EXPLORING THE LIMITS OF NEAR-FIELD FLUORESCENCE MICROSCOPY: TOWARD MOLECULAR-SCALE IMAGING OF BIOLOGICAL SYSTEMS

by

Benjamin Daniel Mangum

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Physics and Astronomy

The University of Utah

May 2010

STATEMENT OF DISSERTATION APPROVAL

The dissertation of [redacted] **Daniel** [redacted]

has been approved by the following supervisory committee members:

Jordan Gerton	, Chair	**March 17, 2010**
[redacted]	, Member	**March 11, 2010**
[redacted]	, Member	**March 11, 2010**
John [redacted]	, Member	**March 11, 2010**
Markus Babst	, Member	**March 11, 2010**

and by [redacted], Chair of

the Department of [redacted] **and** [redacted]

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

This dissertation describes the advancements made towards the implementation of Tip-Enhanced Fluorescence Microscopy (TEFM) in imaging biological specimens. This specialized type of microscopy combines the chemical specificity of optical microscopy techniques with the resolution of atomic force microscopy (AFM). When an AFM probe is centered in the focal spot of an excitation laser with axial polarization, the probe concentrates the optical field such that it can be used to induce nanometer scale fluorescence.

The physical mechanisms of this optical field enhancement are set forth in detail. The feasibility of this technique for imaging bimolecular networks is discussed in regard to the requirements for adequate image contrast, as well as for obtaining field enhancement in aqueous environments. A semianalytical model for image contrast for TEFM has been developed. This model shows that using demodulation techniques greatly increases the image contrast attainable with this technique, and is capable of predicting the requisite enhancement factors to achieve imaging of biomolecular networks at good contrast levels. This model predicts that signal enhancement factors on the order of 20 are needed to image densely packed samples.

This dissertation also highlights a novel tomographical imaging approach. By timestamping the fluorescence photon arrival times, and subsequently correlating them to the timestamped motion of a vertically oscillating probe, a three-dimensional map of tip-sample interactions can be constructed. The culmination of these advancements has led to the ability to map the interactions between single carbon nanotubes and single fluorescent nanocrystals (quantum dots). Various attempts at using TEFM in water have been thus far unsuccessful. Several explanations for this shortfall have been identified—understanding these shortcomings has helped to identify the optimal excitation conditions for field enhancement.

For my patient wife, Dawn.

CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

First and foremost I would like to thank Dr. Jordan Gerton, who has been equal parts advisor and friend. He has provided a positive work environment—it has always been a pleasure to go to work with Jordan as a boss. His efforts in procuring grant money to keep our lab running and the paychecks coming are truly appreciated. Thanks especially for allowing me so much creative freedom in my research, giving me a chance to reason out what the logical next step might be.

Many thanks go out to the various lab members I have worked with throughout my time in the lab. Thanks to my predecessors in the lab, Chang-an Xie, Chun Mu, and Jon Cox, who helped train me on the system. Several undergraduates in the lab offered help along the way: Jason Ho, Jeff Jensen, and Rachel Anderson. Thanks also to the more recent students, Ben Martin, Carl Ebeling and Anil Ghimire, for their helpful discussions and willingness to help in any way possible. A special thanks to Carl who has also provided his expertise in the preparation and plotting of many of the figures presented in this work. Charles McGuire did the thankless dirty work in the lab of providing a never-ending supply of glass coverslips, clean glassware, and custom machined parts, as well as for running my time-consuming COMSOL simulations. I appreciate Jessica Johnston for her nanotube troubleshooting and help in preparing a recent book chapter, which was a starting point for this dissertation. Thanks to graduate student collaborators in the biology department: Shigeki Watanabe for his worm samples and Betsey Ott has been a top notch consultant for quite some time. Last but not least, thanks to Eyal Shafran, who has been the cause of and solution to many of the problems in the lab. He has been instrumental in the development of our data acquisition and analysis programs, for which I am indebted. I have always enjoyed our many heated arguments, which in the end have yielded both better science and hopefully

better scientists.

I could not have made it this far in my academic career with an always supportive family. Thank you to my parents for providing my family a home. Thanks also to loving grandparents who have always shown interest in my work. A special thanks to my sister Liz, who has been an excellent proofreader. My super-great wife Dawn, who has always supported me in all my endeavors, deserves thanks most of all for at least pretending to care about the details of what I do in the lab each day. Her willingness to take care of life at home when I was working long hours has definitely not gone unnoticed.

CHAPTER 1

INTRODUCTION

This dissertation will discuss my research involving a rather specialized form of microscopy, which falls into the general category of near-field optics. While the term "microscope" is familiar to most people, the type of instrument I use may seem quite far removed from the mental picture most people envision. In that regard, I will give a brief overview of what exactly constitutes a microscope and how my device fits into that framework. This chapter will provide a brief introduction to near-field microscopy, accessible to both new students in the lab as well as non-physicists, who will be able to gain a better understanding of my research. The following chapters will detail the experimental apparatus used and the basic principles behind near-field microscopy; these chapters will require some background in physics/optics. The latter chapters in this work will include much of the results, data, and conclusions I have produced while working over the last several years.

1.1 Motivation

Biological cells fabricate and assemble proteins and other biomolecules into diverse networks with striking complexity and functionality. Such networks are critical components in the complicated machinery of the cell as they participate in a host of cellular functions and are important for protection against many diseases. In general, the structure of the networks plays a large role in their functionality. Thus, nanoscale-resolution imaging of intact networks in their native conformation should yield information that could create the ability to optimize specific cellular functions, to engineer new functions, and to strengthen a cell's defense against disease. Further, these networks are prototypical nanosystems and should be studied in detail

for insight into the rational design of synthetic molecular systems for a multitude of technological applications. To study biomolecular systems in this context, it is crucial to observe their molecular machinery at work in a physiologically relevant environment. Currently, there are no techniques that can accomplish this.

While existing techniques such as electron microscopy, x-ray crystallography, and nuclear magnetic resonance spectroscopy can yield structural information with exquisite detail, these techniques are not well suited for in vitro studies of complex molecular networks. In comparison, optical microscopy is minimally perturbative and is routinely used under physiological conditions. Fluorescence microscopy in particular is very powerful for studying biological systems because it can be used to detect and image single molecules and to distinguish between the chemical identities of those molecules based on their spectrum (color). The major limitation of traditional optical microscopy is the limited resolution ($\sim$250 nm) imposed by classical light diffraction.

A new type of imaging system was organized at Caltech in the lab of Stephen Quake several years ago that was able break the diffraction limit and achieve optical resolution below 10 nm [1, 2, 3]. That system combined the benefits of scanning probe microscopy and optical microscopy. This work is really an extension of their original work. Our goal in the lab has been to further develop and refine a similar microscope system that is capable of imaging biomolecular networks at physiologically relevant conditions, at a resolution comparable to that of single proteins (< 10 nm).

1.2 What Is a Microscope?

The term microscope comes from the Greek *mikrós*, meaning "small", and *skopeîn*, meaning "to see," or simply a device to see small things. In science terminology the prefix "micro" refers to a unit of measure meaning 10^{-6}, or one part in a million; one micrometer (a.k.a. micron) is one millionth of a meter. To appreciate how small this is, the width of a human hair is about 100 micrometers (100 μm), a red blood cell has a diameter of about 10 μm, and spider silk has a

width of 3 μm. In this respect, items that are roughly a "micro"-meter in size are too small to be seen without the aid of a "micro"-scope. In fact, using the same convention, the near-field microscope employed in our lab can really be considered to be a "nanoscope" in that it allows for visualization of things on the scale of one nanometer (10^{-9} m or one billionth of a meter).

Perhaps the simplest of all microscopes is a magnifying glass. Light reflecting off a sample travels through a lens where the resulting image is magnified. Much more complicated systems using multiple lenses have been developed to achieve very high magnifications; the magnification of most commercially available microscopes can easily be over 1000×, meaning that the object would appear 1000 times larger under the microscope than it would with the naked eye.

1.3 Resolution

As mentioned, commercially available light microscopes can easily provide magnification of over 1000×; however, a related but separate issue is that of resolution. Resolving power for a lens refers to its ability to distinguish detail. The classic example is that of a car's headlights. As an oncoming vehicle approaches at night from very far away the headlights are noticeable; however, if the car is sufficiently far away it is impossible to tell if it is a car with two headlights or rather a motorcycle with one. As the vehicle approaches, it can be seen that the two lights gradually start to appear to be separate. This phenomenon is due to the limited resolving power of the human eye.

Many people have some experience with digital cameras; once a picture is taken the image can be magnified to your heart's content on a computer. However, at some point the fine details become less clear. As seen in Figure 1.1, a portion of the image has been magnified so it is seven times larger than the original. The magnified portion of the image, while larger, provides no new information about the object. Magnification can be essentially unlimited and is easily achieved, although without resolution to go along with it, magnification can be useless. Magnification of digital images can be done after the fact on a computer. Resolution on the other

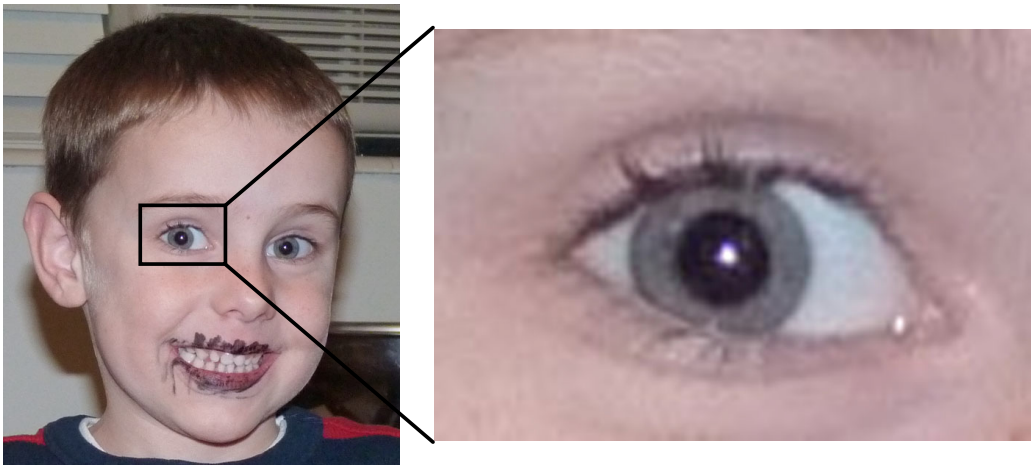

Figure 1.1. Demonstration of the difference between magnification and resolution. The picture of my son on the left is at full size, while on the right his eye has been magnified by a factor of 7. Notice that the magnified image, while larger, does not provide any extra detail not already seen in the original image.

hand is expensive, and relies upon the lens used to take the picture. Likewise, in microscopes magnification is cheap: rather inexpensive systems can boast high magnification. On the other hand, resolution is expensive: it is determined only by the lens closest to the sample, which is responsible for gathering light (the information) and is called the objective lens. Now we have some small sample that we would like to put under our microscope to observe, such as very small grains of sand spread out on a piece of glass. Should we want to look in the microscope to count the grains of sand, an important question to ask is: how powerful of a microscope is needed? The answer is based not strictly on the size of the grains but also the spacing. For counting, we are not interested in the structure of each grain; we simply want to know how many there are. To accurately count them all, the resolving power of the microscope will need to be better than the smallest distance between two grains, lets call it ΔL. Lord Rayleigh calculated that the best resolution a microscope can give is based on the properties of its objective lens. Namely,

$$\Delta L = \frac{0.61\lambda}{NA}, \tag{1.1}$$

where λ is the wavelength, or color, of the light being used, and NA is the Numerical Aperture of the lens. More often, this resolution limit is expressed in terms of the Numerical Aperture (NA) of the objective lens. Simply put, the numerical aperture defines the light gathering ability of a lens. Again think of a magnifying glass, this time used to burn ants. While a bit sadistic, it is easy to visualize. To efficiently burn an ant, a high NA lens is desirable as it would gather a lot of sunlight and concentrate it to a tight (and deadly) spot. A more technical definition of NA is actually:

$$NA = n \sin\theta, \tag{1.2}$$

where n is the index of refraction of the media the lens operates in[1], and θ is the maximum angle from which light rays can enter the lens. Numerical aperture is related to the diameter of the lens D, as well as its focal length (f):

$$NA \approx n\frac{D}{2f} \tag{1.3}$$

This can be visualized via Figure 1.2; formally $\tan\theta = r/f$, where r is the radius of the lens, θ is the angle indicated, and f is the focal length of the lens. For small angles, $\tan\theta \approx \sin\theta$, which leads to the solution of 1.3. Thus Rayleigh's formula can also be written as:

$$\Delta L = \frac{1.22\lambda f}{D}, \tag{1.4}$$

which assumes imaging in air ($n_{air} = 1.0$). Camera buffs will recognize the ratio f/D as the $f/\#$ of the lens.

Notice that the diameter of the lens is important—the bigger the diameter, the better the achievable resolution. This is precisely why professional camera equipment seems to be so large and produces such nice results. Of course as the size of a lens is increased, the cost of the lens rises even faster. Producing and polishing high quality optics gets more expensive as the size increases, as the smoothness of

[1]The index of refraction can be thought of as a measure of the ability of a material to bend light. Technically it is a measure of how fast light travels in that media (v) relative to the speed of light in a vacuum (c); $n = c/v$.

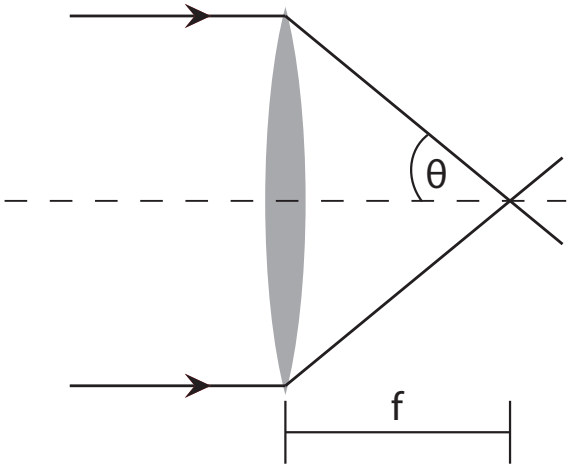

Figure 1.2. Raytracing diagram for a simple lens. Rays parallel to the optic axis (dashed line) are focused at a distance away from the lens called the focal point, f. The maximum angle θ is given by rays originating from the extremities of the lens.

a quality lens should be much less than the wavelength of light, or should not have bumps over several tens of nanometers.

The NA of a microscope objective is a very important parameter in microscopy, as it ultimately determines the maximum achievable resolution. As seen in Equation 1.2 it is limited by the index of refraction of the media, even if all possible light is collected $\theta = 90^\circ \Rightarrow \sin\theta = 1$. The most common media used in microscopy are air ($n_a = 1.0$), water ($n_w = 1.33$), and oil ($n_o = 1.5$), and thus the best objectives have NA's of 0.95, 1.20, and 1.49 for air, water, and oil respectively. Wavelengths for visible light range from 400 nm (blue) to 700 nm (red). Assuming blue light and an oil immersion numerical aperture of $NA = 1.49$, this limits the resolution of the best optical microscopes to $\sim 164nm$. In practice such values are rarely seen as often the wavelength of light employed is much longer; more commonly, the resolution limit is generally held to be around 250 nm.

In this manner, the resolution capabilities of a microscope is ultimately determined by the light gathering abilities of the objective lens. As the NA of the microscope objective is increased, more of the light information from the sample is collected—allowing for a more complete image reconstruction of the sample at the detector. Thus as the NA of a microscope objective is increased the better the

resolution will be; however, due to diffraction, image patterns become distorted.

1.3.1 Diffraction

The ability of all waves, be they water, sound, light etc., to bend around edges is known as diffraction. To quote the famous text by Landau and Lifshitz [4]:

> Diffraction phenomena can be observed, for example, if along the path of propagation of the light there is an obstacle—an opaque body (we call it a screen) of arbitrary form. If the laws of geometrical optics were strictly satisfied, there would be beyond the screen regions of "shadow" sharply delineated from regions where the light falls. The diffraction has the consequence that, instead of a sharp boundary between light and shadow, there is a quite complex distribution of the intensity of the light. These diffraction phenomena appear the more strongly the smaller the dimensions of the screens and the apertures in them, or the greater the wavelength.

Diffraction phenomena are visible in many everyday occurrences. One example is that of waves entering a marina as seen in Figure 1.3. Marinas have rock barriers shielding the ships docked inside from large waves. As planar waves enter the marina through a narrow opening, the waves are diffracted, creating a complicated pattern filling the entire marina. Notice how the waves stop traveling with planar wave fronts upon entering the marina; due to diffraction, the wave fronts propagate in a circular pattern.

