.edu/Pubs/TechRpts/2015/EECS-2015-122.html

May 15, 2015

**Metal Optics Based nanoLEDs:**
**In Search of a Fast, Efficient, Nanoscale Light Emitter**

By

Michael Scott Eggleston

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy
in

Engineering – Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ming C. Wu, Chair
Professor Eli Yablonovitch
Professor Xiang Zhang

Spring 2015

Abstract

Metal Optics Based nanoLEDs:

In Search of a Fast, Efficient, Nanoscale Light Emitter

by

Michael Scott Eggleston

Doctor of Philosophy in Engineering – Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Ming C. Wu, Chair

Since the invention of the laser, stimulated emission has been the de facto king of optical communication. Lasers can be directly modulated at rates as high as 50GHz, much faster than a typical solid state light-emitting diode (LED) that is limited by spontaneous emission to <1GHz. Unfortunately, lasers have a severe scaling problem; they require large cavities operated at high power to achieve efficient lasing. A properly designed LED can be made arbitrarily small and still operate with high-efficiency. On-chip interconnects is an area that is in desperate need of a high-speed, low-power optical emitter that can enable on-chip links to replace current high-loss metal wires. In this work, I will show that by utilizing proper antenna design, a nanoLED can be created that is faster than a laser while still operating at >50% efficiency.

I start by formulating an optical antenna circuit model whose elements are based completely off of antenna geometry. This allows for intuitive antenna design and suggests that rate enhancements up to ~3,000x are possible while keeping antenna efficiency >50%. Such a massive speed-up in spontaneous emission would enable an LED that can be directly modulated at 100's of GHz, much faster than any laser.

I then use the circuit model to design an arch-dipole antenna, a dipole antenna with an inductive arch across the feedgap. I experimentally demonstrate a free-standing arch-dipole based nanoLED with rate enhancement of 115x and 66% antenna efficiency. Because the emitter is InGaAsP, a common III-V material, I experimentally show that this device can be easily and efficiently coupled into an InP waveguide. Experimental

coupling efficiencies up to 70% are demonstrated and directional antennas are employed that offer front to back emission ratios of 3:1.

Finally, I show that a nanoLED can still have high quantum yield by using a transition metal dichalcogenide, $WSe_2$, as the emitter material. By coupling a monolayer of $WSe_2$ to a cavity-backed slot antenna, I demonstrate a record rate enhancement for a solid state emitter of 320x. In addition, the nanoscale devices (30nm x 250nm) have a quantum yield comparable to an unprocessed $WSe_2$ monolayer. Such a fast, efficient, nano-emitter not only has the ability to reduce power consumption in central processing units (CPUs) by orders of magnitude but may also revolutionize integrated sensing and imaging applications at the nanoscale.

The dissertation of Michael Scott Eggleston, titled Metal Optics Based nanoLEDs:

In Search of a Fast, Efficient, Nanoscale Light Emitter, is approved:

Chair: _____ Date _____

_____ Date _____

_____ Date _____

University of California, Berkeley

To Peanut and Charlie

# Table of contents

# List of Figures

# List of Tables

# Acknowledgements

I would like to thank my advisor Professor Ming C. Wu for his support and guidance throughout my six year tenure as a graduate student at Berkeley. He has had the uncanny ability to give me direction when I'm lost yet give me leeway when I think I know what I'm doing. His mentorship in both academic matters and career guidance has truly been unparalleled.

I would also like to thank Professor Eli Yablonovitch for his advice and insight over the years. Our many discussions have been truly inspirational and have helped shape both my past and future research. I would like to thank Professor Constance Chang-Hasnain and Professor Xiang Zhang for serving on my qualification and thesis committees. I would also like to thank Professor Ali Javey for our many fruitful collaborations.

There are so many students and post-docs who have helped me along the way. Specifically I'd like to thank Dr. Amit Lakhani and Dr. Erwin Lau for mentoring me when I first came to Berkeley and helping me get on the right track to become a successful Ph.D. student. I'd also like to thank Dr. Matteo Staffaroni, Dr. Tae Joon Seok, Dr. Nikhil Kumar, Kevin Messer, and Seth Fortuna for their instrumental contributions to this work as well as their friendship outside of the lab. I thank the entire Ming Wu group for their constant support including Dr. Arash Jamshidi, Dr. Justin Valley, Dr. John Wyrwas, Dr. Karen Grutter, Shao Ning Pei, Dr. Alex Grine, Ryan Going, Dr. Niels Quack, Dr. Byung-wook Yoo, Sangyoon Han, Phil Sandborn, and Dr. Chang-Kyu Kim. I'd also like to thank all of the other students who have helped me along the way: Dr. Sapan Agarwal, Vidya Ganapati, Jared Carter, Samarth Bhargava, James Ferrara, Dr. Frank Rao, Dr. Owen Miller, Sujay Desai, Henry Barrow, and Robert Schneider. Without this great team of people at my side I never could have accomplished what I have.

The staff of the Marvell Nanofabrication Laboratory have my many thanks for their tireless work that has allowed me to pursue my research. They have made the Nanolab a wonderful place to work and it truly saddens me that I may never again have access to such an amazing workshop.

I would like to thank my parents and sisters for helping me to get where I am today. No one could ask for a better support structure then them; I am truly blessed to have them as my family. Finally, I would like to thank Anita for making my life complete; finding you was my greatest discovery.

# Introduction

*"A scientist can discover a new star, but he cannot make one. He would have to ask an engineer to do it for him."*

*-Gordon L. Glegg*

Since the invention of the laser in the 1960's, stimulated emission has been the de facto high-speed champion. With possible direct modulation rates up to ~50GHz, lasers form the backbone of our high-speed modern global communication network. Despite their speed, lasers have found limited use in short-scale communication networks, especially at length scales below a meter. The primary reason for this is the difficulty to efficiently scale lasers to small sizes and low power.

LEDs on the other hand, know no size limits. Atoms, molecules, and quantum dots are all very efficient light emitters that are nanoscale in nature. However, when compared to lasers, their emission rates are embarrassingly slow. Solid-state LEDs have a maximum direct modulation rate around 300MHz, much too slow for our modern computational and communication needs. Luckily, this rate is not etched in stone. Through proper engineering of a light emitter's environment, even the slowest of photon emission processes can be sped up by several orders of magnitude.

In this dissertation we will visit the physics behind stimulated and spontaneous emission and see how we can make spontaneous emission just as fast, if not faster than stimulated emission. In Chapter 1 I discuss the basics of spontaneous emission enhancement, especially in regards to optical-antenna based enhancement. In Chapter 2 I develop a circuit model that gives intuitive insight into proper optical antenna design and allows for fast device optimization without reliance on lengthy numerical calculations. In Chapter 3 I discuss how to overcome large gap capacitance with the arch-antenna. Experimental results from fabricated free-standing arch-antennas are presented and compared to expected results from our simple circuit model. In Chapter 4 I integrate the arch-antenna into an InP waveguide, and design a structure that offers high optical coupling and directional emission. In Chapter 5 I will introduce the slot antenna and cavity-backed slot antenna and derive circuits to describe their radiative properties. Finally, in Chapter 6 I discuss the use of new material systems for higher-efficiency nanoLEDs. Antennas are designed for use with $WSe_2$, a transition-metal-dichalcogenide light emitter with near-ideal surfaces. Experimental results show large spontaneous emission enhancement and high internal quantum yield.

# Chapter 1

# Light for Communication

When most people think of an LED, their thoughts turn to modern lighting where the LED is gaining ground over the inefficient and antiquated incandescent light bulb. Through the simple act of turning on a light bulb, instantly our eyes can detect an enormous amount of information about the world around us. There is little else in this world that can bring so much information at such great speed; light is truly the champion of communication. While light has been used in this most basic fashion since the dawn of time, more recently our technological advances have made light ever more needed as we are now immersed in a world where information is everywhere and demand on communication grows exponentially.

While many applications exist that could use a fast, efficient, nanoscale emitter, my focus will be on short-range optical communication. Equally important, and therefore in need of mention, are applications in biological and chemical sensing where integrated optical nano-emitters could revolutionize the detection of nanoparticles and probe the inner workings of the brain and the human body.

In this chapter I will briefly describe the historical trends in optical communication and point towards the requirements for future systems. I will then discuss the physical mechanisms that set photon emission rates and how we might go about engineering them to be much faster.

## 1.1 Optical Communication

The first global communication system was the telegraphy system that fully encircled the world by 1902. It consisted of metal wires stretching across the globe, connecting stations filled with Morse-code wielding telegraph operators. The bandwidth of these systems were very low, with long distance submarine cables only capable of ~10 words per minute. These systems were later replaced by higher data-rate telephone cables in the 1950s that used in-line repeaters to boost communication speed. One of the main issues with using metal wires for long-distance communication is the massive ohmic losses, shown in Figure 1.1(a) for modern coaxial cables, that propagating signals experience.

To keep up with increasing demand for bandwidth – especially since the creation and widespread adoption of the internet – long-range communications was taken over by optical fiber. As

demonstrated in Figure 1.1(b), optical fiber offer orders of magnitude lower optical loss with incredibly high bandwidth not attainable in metal wires.

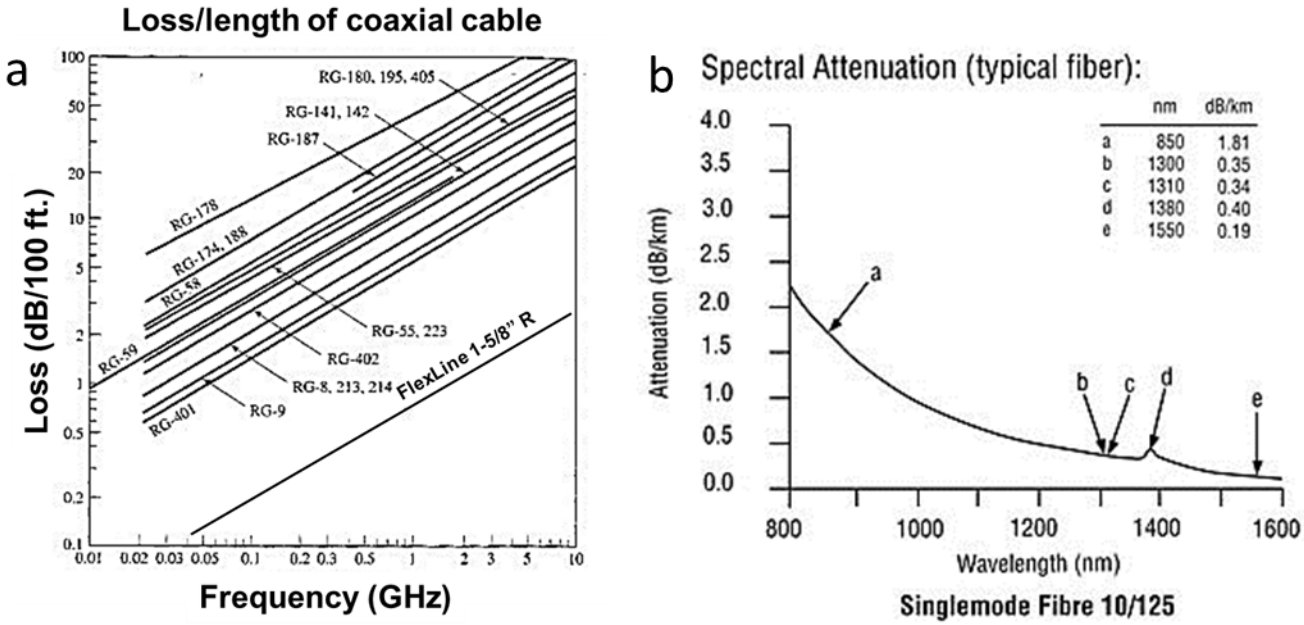


**Figure 1.1** Comparison of propagation losses in (a) coaxial cables and (b) optical fiber. Reproduced from [1].

As shown in Figure 1.1(a), higher data rates over metal wires incur increasingly higher losses. However, optical connections have always proven to be much more complicated and therefore more expensive to implement then electrical lines. The result is a slow takeover of communications from metal to optical links as data rates increase. At a certain speed, metal is no longer economical at which point optical links are implemented. This trend, nicely depicted in Figure 1.2, shows how optics has been able to take over at shorter and shorter distances as our thirst for communication has increased. The question does not really seem to be *will* optical interconnects be used on chip, but rather *when*.

The exclusive use of metal interconnects on chips has already had severe repercussions for the modern computer. Over a decade ago, in 2004, power dissipated in the metal interconnects on CPUs reached 50% of total consumed power on a chip [2]. That number is expected to rise to ~80% and by all accounts may have already reached this level. To help mitigate this massive amount of power consumption, chip designers have capped on-chip clock-speeds to ~2-3GHz and have instead gone to multi-core architecture. While this has for the time being allowed for continued scaling of total CPU computational power, it has not solved the interconnect problem. In addition, problems that cannot be parallelized no longer benefit from newer technology nodes.

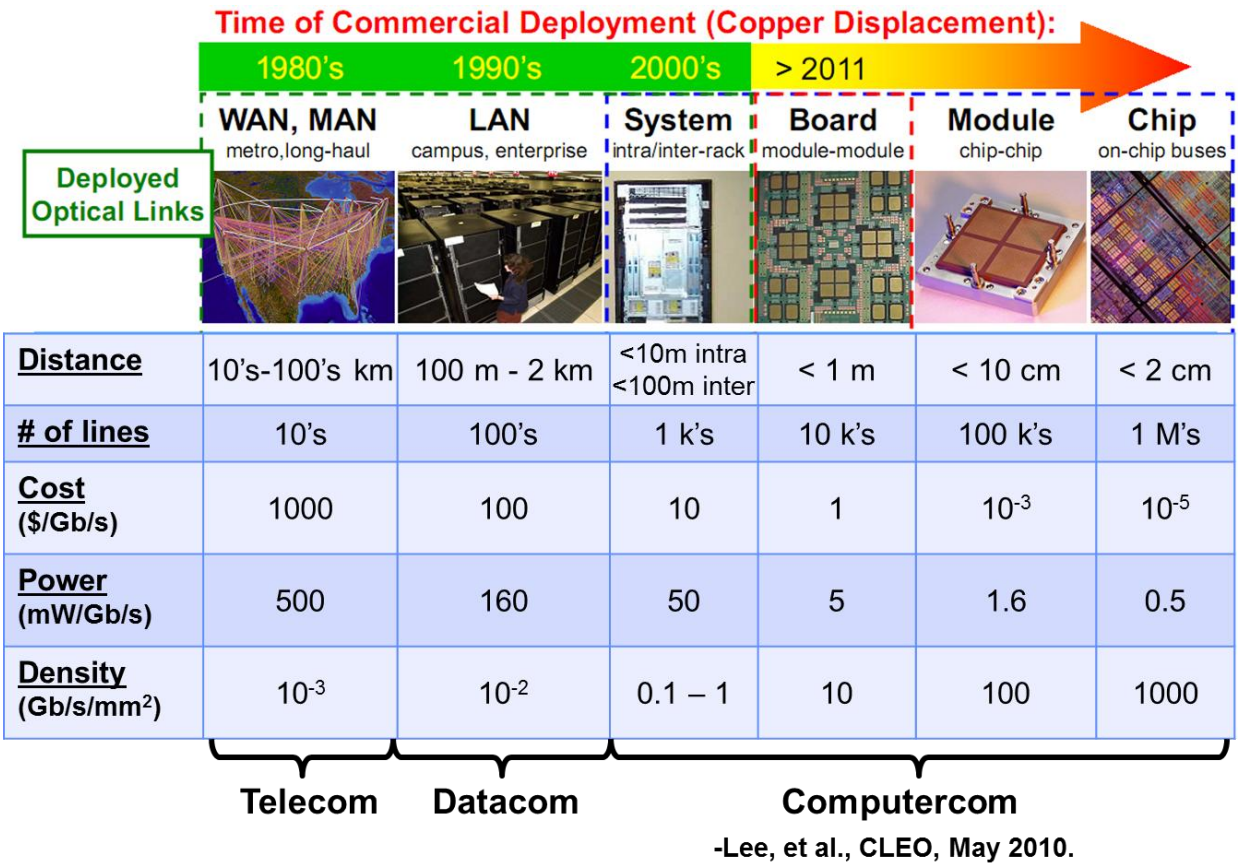


**Figure 1.2** Trend of Optical Communication moving to smaller scales. Reproduced from [3].

ITRS projections for 2022 estimate that on-chip interconnects will only have ~30 fJ/bit of energy available to them [2]. To see how far that will get you with a metal interconnect we can do a quick calculation of how much energy it takes to send one bit of data over a wire. For a transistor to signal to another transistor on chip, it must charge up the capacitance of the metal wire connecting the two. Wires have an intrinsic lower limit capacitance of about 2pF per centimeter [2]. The energy required to charge a line to a 1V signaling level is then:

$$\frac{Energy}{bit} = CV^2 = \frac{2pF}{cm} * 1V^2 = 2pJ/cm \tag{1.1}$$

If only 30 fJ/bit is available then the signal can only go 150μm, obviously far too short to communicate across a 1 cm chip. The only parameter that can be changed is the signaling voltage, which is limited by the threshold voltage of modern metal-oxide-semiconductor field-effect transistors (MOSFETs). Optical interconnects can completely circumvent this problem because optical waveguides don't have capacitance.

To calculate the fundamental lower limit of an optical link, we'll need to use a little photon statistics. On a transistor to transistor level there is no such thing as error correction, so we can't

tolerate any bit errors. If we had a bit error rate (BER) of $10^{-18}$ with $10^9$ transistors on a chip operating at $10^9$ Hz, then we would expect an error somewhere on the chip every second. It is very unlikely (if not impossible) for a CPU to actually be run at 100% like this, but to be safe lets add two orders of magnitude for a BER of $10^{-20}$. Now if I have a photo-detector that can detect single photons at 100% efficiency, the only limit to my BER is whether or not a photon is actually emitted from my optical emitter within the time frame of a bit. Most light emitters follow Poisson statistics, meaning that the probability of not emitting any photons if an average of μ photons are emitter per bit is:

$$P(N = 0) = \frac{\mu^0 e^{\mu}}{0!} = e^{\mu} \tag{1.2}$$

So if we want this probability to be less than $10^{-20}$:

$$10^{-20} = e^{\mu} \Rightarrow \mu \geq 46 \tag{1.3}$$

If we use on-off keying, then the presence of a photon is a "1" and no photons is a "0". To maintain our high BER a "1" bit, according to equation (1.3), must contain an average of 46 photons. Luckily a "0" bit contains no photons so on an average we need 23 photons per bit. If our photons have an energy of 0.8 eV (1500nm wavelength), then the average energy per bit is then just 3 aJ/bit! This is four orders of magnitude lower than the total budget, which leaves room for non-ideal devices, waveguide loss, and less than perfect detectors.

## 1.2 Contenders for a Fast, Efficient, Nanoscale Emitter

Even though the quantum limit for optical communication efficiency is about 3 aJ/bit, current state of the art optical links use closer to ~1 pJ/bit [4]. A device that we could use to communicate on a chip and get to this fundamental limit would need to fit three criteria. First, it needs to be fast. Current CPU's run in excess of 3GHz; a competing optical technology would need to at minimum reach this speed, if not into the many tens of GHz or faster. Second, the device must be efficient or it will never be able to reach low energy per bit operation. Finally, the device must be nanoscale. Ultimately size will not only limit where you can put the device, such as integrated with a <100nm sized transistor, but also how low of power it can operate at.

The first device that might come to mind is a laser, which has been the traditional work horse for high-speed energy-efficient optical links. Dramatic advances in semiconductor processing and packaging technology have been able to shrink the laser to smaller sizes, bringing its advantages of high power efficiency and speed to shorter and shorter ranges [5]. However, lasers require a large photon density in order to both sustain lasing and reach fast modulation rates, as I will show in section 1.3. To achieve high photon densities, high-Q cavities are typically used to trap photons for a long period of time. Dielectric cavities such as photonic crystals [6–8] and microdisks [9,10] have been able to push lasers to the diffraction limit. The record for most

energy-efficient laser is currently held by a photonic crystal laser operating at a mere 4.4 fJ/bit [8]. Even this heroic demonstration still only achieves an operating efficiency of 5% and a size >30μm$^2$.

Metal has been used to shrink lasers below the diffraction limit and into the nanoscale [11–15]. Unfortunately, high losses in metal-based cavities have made it exceedingly difficult to achieve efficient lasing at these tiny dimensions [16]. Most demonstration come from devices on the order of a micron or more. It is still unknown whether metal optics will ever be able to yield an efficient laser at a sufficiently small size.

Another device that has received very little attention in the optical interconnect world is the light emitting diode (LED). LEDs are not limited by low-Q cavities and can operate efficiently well below threshold. If well engineered, they can be made very small and still maintain good characteristics. Unfortunately, their modulation speeds are limited by the relatively slow process of spontaneous emission; relegating typical solid-state LEDs to a maximum 300MHz modulation rate. The rate of spontaneous emission, however, is not set in stone. Through careful engineering of the optical environment, spontaneous emission could be faster than a laser [17].

## 1.3 Spontaneous and Stimulated Emission

Before we can design a fast emitter, we need to know what determines photon emission rate. There are many ways to derive the rates of stimulated and spontaneous emission. The two most common ways are by use of Einstein's A and B coefficients [18] and by using Fermi's Golden Rule, which can be derived from time-dependent perturbation theory [19]. I prefer to use Fermi's Golden Rule because I believe it gives deeper insight into spontaneous emission and clearly shows how it is intrinsically linked with stimulated emission. Fermi's Golden Rule gives the transition rate between an initial (*i*) and final (*f*) state in units of number per unit time per unit volume as:

$$R_{fi} = \frac{2\pi}{\hbar}\left|M_{fi}\right|^2 \rho(\hbar\omega) \tag{1.4}$$

Where $M_{fi}$ is the matrix element describing the transition from initial to final state, $\rho(\hbar\omega)$ is the optical density of states, and $\hbar\omega$ is the energy corresponding to the transition between the initial and final states. By and far the most common source of light is from dipole transitions, so I will focus on these transitions in this dissertation. The matrix element for a dipole transition can be written as:

$$M_{fi} = \left\langle \left(n_p + 1\right), f \middle| q\boldsymbol{x} \cdot \boldsymbol{\mathcal{E}} \middle| n_p, i \right\rangle \tag{1.5}$$

Here $q\boldsymbol{x}$ is the dipole moment, $\boldsymbol{\mathcal{E}}$ is the electric field interacting with the dipole, and $n_p$ is the number of photons of frequency $\omega$ in the field. Using the notation of second quantization (and

using MKS units since this is an Engineering dissertation) we can write the electric field operator as:

$$\boldsymbol{\mathcal{E}} = i\sqrt{\frac{\hbar\omega}{2\epsilon V}}\ \ \left(\boldsymbol{a}^{\dagger}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} + \boldsymbol{a}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\right)\hat{\boldsymbol{e}} \tag{1.6}$$

$\epsilon$ is the dielectric permittivity of the surroundings, $V$ is the volume of space the dipole is in, $\boldsymbol{a}^{\dagger}$ and $\boldsymbol{a}$ are the creation and destruction operators respectively, and $\hat{\mathbf{e}}$ denotes the polarization direction. Plugging this into equation (1.5):

$$M_{fi} = \left\langle \left(n_p + 1\right), f \middle| q\boldsymbol{x}\cdot i\sqrt{\frac{\hbar\omega}{2\epsilon V}}\ \ \left(\boldsymbol{a}^{\dagger}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} + \boldsymbol{a}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\right)\hat{\boldsymbol{e}} \middle| n_p, i \right\rangle \tag{1.7}$$

Since we are interested in emission and not absorption, we can drop the term associated with the destruction operator (i.e. the term that takes a photon out of the field). We will also assume that the dipole and electric field are in the same direction. We can simplify the equation by separating the wave-function into a term for the field and a term for the electrical states and then use the electric dipole approximation [19] due to the fact that the dipole transition length is much smaller than a wavelength ( x << 2π/k):

$$M_{fi} = i\sqrt{\frac{\hbar\omega}{2\epsilon V}}\left\langle \left(n_p + 1\right), f \middle| qx\left(\boldsymbol{a}^{\dagger}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right) \middle| n_p, i \right\rangle$$

$$= i\sqrt{\frac{\hbar\omega}{2\epsilon V}}\left\langle \left(n_p + 1\right) \middle| \boldsymbol{a}^{\dagger} \middle| n_p \right\rangle \langle f|qx|i\rangle \tag{1.8}$$

The first bra-ket can now be easily evaluated:

$$M_{fi} = i\sqrt{\frac{\hbar\omega}{2\epsilon V}}\sqrt{\left(n_p + 1\right)}\langle f|qx|i\rangle \tag{1.9}$$

It's interesting to note that even if there are no photons in the field ($n_p$ = 0) the matrix element – and therefore the electric field – is non-zero. This comes from the fact that even with no photons present there are still zero-point fluctuations in the field. These fluctuations are often described as a background of photons randomly being created and destroyed, which when time averaged gives an average energy of $\hbar\omega/2$ in the empty field. This plays a crucial role in spontaneous emission as discussed later.

We must now multiply equation (1.9) by its complex conjugate to get $|M_{fi}|^2$:

$$|M_{fi}|^2 = \frac{\hbar\omega}{2\epsilon V}(n_p + 1)|\langle f|qx|i\rangle|^2 = \frac{\hbar\omega}{2n^2\epsilon_0 V}(n_p + 1)|\langle qx\rangle|^2 \tag{1.10}$$

Using *n* as the refractive index of the dielectric medium and employing the expectation value $\langle qx\rangle$ for the matrix element. The final piece we need to evaluate equation (1.4) is the optical density of states. I won't derive it here but it can be found in most optics text books [18]:

$$\rho(\hbar\omega) = \frac{n^3(\hbar\omega)^2}{\pi^2\hbar^3c^3} = \frac{n^3\omega^2}{\pi^2\hbar c^3} \tag{1.11}$$

Taking into account that only 1/3 of those states are in the $\hat{\boldsymbol{e}}$ direction, we can plug equation (1.10) and (1.11) into equation (1.4) which gives us:

$$R_{fi} = \frac{2\pi}{\hbar}\frac{\hbar\omega}{2n^2\epsilon_0 V}(n_p + 1)(qx_0)^2\frac{n^3\omega^2}{3\pi^2\hbar c^3} = \frac{\omega^3 n}{3\pi\epsilon_0\hbar c^3 V}(n_p + 1)|\langle qx\rangle|^2 \tag{1.12}$$

For an emitting volume V this simplifies to a rate of:

$$\frac{1}{\tau} = \frac{\omega^3 n}{3\pi\epsilon_0\hbar c^3}(n_p + 1)|\langle qx\rangle|^2 \tag{1.13}$$

The beauty of this derivation is that equation (1.13) has both the contributions of spontaneous *and* stimulated emission. If there are no photons in the field then the emission cannot be stimulated. For $n_p$ = 0, this leaves only spontaneous emission, which can be thought of as *emission stimulated by the vacuum field*. The rate of spontaneous emission is therefore:

$$\frac{1}{\tau} = \frac{\omega^3 n}{3\pi\epsilon_0\hbar c^3}|\langle qx\rangle|^2 \tag{1.14}$$

Instantly we can see why spontaneous emission might be more attractive for low-energy systems. To get stimulated emission you need photons to already exist in your device. In practice, lasers need to operate well above threshold – the point where there is more stimulated emission than spontaneous emission – in order for there to be enough photons to make the emission fast. It also demonstrates that a directly modulated laser can only be fast if the modulation depth is small, i.e. it never comes close to turning off.

## 1.4 The Purcell Effect

Now that we have an understanding of what determines spontaneous emission rates, we can start work on enhancing that rate. In the 1940's, Edward Purcell discovered that emitters placed within a cavity decay at an increased rate [20]. The subsequently termed "Purcell Effect" is an often used tool to predict the spontaneous emission rate of an optical emitter placed within a

cavity. It is in effect, an extension of Fermi's golden rule that instead of directly solving for emission rate, gives a ratio, *F*, of rate enhancement compared to an emitter in free space. To do this we must determine the density of states of an arbitrary "cavity" in which we can place our emitter. If we assume the cavity supports an optical mode of effective mode volume, *V*, a quality factor, *Q*, and has a Lorentzian line shape, than the density of states becomes:

$$\rho_{cavity}(\hbar\omega) = \frac{1}{V}\frac{\mathrm{d}N}{dE} = \frac{1}{V}\frac{2}{\pi}\frac{1}{\hbar\Delta\omega} = \frac{2}{\pi\hbar\omega}\frac{Q}{V} \tag{1.15}$$

Where we have made the substitution for number of states per unit energy, dN/dE, as the value of the lorentzian line shape at its maximum, $2/\pi\hbar\Delta\omega$. We can then substitute the spectral width of the cavity, $\Delta\omega$, with the cavity $Q = \omega/\Delta\omega$. From this equation it is straightforward to derive the Purcell Factor, F, giving the increased radiation rate of our emitter in this cavity:

$$F = \frac{\frac{2\pi}{\hbar}\left|M_{fi}\right|^2\rho_{cavity}(\hbar\omega)}{\frac{2\pi}{\hbar}\left|M_{fi}\right|^2\rho_{free}(\hbar\omega)} = \frac{\frac{2}{\pi\hbar\omega}\frac{Q}{V}}{\frac{\omega^2}{\pi^2\hbar c^3}} = \frac{3\lambda^3}{4\pi^2}\frac{Q}{V} \tag{1.16}$$

Equation (1.16) can be simplified even further if we normalized the mode volume to a cubic half-wavelength, $V_n = V/(\lambda/2)^3$. A normalized mode volume of one corresponds to the smallest modal volume we can get with traditional optical cavities, such as an enclosed box of mirrors.

$$F = \frac{6}{\pi^2}\frac{Q}{V_n} \tag{1.17}$$

We can see from this equation we need very high Q-factors to achieve large enhancements in wavelength or larger scale devices. An alternative is to engineer the mode volume to be significantly sub-wavelength so that the denominator in equation (1.17) is sufficiently small. If we assume that our emitter can modulate at 1GHz in free space, than by placing it in a cavity its maximum modulation rate, following equation (1.17), is shown in Figure 1.3. Note that at large Q the dependence on mode volume saturates and *F* becomes completely dependent on the Q factor. This is due to the modulation rate being determined by the cavity lifetime, $Q/\omega$, rather than the decay time of the emitter itself. Another way to say this is that if the Q is too high, our photon is emitted from our source very quickly, but it is then stuck in the cavity for a very long time. For each mode volume there is some optimum value of Q at which we can achieve the highest modulation rate.

It should be noted that the plot in Figure 1.3 assumes our optical emitter emits as a delta function in frequency. In a real system it will have some bandwidth $\Delta\omega_{emitter}$. If our Q factor is too high, $\omega/Q < \Delta\omega_{emitter}$, the cavity density of states will not fully overlap our emitter density of states and the modulation bandwidth predicted in the high Q regime of Figure 1.3 is no longer valid [21].

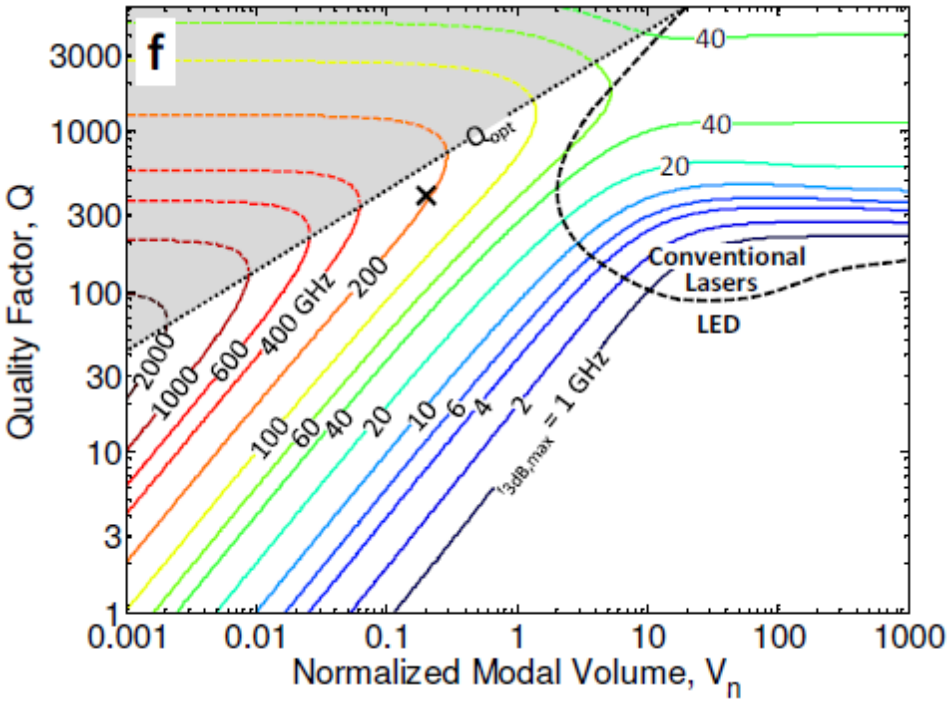


**Figure 1.3** Modulation bandwidth of an optical emitter placed in a cavity with quality factor Q and normalized mode volume $V_n$. Reproduced from [17].

Figure 1.3 gives us some insight into the type of cavity we should design. In general, the highest modulation rates are obtainable in the very low mode volume regime. Meanwhile, the high-Q regime that lasers normally occupies, offers many disadvantages such as long cavity lifetimes and very narrow density of states. Unfortunately, it gives us almost zero guidance into how to design such a cavity. What does a mode with significantly sub-wavelength dimensions look like? While we realize one potential idea is to use metal optics to squeeze the light into a small volume, how do we determine the "mode volume" of a non-enclosed metal structure such as an optical antenna or even a flat metal sheet?

In addition to the ambiguity of mode volume in deep sub-wavelength structures, the traditional Purcell mindset offers little to us as a device designer. Equation (1.17) can mislead us to believe that Q and mode volume are two independent variables we can tune to maximize enhancement. This, however, is almost never the case. So how do we design, much less characterize our structure to give us large spontaneous emission rate enhancements? Luckily, the problem of radiating energy at high rates was solved a very long time ago by RF antenna designers.

## 1.5 Modeling Spontaneous Emission as an Antenna

Let's now forget the previous section and start from scratch. We have an optical emitter that can be characterized by some transition matrix element. In almost all cases, this matrix element will take the form of a dipole, $qx_o$, where $x_o$ is the characteristic length describing the positional

change in our emitter wavefunction between the excited state to the ground state. In general, this value needs to be evaluated using quantum mechanics and the full knowledge of the wavefunctions of the emitter we are using. It turns out if we just want enhancement, the actual value doesn't really matter. However, for a solid state emitter we can assume this value is on the order of a lattice spacing (~5Å), since optical emission is a local process.

Now that we have the dipole moment of our emitter, we can borrow a little antenna theory to determine how fast the power will be radiated. For a dipole of length $x_o$, the radiated power is given as [22]:

$$P_{rad} = \frac{\pi}{3} Z_0 \left(\frac{x_0}{\lambda_0}\right)^2 (q\omega)^2 \tag{1.18}$$

Obviously the power will be influenced by how many photons are actually being emitted, which will be partially determined by how many electron-hole pairs are excited in our material. It is then much easier to write this as emission rate per photon by dividing by the energy per photon, $\hbar\omega$:

$$\frac{1}{\tau_{rad}} = \frac{P_{rad}}{\hbar\omega} = \frac{\pi}{3\hbar\omega} Z_0 \left(\frac{q\omega x_0}{\lambda_0}\right)^2 = \left(\frac{2\pi^2}{3}\right)\left(\frac{q^2}{\hbar}\right)(Z_0)\frac{x_0^2}{\lambda_0^3} c \tag{1.19}$$

This can be compared to equation (1.14) to see that the two forms are identical if we set $\langle x \rangle = \frac{x_0}{2}$.

By writing equation (1.19) in this form we can instantly see the factors important to the rate of emission. The first three terms are all fundamental constants, where $q^2/\hbar$ is the fine structure constant and $Z_0$ is the impedance of free space. The only variable parameters are the dipole length, $x_o$, and the emission wavelength, $\lambda_o$ which are both material dependent. In fact it is the ratio of the square of the dipole length to the wavelength cubed multiplied by the speed of light that sets the emission rate. Since there is a huge mismatch between $x_o$≈5Å and $\lambda_o$≈1μm, this rate is very slow for most solid-state emitters, ~1ns. This corresponds to a very large $Q=\omega\tau \approx 300{,}000$.

This result should not be too surprising to us considering Wheeler's limit [23], which states that the lowest obtainable Q for a small antenna is:

$$Q > \frac{3}{4\pi^2}\left(\frac{\lambda_0}{a}\right)^3 \tag{1.20}$$

Where $a$ is the longest dimension of the antenna; in our case $a = x_o$. So really the spontaneous emission rate from a semiconductor material is slow because its dipole length is far too small to act as an efficient radiator.

This formalism now gives us some idea on where we can start our device design to enhance the spontaneous emission rate. The dipole length is the only parameter in equation (1.19) that we can

actually change (assuming we've already picked our emission wavelength). From equation (1.20) we also know that this dipole is far too short to be a good radiator. This makes our choice of design very simple; all we have to do is make it longer! We can't, however, change the material property $x_o$ itself (at least not by any appreciably amount), so any method we use to increase the dipole length must be *external* to the material. This is exactly what antennas have been doing for over a century: providing a large dipole moment to small circuits to enhance the radiation of their electric energy into free space!

What's truly amazing is that nearly 100 years went by before antennas began to be used to help extract optical frequency radiation from very small sources [24] such as dye molecules [25–31] and quantum dots [32–37]. Unfortunately, many attempts to use metal to enhance spontaneous emission rates have relied on surface plasmon [32,38–41] or gap plasmon structures [29,30] that are far from ideal antennas, resulting in low radiation efficiencies and large ohmic-losses. In addition, early work often achieved high enhancement by attaching small antennas to a scanning tip to excite emitters dispersed on a surface [42–45], yet these type of structures are far from integrated or mass producible. Work with non-quantum dot based semiconductors [36,38–42,46–50] have typically suffered from poor antenna-emitter coupling.

The remainder of this thesis will focus on the design and implementation of a semiconductor based nanoLED with antenna enhanced spontaneous emission. While experiments with dye molecules typically have achieved the highest rate enhancements due to ease of fabrication and high quantum yield emitters, they cannot be directly modulated and therefore would not be useful as an integrated nanoemitter. By utilizing top-down fabrication methods and antenna designs, we help insure that the devices we design will allow for large scale fabrication, crucial for any nanoemitter to be used on a chip-scale network.

# Chapter 2

# Optical Antenna Circuit Model

At this point we've realized that if we want to make spontaneous emission faster, we must provide a proper external antenna to our optical emitters to help them radiate into free space. But how do we know what a "proper" antenna is? In this chapter I will go through the derivation of a circuit model to describe the radiation characteristics of a linear dipole antenna. The goal here is to create a model that will give us intuitive insight into how to design our antenna. To do this we must do our best to relate everything to something physical, then our model will have the power to tell us what to physically change in order to get the results we want. Our model will include three main components: dipole current induced in the antenna, reactive response of the metal antenna, and resistances i.e. loss pathways within the antenna. Although the model is general in terms of wavelength, all figures and calculated values assume a resonant frequency of 200THz (1500nm) – consistent with the peak emission wavelength of InGaAsP – unless otherwise noted.

## 2.1 Coupling to an Optical Dipole Emitter

The first step in designing our antenna is to determine how we are going to couple our optical emitter to an antenna. Since our antenna will be very close to the emitter but not actually electrically connected, our two choices are to employ either an inductive or capacitive coupling. Since it is an electric dipole we are trying to couple, capacitive coupling seems like the natural choice.

The key to good capacitive coupling is to place the electric dipole so it sees the highest possible electric field from the antenna mode it is trying to couple to. The fundamental mode of a dipole antenna is shown in Figure 2.1(a). The largest electric field is at the tips of the antenna where the field lines crowd together, therefore this could be a good place to put our emitters. However, as derived in equation (2.3), this sort of structure is much like a lightning rod. To get higher fields, therefore better coupling, the ends need to be more and more pointed. This turns out to be very difficult to do at the nanoscale, especially with soft metals like gold and silver.

Another strategy would be to place the emitter near the center alongside the wire, however we can see that the electric field here is very weak and therefore the coupling will not be strong. Really, the best location for good coupling would be at the very center of the antenna, *inside* the

metal. Unfortunately, this region is obviously not accessible to our emitter. However, by cutting the wire in two and forming a small gap between the two halves, we can maintain the same antenna mode as the unbroken wire while moving the high-field region outside the metal (Figure 2.1(b)). This high-field region is called the feedgap, and is a common perturbation of dipole antennas. As shown in Figure 2.1(b), the electric-field in the gap is 4x stronger than the field near the tips. This will give a 16x better coupling as well as allow for even higher coupling with smaller gaps.

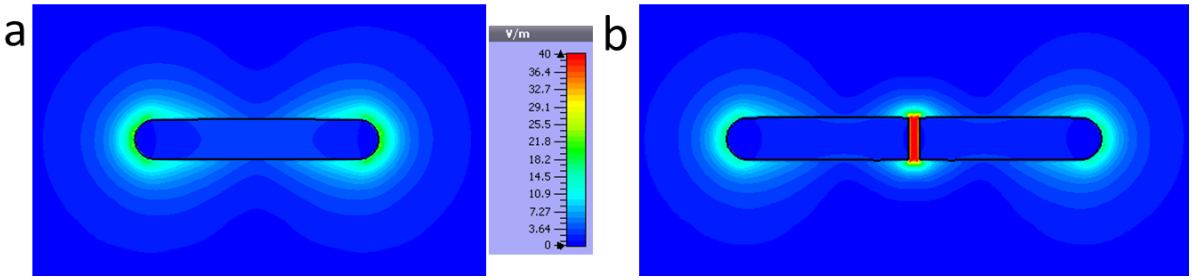


**Figure 2.1** Electric field profile of a dipole antenna both without (a) and with (b) a feedgap.

Now that we have two candidate positions for dipole coupling, shown in Figure 2.2, we can now calculate their coupling strength. To do this, we will employ Ramo's theorem [51], also known as Ramo-Schockely's theorem. In simple terms, this theorem states that the current induced on a pair of metal electrodes by a moving charge can be determined by the charge's velocity and the DC electric field it would see if the electrodes were charged to a voltage $V$. The velocity of an oscillating dipole is merely $\omega x_o$ if its motion is sinusoidal at frequency $\omega$. We are then just left to find the electric field of the specific geometry we are using to couple. In Figure 2.2(a) we can see the feedgap forms a parallel-plate capacitor with plate spacing $d$. We know that for a parallel-plate capacitor the electric field in the gap is a constant $V/d$. Employing Ramo's theorem:

$$I_0 V = qv\mathcal{E}$$

$$I_0 V = q\omega x_0 \frac{V}{d}$$

$$I_0 = \frac{q\omega x_0}{d} \tag{2.1}$$

From equation (2.1) we can see that for a parallel-plate type coupling an emitter at any point in the gap will induce a current on the antenna proportional to $x_o/d$. Therefore to get as large as coupling possible, we will want to shrink the gap to as small as possible. The implications of this will be explored in section 2.4.

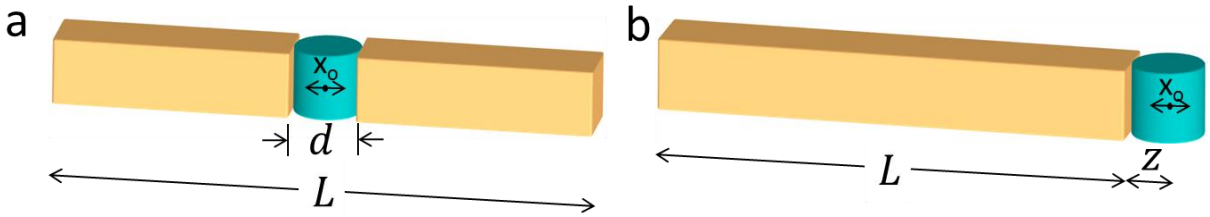


**Figure 2.2** Coupling of a semiconductor with dipole length $x_o$ to an optical antenna of length L. (a) parallel-plate type coupling in the antenna feedgap. (b) end-coupling to an antenna lacking a feedgap.

In general, this procedure can be used for any geometric configuration. As another demonstration let's look at the case shown in Figure 2.2(b). The electric field off the end of the wire can be approximated as a single plate capacitor. If we assume it is round with a radius r, we can approximate the electric field as that given by a disk of charge [52]:

$$\mathcal{E} = \frac{V}{r}\left[1 - \frac{z}{\sqrt{z^2 + r^2}}\right] \tag{2.2}$$

Where z is the distance from the end of the antenna to the center of the dipole along the antennas major axis. We can then calculate the induced current as:

$$I_0 = 2q\omega\frac{x_0}{r}\left[1 - \frac{z}{\sqrt{z^2 + r^2}}\right] \tag{2.3}$$

We can see that now the coupling, $x_o/r$, is inversely proportional to the radius of the wire. Also, the coupling is fairly strong near the surface, but falls of quickly as z becomes comparable to the wire radius. This equation is very similar to that of a lightning rod [53]. Unfortunately, this type of coupling not only requires the emitter to be very close to the antenna, but also requires a very thin wire to be used. As will be discussed in section 2.3 and 2.4, this will significantly increase both the kinetic inductance and ohmic losses in the antenna, leading to much lower efficiencies.

## 2.2 Inductance and Capacitance of Metal Structures

Every metallic structure is resonant at some frequency and is, in effect, an LC resonator with distributed inductance L and capacitance C. The oscillating ac currents associated with the LC resonators can radiate electromagnetic energy. Thus all metal objects act as antennas to some degree, converting ac currents to radiated power. The lowest LC resonant frequency of a metallic object is usually related to its size. That is only first of many electromagnetic resonances that extend all the way to high optical frequencies, at which point the resonances could be heavily influenced by kinetic inductance and assume some plasmonic character.

We shall stick to the lowest LC resonant frequency as we construct our circuit model. Doing this allows us to view our distributed inductance and capacitance as lumped components. While we obviously lose information about higher order modes, it greatly simplifies the circuit. Note that if we did want to model higher order modes, we would merely need to break our lumped components into several distributed pieces.

The lumped inductance ($L_f$ , $L_k$) and capacitance ($C_A$) values of a simple wire are fairly easy to calculate, and if you are interested in doing so I highly recommend reading *Sokolnikoff and Friis* [22]. These values can be found in Table I and are also shown schematically on a metal wire in Figure 2.3(a). Remarkably, both the inductance and capacitance are almost linearly proportional to the length of the wire. Each parameter also carries a logarithmic term that depends on the length, but this varies little within the normal parameter space of possible antenna lengths.

In metal optics, a lot of attention is often given to 'plasmonics' for its ability to confine light to deep sub-wavelength volumes. Plasmons are just alternating currents in metal at optical frequencies. The physics behind plasmons and alternating current at low frequencies are almost identical. The only difference at optical frequency comes from the fact that electrons have mass, so when electric fields induce charges to move within a metal, part of the field energy is stored in the kinetic energy of the electrons. This phenomenon takes the form of an inductance, and comes from the imaginary part of the ohmic resistance [54]. Because this inductance is a result of kinetic energy instead of magnetic energy, it is called *kinetic inductance*. Although this is also true at low frequency, the electron kinetic energy at low frequencies is so low it can be neglected.

Looking at Table I we can derive the fundamental resonance frequency of a wire:

$$\omega_0 = \frac{1}{\sqrt{LC_A}} = \frac{1}{\sqrt{\left(\frac{l_{eff}}{A} Im\{\rho\} + \frac{\mu_o}{\pi^2} l \times \ln\left(\frac{l}{r}\right)\right)\frac{\epsilon_A l}{\ln\left(\frac{l}{r}\right)}}} = \frac{1}{l}\frac{1}{\sqrt{\left(\frac{Im\{\rho\}}{A} + const\right) const}} \qquad (2.4)$$

Where I have taken the approximation $l_{eff} = l$. We can instantly see that for a given frequency $\omega_o$ we can confine the mode to a length of wire $l$:

$$l = \frac{1}{\omega_0}\frac{1}{\sqrt{\left(\frac{Im\{\rho\}}{A} + const\right) const}} \qquad (2.5)$$

We can now increase the kinetic inductance (i.e. make it more 'plasmonic') either through using a material with higher $Im\{\rho\}$ or by shrinking the cross-sectional area. By doing this, the mode centered at $\omega_o$ is confined to an ever decreasing size $l$ and we therefore show that through 'plasmonics' we can get highly focused fields. If we care about the efficiency of this device, however, we must take a closer look at the whole picture. We can see in Table I that the kinetic

inductance, $L_k$, has the same dependence on geometry as ohmic resistance. In general, tuning a structure for higher kinetic inductance, i.e. making it more plasmonic, increases the ohmic losses and therefore decreases overall efficiency. It is very important to understand that optical antennas are not fundamentally plasmonic. In fact, making them more plasmonic can often reduce their effectiveness.

**Table I: Dipole Antenna Circuit Parameters**: Circuit parameters of wire antennas; where r = wire radius, $l$ = antenna length, $l_{eff}$ = effective antenna length accounting for current→0 at ends; for a half-wave antenna $l_{eff}$ =0.64$l$; $A$=antenna wire cross-sectional area, $Z_o \equiv \sqrt{(\mu_o/\varepsilon_o)}$, the impedance of free space; $x_o$=optical ac peak dipole moment length centered in a gap-spacing $d$; $\beta d$ is the diameter on the adjacent electrodes over which the dipole currents spread, ($\beta$=1.6 for flat electrodes); $n_A$ & $\epsilon_A$, the refractive index and dielectric constant of the medium surrounding the antenna; $\epsilon_m$=metal relative dielectric constant; $\epsilon_g$=dielectric constant in the gap.

| Gap Capacitance | | Current Induced in Antenna | |
|---|---|---|---|
| Blunt Tips [55] | $C_{gap} = \frac{\epsilon_g A}{d}$ | Parallel plate current induced by ac dipole [51] | $I_0 = \frac{q\omega x_o}{d}$ |
| Round Tips [56] | $C_{gap} = \frac{\pi\epsilon_g r}{2}\left[ln\left(\frac{r}{d}\right) + 2\gamma\right]$ | | |
| Lumped Reactance Terms | | Resistance Terms | |
| Faraday [22] Inductance | $L_f = \frac{\mu_o}{\pi^2} l \times \ln\left(\frac{l}{r}\right)$ | Radiation [22] Resistance | $R_{rad} = \frac{2\pi}{3} Z_o \left(\frac{l_{eff}}{\lambda}\right)^2 n_A$ |
| Kinetic [54] Inductance | $L_k = \frac{l_{eff}}{A} Im\{\rho\} = \frac{l_{eff}}{A} Re\left\{\frac{1}{\omega^2 \epsilon_o (1-\epsilon_m)}\right\}$ | Ohmic [54] Resistance | $R_{ohmic} = \frac{l_{eff}}{A} Re\{\rho\} = \frac{l_{eff}}{A} Im\left\{\frac{1}{\omega \epsilon_o (1-\epsilon_m)}\right\}$ |
| Antenna [22] Capacitance | $C_A = \frac{\epsilon_A l}{\ln\left(\frac{l}{r}\right)}$ | Spreading [57] Resistance [54] | $R_{spread} = \frac{1}{\beta d} Im\left\{\frac{1}{\omega \epsilon_o (1-\epsilon_m)}\right\}$ |

A simple wire, as discussed in the previous section, is difficult to couple to. In general, much better coupling can be achieved if a small gap, referred to as the feedgap, is formed at the center of the wire. Though this perturbation seems small, it introduces a somewhat more complicated network of capacitances in our circuit model, as shown in Figure 2.3(b). The first addition is the capacitance added from the parallel-plate geometry formed at the gap, called $C_{gap}$, which bridges the two halves of the wire. In addition, there is also stray capacitance in parallel to $C_{gap}$ written as $\alpha C_A$. This capacitance takes into account the distributed capacitance across the gap that couples the two arms. Before the gap was added, this term was effectively shorted out, leaving the outer capacitance the dominant term. In the presence of the gap it is no longer shorted and therefore plays a significant role. For all subsequent analysis, the constant α will be assumed to be one. This is not necessarily always the case, especially if the wire is non-uniform along its length.

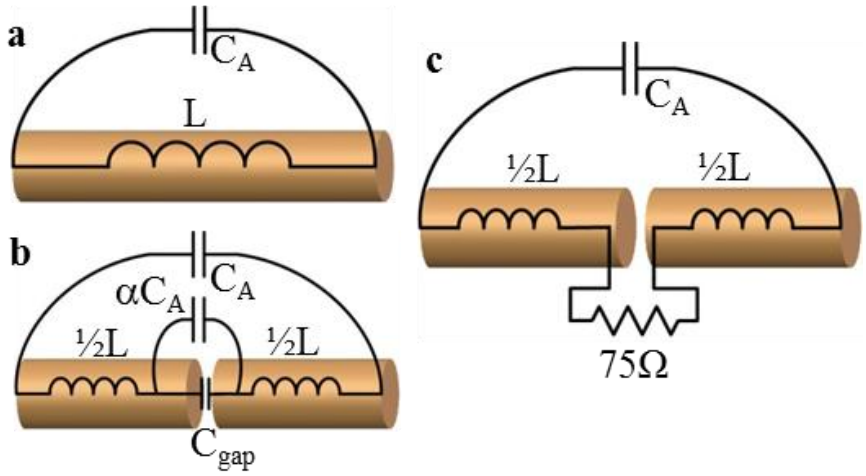


**Figure 2.3** Circuit model of (a) simple wire, (b) a wire with a small gap in the center, and (c) a wire with a feedgap fed by a 75Ω line.

Finally, it is important to touch on the structure most often encountered in RF antennas, shown in Figure 2.3(c). Typical RF antennas are driven with a 75Ω line at the feedgap, which is significantly lower resistance than either $C_{gap}$ or $\alpha C_A$ at typical RF antenna dimension. This effectively shorts out the gap, making it look much more similar to Figure 2.3(a) than to Figure 2.3(b). The resonance frequencies of these three antenna geometries are shown in Figure 2.4.

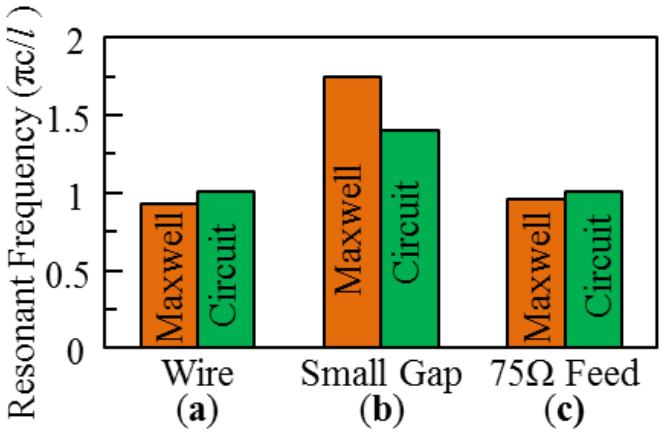


**Figure 2.4** Resonant frequency of the three antenna geometries shown in Figure 2.3. The orange bar corresponds to the resonance frequency predicted with a full 3D FDTD simulation and the green bar is from the circuit model.

For both the wire antenna and the small gap antenna fed with a 75Ω line, the circuit model is very close to the actual resonance frequency predicted with a full 3D FDTD simulation. For the small gap antenna the circuit model predicts a much higher frequency than a wire antenna of similar length. This trend is also seen in the simulation, though the error between the circuit

model and simulation is slightly higher than before, ~20% compared to ~3% for the wire and 75Ω feed. A closer look at the small gap antenna reveals an interesting dynamic for different gap spacing and wire diameters (Figure 2.5).

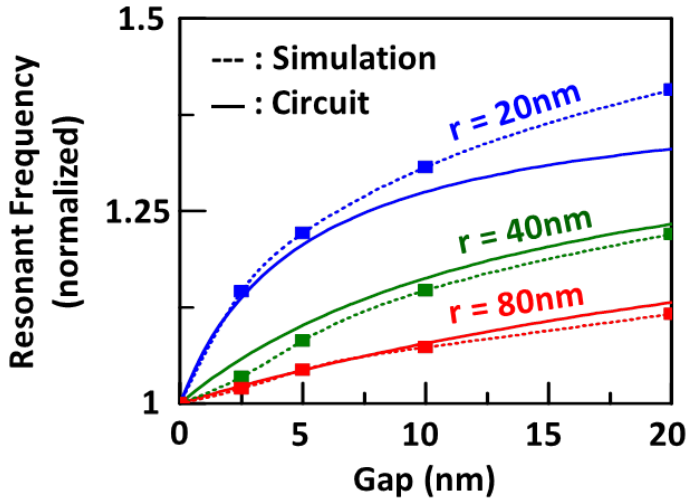


**Figure 2.5** Resonant frequency versus gap spacing for an antenna tuned to ~200THz. The solid line is what the circuit in Figure 2.3(b) predicts and the solid squares connected by a dotted line are from a full 3D FDTD simulation. Three different wire radii are used: 20nm (blue), 40nm (green), and 80nm (red).

As the gap goes to zero the gap impedance also goes to zero and in all cases the dipole effectively looks like a single metal wire. For the wider antennas, this gap capacitance is much larger and therefore the gap looks shorted out even at larger gap dimensions, especially compared to the thinner wires. At large gap spacing, specifically when $d > r$, the parallel plate capacitor approximation breaks down and the predicted resonance frequency using the circuit model starts to deviate from actual simulated values.

This model ignores the fact that the ends of the wires have a considerable amount of capacitance, especially for the wide wires. This term will ultimately red shift the antenna resonance frequency as shown in Figure 3.4(a). However, since we are primarily interested in the effect of gap width on antenna performance, we can ignored it because the end capacitance is constant for all gap spacings. When fabricating antennas we must remember to make them shorter than what our circuit model predicts since this neglected capacitance makes their electrical length longer than their physical length.

## 2.3 Loss in Metal

The most important part of the antenna is of course the loss pathways, since this will determine where the energy fed into it ultimately goes. These loss pathways take the form of resistors in our

circuit model, since they dissipate energy. The three main resistances, shown in Figure 2.6 and Table I, are the radiation resistance ($R_{rad}$), ohmic resistance ($R_{\Omega}$), and spreading resistance ($R_{spread}$).

Radiation resistance is a desirable loss mechanism as this represents the energy radiated into free space. Derived in most antenna books [22,58] and shown in Table I, it is a function of the ratio of antenna length to emission wavelength, squared, as well as the refractive index of the medium it is in. We will always assume the antenna is in vacuum unless explicitly stated otherwise. It may seem like we always want a large radiation resistance to maximize the radiated power. However, as we will explore in section 2.4, this is not always the case and a full circuit analysis is usually required.

Ohmic resistance gives rise to unfavorable energy lost as heat in the metal. For a simple wire antenna this resistance takes the familiar form, $R_{\Omega} = \rho\frac{l_{eff}}{A}$. This has a strong dependence on geometry. By making the antenna wider the ohmic resistance is reduced and vice versa. As mentioned in section 2.2 this is the same functional form as the kinetic inductance.

Note that both the radiation resistance and ohmic resistance use the effective length ($l_{eff}$) of the antenna instead of the physical length. This is because for most antennas the current in the metal arms is not constant, but actually decreases towards zero at the ends of the wire. Since there is less current at the ends of the wire there is less loss there too. To take this into account, we must find the average current in the arms, or conversely, the 'effective length' of the antenna:

$$l_{eff} = \frac{\int_0^{l/2} I(x)dx}{I_0} \tag{2.6}$$

Where $I_o$ is current induced on the antenna at the feedgap. For a wire with a sinusoidal current distribution, $l_{eff} = \frac{2}{\pi}l \approx 0.64l$.

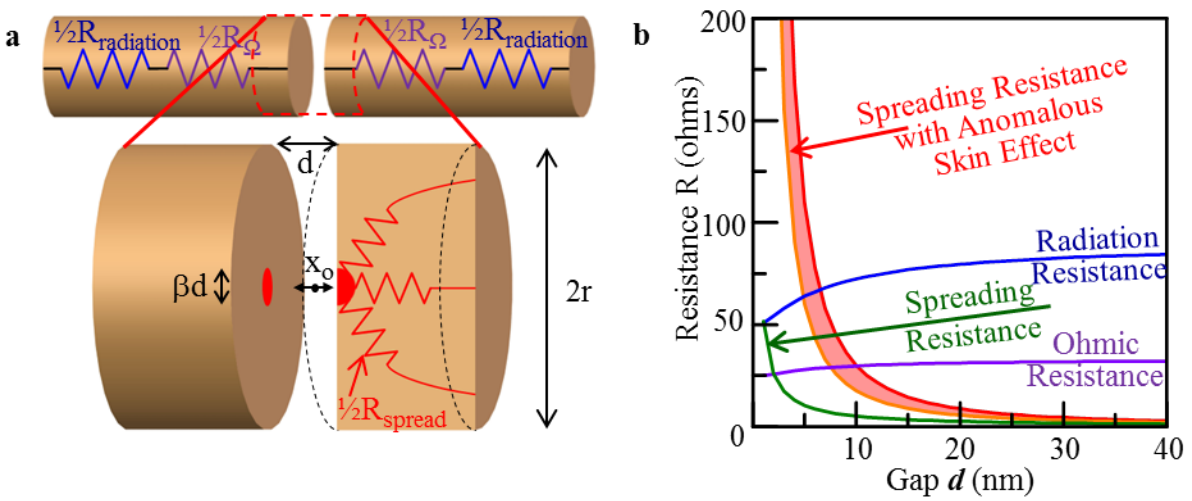


**Figure 2.6** Optical Antenna Resistance. (a) Schematic of a linear antenna (Au wire radius $r$=20nm) with a blunt vacuum gap showing the radiation resistance (blue), Ohmic resistance (purple), and a zoom-in of the gap region highlighting the spreading resistance (red). (b) Plot of the radiation resistance (blue), ohmic (purple), and spreading resistance (green) of a wire antenna as a function of gap spacing. For dipoles centered between flat electrodes the current is spread over a diameter $\beta d$=1.6$d$. The effect of the anomalous skin effect is shown in the shaded red region bounded by $\delta$=1 (orange) and by $\delta$=0.5 (red). The antenna length $l$ is adjusted to maintain 200THz resonant frequency.

As discussed in section 2.1, for a parallel plate capacitor the amount of current coupled into the antenna is: $I_o=q\omega|x_o|/d$. It is therefore advantageous to shrink the gap $d$, to very small dimensions to increase the current coupled into the antenna. Unfortunately, at very small gap spacing, the fields of the driving dipole do not have time to expand out before hitting the metal and are confined to a very small region, $\beta d$, on the antenna tips. These fields give rise to current crowding prior to the current spreading out to the main antenna arm. This very high concentration of current at the feed-gap of the antenna gives rise to an additional loss mechanism, spreading resistance. Although negligible at large gap spacing, spreading resistance becomes the dominant loss mechanisms for small gaps (Figure 2.6). A dipole that is a short distance $d$/2 from a metal surface will produce a corresponding current distribution spread over a diameter ~$\beta d$, resulting in a total spreading resistance $R_{spread}=\rho/\beta d$ for both antenna arms in series. For dipoles centered between flat electrodes we find the dimensionless geometrical parameter $\beta$=1.6.

The concentrated current also experiences a much shorter mean free path than electrons in the bulk of the antenna due to the anomalous skin effect [59], essentially surface collisions. This increases the effective metal resistivity in the concentrated current region, by the factor $(l_e+\delta d)/\delta d$ further exacerbating the spreading resistance, where $l_e$ is the electron mean free path in the bulk metal, and $\delta d$ is the surface collision mean free path, which scales with the

concentrated current region size. Since δ requires a complicated non-local electrodynamic calculation, Figure 2.6 simply plots the range $0.5<\delta<1$.

Figure 2.6 shows how these loss mechanisms change with gap spacing. When the anomalous skin effect is included, spreading resistance eventually dominates all other loss mechanisms as the gap *d* is diminished. This unavoidable loss mechanism will ultimately limit the maximum efficiency of the antenna at high spontaneous emission enhancement.

## 2.4 Circuit Analysis of a Linear Dipole Antenna

Incorporating the resistive components into our circuit diagram yields the form shown in Figure 2.7(a) and simplified in Figure 2.7(b).

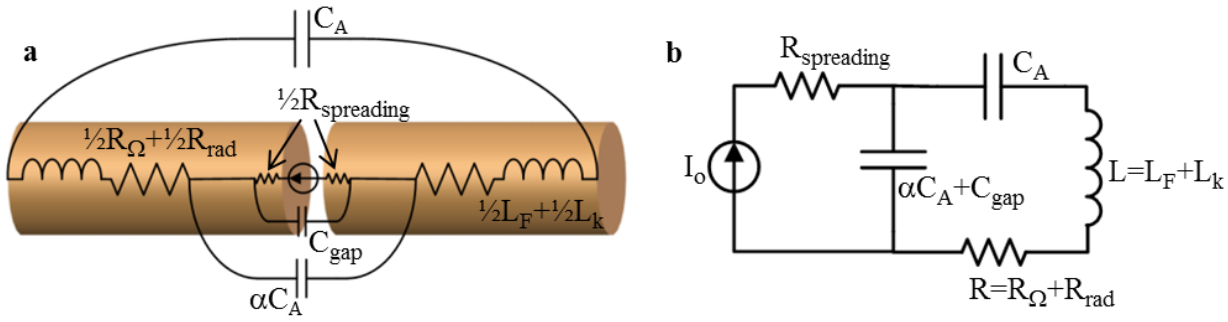


**Figure 2.7** Optical Antenna Circuit Model. (a) Circuit components of an optical antenna, where $R_\Omega$ is the Ohmic resistance, $R_{rad}$ is the radiation resistance, $L_k$ is the kinetic inductance, $L_F$ is the Faraday inductance, $C_{gap}$ is the internal gap capacitance=$\varepsilon_o\pi r^2/d$ for blunt tips of radius r, $C_A$ is the external tip-to-tip capacitance, while $\alpha C_A$ is the stray capacitance between internal tips. We find $\alpha C_A \sim C_A$. (b) The same circuit in a simplified schematic.

In its reduced form, the antenna circuit model is fairly simple to analyze. The total Q of the circuit can be found by turning off the source, at which point the circuit simplifies to a simple series RLC circuit with:

$$Q = \frac{1}{R}\sqrt{\frac{L}{C_A}\left(\frac{1+\alpha+\frac{C_{gap}}{C_A}}{\alpha+\frac{C_{gap}}{C_A}}\right)}, \qquad \omega_0 = \sqrt{\frac{1}{LC_A}\left(\frac{1+\alpha+\frac{C_{gap}}{C_A}}{\alpha+\frac{C_{gap}}{C_A}}\right)} \tag{2.7}$$

Where we have written the total resistance $R=R_{rad}+R_\Omega$. The Q is useful to determine the total radiating current, that is, the total current that flows through the radiation resistance. It can be found that:

$$I_{rad} = \frac{I_o Q}{1 + \alpha + \frac{C_{gap}}{C_A}} \tag{2.8}$$

Where $I_o$ is the current induced on the antenna by the oscillating optical dipole. It is then straightforward to write the total radiated power:

$$P_{rad} = I_{rad}^2 R_{rad} = \left( \frac{I_o Q}{1 + \alpha + \frac{C_{gap}}{C_A}} \right)^2 R_{rad} \tag{2.9}$$

The total antenna efficiency including losses from spreading resistance is then:

$$\begin{aligned} Efficiency &= \frac{P_{rad}}{P_{rad} + P_\Omega + P_{spreading}} = \frac{R_{rad}}{R_{rad} + R_\Omega + \left(\frac{I_o}{I_{rad}}\right)^2 R_{spreading}} \\ &= \frac{R_{rad}}{R + R^2 \frac{C_A}{L} \left(\alpha + \frac{C_{gap}}{C_A}\right) \left(1 + \alpha + \frac{C_{gap}}{C_A}\right) R_{spread}} \end{aligned} \tag{2.10}$$

Which, neglecting the spreading resistance, $R_{spread}$, the efficiency is that of a normal resistive voltage divider.

Now that the radiated power is known, it simply has to be ratioed to the power radiated by a dipole in free space (equation (1.18)) to get the total spontaneous rate enhancement:

$$Enhancement = \frac{P_{rad}}{P_0} = \frac{\left(\frac{q\omega_o x_0}{d}\right)^2 Q^2 \left(\frac{1}{1+\alpha+\frac{C_{gap}}{C_A}}\right)^2 R_{rad}}{\frac{2\pi}{3} Z_o \left(\frac{x_o}{\lambda_0}\right)^2 (q\omega_0)^2}$$

$$= \frac{\left(\frac{\lambda_o}{d}\right)^2 \frac{1}{\omega_0^2 R^2} \left(\frac{1+\alpha+\frac{C_{gap}}{C_A}}{\alpha C_A + C_{gap}}\right)^2 \left(\frac{1}{1+\alpha+\frac{C_{gap}}{C_A}}\right)^2 R_{rad}}{\frac{2\pi}{3} Z_o}$$

$$= \frac{3}{2\pi}\left(\frac{\lambda_0}{d}\right)^2 \frac{\left(\frac{1}{\omega_o(\alpha C_A + C_{gap})}\right)^2 R_{rad}}{Z_0 R^2}$$

$$= \frac{3}{2\pi}\left(\frac{\lambda_o}{d}\right)^2 \frac{Z_{gap}^2 R_{rad}}{Z_o (R_{rad} + R_\Omega)^2} \qquad (2.11)$$

Where we have written, $Z_{gap} \equiv 1/j\omega(C_{gap}+\alpha C_A)$, which is the total impedance across the gap from both the local and distributed capacitance terms.

I have purposefully written equation (2.11) as it is to highlight the main parameters critical to get large rate enhancements. First of all, a small gap is required, as can be seen in the first term representing dipole coupling to the antenna $(\lambda_o/d)^2$. This comes from the fact that radiated power is proportional to current squared, and therefore the $d$ in equation (2.1) is squared (the rest of the constants in equation (2.1) get cancelled out when we normalize to a free space dipole).

The second term in equation (2.11) is effectively an impedance ratio describing the amount of current lost to the shunt pathway ($Z_{gap}$) shown in Figure 2.7(b). To maximize this term you must first balance the ohmic and radiation losses (i.e. make $R_{rad} = R_\Omega$), as is typical in antenna design [60]. Higher enhancement can be obtained by decreasing the total resistance, which effectively increases the antenna Q. However, in most antenna designs the largest obtainable Q will be determined the ohmic loss. In general, $R_{rad}$ and $R_\Omega$ are difficult to tune independently since they both, as well as the antenna resonance, depend on the antenna length. The only free variable is the antenna diameter, which can be made larger or smaller to tune to ohmic losses. The diameter, however, can have a pronounced effect on the gap impedance, $Z_{gap}$.

Maximizing the second term in equation (2.11) also requires maintaining a large $Z_{gap}$. Since there is little which can be engineered about the inner distributed capacitance, this typically means making sure $C_{gap} < \alpha C_A$. If $d$ is made so small that $C_{gap} > \alpha C_A$, then the enhancement no longer

increases as $1/d^{2}$. As an example we can take the case of a parallel plate type coupling as shown in Figure 2.7(a). If the gap capacitance dominates we can rewrite equation (2.11):

$$\begin{aligned} Enhancement &= \frac{3}{2\pi}\left(\frac{\lambda_o}{d}\right)^2 \frac{\left(\frac{1}{\omega_0 C_{gap}}\right)^2 R_{rad}}{Z_o(R_{rad}+R_{ohmic})^2} \\ &= \frac{3}{2\pi}\left(\frac{\lambda_o}{d}\right)^2 \left(\frac{d}{\epsilon_g A}\right)^2 \left(\frac{\lambda_o}{2\pi c}\right)^2 \frac{R_{rad}}{Z_o(R_{rad}+R_{ohmic})^2} \\ &= \frac{3\lambda_o^4}{8\pi^3 c^2}\left(\frac{1}{\epsilon_g A}\right)^2 \frac{R_{rad}}{Z_o(R_{rad}+R_{ohmic})^2} \end{aligned} \quad (2.12)$$

We see that when $C_{gap} > \alpha C_{A,}$ the enhancement is no longer a function of gap spacing. To achieve the maximum enhancement possible we must engineer both a small feedgap and a large gap impedance.

## 2.5 Comparison of Different Antenna Geometries

With equations (2.10) and (2.11) for efficiency and enhancement, we can now easily compare different antenna geometries without doing lengthy FDTD simulations. Figure 2.8 shows the gap impedance and rate enhancement of four different antenna geometries: the simple dipole antenna with parallel-plate gap (black), the simple dipole antenna with a gap filled with dielectric material typical of a semiconductor emitter, index $n$=3.4 (red), a dipole antenna with high index gap and rounded tips (blue), and the arch-antenna with high index gap (green). The solid curves are from the circuit model in Figure 2.7, while the dotted lines connecting square dots represent a 3D Finite Difference Time Domain simulation of the same structure (Lumerical and CST Microwave Studio). Also shown in Figure 2.8(b) is a dotted black line representing the ideal rate enhancement trend of $(\lambda_o/d)^2$.

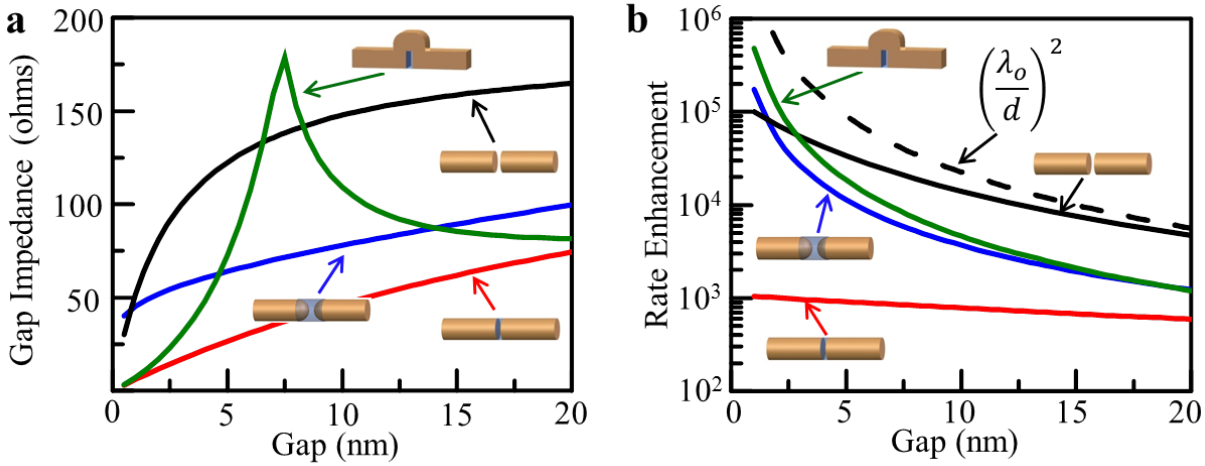


**Figure 2.8** Comparison of four different dipole antenna geometries. (a) Gap impedance ($Z_{gap}$) and (b) rate enhancement as a function of gap spacing for a simple dipole (black), dipole with high index gap (red), dipole with high index gap and rounded tips (blue), and arch-dipole with high index gap (green).

First let's look at the basic case of a dipole antenna with a vacuum filled parallel-plate gap. At large gap spacing $Z_{gap}$ is large, limited by the inner distributed capacitance. However, as the gap shrinks down the impedance starts to drop and goes towards zero as the gap approaches zero. Comparing this to its rate enhancement in Figure 2.8(b), we see that when the impedance decreases at small gaps, the enhancement deviates from the ideal $(\lambda_o/d)^2$ behavior. A real antenna, however, needs to be fed by an optical emitter. Most optical emitters, such as semiconductors, are high index materials. When the gap is filled with such a high index material, $C_{gap}$ significantly increases and becomes the dominant gap impedance term. $Z_{gap}$ plummets as shown in the red curves and the corresponding enhancement is nearly constant with gap spacing. The large $C_{gap}$ capacitance of a rectangular gap at narrow spacing effectively shorts out the current being driven into the antenna. The shorted current never sees the radiation resistance if $|1/\omega C_{gap}| < R_{rad}$.

We can remedy the drastic shunting of current across the gap by reducing gap capacitance; for example by adopting hemispherical gap tips. If the tip radius of curvature is the same as the radius of the wire, the gap capacitance on close approach [56] is that of two spheres : $C_g = \frac{\pi\epsilon_0 r}{2}\left[ln\left(\frac{r}{d}\right) + 2\gamma\right]$. This is only weakly logarithmically dependent on the $r/d$ ratio. In Figure 2.8, the rounded inner tips (blue) partially compensate for the higher capacitance semiconductor filling. $Z_{gap}$ stays moderately large, even at small gaps, and the enhancement goes back to following the $(\lambda_o/d)^2$ trend.

The rate enhancement, however, is only half of what we're interested in. Now that we have a structure with good a good gap scaling trend we can now look at what sort of antenna efficiency

we can expect. To show the importance of both enhancement and efficiency, the two are plotted side-by-side for the case of hemispherical tips at the gap in Figure 2.9. The blue line plotted on the left-side axis is the spontaneous emission rate enhancement reproduced from Figure 2.8(b). I have also added data points taken from a full 3D maxwell solver shown as blue squares connected with a dotted line. We can see that the circuit model lines up almost exactly with the Maxwell solver.

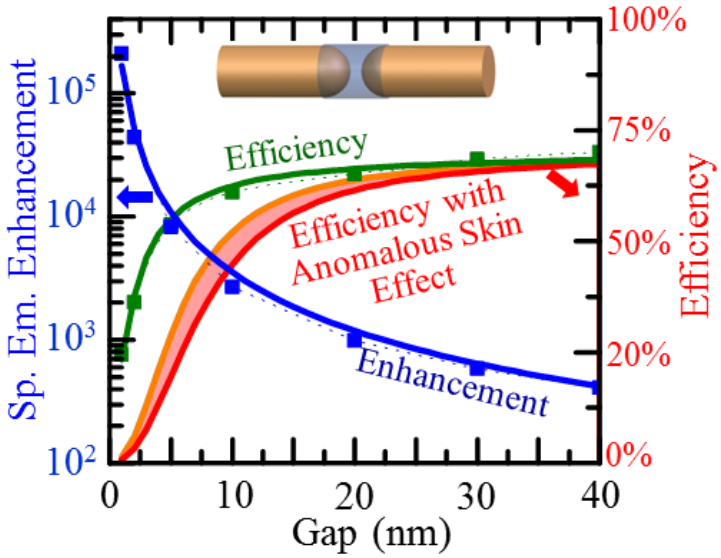


**Figure 2.9** Enhancement (blue) and Efficiency (green) of an Au-wire antenna (radius *r*=20nm) with rounded inner tips and a semiconductor filled gap, (light blue, refractive index=3.4). The harmful effect of anomalous skin effect is plotted as the shaded red region bounded by $\delta$=1 (orange) and by $\delta$=0.5 (red). Solid lines are derived from the circuit model. Solid boxes connected by dotted lines are obtained from 3D FDTD simulations (CST, Lumerical). The antenna length *l* is adjusted to maintain 200THz resonant frequency. For rounded electrodes, the current spreading factor $\beta$~0.5 is a weak function of *d/r* in this range.

Next, we can look at the efficiency of this antenna plotted on the right-side axis. Neglecting the anomalous skin effect, the efficiency is plotted as a solid green line for the circuit model results and green squares for the full 3D Maxwell solver result. Note that once again, there is very good agreement between simulation and the circuit model. It should be noted, however, that one parameter was used to fit this data instead of obtaining an analytical solution. That factor is $\beta$ from Table I in the spreading resistance formula, accounting for the size of the electric field spot on the antenna tips. $\beta \approx 1$, and in this case a $\beta = 0.5$ is used to get the best fit. It is intuitive that $\beta$ would be smaller for rounded tips since the electric field will be more concentrated. As the gap shrinks down and the dipole gets closer to the metal it will look increasingly flat, and therefore $\beta$ will increase. In the regime plotted we do not vary $\beta$.

It is very apparent in Figure 2.9 that the efficiency begins to decrease significantly at small gap spacing, consistent with the trend for spreading resistance from Figure 2.6(b). With spreading resistance taken into account, the shaded region bounded by the red and yellow curves, we can instantly see that the efficiency drops below 50% at a 10nm gap size. While our actual design criteria may vary, we set 50% as the lowest efficiency tolerable for our efficient nanoemitter. This means that the highest enhancement we can expect is that from a 10nm gap dipole antenna, which from Figure 2.9 is ~2,500x. This means that potentially our LED originally operating at 200MHz could now operate at 500GHz!!

In practice, rounding the inner tips of a dipole antenna can prove quite challenging. An alternative way to manage the gap impedance is by introducing an inductor across the gap to "resonate out" the gap capacitance. This structure, called the arch-dipole antenna, or arch-antenna for short, is discussed in detail in Chapter 3. Figure 2.8(b) shows that this structure can provide similar if not better enhancement than the rounded tip dipole.

# Chapter 3

# The Free-Standing Arch-Dipole Antenna

As discussed in the previous section, one of the most critical elements of designing a good optical antenna is mitigation of large gap capacitance. While rounding or tapering the antenna towards the feedgap achieves this nicely on paper, both tend to be very difficult to fabricate. Fabricating a simple dipole antenna with a gap spacing <20nm can prove challenging, even with ebeam lithography. This is complicated even further with the need to align this tiny gap to a similar sized optical emitter. Once the need for a complicated gap structure is factored in, this becomes a very difficult problem to solve.

For this reason we developed the arch-dipole antenna structure. The arch-dipole offers a feedgap that is self-aligned to the optical emitter and also an inductive arch which can effectively mitigate gap capacitance. In this chapter I will present a modified circuit model to describe the arch-dipole, compare this model to simulated results, discuss effective strategies for arch-dipole design, and present measurement results from actual fabricated devices.

## 3.1 Modification of the Simple Dipole Circuit Model

The arch-dipole antenna, pictured in Figure 3.1, is a slightly modified simple dipole antenna. As usual, the emitter is placed in the feedgap, however, the two antenna arms are connected with a short metal arch over the gap. This metal arch carries a small amount of inductance and resistance. Since the arch goes over the gap, the added inductance and resistance terms are in parallel to the gap capacitance. These two new terms can be added to the circuit model in Figure 2.7(b) to create the modified circuit describing the arch-dipole, shown in Figure 3.2.

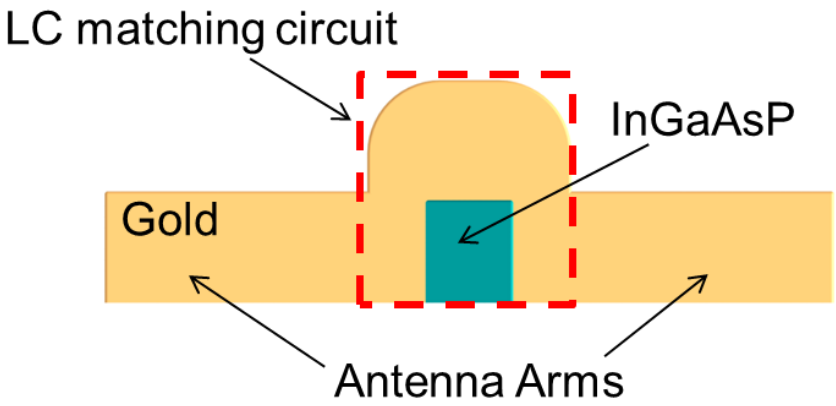


**Figure 3.1** Simple schematic drawing of an arch-dipole antenna.

If poorly designed, this small inductance and resistance can easily create a new shunt path, further reducing the current in the radiation resistor. However, if properly tuned to the same impedance as the gap capacitance, the arch inductance and gap capacitance will look like an effective open circuit. This completely removes both terms from the circuit, leaving only the inner distributed capacitance as a shunt path for current. It should be noted that the inner distributed capacitance cannot be resonated out since it is not a localized element.

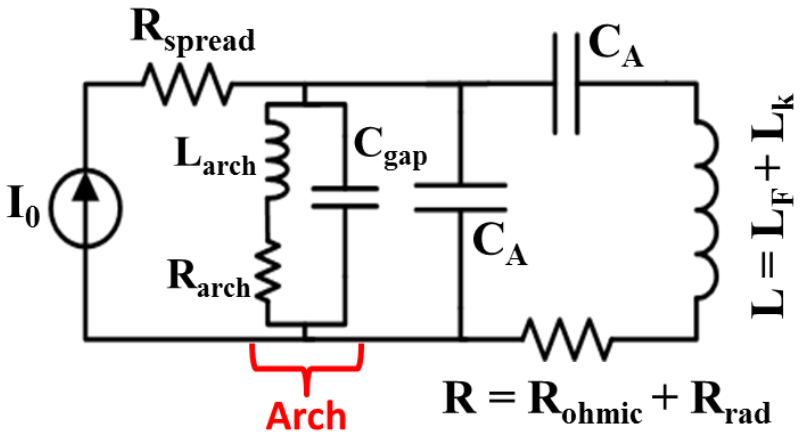


**Figure 3.2** Circuit diagram of arch-dipole antenna.

Since adding in the arch adds another LC network to the circuit model, it is not surprising that an additional mode is also added. The two new modes, depicted in Figure 3.3, split the ordinary dipole mode (blue) into two peaks; one at higher frequency and one at lower frequency. The lower frequency mode has current going through the arch ($L_{arch}$) in phase with the current going through the antenna arms (L, R) and I will hereafter denote this as the *symmetric* mode. Likewise, the higher frequency mode has the current in the arch 180° out of phase with the current in the antenna arms, which I will call the *anti-symmetric* mode.

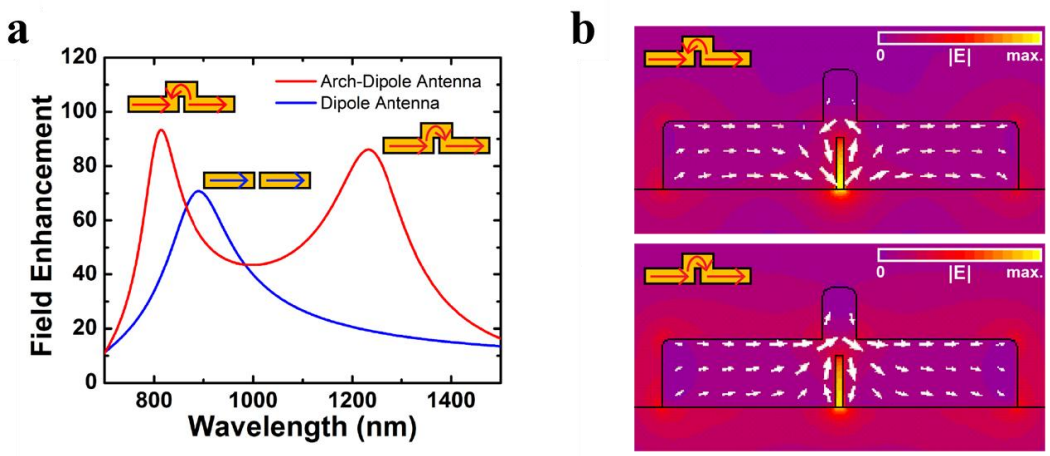


**Figure 3.3** Modes of an arch-dipole antenna. (a) Field enhancement in the feedgap as a function of wavelength for the arch-dipole (red) and simple dipole (blue) antenna. (b) The current and electric field of the two arch-dipole modes. Reproduced from [61].

In general, I have usually found the anti-symmetric mode allows for longer antenna arms to be used than with the symmetric mode, which often leads to higher enhancement and better efficiency. However, the anti-symmetric mode is also much more dependent on the exact geometry of the metal arch, while the symmetric mode is mainly determined by the overall antenna length. Since antenna length is usually much easier to control than the height, width, and metal coverage of the arch, the lower frequency mode is much easier to fabricate controllably.

The exact analytical form for the arch inductance can be somewhat difficult to derive. We can approximate it by using the simplified form for the faraday inductance of a round solenoid $L_{f\_arch} = \frac{\mu_0 A_{gap}}{w}$. Where $A_{gap}$ is the cross-sectional area of the gap and $w$ is the width of the antenna. The resistance and kinetic inductance can be approximated by taking the length of the current path through the arch and the cross-sectional area of the current flow:

$$R_{arch} + j\omega L_{k\_arch} = \rho \frac{l}{A} = \rho \frac{d + h_{arch}}{w * \delta_{skin}} \tag{3.1}$$

Where $h_{arch}$ is the height of the arch and $\delta_{skin}$ is the skin depth.

### 3.1.1 Comparing Circuit Model to Simulation

In Figure 3.4(a) I have plotted the rate enhancement versus frequency predicted both by simulation (red curve) and with the circuit model (blue curve) for a simple dipole antenna with a 10nm wide blunt-gap filled with high-dielectric material. To compensate for the extra capacitance from the ends of the antenna neglected in the circuit model as mentioned in section 2.2, the simulated antenna was made to be shorter than the length given for the circuit model (300nm long versus 420nm).

The circuit model does a pretty good job at predicting the shape and amplitude of the antenna resonance. The peak enhancement is 295x for the circuit model and 422x for the full-wave simulation, giving an error of ~30%. In this regime (high-index, narrow gap) the exact value of the gap capacitance plays a critical role in determining the peak enhancement. Such close agreement between simulation and the circuit model validates that our approximation of using parallel plate capacitance for the gap impedance is valid.

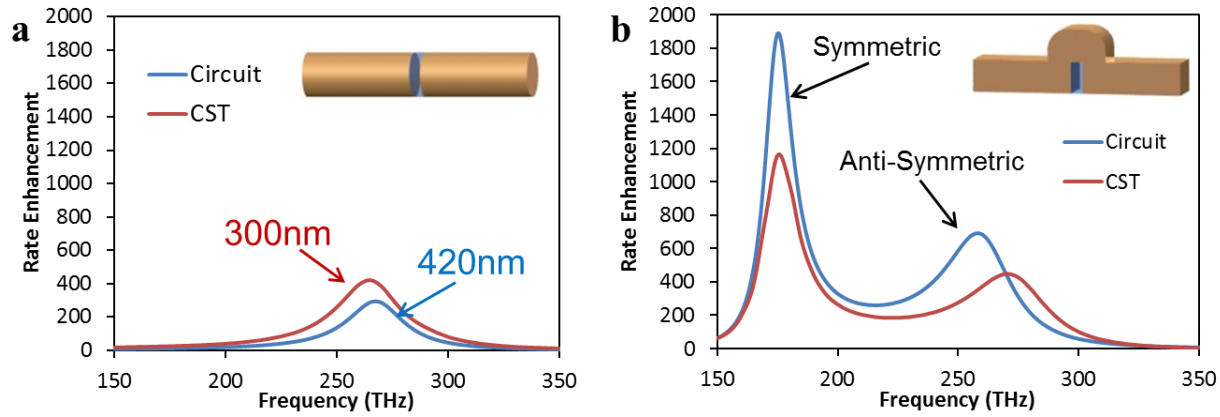


**Figure 3.4** Comparison of rate enhancement calculated with the circuit model (blue) for a 420nm long antenna versus simulated with CST (red) for a 300nm long antenna. (a) A simple dipole antenna with 10nm wide blunt gap filled with high-dielectric material. (b) An arch-dipole antenna with a 10nm wide, 30nm tall gap filled with high-dielectric material.

The dipole antenna can be converted to the arch-dipole by adding the metal arch over the active material. For simplicity, the structure shown in Figure 3.4(b) has a rectangular cross-section of 40nm by 40nm versus a round cross-section of r = 20nm for Figure 3.4(a). The single resonance splits into the symmetric and anti-symmetric modes as expected, with the symmetric mode giving the larger enhancement ~4x larger than that of the simple dipole antenna. Although the general shape of the arch-antenna resonance is predicted correctly with the circuit model, the peak value for enhancement is overestimated by 50-60% for both modes. This error is due to the difficulty in estimated the precise value of the inductance and resistance of the arch. In general the circuit model will give the approximate resonance and enhancement to expect for a given geometry, at which point simulations around this point are needed to optimize the structure. A better analytical solution for the arch impedance would allow for more accurate predictions.

### 3.1.2 Guidelines for Proper Arch-Dipole Design

Design of a dipole antenna is fairly straight-forward. The gap spacing is usually determined by the smallest feature size that can be reliably fabricated. This leaves the cross-sectional area as the only free parameter, since the length must be chosen to achieve the right resonance frequency. The cross-sectional area tends to have only a minor effect on enhancement. If it is too small there

is too much ohmic resistance and the antenna becomes inefficient. If it is too wide then the gap shorts out and enhancement is lost. Anything in-between usually works about the same.

Things become a little more complicated with the arch-dipole antenna. The arch not only splits the resonance into two modes, but also adds an additional degree of freedom: arch height. This allows the resonance to be tuned by altering either the length of the antenna or the arch dimensions. This opens up a huge design space where a large number of combinations of arch height and antenna length all give the same resonance, but many different values for enhancement.

As a general design rule I have found that the largest enhancement and highest efficiencies are achieved when designing the structure like an ideal dipole-antenna. Since the arch is only there to mitigate gap capacitance, then it should be designed to have the same magnitude impedance as the gap capacitance. This gives the condition if we assume that the faraday inductance is dominant:

$$\begin{aligned} \left|\omega L_{f\_arch}\right| &= \left|\frac{1}{\omega C_{gap}}\right| \\ \omega \frac{\mu_0 h_{arch} d}{w} &= \frac{d}{\omega \epsilon_g w * h_{arch}} \\ \omega \frac{\mu_0 h_{arch}}{1} &= \frac{1}{\omega \epsilon_g * h_{arch}} \\ h_{arch}^2 = \frac{1}{\omega^2 \mu_0 \epsilon_0 n_g^2} &= \left(\frac{\lambda}{2\pi n_g}\right)^2 \\ h_{arch} &= \frac{\lambda}{2\pi n_g} \end{aligned} \tag{3.2}$$

So for an emitting wavelength of 1500nm and a refractive index of 3.4, the arch should be ~70nm. This of course neglects the kinetic inductance which is not negligible in the optical regime. For a gold antenna at 1500nm the two inductances are about equal, which leads to a desired arch height of ~35nm.

## 3.2 Measuring Spontaneous Emission Enhancement

For almost all the optical measurements presented in this dissertation, fabricated structures were characterized by measuring the amount of light they emitted. Increased spontaneous emission intensity, in competition with non-radiative losses, directly measures spontaneous emission enhancement. This is actually a more reliable measure of spontaneous emission rate than lifetime measurements, since excited state lifetime can be influenced by emitters coupling to surface

plasmons or optical cavities, which may not radiate, and with the energy still being lost as heat [30,38].

The amount of light collected from an emitter can be described as:

$$PL = P * \eta_{internal} * \eta_{extraction} * \eta_{collection} \tag{3.3}$$

Where P is the amount of electron-hole pairs generated in the emitter, $\eta_{internal}$ is the internal quantum efficiency (fraction of electron-hole pairs that create a photon), $\eta_{extraction}$ is the extraction efficiency (fraction of photons that are radiated into free-space), and $\eta_{collection}$ is the photon collection efficiency (fraction of photons emitted into free-space actually collected with the microscope objective). In our experiments the dominant non-radiative mechanism is surface recombination. The internal quantum efficiency of the emitter is then equal to the ratio of the radiative recombination rate to the surface recombination rate:

$$\eta_{internal} = \frac{1/\tau_{rad}}{1/\tau_{rad} + 1/\tau_{nrad}} \approx \frac{1/\tau_{rad}}{1/\tau_{nrad}} = \frac{\tau_{nrad}}{\tau_{rad}} \tag{3.4}$$

If the pumping, surface recombination, and far-field emission patterns are kept constant for both bare and antenna coupled ridges, then an increase in optical emission is therefore a direct measurement of the increased rate of spontaneous emission into the far-field.

$$\frac{PL_{antenna}}{PL_{bare}} = \frac{P_{antenna} * \tau_{nrad}/\tau_{rad\_anenna} * \eta_{collection}}{P_{bare} * \tau_{nrad}/\tau_{rad\_bare} * \eta_{collection}} = \frac{1/\tau_{rad\_anenna}}{1/\tau_{rad\_bare}} \tag{3.5}$$

While the non-radiative losses are undesirable in a device, they are a type of clock for fundamental measurements.

To ensure that non-radiative rates dominate, the InGaAsP surface was not treated with any form of passivation layer. Bare InGaAsP layers have been rigorously studied before and have been shown to have a surface recombination velocity, $v_s = 3x10^4$ cm/s [62,63]. In addition, the ridges are covered in a 3nm $TiO_2$ layer that experimental measurements have shown increases the surface recombination rate by an additional factor three. Considering the ~35nm width of the InGaAsP ridges, this corresponds to a surface recombination lifetime that can be calculated as:

$$\frac{1}{\tau_{nrad}} \approx \frac{A_s}{V} v_s = \frac{(4 * 150 + 2 * 35) * 35nm^2}{35 * 35 * 150nm^3} 9x10^4 cm/s = \frac{1}{9ps} \tag{3.6}$$

As expected, the non-radiative lifetime of <9ps is significantly shorter than InGaAsP's un-enhanced spontaneous emission lifetime of >10ns.