In optics, the ability of a lens to focus light to a spot depends in part on the diffraction limit. Since light waves impinging on a lens are diffracted, the light can not be focused to an infinitely tight spot. Consider how the light impinging on the edges of the lens would tend to bend around the edge. Due to the bending of the light around the edges of a lens, a particular pattern known as an Airy pattern emerges (cf. Fig. 1.4), which has a central bright spot surrounded by faint rings of lower intensity.

This diffraction limit goes both ways: it limits the ability of a lens to focus light, but also the ability of a lens to image small objects. This means that an isolated infinitesimally small object when imaged by even the best of microscopes would appear as an Airy pattern, whose dimensions would depend on the objective

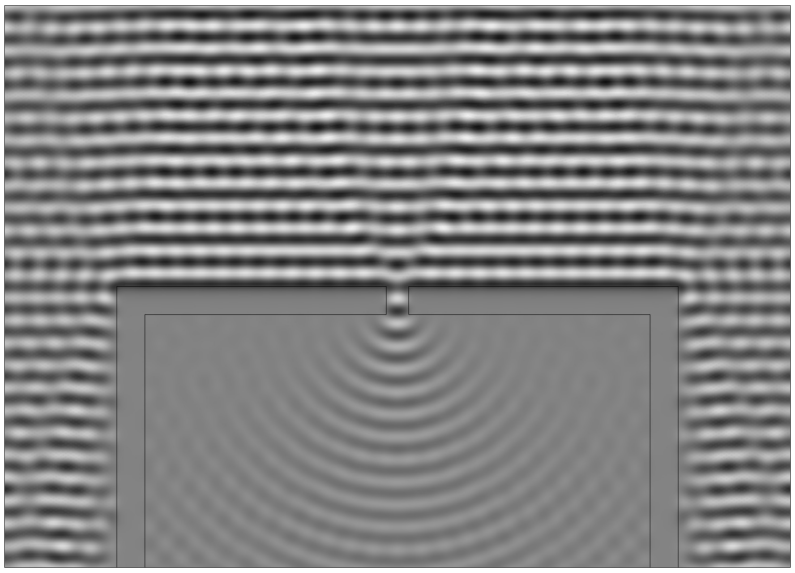

Figure 1.3. Computer generated diffraction pattern of plane waves entering a marina. Notice how the wave diffracts through the entrance of the marina. Also observable is diffraction around the outer corners of the rock walls.

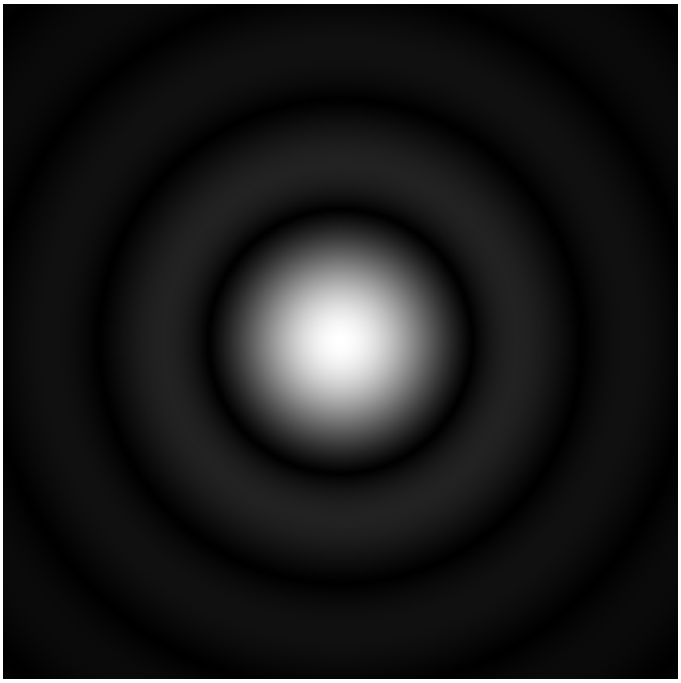

Figure 1.4. An example of an Airy pattern, representative of a typical signal in a diffraction limited imaging system. The diameter of the central disk is given by Eq. 1.5. Here a nonlinear colorscale is used to give emphasis to the outer rings.

lens being used. In fact, the full width of the central disk of the airy pattern has a diameter (D) given by:[2]

$$D = \frac{1.22\lambda}{NA}. \tag{1.5}$$

For example, there is a 3-nm large single protein that is a sitting on a piece of glass that we would like to image. We put it under a microscope using a very nice objective (NA = 1.4) using red light, and the image we would get is not a 3-nm spot, but an Airy pattern (as seen in Fig. 1.4) with a diameter of $D = 1.22 \cdot 600nm/1.4 \approx 525nm$.

As Ernst Abbe was the first to estimate the theoretical diffraction limit (1872), it often bears his name (Abbe Limit). He calculated the smallest spot size would have a radius r given by:

$$r = \frac{0.61\lambda}{NA}. \tag{1.6}$$

The factor 1.22 is specific to circular openings, and takes complicated mathematics to be able to derive. If this equation seems very similar to Rayleigh's, it is because Rayleigh used the diffraction limited spot size to determine the resolution limit.

1.3.2 Resolution Criteria

Lord Rayleigh used the known result from the Airy function and put a limit on what was considered resolvable. The limit he decided upon, for sake of simplicity, was to say when two identical Airy patterns are closer than the radius of the central disk, they are no longer resolvable. This is demonstrated in Figure 1.5 where two Airy patterns are brought increasingly close together. As seen, the point at which they can no longer be distinguished as two separate objects is actually a little ambiguous. In fact other criteria for minimally resolvable objects have also been established. For example the Sparrow criterion states that the two objects are resolvable as long as there is a saddle point between them. In other words, as long as the combined intensity profile contains a central dip, the objects are said

[2]There is often some confusion with the diameter of the diffraction limited spot. The diameter D given here is correct, often the Full Width Half Maximum (FWHM) of the spot is reported, which is close to the *radius* of the spot r, but not quite: $d_{FWHM} \approx 0.84 \cdot r$ for an Airy profile.

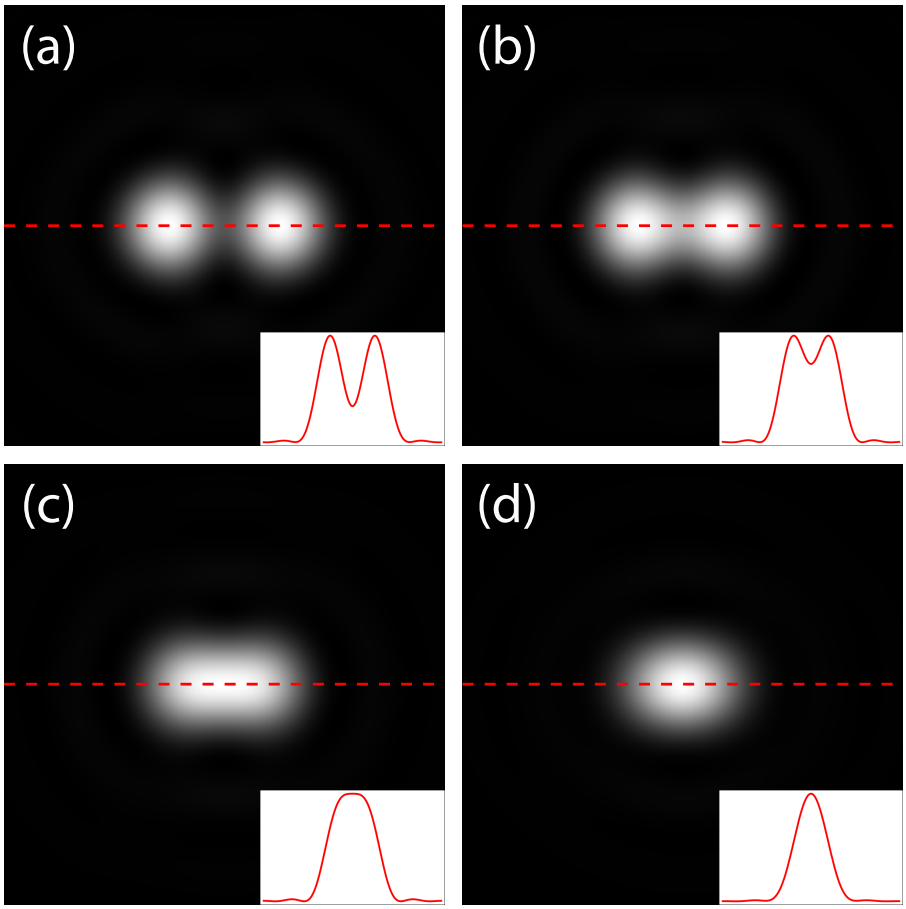

Figure 1.5. Examples of two identical Airy disks separated by various distances in terms of the radius of the disk $r = 0.61\lambda/NA$. Panel (a): Airy disks separated by $1.3 \times r \Rightarrow$ clearly resolved. Panel (b): Airy disks separated by $r \Rightarrow$ the Rayleigh Criterion is just met. Panel (c): Airy disks separated by $0.77 \times r \Rightarrow$ the Sparrow criterion is just met. Panel (d): Airy disks separated by $0.5 \times r \Rightarrow$ unresolved.

to be resolvable. Without worrying about the mathematical definition of saddle point—the criterion is also illustrated in Figure 1.5(c). The Sparrow criterion is widely regarded as a more natural definition of resolution. The diffraction limit puts a lower bound of $\sim$ 250 nm, meaning that features smaller than this cannot be resolved. This presents a major hurdle for scientists trying to understand the world on the nanoscale.

1.4 Classification of Microscopes

A microscope that relies on light and utilizes lenses to magnify an image is an optical microscope or light microscope. While this may seem trivial, there are actually many different types of microscopes, some of which do not rely on light at all. Trying to categorize all existing microscopes would be a daunting task in that there are myriad variations and combinations of different techniques; however, in general there are three main classes: optical (or light) microscopes, electron microscopes, and scanning probe microscopes. Electron microscopes use beams of electrons that are directed onto a sample. Images can be acquired by measuring the electrons that either scatter off the sample or penetrate through it. Scanning probe microscopes use sharp probes that scan over the surface of a sample to record the topography (among other things). Each of these has specific advantages and disadvantages. Light microscopes offer supreme chemical specificity, allowing for distinction of individual types molecules within dense ensembles of biological networks, however, with limited spatial resolution. Electron microscopes can offer excellent resolution, their downfall being a lack of chemical specificity and incompatibility with live biological specimens. Scanning probe microscopes provide resolution on almost the same scale as electron microscopy, and can be compatible with live biological specimens, but yet can only be used to probe the surface of a sample. This section will discuss how it is that light microscopes achieve their high chemical specificity, and the importance of having such a trait.

1.4.1 Fluorescence

Many types of optical microscopes rely on a physical process called fluorescence. Fluorescence refers to an object's ability to absorb light, only to re-emit the light at a later time. Generally speaking the light that is re-emitted is not even the same color as the absorbed light: it is less energetic (more red) and is called red-shifted. Furthermore, the light that is re-emitted is at a later time—just a very brief time later - on the order of nanoseconds (10^{-9} s). Objects that exhibit this property are fluorescent, and the part of the object where this process actually takes place is

called a fluorophore.

Many substances can be fluorescent; dye molecules, single proteins, and nano-sized crystals called quantum dots (QDs) are some of the most important fluorescent substances used in scientific research. The real power of fluorescence microscopy is that these fluorophores can be attached to virtually anything with great specificity. The attachment of quantum dots or dye molecules can be done through chemical means, where after a series of reactions, some specimen of interest can be "tagged" or "labeled" with a fluorophore. Scientists have also figured out how to label proteins with other fluorescent proteins using genetic methods. In fact, this technique is so important that the pioneers of this method were recently awarded the 2008 Nobel prize in chemistry.

Some types of jellyfish (among other animals) naturally produce fluorescent proteins in their body called Green Fluorescent Protein (GFP). Scientists have the ability to remove this jellyfish DNA, which contains the code for GFP, and attach it to the DNA of some other cell. In this manner scientists can determine the function of certain parts of cells with extremely high specificity. Biologists interested in the function of a certain part of a cell - perhaps a cell membrane protein, could take the jellyfish DNA for GFP attach it to the DNA that encodes for the cell membrane protein, and inject this new DNA construct into an embryo. As it develops all the cell membranes would also have GFP attached, and thus also be fluorescent. This genetically modified cell/animal is known as a mutant. At this point, the mutant is put under a microscope, it is illuminated with blue light, and the tagged cell membrane proteins will fluoresce green. Thus, the fluorescent green light will show the location of the cell membrane proteins.

1.4.2 Chemical Specificity

Given this unique ability to label virtually any sample of interest, fluorescence microscopy has an extremely high chemical specificity. Some microscopes offer very high resolution, but with low specificity. Scanning probe microscopes can provide very high resolution ($\sim$ 1 nm) details of surface topography, but without some advanced knowledge of what to expect, all that is really seen is just bumps on a

surface. As scanning probes do not use light, they are not subject to the diffraction limit. The resolution of scanning probes depends only upon the shape of the probe used. Extremely sharp (narrow) probes can accurately provide information on the smallest contours of a surface. Dull (fat) probes may not be able to fit into small crevasses, and tend to broaden and average out surface details. While scanning probe microscopy can offer exquisite detail of a surface, the microscope itself cannot determine what the bumps mean. This can be problematic as there is no way to differentiate between two similar sized objects. There is no way to determine which object is which. However, one nice feature of scanning probe microscopy is that it can be compatible with biological specimens, even live cells.

Electron microscopes can yield even better resolution (< 1 nm), but again, with low chemical specificity. From a physicist's standpoint, electrons can also be thought of as waves, with an extremely small effective wavelength that depends on their velocity; commercially available systems can easily reach effective wavelengths of ~ 0.004 nm. Due to this small wavelength tremendous resolution can be obtained, but the information conveyed is primarily about the density of the sample. Again causing objects of similar sizes and densities to be indistinguishable. Unfortunately, electron microscopy is extremely incompatible with imaging live biological specimens. To provide adequate image contrast, cells must be fixed and embedded in plastics and treated with harsh chemicals, thus live cell imaging is completely out of the question.

The great virtue of light microscopy is that extremely high chemical specificity is obtainable through fluorescence labeling; different objects of similar sizes can be labeled with different colored fluorophores, making them now distinguishable. The ideal microscope then is one that can take advantage of the tremendous chemical specificity afforded by fluorescence techniques, that is not subject to the limits of diffraction, as detailed in Figure 1.6. This is the goal that I have been working toward throughout my graduate career: to combine the utilities of scanning probe microscopy with light microscopy to allow for imaging of biological samples. The hybrid technique is dubbed Near-Field Scanning Optical Microscopy (NSOM or

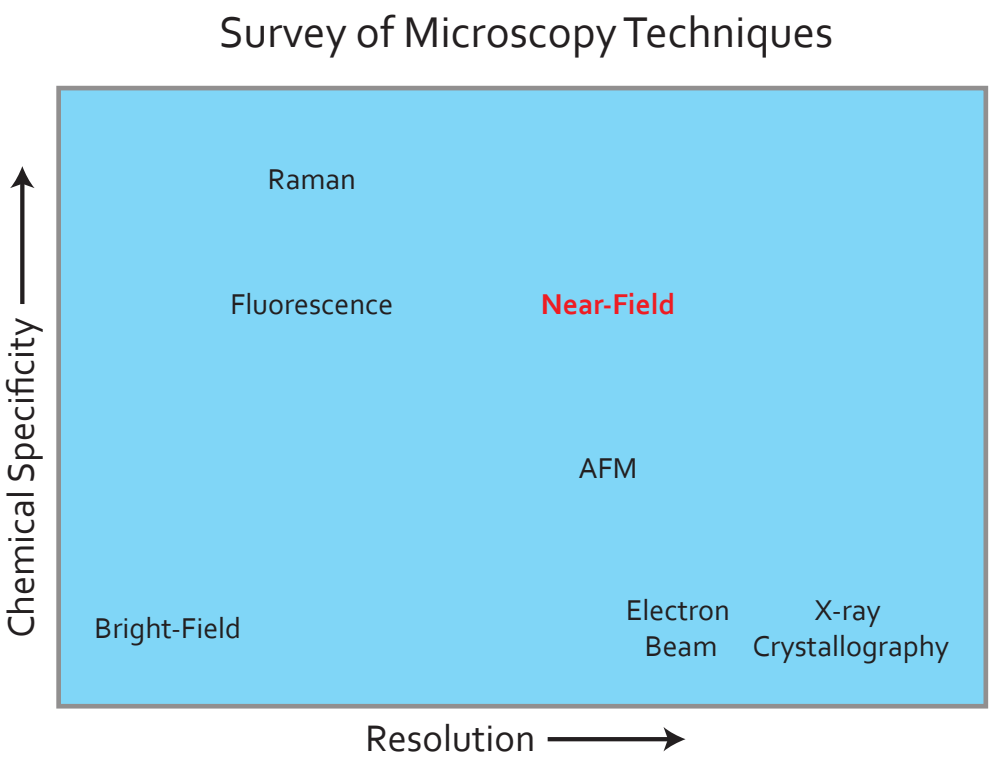

Figure 1.6. Several various microscopy techniques are compared in a plot of chemical specificity vs. resolution. Increasing chemical specificity is up, while increasing resolution is to the right. Thus, the ideal microscope would be found at the top-right corner.

SNOM). Its relation to other microscopy techniques is illustrated in Figure 1.7.

1.5 Breaking the Diffraction Barrier

While the diffraction limit is still in full force for all optical microscopes, scientists have come up with many clever ways to eke out all the resolution possible from a system. The most obvious techniques involve not breaking the diffraction limit, but simply pushing the boundaries of it. One way this is accomplished is by reducing λ, using bluer light, even going to UV or X-Ray wavelengths. Shorter wavelengths require specialized optical components and light sources, as most optical glass absorbs strongly at these wavelengths. Increasing the NA of a system also will lead to better resolution, but again this is limited by the index of refraction of materials. By moving to specialized high index of refraction oils and

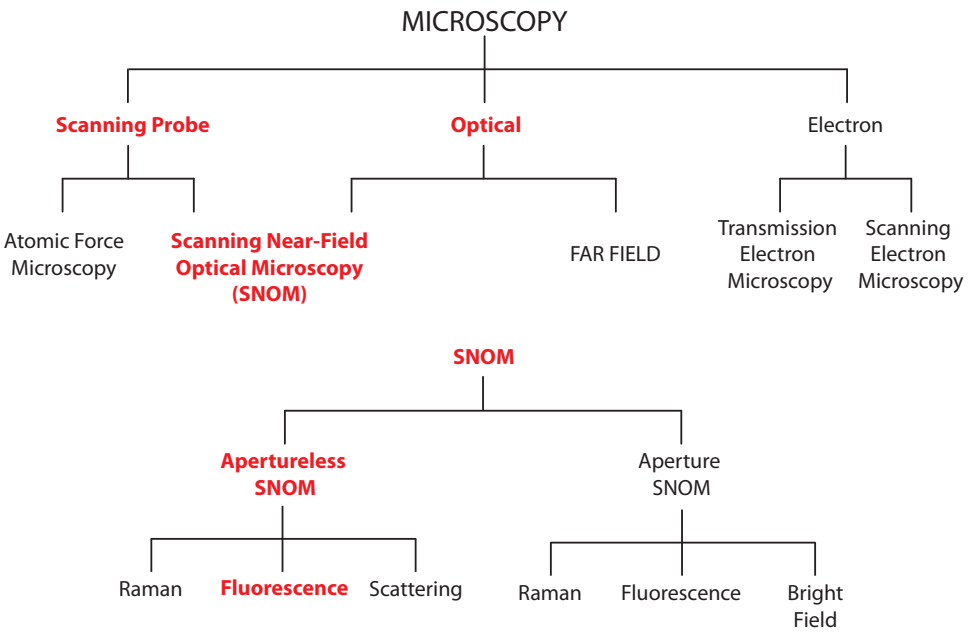

Figure 1.7. Tree of related microscopy techniques.

glass, microscope objectives can be found at ever increasing values of NA. Other techniques involve the simultaneous use of two objectives (one on either side of the sample) to make further gains in increasing the NA of the system [5].

Microscopy techniques that actually break the diffraction limit are also actively being pursued. These include using tailored laser beams that have complex intensity profiles, whose envelope function is diffraction limited, but has regions within the beam profile that can be much smaller than the diffraction limit [5]. Other ideas rely on the fact that we can locate the center of an object very well, but as usual this is only possible if we can resolve it, i.e., tell that there is only one object present. The method then is to only look at fluorophores one at a time, by turning them on in either a random or controlled fashion, and then measuring the location before a neighboring fluorophore is subsequently activated.

1.5.1 Near-Field Scanning Optical Microscopy

Another class of optical techniques that beat the diffraction limit are included in the realm of near-field microscopy The diffraction limit is actually only a far-field phenomenon. In this regard, any technique that is a near-field technique is inherently not subject to the far-field diffraction limit. The prefixes "near"and "far" refer to a distance relative to the wavelength of light employed. Thus "near" field can signify that either the sample is illuminated by a light source much less than a wavelength away ($d << \lambda$) or that a detector is much less than a wavelength away from the sample. Another term for the near-field region is the Fresnel zone. Conversely, far-field phenomena refers to both an illumination source and detector that is at a distance much longer than a wavelength away $d >> \lambda$. Far-field diffraction is also known as Fraunhofer diffraction. For example, in the types of microscopy discussed thus far, a lens is placed quite far, relatively speaking, from a sample in order to collect light (information) coming from the sample. As the wavelength of visible light is 400 - 700 nm, even a sample placed 1 cm from a lens is 20,000 wavelengths away, and thus would have far-field detection. If somehow a local excitation source were placed within several tens of nanometers from the sample it would have near-field illumination.

Again, even within the realm of SNOM there are many variations. NSOM is a relatively new field, having really only existed in practice for the last $\sim$15 years, and naming conventions are still being developed. NSOM is the acronym most commonly used in North America, while SNOM is more popular in Europe. There are two main branches of NSOM: aperture type and apertureless (cf. Fig 1.7). Both rely on a particular scanning probe technique—atomic force microscopy (AFM) for fine control of a probe.

Aperture type NSOM involves a hollow AFM tip with an opening at the distal end of 50-80 nm wide. Light is focused down the center of the tip until only some small fraction exits the narrow aperture as much is lost as it travels down the tip. By scanning this hollowing tip over a sample, a local light source provides near-field detection. Either far-field or near-field detection can be employed with

this technique.

Apertureless NSOM, relies on a solid AFM tip aligned within the central focus of an excitation beam. With the appropriate choice of tip, and by carefully tailoring the polarization of the laser beam, the tip can amplify the optical fields around it, such that a region of extra intensity can occur at the apex of the tip. In this way, the apex of the tip acts as a local light source as it is scanned over a sample, capable of providing near-field illumination. Detection of light signals in this case is always done in the far-field.

Unfortunately both types of NSOM begin with the letter *A*, so the acronym "ANSOM" has been used to refer to both types of techniques, although it is slightly more common to use ANSOM in reference to apertureless-NSOM. Within each branch of NSOM there are several variations still of each technique. As we are interested in combining the attributes of fluorescence microscopy with AFM, the particular flavor we practice in our lab is fluorescence-apertureless-NSOM. The term FANSOM, has been coined and used in a limited way, however, to avoid any ambiguity with a related aperture type technique, the preferred terminology in our lab is Tip-Enhanced Fluorescence Microscopy or TEFM.

1.6 TEFM

Figure 1.8 represents the basic operating principles of TEFM. An AFM probe is aligned into the center of a focused laser. The AFM tip is vertically oscillating, constantly tapping the sample. The tip and laser are aligned, then the sample is scanned between the two. As the sample scans under the tip, the topography is recorded by monitoring the motion of the tip. Simultaneously fluorescence signals coming from the sample are also collected. Additionally, the AFM is able to report extra information beyond simply the height, a relative measure of how hard/soft or sticky a sample is also simultaneously acquired in a separate data channel.

If the tip material and laser polarization are carefully controlled, a region of enhanced local field surrounds the apex of the tip. A fluorescent sample sits atop a glass coverslip. As the sample is in the far-field focused laser spot, it will be

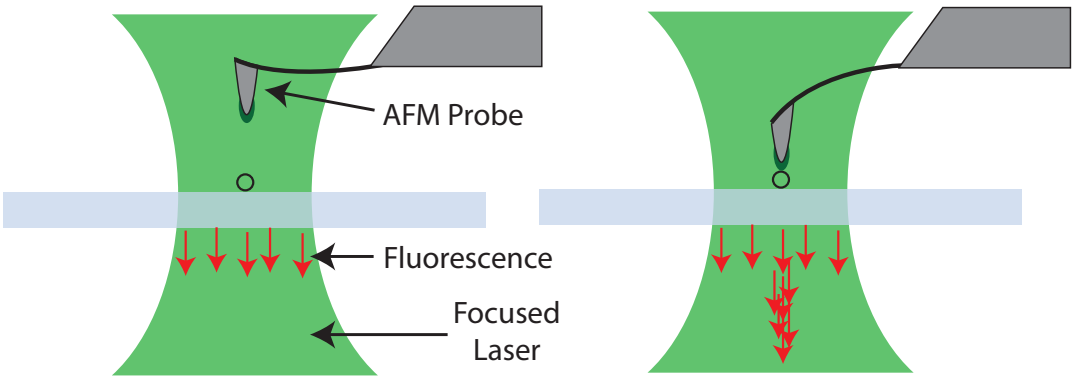

Figure 1.8. Cartoon overview of TEFM. A sharp tip is positioned within the focus of an excitation beam, which causes an enhanced local field at the apex of the tip. The sample (circle) is being illuminated by the far-field background signal and emits fluorescence photons (downward arrows). When the tip apex is near the sample, the local field associated with the tip apex leads to extra fluorescence signal but at a much higher resolution.

emitting photons, which will produce a diffraction limited image. As the tip is brought in close proximity to the sample the "extra" field around the tip apex also illuminates the sample, which in turn gives off more fluorescence signal. Again, the properties of the tip are very important in order to be able to observe any field enhancement. Some materials, such as metals, can lead to a local reduction in fluorescence signal via a process called quenching. Actually, near-field signals can come from either enhancement or quenching of the far-field fluorescence signal. TEFM signals contain both a diffraction limited far-field background signal as well as a near-field signal, whose resolution is only limited by the sharpness of the tip.

In terms of optical resolution, TEFM is the world-record holder, coming in at under 10 nm [2, 3]. My advisor, Jordan Gerton, and co-workers in Stephen Quake's group at Caltech have demonstrated the extreme resolution capabilities of TEFM by imaging short strands of DNA labeled with fluorescent dye molecules at either end. By measuring the distance between the two fluorophores, and using a known code of DNA 60 base pairs in length, they were able to determine the length per base pair of that particular kind of DNA, as shown in Figure 1.9. While there

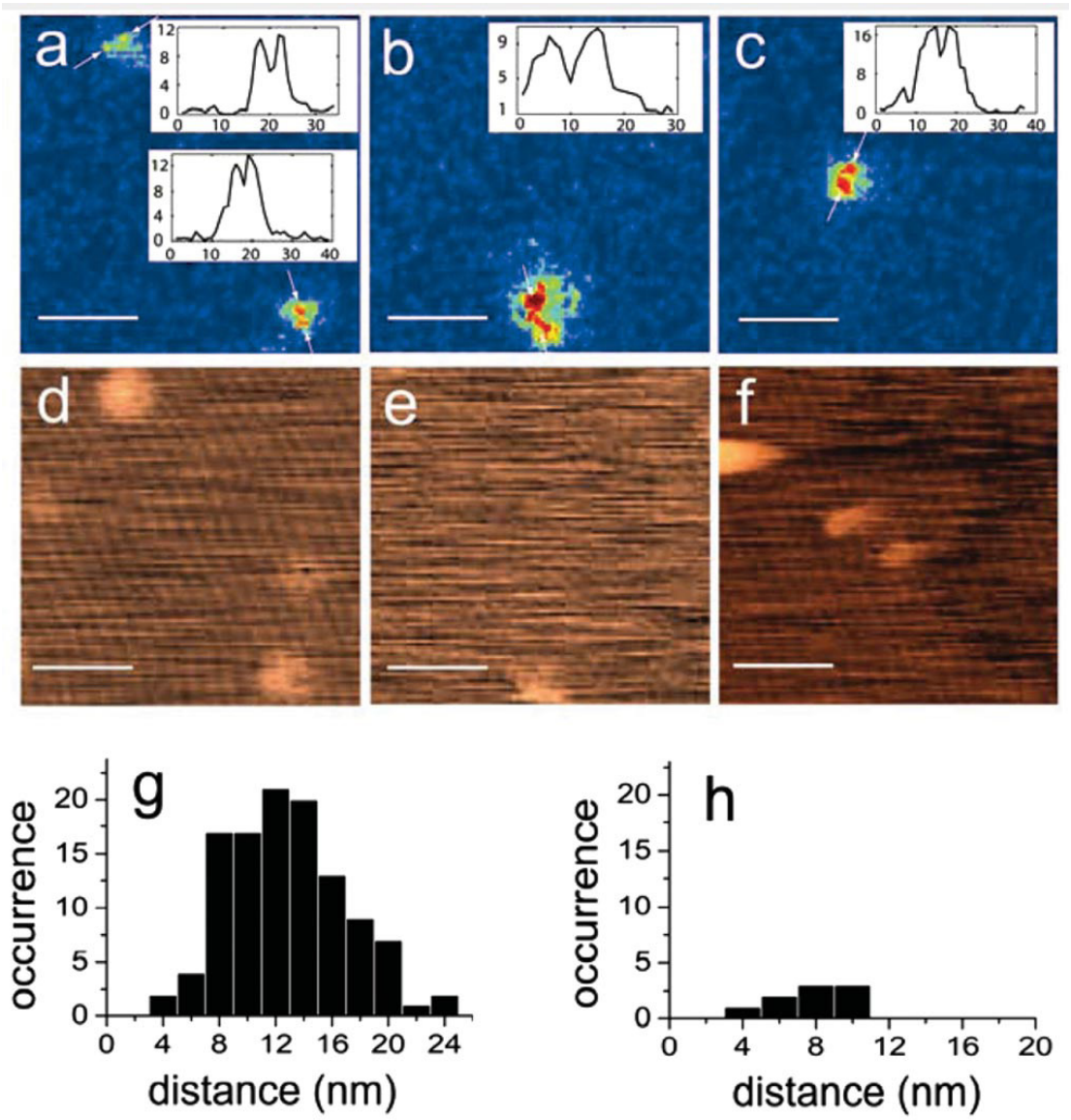

Figure 1.9. Short fragments of DNA labeled with fluorescent dye molecules on either side, creating DNA "dumbbells." These dumbbells were imaged with TEFM, resulting optical images are shown in (a)-(c) with profiles shown as insets. The corresponding AFM topography images are shown in (d)-(f). Scale bars are 50 nm. The length of the DNA chain was determined optically, by the distance between fluorescent centers. Panel (g) shows a histogram of these distances. Panel (h) represents a control experiment, where some single dye molecules exhibit double lobe artifacts - the resulting distances between the artifactual lobes are plotted here as a histogram. Reprinted with permission from reference [3]. Copyright (2006) by The American Physical Society.

are obvious limitations to TEFM in the fact that it is a surface technique, such demonstrations have shown the great promise that TEFM has towards unraveling an untold number of biological secrets.

Just as AFM resolution depends on the sharpness of the probe being used, TEFM resolution scales with tip size. Thus a sharper probe (assuming it is made of the appropriate material) leads to both higher AFM and optical resolution. The sharpness of such AFM probes is often described by the radius of curvature at the apex, as at this scale there really is no such thing as a sharp corner. Commercially available AFM tips can have radii of curvature less than 10 nm. Figure 1.10 shows an SEM image of a typical AFM tip used in our lab.