The $TiO_2$ layer is used to prevent direct contact between the InGaAsP and the gold antennas. This ensures that the only change in non-radiative recombination will be from ohmic losses in the antenna and not from changes in surface recombination. Although chemical surface passivation or cladding layers would greatly increase the light emission from this structure, they would impair the proper measurement of antenna properties which is our main goal for this experiment.

It is also necessary to eliminate mechanisms that enhance external quantum efficiency without increasing the spontaneous emission rate, most notably improved light extraction. For a traditional LED, the majority of light gets trapped in the semiconductor and only a very small emission angle can escape (Figure 3.5(a)).

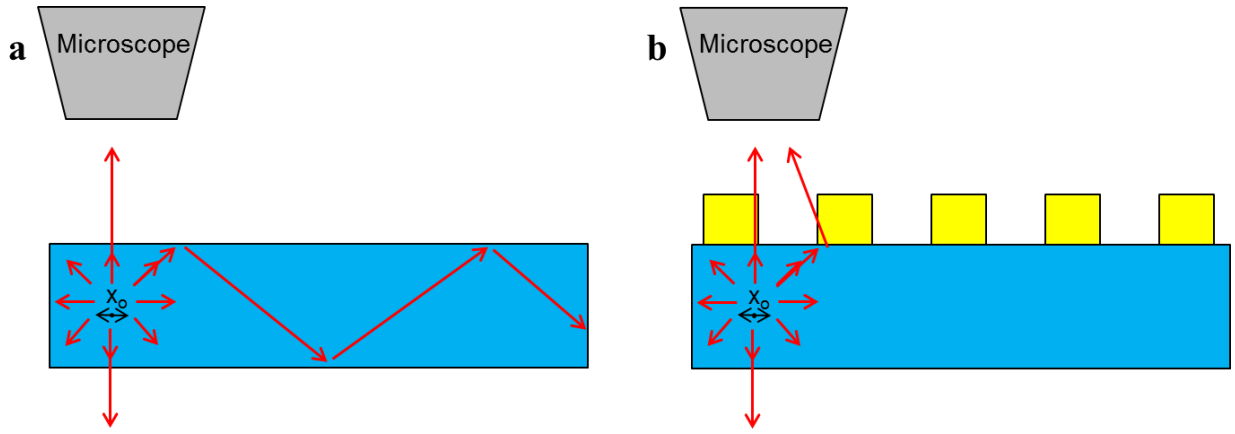


**Figure 3.5** (a) Schematic of an optical dipole emitting light in a high index substrate. Only a small percentage of light escapes the substrate and is collected by the microscope. (b) Metal structures on the surface of the high index substrate help scatter out trapped light, increasing the amount of light observed. Note that in both cases the actual radiative lifetime is the same.

By using a textured surface or grating, the light extraction can be increased by up to $4n^2$, where $n$ is the refractive index of the semiconductor [64]. Arrays of optical antennas can provide scattering extraction enhancement, thereby increasing light output, but without increasing the emission rate as demonstrated in Figure 3.5(b). By completely removing the substrate and only leaving very small, sub-wavelength semiconductor islands, the possibility of scattering light extraction enhancement is removed.

## 3.3 Free-Standing Arch-Dipole

A schematic of the arch-dipole antenna devices actually fabricated and tested is shown in Figure 3.6 (the fabrication process is described in detail in section 3.3.1). The device consists of an InGaAsP ridge (lattice matched to InP, peak PL wavelength of 1500nm) 150nm long, 34nm wide, and 35nm tall conformally coated in a 3nm thick layer of $TiO_2$. A gold bar 50nm wide and

45nm thick is deposited perpendicularly over the ridge forming the arch-antenna. The length of the antenna is varied to tune the resonance frequency.

The entire structure is bonded to a quartz carrier wafer with UV curable epoxy (Norland Optical Adhesive 81, n = 1.56) and then the high-index InP substrate is completely removed. This leaves the final devices, both with and without an antenna, embedded in the very top of a layer of epoxy. Removal of the high-index substrate allows the device to be measured without complications from light trapping effects.

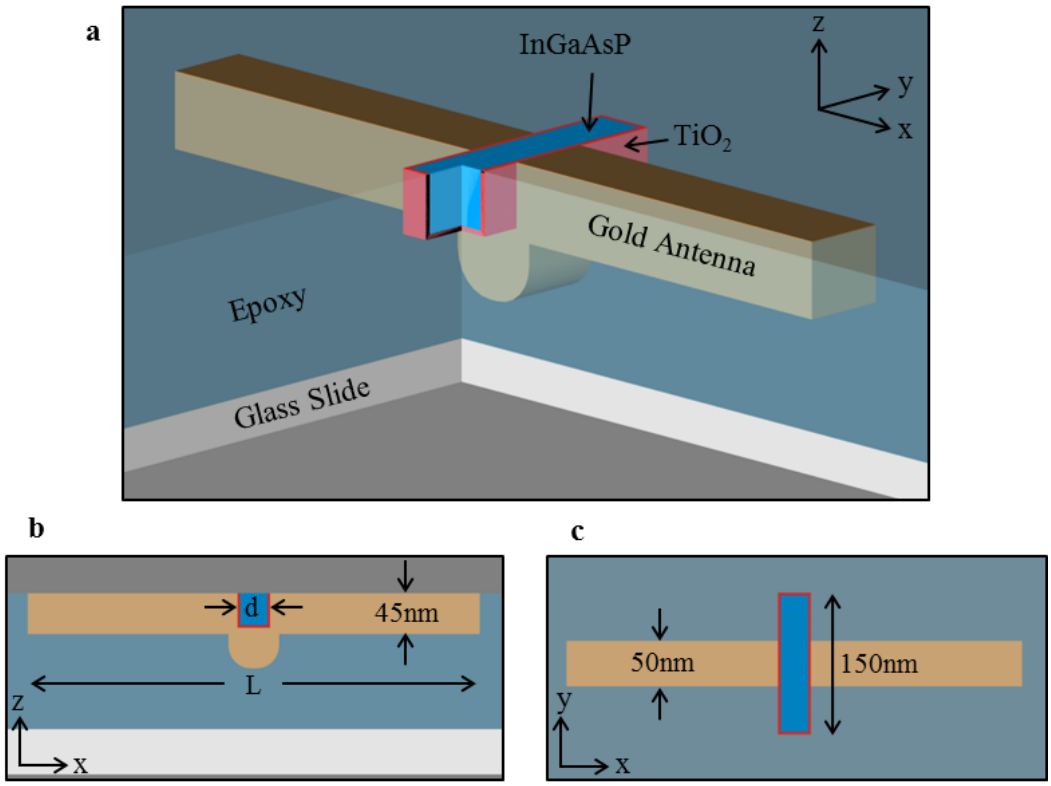


**Figure 3.6** nanoLED Structure. a, Cutaway perspective view of nanoLED structure bonded to a glass slide with epoxy. b, Side-view of antenna structure and c, top-view of antenna structure. Thickness of epoxy and glass slide are not shown to scale.

There are several important insights about the device can be seen from the optical mode of the arch-dipole structure shown in Figure 3.7. A cross-section of the antenna along its long axis (Figure 3.7(a)) shows a large electric field inside the thin $TiO_2$ coating the InGaAsP ridge. This gap configuration is effectively two capacitors in series; therefore the total gap field is divided across the $TiO_2$ and InGaAsP according to their relative thicknesses and dielectric constants. This reduces the electric-field the optical dipoles in the InGaAsP see, reducing their capacitive coupling with the antenna. We can modify equation (2.1) to take this effect into account by replacing the gap spacing, *d*, with a new effective gap width similar to the way effective oxide thickness is calculated for a MOSFET:

$$\frac{I_{0\_wTiO_2}}{I_0} = \frac{d_{InGaAsP}}{d_{InGaAsP} + \frac{\epsilon_g}{\epsilon_{TiO_2}} d_{TiO_2}} \tag{3.7}$$

For a 34nm ridge (n = 3.4) with a total of 6nm of $TiO_2$ (n = 2.4), this corresponds to a coupled current decreased to 74%. Since enhancement is proportional to current squared, the enhancement of this antenna will be reduced to ~1/2 of what would be expected from a 40nm wide gap.

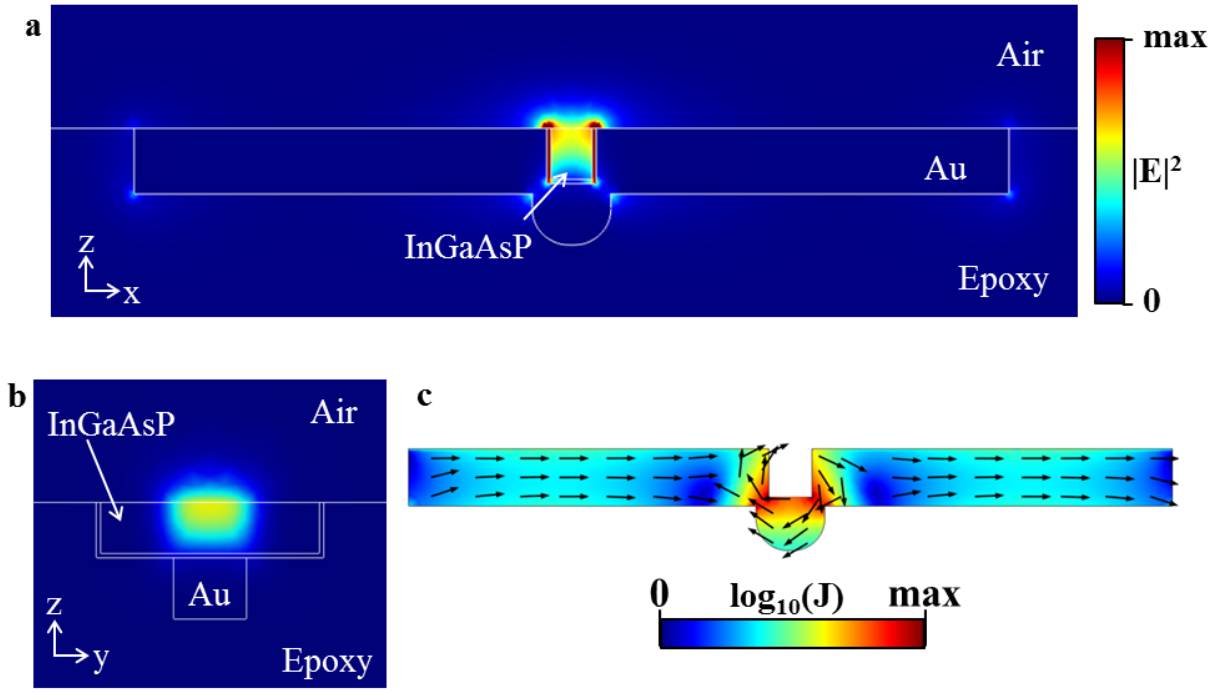


**Figure 3.7** Simulation of antenna-coupled InGaAsP Ridge. a, b, zx-plane and zy-plane cross-sectional view of $|E|^2$ profile for antennas anti-symmetric mode. c, Current density of the anti-symmetric mode. Arrows indicate vector direction of flow and color intensity indicates current density.

A cross-section of the device along the long-axis of the ridge is shown in Figure 3.7(b). It is apparent that the optical mode does not extend outside of the area directly below the gold arch, meaning the parts of the ridge sticking outside of the arch won't optically couple to the antenna. Only ~1/3 of the ridge is covered by the arch; carriers in the remaining 2/3 will have to diffuse under the arch to see any spontaneous emission rate enhancement. Unfortunately, the diffusion length is insufficient compared to the ridge length. Since the recombination lifetime is limited to ~10ps by surface recombination, the diffusion length for holes in InGaAsP can be calculated as ~16nm using one-tenth the bulk hole mobility of $10cm^2/Vs$ [65]. Consequently, a large portion of the carriers will never see the antenna arch before recombining. Given that the exposed InGaAsP arms are ~50nm long, compared to a diffusion length of ~16nm, only ~32% of the

carriers will diffuse to the antenna hotspot. Ideally the ridge would be the same length as the antenna width, however in these experiments we are limited by alignment tolerances.

In addition, the entire height of the ridge is not uniformly enhanced. The mode is concentrated towards the bottom of the ridge, with poorer dipole coupling towards the top. This issue could be solved by using a spacer (such as a high index cladding layer) above the active region to define the part of the ridge with poor coupling. Doing so could slightly increase our measured enhancement, but would significantly complicate the fabrication process.

### 3.3.1 Fabrication Process

Antennas were fabricated on a metal-organic chemical vapor deposition (MOCVD) grown (100) InP epitaxial wafer using top-down semiconductor processing techniques. Processing can be broken down into four main steps: (1) alignment mark deposition; (2) InGaAsP ridge formation; (3) Antenna deposition; (4) Substrate removal. All lithography was done using e-beam lithography with thin (~80nm) positive resist.

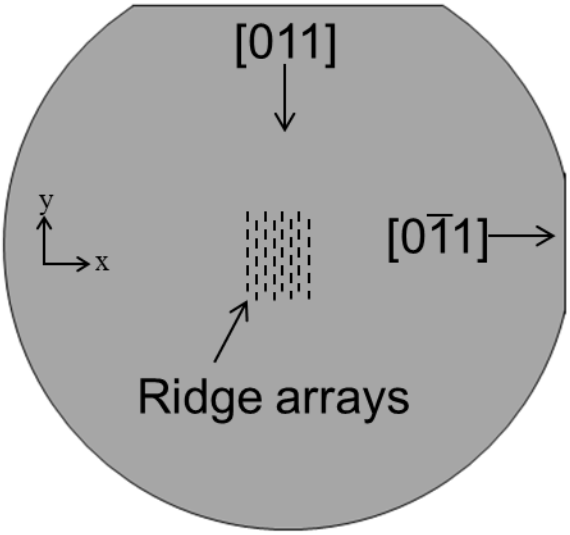


**Figure 3.8** Schematic of a (100) InP Wafer. Crystal directions of the (100) InP epiwafer used for fabrication. Dashes in the center demonstrate the orientation of InGaAsP ridges etched into the wafer (not to scale).

First, alignment marks consisting of a 5nm titanium adhesion layer and a 25nm gold layer were formed using a liftoff process on top of a 35nm thick InGaAsP layer. Gold was used because it is resistant to the chemical etching used in later processing steps. Next, thin titanium islands (7nm thick, 200nm long, 120nm wide) were patterned into 20um x 20um arrays with a 700nm staggered pitch and deposited with e-beam evaporation and lift-off. The titanium islands were aligned along crystal planes so that the long axis was parallel to the [011] direction (Figure 3.8); these were used as a hard mask to etch ridges into the InGaAsP active layer. A dilute piranha solution (1:8:100 $H_2SO_4$:$H_20_2$:$H_20$) was used to slowly wet etch the InGaAsP layer at a rate of

~5nm/s in the [100] direction, ~2nm/s in the [011] direction, and ~3.5nm/s in the [0 $\overline{1}$ 1] direction. Due to the anisotropy of the etchant, the sidewalls parallel to the [011] are fairly vertical for moderate aspect ratios (~1:1). The etching is done in small steps (a few seconds of etching at a time) and then checked under SEM. Because the hardmask layer is so thin, it is semi-transparent to the electron beam and the width of the underlying ridge can be determined. Etching continues until the ridge is etched to a width of ~30nm and then the titanium hardmask is removed with a quick dip in 49% hydrofluoric acid.

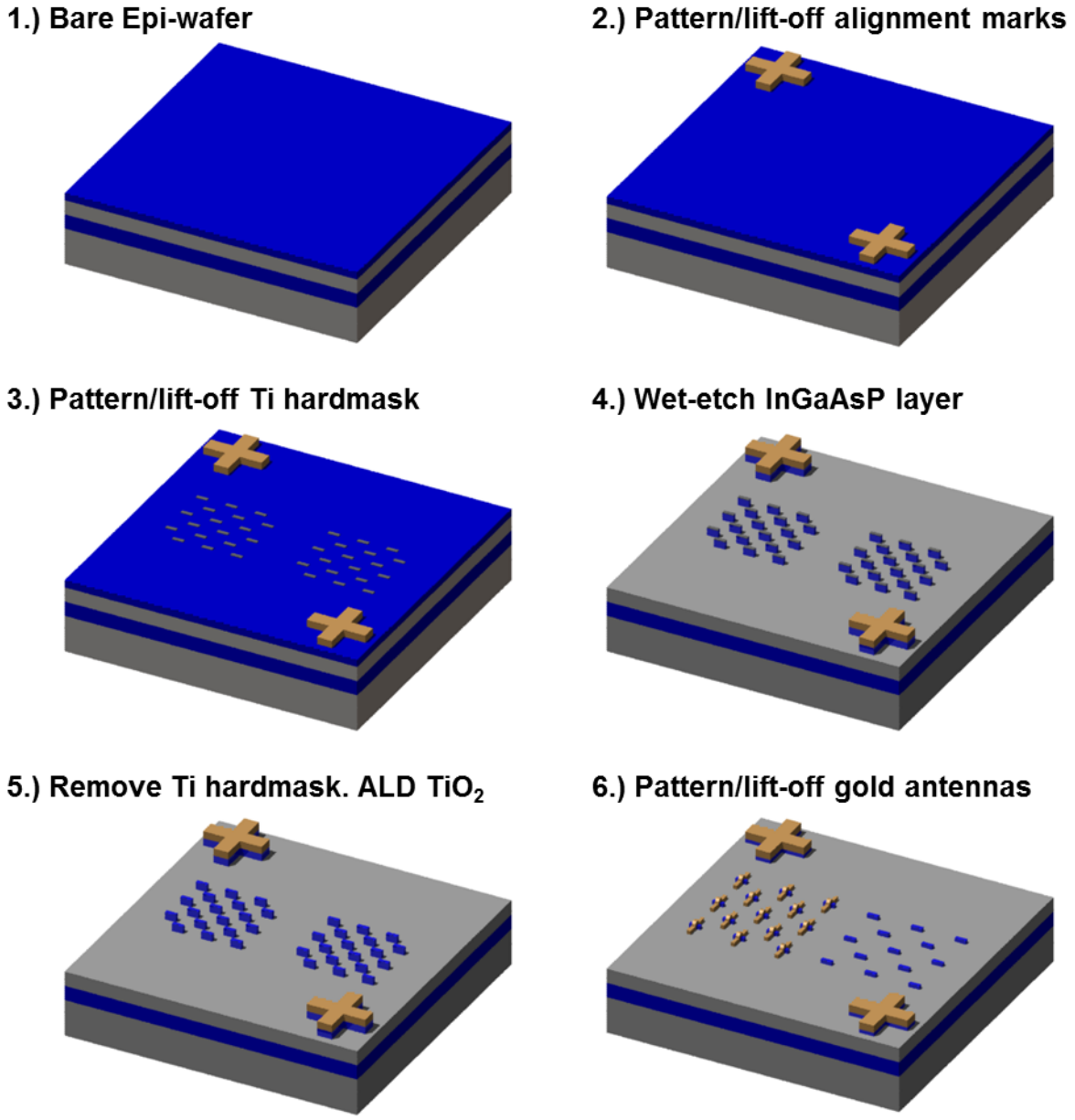


**Figure 3.9** Arch antenna fabrication process flow. Process flow used to etch the InGaAsP ridges and deposit antennas over them.

Immediately after removal of the titanium hardmask, the sample is loaded into an atomic layer deposition (ALD) machine and pumped down to medium vacuum (~1Torr). A conformal 3nm layer of $TiO_2$ is then deposited on the sample at 150$^{o}$C using ALD. This thin oxide layer provides a uniform surface for all the ridges and prevents direct contact with the metal antenna deposited in the subsequent step.

Next, 50nm wide antennas with lengths varying from 400nm to 800nm are patterned perpendicularly over the InGaAsP ridges. 2nm of germanium is then evaporated at a rate of 0.8 Å/s followed by 40nm of gold deposited at a rate of 1.5 Å/s. The slow deposition rates yield polycrystalline metal but offer superior conformal coverage compared to quickly evaporated metal. The germanium layer acts as a wetting layer for the gold to ensure connection of the metal arch over the ridge. The metal is then lifted off leaving the antenna coupled InGaAsP ridges. SEM images are taken at this point to validate successful fabrication of the arch antenna structure.

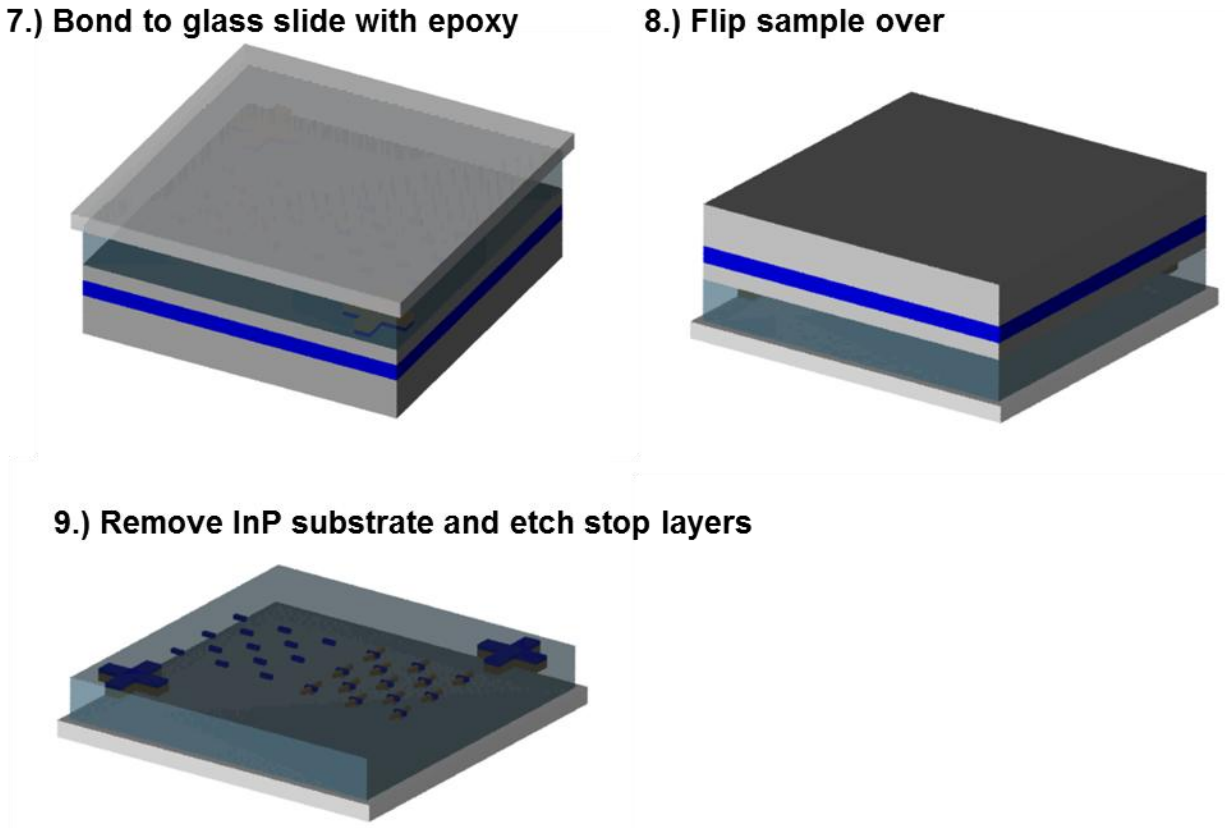


**Figure 3.10** Flip-chip bonding process flow. Following completion of the arch antenna structures, the chip is bonded to glass with epoxy and the substrate removed.

After completion of the arch-antenna structure, the entire chip is flipped over and bonded to a quartz carrier with UV curable epoxy (NOA 81). After UV curing, the epoxy is left to hardbake

for 12 hours at 50°C. The InP substrate is then mechanically lapped with fine-grit sandpaper to a thickness of ~50um. The remaining 50um is then etched off with a 1:1 solution of Hydrochloric Acid:Phosphoric Acid at 80°C and stops on a 150nm thick InGaAs etch stop layer. The InGaAs layer is removed with a 1:1:10 solution of $H_2SO_4$:$H_20_2$:$H_20$ leaving only the InP waveguides embedded in the epoxy.

### 3.3.2 Optical Measurement Results

20um by 20um arrays of InGaAsP ridges were fabricated with a 700nm staggered pitch as shown in Figure 3.11(b). To excite carriers in the semiconductor ridges, a Ti:Sapphire laser with 720nm center wavelength, 120fs pulse width, 13.3MHz repetition rate, and 20uW average power was focused onto the sample using a μ-PL microscope setup drawn schematically in Figure 3.11(a). The spot size of the pump laser was ~2um and therefore roughly 4 devices were probed at a time. The pump was polarized in the y-direction to prevent resonant pumping. Emitted light was collected with the same objective then filtered with a polarizer to discriminate between light emitted polarized perpendicular (y-direction) and parallel (x-direction) to the antenna. Any reflected laser light was completely filtered out using a dichroic mirror and long-pass optical filter.

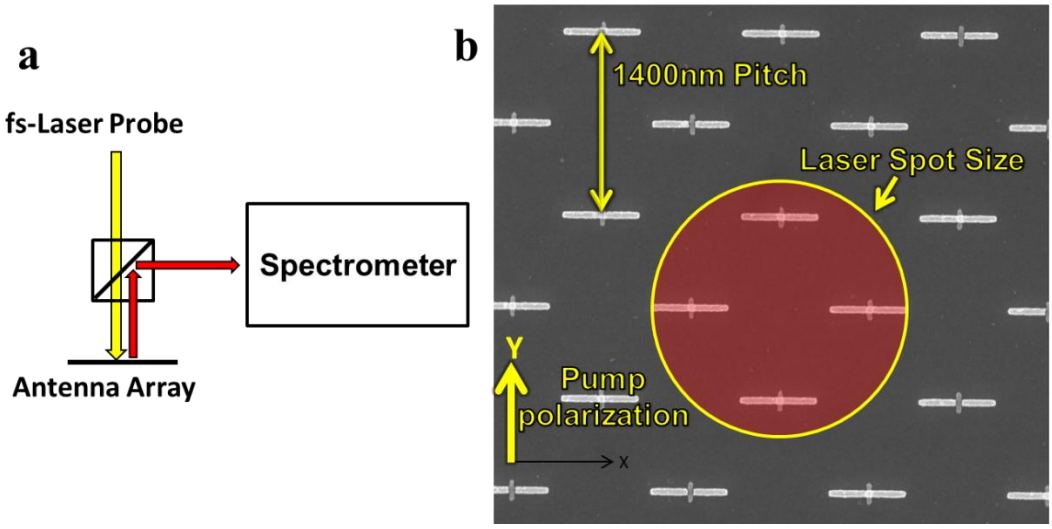


**Figure 3.11** (a) Schematic of the scanning μ-PL microscope. A laser is focused onto the antenna array to inject carriers in the InGaAsP ridges. Light emitted from the devices is collected with the same objective and fed into a spectrometer and CCD. (b) SEM of a fabricated antenna array before substrate removal. The red-shaded yellow circle represents the approximate spots size of the pump laser.

PL scans were performed on the antenna arrays, with each array featuring a different antenna length. Figure 3.12(a) shows the intensity of light emitted polarized in the x-direction for different length antennas and bare ridges. Only the 20 brightest points for each array are shown

(each scan has ~200 points per array). Enhancement is largest for 400nm long antennas, ~35x stronger than the bare ridge, and slowly decreases with increasing antenna length.

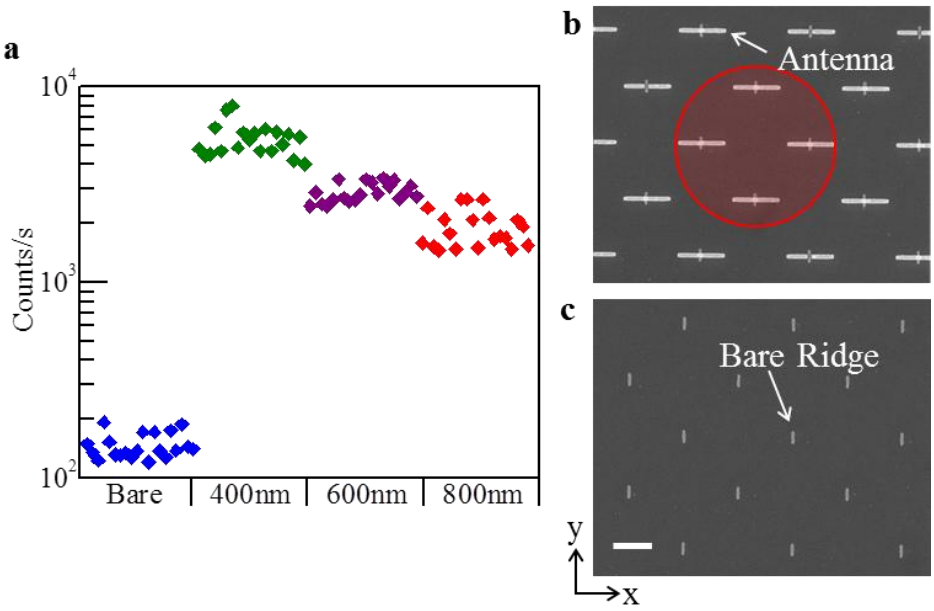


**Figure 3.12** Integrated photoluminescence parallel to antenna long axis. a, Scatter plot showing PL intensity from the brightest 20 points in a scan of individual 20um x 20um InGaAsP ridge arrays coupled to different length antennas. b,c, SEM of antenna-coupled and bare ridge arrays. Red circle outlines approximate spots size of pump laser. (scale bar, 500nm)

Figure 3.13 shows the photoluminescence spectrum collected for the two different emission polarizations for both bare and antenna-coupled ridges. A 400nm long, 50nm wide antenna has very little effect on the light emitted polarized in the y-direction (Figure 3.13(d)), however an improvement of 35x is seen for light emitted polarized in the x-direction (Figure 3.13(e)). The enhancement is very broadband, spanning almost 200nm of spectrum, indicating an antenna Q on the order of ~5. As the antenna arms are lengthened the antenna resonance red-shifts and a corresponding shift in the enhanced spectrum can be observed. The red-shift of the antenna resonance is not linear with antenna length due to the effect of the LC matching arch which has a large effect on the resonance wavelength. Longer antennas show strong peak enhancement, indicating the lower total intensity of light coming from longer antennas can be attributed to the spectral mismatch between their antenna resonance and the spontaneous emission spectra of InGaAsP.

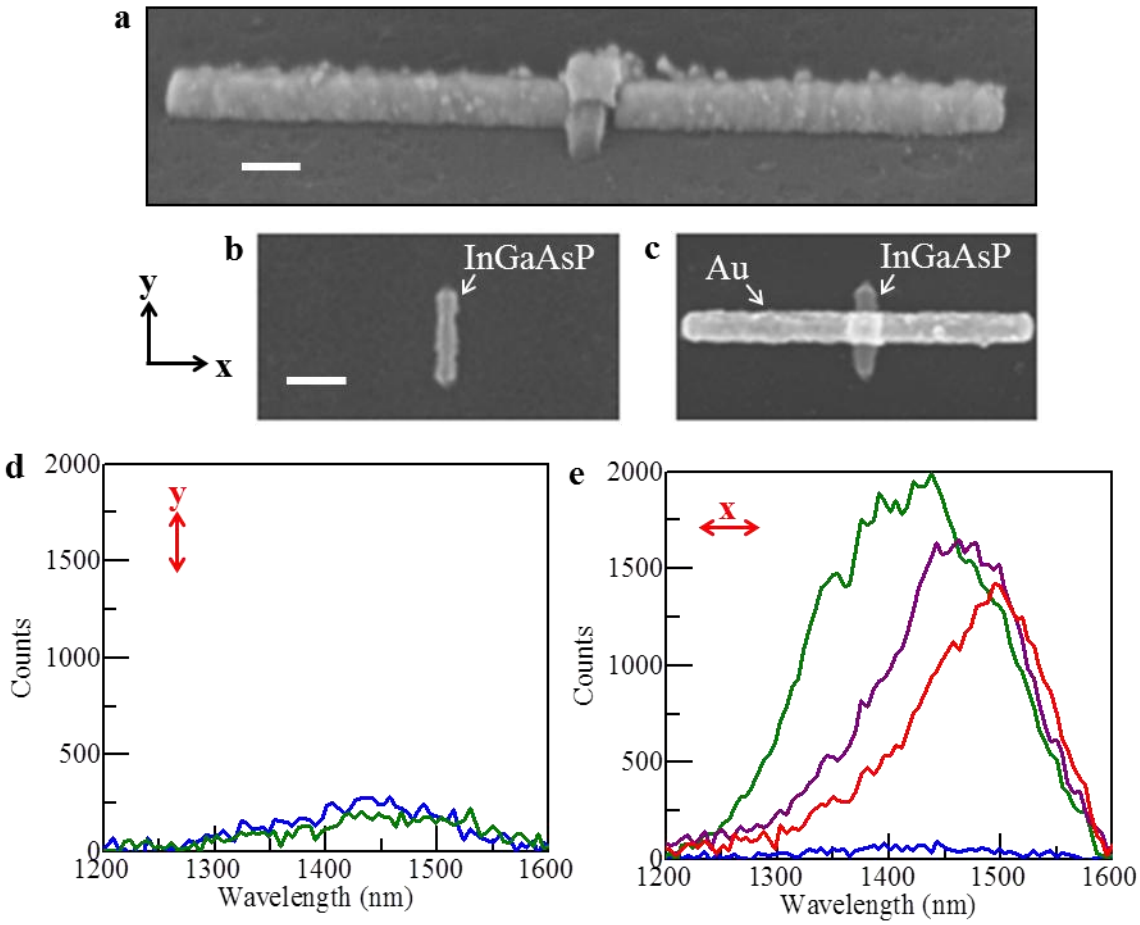


**Figure 3.13** Photoluminescense from InGaAsP nanoridge arrays. a, Perspective SEM of antenna-coupled ridge before substrate removal (scale bar, 50nm). b,c, Top-down SEM image of antenna-coupled and bare ridge (scale bar, 100nm). d, Emitted PL polarized in the y-direction for bare ridge (blue) and antenna-coupled ridge (green). e, PL polarized in the x-direction for bare ridge (blue) and from ridges coupled to different length antennas: 400nm (green), 600nm (purple), and 800nm (red) in length.

The maximum increase in emitted light measured was 35x for the 400nm long antennas. As discussed in Section 3.3, only ~32% of the carriers see the hotspot of the antenna. Taking this into account, the 35x PL increase observed corresponds to a rate enhancement of ~105x. The expected enhancement for this device can be found by taking the ideal enhancement for a properly designed dipole-antenna with 40nm gap from Figure 2.9 and applying the voltage shunting factor of 1/2 discussed in section 3.2. This yields an expected enhancement of 210x, which also agrees well with full-wave simulations of our fabricated structure. The discrepancy between the expected value of 210x and the measured value of 105x could be associated with spatial averaging away from the optimal central spot under the arch antenna. Simulations give an antenna efficiency of 66%.

# Chapter 4

# Waveguide Integrated nanoLED

With a solid footing in the theory and design of devices with enhanced spontaneous emission rates, it's time to turn our attention to figuring out how to use the emitted light. If we really want a fast, efficient, nanoemitter for integrated communication and/or sensing applications, then more likely than not we are going to have to couple the light into a waveguide.

In this chapter I will discuss some simple strategies for efficiently coupling the light emitted from an optical antenna into a waveguide. The only requirement we have on our antenna is that it must have a dipole emission pattern. It turns out that this type of emitter, if sufficiently smaller than the wavelength of emission, can be efficiently coupled into both dielectric multi-mode and single-mode waveguides. I will show simulated coupling efficiencies into a single-mode waveguide as high as 78% and experimental coupling efficiencies up to 70% using an arch-dipole based antenna on an integrated InP multi-mode waveguide.

## 4.1 Trapping Light

We can start by looking at the far-field radiation pattern of a dipole emitter sitting in free space. A cross-section of the power radiated from an arch-dipole antenna in free space is shown in Figure 4.1(b). Perpendicular to the long axis of the antenna, the power radiated is isotropic. Parallel to the long axis, however, there is a null in the far-field emission. This creates the classic "donut" shaped far-field pattern characteristic of dipole emitters [22].

As derived in equation (1.14), the dipole will emit into its surroundings proportional to the refractive index of the dielectric medium, $n$. Therefore, by just placing our dipole emitter on top of a high index substrate with $n = 3.1$, we should expected ~68% of the light to be coupled into the substrate. This case is shown in Figure 4.1(c,d) for an arch-dipole antenna sitting on an InP substrate. It should be noted that in these simulations the dipole was placed in the center of a 35nm tall InGaAsP ridge. The exact height of the dipole above the high-index substrate does effect the coupling efficiency, but because the ridge itself is deeply sub-wavelength ($< \lambda/10$), placing the dipole at 17.5nm is a good approximation of coupling from the whole ridge.

As is well known in bulk LED's, most of the light coupled into the high index substrate will be trapped and not be able to radiate into free space [66]. By restricting the thickness and width of

the InP substrate, an optical waveguide can be formed which utilizes this concept of light trapping to guide light around a chip.

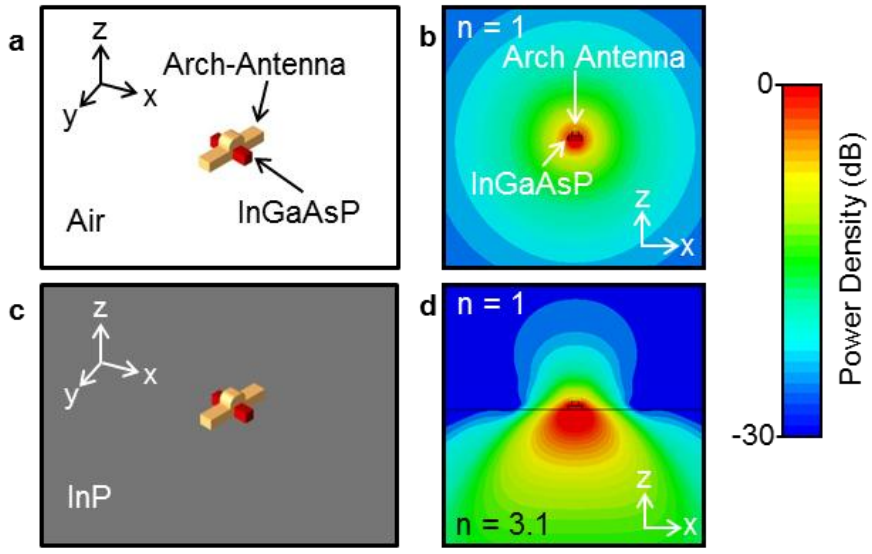


**Figure 4.1** Perspective view and power density emitted from a nanoLED in air (a,b) and on an InP Substrate (c,d).

Figure 4.2(a) depicts an arch-antenna sitting on a slab of InP of finite width and thickness to form a waveguide. The waveguide is cladded in –z direction by vacuum (n=1) and in the +z direction by a low index material (n=1.56) similar to oxide or epoxy. To determine the effect of waveguide thickness on antenna emission patterns, time domain simulations (CST Microwave Studio) were performed using a dipole excitation at the center of an InGaAsP ridge 150nm long, 40nm wide, and 35nm tall. The waveguide was made sufficiently long (±10um) so the total power flowing out of the simulation domain in the ±x direction could be taken as the total power coupled into the waveguide. The waveguide coupling efficiency was then calculated as the ratio of power coupled into the waveguide to the total radiated power, and the confinement factor was calculated as the ratio of power confined within the waveguide to total power coupled into the waveguide mode. With the waveguide width kept at a constant 500nm, the thickness was varied and the subsequent coupling efficiencies and confinement factors were calculated and are plotted in Figure 4.2(b). All simulation results are for a wavelength of 1300nm.

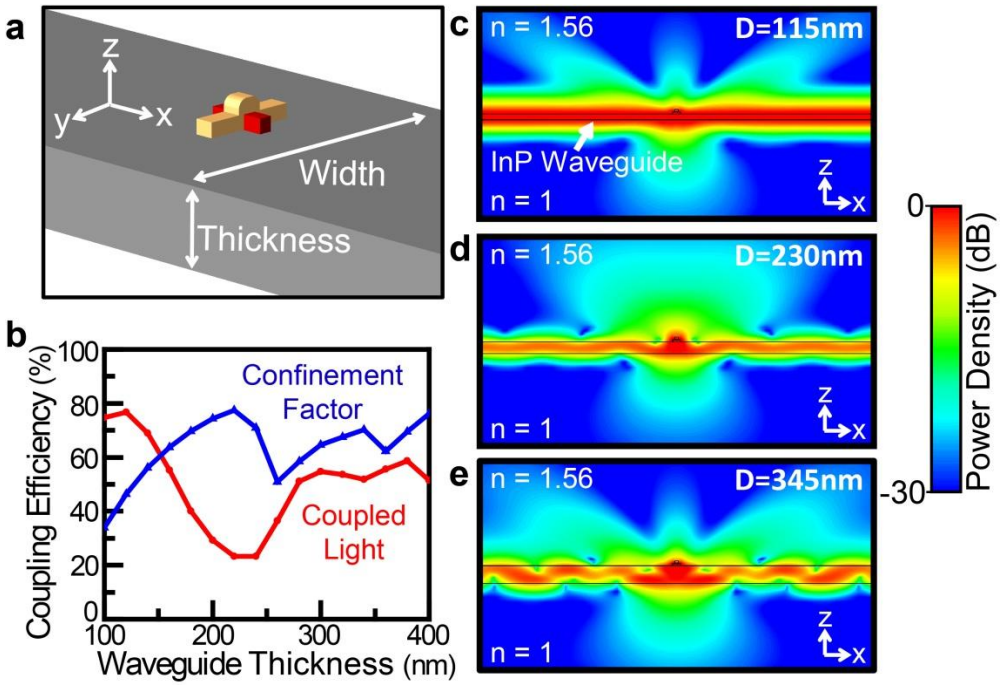


**Figure 4.2** (a) Perspective view of the arch-antenna nanoLED sitting on an InP waveguide. (b) Coupling efficiency (red) and waveguide confinement factor (blue) of light emitted from the nanoLED for waveguides of varying thickness. Power density plot of light emitted from the nanoLED into a (c) 115nm, (d) 230nm, and (e) 345nm thick waveguide.

Peak coupling efficiency occurs when the waveguide is ~115nm thick, which is approximately λ/4. This implies that maximum coupling occurs when light emitted in the –z direction reflects off the bottom interface of the waveguide and destructively interferes with light emitted in the +z direction. This can be seen in Figure 4.2(c) where light emitted normal to the waveguide into the +z direction is attenuated and the principle +z emission is at large angles. Oppositely, the minimum coupling efficiency occurs when the waveguide is ~230nm thick, or approximately λ/2. This corresponds to when light reflecting off the bottom interface of the waveguide constructively interferes with light emitted in the +z direction. Correspondingly, this can be seen in Figure 4.2(d) where there is a strong intensity of light emitted normal to the waveguide in the +z direction. The coupling efficiency peaks again around 3λ/4, ~345nm, for which its emission pattern shown in Figure 4.2(e) is similar to the λ/4 waveguide, except at this thickness the waveguide is multimode. Further increases to the waveguide thickness have little effect on the coupling efficiency.

In addition to the thickness of the waveguide, we can also tune the waveguide width. Three waveguide thicknesses - 115nm, 230nm, and 320nm - were investigated with widths varying from 500nm to 3um. The results, plotted in Figure 4.3, show that the width has much less influence on coupling and confinement than the waveguide thickness. In general, varying the width causes less than a ±10% change in either coupling efficiency or confinement. For almost all cases the 115nm deep waveguide had the highest coupling and lowest confinement, the

230nm deep waveguide had the lowest coupling and highest confinement, and the 320nm waveguide was in between. While in general a 3λ/4 waveguide has lower coupling efficiency than a λ/4 waveguide, the higher confinement could be advantageous as it could decrease scattering losses as light propagates down the waveguide.

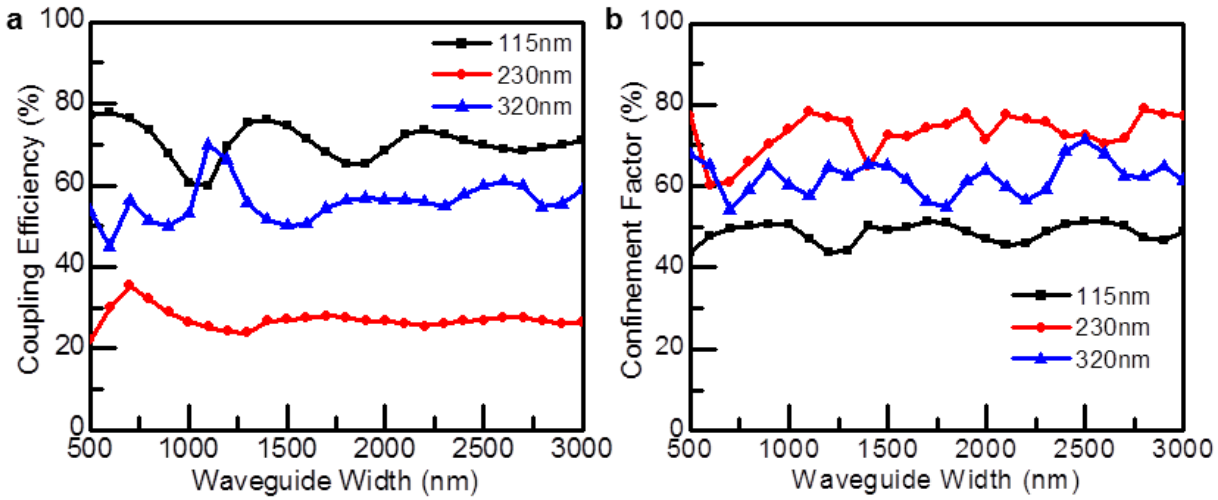


**Figure 4.3** (a) Coupling efficiency and (b) confinement factor of 115nm (black), 230nm (red), and 320nm (blue) thick waveguides for varying widths.

For the remainder of the chapter I will focus on antenna structures coupled to a 320nm thick multimode waveguide. This thickness offers high optical confinement as well as coupling efficiencies up to 70% depending on the width of the waveguide. Coupling to a single-mode waveguide would also be of great interest, especially for a longer range high-speed link. However, the thin dimensions required for good coupling (~115nm) provide difficulty both in propagation loss as well as potential mechanical weakness for the free-standing structures demonstrated later. Also, at the time of these experiments we only had access to epitaxial material with 320nm thick InP layers.

## 4.2 Directional Emission

Figure 4.4(b) depicts the power density of light emitted from an arch-antenna sitting on an InP waveguide 320nm thick and 880nm wide. Due to the symmetry of the structure, an equal amount of light is coupled in both the +x and –x direction along the waveguide. For a point to point optical link, it may be advantageous to have the light directed in a certain direction down the waveguide. This would have the potential to reduce cross-talk, allow for a communication scheme where signals are only detected depending on the direction they come from, or allow for coarse wavelength division multiplexing.

There are two main approaches to realize directional emission: a directional antenna could be used in place of the arch-antenna, or one direction of the waveguide could be truncated with a

reflective element such as a dielectric or metal mirror. Truncating the waveguide will reduce the robustness of the waveguide, obviously prohibiting any components to be placed behind the emitter. For this reason I will first consider the case of using a directive antenna and then briefly touch on the use of a truncated waveguide.