As mentioned, a raw, TEFM image consists of both near-field and far-field components, which makes such raw images messy and difficult for non-experts to interpret. Much work has been devoted to removal (or at least suppression) of the far-field background signal, as high levels of far-field background signal lead to worse

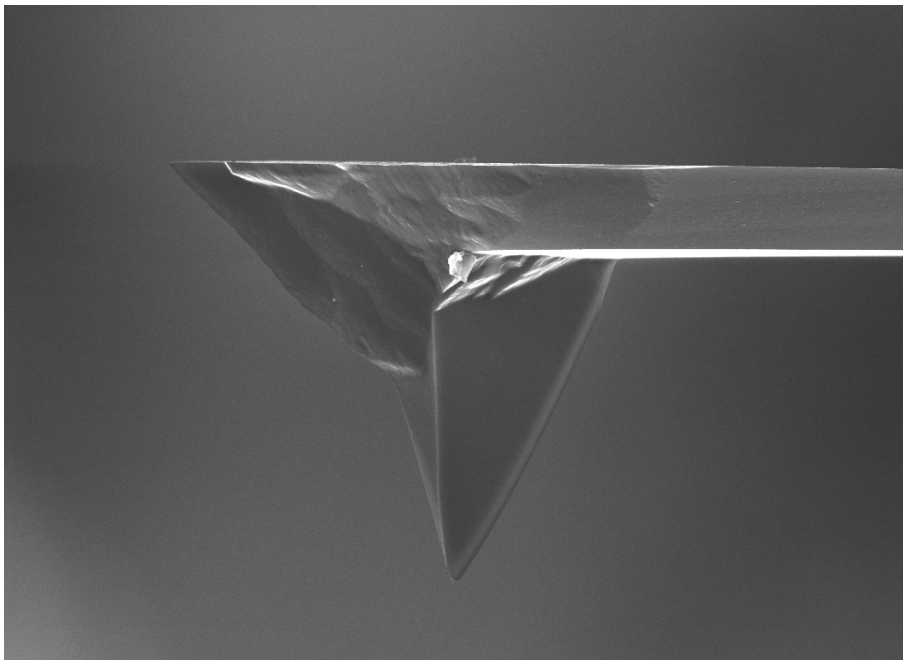

Figure 1.10. Scanning electron microscope (SEM) image of a typical AFM tip. This type of tip (silicon) is also capable of exhibiting strong field enhancement at the apex under appropriate illumination conditions.

image contrast. Some rather straightforward and very effective techniques such as lock-in demodulation have readily been applied to tackling this problem as seen in Figure 1.11. Other more sophisticated background suppression (or near-field isolation) techniques have also been of particular interest in our lab's research. A great deal of my work has dealt with understanding and devising means of suppressing these unwanted background signals.

1.7 Summary

As mentioned we have developed a fluorescence microscope in our lab that is capable of nanoscale resolution imaging of single fluorescent molecules (fluorophores) [2, 3, 7, 6, 8], which we call a tip-enhanced fluorescence microscope (TEFM). To date, we have primarily used this instrument to image isolated fluorophores in air, and have repeatedly demonstrated spatial resolution of ∼10 nm. The objective of my project has been to adapt the microscope for imaging in an aqueous

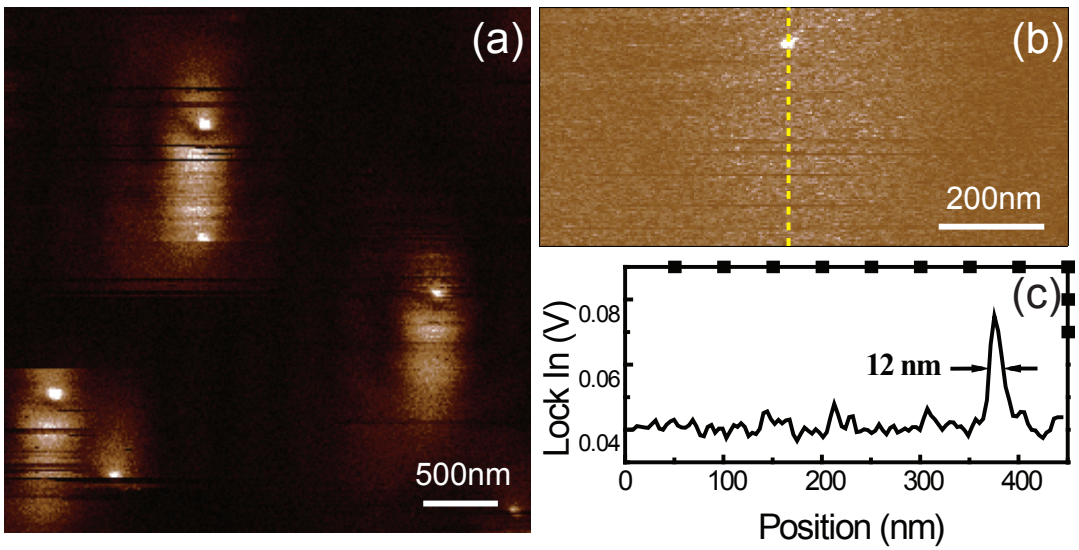

Figure 1.11. TEFM images of quantum dots on a glass surface using a silicon AFM probe. (a) Fluorescence image of several isolated quantum dots with no lock-in amplification. The size of the far-field spot is about $1\mu m \times 0.5\mu m$. (b) Lock-in magnitude image of an isolated quantum dot. (c) Profile specified by dashed line in (b). Reprinted with permission from reference [6], © 2008 IEEE.

environment, and further, to optimize its performance to enable imaging of protein networks in planar membranes assembled onto glass coverslips. To accomplish this, two benchmarks must be met. First, the contrast of the microscope must be sufficient such that individual molecules within a dense ensemble can be resolved. Second, the microscope must be made to work efficiently in water. We have made substantial progress in the first area and are now addressing the second. These have been the primary objectives of my research in the lab; some of the successes in reaching these objectives, as well as some of the remaining hurdles, will be discussed throughout the body of this dissertation.

Chapter 2 will provide the background theory necessary to understand what an optical near-field is, and how it is created. The concept of field enhancement at a tip apex will be explored in depth. Additionally, fluorescence quenching and its role in near-field optics will also be described in detail. Finally, it also discusses optimization of probe geometries/permittivities for maximal results. Chapter 3 will discuss the experimental setup required to make TEFM a reality. This includes both the hardware requirements and some of the basic theory needed to appreciate the design of the instrument. This is a very technical chapter, explaining the particular intricacies of our imaging system, and is primarily intended for new students in the lab.

TEFM as an imaging system will be discussed in Chapter 4. More particularly, attention will be focused on image contrast, and how demodulation techniques can be used to increase contrast. The primary focus will be on the near-field enhancement of fluorescence, but some discussion on fluorescence quenching as a contrast mechanism is also included. This chapter has provided an important framework for the way in which we describe and discuss our system. Chapter 5 describes a particular data acquisition technique that allows for three-dimensional mapping of near-field interactions. This near-field tomography, as it is called, has become an important addition to the lab, allowing for unlimited post-processing possibilities of any data we collect in the lab.

Results from various attempts at imaging in aqueous environments are shown

in Chapter 6. Thus far, we have had extremely limited success in using TEFM in water. Many explanations of why TEFM in water has been so elusive are given. As a consequence of this analysis, the optimal excitation mode is also described. Other methods for imaging biological samples at nanoscale resolution are also discussed. Finally a possible way forward for using TEFM in water is sketched out. Chapter 7 presents the latest applications of utilizing TEFM in conjunction with carbon nanotube (CNT) tips. The extreme precision of our measurement techniques are highlighted by measuring CNT-quantum dot interactions. Finally, a discussion of important future experiments using CNTs is included.

CHAPTER 2

ENHANCEMENT VS. QUENCHING

Thus far I have given a basic overview of tip-enhanced fluorescence microscopy and discussed some of the requirements for achieving optical near-fields, namely that aligning a near-field probe in the center of a laser beam with axial polarization leads to an enhanced field at the apex of the probe. This chapter explores the particular mechanism for this creation of optical near-fields and the factors that influence both the strength and extent of such fields. In addition to discussing the factors influencing enhancement, this chapter also explores the opposite of enhancement: quenching. Briefly, quenching is a process leading to decreased fluorescence signals as the presence of a metal structure increases the nonradiative decay rates of a fluorophore. Quenching is an extremely important factor in near-field microscopy in that any fluorescence signal obtained in TEFM is the net result of a competition between enhancement and quenching. Experimentally these two effects can be difficult to separate, as they can often be competing on similar length scales. To first order, field enhancement relies on only the particular details of the tip and incident field. Quenching, on the other hand, is somewhat more complicated. Quenching is not a decrease in local field strength, but rather it is described by nonradiative fluorescence decay channels, thus its effects cannot be determined without a knowledge of the current state of the fluorophore. Factors determining the extent of quenching by a tip then include the initial radiative and nonradiative decay rates, quantum yield, tip-material, and tip-geometry. While this chapter will discuss enhancement and quenching separately, the focus will be on the importance of appreciating the interplay between the two.

2.1 Setup

In order to properly ubicate the relevance of the theory presented in this chapter, it is first necessary to understand the experimental setup we employ. The basic microscope schematic is shown in Figure 2.1. Briefly, illumination and detection are achieved via the same inverted microscope configuration. Axial polarization of the excitation laser is achieved by either a radial polarization state or evanescent illumination from blocking all but a small wedge of supercritical rays at the back aperture of the objective. The tip can be operated in contact mode, where is is simply dragged along the surface, or tapping mode, where it oscillates vertically, only tapping the surface intermittently.

2.2 Enhancement

When a dielectric material is placed in a uniform electric field, the applied field polarizes the material as the electrons and ions migrate toward their respective sides of the dielectric [9]. This charge separation creates an induced electric field, and the total field near the surface of the material is the superposition of both the applied field and the induced field. For a dielectric sphere, the induced field can be obtained analytically, as shown in Equation 2.1, and illustrated in Figure 2.2. Here, a uniform static field is applied along the vertical axis, but the calculation is also valid for an oscillating field with vertical polarization if the size of the sphere is much smaller than the wavelength. In this quasi-static approximation, retardation effects can be neglected and at each point in time, the applied field can be considered uniform. At the vertical poles of the sphere, the total electric field is enhanced relative to the applied field, while the total field is reduced along the horizontal equator.

$$\vec{E_{out}}(r,\theta) = \vec{E_0} \cdot \cos\theta \left[1 + 2\frac{\epsilon_r - 1}{\epsilon_r + 2}\frac{R^3}{r^3}\right]\hat{r} + \vec{E_0} \cdot \sin\theta \left[-1 + \frac{\epsilon_r - 1}{\epsilon_r + 2}\frac{R^3}{r^3}\right]\hat{\theta} \quad (2.1)$$

From Equation 2.1 it can be seen that the maximum electric field strength occurs at the poles of the sphere, and is given by:

$$E_{max} = E_0\left(1 + 2\frac{\epsilon_r - 1}{\epsilon_r + 2}\right), \quad (2.2)$$

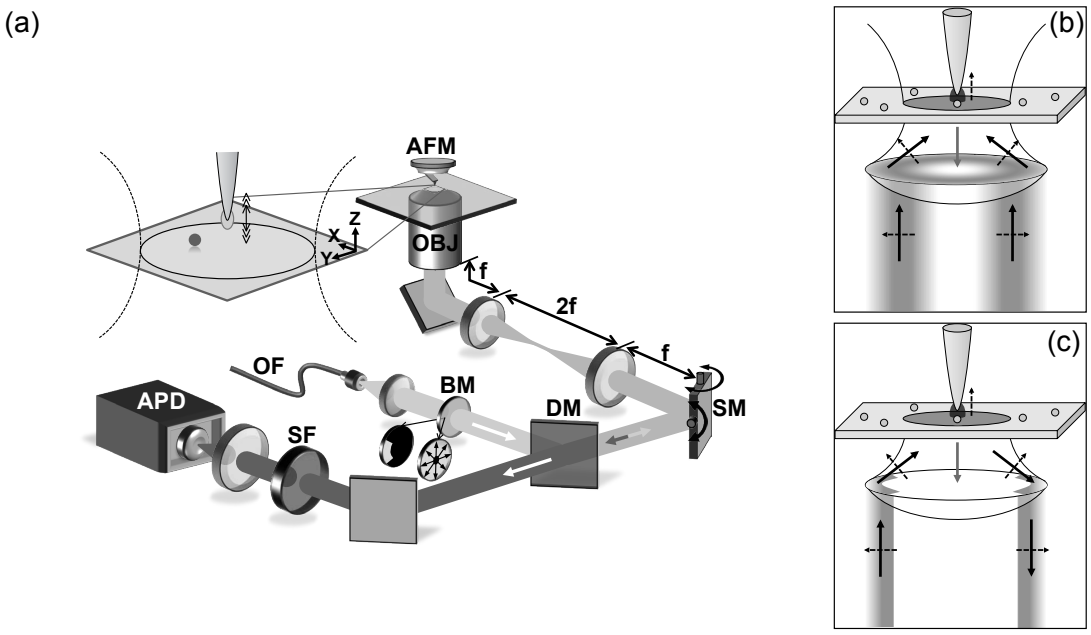

Figure 2.1. Experimental realization of a fluorescence apertureless-NSOM setup (a). An excitation beam exits an optical fiber (OF) and goes through a beam mask (BM) in either a wedge or radial configuration. The excitation beam is reflected by a dichroic mirror (DM) and off a scanning mirror (SM) before being focused through a microscope objective (OBJ). Signals are collected through the same path and directed onto an avalanche photodiode (APD) after passing through the appropriate spectral filters (SF). Panels (b) and (c) show ray diagrams for a radial and wedge beam mask respectively. Solid arrows show the direction of beam propagation (dark arrows for excitation and lighter arrows for emission), while dashed arrows represent the polarization direction.

where E_0 is the applied field and ϵ_r is the permittivity of the sphere relative to that of the surrounding medium: $\epsilon_r = \epsilon_{dielectric}/\epsilon_{media}$. In principle, a dielectric sphere can be used as an apertureless NSOM probe if it can be scanned in close proximity to a sample. In this case, the enhanced field at the distal pole of the sphere can increase the optical response. This response is generally proportional to the optical intensity or higher orders thereof depending on the particular scattering process [10], the expression in Equation 2.2 must be raised to an appropriate power to find the expected enhancement in the scattering rate. For dielectrics, the magnitude of this effect in some sense is independent of the size of the sphere. As shown in Figure 2.3,

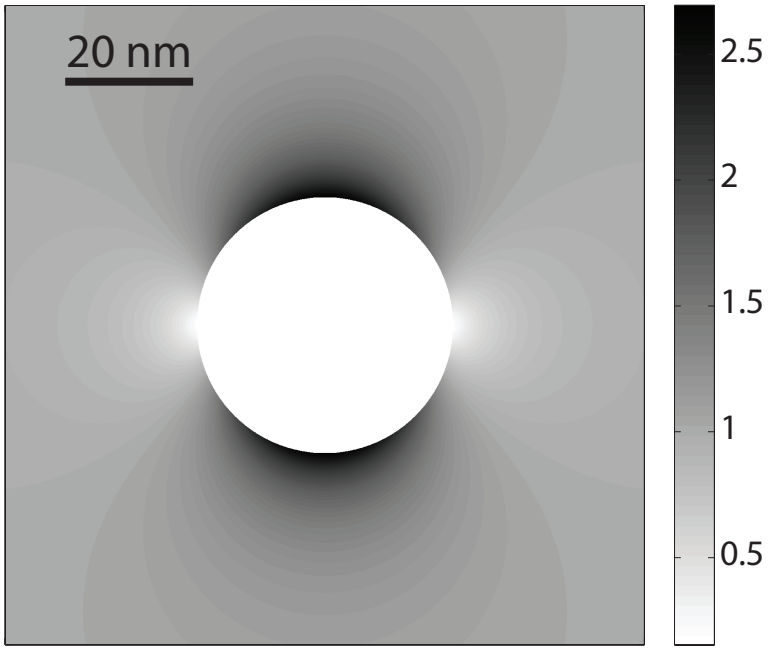

Figure 2.2. A dielectric sphere (Si) placed in a vertically oriented uniform electric field. The dark regions correspond to elevated electric fields. The incident field has a strength of $E_0 = 1V/m$ as can be seen on the scale shown.

the peak field enhancement for a dielectric sphere is identical, regardless of particle radius, the caveat being that the field decays more slowly. A slowly decaying field around a near-field probe leads to a decrease in optical resolution, an unwanted effect. Note also that it is in this manner that near-field resolution is a function of tip-sharpness rather than wavelength—in general, sharper tips have a steeper field decay.