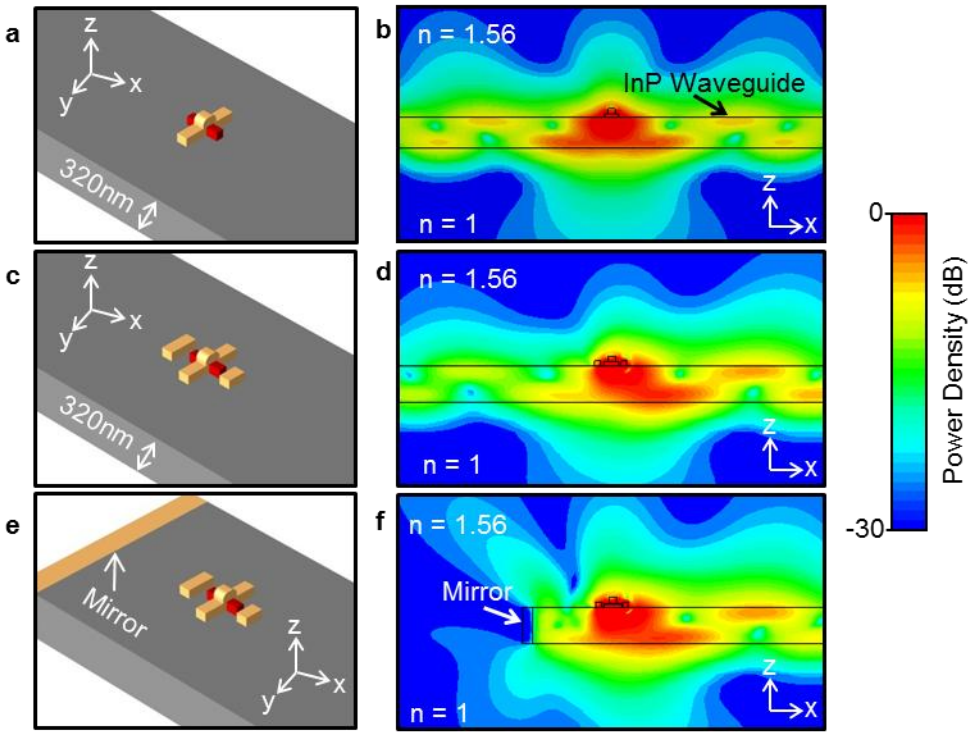


**Figure 4.4** Perspective view and power density emitted from a nanoLED on a 320nm thick InP waveguide (a,b) on a waveguide with a passive reflector and director (c,d) and with a waveguide truncated with a mirror (e,f).

Yagi-Uda antenna structures have been shown to be effective in directing optical light in a preferred direction [34,67–69]. The arch-antenna can be used as the driven element in a Yagi-Uda structure by adding a passive reflector element behind it and a director element in front as depicted in Figure 4.4(c). The simplest antenna element to fabricate is a single gold bar without a feedgap. In order for the light scattered from the reflector and director to have the correct phase shift to constructively interfere in the forward direction, the reflector and director are slightly detuned from resonance. The reflector is made slightly longer than the resonant length while the director is made slightly shorter than the resonant length [70]. The resonant length of a simple gold bar, however, is not the same as an arch-antenna which has additional reactance from the metal arch over the ridge as discussed in Chapter 2 and Chapter 3. Simulations show that a 50nm wide, 40nm thick gold bar sitting on InP must be 125nm long to be resonant at a wavelength of 1300nm, compared to 250nm long for an arch-antenna of the same width and thickness. This will cause both the reflector and director elements to be shorter than the driven element. While it is counter-intuitive that a small element can efficiently direct emission from a longer driven

element, it is well known that small linear antennas have a scattering cross-section independent of length [22,70]. Therefore even a short element can affect fields within an area much larger than their physical size.

A parametric simulation sweep of reflector and director length and spacing was performed to optimize emission into the +x direction. The minimum element spacing was set by the length of the InGaAsP ridge extending out from the arch-antenna, ~50nm depending on alignment tolerances. Simulations find that for a director length (spacing) of 75nm (105nm) and a reflector length (spacing) of 135nm (125nm) a front to back emission ratio of 3:1 can be achieved with a coupling efficiency of 68%. Figure 4.4(d) demonstrates that for a well-tuned Yagi-Uda configuration, the arch antenna excites the same waveguide mode as the non-directional case, but now light is preferentially emitted in the +x direction.

Even when a directional antenna is used, a significant portion of the emitted light couples into the –x propagation direction in the waveguide (17%). This can be reduced by adding more director elements in front of the active-antenna. However, the simplest method to completely eliminate this propagation direction is to truncate the waveguide with a reflector in the –x direction as shown in Figure 4.4(e). If the waveguide is truncated the correct distance behind the antenna with a metal reflector then light coupled in the –x direction will reflect off the mirror and constructively interfere with the light coupled into the +x direction. Using this approach in addition to a Yagi-Uda antenna, simulations show 78% of the light can be coupled into the +x direction of a 320nm thick waveguide truncated with a gold mirror 550nm behind the driven element. This reflector does not have to be metallic, an InP based DBR structure could also be used which would simplify fabrication.

## 4.3 Demonstration of nanoLEDs Coupled to an InP Waveguide

Now that I have demonstrated large coupling efficiencies are theoretically achievable into an InP waveguide, it is time to experimentally demonstrate the structure shown in Figure 4.2(a). Luckily, most of the fabrication techniques developed for the arch-dipole antenna can also be used for this structure. The only major modification is the addition of a waveguide etching step. The following sections detail the fabrication and testing of an InP waveguide coupled nanoLED.

### 4.3.1 Fabrication of Waveguide Coupled Devices

To experimentally validate the coupling efficiencies calculated in Figure 4.2 and Figure 4.4, a nanoLED was fabricated on a 320nm thick InP waveguide. Antennas were first fabricated on a metal-organic chemical vapor deposition (MOCVD) grown (100) InP epitaxial wafer using top-down semiconductor processing techniques. Processing can be broken down into five main steps: (1) alignment mark deposition; (2) InP waveguide patterning (3) InGaAsP ridge formation; (4) Antenna deposition; (5) Substrate removal. All lithography was done using e-beam lithography.

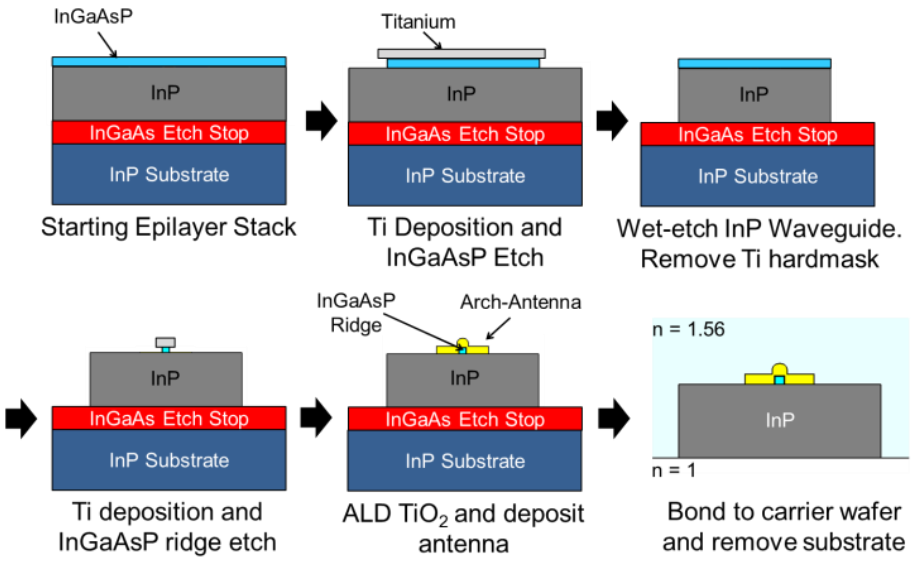


**Figure 4.5** Abbreviated process flow for fabricating a nanoLED on an integrated InP waveguide (ebeam alignment mark deposition not shown).

First, alignment marks consisting of a 5nm titanium adhesion layer and a 25nm gold layer were formed using a liftoff process on top of a 35nm thick InGaAsP layer. Gold was used because it is resistant to the chemical etching used in later processing steps.

Next, thin titanium strips (7nm thick, 50um long, varying width) were patterned and deposited using e-beam lithography followed by e-beam evaporation and lift-off. The titanium islands were aligned along crystal planes so that the long axis was parallel to the [011] direction; these were used as a hard mask to etch ridges into the InGaAsP active layer. A dilute piranha solution (1:8:100 $H_2SO_4$:$H_20_2$:$H_20$) was used to slowly wet etch the InGaAsP layer. Due to the anisotropy of the etchant, the sidewalls parallel to the [011] are fairly vertical for moderate aspect ratios.

The InGaAsP layer is then used as a near perfect hard mask to etch the 320nm InP layer below it without undercut. This allows for very narrow waveguides with smooth vertical sidewalls to be created. At the ends of the wavguides, the <100> direction does undercut the InGaAsP mask, forming a tapered end facet as shown in Figure 4.6.

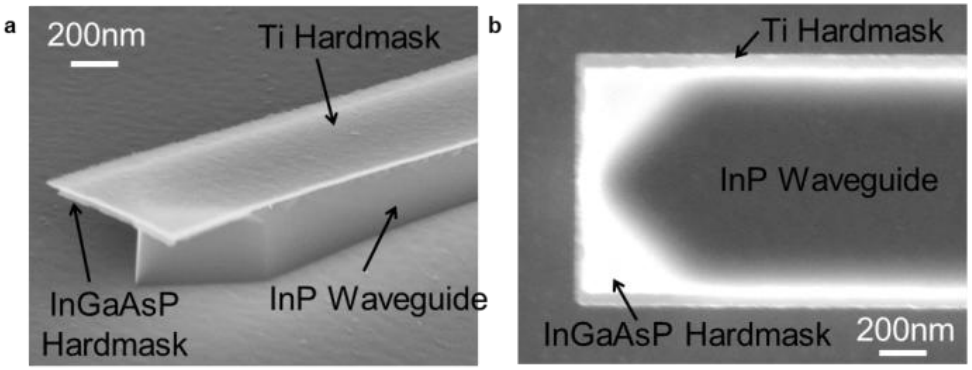


**Figure 4.6** (a) Perspective view and (b) top-down view of end facet of etched InP waveguide with InGaAsP and titanium hard masks still remaining.

After stripping the titanium mask with a quick dip in 49% hydrofluoric acid, a new titanium hardmask (7nm thick, 200nm long, 120nm wide) is patterned and deposited on the InGaAsP layer to define the InGaAsP nano-ridges. A dilute piranha solution (1:8:100 $H_2SO_4$:$H_20_2$:$H_20$) was then used to slowly wet etch the InGaAsP layer. The etching is done in small steps (a few seconds of etching at a time) and then checked under SEM. Because the hardmask layer is so thin, it is semi-transparent to the electron beam and the width of the underlying ridge can be determined. Etching continues until the ridge is etched to a width of ~35nm and then the titanium hardmask is removed with a quick dip in 49% hydrofluoric acid.

Immediately after removal of the titanium hardmask, the sample is loaded into an atomic layer deposition (ALD) machine and pumped down to medium vacuum (~1 Torr). A conformal 3nm layer of $TiO_2$ is then deposited on the sample at 150°C using ALD. This thin oxide layer provides a uniform surface for all the ridges and prevents direct contact with the metal antenna deposited in the subsequent step.

Next, 50nm wide antennas are patterned perpendicularly over the InGaAsP ridges. Reflector and director elements are also patterned for Yagi-Uda structures. 2nm of germanium is then evaporated at a rate of 0.4 Å/s followed by 40nm of gold deposited at a rate of 1.2 Å/s. The slow deposition rates yield polycrystalline metal but offers superior conformal coverage compared to quickly evaporated metal. The germanium layer acts as a wetting layer for the gold to ensure connection of the metal arch over the ridge. The metal is then lifted off leaving the antenna coupled InGaAsP ridges. SEM images are taken at this point to validate successful fabrication of the arch-antenna structure.

After completion of the arch-antenna structure, the entire chip is flipped over and bonded to a quartz carrier with UV curable epoxy (NOA 81). After UV curing, the epoxy is left to hardbake for 12 hours at 50°C. The InP substrate is then mechanically lapped with fine-grit sandpaper to a thickness of ~50um. The remaining 50um is then etched off with a 1:1 solution of Hydrochloric Acid:Phosphoric Acid at 80°C and stops on a 150nm thick InGaAs etch stop layer. The InGaAs

layer is removed with a 1:1:10 solution of $H_2SO_4$:$H_20_2$:$H_20$ leaving only the InP waveguides embedded in the epoxy.

For the 3um wide waveguides an InGaAsP hardmask was not required, instead the waveguides were patterned using only a titanium hardmask followed by etching in 1:3 HCL:Phosphoric Acid. This was done on the backside of the sample after substrate removal in contrast to the narrower waveguides which were etched before substrate removal. Etching from the bottom side of the wafer inverts the crystal plane orientation resulting in sloped sidewalls and vertical end facets. A comparison of a 3um waveguide and an 880nm waveguide is shown in Figure 4.7.

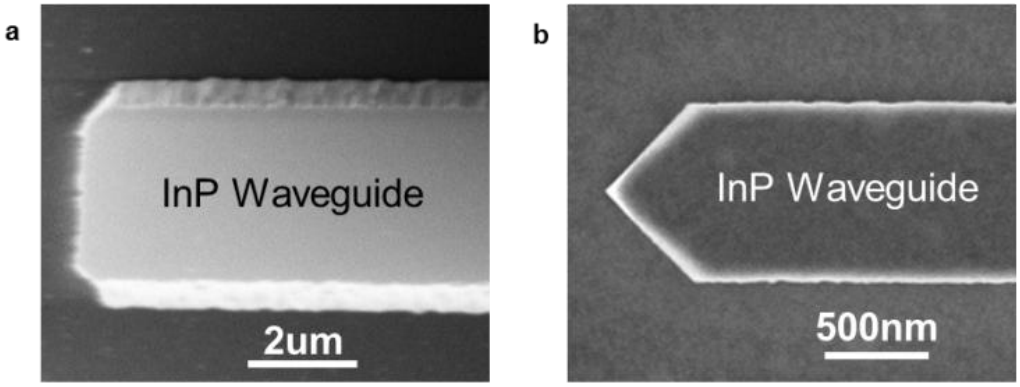


**Figure 4.7** End facet of a (a) 3um wide waveguide shown on epoxy after substrate removal and a (b) 880nm wide waveguide shown on InGaAsP etch stop before substrate removal.

### 4.3.2 Measuring Coupling Efficiency

Optical emission measurements were performed by optically injecting carriers into the InGaAsP ridge using a Ti:Sapphire laser with center wavelength of 720nm that was focused onto the ridge with a 100x 0.8NA objective. The laser is polarized perpendicular to the antenna to minimize pump enhancement in the antenna-coupled ridge. The InP waveguide is absorbing at the pump wavelength which allows for a greater amount of carriers to be injected into the InGaAsP and also helps minimize the effect of the antenna on pumping conditions.

Light emitted from the structure was either collected with the same objective used to pump the sample or through a second 100x 0.8NA objective on the opposing side of the pump objective as shown in Figure 4.8(a). By using two identical objectives, the pump can be kept constant and emitted light can be collected from both the top and bottom side of the sample. Collected light was either directly detected with a linear InGaAs CCD to observe the spatial pattern of emitted light, or fed through a spectrometer first to determine spectral information.

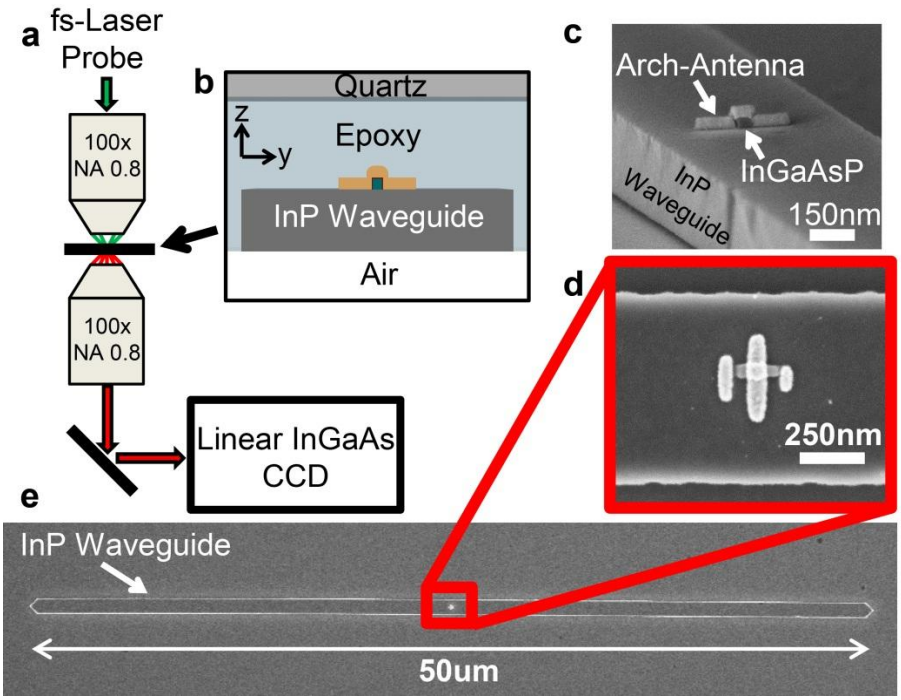


**Figure 4.8** (a) Schematic of measurement setup. (b) Cross section of nanoLED on an InP waveguide bonded to a quartz handle wafer with epoxy. (c) Perspective and (d) top down SEM of fabricated nanoLED structure on a 320nm thick InP layer. (e) Top view SEM of fabricated nanoLED on a 50um long InP waveguide.

Light emitted by the InGaAsP ridge is either coupled into the InP waveguide or radiates out into free space. The coupled light will travel down the waveguide and when it reaches the end facet the majority will scatter out (>70%) while the remainder will be reflected back into the waveguide and bounce back and forth between the two ends until it eventually scatters out. Sidewall roughness can be ignored since even a very lossy propagation loss of 10dB/cm would only yield 1.1% loss from one end of the waveguide to the other. By imaging the waveguide with the InGaAs CCD, it can be determined from where light is being emitted. Light that is not coupled to the waveguide will come directly from the antenna at the center of the waveguide while the coupled light will be scattered out the ends of the waveguide.

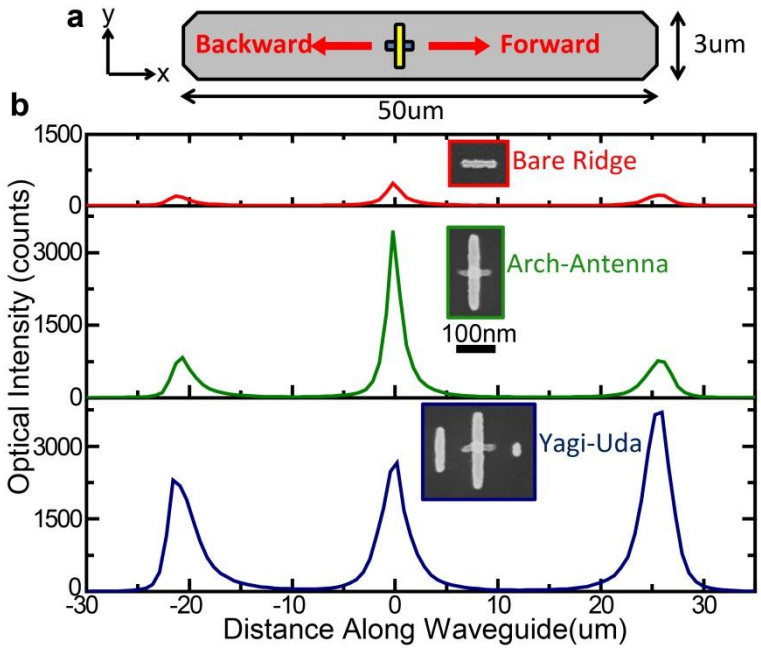


**Figure 4.9** (a) Schematic drawing of waveguide with single device at the center indicating forward and backward directions. (b) Light intensity as a function of distance along the waveguide for a bare ridge (red) 250nm long arch-antenna (green) and a Yagi-Uda antenna (blue).

The spatial image captured by the CCD of light emitted from a 3um wide waveguide structure is shown in Figure 4.9(b). At the very center of the waveguide, where the single device is located, there is a large peak corresponding to light that is coming directly from the device without coupling into the waveguide. There is then very little light observed until the end of the waveguide 25um from the center, corresponding to light that has been coupled into the waveguide and then scattered out the end. For the bi-directional structures the central peak is the largest since the coupled light is split into two peaks at either side of the waveguide. For the Yagi-Uda antenna, more light is emitted in the forward waveguide direction than in the backwards direction. A front-to-back ratio of ~1.6:1 can be calculated from Figure 4.9(b). This ratio is most likely limited by reflections at the end facet of the waveguide. Narrower waveguides with tapered end facets showed front-to-back ratios as high as 3:1, in good agreement with simulated values.

To estimate the coupling efficiency, the number of photons collected from the two ends of the waveguide is compared to the total number of photons collected. Doing this calculation based on the data in Figure 4.9(b) yields coupling efficiencies of 57.5%, 43.9%, and 71.5% for the bare, arch-antenna coupled, and Yagi-Uda coupled cases respectively. Evaluation of the simulated farfield emission patterns (Appendix B) suggest this gives a reasonable estimation of coupling efficiency. Similar devices within the same sample show similar coupling results, with fabrication variability having a larger effect on total light collected.

A spatial map of light emitted from nanoLEDs coupled to an 880nm wide waveguide are shown in Figure 4.10. Although the measured coupling efficiencies for these structures were slightly lower than for the 3um wide waveguides (30%, 40%, and 54% for a bare, arch-antenna, and Yagi-Uda antenna resptectively), the Yagi-Uda structure showed a front-to-back ratio of 3:1. This is most likely due to minimized internal waveguide reflections in the narrower waveguide from the tapered end facets as shown in Figure 4.7.

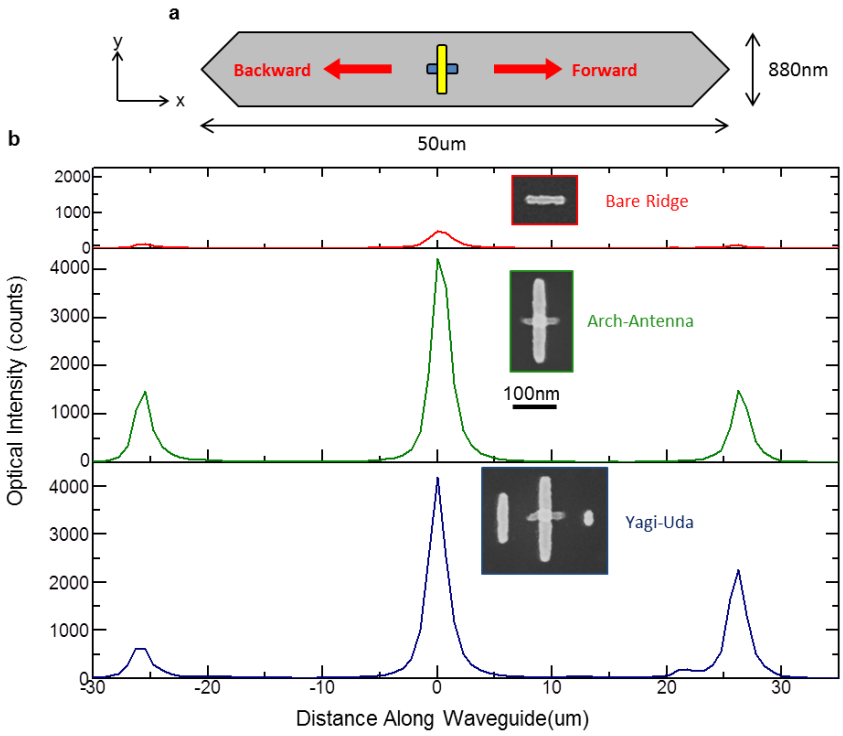


**Figure 4.10** (a) Schematic drawing of an 880nm wide waveguide with a single device at the center indicating forward and backward directions. (b) Light intensity as a function of distance along the waveguide for a bare ridge (red) 315nm long arch-antenna (green) and a Yagi-Uda antenna (blue).

### 4.3.3 Measurement of Rate Enhancement

The measurements performed in the previous section not only confirm high coupling efficiencies, but also show that the antenna (both arch-dipole and Yagi-Uda) is still functioning to enhance the spontaneous emission rate. In Figure 4.9(b), although the bare ridge and arch-antenna show similar coupling efficiencies, ~5x more light is emitted from arch-dipole coupled ridges. The Yagi-Uda structure also shows large enhancement, ~11x more total light emitted compared to the bare ridge.

Antenna-coupled devices coupled to an 880nm wide waveguide shown in Figure 4.10(b) also show enhanced emission. To better compare the enhanced emission, the emitted light was fed through a spectrometer and read out on an InGaAs CCD. Figure 4.11 shows the spectral

emission the devices shown in Figure 4.10. The arch-antenna exhibits about 12x enhanced emission compared to the bare ridge, while the Yagi-Uda emission is enhanced by about 9x. To achieve both high enhancement and large front-to-back ratios, the driven element length, arch height, reflector and director spacing and lengths must all be optimized. Since only a small subset of combinations of these six parameters was actually fabricated, it is very unlikely any one device is much optimized.

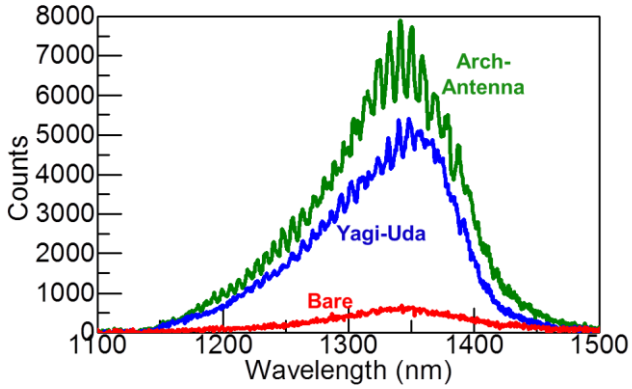


**Figure 4.11** Emission spectrum from a bare ridge (red), arch-antenna (green), and a Yagi-Uda antenna (blue) coupled to a 880nm wide InP waveguide.

The high frequency oscillations in the spectrums shown in Figure 4.11 can be attributed to reflections at the end facets of the waveguides creating weak fabry-perot resonances in the waveguide. These oscillations were much less pronounced in the 880nm wide waveguides compared to the 3um wide waveguides.

Using the increase in light emitted as a measure of rate enhancement is only valid if the enhanced and reference case are dominated by non-radiative recombination, the antenna does not alter the collection efficiency, and the presence of an antenna does not alter the pumping conditions. The first was proven true in Section 3.2 and the second is shown to be valid in Appendix B.

To show that a similar number of carriers are pumped into both bare and antenna coupled devices, 3D time-domain simulations were performed to simulated power absorbed in the devices. The bare and antenna coupled structures were illuminated with a plane wave and the amount of power absorbed in the ridge and in the InP waveguide was calculated. Since carriers can diffuse into the ridge from the InP waveguide, it acts as a reservoir of carriers. To get an estimate of the amount of carriers generated in each layer, an integration of power absorbed in the InP was calculated over hemispheres of equal distance from the center of the ridge.

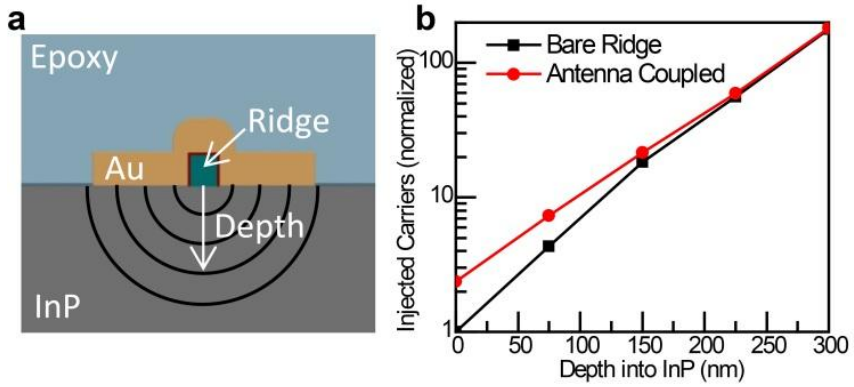


**Figure 4.12** Integrated number of carriers created as a function of depth into InP waveguide. Normalized to the number of carriers generated in the bare ridge of InGaAsP (0nm on the plot).

Figure 4.12(b) shows the number of carriers absorbed as a function of depth into the InP layer. In the ridge itself, the antenna enhances the absorption by up to 2.4x, though this number could be smaller in a non-ideal structure (such as one with rounded edges). The amount of carriers generated in the InP within 75nm of the bottom of the InGaAsP in addition to the ones generated in the InGaAsP is 1.7x larger than the case without an antenna. The total number of carriers generated in the ridge and within 150nm of the ridge, is almost identical for both cases. Carrier recombination in the thin 320nm InP waveguide will be dominated by the surfaces. Since InP is a relatively efficient material with good surfaces [71], it is reasonable to estimate a carrier diffusion length >> 100nm, which suggests most of the carriers recombining in the InGaAsP will have actually have been generated in the InP layer. This agrees well with experimental observation of significantly brighter emission when pumping the structure with a wavelength shorter than the bandgap of InP. Although a slight increase in pumping may be observed with an antenna, it is on the order of ~1x.

To predict the increase in spontaneous emission rates caused by the presence of the gold antenna, time domain simulations (CST Microwave Studio) were performed of both the bare and antenna clad InGaAsP ridges. For a ridge 45nm wide, 35nm tall, 150nm long coated in 3nm of $TiO_2$ and sitting on a 320nm thick by 3um wide InP waveguide, 87x more power was radiated when it was coupled to a 250nm long arch-antenna compared to when it was bare. Due to the wider gap dimensions (~51nm vs 40nm) and presence of the InP waveguide, this is less than half the enhancement expected from the free-space arch-dipole from Chapter 3. The arch-height of the arch-dipoles was not re-optimized in the presence of the high-index waveguide. Doing so would most likely yield rate enhancements up to ~140x for the same gap width.

The efficiency of the arch-dipole antenna was simulated to be 66.4%. For an antenna in a Yagi-Uda formation, a similar enhancement of 79x was obtained with an antenna efficiency of 53.9%. These values, however, are an optimistic situation where the emitting dipole is perfectly placed within the antenna gap. As was the case for the free-standing structures, only ~1/3 of the ridge was covered by the antenna. The maximum observed emission enhancement of 12x therefore

corresponds to a rate enhancement of ~36x. The discrepancy between the expected value of 87x and the measured value of 36x could be a combination of spatial averaging away from the optimal central spot under the arch antenna as well as fabrication imperfections such as small metal grain size, imperfect metal coating of the ridge, and size variations in the ridge (±5nm).

# Chapter 5

# The Slot Antenna

So far our discussion of optical antennas has focused on the dipole antenna. While it is one of the most popular types of antennas for its simplicity, it is far from the only good antenna design. Using a dipole antenna has some obvious limitations with a 2D semiconductor, which we will explore in greater detail in Chapter 6. Before, we used the height and width of the emitter to define a self-aligned feedgap for the arch-dipole antenna. This strategy will obviously not work for a 2D material. In addition, we want to etch the material so it only sees the hotspot of the antenna, which for a dipole antenna would be < 50 x 40 $nm^2$. To reach the highest PL enhancement possible we would like to avoid uncoupled material like the wings that stuck out for the arch-dipole in Chapter 3.

Another popular antenna design that can solve many of these issues is the slot antenna, shown schematically in Figure 5.1(b). The slot antenna is the dual of the simple dipole antenna shown in Figure 5.1(a). In the 1800's, French physicist Jacques Babinet showed that similar diffraction patterns are produced by complimentary metal structures. In other words, the scattering of electromagnetic energy from a bar of metal surrounded by a plane of dielectric is equivalent to the scattering from the same size bar of dielectric surrounded by a plane of metal. We can therefore expect very similar antenna properties between the two structures shown in Figure 5.1.

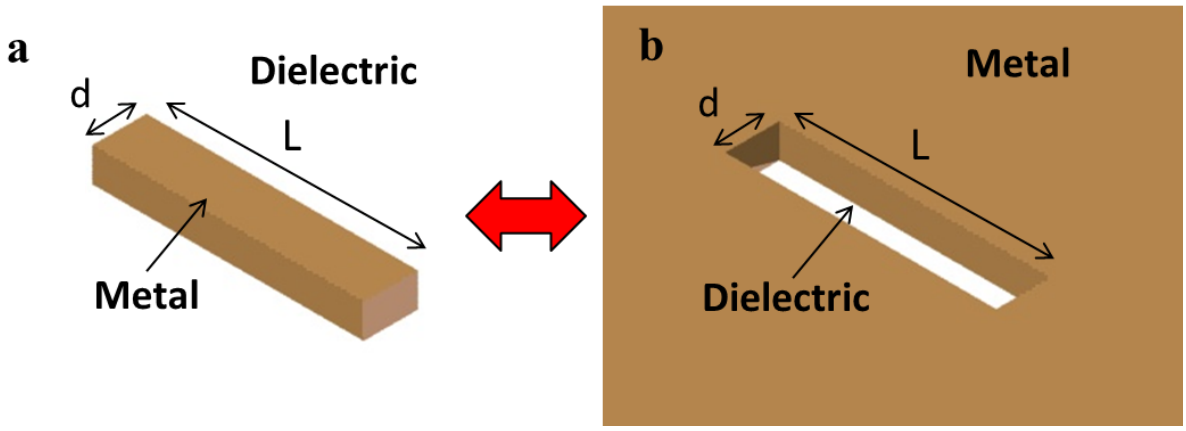


**Figure 5.1** Schematic diagram of (a) a simple dipole antenna and its dual structure (b) the slot antenna.

Furthermore, although the farfield diffraction patterns are similar, the fields of the two structures are actually inverted:

$$\mathcal{E}_{slot} = H_{dipole}$$
$$H_{slot} = -\frac{\mathcal{E}_{dipole}}{Z_0^2} \tag{5.1}$$

This is actually a very good thing for us since this moves the high electric field into the slot portion of the antenna, as shown in Figure 5.2. This provides a much larger area than a feedgap in which to put our optical emitter for good coupling and also allows for a simplified self-aligned etch that will be discussed in section 6.3.1. In the next few sections I will show that the slot antenna offers the following benefits: elimination of a reactive gap shunt path, strong coupling to 2D materials, simple fabrication, and directional emission.

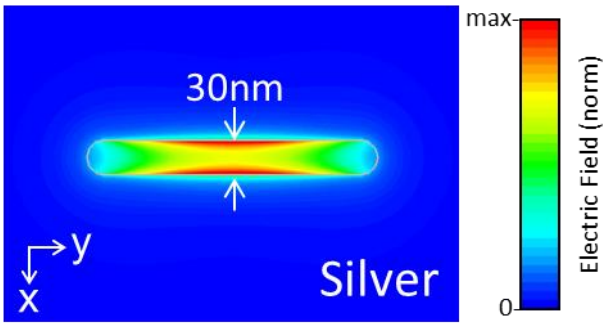


**Figure 5.2** Electric field of the fundamental mode of a slot antenna 30nm wide and 250nm long.

## 5.1 Circuit Theory of an Optical Slot Antenna

Since we already have a well-developed understanding of the dipole antenna, it is fairly straight forward to derive the circuit for its dual, the slot antenna. It is important to remember that the dual of the slot antenna is the dipole antenna, not a dipole antenna with a feedgap. The dipole antenna with a feedgap is actually a perturbation of a true dipole antenna (albeit a necessary one). We will then start with the circuit for a simple dipole antenna that is *l* long, *d* wide, and *t* thick as shown in Figure 5.3. Note that I have removed the source from the antenna (including the spreading resistance which manifests due to the source). I have done this because we only want to look at the antenna mode itself right now, we'll figure out how to put in the source later.

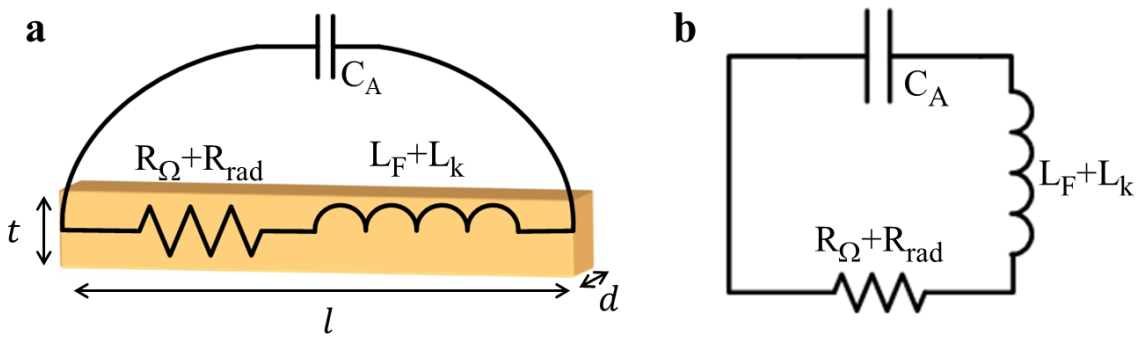


**Figure 5.3** (a) Schematic of a simple dipole antenna with lumped circuit components. (b) Simplified circuit diagram of a dipole antenna.

In the 1940's Booker extended the work of Babinet to dipole and slot antennas, showing that their respective impedances are related by [72]:

$$Z_{slot}Z_{dipole} = \frac{Z_A^2}{4} \tag{5.2}$$

I have used $Z_A$ as the impedance of the dielectric surrounding the final slot antenna($Z_o/n_A$). We can use this equation to transform the circuit elements in Figure 5.3 to their respective slot antenna forms. Unfortunately, Babinet's theorem only applies to an infinitely thin lossless metal plane; meaning this transformation will only work for the faraday inductance, antenna capacitance, and radiation resistance as shown in Table II.

**Table II: Babinet Transformed Slot Antenna Circuit Parameters**: Circuit parameters of dipole antenna transformed to their complimentary forms for the slot antenna using Babinet's Principle [72].

| **Dipole Antenna** | | | **Slot Antenna** |
|---|---|---|---|
| Babinet Transformation Terms | | | |
| Faraday Inductance | $L_f = \frac{\mu_o}{2\pi} l_{eff} * \ln\left(\frac{2l}{t}\right)$ | $\leftrightarrow$ | $C_A^{slot} = \frac{2\epsilon_A l_{eff}}{\pi} \ln\left(\frac{2l}{t}\right)$ |
| Antenna Capacitance | $C_A = \frac{\pi\epsilon_A l_{eff}}{2\ln\left(\frac{2l}{t}\right)}$ | $\leftrightarrow$ | $L_f^{slot} = \frac{\pi\mu_o l_{eff}}{8 * \ln\left(\frac{2l}{t}\right)}$ |
| Radiation Resistance | $R_{rad} = \frac{2\pi}{3} Z_o \left(\frac{l_{eff}}{\lambda}\right)^2 n_A$ | $\leftrightarrow$ | $R_{rad}^{slot} = \frac{3}{8\pi n_A^3} Z_o \left(\frac{\lambda_0}{l_{eff}}\right)^2$ |

Note that I have changed the equations for faraday inductance and antenna capacitance to depend on $l_{eff}$ instead of $l$ as was the case for a dipole antenna in Table I. This is because I had assumed

a sinusoidal current distribution in that case, now I am making the equations more general. In general $l_{eff}$ can be approximated by assuming a sinusoidal current distribution that follows the electric field distribution in the slot. Boundary conditions require the electric field to go to zero near the end of the slot plus a skin depth, $\delta_{skin}$, penetration into the metal. $l_{eff}$ can then be calculated as the average current:

$$l_{eff} = \frac{1}{I_0}\int_{-\infty}^{\infty} I(x)dx = \frac{1}{I_0}\int_{-\delta_{skin}}^{l+\delta_{skin}} I_0 \cos\left(\frac{\pi}{2}\frac{x}{l+\delta_{skin}}\right)dx = \frac{2}{\pi}(l+2\delta_{skin})$$

$$\frac{l_{eff}}{l} = \frac{2}{\pi}\left(1+\frac{2\delta_{skin}}{l}\right) \tag{5.3}$$

Longer antennas will therefore have the characteristic $\frac{2}{\pi}l$ effective length of a half-wave dipole while shorter antennas will have a longer effective length.

Many of the entries in Table II seem to look quite familiar. The faraday inductance transforms to a capacitance that looks very similar to our original antenna capacitance and the antenna capacitance transforms to something very similar to our original faraday inductance. Note that the pre-factors for the respective reactive terms change, so the transformation effectively changes the weight of the respective inductive and capacitive terms.

We are now left to figure out the ohmic resistance of the slot antenna and therefore also the kinetic inductance. Unfortunately, deriving the ohmic resistance in a non-lossless slot antenna is quite a challenging task [73]. To the best of my knowledge, there is no closed-form analytical solution to this problem. We can start however, by transforming the ohmic resistance of the dipole antenna as we did the other terms in Table II, keeping in mind that this isn't strictly "valid". Doing so yields:

$$R_{ohmic}^{slot} = \frac{Z_0^2}{4}\left(\frac{l_{eff}}{td}Re\{\rho\}\right)^{-1} \propto \frac{Z_0^2}{4 * Re\{\rho\}}\frac{td}{l_{eff}} \tag{5.4}$$

This term, however, doesn't seem to take into account the fact that current in a metal plane can spread out. While the size of the slot will most likely cause current crowding that leads to resistance similar to what is expressed in equation (5.4), we need to add a term that accounts for this current spreading. If we picture the current starting in a disk of charge around our slot of radius *a* and spreading out to a radius *b*, it will have a corresponding resistance of:

$$R_{planar} = Re\{\rho\}\int_a^b \frac{dr}{t * 2\pi r} = \frac{Re\{\rho\}}{2\pi t}\ln\left(\frac{b}{a}\right) \tag{5.5}$$

Which then transforms into the slot equivalent form of:

$$R_{planar}^{slot} = \frac{Z_0^2}{4n_A^2}\left(\frac{Re\{\rho\}}{2\pi t}\ln\left(\frac{b}{a}\right)\right)^{-1} \propto \frac{Z_0^2 t}{n_A^2 * Re\{\rho\}} \tag{5.6}$$

Neither of the terms in equations (5.4) or (5.5) really correctly take into account the shape of the slot and current distribution, but their functional forms should be correct. For this reason I have used a proportional sign since we don't know the actual pre-factors. While further investigation is required for a better understanding of the underlying physics, for now we will use the pre-factors that seem to fit well with simulated results. These terms are given in Table III and will be discussed further in section 5.2.

**Table III: Resistive and Local Slot Antenna Circuit Parameters**: Circuit parameters of dipole antenna transformed to their complimentary forms for the slot antenna.

| **Dipole Antenna** | | | **Slot Antenna** |
|---|---|---|---|
| Transformed Ohmic Terms | | | |
| Ohmic Resistance | $R_{ohmic} = \frac{l_{eff}}{td} Re\{\rho\}$ | $\leftrightarrow$ | $R_{ohmic}^{slot} = \left[\left(\frac{l}{l_{eff}}\right)^4 \frac{n_A}{n_g}\right] \frac{Z_0^2}{4 * Re\{\rho\}} \frac{td}{l_{eff}}$ |
| Kinetic Inductance | $L_k = \frac{l_{eff}}{td} Im\{\rho\}$ | $\leftrightarrow$ | $C_k^{slot} = \left[\left(\frac{l_{eff}}{l}\right)^4 \frac{n_g}{n_A}\right] \frac{4 * Im\{\rho\}}{Z_0^2} \frac{l_{eff}}{td}$ |
| Ohmic Planar Resistance | $R_{planar} = \frac{Re\{\rho\}}{2\pi t}\ln\left(\frac{b}{a}\right)$ | $\leftrightarrow$ | $R_{planar}^{slot} = \frac{Z_0^2 t}{4\pi n_A^2 * Re\{\rho\}}$ |
| Kinetic Planar Inductance | $L_k = \frac{Im\{\rho\}}{2\pi t}\ln\left(\frac{b}{a}\right)$ | $\leftrightarrow$ | $C_{kplanar}^{slot} = \frac{4\pi n_A^2 * Im\{\rho\}}{Z_0^2 t}$ |
| Local Terms | | | |
| Gap Capacitance | $C_{gap} = \frac{\epsilon_g td}{d}$ | $\leftrightarrow$ | $C_{gap} = \frac{\epsilon_g t l_{eff}}{d}$ |

Several of the terms transform into a very peculiar looking form. For example, the kinetic inductance term becomes a capacitive term. Also, the radiation and ohmic resistance terms now follow the opposite trend of what they did (i.e. a longer length slot has a lower resistance). The reason for this becomes clear when we put these parameters into the transformed circuit model, shown in Figure 5.4. Since the dipole antenna is a series RLC network, it makes sense that by transforming it through equation (5.2) it becomes a *parallel* RLC network.

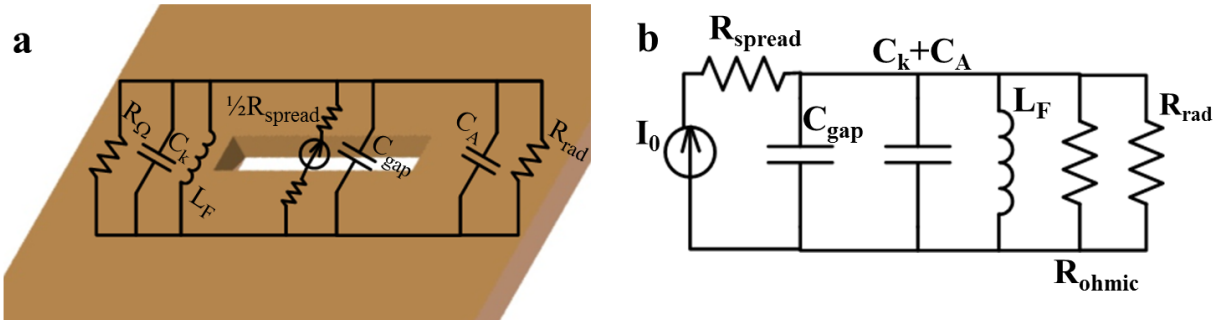


**Figure 5.4** (a) Schematic of a slot antenna superimposed with its lumped circuit elements. (b) Simplified circuit for the slot-antenna.

The dipole source will go in the slot of the antenna where the high-field region is, which as pictured in Figure 5.4(a), is in parallel to all the other terms. Finally, we have to add in the two local circuit elements that were not in the dipole antenna circuit. First, the spreading resistance goes in series with the current source as it did in our previous circuit model in Figure 2.7. Second, there is the non-negligible parallel plate capacitance going across the slot which I will refer to as the gap capacitance. This term does not come from the Babinet transformation since that theorem assumes an infinitesimally thick film, which would make $C_{gap} \to 0$. In a real antenna, however, there is some thickness to the metal and therefore some capacitance. Traditional microwave theory still is able to ignore the very small contribution of gap capacitance. If we write the combination of $C_{gap}$ plus $C_A^{slot}$ we can see why:

$$C_{gap} + C_A^{slot} = \frac{\epsilon_g t l_{eff}}{d} + \frac{2\epsilon_A}{\pi} l_{eff} * \ln\left(\frac{2l}{t}\right) = \epsilon_0 l_{eff}\left(n_g^2 \frac{t}{d} + n_A^2 \frac{2\ln\left(\frac{2l}{t}\right)}{\pi}\right) \qquad (5.7)$$

The second term in the parenthesis in equation (5.7) is approximately ~2, leaving the total capacitance equal to, $C_{total} \approx \epsilon_0 l_{eff}\left(2 + \frac{t}{d}\right)$. In a typical slot antenna on a PCB the metal thickness is ~35μm with a minimum trace width of >100μm. This means the ratio of *t*/*d* will be quite small and the gap capacitance can be safely ignored. In the optical regime this is not the case as the gap dimension will usually be on the same order of magnitude as the metal thickness and may be filled with a high index dielectric. As the gap is scaled down for higher enhancement, the ratio of *t*/*d* will actually be much larger than one in order to maintain a low loss antenna and the gap capacitance therefore becomes important.

One of the most striking differences between the circuit in Figure 5.4(b) and the circuit in Figure 2.7 is the fact that there is no shunting term. On resonance, the slot antenna will simplify to the circuit shown in Figure 5.5.

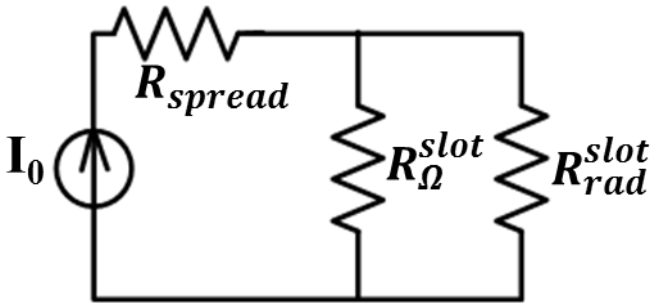


**Figure 5.5** Simplified circuit of a slot antenna on resonance.

Here I have used the symbol $R_{\Omega}$ to denote the total ohmic resistance:

$$R_{\Omega}^{slot} = \left(\frac{1}{R_{ohmic}^{slot}} + \frac{1}{R_{planar}^{slot}}\right)^{-1}$$

$$= \left[\frac{1}{\left[\left(\frac{l}{l_{eff}}\right)^4 \frac{n_A}{n_g}\right]\frac{Z_0^2}{4 * Re\{\rho\}}\frac{td}{l_{eff}}} + \frac{1}{\frac{Z_0^2 t}{4\pi n_A^2 * Re\{\rho\}}}\right]^{-1}$$

$$= \frac{Z_0^2 t}{4 n_A Re\{\rho\}}\left[\frac{l_{eff}}{d}\left(\frac{l_{eff}}{l}\right)^4 n_g + \pi n_A\right]^{-1} \tag{5.8}$$

Unlike the dipole antenna, a large gap capacitance does not shunt current away from the radiation resistor. Instead, it just shifts the resonance frequency. To maintain a particular resonance frequency, the length of a dipole antenna only changes by a factor of two between a shorted gap and a very large gap. Assuming a slot in vacuum with $l_{eff} = l$, the resonance frequency of a slot antenna is given by:

$$\omega_0 = \sqrt{\frac{1}{LC}} \approx \sqrt{\frac{1}{\frac{\pi\mu_o l}{8 * \ln\left(\frac{2l}{t}\right)}\left[\epsilon_0 l\left(\frac{t}{d} + \frac{2\ln\left(\frac{2l}{t}\right)}{\pi} + \frac{4Im\{\rho\}}{Z_0^2 td}\right)\right]}}$$

$$= \frac{c}{l}\sqrt{\frac{8 * \ln\left(\frac{2l}{t}\right)}{\frac{t}{d} + \frac{2\ln\left(\frac{2l}{t}\right)}{\pi} + \frac{4Im\{\rho\}}{Z_0^2 td}}} \tag{5.9}$$

To maintain resonance while scaling down the gap, the length must be adjusted according to:

$$l = \frac{c}{\omega_0}\sqrt{\frac{8 * \ln\left(\frac{2l}{t}\right)}{\frac{t}{d} + \frac{2\ln\left(\frac{2l}{t}\right)}{\pi} + \frac{4Im\{\rho\}}{Z_0^2 td}}} \tag{5.10}$$

In the limit where $d \to 0$, $l$ also goes to zero. So scaling down the gap of a slot antenna does not directly increase current shunting, it merely changes the required length of the slot. This does not mean the slot antenna does not suffer from current shunting. Instead, as can be seen in Figure 5.5, the current shunting path is through the ohmic resistance. This means that high ohmic loss will not only decrease efficiency, but also decrease the maximum enhancement.

To explore the relationship between ohmic resistance and enhancement we can analytically derive the equations for rate enhancement and efficiency. On resonance this is a very easy task since our circuit is comprised of just three resistors, of which the spreading resistance plays no role in determining enhancement:

$$Enhancement = \frac{I_0^2\left(\frac{R_\Omega^{slot}}{R_\Omega^{slot} + R_{rad}^{slot}}\right)^2 R_{rad}^{slot}}{(q\omega)^2 R_{dip}} = \frac{\left(\frac{q\omega x_0}{d}\right)^2 \frac{3}{8\pi}\frac{Z_0}{n_A^3}\left(\frac{\lambda_0}{l_{eff}}\right)^2 \eta^2}{(q\omega)^2 \frac{2\pi}{3} Z_0 \left(\frac{x_0}{\lambda_0}\right)^2}$$

$$= \frac{9}{16\pi^2 n_A^3}\frac{\lambda_0^4}{l_{eff}^2 d^2}\eta^2 \tag{5.11}$$

Here I have used $\eta$ as the efficiency of the antenna disregarding any spreading resistance. We can think of this as the intrinsic modal efficiency of the antenna, i.e. its efficiency without regard to how it is excited. The total antenna efficiency will of course include spreading losses. To differentiate the two I will write the total antenna efficiency as $\eta_A$.

Equation (5.11) shows several very interesting features that affect rate enhancement. It is proportional to $\lambda^4$, indicating longer wavelengths can benefit the most. It is also just as sensitive to the slot length as it is the gap spacing. Since the slot length is proportional to the emission wavelength, the enhancement is proportional to $\left(\frac{\lambda_0}{d}\right)^2$ similar to the dipole antenna. Arguably the most interesting feature is that the enhancement is proportional to the efficiency squared! This highlights how important efficiency is to this device concept. Not only will a more efficient device use less power, but it will also have higher enhancement.

An unexpected outcome is that the enhancement in inversely proportional to the square of the slot length. The whole point of coupling to an antenna is to increase the effective dipole length of our emitter. So why does equation (5.11) suggest we want a short slot? This is because shorter antennas have higher Q, which as the Purcell Effect derived in equation (1.16) shows, causes

larger enhancement. Of course we need to be careful here since the efficiency also depends on length:

$$\eta = \frac{R_{ohmic}^{slot}}{R_{ohmic}^{slot} + R_{rad}^{slot}} = \left[1 + \frac{R_{rad}^{slot}}{R_{ohmic}^{slot}}\right]^{-1}$$

$$= \left\{1 + \frac{\frac{3}{8\pi n_A^3} Z_o \left(\frac{\lambda_0}{l_{eff}}\right)^2 \left[\frac{l_{eff}}{d}\left(\frac{l_{eff}}{l}\right)^4 n_g + \pi n_A\right]}{\frac{Z_0^2 t}{4 n_A Re\{\rho\}}}\right\}^{-1}$$

$$= \left\{1 + \frac{3}{2\pi n_A^4} \frac{Re\{\rho\}}{Z_0 t}\left(\frac{\lambda_0}{l_{eff}}\right)^2 \left[\frac{l_{eff}}{d}\left(\frac{l_{eff}}{l}\right)^4 n_g + \pi n_A\right]\right\}^{-1} \quad (5.12)$$

There are two main limiting regimes of the ohmic resistance: local resistance dominated or planar resistance dominated. If the antenna is in a low index medium but has a high index slot or small gap then the local resistance dominates and equation (5.12) becomes:

$$\eta = \left\{1 + \frac{3 n_g}{2\pi n_A^4} \frac{Re\{\rho\}\lambda_0^2}{Z_0 l_{eff} t d}\left(\frac{l_{eff}}{l}\right)^4\right\}^{-1} \quad (5.13)$$

If the resistivity is high, or the slot volume small, than the second term becomes the dominant term and the efficiency is proportional to the length of the slot. This is only the case when the efficiency is below 50%. In this low-efficiency regime the enhancement becomes:

$$Enhancement = \frac{9}{16\pi^2 n_A^3} \frac{\lambda_0^4}{l_{eff}^2 d^2}\left(\frac{2\pi n_A^4 Z_0 l_{eff} t d}{3 n_g Re\{\rho\}\lambda_0^2}\right)^2 = \frac{n_A^5 Z_0^2 t^2}{4 n_g^2 Re\{\rho\}^2}$$

$$= \frac{n_A^5}{4 n_g^2}\left(\frac{Z_0}{R_{sheet}}\right)^2 \quad (5.14)$$

Equation (5.14) shows that if the antenna is very inefficient than the geometry doesn't matter at all! The enhancement will simply be a function of the metal's sheet resistance.

The second regime for ohmic resistance is when the antenna is in a high-index medium with a low-index gap, at which point the planar resistance for the spreading current dominates (not to be confused with the spreading resistance!). This would be the case of a slot antenna sitting on a high-index substrate like silicon or InP. The efficiency then becomes:

$$\eta = \left\{1 + \frac{3}{2n_A^3}\frac{Re\{\rho\}}{Z_0 t}\left(\frac{\lambda_0}{l_{eff}}\right)^2\right\}^{-1} \tag{5.15}$$

Now the efficiency doesn't depend on the gap at all, merely the length of the slot and the metal thickness. If the second term dominates then the overall enhancement becomes:

$$Enhancement = \frac{9}{16\pi^2 n_A^3}\frac{\lambda_0^4}{l_{eff}^2 d^2}\left(\frac{2n_A^3 Z_0 t}{3Re\{\rho\}}\left(\frac{l_{eff}}{\lambda_0}\right)^2\right)^2$$

$$= \frac{n_A^3}{4\pi^2}\left(\frac{l_{eff}}{d}\right)^2\left(\frac{Z_0}{R_{sheet}}\right)^2 \tag{5.16}$$

In this low-efficiency regime the enhancement is still inversely proportional to the square of the sheet resistance; however we still benefit from a small gap spacing. In general we will always avoid the low-efficiency regime. That means we can follow equation (5.11). In our design we will want to create a short, narrow slot, in a low-index medium that maintains high efficiency to get the highest enhancement possible.

## 5.2 Comparing Circuit to Simulation

Now that we have developed a circuit model for the slot antenna we need to compare it to 3D time domain simulations to see if it actually predicts the correct results. The first thing to test is the validity of equation (5.11). Figure 5.6 compares the simulated enhancement for three different test cases against the predicted enhancement using equation (5.11) for slot antennas resonant at 200THz. The only free-variable was the pre-factor for $l_{eff}$. A half-wavelength dipole typically has $l_{eff} = 0.637l$ while an infinitesimal dipole has $l_{eff} = l$. The fitted parameters for the slot antenna in vacuum, with high-index gap, and in high-index surrounding were 0.74, 0.9, and 1 respectively. Using equation (5.3) we can calculate an effective length for these antennas of 0.72, 0.89, and 0.95 respectively, in good agreement with the fitted values. This confirms equation (5.3) does indeed give a good approximation of effective length.

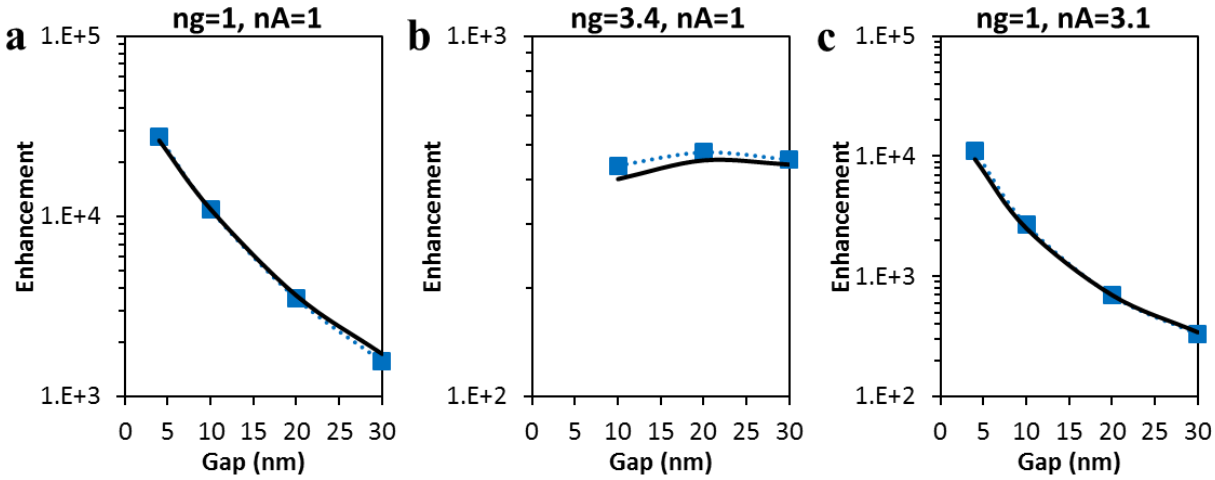


**Figure 5.6** Enhancement calculated using equation (5.11) using simulated length and efficiency for a slot antenna with (a) low-index (n=1) slot in vacuum, (b) high-index (n=3.4) slot in vacuum, and (c) low-index (n=1) slot in a high-index surrounding (n=3.1). Solid black line is from analytical solution and solid blue squares are from 3D simulation. $l_{eff}/l$ was 0.74, 0.9, and 1 respectively.

Figure 5.6 verifies that our analytical solution for enhancement does indeed hold true over a variety of different cases. The next equation to check is equation (5.12) for the antenna efficiency. Using the simulated resonance length, the subsequent efficiency was calculated using equation (5.12) and this was in turn used to calculate enhancement using equation (5.11). These results are plotted in Figure 5.7 and compared to simulated results.

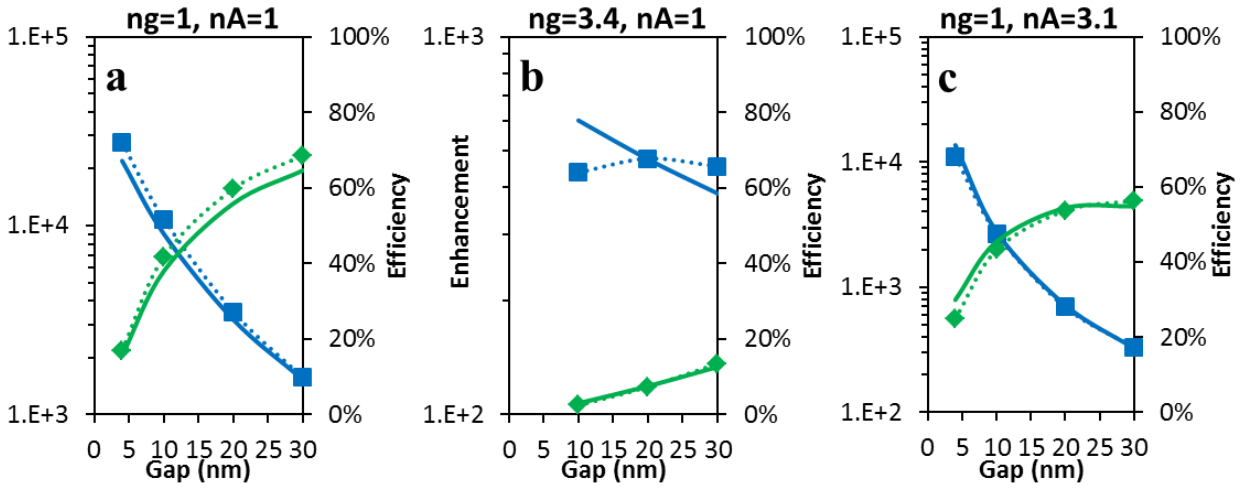


**Figure 5.7** Efficiency and enhancement calculated using equations (5.12) and (5.11) using simulated length for a slot antenna with (a) low-index (n=1) slot in vacuum, (b) high-index (n=3.4) slot in vacuum, and (c) low-index (n=1) slot in a high-index surrounding (n=3.1). Solid lines are from analytical solution and solid squares are from 3D simulation. $l_{eff}/l$ was 0.74, 0.9, and 1 respectively.

The simulated values almost exactly match the analytical solution, confirming that our equations for both radiation resistance and ohmic resistance work well. The pre-factors for the ohmic resistance were chosen to get the best fit for these curves. Other combinations of refractive index for the slot and surrounding medium were chosen and those results (not shown here) also agree very well with the analytical solution.

The final test of our circuit model and analytical component values for the slot antenna is whether or not the reactive terms predict the correct resonant slot length. Since the pre-factors for the kinetic inductance have to be the same as the ohmic resistance, correct prediction of the resonant length will help confirm the pre-factors we've chosen.

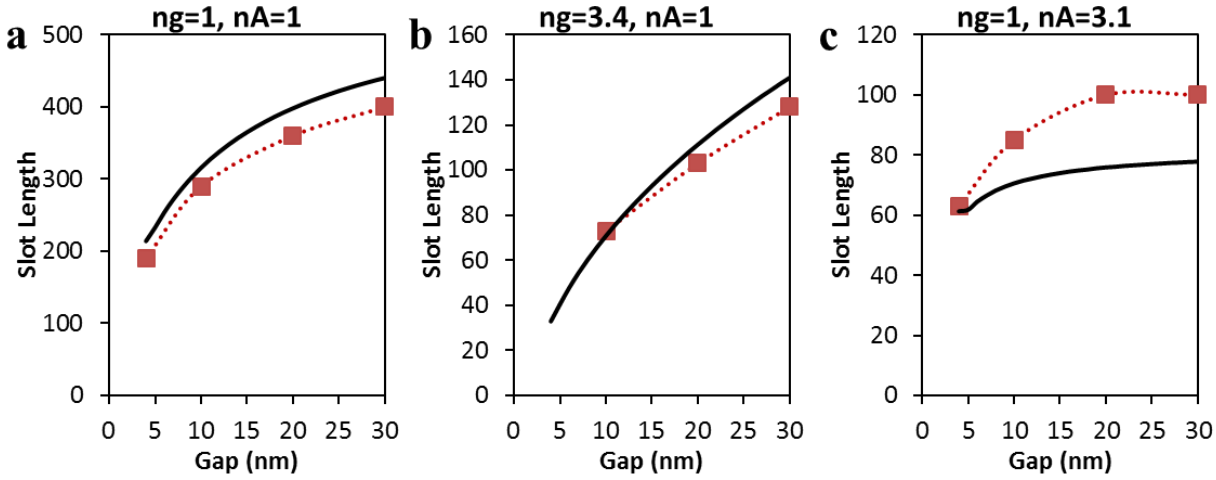


**Figure 5.8** Comparison of simulated resonant slot length with length predicted by the circuit model in Figure 5.4 for a slot antenna with (a) low-index (n=1) slot in vacuum, (b) high-index (n=3.4) slot in vacuum, and (c) low-index (n=1) slot in a high-index surrounding (n=3.1). Solid black lines are from the circuit model and solid red squares are from 3D simulation. $l_{eff}/l$ was 0.74, 0.9, and 1 respectively.

Although not as accurate as the efficiency and enhancement formulas, the predicted resonant lengths are very close to the actual values obtained with time-domain simulations. The predicted values for both the low-index antenna and high-index gap antenna show about ~10% offset from actual values. This is quite remarkable considering the complex geometry involved, but it helps confirm the validity of the circuit parameters used in Table II and III. The largest error was the antenna in a high-index medium, with deviations as large as ~25% from actual values. It should be noted that this antenna also suffered from a secondary resonance very near the fundamental resonance which made it difficult to discern precisely the resonant slot length.

At this point I would like to point out that the slot antenna with a high index gap had significantly lower efficiency and enhancement than the antenna with a low index gap; the same result we saw in the dipole antenna. Even though there is no shunting capacitor, the high-index gap caused the resonant slot length to be so short that the efficiency, and therefore enhancement,

suffers. The slot length had to be shortened to compensate for the extra capacitance caused by the high-index material. One solution would be to switch to a lower resistivity metal, such as silver, that can raise the efficiency even with a small slot volume. Another approach might be to add an inductive arch over the slot just as we did with the dipole to mitigate gap capacitance. This structure is already a well-known antenna known as the cavity-backed slot antenna and offers several advantages over the simple slot antenna that will be discussed in the following section.

## 5.3 Cavity-Backed Slot Antenna

A cavity-backed slot antenna is simply a slot antenna where the slot in the metal is only open on one side, as shown schematically in Figure 5.9(b). Blocking off one side of the slot effectively gives us a "one-sided" slot antenna. We can compare this to a normal slot antenna shown in Figure 5.9(a). When current couples into the slot antenna it flows along the surface of the metal within a skin depth, $\delta_{skin}$. If the metal thickness $t > \delta_{skin}$ than it is quite obvious the current flows in parallel on two sides of the metal (also $\delta_{skin}$ should be substituted for t in the circuit impedance values). If the $t < \delta_{skin}$ the current still flows in parallel paths on the two sides of the metal, but the two currents start to spatially overlap.

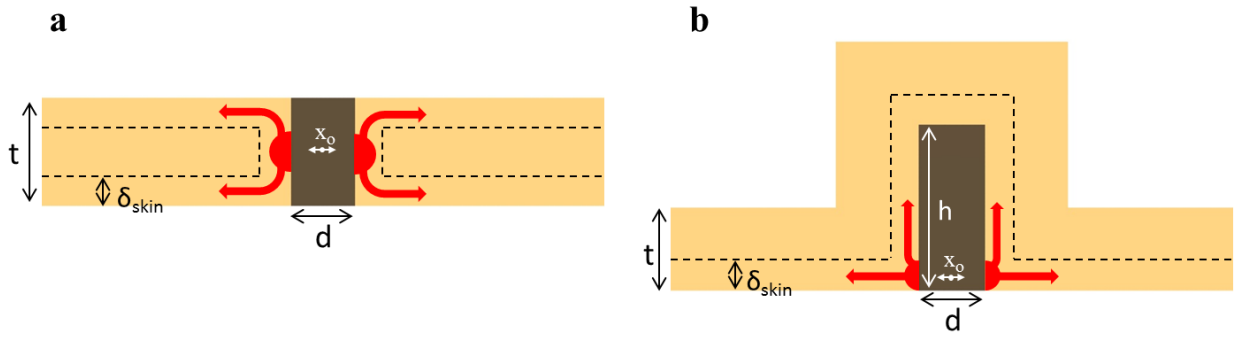


**Figure 5.9** Cross-section of (a) slot-antenna and (b) cavity-backed slot antenna showing direction of flow of current. Note that the current is confined to the skin depth of the metal.

In the cavity-backed slot antenna, the current only flows on one side of the metal. This is an important difference because the component values of this antenna will vary slightly from a normal slot antenna. Instead of the current flowing in two parallel paths, it all flows in one, so the corresponding antenna impedances will all double [22]. The exception of course is the gap capacitance and the spreading resistance since these local impedances don't change, as well as the faraday inductance which only depends on the total current in the antenna and not the specific geometry. These new component values are reflected in Table IV.

**Table IV: Circuit Components of One-Side Slot Antenna**: Transformed impedances to be used in the cavity-backed slot antenna circuit model based off of impedances from the slot antenna.

| Lumped Terms | | | |
|---|---|---|---|
| Faraday Inductance: | $Z_{Lf}^{CBslot} = Z_{Lf}^{slot}$ | Radiation Resistance: | $R_{rad}^{CBslot} = 2 \times R_{rad}^{slot}$ |
| Kinetic Capacitance: | $Z_{Ck}^{CBslot} = 2 \times Z_{Ck}^{slot}$ | Ohmic Resistance: | $R_{ohmic}^{CBslot} = 2 \times R_{ohmic}^{slot}$ |
| Planar Kinetic Capacitance: | $Z_{Ckplanar}^{CBslot} = 2 \times Z_{Ckplanar}^{slot}$ | Planar Ohmic Resistance: | $R_{planar}^{CBslot} = 2 \times R_{planar}^{slot}$ |
| Antenna Capacitance: | $Z_{Ca}^{CBslot} = 2 \times Z_{Ca}^{slot}$ | | |
| Local Terms | | | |
| Gap Capacitance: | $Z_{Cgap}^{CBslot} = Z_{Cgap}^{slot}$ | Spreading Resistance: | $R_{spread}^{CBslot} = R_{spread}^{slot}$ |

With the entries in Table IV we can now start to construct a circuit model for the cavity-backed slot antenna. We could model the cavity portion as a lumped inductance and capacitance as we did for the arch-dipole antenna. Another approach, as depicted in Figure 5.10, is to describe the cavity as a shorted transmission line stub connected in parallel with the gap impedance. This is a slightly more robust modeling method and will allow us to calculate higher order cavity modes.

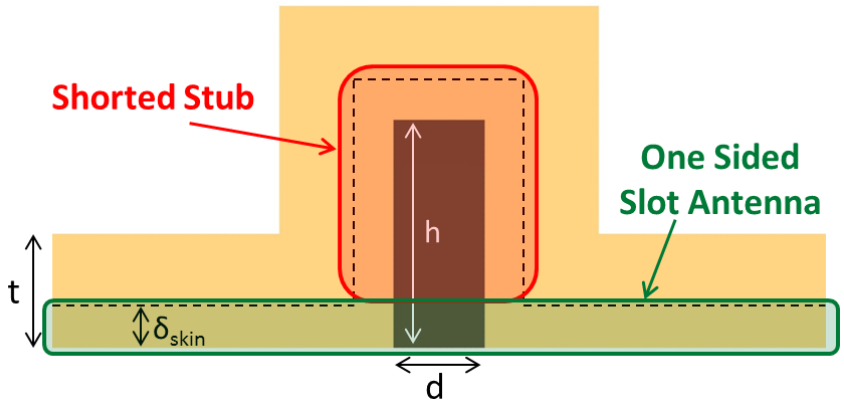


**Figure 5.10** Cross-section view of a cavity-backed slot antenna. The one-sided slot antenna portion is highlighted in green and the cavity, which can be modeled as a shorted stub, is highlighted in red.

The impedance of a shorted stub is well known [74] and given in Table V. For our cavity we can approximate it as a parallel plate transmission line and use the corresponding distributed inductance, capacitance, and resistance. Note that because the cavity is actually fully enclosed the electric field must go to zero at the ends of the long axis of the cavity. This yields the boundary condition: $sin\left(\frac{n\pi}{l+2\delta_{skin}}\right) = 0$, where the mode number n is an integer.

**Table V: Impedance of a Shorted Stub**: Impedance of the cavity modeled as a shorted stub. Note the resistance and reactance terms are given per unit length.

| | | | |
|---|---|---|---|
| Stub Impedance | $Z_{stub} = Z_{cav} * tanh[\gamma * (h - \delta_{skin})]$ | Ohmic Resistance | $R_{cav} = \frac{Re\{\rho\}}{\delta_{skin} l_{eff}}$ |
| Complex Propagation Constant | $\gamma = \alpha + i\beta$ <br> $\alpha = \frac{R_{cav}}{Z_{cav}},\ \beta = \sqrt{\kappa^2 - \left(\frac{n\pi}{l + 2\delta_{skin}}\right)^2}$ | Faraday Inductance | $L_{cavf} = 2\mu_0 \frac{d}{l_{eff}}$ |
| Cavity Wavenumber | $\kappa = \omega\sqrt{L_{cav} C_{cav}}$ | Kinetic Inductance | $L_{cavk} = \frac{Im\{\rho\}}{\delta_{skin} l_{eff}}$ |
| Characteristic Impedance | $Z_{cav} = \frac{\kappa}{\beta}\sqrt{\frac{L_{cav}}{C_{cav}}}$ | Capacitance | $C_{cav} = \frac{n_g^2 \epsilon_0 l_{eff}}{d}$ |

The impedance of the cavity is given as a single value, $Z_{stub}$, which can be both real (resistive) and imaginary (reactive). Above the cut-off frequency the stub will look inductive if $(h - \delta_{skin}) < \frac{1}{\beta}$ and capacitive if $(h - \delta_{skin}) > \frac{1}{\beta}$. Below cutoff the stub will look inductive.

Using the parameters in Table IV and Table V we can construct the cavity-backed slot antenna circuit model, shown in Figure 5.11. I have simplified the circuit, showing $Z_{slot}$ as the parallel combination of the impedances given in Table IV. The circuit is nearly identical to the slot antenna in Figure 5.4(b) except with the addition of the cavity impedance.

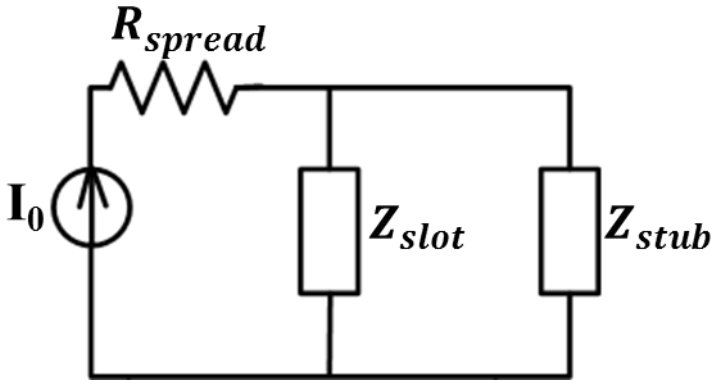


**Figure 5.11** Simplified circuit of a cavity-backed slot antenna. $Z_{stub}$ is given in Table V and $Z_{slot}$ is the parallel combination of the terms given in Table IV.

On resonance all of the reactive terms will cancel out and the rate enhancement is given by:

$$Enhancement = \frac{I_0^2 \left( \frac{R_\Omega^{slot} + Re\{Z_{stub}\}}{R_\Omega^{slot} + Re\{Z_{stub}\} + R_{rad}^{slot}} \right)^2 R_{rad}^{slot}}{(q\omega)^2 R_{dip}}$$

$$= \frac{\left(\frac{q\omega x_0}{d}\right)^2 \frac{3}{4\pi} \frac{Z_0}{n_A^3} \left(\frac{\lambda_0}{l_{eff}}\right)^2 \eta^2}{(q\omega)^2 \frac{2\pi}{3} Z_0 \left(\frac{x_0}{\lambda_0}\right)^2} = \frac{9}{8\pi^2 n_A^3} \frac{\lambda_0^4}{l_{eff}^2 d^2} \eta^2 \qquad (5.17)$$

This is nearly identical to the enhancement of a slot antenna given in (5.11) except the larger radiation resistance of the cavity-backed slot antenna makes the pre-factor twice as large. This means that a cavity-backed slot antenna with the same dimensions and efficiency as a slot antenna will actually provide twice the rate enhancement. As long as the cavity doesn't provide much loss, this is a fairly easy way to get larger enhancements. The added impedance of the cavity also allows for a much larger range of slot lengths to be engineered.

To demonstrate the accuracy of this model, parameter sweeps were done of several different cavity-backed slot antenna structures. The simplest structure, shown in Figure 5.12(d), consists of a gold antenna in vacuum with a cavity filled with vacuum. The gap was kept at a constant 30nm and the length and height of the slot were swept from 300nm to 800nm and 200nm to 800nm respectively. The corresponding resonance frequency, enhancement, and efficiency is shown in Figure 5.12.

At short slot lengths the resonance is nearly independent of cavity height. While at longer slot lengths the cavity height has the largest effect on resonance. The highest enhancement is found somewhere between these two regimes. Note that the curves appear jagged due to a limited number of data points (a total of 156). The reason for the shape of these curves, however, is not immediately apparent. We can turn to the circuit model to get a better understanding of what is causing the performance we are seeing.

The resonance length of the single-side slot antenna for this geometry operating at 200THz is ~650nm. By making the slot smaller the antenna becomes capacitive. The cavity is inductive below cutoff and at short heights which when put in parallel with the antenna impedance brings the antenna into resonance. Below cutoff the inductance of the cavity doesn't change much with height (since it is an evanescent mode) so the length mainly determines the resonance. Above cutoff the inductance of the cavity becomes a strong function of cavity height, so the resonance is mainly set by the cavity height.

The efficiency at 200THz, plotted in Figure 5.12(c) is highest when the antenna is on resonance because the reactive impedance of the antenna is very large, reducing shunt pathways. Above the cutoff length energy can flow into the cavity and so higher cavities provide more loss (at least

until the second mode comes in). Below the cutoff length energy cannot flow very deeply into the cavity and the major loss mechanism comes from spreading resistance. This is because at short slot lengths current is shunted through the reactive components and does not see the radiation resistance.

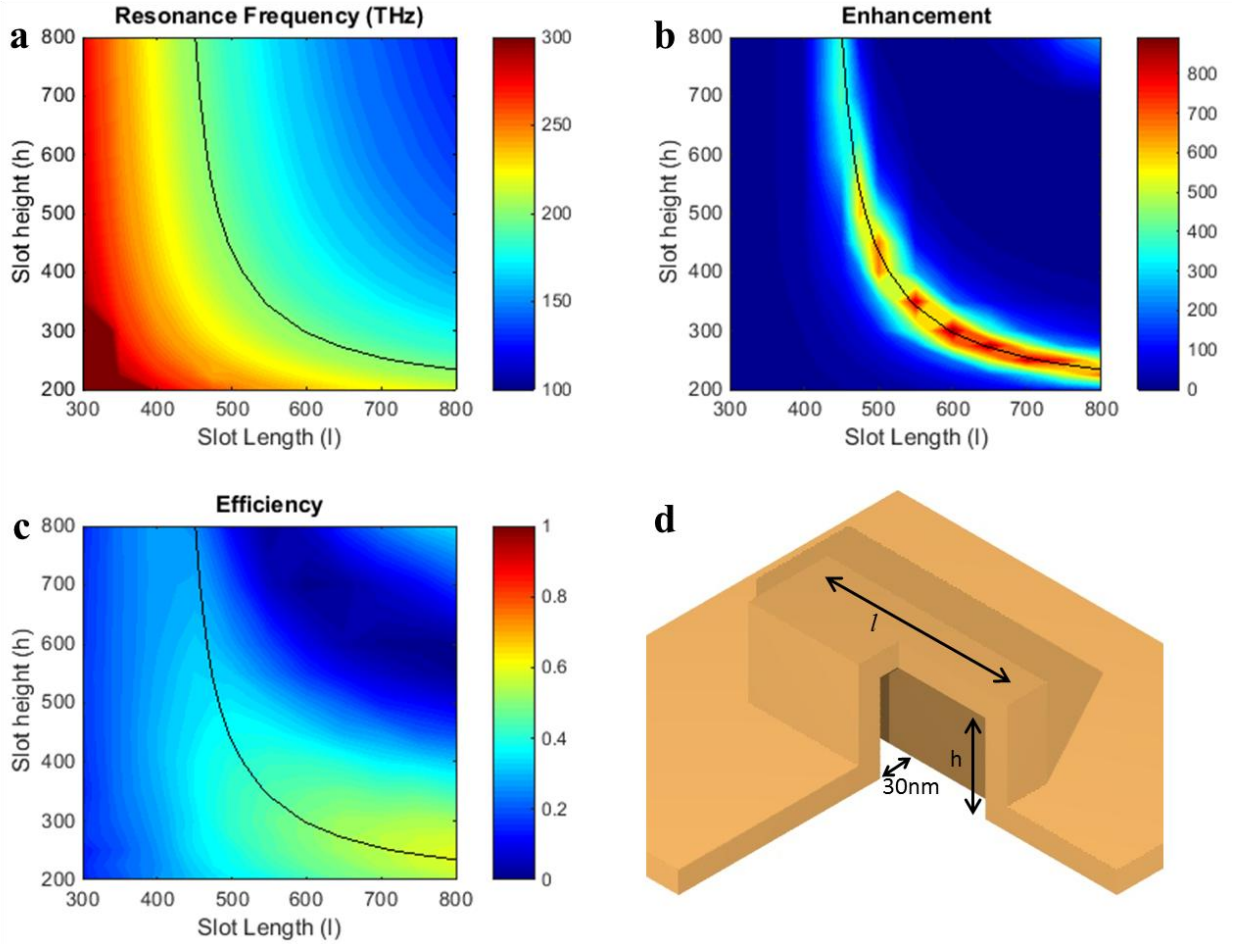


**Figure 5.12** Simulated values for a cavity-backed slot antenna in vacuum with vacuum filled slot. Simulated (a) resonance frequency, (b) enhancement at 200THz, and (c) efficiency at 200Thz as a function of slot length and height. (d) Schematic drawing of structure with a 30nm gap.

The circuit model gives very good agreement with simulated values, as can be seen by comparing the circuit results in Figure 5.13 with the simulated results in Figure 5.12. The circuit slightly underestimates the maximum enhancement, ~700x versus the actual ~850x. This is not because equation (5.17) is incorrect, but because the circuit doesn't perfectly estimate the resonance frequency and efficiency. Regardless, the maximum error between circuit and simulation is on the order of ~10% for the resonance frequency and efficiency.

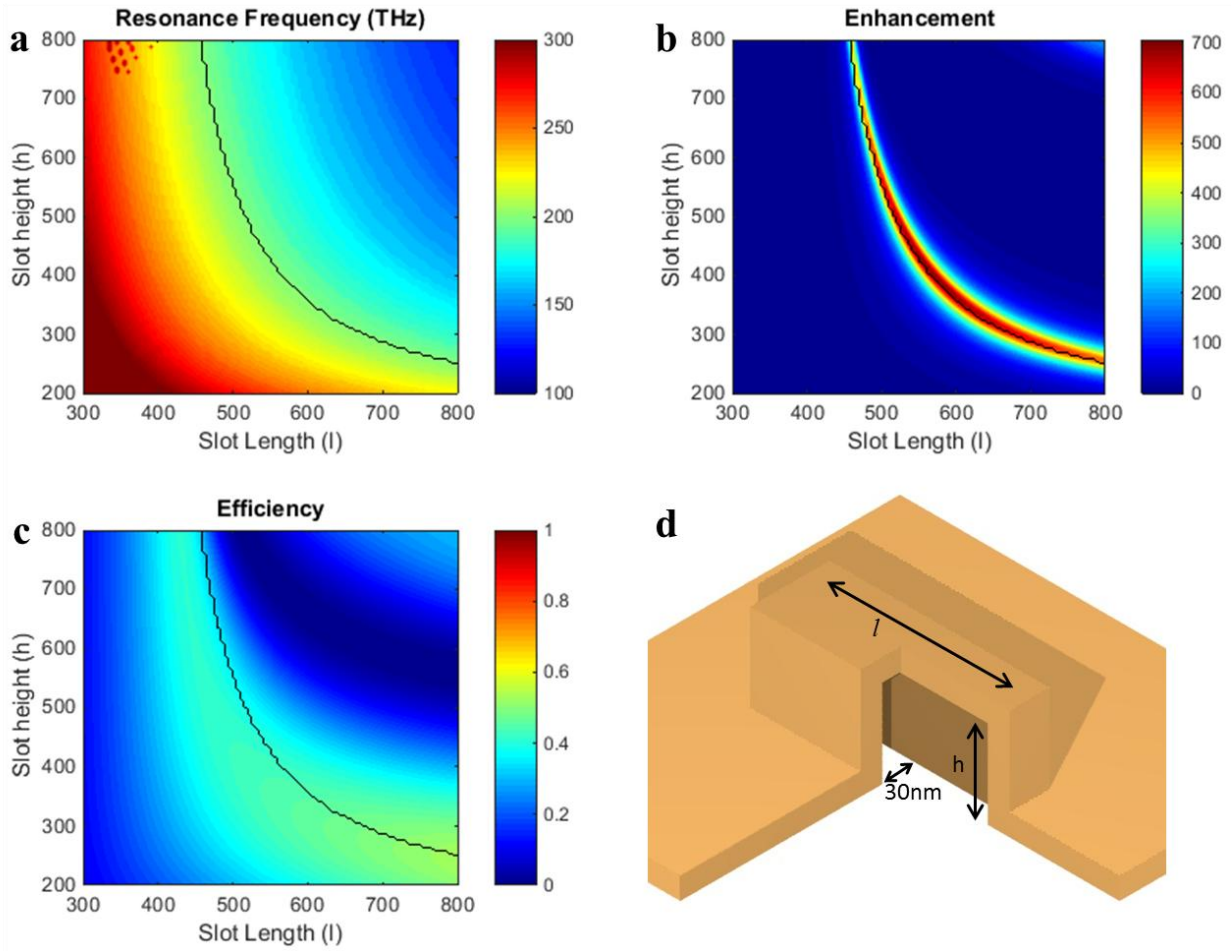


**Figure 5.13** Circuit model derived values for a cavity-backed slot antenna in vacuum with vacuum filled slot. Predicted (a) resonance frequency, (b) enhancement at 200THz, and (c) efficiency at 200Thz as a function of slot length and height. (d) Schematic drawing of structure with a 30nm gap.

Next, a gold antenna was simulated that is sitting on InP (n = 3.17) with a cavity filled with InGaAsP (n = 3.55). In addition, there is a 5nm conformal $Al_2O_3$ (n = 1.63) coating on the inside of the cavity. This structure is modeled off the electrically injected structure from [75]. To model the complicated gap structure, a simple substitution was made for all the circuit parameters to model the effective electrical gap width: $d = d_{InGaAsP} + d_{Al2O3}\frac{n_{InGaAsP}^2}{n_{Al2O3}^2}$. The low index of the $Al_2O_3$ causes a severe penalty to the gap, making it electrically much larger than its physical size. The simulated results for this structure are given in Figure 5.14. The shape of these curves are very similar to the in-vacuum case, however the large index of the InP surroundings significantly lowers the enhancement to a maximum of ~50x. The efficiency, however, is substantially larger.

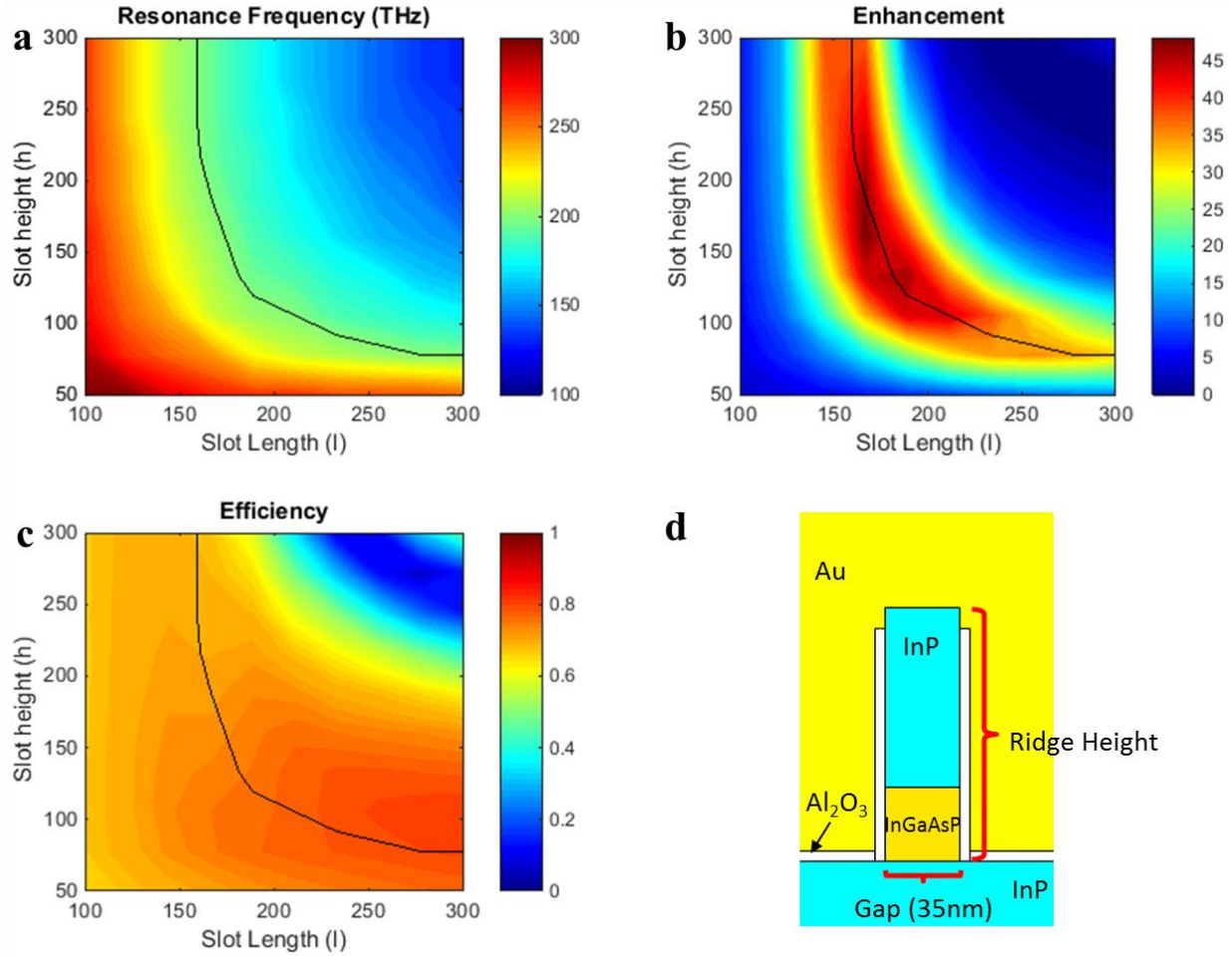


**Figure 5.14** Simulated values for a cavity-backed slot antenna on InP with InGaAsP filled slot. Simulated (a) resonance frequency, (b) enhancement at 200THz, and (c) efficiency at 200Thz as a function of slot length and height. (d) Schematic drawing of the structure with a 35nm gap and a 5nm coating of $Al_2O_3$. Data reproduced from [76].

The circuit model again gives very similar results to the simulated value. While there is some difference in the exact value of efficiency and enhancement, the error again is very small.

With the development of the circuit model, an arbitrary cavity-backed slot antenna can now be designed very rapidly without the need for heavy reliance on simulations. Most importantly, the circuit model can be used to understand major effects geometry and refractive indices have on antenna performance. The cavity-backed slot antenna will become a very important structure in the next chapter where it will be used to achieve record rate enhancements in 2D semiconductors.

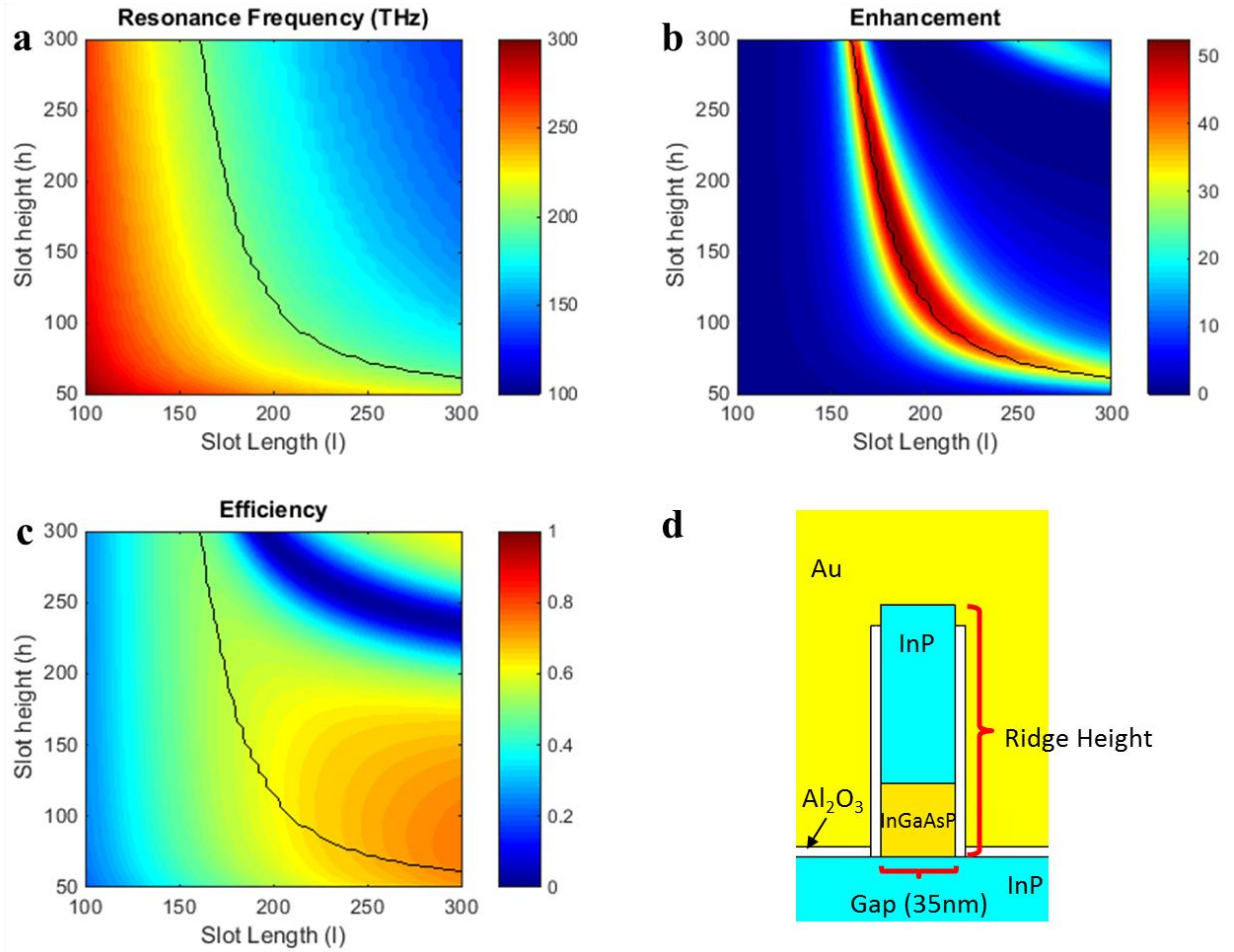


**Figure 5.15** Circuit model derived values for a cavity-backed slot antenna on InP with InGaAsP filled slot. Predicted (a) resonance frequency, (b) enhancement at 200THz, and (c) efficiency at 200Thz as a function of slot length and height. (d) Schematic drawing of the structure with a 35nm gap and a 5nm coating of $Al_2O_3$.

# Chapter 6

# Achieving Higher Enhancement and High Quantum Yield

At this point I have demonstrated that optical antennas can greatly increase the rate of spontaneous emission from semiconductor materials while maintaining high efficiency, even when coupled into low-loss dielectric waveguides. However, for every experimental demonstration up to this point, the overall quantum yield of the devices has been very low (< 1%). The final step towards a demonstration of an efficient nano-emitter is demonstrating that not only can the antenna be efficient, but the overall quantum yield can be high as well.