For dielectric materials, the peak intensity can be enhanced by at most a factor of nine (when $\epsilon_r \rightarrow \infty$) for this spherical geometry. Although the spherical geometry can be solved analytically, it often does not accurately approximate the shape of many AFM tips. Furthermore, elongated geometries can yield significantly larger field enhancement. Figure 2.4 demonstrates this effect by plotting the maximum intensity enhancement from a three-dimensional finite element calculation of prolate spheroids of increasing semiaxis ratios. In fact, Bohn et al. have shown that as the long axis of the spheroid approaches infinity, the intensity enhancement adopts the

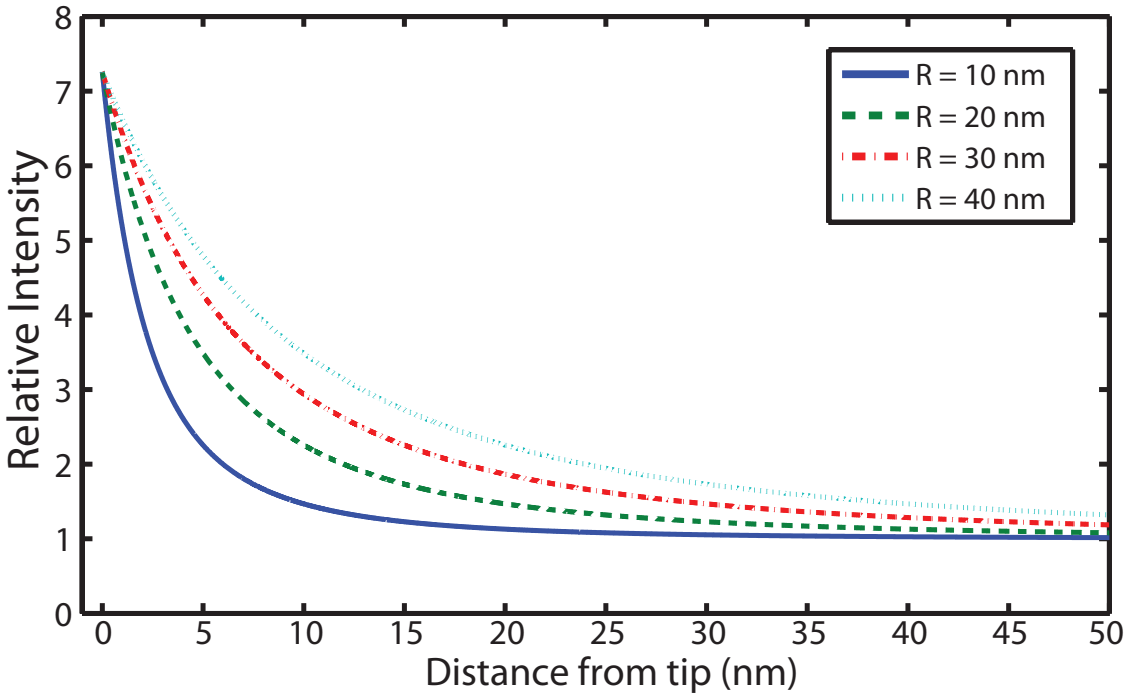

Figure 2.3. Enhanced intensity decay of silicon ($\epsilon_{Si}^{543nm} = 17.6 + 0.12i$) spheres. The field decay for several sized dielectric spheres is plotted taking $\theta = 0$ from Eqn. 2.1 and then plotting for $r \geq R$. The peak field enhancement is independent of radius. Larger particles lead to slower field decay and thus would yield poor optical resolution when used as a near-field probe.

analytic form [11]:

$$E_{peak} = E_0 \cdot \epsilon_r \Rightarrow I_{peak} = E_0^2 \cdot \epsilon_r^2, \tag{2.3}$$

which is reproduced well in the numeric calculations shown in Figure 2.4. This geometry-dependent electric-field enhancement, where the highest enhancement values occur in regions of highest curvature is known as the lightning-rod effect. It is important to remember that Equations 2.2 and 2.3 are strictly valid for static electric fields, and thus can only be used for optical fields (resulting from either continuous wave or pulsed lasers) when retardation effects can be neglected, namely when the size of the object is much smaller than the light wavelength. This, of course, prohibits rigorous application of Equation 2.3 for commercial AFM tips in NSOM, as the entire structure of a commercial tip, while very sharp at the apex, is an extended structure many times larger than optical wavelengths ($\sim 10\ \mu m$).

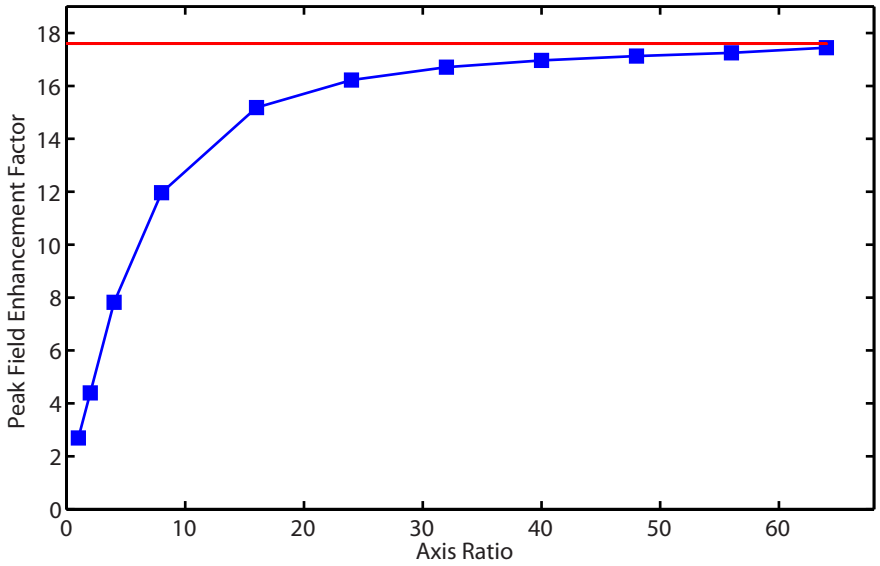

Figure 2.4. A three-dimensional electrostatic finite element calculation of the maximum field enhancement for prolate spheroids of varying aspect ratios. Spheroids were given the permittivity of silicon (17.6+0.12i @ 543 nm). Aspect ratio is calculated as the semi-major axis divided by the semi-minor axis of the prolate spheroid (short axis was kept constant). The incident field is applied parallel to the long axis of the spheroid. The horizontal line is the theoretical limit as the ratio approaches infinity as found in Eq. 2.3.

Nothing prohibits applying Equations 2.2 and 2.3 to metal nanoparticles as long as the size of the particle is no larger than the optical skin depth, $\delta = \lambda/(4\pi\sqrt{\epsilon})$; otherwise, the bulk electrons are shielded and only the surface electrons experience the external driving field. In general the dielectric function is complex valued and also frequency dependent, $\epsilon_r(\omega) = \epsilon_r'(\omega) + \epsilon_r''(\omega)$. Here the imaginary part relates to light absorption, which ultimately results in energy loss through ohmic dissipation (Joule heating) [12]. Metals generally have negative values for the real parts of their dielectric function throughout the visible spectrum, which also raises the possibility of exciting plasmon resonances (cf. Eq. 2.1 for $\epsilon_r = -2$) that can massively increase the enhancement factor (see below).

More precise predictions of the field enhancement for realistic tip geometries, such as the conical or pyramidal tips that are commercially available, can be obtained by solving Maxwell's equations on a discrete grid using a finite element anal-

ysis program such as COMSOL [13] or CST Microwave Studio [14]. Alternatively, a discrete dipole approximation (DDA) [15] or multiple multipoles method can also be used [16]. These previous calculations have predicted intensity enhancement factors in excess of 1,000 for metal tips and around 225 for dielectric tips [16, 11]. The previous calculations by Bohn et al. using spheroids have been readily duplicated (as evidenced by Fig. 2.4), and a slightly more realistic probe geometry is now used. Using a finite element solver (COMSOL) I have also calculated the expected field enhancement for a dielectric tip as shown in Figure 2.5.

A silicon tip is represented as a conical probe with an aspect ratio of 4:1 (equivalent to a half cone angle of 14°). Near the apex of the probe the cone has been truncated and replaced with a symmetric quadratic Bézier curve with an identical aspect ratio in a continuous fashion. The sharpness of the tip can be considered to have a radius of 10 nm; the tip height is 10 μm. The maximum intensity enhancement factor of such a probe was calculated to be nearly 3000.

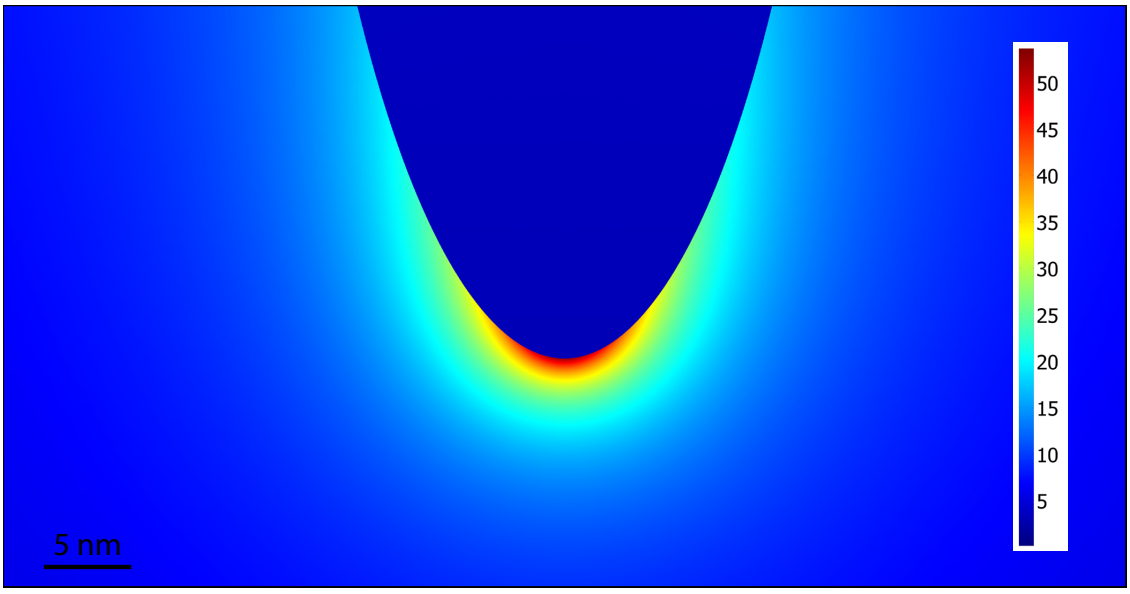

Figure 2.5. A 3D electrostatic calculation of field enhancement around a near-field probe was performed in Comsol. A small portion of the simulation is viewable here. The modeled Si tip has a radius of 10 nm, a height of 10 μ m, and an aspect ratio of 4:1. The scale bar indicates the magnitude of the field, where the incident field is 1 V/m.

Such large enhancement factors have never been observed. For example, to the best of our knowledge, the largest signal enhancement factor reported for silicon tips is ~ 20 for fluorescence measurements of 4-nm diameter quantum dots [2]. This discrepancy between predicted and observed enhancement values could also be due to rapid growth of oxide layers, which have a smaller permittivity, or to irregular geometry. Furthermore, quenching, redirection of fluorescence, and interference effects can also contribute to such a discrepancy. Clearly the geometry of the probe in all aforementioned calculations is quite different from real probes. Furthermore, keeping the apex of the tip the same size, while changing the tip height can lead to vastly different values in calculated enhancement factors as shown in Figure 2.6. More properly, the amount of tip-enhancement in such electrostatic calculations

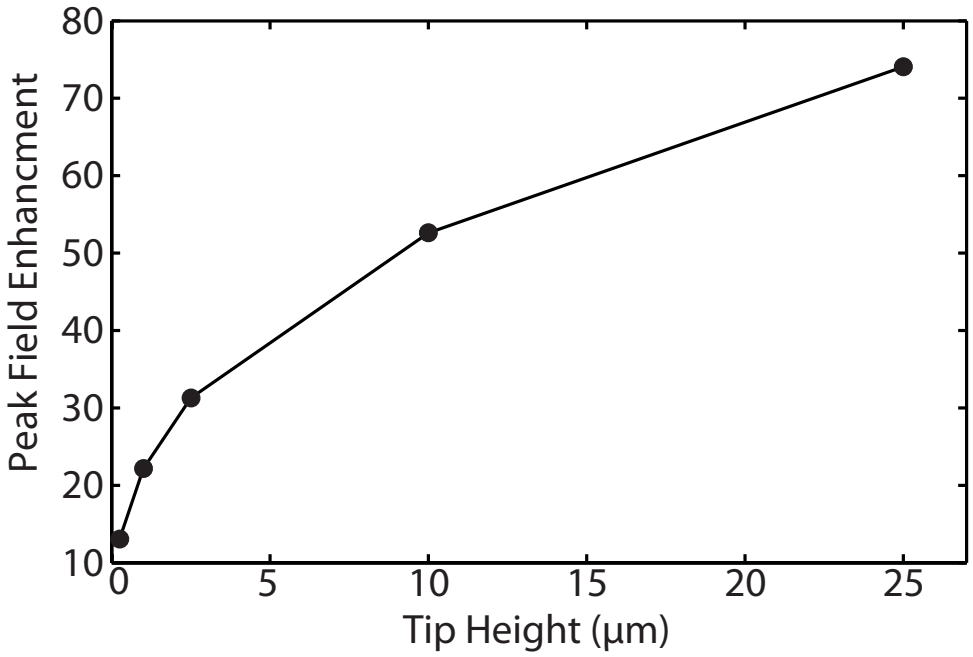

Figure 2.6. Peak field enhancement as a function of tip height. A 3D electrostatic calculation of field enhancement around a near-field probe was performed in Comsol. A silicon tip is modeled by a conical probe with an aspect ratio of 4:1. The apex of the probe the cone has been truncated and replaced with a symmetric quadratic Bézier curve. The peak enhancement factor under the tip is plotted as a function of the tip height—keeping the geometry of the apex constant.

scales with the volume. This can be troublesome as many calculations keep the vertical extent of the probe to a minimum in order to lower computational demands. As evidenced by the drastically different results reported in the literature, these calculated enhancement values must be used as only very rough estimates.

2.3 Plasmon Resonances

The primary optical response of metals in the presence of a driving field is determined by the motion of their conduction electrons, which have an effective mass characteristic of the particular material. Neglecting the response of the valence electrons (i.e., interband transitions), we can then apply the Drude-Sommerfeld model for the free electron gas and solve for the frequency-dependent permittivity:

$$\epsilon_{Drude} = 1 - \frac{\omega_p^2}{\omega^2 + \Gamma^2} + i\frac{\Gamma\omega_p^2}{\omega\left(\omega^2 + \Gamma^2\right)}, \tag{2.4}$$

where $\omega_p = \sqrt{ne^2/m_e\epsilon_0}$ is the bulk plasma frequency, n is the free-electron density, e and m_e are the electron charge and effective mass, $\Gamma = v_f/l$ is a damping term, and v_f and l are the Fermi velocity and mean free path of the electrons. For ballistic charge transport (i.e., $\Gamma = 0$), the dielectric function is zero at the plasma frequency, positive at high frequencies (i.e., UV), and negative at low frequencies (i.e., visible and IR wavelengths). Thus, bulk metals are largely reflective below the plasma frequency and largely transparent above it [17]. A plasmon at the surface of a metal is known as a surface plasmon, and from Maxwell's equations it can be shown that surface plasmons can exist only at an interface of two materials with permittivities ϵ_1 and ϵ_2 that simultaneously satisfy [18]:

$$\epsilon_1(\omega) \cdot \epsilon_2(\omega) < 0 \epsilon_1(\omega) + \epsilon_2(\omega) < 0. \tag{2.5}$$

These conditions are easily met at a metal-dielectric interface, when a metal has a sufficiently large (and negative) permittivity compared to the dielectric; at such an interface it can be shown that the field decays exponentially into both materials away from the interface [9]—it is in this way that a surface plasmon derives its name. Plasmons and surface plasmons have been studied extensively and more details can be found in a variety of sources [9, 17, 19, 20, 18].

Surface plasmons play a large role in near-field optics as a material excited resonantly can generate enormous field enhancements beyond those expected from the lightning rod effect alone. Optimal field enhancement requires the right combination of material, excitation wavelength, and also extremely importantly, geometry. For instance, if a subwavelength sized metal sphere is illuminated with an excitation source near its plasmon resonance frequency, localized surface plasmons can greatly increase the enhancement factor. As the spherical particle is elongated along its polarization axis the field can be enhanced even further [21]. However, while elongated particles lead to larger enhancement as compared to their spherical counterparts made of the same metal, the excitation wavelength needed to achieve the maximum enhancement may be quite different for the two geometries. As nanoparticles become elongated, the extinction spectrum can redshift significantly [15]. Furthermore, plasmon resonances for metal nanostructures redshift with increasing size [22]. Aside from having different extinction spectra, the particle size also matters in terms of the magnitude of obtainable field enhancement: if the particle is larger than the skin depth, the inner electrons will be shielded, resulting in reduced enhancement. For smaller particles with large surface area-to-volume ratios, electron collisions with the surface become a large source of plasmon damping, thus reducing plasmonic field enhancement [23, 24]. These geometry and size effects can collude to make predictions about field enhancement from metal nanoparticles challenging; for example, just because an 80-nm gold sphere leads to reasonable enhancement at $\lambda = 633$ nm, it should not be presumed that a much bigger and elongated gold tip should yield similar performance at the same wavelength.

Bulk and surface plasmons can also play a major role in energy transfer. In particular, a photoexcited particle (e.g., quantum dot or fluorophore) that would normally relax via radiative channels (fluorescence), can instead nonradiatively transfer its internal energy to a nearby plasmon-active structure. This results in a sharp reduction in the detected fluorescence rate (quenching), and also in the fluorescence lifetime [1, 25], as nonradiative decay channels become predominant. Recently, several groups have directly observed the competition between

field enhancement and fluorescence quenching when using metal tips in NSOM [12, 26, 27, 13, 24]. The net signal depends on the details of this competition, which in turn depends strongly on the tip geometry and size, the light wavelength and polarization, and the distance between the sample and tip.

2.4 Optical Antennas

The combination of the lightning-rod effect and plasmon resonances naturally leads to the concept of designing nanostructures with strong, shape-specific resonances to drastically enhance the optical field. This is, in fact, a description of an optical antenna, which like their radio or microwave analogs, can be used to convert free propagating electromagnetic waves into localized fields and vice versa. The design and implementation of antennas for optical frequencies is highly desirable for a number of applications including NSOM, biochemical sensing, display technology, etc., and is a very active area of research. While any near-field probe can be considered an antenna inasmuch as it can locally focus light, we will more rigorously use the term "antenna" to describe a device that exhibits shape-specific resonances, which implies that they are made of metal.

There are a number of difficult challenges associated with scaling down antennas from the macro to the nanoscale as needed for optical field enhancement. For example, it is difficult to fabricate structures of this size using conventional lithography, so specialized techniques such as focused ion beam milling or electron beam lithography must be used. Furthermore, at optical frequencies, charge transport in nanoscale metallic structures can suffer from a number of damping mechanisms, in contrast to the ideal conductors envisioned in antenna theory for the microwave and radio wave regions of the electromagnetic spectrum. Thus, the design of efficient antennas at optical frequencies requires new theories, or at least rigorous adaptation of existing microwave theories, to account for this non-ideal behavior.

At optical frequencies, the skin depth of a metal can be of the same order of magnitude as the antenna size. The penetration of electromagnetic waves into the antenna creates electron oscillations inside the metal, which tends to push

the antenna resonance toward a higher frequency and thus a shorter effective wavelength. For example, van Hulst and co-workers have shown that the resonant length of a linear monopole antenna is significantly shorter than predicted for an ideal conductor using classical antenna theory [28]. Novotny modeled the antenna as a strongly coupled plasma in order to determine the relation between the external and effective wavelengths [29]. He found that the effective wavelength, λ_{eff} , is related to the external wavelength, λ, according to:

$$\lambda_{eff} = n_1 + n_2 \frac{\lambda}{\lambda_p} \tag{2.6}$$

where n_1 and n_2 are complicated expressions depending on the geometric and static dielectric properties of the antenna, and λ_p is the plasma wavelength. The shorter effective wavelength predicted by Novotny can be several times smaller than the free-space wavelength [30]. This correction to the wavelength is very important for antenna design since it implies that optimized antenna sizes should be shorter than what traditional antenna theories project and are dependent on the shape of the tip and the properties of the metal.

A number of antenna designs have been used in NSOM to obtain large and confined field enhancement, thus obtaining optical resolution beyond the diffraction limit. For example, the monopole antenna mentioned above [28] was accomplished with the tip-on-aperture approach developed by Guckenberger and co-workers [31, 32]. The antenna was driven with light emerging from the aperture with polarization along the antenna axis. In this work, the antenna resonances were mapped by scanning an antenna over a single molecule while monitoring its rate of fluorescence emission, and repeating the experiment for different length antennas. The observed fluorescence rate for similarly oriented molecules increased dramatically for the optimal antenna length. In a second experiment, the polarization of the florescence emission from single molecules was monitored while scanning an antenna in proximity to the sample plane [14]. When the antenna was directly over a molecule, its emission pattern changed to that of the coupled antenna-molecule system, illustrating that it is possible to redirect the dipole emission of a single quantum emitter to match that of a near-field antenna.

Another simple antenna geometry commonly employed is a single gold nanosphere attached to the end of a dielectric probe, such as a pulled glass fiber or an AFM tip [12, 26]. The spherical geometry yields plasmon resonance modes with strong field enhancement at the poles of the sphere along the polarization direction. These antennas have been used as described above to image single molecules and to study the competition between field enhancement and fluorescence quenching. As above, the emission rate of single molecules was recorded as the spherical nanoantenna was scanned in close proximity. As the antenna approached a molecule, the emission of the molecule initially increased due to field enhancement. At very short range ($\sim$ 10 nm), fluorescence quenching overwhelmed this enhancement, leading to a reduction in signal. Under similar illumination conditions using nonresonant, gold-coated AFM tips, only quenching was observed, demonstrating the importance of resonance effects [24].

2.5 Fluorescence Quenching

When an apertureless NSOM probe is applied to a fluorescent sample, the field enhancement mechanisms discussed above can cause an increase in the detected fluorescence signal. Simultaneously, the presence of the tip can decrease the detected fluorescence signal in a variety of different ways. As mentioned, the tip can change the angular distribution of the emission away from the collection angle of the objective. Interference effects from direct laser illumination and excitation light scattered from the tip can also lead to either an increase or decrease in the local field value. Furthermore, the tip can also open up additional channels for photo-excited fluorophores to relax back to the electronic ground state nonradiatively, thereby quenching the fluorescence. This is particularly important for metallic NSOM probes, which can respond to a wide range of wavelengths via dipole-dipole coupling similar to fluorescence (Förster) resonance energy transfer (FRET). In this process, energy is transferred from fluorophore to tip via exchange of a virtual photon, which in turn creates an excitation in the metal where the energy is rapidly dissipated as heat within the tip. Thus, as a tip approaches a fluorescent sample, the local

nonradiative relaxation rate (γ_{NR}) increases, thus decreasing the apparent quantum yield, resulting in a suppression of the detected emission [33, 34]. This fluorescence quenching may be accompanied by a change in the radiative rate γ_R, where both γ_{NR} and γ_R depend on the orientation of the molecule transition dipole moment relative to the probe geometry [34, 35]. This means that the total fluorescence lifetime, $\tau = (\gamma_R + \gamma_{NR})^{-1}$, is altered near a metal surface. By using pulsed lasers and time-correlated single photon counting (TCSPC), it is possible to measure τ directly, even for single molecules, by building up a histogram of fluorescence photon delay times following a laser pulse. This lifetime can then be plotted pixel by pixel to build up an image, where the value of each pixel represents the fluorescence lifetime of the corresponding location. Since a metal tip will alter the lifetime, it can be used as a contrast mechanism in NSOM, as has been demonstrated previously [25, 36]. This type of imaging can provide a great deal of information about near-field interactions between the tip and sample as it provides simultaneous topography, fluorescence, and lifetime data.

There are two general cases to consider with regard to quenching with metal tips: 1) quenching by tips with well-defined, closed geometries such as the spherical, or monopole antennas described above, and 2) quenching by tips with open geometries such as the metal-coated pyramidal AFM probes available commercially. Closed geometries support localized plasmons with well-defined and relatively narrow resonance frequencies determined by the detailed geometry and material of the probe. In this case, the fluorescence quenching efficiency should depend very sensitively on the emission wavelength of the fluorophore, with maximum quenching at the plasmon resonance frequency. For such closed geometry tips, however, the field enhancement is also strongly wavelength dependent, so the competition between enhancement and quenching, and thus the net fluorescence signal, is delicately balanced and can be difficult to predict. For molecular fluorophores, which exhibit relatively small Stokes shifts, the excitation and emission wavelengths may only be as little as 10-20 nm apart, in which case enhancement and quenching can both be quite strong near the plasmon resonance. Colloidal quantum dots, on the other

hand, have a broad absorption profile extending into the UV. Thus it is possible, in principle, to have a large imbalance between field enhancement and quenching, as the excitation laser can be tuned far to the blue of a plasmon resonance, while the quantum dot emission can be at the resonance frequency. Varying the excitation wavelength toward the resonance frequency would then make it possible to study the competition between quenching and enhancement in detail. To our knowledge, such an experiment has not yet been reported.

For the case of tips with open geometries, such as metal-coated pyramidal or conical probes, the situation can be quite different. These probes do not support localized plasmon resonances but can still dissipate energy via damping of traveling plasmon waves. Issa and Guckenburger recently attempted to quantify this effect through the use of finite element analysis using COMSOL [13]. Again it is important to distinguish between two effects: local energy transfer and plasmonic energy losses. True quenching originates from a Förster like local energy transfer from an emitter to a proximate metal structure where energy is locally dissipated through ohmic losses. Generally speaking, field enhancement is proportional to the real part of the effective polarizability (e.g., Equation 2.2, quantity in parentheses), while the imaginary part of the permittivity, ϵ'', is a predictor for the quenching efficiency [12]. Plasmonic losses on the other hand, can occur even in a lossless material, $\epsilon'' = 0$. Energy can be transfered from an emitter to a surface plasmon on the metal structure; the energy need not be dissipated locally, the traveling plasmon may move away where it may eventually be dissipated.

Figures 2.7 and 2.8 from reference [13] show the extent of this effect. Figure 2.7 shows an approach curve as a silver tip vertically approaching a single emitting dipole as a function of initial quantum yield, q_0. For large initial quantum yields, as the tip approaches the apparent quantum yield q decreases. The reverse is true when the initial quantum yield is already low; an increase in the radiative rate can be observed for intermediate tip-sample separations. Other important parameters include the fraction of power radiated into the lower half-space, $\gamma_r^{\downarrow}/\gamma$, the amount of power dissipated via surface plasmon polaritons, γ_{spp}/γ, and the amount of power

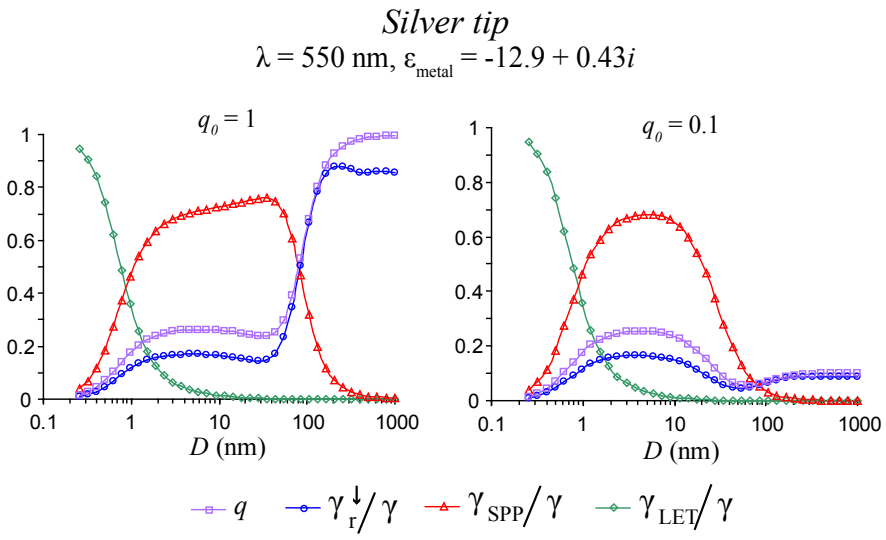

Figure 2.7. Role of plasmonic energy dissipation and initial quantum yield in a lossy metal. Parameters plotted include the apparent quantum yield, the radiative decay rate into the lower half-space, $\gamma_r^{\downarrow}/\gamma$, the local energy transfer rate, γ_{let}/γ, and the nonradiative surface plasmon polariton coupling rate, γ_{spp}/γ. Reprinted with permission from [13].

dissipated by local energy transfer, γ_{let}/γ. Regardless of initial quantum yield, at large tip-sample separation distances surface plasmon polariton coupling can be quite strong, while at shorter distances, the primary source of losses stems from local energy transfer.

The same calculation as in Figure 2.7 is again repeated, but now assuming Ag has no losses, $\epsilon'' = 0$, as shown in Figure 2.8. Such a situation requires that the local energy transfer rate be exactly zero. Despite having $\gamma_{let} = 0$, the apparent quantum yield still drops precipitously as the tip approaches the sample. This drop in q is due entirely to the transfer of energy to surface plasmon polaritons to the metal tip. As the plasmons propagate away from the sample they carry away energy also. Whether or not the plasmons are rapidly dissipated is a moot point as there is power flow directed away from the sample. It is of course possible that the surface plasmons could scatter off some irregularity on the tip surface and emit a real photon; however, this may or may not be detectable depending on the propagation direction of photon, the wavelength of the photon, and how high up

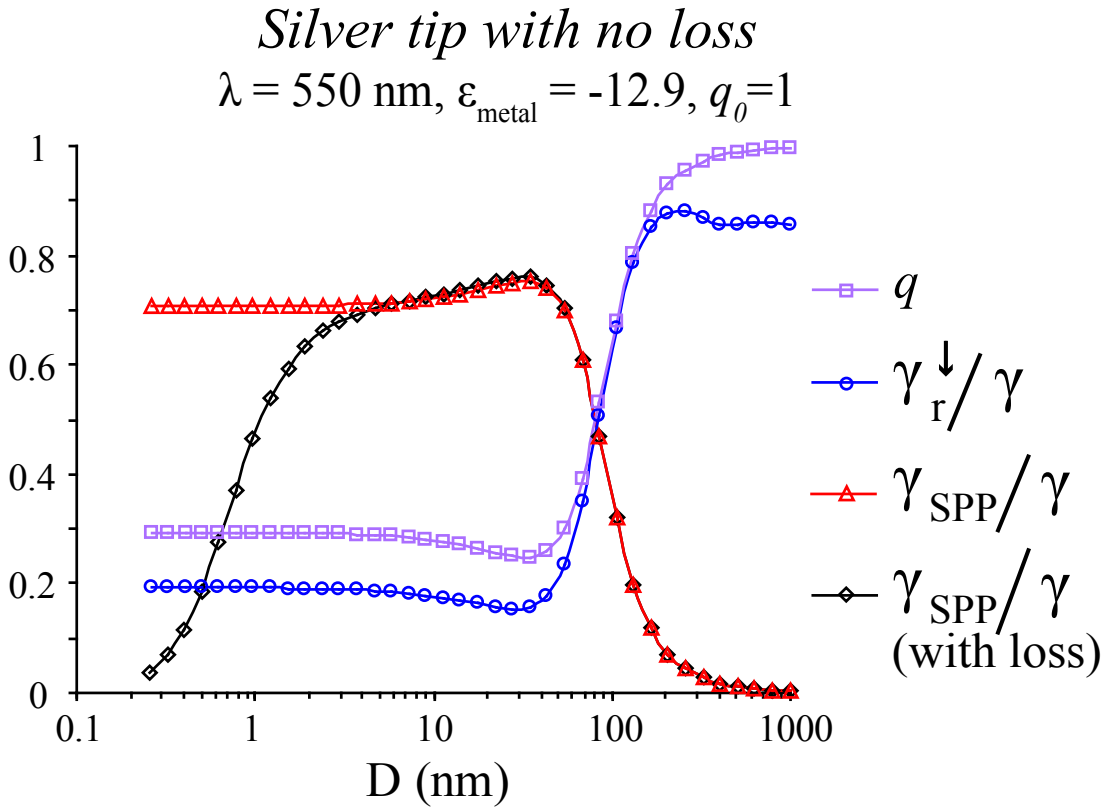

Figure 2.8. Role of plasmonic energy dissipation in a lossless metal. Parameters plotted include the apparent quantum yield, the radiative decay rate into the lower half-space, $\gamma_r^{\downarrow}/\gamma$, the nonradiative surface plasmon polariton coupling rate, γ_{spp}/γ, with no losses, and with losses as shown in Fig. 2.7 for comparison. Note that the the local energy transfer rate, γ_{let}/γ, is exactly zero for all distances and is thus not plotted. Reprinted with permission from [13].

the tip the photon originated.

The difference between the response of the fluorescence signal using tips with closed and open geometries has been demonstrated in a number of recent experiments. For example, Novotny and co-workers [12, 27] have shown that for spherical gold tips, field enhancement is clearly evident leading to a strong enhancement in the fluorescence signal from single molecules, which is mitigated by quenching only when the tip is brought to within $\sim$ 5 nm of a molecule. Using gold-coated pyramidal tips, on the other hand, switches the balance strongly in favor of quenching, often leading to a complete lack of observable enhancement in the fluorescence

signal at any distance scale, as demonstrated by a number of groups [1, 25, 26, 24].