In this chapter I will discuss the problem of low quantum yield from nano-structured semiconductors. First, I will discuss best case and practical case scenarios for III-V semiconductors and then turn my attention towards a new class of semiconductors, transition metal dichalcogenides (TMDCs), which show promise for both high efficiency and nanoscale dimensions. I will then develop the theory needed to use a new type of optical antenna, the cavity-backed slot antenna. Not only does this antenna allow for large enhancement rates and easy fabrication, but it also couples well to the inherently planar structure of TMDCs. Using a TMDC, $WSe_2$, I then show rate enhancements up to 380x with overall quantum yield comparable to unprocessed material.

## 6.1 Quantum Yield at the Nanoscale

The reason nano-structures suffer from such low quantum yield is because traditional semiconductors suffer from very poor surfaces. When a semiconductor is exposed to air the crystal structure obviously has to terminate. At these termination points the crystal is no longer perfect, and for III-V materials these lattice defects cause very fast surface recombination. As the semiconductor is scaled down the surface to volume ratio goes up, meaning more and more surfaces.

To derive the rate of surface recombination, we can think of the carriers in our material moving about near the edge of the material at the thermal velocity, $v_{thermal}$. The thermal velocity is the

speed corresponding the the kinetic energy of an electron associated with its thermal energy. At room temperature (~300K), the thermal energy is $kT$ = 25.7 meV. The thermal velocity is therefore:

$$\frac{1}{2}mv_{thermal}^2 = kT$$
$$v_{thermal} = \sqrt{\frac{2kT}{m}} \approx 10^7\ cm/s \tag{6.1}$$

As the carriers move towards the surface they can either hit a defect and recombine or elastically scatter back into the material. We can then assign a fraction to how many of the carriers will hit a defect, $\Gamma$. If every carrier hits a defect then $\Gamma$ = 1. If all of the carriers reflect back then $\Gamma$ = 0. This fraction can be calculated if the capture cross-section and density of defects is known. Since these defect must sit within the bandgap in order to assist in recombination (since higher energy states wouldn't ever be favorable), then whether or not they will be filled already, and therefore can't capture another carrier, is determined by the Fermi level at the surface. The kinetics of this can get very complicated and the exact number depends on the interface quality and carrier concentration. For simplicity we will assume that $\Gamma$ is constant for any particular material system we are using. I need to stress, however, that while this is good enough for a simple approximation, this assumption is not actually true!

Using $\Gamma$, we get a value that is usually reported for materials, the surface recombination velocity:

$$v_{surface} = \Gamma * v_{thermal} \tag{6.2}$$

Using this number we can then quickly and easily compare different materials on how well their surfaces behave. GaAs is a very common material using in lasers. It also has notoriously bad surfaces, with $v_{surface} = 10^6$ cm/s [77]. Another common semiconductor is silicon, has one of the best surfaces of any known material at $v_{surface} = 10^4$ cm/s when cladded with native oxide or as low as ~1 cm/s for a properly annealed thermal oxide! Unfortunately, silicon doesn't have a direct bandgap so is a pretty lousy light emitter.

Once we have the surface recombination velocity, we can quickly calculate the surface recombination lifetime of a carrier in the material. If the layer of semiconductor has a thickness of $W$ and diffusion coefficient $D$, then a good approximation of the surface lifetime is [78]:

$$\tau_{surface} = \frac{W}{2v_{surface}} + \frac{1}{D}\left(\frac{W}{\pi}\right)^2 \tag{6.3}$$

For very thin dimensions the first term dominates and we see that the surface recombination rate is directly proportional to the surface velocity. The surface recombination rate is also commonly written as [79]:

$$\frac{1}{\tau_{surface}} = \frac{A}{V} v_{surface} \tag{6.4}$$

Where $A$ is the total surface area and $V$ is the total volume. Equations (6.4) and (6.3) yield the same result in thin layers where $A = 2L^2$ and $V = W*L^2$.

We can now look at the case of a semiconductor sandwiched in the feedgap of an optical antenna; the band diagram of which is shown in Figure 6.1. Because metal has no bandgap, i.e. there is a continuum of energy levels, carriers in the semiconductor can freely flow into the metal. Therefore in this case we have our worst case scenario, $\Gamma$=1.

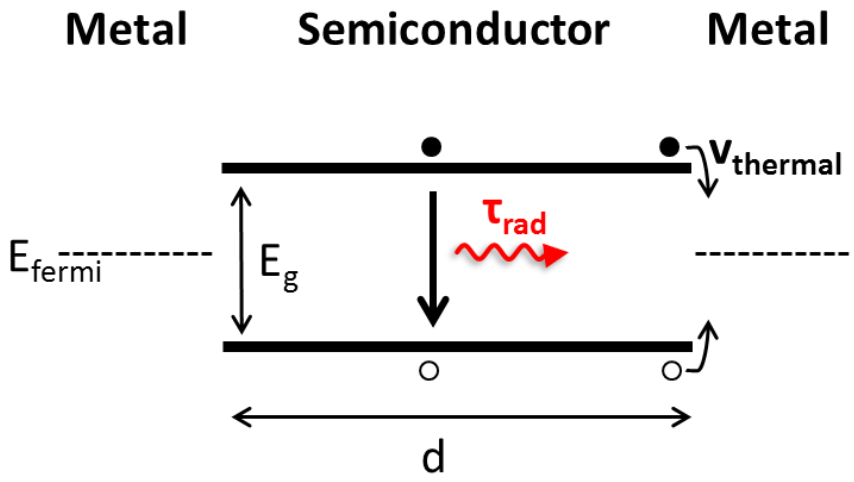


**Figure 6.1** Band diagram of a semiconductor emitter sandwiched between two pieces of metal.

If this semiconductor is nanoscale in dimensions, it will obviously have a very low quantum yield. Let's say it has the right dimensions to fit in the feed-gap of our dipole antenna from Figure 2.9. For simplicity we'll make it a cylinder $d$ = 20nm long and a radius of 20nm. Assuming an external quantum efficiency of one (i.e. no light trapping), the total quantum efficiency is just the internal quantum efficiency (iQE). Therefore, with a fast radiative lifetime $\tau_{rad}$ = 1ns:

$$QE = \frac{\frac{1}{\tau_{rad}}}{\frac{1}{\tau_{rad}} + \frac{1}{\tau_{surface}}} = \frac{\frac{1}{\tau_{rad}}}{\frac{1}{\tau_{rad}} + \frac{A}{V} v_{surface}} = \frac{\frac{1}{1ns}}{\frac{1}{1ns} + \frac{4\pi(20nm)^2}{2\pi(20nm)^3} 10^7 cm/s} \quad (6.5)$$

$$= 10^{-4}$$

As a general rule of thumb, if it's easiest to write the quantum yield in scientific notation, it's too low. When we put this material in the feedgap of an optical antenna this number will go back up. With rate enhancement the new quantum efficiency becomes:

$$QE = \frac{\frac{F}{\tau_{rad}}}{\frac{F}{\tau_{rad}} + \frac{A}{V} v_{surface} + \frac{F}{\tau_{rad}} \left( \frac{1}{\eta_{antenna}} - 1 \right)} \quad (6.6)$$

Where the third term in the denominator comes from ohmic losses in the antenna. This is plotted in Figure 6.2 by using the antenna efficiency and enhancement from Figure 2.9.

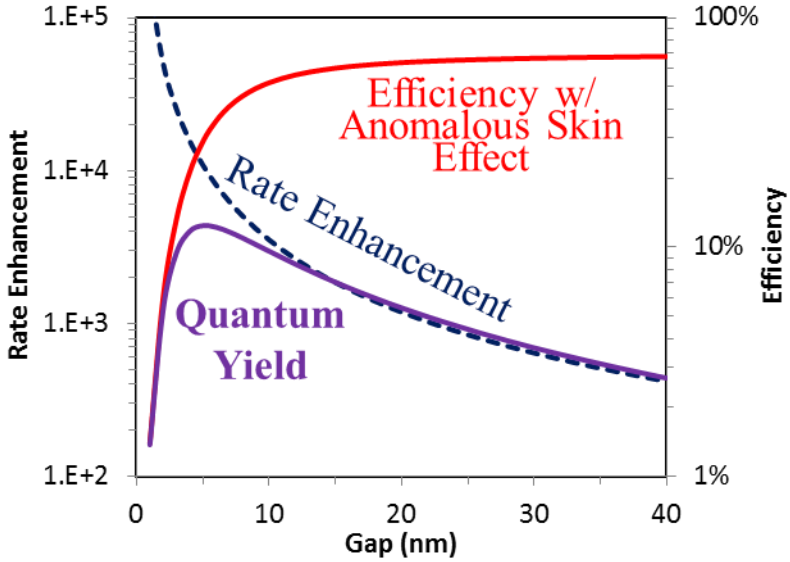


**Figure 6.2** Quantum yield (purple) of an emitter placed in the feedgap of the antenna shown in Figure 2.9. Also replotted here is rate enhancement (blue dotted line) and antenna efficiency (red).

Note that the efficiency is plotted on a logarithmic scale. For large gap spacing the quantum yield linearly follows the rate enhancement. However, as the gap shrinks down the surface recombination stops being the dominant source and instead the poor efficiency from high spreading resistance dominates. Thus there is a peak quantum yield of ~10% at a gap spacing of about 7nm. Remember that this is our best case scenario, so if this device were actually made it probably wouldn't fare to well. This points out the obvious problem with coupling

semiconductors to metal: the quantum yield is always very low. This doesn't bode well for our fast and *efficient* nanoemitter.

The solution to this problem is to use some sort of cladding material between our semiconductor and the antenna to prevent loss of carriers. The modified band structure for this configuration is shown in Figure 6.3.

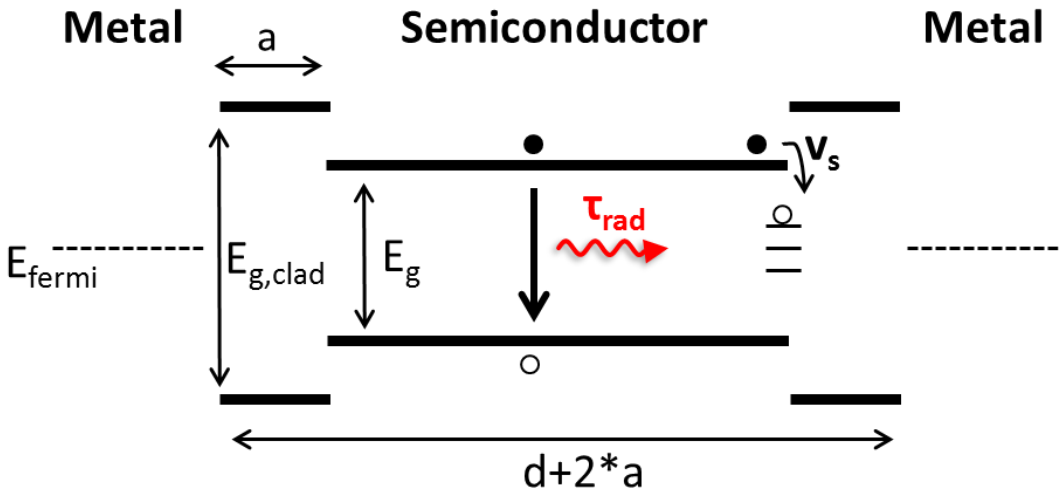


**Figure 6.3** Band diagram of cladded semiconductor sandwiched between two pieces of metal.

By using a material with a larger bandgap, we can prevent recombination of carriers into the metal surface. This is why the $TiO_2$ layer was used in our experiments with the InGaAsP. However, if the interface between the cladding material and semiconductor is not good then we'll still have fast surface recombination at that interface. For $TiO_2$ on InGaAsP the surface velocity is ~$10^5$ cm/s, which although two orders of magnitude better than the worst case scenario, is still pretty bad. Ideally we'd want to use something like InP, which has an interfacial surface velocity of ~50 cm/s with InGaAsP. However, the barrier in InP is not very high so if we don't make it thick enough carriers can just tunnel through or jump over the barrier into the metal. If we make the cladding thicker, however, we decrease our rate enhancement because we make our gap bigger. In the end we are left with a tricky problem where we have to balance material properties with antenna design to actually engineer an efficient structure. Because of the complicated nanoscale geometry required, such a device will most likely be very difficult to actually fabricate. This is also a problem with nanolasers which has been partially overcome by advanced lithography, etching, and regrowth techniques [8].

## 6.2 Transition Metal Dichalcogenides

Another solution to achieve high quantum yield in nanoscale emitters is to use an emitting material that has very low or no defects on the surface. Any three dimension crystal will have

surface defects when made nanoscale because of dangling bonds at the surface. However, a two dimensional material, such as graphene [80], may only have dangling bonds in two dimensions. Even smaller dimensional materials could have even better surfaces at the nanoscale (such as dye molecules which are essentially zero dimensional and can have near unity quantum yield). A two dimensional material may suffice for our case since we only need strong confinement in a single direction (the feedgap) to get good coupling to an antenna.

Transition metal dichalcogenide (TMDC) materials such as $MoS_2$ and $WSe_2$ have emerged as an interesting new optoelectronic material [81–83]. TMDCs, like graphene, consist of 2D layers bonded together with Van Der Waals forces, as shown in Figure 6.4. Monolayers of many types of TMDCs exhibit semiconductor properties with a direct bandgap and near ideal surfaces due to the lack of dangling bonds at their top and bottom surfaces. This results in many of these materials having high intrinsic photoluminescence quantum-yield [81,84]. While the best material currently available has quantum yield on the order of ~1%, this appears to be limited by defects in the crystal structure. As this material system matures and material quality improves, this number will most likely increase. Theoretically the only limiting point on the crystal is the edges which do contain dangling bonds.

The unique properties of TMDCs have allowed for a wide range of optoelectronic devices to be demonstrated including LEDs [85–87], solar cells [88], and photo-detectors [89]. Their strong light emission, near-ideal surfaces, and intrinsically atomic layer thickness make TMDCs very attractive candidates for use in an optical nanoemitter.

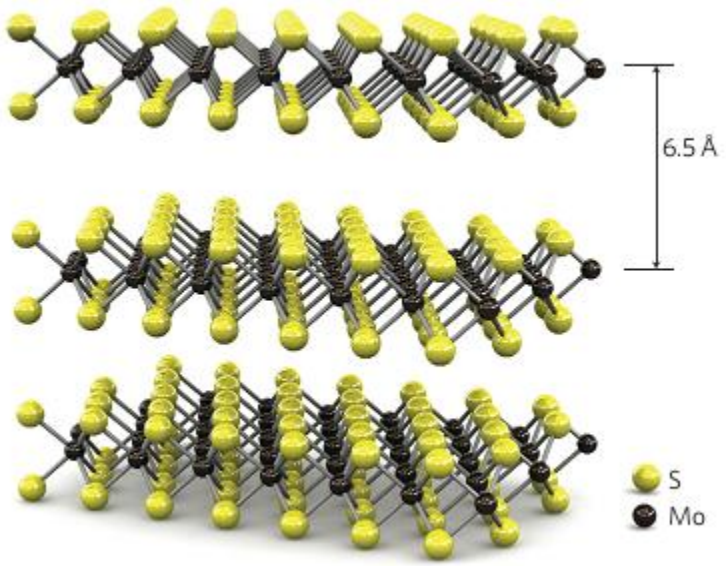


**Figure 6.4** Layer structure of $MoS_2$, reproduced from [90]. The interlayer thickness is ~6.5Å; bulk structures held together by Van Der Waals forces.

The first TMDC to be extensively studied was $MoS_2$ [81]. The quantum yield as a function of number of layers is plotted in Figure 6.5(a), with monolayers exhibiting an iQE of ~0.5%. Since then, materials such as $WS_2$ and $WSe_2$ have shown even brighter emission (Figure 6.5(b)). The exact reason for the brighter emission is still not understood, but regardless such materials offer even higher intrinsic quantum yield than $MoS_2$.

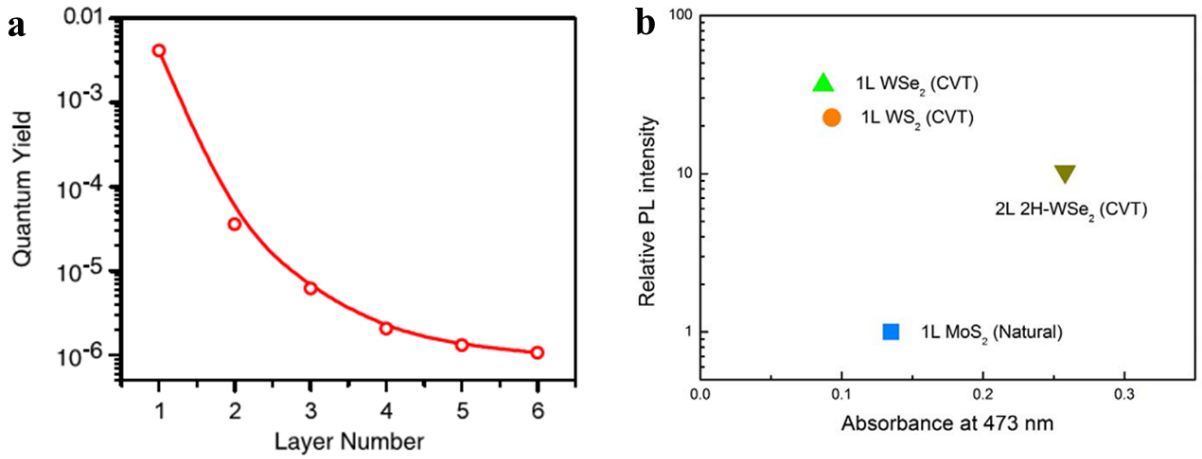


**Figure 6.5** (a) Quantum yield of $MoS_2$ as a function of the number of layers, reproduced from [81]. (b) Relative PL intensity from $WSe_2$, $WS_2$, and $MoS_2$ monolayers, reproduced from [84].

Carrier lifetimes in these materials seem to vary significantly. Reports on $MoS_2$ place the non-radiative trap-dominated recombination rate at ~3ps [91] and for $WSe_2$ in lifetimes have been reported in the range of 18ps [92] to 36ps [93]. These numbers, coupled with reports of quantum-yield, suggest radiative lifetimes on the order of ~1ns.

For our purposes, a material that emits at longer wavelength is much easier to work with both because enhancement is proportional to $\lambda^2$ and because lower frequency antennas suffer less kinetic inductance and therefore are less lossy. Of the materials shown in Figure 6.5(a), $WSe_2$ is both the brightest and has the longest wavelength of emission, ~750nm. For these reasons I have chosen to use this material as our initial candidate for a TMDC based nanoLED.

## 6.3 $WSe_2$ Coupled Cavity-Backed Slot Antenna

In this section we will explore the use of a cavity-backed slot antenna for use as an optical antenna for enhancing emission from $WSe_2$. This structure offers many advantages including: directing light into a single direction which allows for higher collection efficiency, simpler fabrication process than previous antenna structures, enables the use of a low-index gap material, and couples well to a planar emitter like $WSe_2$.

### 6.3.1 Fabrication Process

In this section I will discuss the final device design and fabrication for the $WSe_2$ coupled to a cavity-backed slot antenna. For our $WSe_2$ based device, I have used silver for the antenna instead of gold like in my previous experimental work. Silver offers both lower kinetic inductance and resistivity in the near-IR (750nm) compared to gold. For this reason, it can be used to create slot antennas with higher efficiency and therefore higher enhancement. Experimental work has also shown that gold starts to fluoresce in the visible, which can be enhanced with an antenna and provides an unfortunate background signal.

A simplified schematic of the final structure is shown Figure 6.6(a). It consists of a silicon oxide slot embedded in a plane of silver, with a nanostrip of $WSe_2$ at the very top of the slot. An SEM of a final slot antenna array is shown in Figure 6.6(b). Slots are patterned with a 700nm pitch with lengths varying from 200nm to 350nm.

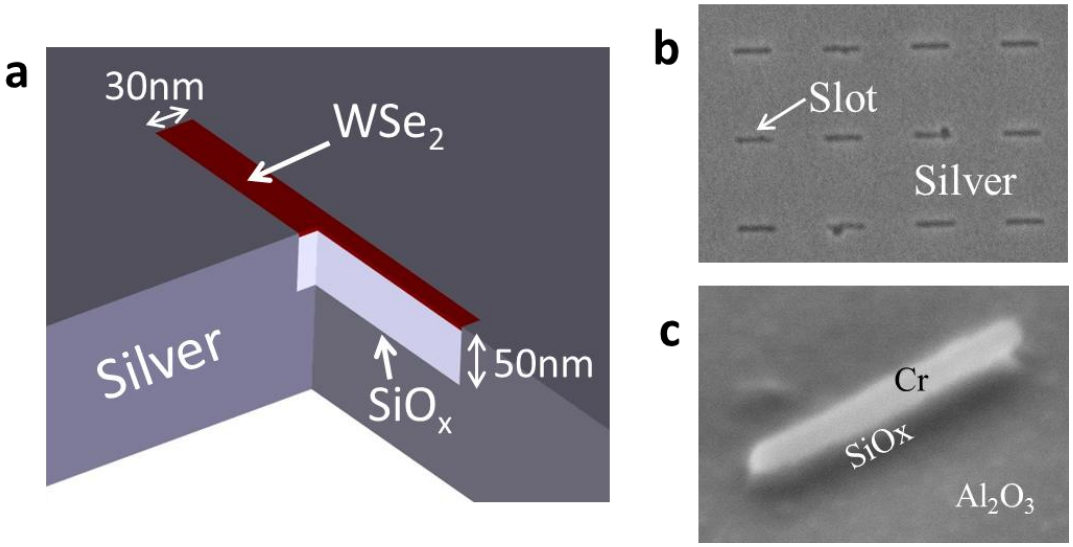


**Figure 6.6** (a) Perspective cut-through schematic of WSe2 monolayer coupled to an optical slot antenna. (b) Top down SEM of fabricated devices showing slots in silver plane. (c) Perspective SEM of oxide slot before metal deposition.

The fabrication process for this structure is outlined in Figure 6.7. First, $WSe_2$ is exfoliated on a silicon wafer that has been coated with a 270nm thick $SiO_2$ layer. This thickness gives high contrast of exfoliated $WSe_2$ flakes so the number of layers can be determined simply by its color [94]. Monolayers are checked with PL to verify they are monolayers and then are transferred to an InP carrier using the method described in [95]. The InP carrier consists of a bulk InP substrate with a ~150nm MOCVD grown InGaAs etch stop layer on top. Prior to transfer, a 20nm thick layer of $Al_2O_3$ is deposited using ALD on top of the InGaAs layer. The $Al_2O_3$ layer provides contrast so that the monolayer is still visible after transfer, acts as an etch stop during etching, and protects the silver antenna from tarnishing.

After the flake is transferred to the InP carrier it is coated in 40nm of $SiO_x$ using ebeam evaporation. Evaporated oxide is typically very low in quality and therefore is most likely not stoichiometric. A thin chrome strip, ~30nm wide and ~250nm long is patterned on the oxide using PMMA A2 and ebeam lithography. The chrome is then deposited with ebeam evaporation and lifted off in acetone. The target thickness of the chrome is 7nm, though AFM measurements suggest it is closer to 10nm.

The chrome acts as a hardmask during a 45sec $SF_6$ RIE (reactive ion etch) of the $SiO_x$. The etch completely removes all the oxide and the $WSe_2$ not protected by the chrome. This creates a 40nm tall ridge of $SiO_x$ with the thin layer of chrome on top. $SF_6$ has a very high etch selectivity over $Al_2O_3$, so it acts as an effective etch stop layer. The etch is precisely timed to prevent over-etching that may cause footings under the oxide ridge. Timings are calibrated with a dummy sample etch before every device etch.

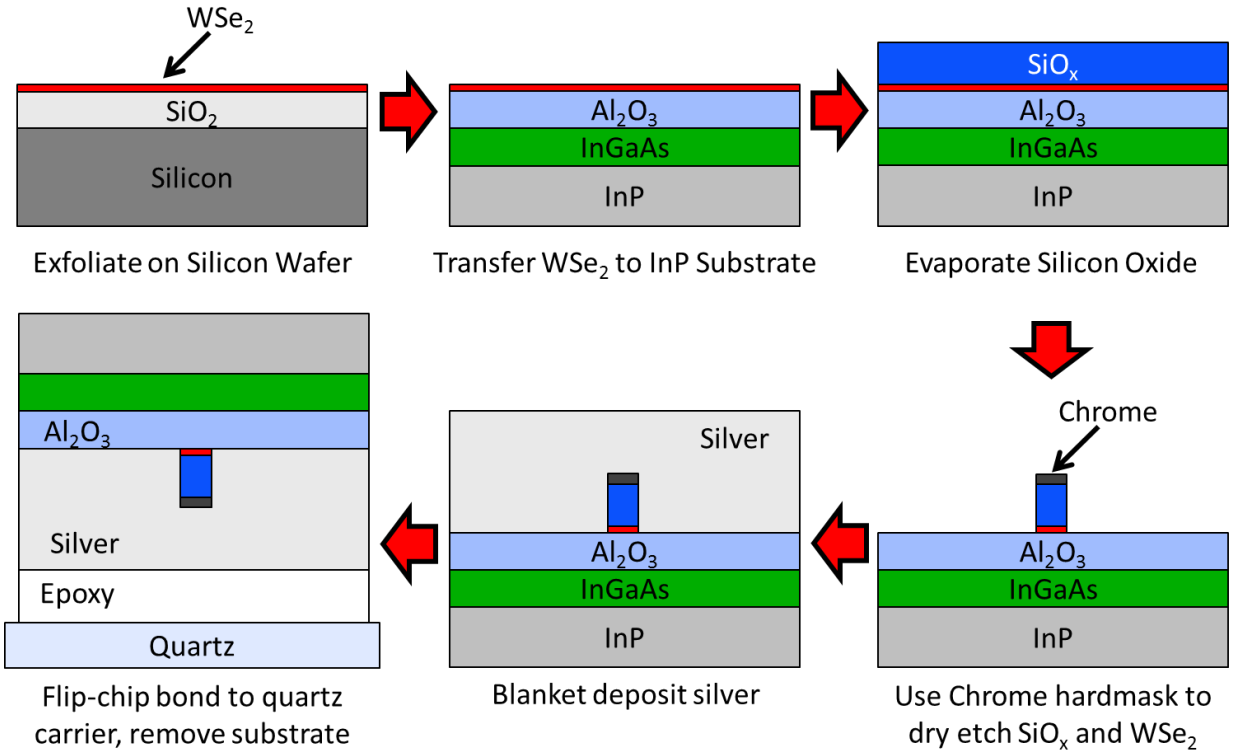


**Figure 6.7** Fabrication process for a $WSe_2$ monolayer coupled to a cavity-backed slot antenna.

At this stage the $WSe_2$ has been etched to its final dimensions, exactly the same size as the slot of the antenna. For antenna-coupled structures the entire chip is then coated with a blanket evaporation of silver which encapsulates the $SiO_x$ and forms the slot antenna. The silver is coated in a 20nm layer of titanium to protect the silver as well as promote adhesion with the epoxy. The metal evaporation step is skipped for bare samples.

Finally, the entire chip is bonded to a quartz carrier with NOA 81 UV epoxy. The chip is left to cure for 12 hours at 50°C to harden the epoxy. Finally the InP substrate is removed by first lapping with 1000 grit sandpaper down to ~75µm followed by a wet etch in 1:1 HCL:Phosphoric Acid at 80°C. This etchant stops on the InGaAs layer with is removed with a subsequent 60 second etch in 1:1:10 $H_2SO_4$:$H_2O_2$:$H_2O$. This etchant stops on the $Al_2O_3$ layer, which is left intact to protect the silver below.

### 6.3.2 Simulated Enhancement

Full-wave time domain simulations (CST Microwave Studio) show that a dipole emitter in the $WSe_2$ layer at the very center of the slot experiences a peak emission rate enhancement into the far-field of 420x with an antenna efficiency of 54%. As shown in Figure 6.8, this is not the point of highest enhancement; as the dipole gets closer to the side of the slot the enhancement increases slightly. Due to boundary conditions, however, the electric field has to fall to zero at the ±y ends of the slot, decreasing coupling at these outer edges. The center of the slot, therefore, gives a fairly accurate representation of the spatially averaged rate enhancement in the flake.

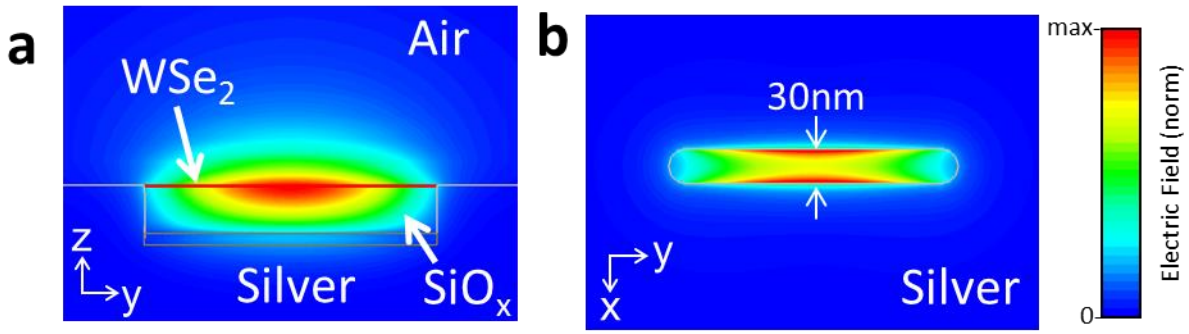


**Figure 6.8** Electric-field profile of the cavity-backed slot antenna fundamental mode for the (a) cross-section along the long axis of the slot and (b) cross-section along the center of $WSe_2$ in the xy plane.

The exact peak enhancement and resonance frequency is highly dependent on the exact geometry of the slot and the exact dielectric constants of its surroundings. Figure 6.9 depicts the simulated enhancement for a 30nm wide, 40nm tall, and 250nm long slot as the slot height (a), width (b), and length (c) are varied.

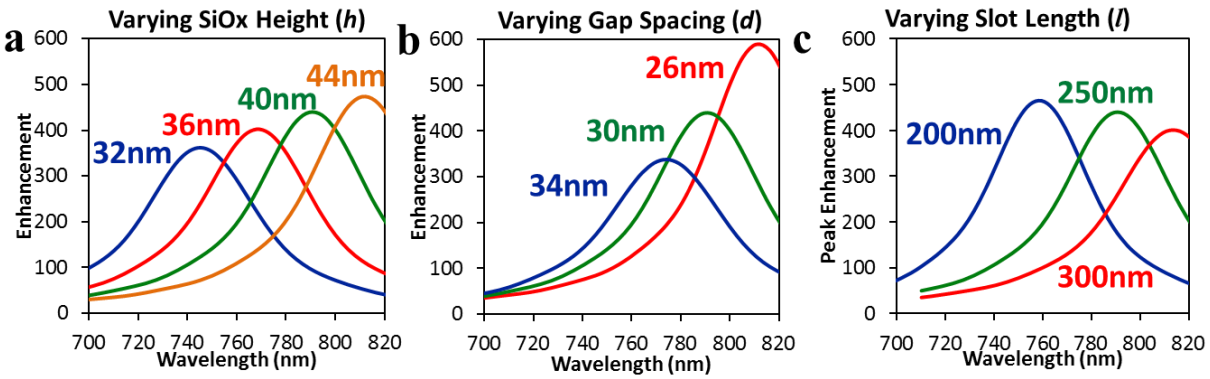


**Figure 6.9** Simulated enhancement of a silver cavity-backed slot antenna with a $SiO_x$ slot 30nm wide, 40nm tall, and 250nm long. There is also 10nm of chrome on top of the slot. Plots show enhancement with varying (a) $SiO_x$ height, (b) gap spacing, and (c) slot length.

The slot height and width and height have a similar effect on the antenna resonance, causing a resonant wavelength shift of -4.7 nm/nm and +5.5 nm/nm respectively. The gap, not surprisingly, has a much larger effect on enhancement. Changing the gap causes a -6.5% /nm change in enhancement versus 2.2% /nm for the height. The length has the smallest effect on either resonance or enhancement, causing a +0.55 nm/nm and -0.2% / nm change respectively.

This demonstrates how just a few nanometer change in any dimension can cause a large effect on the final antenna resonance. Process variation will inevitably causes defects and small changes from slot to slot in fabricated devices which can cause the resonance of individual antennas in an array to vary significantly from their ideally identical neighbor.

### 6.3.3 Optical Measurement Results

CW and time-resolved PL measurements were performed to experimentally measure the rate enhancement provided by the slot antennas. CW measurements were performed by generating carriers in the $WSe_2$ with a 532nm continuous-wave laser polarized parallel to the slot with a μ-PL microscope system. Optical emission was collected with a 0.8NA 100x objective, passed through a spectrometer and readout on a Princeton Instruments liquid nitrogen cooled silicon CCD. Collected optical emission from three different antenna-coupled devices and a bare reference sample are shown in Figure 6.10. The off-resonant antenna (orange) is 20x brighter than the bare sample, the near-resonant antenna (green) is 190x brighter, and the on-resonant antenna device is 700x brighter (blue). When the difference in pumping and collection efficiency is taken into account, the brightest intensity antenna has a quantum yield comparable to an unprocessed monolayer.

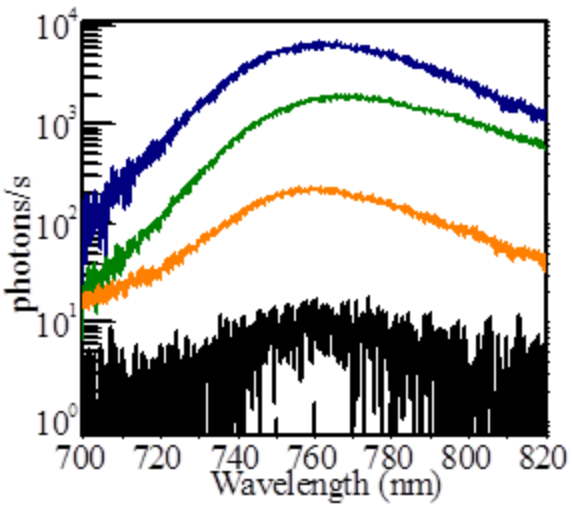


**Figure 6.10** Raw number of photons per second collected for an on-resonance (blue), near-resonance (green), and off-resonance (orange) antenna-coupled device compared to a bare (black) etched $WSe_2$ monolayer.

To determine rate enhancement from CW PL measurements, the emitting material's recombination must be non-radiative rate limited, equal pumping conditions must occur in both the antenna coupled and bare case, and the far-field emission pattern of both bare and antenna-coupled structures must be taken into account. Light emitted per unit area from the bare etched sample is two orders of magnitude weaker than the brightest bare unprocessed flakes. This signifies that the edges exposed during the etch process significantly increase the non-radiative recombination rates. Since the same etching process is used for all samples, this pins the total recombination rate of all the devices at the same edge-recombination limited rate. Edge passivation or cladding could potentially increase the quantum yield of future devices, but for this study the edges are left unpassivated to allow for CW measurement of the radiative rate enhancement.

To confirm the devices were indeed limited by the same edge-recombination lifetime, time resolved measurements were taken by pumping the devices with a 635nm Ti:Sapphire pulse and measuring their photon decay using a Hamamatsu streak camera with a 1ps time resolution. An unprocessed flake sitting on a 270nm $SiO_2$ layer on a Si substrate was fit with a bi-exponential decay (inset Figure 6.11) with lifetimes of 10ps and 85ps for the short and long component respectively. I have measured a range of quantum yield from our monolayer flakes varying by 5x between the weakest and brightest flakes measured. Lower quantum yield flakes most likely suffer from a faster non-radiative recombination rate. The unprocessed flake measured here exhibited an eQE ~4x lower than the brightest flake, suggestion the non-radiative lifetime in our $WSe_2$ flakes is as high as 40ps. Lifetime reports from literature vary, with a fast exponential decay between 18 and 36ps reported at room-temperature [92,93], in good agreement with our result.

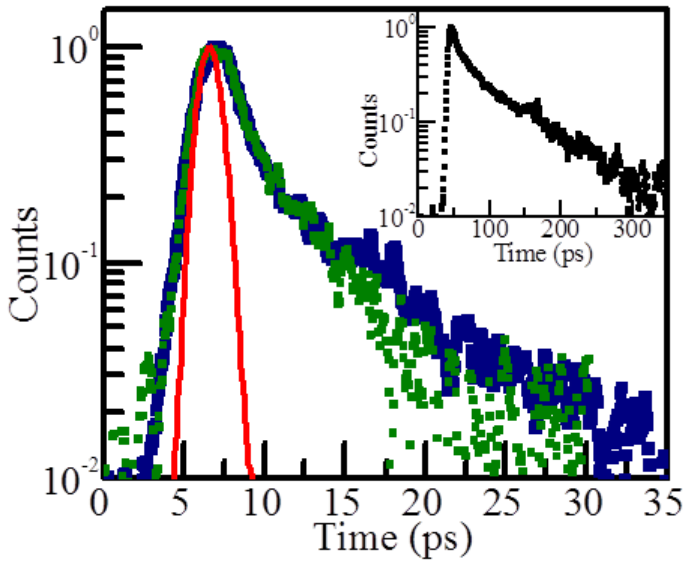


**Figure 6.11** Time resolved PL of an unprocessed flake (black, inset), an on-resonance antenna (blue), and a near-resonance antenna (green). The instrument response function is shown in red.

The decay curve for both the on-resonance (blue) and near-resonance (green) devices are shown in Figure 6.11. The curves were fit by convolving the instrument response with a bi-exponential decay. Both have very similar decay constants, with time constants of 1ps and 7ps for on-resonance and 1ps and 5ps for near-resonance. The on-resonance antenna, however, emits almost 4x more light. The near identical decay constants of both devices in the presence of significantly different intensity of emitted light confirm that the non-radiative rates are in fact dominant.

Next, the effect of the silver antenna on pumping conditions was simulated using 3D time domain simulations (CST Microwave Studio). Because the laser pump is polarized perpendicular to the antenna resonant mode, very little enhancement is observed. Simulations predict a 1.6x pump enhancement for antenna coupled devices versus bare etched flakes. Because pumping is far off resonance, as confirmed with scattering measurements in Figure 6.13, this enhancement is minimally effected by minor changes in the slot height and length.

The difference in collection efficiency between the antenna-coupled and bare devices must also be taken into account. Simulations of bare etched samples give a collection efficiency into a NA=0.8 objective of 17.2%. The collection efficiency of the antenna-coupled device was simulated as 48% due to its high directionality into the +z direction. The antenna therefore provides a 2.8x increase in collection efficiency.

Finally, only dipoles perpendicular to the slot long-axis will couple to the antenna. Parallel oriented dipoles show no enhancement in simulations because they are orthogonal to the current in the antenna mode. The bare sample has no polarization dependence, suggesting there is no intrinsic polarization in the flake itself. The light emitted from the antennas is therefore only

coming from half the total dipoles in the material. Combining this factor of 1/2 with the increased pumping and better collection efficiency gives a total adjustment factor of 2.2x. The total increased light emission from the antenna-coupled devices therefore corresponds to an average rate increase of 317x, 86x, and 9.4x for the on-resonance, near-resonance, and off-resonance antennas respectively.

The shape of the antenna resonances can be clearly seen by dividing the adjusted light emission from the antenna coupled devices by the light emitted by a bare sample, as shown in Figure 6.12(a). A clear peak at 780nm can be seen in the on-resonant PL enhancement curve (blue), while the near-resonant antenna appears to be resonant at a wavelength slightly longer than 825nm. The peak enhancement of the antenna-coupled devices is slightly higher than the averaged emission enhancement due to the spectral mismatch between the antenna resonance and the exciton wavelength of $WSe_2$. The peak enhancement observed for the on-resonance case is 380x, in good agreement with the 420x predicted with simulation.

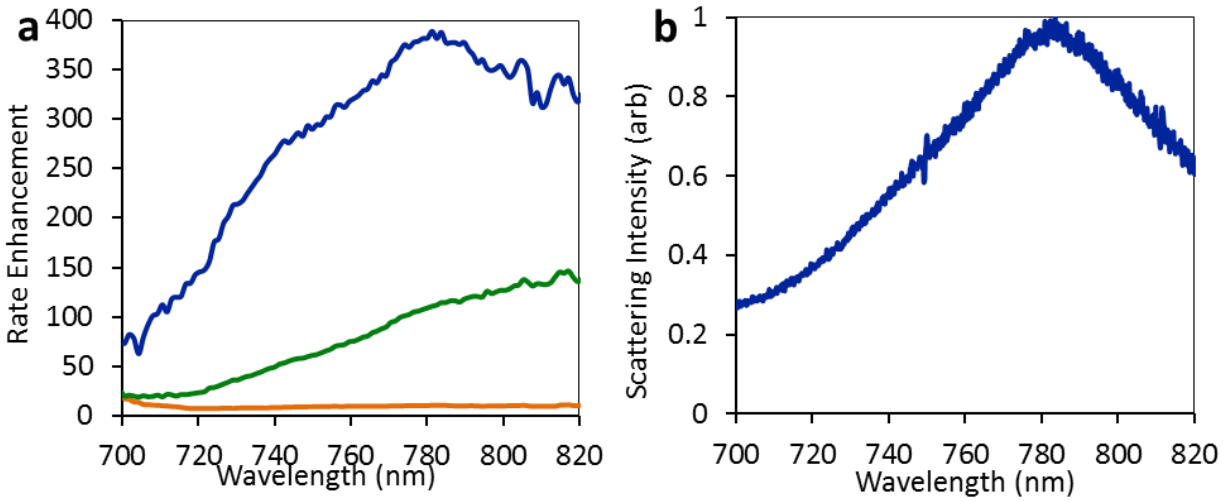


**Figure 6.12** (a) Rate enhancement measured for antenna-coupled devices. (b) Dark-field antenna-polarized scattering intensity from antenna-coupled devices in the wavelength range of PL emission.

To characterize the response of each individual antenna, dark-field scattering measurements were performed on each of the antennas measured in Figure 6.10. Figure 6.12(b) shows the scattering intensity of light polarized in the antenna polarization. Almost exact agreement of the resonance frequency and shape for the on-resonant antenna between both scattering intensity in Figure 6.12(a) and rate enhancement in Figure 6.12(b) can be seen. The intensity of scattered light polarized parallel to the slot is shown in Figure 6.13(b), confirming that the antenna is far from resonant at the pump wavelength of 532nm.

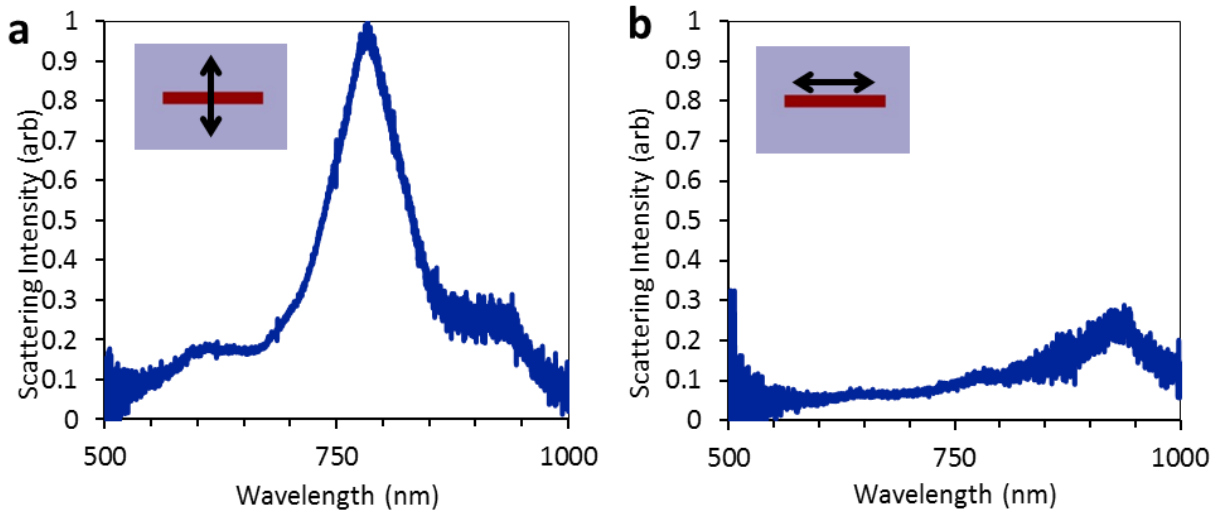


**Figure 6.13** Dark-field scattering intensity for the (a) antenna-polarized and (b) slot-polarized direction from an on-resonance antenna-coupled device for the entire optical spectrum down to 500nm.

Ideally, light emitted from the slot antennas should be polarized perpendicular to the slot by a ratio as high as the rate enhancement. Unfortunately, imperfections in the slot and silver surface create scattering centers that can scatter light into the orthogonal polarization direction. Figure 6.13 shows the scattering intensity from the on-resonant antenna. The ratio of the intensities of the two polarizations on resonance is ~8x, which is just slightly higher than the PL ratio of ~6x. This result is plotted along-side the ratios of the near-resonance and off-resonance antennas in Figure 6.14. This plot shows that that maximum PL ratio is always slightly less than the scattering intensity, suggesting that the low PL ratios observed are indeed a consequence of poorly polarized antennas.

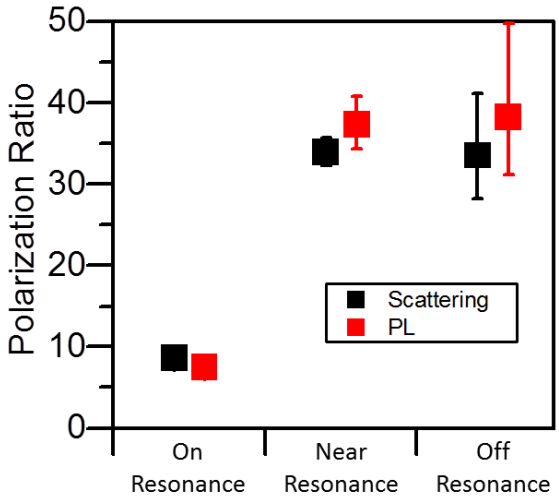


**Figure 6.14** Polarization ratio of both peak PL (red) and scattering intensity (black) for antenna-coupled devices.

The antenna with the largest polarization ratio was the near-resonance device. The total PL emitted from this device as a function of polarizer angle is shown in Figure 6.15(b). There was some drift in the measurement as can see by the number of counts being higher at 180° than at 0°. A cosine squared fit line shows the polarization ratio of this device is ~55x. This is very near the average rate-enhancement measured for this device (89x), suggesting a fairly low concentration of depolarizing defects in the array.

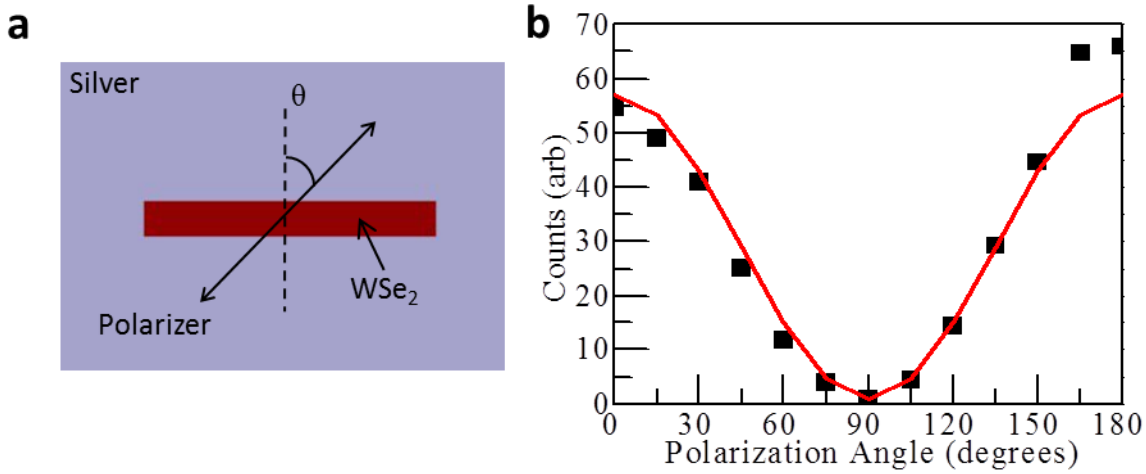


**Figure 6.15** (a) Schematic top-view of slot antenna showing the slot and angle of the polarizer. (b) Total number of counts collected from an antenna coupled device as a function of polarizer angle.

The cause of the poorly polarized antennas can be further investigated by looking at the SEM and scattering measurements of whole antenna arrays. Figure 6.16(a) shows an SEM of the etched slots in the on-resonant antenna array. While many of the slots look near ideal, there are many slots with large defects either on or next to them. The good slots show a high contrast between the two scattering polarizations (Figure 6.16(b) versus Figure 6.16(c)), while the defects are bright in both cases. The presence of these many defects is most likely what causes poor polarization of the antennas since the current in the antenna mode is able to spread out in the metal plane surrounding the slot and interact with these defects.