In any NSOM experiment it is important to remember that the effect of the near-field probe on the sample can be non-trivial, especially for metallic probes. In particular, for fluorescent samples, the net detected signal is affected by field enhancement, quenching, and other possible mechanisms such as the redirection of fluorescence. Thus, determining the field enhancement factor of a particular probe is complicated in that it may be impossible to decouple any observed increase in fluorescence signals with any quenching that may also be occurring. In general, this would require a rather sophisticated model of the system, which is highly dependent on the probe geometry and material.

2.6 Field Enhancement vs. Signal Enhancement

In near-field literature, the term *enhancement* is sometimes used somewhat ambiguously and at times perhaps even incorrectly. Enhancement may refer to either field enhancement or to signal enhancement; however, it is usually incorrect to use these terms interchangeably. To first order, field enhancement deals simply with the interaction of the tip and laser. Signal enhancement, on the other hand, can be proportional to field enhancement, but includes the sum of many other effects such as fluorescence quenching, redirection, interference effects, and very importantly, the field-sample volume overlap. Calculating the peak field enhancement of a tip alone yields a somewhat incomplete picture; that is the spatial extent to which the field overlaps with the volume of the sample must be taken into account to determine the amount of signal enhancement to be expected.

In the same model as Figure 2.5, the intensity-volume overlap is calculated by integrating the intensity contained in several different spherical volumes directly under the tip as shown in Figure 2.9. This represents the integrated intensity that a spherical sample might see when the tip is directly on top of it. This integrated intensity was then normalized to both the integrated intensity under similar illumination in the absence of the tip and the peak intensity enhancement under the tip. The results are summarized in Table 2.1, where the intensity

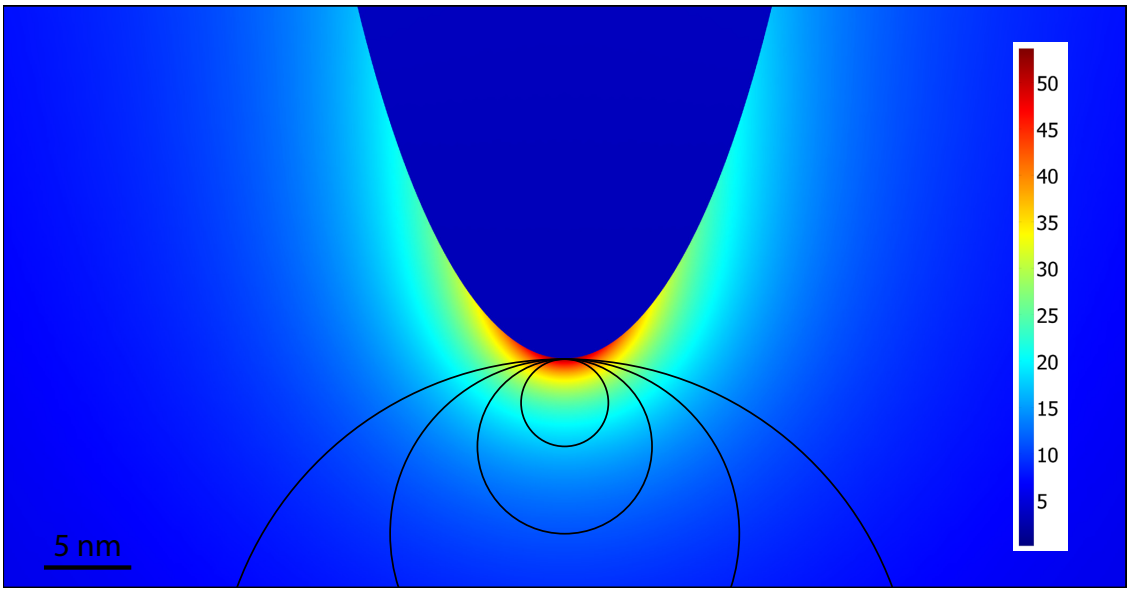

Figure 2.9. Calculation of field-volume overlap integrals. A 3D electrostatic calculation of field enhancement around a near-field probe was performed in Comsol as described in Fig. 2.5. The scale bar represents the magnitude of the field, where the incident field is 1 V/m. Integration of the *intensity* was performed at spheres of various sizes indicated by solid black lines.

enhancement factor is reported as a percent of the peak intensity below the tip. The radii chosen for this calculation correspond to some typical samples measured in the lab: primarily 5 nm diameter quantum dots, and 20 nm diameter dye-doped latex beads. This calculation corroborates the data reported previously by Gerton et al. [2] where a 5 nm diameter quantum dot exhibited roughly 4 times the *signal* enhancement as a 20 nm dye-doped bead, which is similar to the predictions made in Table 2.1.

While to first-order, field enhancement is purely a function of tip-and excitation fields, tip sample coupling can also play an important role. As a tip approaches a sample, it no longer is rigorously correct to discuss the polarizability of the tip alone, but rather, tip-sample coupling effects require an effective polarizability of the tip-sample complex. An analytic solution for tip-sample coupling between a spherical tip and a sample (an infinite plane) has been worked out some time ago by Knoll and Keilmann [37]. In fact for scattering SNOM this tip-sample coupling

Table 2.1. Intensities were integrated over the volumes of spheres of several different radii as shown in Fig. 2.9. This integrated result was normalized to both the value obtained in that region in the absence of a tip and to the peak intensity at the apex of the tip. The normalized intensity enhancement factor is reported as a percent of the peak intensity. The results were normalized in this fashion to ensure they are not sensitive to the particular tip height chosen for the calculation.

Radius	Normalized Intensity Enhancement Factor
20 nm	2%
10 nm	5%
5.0 nm	11%
2.5 nm	23%

is the entire signal of interest [37].

$$\alpha_{eff} = \alpha\left(1+\beta\right)\left(1-\frac{\alpha\beta}{16\pi\left(a+z\right)^3}\right)^{-1} \tag{2.7}$$

where α is the polarizability of the tip (or the tip modeled as a sphere),

$$\alpha = 4\pi a^3 \frac{\epsilon_1 - 1}{\epsilon_1 + 2} \tag{2.8}$$

β is the dielectric response function of the sample, $\beta = (\epsilon_2 - 1)/(\epsilon_2 + 1)$, z is the tip-sample separation distance, a is the radius of the tip, ϵ_1 is the permittivity of the tip, and ϵ_2 is the permittivity of the sample. It must be remembered that this calculation assumes the sample is an infinite plane; deviations to these approximations of course lead to nonanalytic results, however, such a calculation may still be useful.

Also included in the signal enhancement is any redirection of fluorescence signals by the tip. It has been demonstrated that the presence of a tip can redirect the fluorescence emission of a sample [38]. Such redirection can be amplified if the probe exhibits a strong antenna effect [14]. Aside from redirection, the tip can also produce long range interference effects, also contributing to an observed signal enhancement. As the light scatters from a probe interferes with the direct laser illumination, complicated tip-geometry dependent patterns can emerge such that the fluorescence signal can be either increased or decreased corresponding to regions of constructive or destructive interference from the tip [39, 40, 41, 42, 24] With

the signal enhancement being a function of so many differing and widely variable parameters, it is often unclear a priori how a given tip might effect the observed near-field signal.

2.6.1 Examples

This interplay between field-enhancement and signal enhancement is well illustrated through a few examples. In particular we have observed that platinum coated tips can lead to interesting results. Using a platinum/iridium coated Si tip with a vertical wedge illumination, we expect to obtain some field enhancement at the tip apex. However, due to the reasons just stated, we also expect strong quenching from such probes. Figure 2.10 plots the fluorescence signal (normalized to the far-field signal) as a Pt/Ir tip is lowered down onto an elongated CdSe/ZnS quantum dot, which is nominally 4×9nm. As is shown the result shows no evidence of enhancement at any distance, but rather shows strict quenching.

For comparison, the same tip was also used to generate an approach curve on a 20 nm diameter dye-doped latex bead, Figure 2.11. Since the bead has a much larger

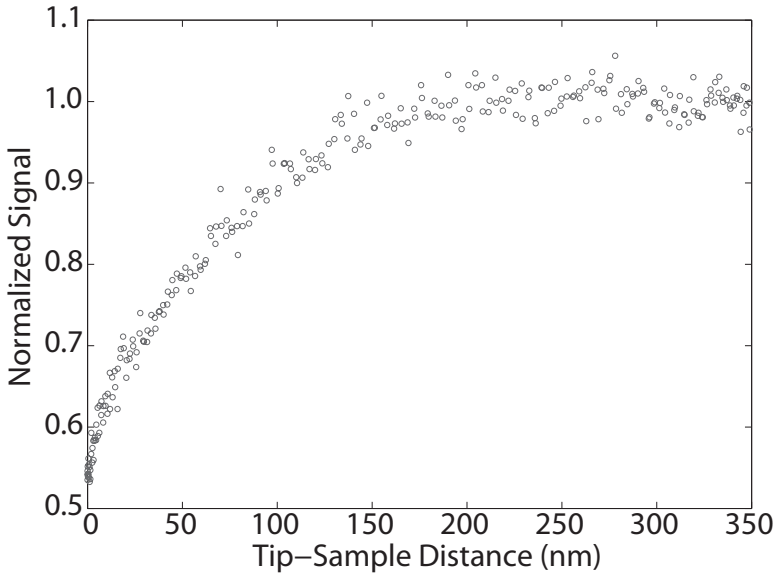

Figure 2.10. Approach curve of a Pt/Ir tip on a 4×9 nm CdSe/ZnS quantum dot. Vertical wedge illumination (543 nm) was used such that some *field* enhancement would be expected.The fluorescence rate has been normalized to the far-field value.

volume we may expect that any enhancement effects would be reduced. However, as seen in Figure 2.11 there is some "recovery" of the fluorescence signal at short length scales. Note that this recovery is not present when using horizontal wedge illumination. Since fluorescence quenching is expected to dominate all interactions at the shortest length scales, this result can be even more perplexing. The explanation is that the sample is an extended object with a roughly uniform distribution of fluorophores throughout. As the tip approaches some fraction of the fluorophores are quenched to different extents while others are more preferentially enhanced. This of course is evidence that there is indeed some field enhancement present in this setup, which was not at all apparent when using the quantum dot. These two competing effects occur at roughly the same length scales [12] and thus can be very difficult to sort out experimentally The overall message is that the total near-field signal collected can be much more complicated than estimating the field enhancement a tip may produce.

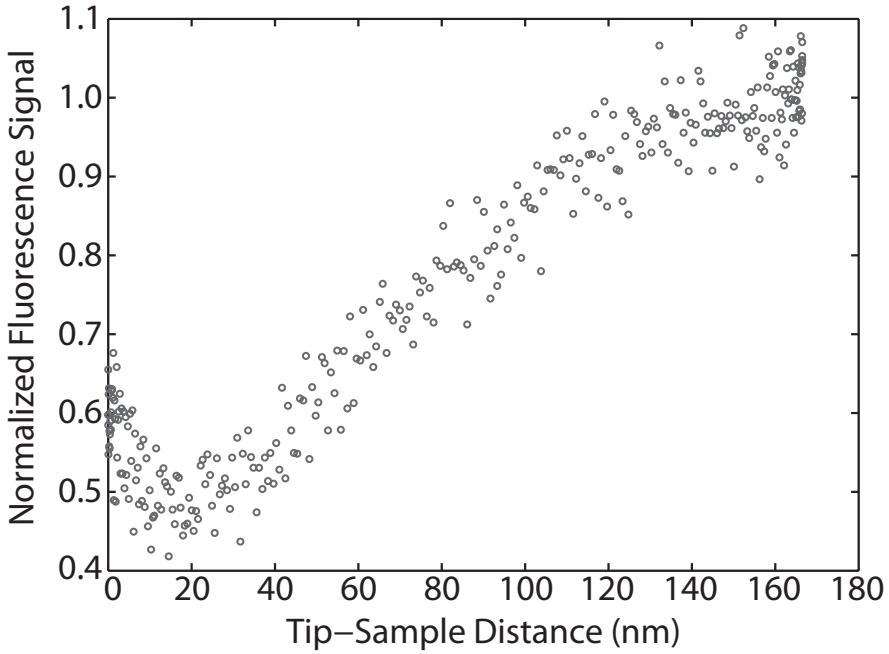

Figure 2.11. Approach curve of a Pt/Ir tip on a 20 nm diameter dye-doped latex bead. Vertical wedge illumination (543 nm) was used such that some *field* enhancement would be expected. The fluorescence rate has been normalized to the far-field value.

2.7 Tip Selection

In order to obtain the strongest near-field signal, optimization of the tip's enhancement factor is essential. There are generally three different types of tips employed in apertureless-NSOM experiments: commercial cantilever-based AFM tips made of silicon, silicon nitride, or metal-coated silicon; spherical metal nanoparticles attached to the distal end of a pulled glass fiber or commercial cantilever-based AFM tip [43]; and electrochemically sharpened metal wires (gold, tungsten, silver, etc.) [44]. These various probes have different strengths and weaknesses, and the choice of one over another depends in large part on the particular optical process to be employed during the experiment.

The key factor in choosing the most appropriate tip in fluorescence imaging is to optimize the competition between field enhancement and quenching, as described above. Although metal tips can generate extremely high field enhancement factors, particularly near a plasmon resonance frequency, they also quench fluorescence at very short range. This competes with field enhancement, thereby reducing the net fluorescence signal. Generally speaking, field enhancement is proportional to the real part of the effective polarizability (e.g., Eq. 2.2, quantity in parentheses), while the imaginary part of the permittivity, ϵ'', is a predictor for the quenching efficiency [12]. The precise dependence of quenching and enhancement on the complex permittivity is also very sensitive to geometry.

For metals, both the effective polarizability and ϵ'' can become very large in magnitude near a plasmon resonance frequency. Thus it is difficult to predict what the net signal enhancement will be for an arbitrary geometry, and there are only a few rigorous calculations that have been compared with experiment [12, 27, 35], and only for a simple spherical geometry. Nonetheless, all fluorescence experiments thus far with metal tips have exhibited strong quenching, and thus reduced signal, at very short tip-sample separation distances. Tips composed of small metal spheres (attached to dielectric probes) can support strong localized plasmon resonances, and can thus exhibit appreciable signal enhancement outside this quenching zone (see below). Metal-coated tips with extended geometries (e.g., commercial metalized

probes), on the other hand, do not support localized plasmon resonances, and have exhibited no net enhancement of the fluorescence signal at any length scale [1, 25, 26, 24].

Silicon has a complex permittivity of $\epsilon_{Si} = 17.6 + 0.12i$ (at $\lambda = 532$ nm) [45], indicating the potential for good fluorescence signal enhancement, with only minimal quenching. Furthermore, silicon tips can be made quite sharp, particularly compared to metal-coated tips, so the lightning-rod enhancement and resulting increase in spatial resolution should be quite good. Indeed, silicon tips have exhibited fluorescence signal enhancement factors as large as $f \sim 20$ [2], and spatial resolution as small as ~ 10 nm [46, 7, 3, 8, 24, 47]. As described above, this enhancement factor should provide adequate contrast to image even rather complex samples with high background signals. A literature search of dielectric constants at visible frequencies for commonly available materials indicates that silicon gives the largest enhancement factor, although reliable optical constants for some materials are difficult to obtain. Figures 2.12 and 2.13 highlight the permittivities for some of the more common AFM probe materials; Figure 2.12 shows dielectric materials while metals are shown in 2.13.

Silicon has by far the highest ϵ' of the dielectrics; however, it is also the only dielectric to have a nonzero imaginary part ϵ'' at visible wavelengths . The imaginary parts of the permittivity for SiO_2, Si_3N_4, and diamond are exactly zero in this wavelength range and thus are not plotted. Evidently, silicon appears to be the best, or at least most straightforward, probe material for fluorescence imaging, and there are a variety of probe geometries available, most of them cantilever-based. Super-sharp silicon AFM probes have been used to obtain large signal enhancement in the past [7] but they also suffer from rapid wear, which leads to large variations in their performance.