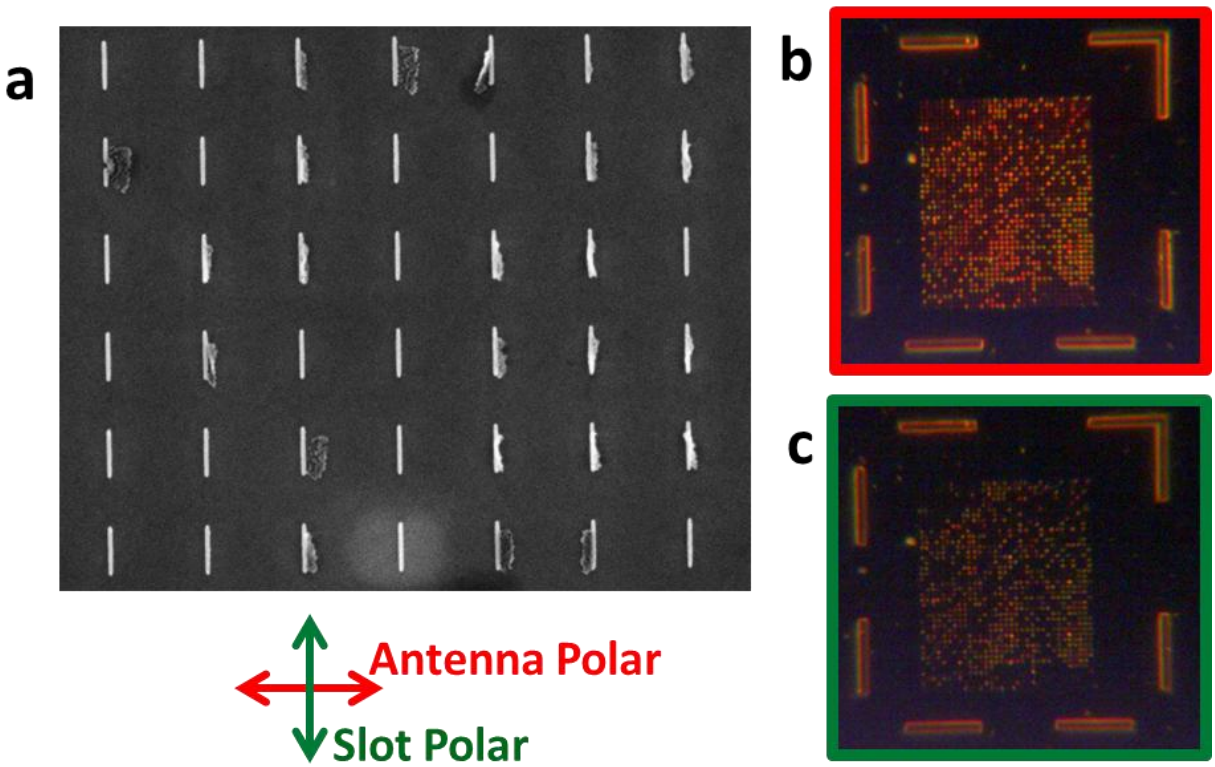


**Figure 6.16** (a) SEM of a test pattern on the on-resonance antenna sample showing the etched ridges. Dark field scattering of the on-resonant antenna chip of light polarized (b) perpendicular to the slot (antenna polar) and (c) parallel to the slot (slot polar).

Figure 6.17(a) shows the SEM of the etched slots in the near-resonant antenna array. This process yielded very ideal looking slots across the chip, but left a slightly roughened surface (~nanometers). The scattering from this array shows very good contrast between the antenna-polarization (Figure 6.17(b)) and the slot-polarization (Figure 6.17(c)). This sample showed a much higher polarization ratio than the on-resonant case, suggesting the large ridge defects cause the most significant de-polarization. The presence of the non-ideally smooth surface most likely causes the less than ideal polarization measured.

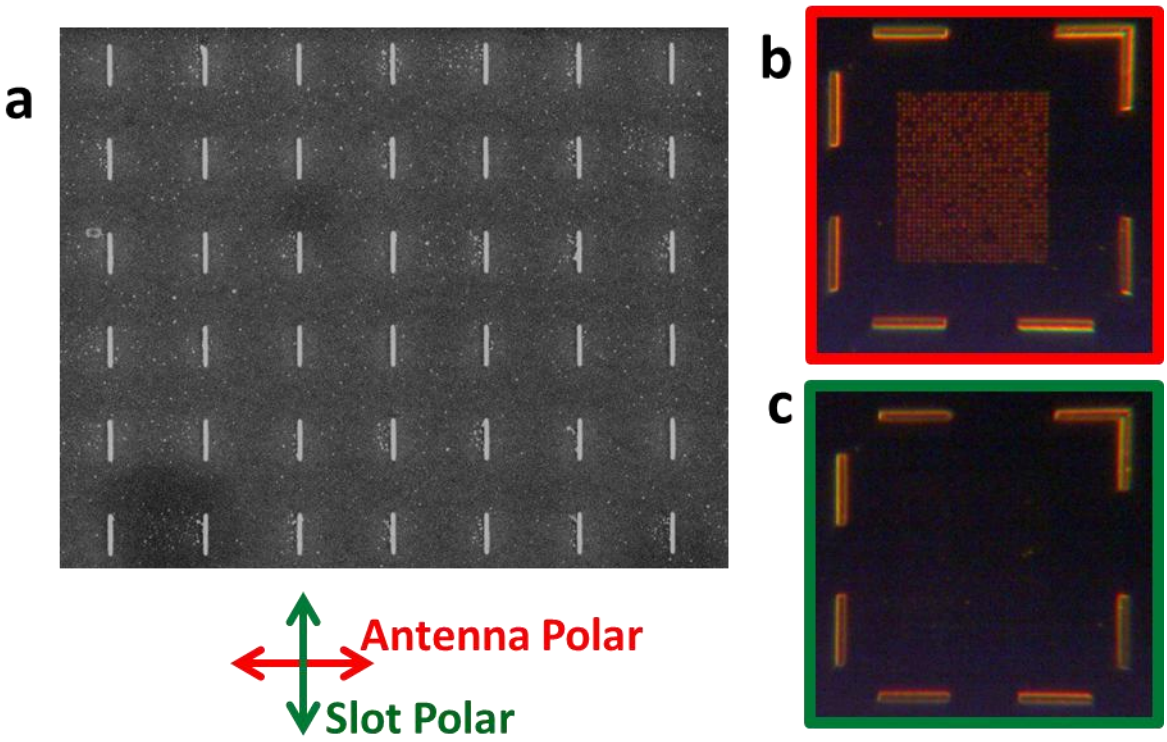


**Figure 6.17** (a) SEM of a test pattern on the near-resonance antenna sample showing the etched ridges. Dark field scattering of the near-resonant antenna chip of light polarized (b) perpendicular to the slot (antenna polar) and (c) parallel to the slot (slot polar).

# Chapter 7

# Conclusions

In this dissertation I began by looking at the energy requirements for future on-chip communication schemes. I showed that we are currently in a regime where the metal interconnect layers dissipate the majority of on-chip power and limit the maximum clock-rate in CPUs. The outlook for future systems is grim if a solution beyond metal wires cannot be found. While optical interconnects using a laser-based light source may seem attractive initially, I found that lasers are much too large and power hungry to be used as an energy-efficient source. LEDs, however, can be truly nanoscale and still be highly efficient. I showed that if the rate of spontaneous emission can be increased by a large enough factor, nanoLEDs would be an ideal candidate for an on-chip source that could reduce CPU power consumption by orders of magnitude.

Metal optics, particularly optical antennas, were then presented as a way of greatly increasing the rate of emission from semiconductor sources. While antennas have been used for over a hundred years in the microwave regime, optical antennas only began to be used in the last decade. Since then there has been widespread misunderstanding of optical antennas and the role of plasmonics in electromagnetic focusing and emission enhancement. I showed that plasmonics is merely a result of electrons carrying kinetic energy and therefore adding additional inductance (kinetic inductance) to metal structures. Unfortunately, because kinetic inductance and ohmic resistance share the same geometric pre-factor, a highly-plasmonic device will also be very lossy. Because of this, metal optics has often been branded as too inefficient and lossy to be useful.

In an effort to design of a proper antenna that provides both large rate enhancements and maintains high efficiency, I developed a simple circuit model based off of device geometry. I showed that the resonance frequency, enhancement, and efficiency of a dipole antenna can be accurately modeled using this lumped element model. The circuit model also gives intuitive insight into device design and the upper limit of rate enhancement. I showed that if the gap impedance can be properly engineered, rate enhancements up to ~3,000x can be achieved while maintaining an antenna efficiency >50%.

With the aid of the circuit model, I presented a new type of dipole antenna, the arch-dipole. I showed that introducing an inductive arch across the feedgap can drastically reduce the shunting effects of large gap capacitance. I then demonstrated a free-standing arch-dipole antenna that

enhanced the rate of spontaneous emission from an InGaAsP nanoridge by 115x, while maintaining an antenna efficiency of 66%.

InGaAsP is an attractive gain material because it is very well understood and highly utilized in long-haul optical interconnects. I demonstrated how an InGaAsP based arch-dipole antenna can be efficiently coupled into an InP waveguide to distribute light around a chip. Theoretical coupling efficiencies as high as 79% were shown with a simple antenna on a $\lambda/4$ thick InP waveguide design. Using a multi-mode waveguide, I demonstrated coupling efficiencies as high as 70% and direction emission with a front to back ratio of 3:1.

While the dipole antenna, particularly the arch-dipole antenna, is a very useful and efficient antenna structure for large rate enhancements, it is difficult to couple to 2D materials. To solve this problem, I develop a circuit theory for the optical slot antenna, the electromagnetic dual of the dipole antenna. I found that the rate enhancement provided by the slot antenna is proportional to its efficiency squared. It also does not suffer as much from large gap capacitance as the dipole antenna.

With the demonstration of large rate enhancements, efficient antenna designs, and efficient waveguide coupling, the last step was to show that a nanoscale emitter can also have a high quantum yield. I discussed the problem with large surface recombination rates of III-V emitters, specifically how difficult it is to passivate dangling bonds in a 3D semiconductor crystal. As an alternative, I presented $WSe_2$, a 2D transition metal dichalcogenide that not only has a direct bandgap but has near ideal surfaces even at the nanoscale. These ideal surfaces allow monolayer $WSe_2$ to have high intrinsic quantum yield on the order of ~1%. Using a cavity-backed slot antenna, I then demonstrated that a peak spontaneous emission rate enhancement of 380x from $WSe_2$. This large rate enhancement offset the detrimental effects of the dry etching used to form the device and resulted in a nanoLED with quantum yield comparable to unprocessed $WSe_2$.

While I have demonstrated that through proper design a metal-optics based LED can be a fast, efficient, and nanoscale emitter, they still have a long way to go before they are widely accepted as such. Primarily, an electrically injected structure needs to be developed that can be directly modulated at high speeds. New materials and surface passivation techniques may also be needed to keep device quantum yield high enough to be beneficial in an energy efficient link. Even if all of these can be achieved, whether they will ever satisfy the requirements for an on-chip link is still a very open question. Even more so, and even less explored, is what impact they may have on integrated sensing and imaging at truly nanoscale dimensions. Optical antennas offer amazing possibilities and are a field that is sure to have many exciting advances in the coming years.

# Appendices

## Appendix A. Polarization of InGaAsP Ridges

At the nanoscale, geometric dimensions play an important role in emission and absorption properties of semiconductor materials. It is important to note that the bare InGaAsP ridges are slightly polarized even in the absence of metal. PL measurements were taken of bare and antenna clad ridges to determine the relative ratio of PL intensity on emission polarization (Figure A.1). Bare ridges preferentially radiate light polarized in the y-direction compared to the x-direction by a ratio of 3.3. Previous studies on semiconductor nanowires [96] have shown that this effect can be explained by leaky wave resonances in the nanostructure that exhibit stronger fields for transverse magnetic (TM) modes than transverse electric (TE) modes. The presence of these leaky wave resonances enhances emission for TM (y-direction) polarized light over TE (x-direction) polarized light in the bare ridges.

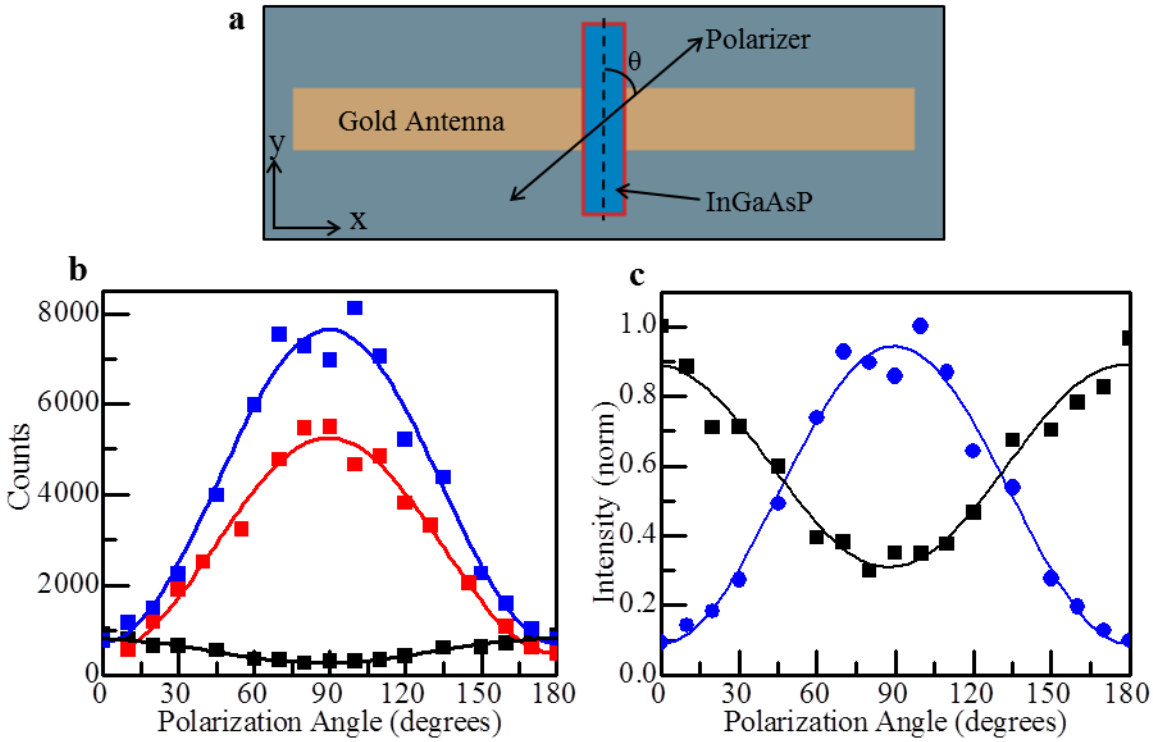


**Figure A.1.** Polarization Dependence of InGaAsP Ridges. a, Top view of antenna-coupled ridge showing angle of polarizer during measurement. b, Total number of integrated counts for bare ridge (black), 600nm antenna (red), and 400nm antenna (blue) as a function of polarizer angle.

Symbols are measured data and solid lines are a sinusoid fitting function. c, Normalized intensity of emission from antenna clad and bare ridge as a function of polarizer angle.

The exact placement of a radiating dipole within the ridge also effects its preferred polarization. Simulations show that an emitter placed near the edge of the ridge is three times more likely to be TE polarized than an emitter placed in the center of the ridge (Figure A.2), the data shown in Figure A.1 is therefore an effective average over all emitters in the ridge. Since the antennas only enhance the x-polarized dipoles in the center of the ridges (Figure 3.7(b)), they enhance the most poorly emitting dipoles.

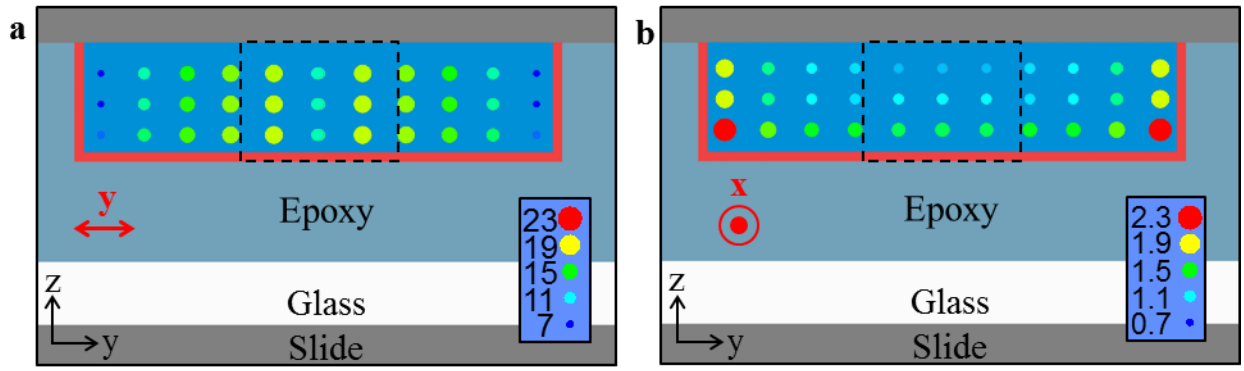


**Figure A.2.** Simulation of power radiated from a dipole emitter within a bare InGaAsP ridge. a,b, Cross-section of a ridge showing the relative power radiated from a dipole emitter placed at different locations with the ridge polarized in the y-direction, x-direction. Red outline of the InGaAsP ridge represents the 3nm $TiO_2$ coating. Black dotted line indicates where the antenna arms would be in if the ridge was antenna-coupled.

# Appendix B. Simulated Collection Efficiency of Waveguide Coupled Devices

The collection efficiency of the measurements performed on the waveguide coupled devices was estimated using 3D time-domain simulations. This was done by simulating both the far-field pattern of the antenna sitting at the center of the InP waveguide and the far-field pattern of light scattered out the end facet of the waveguide.

First the far-field emission pattern of light not coupled to the waveguide was simulated. This was done by simulating a single arch-antenna coupled ridge sitting on a 320nm thick by 3um wide waveguide 20um long. The power coupled into the waveguide was absorbed at the boundary of the simulation domain by using a multi-mode waveguide port. Figure B.1 shows a cross-section of the power flow coming from the nanoLED in the x-z plane.

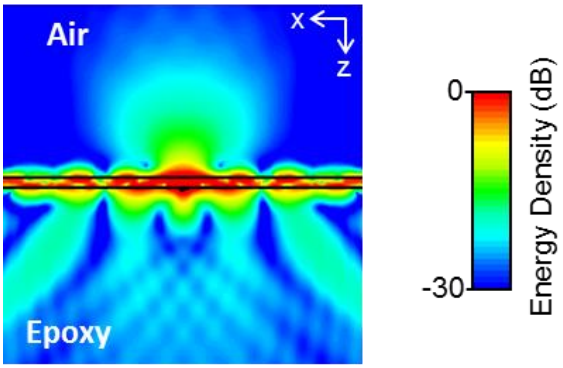


**Figure B.1.** Power flow from a waveguide coupled nanoLED shown in the x-z cutplane.

The collection efficiency was then calculated by integrating the amount of power falling within the collection cone of a 100x 0.8NA objective in the far-field. For measurements from the air-side this angle is ±53.17°, while from the epoxy-side it is only ±23.5°. The reduction in angle on the epoxy-side is due to total internal reflections of high angle rays at the glass-air interface (the epoxy is well index matched to the glass it is bonded to, n≈1.56). The results show that 32% of the light emitted into the air is collected while only 6% of the light emitted into the epoxy is collected. Bare and antenna clad ridges yielded the same collection efficiency. The same procedure was then used for the waveguide end-facet. The power flow of out of the end-facet is shown in Figure B.2.

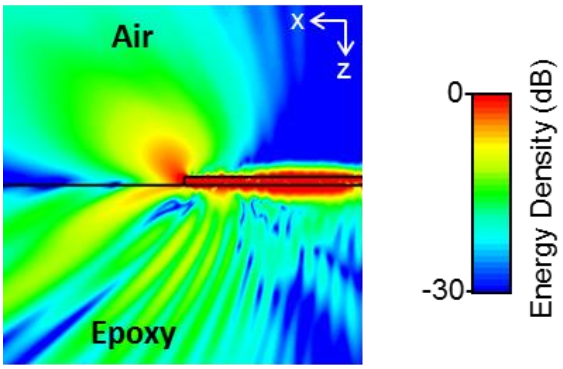


**Figure B.2.** Power flow of light scattered out the end of the waveguide shown in the x-z cutplane.

Simulations of the end facet show that 36% of the light coupled into the air is collected while only 12% of the light coupled into the epoxy is collected. Simulations were also performed with a more tapered output facet similar to that shown in Figure 4.7(b). This facet yielded slightly reduced collection efficiency of 29.5% in air and 7.3% in the epoxy layer. This would lead to an underestimate of the coupling efficiencies for a tapered end-facet.

A summary of the total distribution of emitted light is shown in Figure B.3. Using this data we can now check to see if comparing light collected from the end facets to light collected from the

center of the waveguide is a reasonable way to estimate coupling efficiency (Table B1). A measurement from the air-side would yield (2*3.8%)/(2*3.8%+5%) = 60.3% coupling efficiency and a measurement from the epoxy-side would yield (2*2.2%)/(2*2.2%+1.6%) = 73.3%. Comparing this to the actual coupling efficiency of 59%, we can see that air-side measurements provide a very close estimation of the coupling efficiency while the epoxy-side measurement over-estimates the coupling.

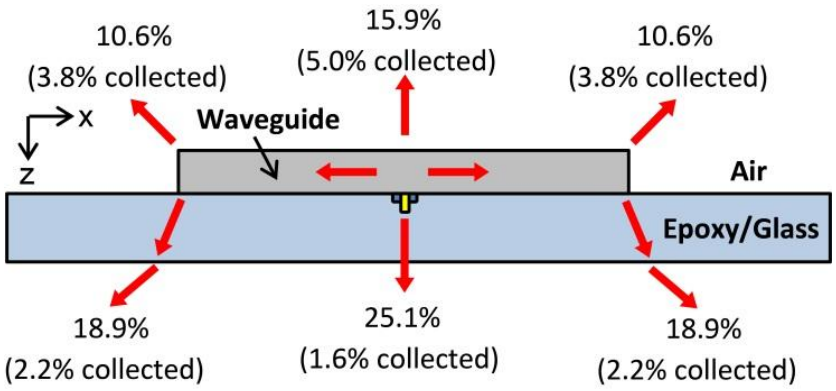


**Figure B.3.** Schematic of the distribution of emitted light from a nanoLED coupled to a 320nm thick waveguide obtained using 3D FDTD simulations.

**Table B1. Comparison of measured Arch-Antenna coupling efficiencies with what those predicted by simulated collection efficiency**

| Structure | Air-Side | Epoxy-Side |
|---|---|---|
| Measured $\eta_{coupling}$ | 43.9% | 50.4% |
| Simulated $\eta_{coupling}$ (using simulated $\eta_{collection}$) | 60.3% | 73.3% |
| Simulated $\eta_{coupling}$ (100% collection) | 59% | |

From Table B1 we can see that the measured values are about 20% lower than what is expected. The same calculation was done for the results obtained for the bare and Yagi-Uda cases and is shown in Table B2. Both the bare ridge and Yagi-Uda show very good agreement with expected values for the air-side measurement. The epoxy-side measurement, however, is lower than expected. This is most likely because epoxy-side measurements are much more susceptible to small changes in the scattered light due to the much lower collection efficiencies. Overall, the much higher collection efficiencies from the air-side offer a much higher confidence of the estimated coupling efficiency.

**Table B2. Comparison of measured and simulated waveguide coupling efficiencies**

| Structure | Simulated | Measured (Air) | Measured (Epoxy) |
|---|---|---|---|
| Bare Ridge | 59% | 57.5% | 48.4% |
| Arch-Antenna | 59% | 43.9% | 50.4% |
| Yagi-Antenna | 68% | 71.5% | 68.1% |

For the 880nm wide waveguide, the end-facet is completely tapered to a point. Simulations show that this type of end-facet has much more in-plane scattering out the end-facets that is not collected from either the air or epoxy side. This could explain why the coupling efficiencies measured in section 4.3.2 are lower than what was theoretically calculated.

# Appendix C. Free-Standing Arch-Dipole Fabrication

This appendix gives step-by-step instructions for fabricating an InGaAsP based arch-dipole antenna. Although optimal arch height is dependent on the exact antenna design, for the antennas discussed in Chapter 3 a 35nm thick InGaAsP layer was used that was lattice matched to InP and had a peak PL of ~1500nm. A typical epitaxial wafer structure for this device is shown in Figure C.1

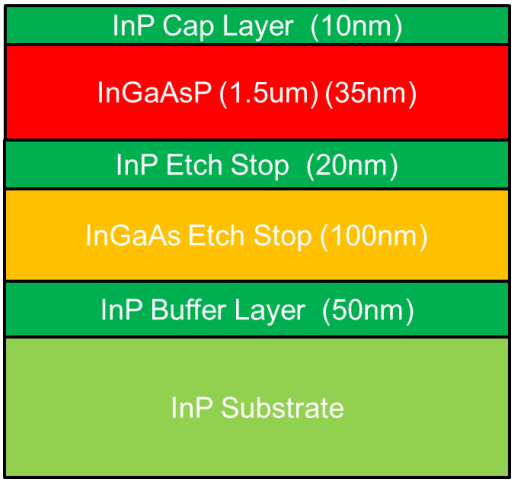


**Figure C.1.** Typical InGaAsP epitaxial layer structure.

The top l0nm thick layer of InP is to help prevent the wafer before processing. Directly below this layer is the 35nm thick InGaAsP active layer, followed by a 20nm InP etch stop, and finally a 100nm InGaAs etch stop layer. An optional 50nm buffer layer of InP is also shown in Figure

C.1. Buffer layers are typical in MOCVD growth to help create a more perfect starting layer than the substrate itself.

To conserve material, the wafer is diced into pieces ~5mm x 7mm. This size is large enough to be handled but doesn't consume too much of the wafer per chip. Rectangular pieces are diced to preserve knowledge of the crystal plane orientation of each individual chip. I typically align the long axis of the chip parallel to the $[0\bar{1}0]$ direction.

**Step 1.) Alignment Mark Deposition**

The first step of the process is to deposit alignment marks on your chip that you can align to in subsequent ebeam writing steps. This will be critical since you need tight alignment between the InGaAsP ridges and the metal antenna.

1.) Etch off the InP protective cap layer in a 1:1 solution of HCL:Phosphoric Acid. This etchant typically etches at ~1um/min, so the etch is over almost immediately. However, residue on the surface can delay the start of the etch. I keep the chip in the etchant for 7-10 seconds.
2.) Spin-on PMMA A2 at 2000rpm for 46 seconds, this will yield a final thickness of 80-90nm. Follow with a 90 second bake on a hot-plate set to 190°C. I always put the chip on a glass slide so it is not actually in direct contact with the hotplate. Be careful to ensure the back of the chip is clean. If PMMA is on the backside it will bond the chip to the glass slide during the bake.
3.) Make a light scratch at the top of the chip as an indicator for subsequent electron-beam (ebeam) writing steps. Load the chip into the ebeam lithography system (in my case Crestec). Since the chip is very small I recommend using aluminum tape instead of clips to hold the chip in place.
4.) After setting the beam current to 50 pA, perform focusing and stigmation adjustment on the ebeam system. The focus can be checked by focusing on a spot on the sample at full magnification while turning the beam sweep off, this burns a hole in the PMMA. Adjust the fine focus and stigmation until the holes burned in the PMMA are perfectly circular and in focus. At this point you can perform "Mark Management" which calibrates the alignment of the ebeam system. This is crucial in order to get good subsequent alignment.
5.) After "Mark Management" completes be sure to make note of the calibration numbers, you can check them against previous runs to make sure it was performed correctly. Then return to the bottom of the scratch on your sample. From this point I always go 1mm down and start writing my alignment mark pattern. I use the alignment mark pattern shown in Figure C.2 which allows for both visibility and tight alignment.

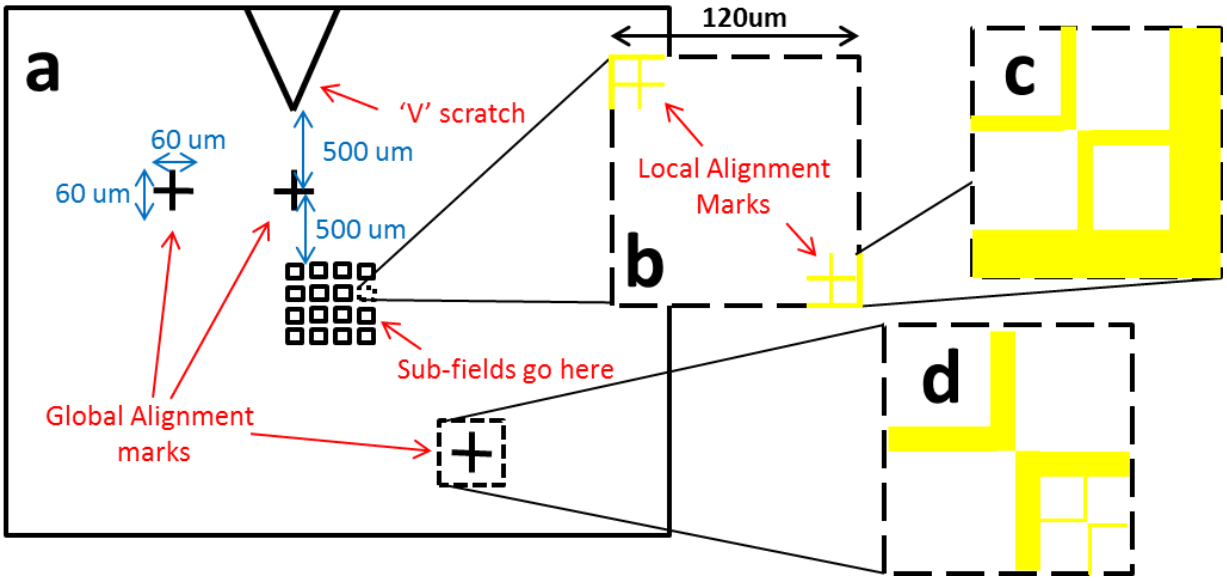


**Figure C.2.** Overview of typical ebeam chip layout. (a) Overview of the whole chip (typically ~5mm x 7mm). Three global alignment marks are used with two local alignment marks in each sub-field. (b) Overview of typical sub-field. 120um size with a local alignment mark in each corner. (c) Local alignment mark. Large "L" is typically 1um line width and small disconnected cross is 100nm line width. (d) Global alignment mark. Large disconnected cross is ~5um linewidth and small alignment mark is ~100nm linewidth.

6.) After exposure, develop the PMMA in 1:3 MIBK:IPA for 60sec. Follow with an IPA rinse, do not use water as I have noticed bad things happening when using water to rinse.
7.) At this point you should be able to see the alignment marks in the PMMA layer in a microscope. If you can, load the sample in an evaporator (I prefer to use ebeam evaporation such as Ultek2). You will want to load both a titanium and a gold source.
8.) Pump the chamber down to at least below 5e-6 Torr. Then evaporate 5nm of titanium at a rate of 0.8Å/s followed by 25nm of gold at 1.2Å/s. You can use slightly thicker gold if you need better contrast in future alignment mark steps. If you need really thick alignment marks (>50nm) you may want to consider using a thicker resist such as PMMA C4 (which can also form good alignment marks). I would avoid thicker adhesion layer since titanium can be attacked in some further processing steps and a thin layer helps reduce the probability of the alignment marks etching off.
9.) After evaporation, unload the sample and put it in a room temperature bath of acetone. The time it takes for lift-off can vary from a few minutes to up to an hour. If the film is having trouble lifting off you can use an acetone squirt gun to help loosen particles. In extreme cases you can use the acetone spray gun located on the right side of msink18, though beware anything not firmly stuck down may come off. I typically will NEVER use a sonicator with an InP substrate because it can easily cleave during sonication.
10.) Inspect the alignment marks under a microscope, they should be complete.

**Step 2.) InGaAsP Ridge Formation**

The next step is to define and etch the InGaAsP nanoridge.

2.1.) Spin-on PMMA A2 at 2000rpm for 46 seconds. Follow with a 90 second bake on a hot-plate set to 190°C.
2.2.) Load the chip into the ebeam lithography system with the "V" scratch in the same direction as when you did ebeam lithography for the alignment marks.
2.3.) After setting the beam current to 50 pA, perform focusing and stigmation adjustment on the ebeam system. At this point you can perform "Mark Management" which calibrates the alignment of the ebeam system. This is crucial in order to get good subsequent alignment.
2.4.) After "Mark Management" completes be sure to make note of the calibration numbers, you can check them against previous runs to make sure it was performed correctly. Then return to the bottom of the scratch on your sample. From this point go 1mm down and find your first global alignment mark. You can use this as a reference to find your other two alignment marks. I like to find them before I start writing since then when the ebeam writer asks you to "find the alignment marks" you already have the coordinates in hand. Be careful not to go over a region where you are going to write!! This is why the global alignment marks are far from the local fields.
2.5.) Pattern arrays of rectangles 240nm x 120nm aligned along crystal planes so that the long axis is parallel to the [011] direction (Figure 3.8), this will yield a final InGaAsP ridge approximately 150nm by 35nm. The dose will have to be changed according to the density of patterned rectangles. For good alignment you will want to use the local alignment marks to re-align each field you write. Be sure to zoom in quickly to avoid exposing the whole field.
2.6.) After exposure, develop the PMMA in 1:3 MIBK:IPA for 60sec. Follow with an IPA rinse.
2.7.) At this point you should be able to see the exposed rectangles in the PMMA layer in a microscope (try darkfield). Next, load the sample in an evaporator (I prefer to use ebeam evaporation such as Ultek2). You will want to load a titanium source.
2.8.) Pump the chamber down to at least below 3e-6 Torr. Then evaporate 7nm of titanium at a rate of 0.2Å/s. Ultek2 is very reproducible, even though the rate may fluctuate a bit.
2.9.) After evaporation, unload the sample and put it in a room temperature bath of acetone. The time it takes for lift-off can vary from a few minutes to up to an hour.
2.10.) Inspect the lifted off titanium islands under the microscope. You might not be able to see them, depending on the quality of the microscope. Make sure all of the fields have lifted off properly.
2.11.) Using a dilute piranha solution of 1:8:100 $H_2SO_4$:$H_20_2$:$H_20$, slowly etch the InGaAsP ridge. This strength will etch the InGaAsP layer (lattice matched to InP, peak emission ~1500nm) at a rate of ~5nm/s in the [100] direction, ~2nm/s in the [011] direction, and ~3.5nm/s in the $[0\ \overline{1}\ 1]$ direction. You can use an even more dilute solution if you want finer control of the etching process. Due to the anisotropy of the etchant, the sidewalls parallel to the [011] are fairly vertical for moderate aspect ratios (~1:1). Perform the etching in small steps (~10 sec at first then a few seconds of etching at a time after that) and then check under SEM. Because the hardmask layer is so thin, it is semi-transparent to the electron beam and the width of the underlying ridge can be determined. Make sure you have test ridges to look at because any ridge imaged in SEM will be coated in carbon which

will mask further etching. Continue etching until the ridges are etched to a width of ~30nm and remove the titanium hardmask with a quick dip in 49% hydrofluoric acid (5 sec should be enough). You can check in SEM to make sure the hardmask is fully removed.

2.12.) After removal of the titanium hardmask, load the sample into an atomic layer deposition (ALD) machine (such as picosun) and pump down to medium vacuum (~1Torr). Deposit a conformal 3nm layer of $TiO_2$ on the sample at 150$^{o}$C. If you use a clean pocket wafer to hold your chip you should be able to see a very slight shadow after deposition. You can also include a dummy chip and perform a contact resistance measurement to ensure the ALD was successful.

**Step 3.) Antenna Deposition**

The final nanofabrication step in the process is the deposition of the metal antenna. The alignment is critical to be able to effectively "cross your 'T's". Crestec is capable of at least ~20nm alignment tolerances if you did everything right.

3.1.) Spin-on PMMA A2 at 2000rpm for 46 seconds, this will yield a final thickness of 80-90nm. Follow with a 90 second bake on a hot-plate set to 190°C. I always put the chip on a glass slide so it is not actually in direct contact with the hotplate. Be careful to ensure the back of the chip is clean. If PMMA is on the backside it will bond the chip to the glass slide during the bake.

3.2.) Load the chip into the ebeam lithography system with the "V" scratch in the same direction as when you did ebeam lithography for the alignment marks. Since the chip is very small I recommend using aluminum tape instead of clips to hold the chip in place.

3.3.) After setting the beam current to 50 pA, perform focusing and stigmation adjustment on the ebeam system. The focus can be checked by focusing on a spot on the sample at full magnification while turning the beam sweep off, this burns a hole in the PMMA. Adjust the fine focus and stigmation until the holes burned in the PMMA are perfectly circular and in focus. At this point you can perform "Mark Management" which calibrates the alignment of the ebeam system. This is crucial in order to get good subsequent alignment.

3.4.) After "Mark Management" completes be sure to make note of the calibration numbers, you can check them against previous runs to make sure it was performed correctly. Then return to the bottom of the scratch on your sample. From this point go 1mm down and find your first global alignment mark. You can use this as a reference to find your other two alignment marks. I like to find them before I start writing since then when the ebeam writer asks you to "find the alignment marks" you already have the coordinates in hand. Be careful not to go over a region where you are going to write!! This is why the global alignment marks are far from the local fields.

3.5.) Next write the actual antennas overlaying perpendicular to the previous InGaAsP ridge mask. I typically shoot for 50nm wide antennas with lengths varying from 400nm to 800nm, though this usually requires the mask to have a narrower width (such as 30-40nm) to actually hit the 50nm size. For good alignment you will want to use the local alignment marks to re-align each field you write. Be sure to zoom in quickly to avoid exposing the whole field.

3.6.) After exposure, develop the PMMA in 1:3 MIBK:IPA for 60sec. Follow with an IPA rinse, do not use water as I have noticed bad things happening when using water to rinse.

3.7.) You may be able to see the exposed antennas in a microscope dark-field, but you may have to compare to an array of bare ridges. Next, load the sample in an evaporator (I prefer to use ebeam evaporation such as Ultek2). You will want to load a germanium and a gold source.
3.8.) Pump the chamber down to at least below 3e-6 Torr. Evaporate 2nm of germanium at a rate of 0.4 Å/s followed by 40nm of gold deposited at a rate of 1.2 Å/s. The slow deposition rates yield polycrystalline metal but offer superior conformal coverage compared to quickly evaporated metal. The germanium layer acts as a wetting layer for the gold to ensure connection of the metal arch over the ridge.
3.9.) After evaporation, unload the sample and put it in a room temperature bath of acetone. The time it takes for lift-off can vary from a few minutes to up to an hour. If the film is having trouble lifting off you can use an acetone squirt gun to help loosen particles. In extreme cases you can use the acetone spray gun located on the right side of msink18, though beware anything not firmly stuck down may come off. I typically will NEVER use a sonicator with an InP substrate because it can easily cleave during sonication.
3.10.) At this point you should take SEM images to validate successful fabrication of the arch antenna structure. This will be your last chance to image your antennas. Hopefully you have patterned arrays of test structures for imaging. Directly imaging devices with SEM can change their properties!!!

**Step 4.) Substrate Removal**

The final step of the process is to remove the high-index InP substrate. This is done by bonding your samples to a carrier substrate and using a combination of lapping and wet etching. A transparent substrate is highly recommended since it will make imaging and PL measurements much easier. However, avoid using glass!! Glass is full of dopants which will happily fluoresce and give off background signal during PL measurements. I would recommend using either Quartz or Sapphire.

4.1.) Thoroughly clean a quartz (sapphire) substrate that has been diced into ~1 cm x 1 cm square pieces with acetone and IPA. Inspect the chip to make sure there is no dust or particles present.
4.2.) Using double-stick tape, tape your antenna sample to a glass slide, InP substrate down. Be sure to tap the tape several times with your gloved finger to reduce its stickiness, otherwise you may never get your sample off the glass slide/tape!!!
4.3.) Place a small drop of UV curable epoxy (Norland Optical Adhesive 81) on top of your sample with a micro-pipette. Makes sure not to touch your sample with the tip because you don't want to scratch your samples. You want enough adhesive to form a full layer across the chip, but not so much that it is excessively thick.
4.4.) Place your clean quartz substrate onto the top of your sample. Slowly move it around to make sure all of the epoxy is evenly distributed. If there are any trapped particles or air bubbles move the quartz left and right to move the defects off your chip (you can trap them on the edges of the InP).
4.5.) Once the epoxy is well distributed between your sample and the quartz substrate load the sample into a UV oven (handling it by the glass slide it is taped to). Be careful that the quartz doesn't slide off (using a thicker epoxy like NOA 81 will reduce this problem)

4.6.) Bake the sample in the UV oven for 2min. Then take it out and make sure nothing went wrong. If there is a problem you can still theoretically get the quartz substrate off and try again since the epoxy isn't fully cured.

4.7.) If the sample looks good, put it back in the UV oven for 10 minutes.

4.8.) After UV baking is complete, slowly peel your sample off of the double stick tape. If you didn't dull the tape enough it will be very difficult to get off and you risk shattering your sample, be careful!!!!

4.9.) Once the chip is freed from the tape let it bake for 12 hours at 50°C on a hotplate. I put it on a glass slide so it doesn't actually touch the hotplate (which are usually always dirty). Baking for this long will fully cure the epoxy and reduce stress in your sample when you remove the substrate.

4.10.) After baking is complete, bond the chip to the metal chuck of a hand lapping tool using crystal wax. You can do this by placing the chuck on a hotplate set to 150°C and melting some crystal wax on it. Then place your sample on the chuck (quartz down) and move it around to get rid of any bubbles in the wax and make sure there is a nice clean bond. Take the chuck off the hotplate and set on a heat sink to return it to room temperature (obviously the chuck will be very hot, so use gloves or tweezers).

4.11.) Load the chuck on the lapping tool and using 1500 grit sandpaper lap the InP substrate down to a thickness of ~50-75μm. Use wet sandpaper and keep a small drip of water running on the sandpaper while lapping. If your lapping tool is accurate you can use it's built in gauge to determine thickness. Otherwise you a height profilometer and measure often!

4.12.) Once there is only 50-75μm left, take the chuck off the lapping tool. Wash off the chuck with water and dry it. Then place it back on the hotplate at 150°C to melt the crystal wax and remove your chip. Clean the rest of the crystal wax off the chuck with acetone and IPA.

4.13.) At this point your sample will look pretty bad and be covered in crystal wax. Thoroughly clean it off with acetone and IPA. After rinsing the chip in acetone I place it in a shallow petri dish of acetone and then scrub both sides (the quartz and InP side) with a q-tip until the surfaces are clean. It is imperative to get rid of all residue that could mask the subsequent wet etch. A >50um InP thickness is plenty to protect your sample from being scratched by the q-tip.

4.14.) With your sample thoroughly cleaned, etch off the remaining 50um of InP in a 1:1 solution of Hydrochloric Acid:Phosphoric Acid at 80°C. You can tell if it's etching because it will give off lots of bubbles. Watch it carefully in the first several minutes and if you notice any areas of the chip not etching gently agitate it with a Teflon pipette. If you don't do this your sample can etch very unevenly and cause problems. You will know when the etching is done because the bubbling will stop and the surface of your chip will be very shiny.

4.15.) Take your chip out of the Hydrochloric Acid:Phosphoric Acid etchant and wash it off thoroughly with DI water. You want to make sure there is no acid remaining on the chip which could degrade the etch selectivity of subsequent steps. Dry the chip with an $N_2$ gun.

4.16.) Next, place your chip in a 1:1:10 solution of $H_2SO_4$:$H_20_2$:$H_20$. This will etch through your InGaAs etch stop layer. I typically etch for ~30sec for a 100nm thick InGaAs etch stop layer. You can tell when it's done because the chip will stop changing colors. You can over-etch because the selectivity with InP is very high ($>10^6$).

4.17.) At this point, the only thing remaining on your chip is a thin InP layer below which are your devices. Etch off this InP layer in a 1:1 solution of Hydrochloric Acid:Phosphoric Acid at room temperature. A 10nm layer will etch off in a few seconds and the chip will become transparent. Don't use heated acid.

Congratulations!! You have successfully fabricated a chip of free-standing arch-dipole antennas coupled to nanoridges of InGaAsP!

# Appendix D. Waveguide Coupled Arch-Dipole Fabrication

This appendix gives step-by-step instructions for fabricating an InGaAsP based arch-dipole antenna on top of an InP waveguide. Although optimal arch height is dependent on the exact antenna design, for the antennas discussed in Chapter 3 a 35nm thick InGaAsP layer was used that was lattice matched to InP and had a peak PL of ~1300nm. A typical epitaxial wafer structure for this device is shown in Figure C.1

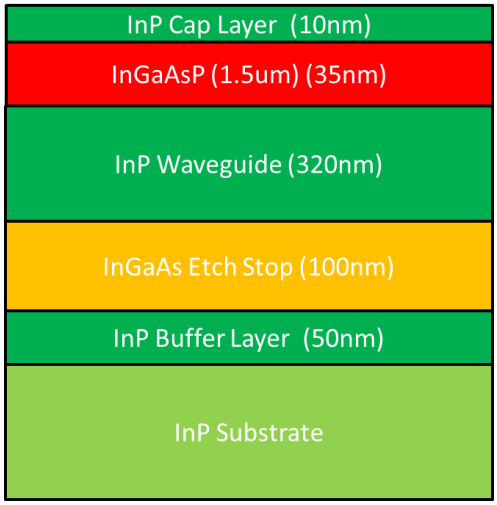


**Figure D.1.** Typical InGaAsP epitaxial layer structure for a waveguide coupled nanoLED.

The top l0nm thick layer of InP is to help prevent the wafer before processing. Directly below this layer is the 35nm thick InGaAsP active layer, followed by a 320nm InP layer that will be patterned into a waveguide, and finally a 100nm InGaAs etch stop layer. An optional 50nm buffer layer of InP is also shown in Figure C.1. Buffer layers are typical in MOCVD growth to help create a more perfect starting layer than the substrate itself.

To conserve material, the wafer is diced into pieces ~5mm x 7mm. This size is large enough to be handled but doesn't consume too much of the wafer per chip. Rectangular pieces are diced to preserve knowledge of the crystal plane orientation of each individual chip. I typically align the long axis of the chip parallel to the $[0\bar{1}0]$ direction.

**Step 1.) Alignment Mark Deposition**

The first step of the process is to deposit alignment marks on your chip that you can align to in subsequent ebeam writing steps. This will be critical since you need tight alignment between the InGaAsP ridges and the metal antenna. These steps are identical to those from Appendix C Step 1.

**Step 2.) InP Waveguide Formation**

The next step is to define and etch the InP waveguide.

2.1.) Spin-on PMMA A2 at 2000rpm for 46 seconds. Follow with a 90 second bake on a hot-plate set to 190°C.
2.2.) Load the chip into the ebeam lithography system with the "V" scratch in the same direction as when you did ebeam lithography for the alignment marks.
2.3.) After setting the beam current to 50 pA, perform focusing and stigmation adjustment on the ebeam system. At this point you can perform "Mark Management" which calibrates the alignment of the ebeam system. This is crucial in order to get good subsequent alignment.
2.4.) After "Mark Management" completes be sure to make note of the calibration numbers, you can check them against previous runs to make sure it was performed correctly. Then return to the bottom of the scratch on your sample. From this point go 1mm down and find your first global alignment mark. You can use this as a reference to find your other two alignment marks. I like to find them before I start writing since then when the ebeam writer asks you to "find the alignment marks" you already have the coordinates in hand. Be careful not to go over a region where you are going to write!! This is why the global alignment marks are far from the local fields.
2.5.) Pattern arrays of rectangles to the desired dimensions of your waveguide (50μm x ~1μm in my case) aligned along crystal planes so that the long axis is parallel to the [011] direction (Figure 3.8). The dose will have to be changed according to the density of patterned rectangles. For good alignment you will want to use the local alignment marks to re-align each field you write. Be sure to zoom in quickly to avoid exposing the whole field.
2.6.) After exposure, develop the PMMA in 1:3 MIBK:IPA for 60sec. Follow with an IPA rinse.
2.7.) At this point you should be able to see the exposed rectangles in the PMMA layer in a microscope. Next, load the sample in an evaporator (I prefer to use ebeam evaporation such as Ultek2). You will want to load a titanium source.
2.8.) Pump the chamber down to at least below 3e-6 Torr. Then evaporate 7nm of titanium at a rate of 0.2Å/s. Ultek2 is very reproducible, even though the rate may fluctuate a bit.
2.9.) After evaporation, unload the sample and put it in a room temperature bath of acetone. The time it takes for lift-off can vary from a few minutes to up to an hour.
2.10.) Inspect the lifted off titanium islands under the microscope. You might not be able to see them, depending on the quality of the microscope. Make sure all of the fields have lifted off properly.
2.11.) Using a dilute piranha solution of 1:8:100 $H_2SO_4$:$H_20_2$:$H_20$, slowly etch the InGaAsP layer. The etching time depends on the exact InGaAsP material ratio, but it will be ~30-45sec to get through the layer. If you over-etch the waveguide will get narrower. You can use this to fine tune the waveguide width.
2.12.) Next, use a 1:3 HCL:Phosphoric Acid solution to etch the InP below the InGaAsP hardmask. The InGaAsP acts as a great mask and the sidewalls should etch very vertically and not undercut the hardmask. If you overetch then the ends of the waveguide will get

shorter. The time to etch through the InP is ~30sec. You can tell if you've etched through because the chip will look very uniform in coloration and very smooth (in SEM). You can also use a profilometer to make sure your waveguides are fully etched.

2.13.) Remove the titanium hardmask with a quick dip in 49% hydrofluoric acid (5 sec should be enough). You can check in SEM to make sure the hardmask is fully removed.

**Step 3.) InGaAsP Ridge Formation**

The next step is to define and etch the InGaAsP nanoridge.

3.1.) Spin-on PMMA C4 at 3000rpm for 60 seconds. Follow with a 90 second bake on a hot-plate set to 190°C. This will yield a ~400nm thick film which is plenty to fully submerge the waveguides.

3.2.) Load the chip into the ebeam lithography system with the "V" scratch in the same direction as when you did ebeam lithography for the alignment marks.

3.3.) After setting the beam current to 50 pA, perform focusing and stigmation adjustment on the ebeam system. At this point you can perform "Mark Management" which calibrates the alignment of the ebeam system. This is crucial in order to get good subsequent alignment.

3.4.) After "Mark Management" completes be sure to make note of the calibration numbers, you can check them against previous runs to make sure it was performed correctly. Then return to the bottom of the scratch on your sample. From this point go 1mm down and find your first global alignment mark. You can use this as a reference to find your other two alignment marks. I like to find them before I start writing since then when the ebeam writer asks you to "find the alignment marks" you already have the coordinates in hand. Be careful not to go over a region where you are going to write!! This is why the global alignment marks are far from the local fields.

3.5.) Pattern a single rectangle at the center of each waveguide. I use roughly a 240nm x 120nm sized mask aligned along crystal planes so that the long axis is parallel to the [011] direction (Figure 3.8), this will yield a final InGaAsP ridge approximately 150nm by 35nm. The dose will have to be changed according to the density of patterned rectangles. For good alignment you will want to use the local alignment marks to re-align each field you write. Be sure to zoom in quickly to avoid exposing the whole field.

3.6.) After exposure, develop the PMMA in 1:3 MIBK:IPA for 60sec. Follow with an IPA rinse.

3.7.) At this point you should be able to see the exposed rectangles in the PMMA layer in a microscope (try darkfield). Next, load the sample in an evaporator (I prefer to use ebeam evaporation such as Ultek2). You will want to load a titanium source.

3.8.) Pump the chamber down to at least below 3e-6 Torr. Then evaporate 7nm of titanium at a rate of 0.2Å/s. Ultek2 is very reproducible, even though the rate may fluctuate a bit.

3.9.) After evaporation, unload the sample and put it in a room temperature bath of acetone. The time it takes for lift-off can vary from a few minutes to up to an hour.

3.10.) Inspect the lifted off titanium islands under the microscope. Make sure all of the fields have lifted off properly.

3.11.) Using a dilute piranha solution of 1:8:100 $H_2SO_4$:$H_20_2$:$H_20$, slowly etch the InGaAsP ridge. You can use an even more dilute solution if you want finer control of the etching process. Due to the anisotropy of the etchant, the sidewalls parallel to the [011] are fairly vertical for moderate aspect ratios (~1:1). Perform the etching in small steps (~10 sec at first then a few seconds of etching at a time after that) and then check under SEM. Because

the hardmask layer is so thin, it is semi-transparent to the electron beam and the width of the underlying ridge can be determined. Make sure you have test ridges to look at because any ridge imaged in SEM will be coated in carbon which will mask further etching. Continue etching until the ridges are etched to a width of ~30nm and remove the titanium hardmask with a quick dip in 49% hydrofluoric acid (5 sec should be enough). You can check in SEM to make sure the hardmask is fully removed. Note that InGaAsP that emits at a shorter wavelength takes longer to etch.

3.12.) After removal of the titanium hardmask, load the sample into an atomic layer deposition (ALD) machine (such as picosun) and pump down to medium vacuum (~1Torr). Deposit a conformal 3nm layer of $TiO_2$ on the sample at 150$^{o}$C. If you use a clean pocket wafer to hold your chip you should be able to see a very slight shadow after deposition. You can also include a dummy chip and perform a contact resistance measurement to ensure the ALD was successful.

**Step 4.) Antenna Deposition**

The final nanofabrication step in the process is the deposition of the metal antenna. The alignment is critical to be able to effectively "cross your 'T's". Crestec is capable of at least ~20nm alignment tolerances if you did everything right.

4.1.) Spin-on PMMA C4 at 3000rpm for 60 seconds. Follow with a 90 second bake on a hot-plate set to 190°C.

4.2.) Load the chip into the ebeam lithography system with the "V" scratch in the same direction as when you did ebeam lithography for the alignment marks. Since the chip is very small I recommend using aluminum tape instead of clips to hold the chip in place.

4.3.) After setting the beam current to 50 pA, perform focusing and stigmation adjustment on the ebeam system. The focus can be checked by focusing on a spot on the sample at full magnification while turning the beam sweep off, this burns a hole in the PMMA. Adjust the fine focus and stigmation until the holes burned in the PMMA are perfectly circular and in focus. At this point you can perform "Mark Management" which calibrates the alignment of the ebeam system. This is crucial in order to get good subsequent alignment.

4.4.) After "Mark Management" completes be sure to make note of the calibration numbers, you can check them against previous runs to make sure it was performed correctly. Then return to the bottom of the scratch on your sample. From this point go 1mm down and find your first global alignment mark. You can use this as a reference to find your other two alignment marks. I like to find them before I start writing since then when the ebeam writer asks you to "find the alignment marks" you already have the coordinates in hand. Be careful not to go over a region where you are going to write!! This is why the global alignment marks are far from the local fields.

4.5.) Next write the actual antennas overlaying perpendicular to the previous InGaAsP ridge mask. I typically shoot for 50nm wide antennas with lengths varying from 400nm to 800nm, though this usually requires the mask to have a narrower width (such as 30-40nm) to actually hit the 50nm size. If you want to make Yagi-Uda's then remember to pattern the reflector and director elements at the same time as the active antenna. For good alignment you will want to use the local alignment marks to re-align each field you write. Be sure to zoom in quickly to avoid exposing the whole field.

4.6.) After exposure, develop the PMMA in 1:3 MIBK:IPA for 60sec. Follow with an IPA rinse, do not use water as I have noticed bad things happening when using water to rinse.
4.7.) You may be able to see the exposed antennas in a microscope dark-field, but you may have to compare to an array of bare ridges. Next, load the sample in an evaporator (I prefer to use ebeam evaporation such as Ultek2). You will want to load a germanium and a gold source.
4.8.) Pump the chamber down to at least below 3e-6 Torr. Evaporate 2nm of germanium at a rate of 0.4 Å/s followed by 40nm of gold deposited at a rate of 1.2 Å/s. The slow deposition rates yield polycrystalline metal but offer superior conformal coverage compared to quickly evaporated metal. The germanium layer acts as a wetting layer for the gold to ensure connection of the metal arch over the ridge.
4.9.) After evaporation, unload the sample and put it in a room temperature bath of acetone. The time it takes for lift-off can vary from a few minutes to up to an hour. If the film is having trouble lifting off you can use an acetone squirt gun to help loosen particles. In extreme cases you can use the acetone spray gun located on the right side of msink18, though beware anything not firmly stuck down may come off. I typically will NEVER use a sonicator with an InP substrate because it can easily cleave during sonication.
4.10.) At this point you should take SEM images to validate successful fabrication of the arch antenna structure. This will be your last chance to image your antennas. Hopefully you have patterned arrays of test structures for imaging. Directly imaging devices with SEM can change their properties!!!

**Step 5.) Substrate Removal**

The final step of the process is to remove the high-index InP substrate. This is done by bonding your samples to a carrier substrate and using a combination of lapping and wet etching. A transparent substrate is highly recommended since it will make imaging and PL measurements much easier. However, avoid using glass!! Glass is full of dopants which will happily fluoresce and give off background signal during PL measurements. I would recommend using either Quartz or Sapphire.

5.1.) Thoroughly clean a quartz (sapphire) substrate that has been diced into ~1 cm x 1 cm square pieces with acetone and IPA. Inspect the chip to make sure there is no dust or particles present.
5.2.) Using double-stick tape, tape your antenna sample to a glass slide, InP substrate down. Be sure to tap the tape several times with your gloved finger to reduce its stickiness, otherwise you may never get your sample off the glass slide/tape!!!
5.3.) Place a small drop of UV curable epoxy (Norland Optical Adhesive 81) on top of your sample with a micro-pipette. Makes sure not to touch your sample with the tip because you don't want to scratch your samples. You want enough adhesive to form a full layer across the chip, but not so much that it is excessively thick.
5.4.) Place your clean quartz substrate onto the top of your sample. Slowly move it around to make sure all of the epoxy is evenly distributed. If there are any trapped particles or air bubbles move the quartz left and right to move the defects off your chip (you can trap them on the edges of the InP).
5.5.) Once the epoxy is well distributed between your sample and the quartz substrate load the sample into a UV oven (handling it by the glass slide it is taped to). Be careful that the quartz doesn't slide off (using a thicker epoxy like NOA 81 will reduce this problem)

5.6.) Bake the sample in the UV oven for 2min. Then take it out and make sure nothing went wrong. If there is a problem you can still theoretically get the quartz substrate off and try again since the epoxy isn't fully cured.

5.7.) If the sample looks good, put it back in the UV oven for 10 minutes.

5.8.) After UV baking is complete, slowly peel your sample off of the double stick tape. If you didn't dull the tape enough it will be very difficult to get off and you risk shattering your sample, be careful!!!!

5.9.) Once the chip is freed from the tape let it bake for 12 hours at 50°C on a hotplate. I put it on a glass slide so it doesn't actually touch the hotplate (which are usually always dirty). Baking for this long will fully cure the epoxy and reduce stress in your sample when you remove the substrate.

5.10.) After baking is complete, bond the chip to the metal chuck of a hand lapping tool using crystal wax. You can do this by placing the chuck on a hotplate set to 150°C and melting some crystal wax on it. Then place your sample on the chuck (quartz down) and move it around to get rid of any bubbles in the wax and make sure there is a nice clean bond. Take the chuck off the hotplate and set on a heat sink to return it to room temperature (obviously the chuck will be very hot, so use gloves or tweezers).

5.11.) Load the chuck on the lapping tool and using 1500 grit sandpaper lap the InP substrate down to a thickness of ~50-75μm. Use wet sandpaper and keep a small drip of water running on the sandpaper while lapping. If your lapping tool is accurate you can use it's built in gauge to determine thickness. Otherwise you a height profilometer and measure often!

5.12.) Once there is only 50-75μm left, take the chuck off the lapping tool. Wash off the chuck with water and dry it. Then place it back on the hotplate at 150°C to melt the crystal wax and remove your chip. Clean the rest of the crystal wax off the chuck with acetone and IPA.

5.13.) At this point your sample will look pretty bad and be covered in crystal wax. Thoroughly clean it off with acetone and IPA. After rinsing the chip in acetone I place it in a shallow petri dish of acetone and then scrub both sides (the quartz and InP side) with a q-tip until the surfaces are clean. It is imperative to get rid of all residue that could mask the subsequent wet etch. A >50um InP thickness is plenty to protect your sample from being scratched by the q-tip.

5.14.) With your sample thoroughly cleaned, etch off the remaining 50um of InP in a 1:1 solution of Hydrochloric Acid:Phosphoric Acid at 80°C. You can tell if it's etching because it will give off lots of bubbles. Watch it carefully in the first several minutes and if you notice any areas of the chip not etching gently agitate it with a Teflon pipette. If you don't do this your sample can etch very unevenly and cause problems. You will know when the etching is done because the bubbling will stop and the surface of your chip will be very shiny.

5.15.) Take your chip out of the Hydrochloric Acid:Phosphoric Acid etchant and wash it off thoroughly with DI water. You want to make sure there is no acid remaining on the chip which could degrade the etch selectivity of subsequent steps. Dry the chip with an $N_2$ gun.

5.16.) Next, place your chip in a 1:1:10 solution of $H_2SO_4$:$H_20_2$:$H_20$. This will etch through your InGaAs etch stop layer. I typically etch for ~30sec for a 100nm thick InGaAs etch stop layer. You can tell when it's done because the chip will become completely transparent and the waveguides should be highly visible.

Congratulations!! You have successfully fabricated waveguide coupled arch-dipole antennas coupled to nanoridges of InGaAsP!

# Appendix E. Cavity-Backed Slot Antenna Fabrication

This appendix gives step-by-step instructions for fabricating a silver cavity-backed slot antenna coupled to a strip of $WSe_2$. Following these steps will yield a ~30nm x 250nm slot that is 40nm tall (SiOx) with a ~10nm strip of chrome at the top of the cavity. The fabrication process for the cavity-backed slot antennas is outlined in Figure 6.7.

**Step 1.) Fabricate Alignment Mark Chips**

The first step of the process is to fabricate alignment mark chips on which flakes of $WSe_2$ can be transferred to. Start with an InP wafer such as the one depicted in Figure C.1, although only a single epitaxial layer is required, the 100nm InGaAs layer. If your wafer has more layers, that's fine, you'll just have to do more etching.

1.1.) Dice up the wafer into small chips, ~6mm x 6mm. The crystal orientation does not matter.
1.2.) Load the chips into an atomic layer deposition (ALD) machine (such as picosun) and pump down to medium vacuum (~1Torr). Deposit 20nm of $Al_2O_3$ on the sample at 150$^{o}$C. If you use a clean pocket wafer to hold your chip you should be able to see a very slight shadow after deposition. You can also include a dummy chip and perform a contact resistance measurement to ensure the ALD was successful.
1.3.) After ALD, spin-on PMMA A2 at 2000rpm for 46 seconds, this will yield a final thickness of 80-90nm. Follow with a 90 second bake on a hot-plate set to 190°C.
1.4.) Make a light scratch at the top of the chip as an indicator for subsequent electron-beam (ebeam) writing steps. Load the chip into the ebeam lithography system (in my case Crestec).
1.5.) After setting the beam current to 50 pA, perform focusing and stigmation adjustment on the ebeam system. Then, perform "Mark Management" which calibrates the alignment of the ebeam system. This is crucial in order to get good subsequent alignment.
1.6.) After "Mark Management" completes be sure to make note of the calibration numbers, you can check them against previous runs to make sure it was performed correctly. Then return to the bottom of the scratch on your sample. From this point I always go 1mm down and start writing my alignment mark pattern. I use the alignment mark pattern shown in Figure C.2 which allows for both visibility and tight alignment.
1.7.) After exposure, develop the PMMA in 1:3 MIBK:IPA for 60sec. Follow with an IPA rinse.
1.8.) At this point you should be able to see the alignment marks in the PMMA layer in a microscope. If you can, load the sample in an evaporator (I prefer to use ebeam evaporation such as Ultek2). You will want to load both a titanium and a gold source.
1.9.) Pump the chamber down to at least below 5e-6 Torr. Then evaporate 5nm of titanium at a rate of 0.8Å/s followed by 25nm of gold at 1.2Å/s. You can use slightly thicker gold if you need better contrast in future alignment mark steps.
1.10.) After evaporation, unload the sample and put it in a room temperature bath of acetone. The time it takes for lift-off can vary from a few minutes to up to an hour.

1.11.) Inspect the alignment marks under a microscope, they should be complete.

**Step 2.) Transfer $WSe_2$ flakes onto sample**

First exfoliate $WSe_2$ onto a 260 nm thick $Si/SiO_2$ substrate and then map it using an optical microscope to find monolayer flakes. You can confirm the flake thickness by photoluminescence measurements. Then transfer $WSe_2$ monolayers onto the InP alignment mark substrate using poly(methyl methacrylate) (PMMA) as the transfer medium [95]. The $Al_2O_3$ layer provides contrast so that the monolayer is still visible after transfer. Map the flake in an optical microscope to get measurements of how far it is from your alignment marks. You will need this information when writing your ebeam mask.

**Step 3.) Form SiOx Slot and Deposit Antenna**

After the flake is transferred to the InP carrier it is coated in 40nm of $SiO_x$ which is etched to form the slot of the slot antenna.

3.1.) Load the sample in an evaporator that allows dielectric evaporation (I prefer to use ebeam evaporation such as edwardseb3). You will want to load a silicon oxide source. Note that the silicon oxide pellets will never really melt together.
3.2.) Pump the chamber down to at least below 5e-6 Torr. Evaporate the silicon oxide at ~1 Å/s. If you go too fast the silicon and oxygen will dissociate and you will just be depositing silicon. Evaporate a thickness corresponding to the depth of your slot, i.e. 40nm, and then unload the sample.
3.3.) Spin-on PMMA A2 at 2000rpm for 46 seconds. Follow with a 90 second bake on a hot-plate set to 190°C.
3.4.) Load the chip into the ebeam lithography system with the "V" scratch in the same direction as when you did ebeam lithography for the alignment marks.
3.5.) After setting the beam current to 50 pA, perform focusing and stigmation adjustment on the ebeam system. At this point you can perform "Mark Management" which calibrates the alignment of the ebeam system. This is crucial in order to get good subsequent alignment.
3.6.) After "Mark Management" completes be sure to make note of the calibration numbers, you can check them against previous runs to make sure it was performed correctly.
3.7.) Pattern arrays of rectangles the size of the slots you want to make, typically 30nm x 250nm, aligned to the $WSe_2$ monolayer. The dose will have to be changed according to the density of patterned rectangles. I typically pattern an array ~20μm x 20μm so the alignment with the actual flake isn't too critical (since the flake is typically ~5μm x 5μm).
3.8.) After exposure, develop the PMMA in 1:3 MIBK:IPA for 60sec. Follow with an IPA rinse.
3.9.) At this point you should be able to see the exposed rectangles in the PMMA layer in a microscope (try darkfield). Next, load the sample in an evaporator (I prefer to use ebeam evaporation such as Ultek2). You will want to load a chromium source.
3.10.) Pump the chamber down to at least below 3e-6 Torr. Then evaporate 7nm of chrome at a rate of 0.2Å/s. Ultek2 is very reproducible, even though the rate may fluctuate a bit.
3.11.) After evaporation, unload the sample and put it in a room temperature bath of acetone. The time it takes for lift-off can vary from a few minutes to up to an hour.

3.12.) Inspect the lifted off chrome islands under the microscope. You might not be able to see them, depending on the quality of the microscope. Make sure all of the fields have lifted off properly.
3.13.) Load the chip into a dry etcher (such as technic-c). Then etch the silicon oxide layer using a 45sec $SF_6$ reactive ion etch. The etch will completely removes all the oxide and the $WSe_2$ not protected by the chrome. This creates a 40nm tall ridge of SiOx with the thin layer of chrome on top. $SF_6$ has a very high etch selectivity over $Al_2O_3$, so it acts as an effective etch stop layer. I would highly recommend etching dummy chip first to calibrate the etch rate. You can tell if your actual chip is etched through because you will no longer be able to see the monolayer. Don't overetch!!
3.14.) Take SEM images of the etched ridges to confirm the etching went well.
3.15.) For antenna-coupled structures the entire chip is then coated with a blanket evaporation of silver which encapsulates the $SiO_x$ and forms the slot antenna. The metal evaporation step is skipped for bare samples. If you are making an antenna coupled structure load it into and evaporator with a titanium and silver source.
3.16.) Wait until the evaporator gets to the lowest possible pressure. Silver quality highly depends on evaporation pressure, you want something ~5e-7 Torr for good metal quality. Start by evaporating ~1nm of titanium at ~0.2Å/s. This will not be a uniform layer but will create pinning sites to help the sliver stick. Then deposit 120nm of silver. I typically start at a rate of ~1.2Å/s and then go to ~2.5Å/s after depositing ~50nm. Cap the entire structure off with 20nm of titanium to help adhesion to epoxy and prevent silver oxidation.

**Step 4.) Remove substrate**

The final step is to remove the InP substrate. To do this follow the instructions in Appendix XX, step 5. Note that when doing substrate removal the entire chip will be a shiny silver color when the InP substrate is fully removed. The 20nm of $Al_2O_3$ should be plenty to prevent the silver from oxidizing or tarnishing. It will also act as an etch stop layer for the 1:1:10 solution of $H_2SO_4$:$H_20_2$:$H_20$.

Congradulations!!! You have created a cavity-backed slot antenna coupled to a $WSe_2$ monolayer!!

# Appendix F. Optical Measurement Setup

The setup used to do all of the optical measurements is shown in Figure F.1. It consists of an Eclipse TE-2000U Nikon inverted microscope with a secondary custom-built microscope on top in place of the traditional white light illuminator. The inverted microscope has two beam splitter cube turrets (BS3 and BS4) which can hold six beam splitter cubes each. Each beam splitter cube can hold a dichroic mirror and two filters, commonly used beam splitter configurations are shown in Table VI and Table VII. The inverted microscope has a nosepiece turret with an adjustable Z-axis for focusing.

The TE-2000U has 4 output ports that can be selected using the Port Selector. A PI SpectraPro 2300i is at the focal plane of one output and is attached to a PI LN-cooled 2D Silicon CCD array. The other ports include an eyepiece, a silicon CCD imaging camera, and an unused port. White

light illumination is supplied through the lower beam splitter cube turret and is also the path of IR light that can be channeled into a secondary PI SpectraPro 2300i that is connected to a PI LN-cooled linear InGaAs array. White light and IR PL are filtered using a 1000nm short-pass dichroic mirror (Thorlabs DMSP1000). Finally, the upper beam splitter turret has a port where free-space laser illumination is brought in.

The upper microscope is composed of two 50:50 pellicle beam splitter cubes as well as an LED illuminator, a silicon CCD imaging camera, and a fiber-coupled laser port used for top side laser probing and alignment. This whole microscope is on a Sutter MP-285 micromanipulator stage that allows for full XYZ motion.

The sample is held in-between the two microscopes with a second Sutter MP-285 stage, allowing the sample to move in the XYZ directions independent of the top and bottom microscopes.

**Table VI: Description of beamsplitters in optical setup.** List of the beamsplitters used in the five beamsplitter locations listed in Figure F.1.

| Optic | Description |
|---|---|
| BS1/2 | 50:50 Pellicle Beamsplitter |
| BS3 | Beamsplitter Cube (2 filters, 1 dichroic mirror) |
| BS4 | Beamsplitter Cube (2 filters, 1 dichroic mirror) |
| BS5 | Thorlabs DMSP1000 (Short-pass Dichroic Mirror) |

**Table VII: Description of beamsplitters in beamsplitter turret.** List of the two most commonly used configurations for the inverted microscope beamsplitter turret.

| Function \ Optic | BS3 | BS4 |
|---|---|---|
| InGaAsP PL | Thorlabs DMLP1180 (Longpass Dichroic) | Gold Mirror |
| $WSe_2$ PL | Semrock FF02-525/40 (Bandpass Filter)<br>Di02-R532 (Dichroic Beamsplitter)<br>BLP01-594R (Long Pass Filter) | Blank<br>**OR**<br>50:50 Beamsplitter (imaging) |

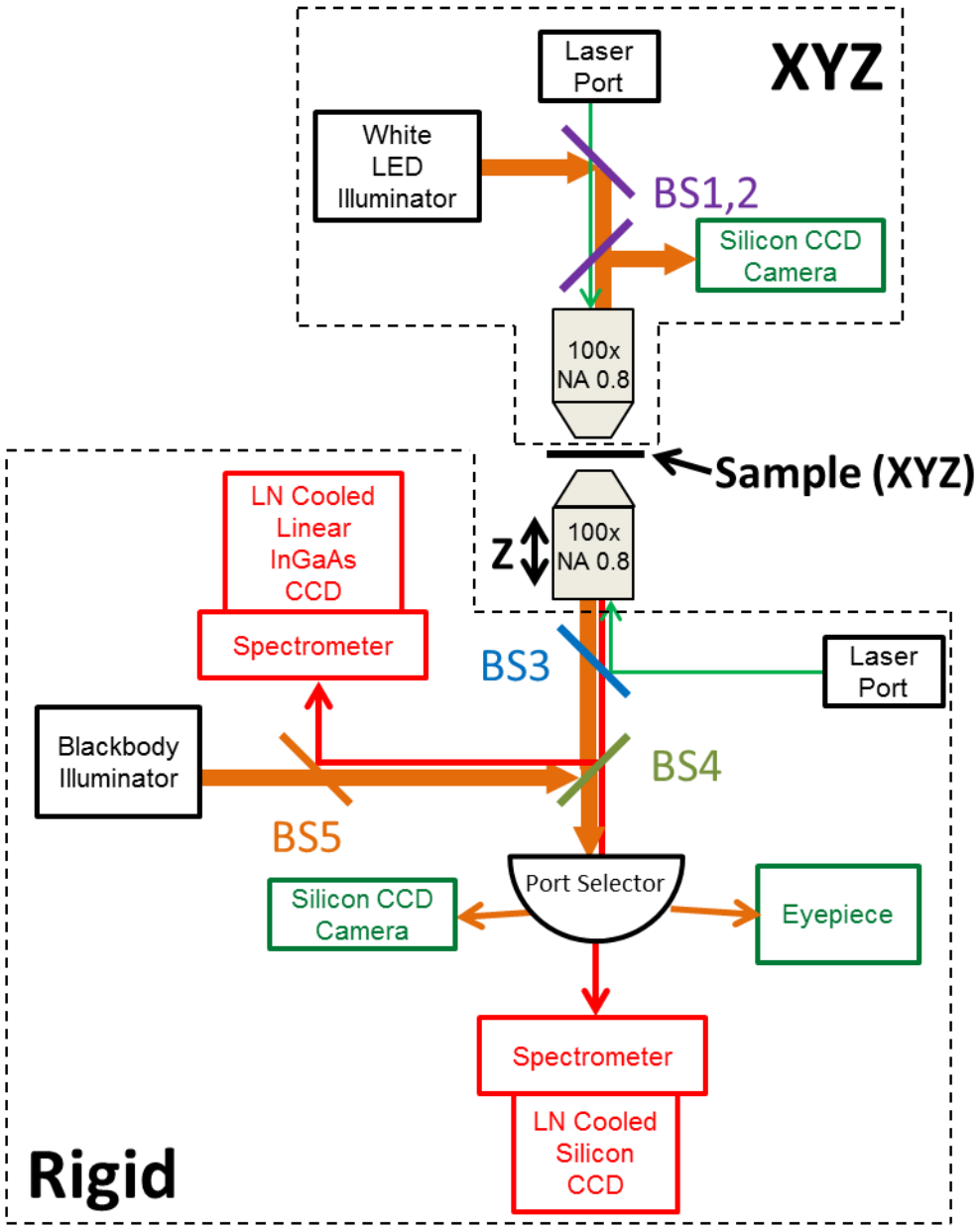


**Figure F.1.** Optical measurement setup consisting of a rigid inverted microscope with a custom-built movable microscope on top.

# References


1. Bruce Robertson, *Loss Measuring Concepts* (Kingfisher International, 2005).
2. D. A. . Miller, "Device Requirement for Optical Interconnects to Silicon Chips," Proc IEEE Spec. Issue Silicon Photonics (2009).
3. B. G. Lee, J. Van Campenhout, A. V. Rylyakov, C. L. Schow, W. M. Green, S. Assefa, M. Yang, F. E. Doany, C. V. Jahnes, R. A. John, and others, "Broadband silicon photonic switch integrated with CMOS drive electronics," in *Conference on Lasers and Electro-Optics* (Optical Society of America, 2010), p. CThJ1.
4. J. B. Heroux, T. Kise, M. Funabashi, T. Aoki, C. Schow, A. Rylyakov, and S. Nakagawa, "Energy-efficient 1060nm optical link operating up to 28 Gb/s," J. Light. Technol. 1–1 (2015).
5. C. Lam, H. Liu, B. Koley, X. Zhao, V. Kamalov, and V. Gill, "Fiber optic communication technologies: What's needed for datacenter network operations," Commun. Mag. IEEE **48**, 32–39 (2010).
6. O. Painter, "Two-Dimensional Photonic Band-Gap Defect Mode Laser," Science **284**, 1819–1821 (1999).
7. H.-G. Park, "Electrically Driven Single-Cell Photonic Crystal Laser," Science **305**, 1444–1447 (2004).
8. K. Takeda, T. Sato, A. Shinya, K. Nozaki, W. Kobayashi, H. Taniyama, M. Notomi, K. Hasebe, T. Kakitsuka, and S. Matsuo, "Few-fJ/bit data transmissions using directly modulated lambda-scale embedded active region photonic-crystal lasers," Nat. Photonics **7**, 569–575 (2013).
9. T. Baba and D. Sano, "Low-threshold lasing and Purcell effect in microdisk lasers at room temperature," IEEE J. Sel. Top. Quantum Electron. **9**, 1340–1346 (2003).
10. J. Van Campenhout, P. Rojo Romeo, P. Regreny, C. Seassal, D. Van Thourhout, S. Verstuyft, L. Di Cioccio, J.-M. Fedeli, C. Lagahe, and R. Baets, "Electrically pumped InP-based microdisk lasers integrated with a nanophotonic silicon-on-insulator waveguide circuit," Opt. Express **15**, 6744–6749 (2007).
11. M. T. Hill, Y.-S. Oei, B. Smalbrugge, Y. Zhu, T. de Vries, P. J. van Veldhoven, F. W. M. van Otten, T. J. Eijkemans, J. P. Turkiewicz, H. de Waardt, E. J. Geluk, S.-H. Kwon, Y.-H. Lee, R. Nötzel, and M. K. Smit, "Lasing in metallic-coated nanocavities," Nat. Photonics **1**, 589–594 (2007).
12. R. F. Oulton, V. J. Sorger, T. Zentgraf, R.-M. Ma, C. Gladden, L. Dai, G. Bartal, and X. Zhang, "Plasmon lasers at deep subwavelength scale," Nature **461**, 629–632 (2009).
13. K. Yu, A. Lakhani, and M. C. Wu, "Subwavelength metal-optic semiconductor nanopatch lasers," Opt. Express **18**, 8790–8799 (2010).
14. K. Ding, Z. C. Liu, L. J. Yin, M. T. Hill, M. J. H. Marell, P. J. van Veldhoven, R. Nöetzel, and C. Z. Ning, "Room-temperature continuous wave lasing in deep-subwavelength metallic cavities under electrical injection," Phys. Rev. B **85**, (2012).
15. M. Khajavikhan, A. Simic, M. Katz, J. H. Lee, B. Slutsky, A. Mizrahi, V. Lomakin, and Y. Fainman, "Thresholdless nanoscale coaxial lasers," Nature **482**, 204–207 (2012).

16. K. Ding and C. Z. Ning, "Fabrication challenges of electrical injection metallic cavity semiconductor nanolasers," Semicond. Sci. Technol. **28**, 124002 (2013).
17. E. K. Lau, A. Lakhani, R. S. Tucker, and M. C. Wu, "Enhanced modulation bandwidth of nanocavity light emitting devices," Opt. Express **17**, 7790–7799 (2009).
18. S. L. Chuang, *Physics of Photonic Devices*, Second (Wiley, 2009).
19. J. J. Sakurai, *Modern Quantum Mechanics*, Revised (Addison-Wesley, 1994).
20. E. Purcell, "Spontaneous Emission Probabilities at Radio Frequencies," Phys. Rev. **69**, 681 (1946).
21. T. Suhr, N. Gregersen, K. Yvind, and J. Mørk, "Modulation response of nanoLEDs and nanolasers exploiting Purcell enhanced spontaneous emission," Opt. Express **18**, 11230 (2010).
22. S. A. Sokolnikoff and H. T. Friis, *Antennas: Theory and Practice* (John Wiley & Sons, Ince., 1952).
23. H. A. Wheeler, "Fundamental Limitations of Small Antennas," Proc. IRE **35**, 1479–1484 (1947).
24. L. Novotny and N. van Hulst, "Antennas for light," Nat. Photonics **5**, 83–90 (2011).
25. A. Kinkhabwala, Z. Yu, S. Fan, Y. Avlasevich, K. Müllen, and W. E. Moerner, "Large single-molecule fluorescence enhancements produced by a bowtie nanoantenna," Nat. Photonics (2009).
26. O. L. Muskens, V. Giannini, J. A. Sánchez-Gil, and J. Gómez Rivas, "Strong Enhancement of the Radiative Decay Rate of Emitters by Single Plasmonic Nanoantennas," Nano Lett. **7**, 2871–2875 (2007).
27. R. M. Bakker, V. P. Drachev, Z. Liu, H. K. Yuan, R. H. Pedersen, A. Boltasseva, J. Chen, J. Irudayaraj, A. V. Kildishev, and V. M. Shalaev, "Nanoantenna array-induced fluorescence enhancement and reduced lifetimes," New J. Phys. **10**, 125022 (2008).
28. M. P. Busson, B. Rolly, B. Stout, N. Bonod, and S. Bidault, "Accelerated single photon emission from dye molecule-driven nanoantennas assembled on DNA," Nat. Commun. **3**, 962 (2012).
29. K. J. Russell, T.-L. Liu, S. Cui, and E. L. Hu, "Large spontaneous emission enhancement in plasmonic nanocavities," Nat. Photonics (2012).
30. V. J. Sorger, N. Pholchai, E. Cubukcu, R. F. Oulton, P. Kolchin, C. Borschel, M. Gnauck, C. Ronning, and X. Zhang, "Strongly Enhanced Molecular Fluorescence inside a Nanoscale Waveguide Gap," Nano Lett. **11**, 4907–4911 (2011).
31. W. Zhang, F. Ding, W.-D. Li, Y. Wang, J. Hu, and S. Y. Chou, "Giant and uniform fluorescence enhancement over large areas using plasmonic nanodots in 3D resonant cavity nanoantenna by nanoimprinting," Nanotechnology **23**, 225301 (2012).
32. J.-H. Song, T. Atay, S. Shi, H. Urabe, and A. V. Nurmikko, "Large Enhancement of Fluorescence Efficiency from CdSe/ZnS Quantum Dots Induced by Resonant Coupling to Spatially Controlled Surface Plasmons," Nano Lett. **5**, 1557–1561 (2005).
33. H. Mertens, J. S. Biteen, H. A. Atwater, and A. Polman, "Polarization-Selective Plasmon-Enhanced Silicon Quantum-Dot Luminescence," Nano Lett. **6**, 2622–2625 (2006).
34. A. G. Curto, G. Volpe, T. H. Taminiau, M. P. Kreuzer, R. Quidant, and N. F. van Hulst, "Unidirectional emission of a quantum dot coupled to a nanoantenna," Science **329**, 930–933 (2010).
35. J. S. Biteen, N. S. Lewis, H. A. Atwater, H. Mertens, and A. Polman, "Spectral tuning of plasmon-enhanced silicon quantum dot luminescence," Appl. Phys. Lett. **88**, 131109 (2006).

36. M. Pfeiffer, K. Lindfors, C. Wolpert, P. Atkinson, M. Benyoucef, A. Rastelli, O. G. Schmidt, H. Giessen, and M. Lippitz, "Enhancing the Optical Excitation Efficiency of a Single Self-Assembled Quantum Dot with a Plasmonic Nanoantenna," Nano Lett. **10**, 4555–4558 (2010).
37. X.-L. Chu, T. J. K. Brenner, X.-W. Chen, Y. Ghosh, J. A. Hollingsworth, V. Sandoghdar, and S. Götzinger, "Experimental realization of an optical antenna designed for collecting 99% of photons from a quantum emitter," Optica **1**, 203 (2014).
38. A. Neogi, C.-W. Lee, H. Everitt, T. Kuroda, A. Tackeuchi, and E. Yablonovitch, "Enhancement of spontaneous recombination rate in a quantum well by resonant surface plasmon coupling," Phys. Rev. B **66**, (2002).
39. K. Okamoto, I. Niki, A. Shvartser, Y. Narukawa, T. Mukai, and A. Scherer, "Surface-plasmon-enhanced light emitters based on InGaN quantum wells," Nat. Mater. **3**, 601–605 (2004).
40. D. Fattal, M. Fiorentino, M. Tan, D. Houng, S. Y. Wang, and R. G. Beausoleil, "Design of an efficient light-emitting diode with 10 GHz modulation bandwidth," Appl. Phys. Lett. **93**, 243501 (2008).
41. C.-H. Cho, C. O. Aspetti, M. E. Turk, J. M. Kikkawa, S.-W. Nam, and R. Agarwal, "Tailoring hot-exciton emission and lifetimes in semiconducting nanowires via whispering-gallery nanocavity plasmons," Nat. Mater. **10**, 669–675 (2011).
42. J. N. Farahani, D. W. Pohl, H.-J. Eisler, and B. Hecht, "Single Quantum Dot Coupled to a Scanning Optical Antenna: A Tunable Superemitter," Phys. Rev. Lett. **95**, (2005).
43. P. Bharadwaj and L. Novotny, "Spectral dependence of single molecule fluorescence enhancement," Opt Express **15**, 14266–14274 (2007).
44. S. Kuhn, G. Mori, M. Agio, and V. Sandoghdar, "Modification of single molecule fluorescence close to a nanostructure: radiation pattern, spontaneous emission and quenching," Mol. Phys. **106**, 893–908 (2008).
45. K.-G. Lee, H. Eghlidi, X.-W. Chen, A. Renn, S. Götzinger, and V. Sandoghdar, "Spontaneous emission enhancement of a single molecule by a double-sphere nanoantenna across an interface," Opt. Express **20**, 23331 (2012).
46. M. Barth, S. Schietinger, S. Fischer, J. Becker, N. Nüsse, T. Aichele, B. Löchel, C. Sönnichsen, and O. Benson, "Nanoassembled Plasmonic-Photonic Hybrid Cavity for Tailored Light-Matter Coupling," Nano Lett. **10**, 891–895 (2010).
47. D. Arbel, N. Berkovitch, A. Nevet, A. Peer, S. Cohen, D. Ritter, and M. Orenstein, "Light emission rate enhancement from InP MQW by plasmon nano-antenna array," Opt. Express **19**, (2011).
48. K. C. Y. Huang, M.-K. Seo, Y. Huo, T. Sarmiento, J. S. Harris, and M. L. Brongersma, "Antenna electrodes for controlling electroluminescence," Nat. Commun. **3**, 1005 (2012).
49. M. S. Eggleston, K. Messer, L. Zhang, E. Yablonovitch, and M. C. Wu, "Optical antenna enhanced spontaneous emission," Proc. Natl. Acad. Sci. 201423294 (2015).
50. K. C. Y. Huang, M.-K. Seo, T. Sarmiento, Y. Huo, J. S. Harris, and M. L. Brongersma, "Electrically driven subwavelength optical nanocircuits," Nat. Photonics **8**, 244–249 (2014).
51. S. Ramo, "Currents Induced by Electron Motion," Proc. IRE **27**, 584–585 (1939).
52. D. J. Griffiths, *Introduction to Electrodynamics*, 3rd ed (Prentice Hall, 1999).
53. C. B. Moore, W. Rison, J. Mathis, and G. Aulich, "Lightning rod improvement studies," J. Appl. Meteorol. **39**, 593–609 (2000).

54. M. Staffaroni, J. Conway, S. Vedantam, J. Tang, and E. Yablonovitch, "Circuit analysis in metal-optics," Photonics Nanostructures - Fundam. Appl. **10**, 166–176 (2012).
55. J. D. Jackson, *Classical Electrodynamics*, Third (John Wiley & Sons, Inc., 1999).
56. J. Lekner, "Capacitance coefficients of two spheres," J. Electrost. **69**, 11–14 (2011).
57. B. Gelmont and M. Shur, "Spreading resistance of a round ohmic contact," Solid-State Electron. **36**, 143–146 (1993).
58. John D. Kraus, *Antennas*, Second (McGraw-Hill, Inc., 1988).
59. P. W. Gilberd, "The anomalous skin effect and the optical properties of metals," J. Phys. F Met. Phys. **12**, 1845–1860 (1982).
60. T. J. Seok, A. Jamshidi, M. Kim, S. Dhuey, A. Lakhani, H. Choo, P. J. Schuck, S. Cabrini, A. M. Schwartzberg, J. Bokor, E. Yablonovitch, and M. C. Wu, "Radiation Engineering of Optical Antennas for Maximum Field Enhancement," Nano Lett **11**, 2606–2610 (2011).
61. T. J. Seok, A. Jamshidi, M. Eggleston, and M. C. Wu, "Mass-producible and efficient optical antennas with CMOS-fabricated nanometer-scale gap," Opt. Express **21**, 16561–16569 (2013).
62. S. Sakai, M. Umeno, and Y. Amemiya, "Measurement of Diffusion Coefficient and Surface Recombination Velocity for p-InGaAsP Grown on InP," Jpn. J. Appl. Phys. **19**, 109–113 (1980).
63. M. Fukuda, "Current drift associated with surface recombination current in InGaAsP/InP optical devices," J. Appl. Phys. **59**, 4172–4176 (1986).
64. E. Yablonovitch, "Statistical ray optics," JOSA **72**, 899–907 (1982).
65. T. P. Pearsall, *GaInAsP Alloy Semiconductors* (John Wiley & Sons, 1982).
66. E. F. Schubert, *Light-Emitting Diodes*, Second Edition (Cambridge University Press, 2006).
67. T. Kosako, Y. Kadoya, and H. F. Hofmann, "Directional control of light by a nano-optical Yagi–Uda antenna," Nat. Photonics **4**, 312–315 (2010).
68. D. Dregely, R. Taubert, J. Dorfmüller, R. Vogelgesang, K. Kern, and H. Giessen, "3D optical Yagi-Uda nanoantenna array," Nat. Commun. **2**, 267 (2011).
69. I. S. Maksymov, I. Staude, A. E. Miroshnichenko, and Y. S. Kivshar, "Optical Yagi-Uda nanoantennas," Nanophotonics **1**, (2012).
70. C. A. Balanis, *Antenna Theory: Analysis and Design*, Third (John Wiley & Sons, Inc., 2005).
71. H. C. Casey and E. Buehler, "Evidence for low surface recombination velocity on n-type InP," Appl. Phys. Lett. **30**, 247 (1977).
72. H. G. Booker, "Slot aerials and their relation to complementary wire aerials (Babinet's principle)," Electr. Eng.-Part IIIA Radiolocation J. Inst. Of **93**, 620–626 (1946).
73. R. C. Hansen, "Slot antenna in a resistive screen," IEEE Trans. Antennas Propag. **46**, 1028–1031 (1998).
74. F. T. Ulaby, *Fundamentals of Applied Electromagnetics*, 6th ed (Prentice Hall, 2010).
75. S. A. Fortuna, M. Eggleston, K. Messer, E. Yablonovitch, and M. C. Wu, "Electrically injected nanoLED with enhanced spontaneous emission from a cavity backed optical slot antenna," in (2014).
76. S. A. Fortuna, *Exploring a New Frontier of Cavity-Backed Slot Antennas in the 21st Century* (UC Berkeley, 2014).
77. E. Yablonovitch, R. Bhat, C. Zah, T. Gmitter, and M. Koza, "Nearly ideal InP/In0. 53Ga0. 47As heterojunction regrowth on chemically prepared In0. 53Ga0. 47As surfaces," Appl. Phys. Lett. **60**, 371–373 (1992).

78. A. B. Sproul, "Dimensionless solution of the equation describing the effect of surface recombination on carrier decay in semiconductors," J. Appl. Phys. **76**, 2851 (1994).
79. L. A. Coldren, *Diode Lasers and Photonic Integrated Circuits*, 2nd ed, Wiley Series in Microwave and Optical Engineering No. 218 (Wiley, 2012).
80. A. K. Geim and K. S. Novoselov, "The rise of graphene," Nat. Mater. **6**, 183–191 (2007).
81. K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, "Atomically Thin MoS2: A New Direct-Gap Semiconductor," Phys. Rev. Lett. **105**, (2010).
82. Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, and M. S. Strano, "Electronics and optoelectronics of two-dimensional transition metal dichalcogenides," Nat. Nanotechnol. **7**, 699–712 (2012).
83. N. Peimyoo, J. Shang, C. Cong, X. Shen, X. Wu, E. K. L. Yeow, and T. Yu, "Nonblinking, Intense Two-Dimensional Light Emitter: Monolayer $WS_2$ Triangles," ACS Nano **7**, 10985–10994 (2013).
84. W. Zhao, Z. Ghorannevis, L. Chu, M. Toh, C. Kloc, P.-H. Tan, and G. Eda, "Evolution of Electronic Structure in Atomically Thin Sheets of $WS_2$ and $WSe_2$," ACS Nano **7**, 791–797 (2013).
85. J. S. Ross, P. Klement, A. M. Jones, N. J. Ghimire, J. Yan, D. G. Mandrus, T. Taniguchi, K. Watanabe, K. Kitamura, and W. Yao, "Electrically Tunable Excitonic Light Emitting Diodes based on Monolayer WSe2 pn Junctions," ArXiv Prepr. ArXiv13121435 (2013).
86. S. Jo, N. Ubrig, H. Berger, A. B. Kuzmenko, and A. F. Morpurgo, "Mono- and Bilayer $WS_2$ Light-Emitting Transistors," Nano Lett. **14**, 2019–2025 (2014).
87. R. Cheng, D. Li, H. Zhou, C. Wang, A. Yin, S. Jiang, Y. Liu, Y. Chen, Y. Huang, and X. Duan, "Electroluminescence and Photocurrent Generation from Atomically Sharp $WSe_2$ /$MoS_2$ Heterojunction *p–n* Diodes," Nano Lett. **14**, 5590–5597 (2014).
88. A. Pospischil, M. M. Furchi, and T. Mueller, "Solar-energy conversion and light emission in an atomic monolayer p–n diode," Nat. Nanotechnol. **9**, 257–261 (2014).
89. O. Lopez-Sanchez, D. Lembke, M. Kayci, A. Radenovic, and A. Kis, "Ultrasensitive photodetectors based on monolayer MoS2," Nat. Nanotechnol. **8**, 497–501 (2013).
90. B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, "Single-layer MoS2 transistors," Nat. Nanotechnol. **6**, 147–150 (2011).
91. H. Shi, R. Yan, S. Bertolazzi, J. Brivio, B. Gao, A. Kis, D. Jena, H. G. Xing, and L. Huang, "Exciton Dynamics in Suspended Monolayer and Few-Layer $MoS_2$ 2D Crystals," ACS Nano **7**, 1072–1080 (2013).
92. Q. Cui, F. Ceballos, N. Kumar, and H. Zhao, "Transient Absorption Microscopy of Monolayer and Bulk $WSe_2$," ACS Nano **8**, 2970–2976 (2014).
93. T. Yan, X. Qiao, X. Liu, P. Tan, and X. Zhang, "Photoluminescence properties and exciton dynamics in monolayer WSe2," Appl. Phys. Lett. **105**, 101901 (2014).
94. M. M. Benameur, B. Radisavljevic, J. S. Héron, S. Sahoo, H. Berger, and A. Kis, "Visibility of dichalcogenide nanolayers," Nanotechnology **22**, 125706 (2011).
95. H. Fang, C. Battaglia, C. Carraro, S. Nemsak, B. Ozdol, J. S. Kang, H. A. Bechtel, S. B. Desai, F. Kronast, A. A. Unal, G. Conti, C. Conlon, G. K. Palsson, M. C. Martin, A. M. Minor, C. S. Fadley, E. Yablonovitch, R. Maboudian, and A. Javey, "Strong interlayer coupling in van der Waals heterostructures built from single-layer chalcogenides," Proc. Natl. Acad. Sci. **111**, 6198–6202 (2014).

96. L. Cao, J. S. White, J.-S. Park, J. A. Schuller, B. M. Clemens, and M. L. Brongersma, "Engineering light absorption in semiconductor nanowire devices," Nat. Mater. **8**, 643–647 (2009).