Metals can also yield exceedingly large field enhancements. Metals can have high enhancement values for two reasons: first, the real part of the permittivity can be very large (albeit usually negative at visible wavelengths), and second, the metal tip can support a plasmon resonance (cf. Fig. 2.13). In the case of a

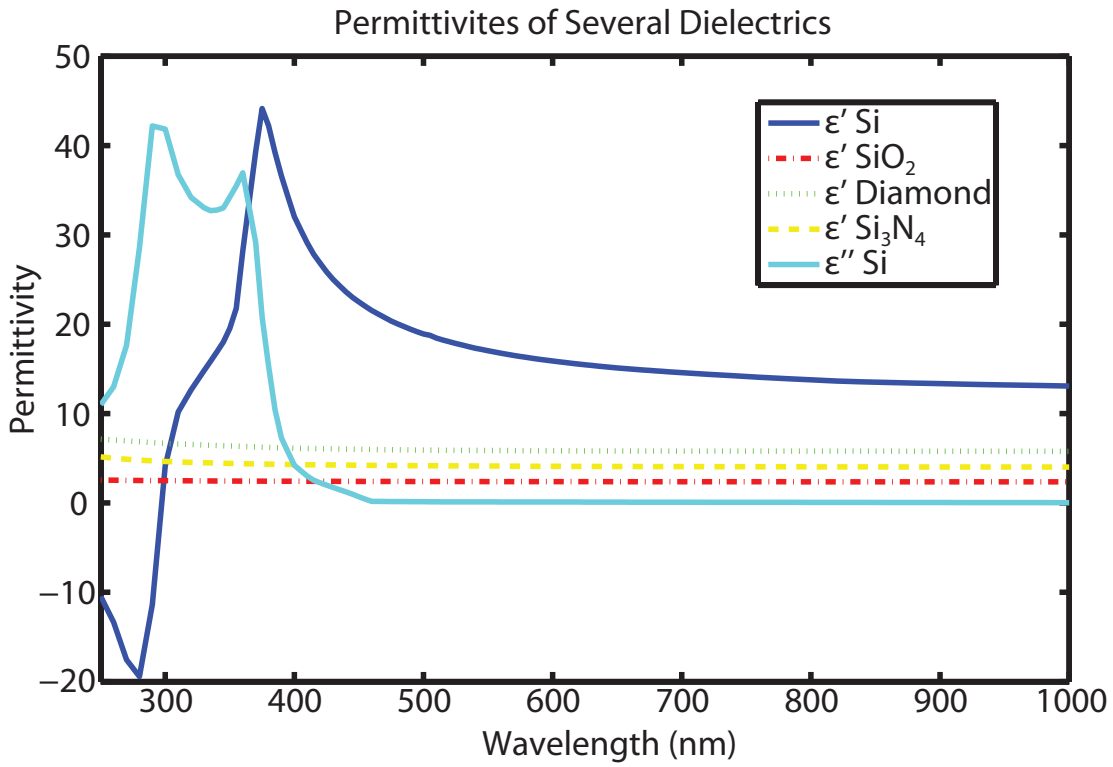

Figure 2.12. Real and imaginary permittivities plotted for some more common dielectric materials used for AFM probes. Permittivity values are found in the following references: Si [45], SiO_2 [48], Si_3N_4 [48], and diamond [49].

sphere, the plasmon resonance occurs when $\epsilon_r = -2$ (cf. Eq. 2.2). Metals also exhibit considerable quenching as the imaginary parts of their permittivities can be considerable (cf. Fig. 2.13).

In our experience, standard silicon "tapping mode" probes yield very acceptable and repeatable results. One problem with silicon probes is the growth of oxide layers, which do not exhibit strong polarizability at optical frequencies, as compared to Si. All silicon tips have some native oxide layer, but as tips age this layer thickens. Using fresh probes alleviates this problem somewhat, but it can be difficult to determine their exact date of manufacture. Storing tips in vacuum chambers is definitely recommended, and some manufacturers have begun shipping probes in hermetically sealed packages of fewer quantities to avoid unnecessary oxidation. Some studies have also indicated that contaminants from gel-pack off-gassing may

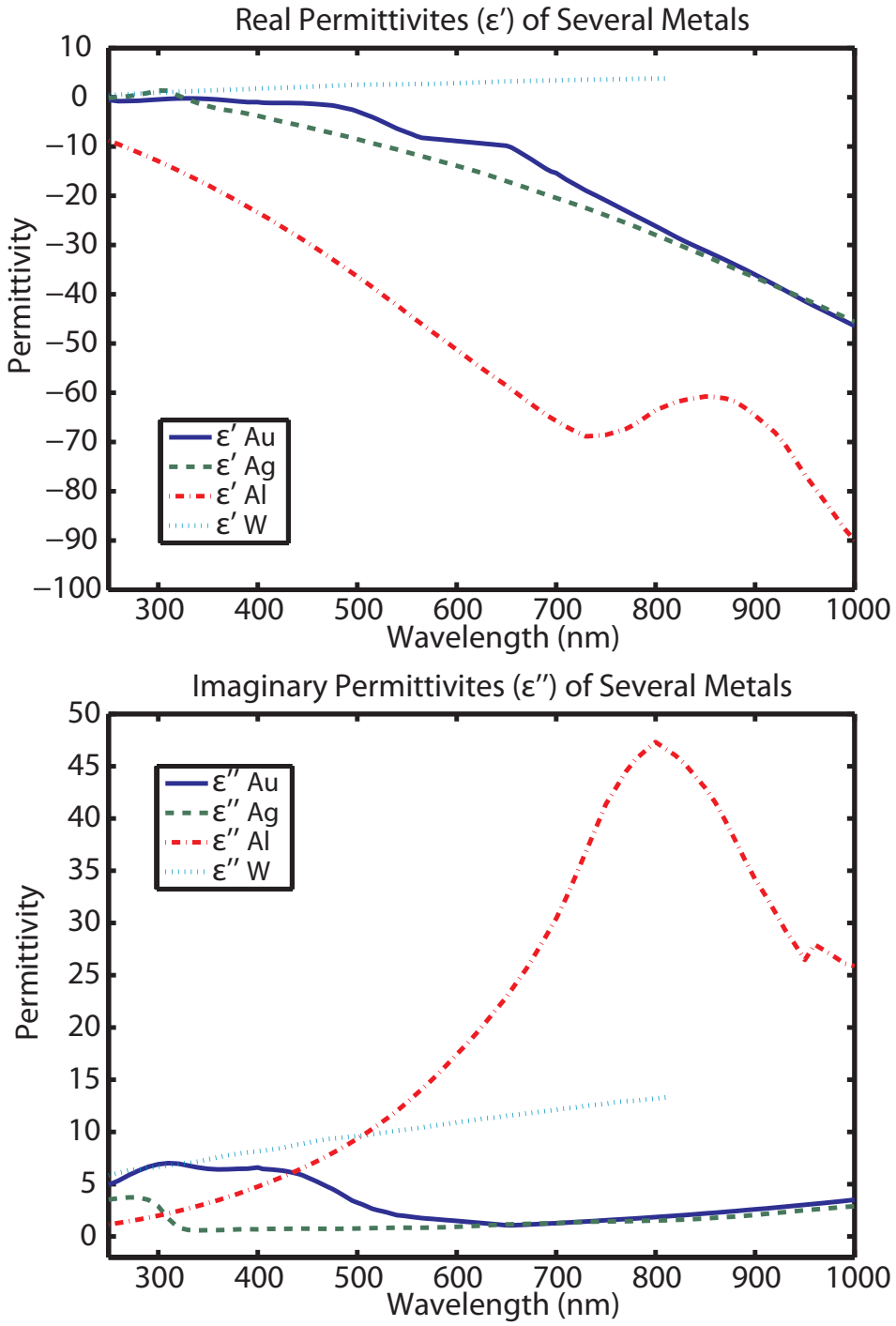

Figure 2.13. Real and imaginary permittivities plotted for some more common metal materials used for AFM probes. Permittivity values are found in the following references: Au [50], Ag [48], Al [48], W [51].

also contribute to a reduction in enhancement factor [46]. It is widely agreed that oxidized or contaminated tips may be "revived" to some extent via etching in hydrofluoric acid (HF) [46, 41]. Standard buffered oxide etch (BOE) procedures prescribe the rate at which silicon oxide layers can be eaten away [52]. This procedure, while effective in removing oxide layers, can also dull the AFM tip, and it should thus be applied carefully and conservatively. Other cantilever-based probes of interest that have currently become available include diamond-like carbon (Mikromasch) and carbon nanotube tips (nanoScience).

While more difficult to produce, metal nanospheres attached to dielectric tips [43] can also yield excellent results [12, 26, 27]. Because these probes can support localized plasmon resonances, they can yield very large field-enhancement factors that can overcome the reduced signal caused by fluorescence quenching beyond some critical tip-sample separation distance. This is clearly the case in Figure 2.14 where an 80-nm gold sphere is used to probe a vertically oriented single molecule; at very close range quenching overpowers the field enhancement. Thus, to optimize image contrast, these tips should be maintained at this critical height from the sample ($\sim$ 5 nm), where the signal enhancement is maximized (cf. [12]). To maintain a constant tip-sample gap, the AFM should be operated in shear-force rather than tapping mode. When using tips composed of spherical nanoparticles, both the size of the particle and the metal to be employed should be chosen so that the localized plasmon resonance frequency is close to the absorption peak of the fluorophore: the resonance frequency is determined by both the size and permittivity of the particle. It has been shown that gold nanoparticles perform better at red wavelengths and silver nanoparticles at blue wavelengths [27]. Smaller particles yield higher resolution but also smaller enhancement [12, 23]: most reports utilize diameters in the 40-80 nm range [12, 26, 18, 27, 53].

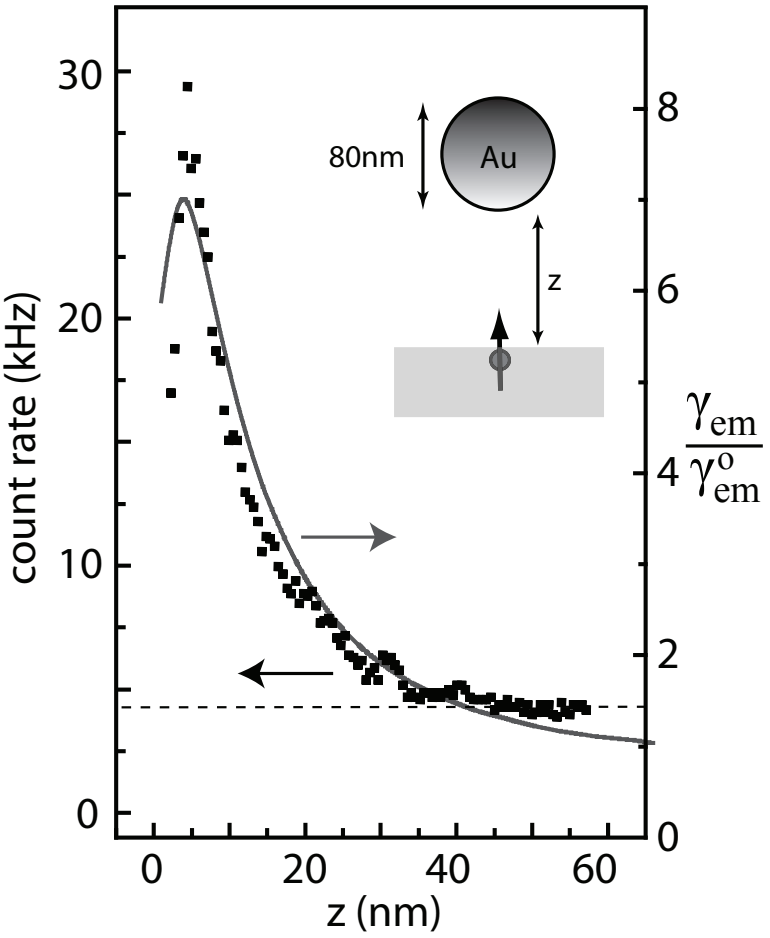

Figure 2.14. Fluorescence rate as a function of particle-surface distance z for a vertically oriented molecule. Dots represent experimentally observed count rates (left axis). The solid curve shows the theoretical normalized emission rate compared to free space, $\gamma_{em}/\gamma^0_{em}$ (right axis). The horizontal dashed line indicates the background level. Reprinted with permission from [12]. Copyright (2006) by The American Physical Society.

CHAPTER 3

EXPERIMENTAL SETUP

Ordinarily a section on the experimental setup may not receive its own chapter, but due to the large amount of effort and constant refinement needed for TEFM, I will dedicate an entire chapter to that end. As alluded to earlier, in order for TEFM to work, a nanoscale probe needs to be aligned to the central ~ 100 nm of a laser spot. Furthermore, the polarization state of this excitation beam needs to be both carefully controlled and alterable.

The microscope system developed in our lab is a multipurpose imaging tool with several distinct functionalities. It has the ability to operate as an inverted confocal microscope in both sample scanning and laser scanning configurations. It can be used as a standalone Atomic Force Microscope (AFM) and most significantly as an apertureless Near-Field Scanning Optical Microscope. This microscope system has been designed in such a way that switching between different imaging modes is meant to be a relatively straightforward task; however, as with any system, increasing functionality necessarily goes hand in hand with increasing complexity. Thus in order to understand the operation of the system, a knowledge of both some basic theory and engineering limitations must be obtained.

3.1 Setup

The basis for the system is an AFM sitting atop a custom built inverted optical microscope (see Fig. 3.1). The AFM is a commercially available system (Asylum Research) that includes a high precision piezo-actuated X-Y scanning stage and an AFM head. This AFM head controls the movement of nanoscale probes that scan the sample and is capable of moving the AFM probe only in the vertical (Z) dimension. As seen in Figure 2.1, the inverted optical microscope has a choice

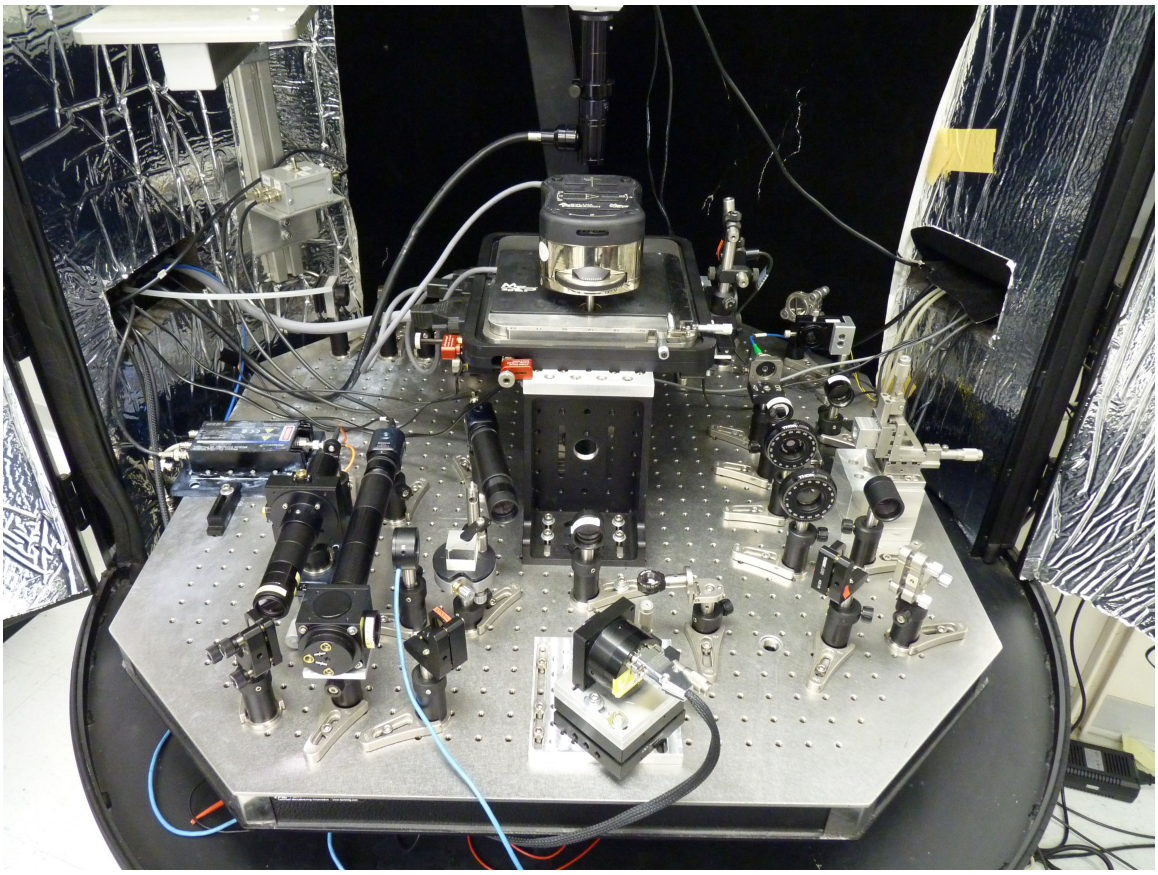

Figure 3.1. Photograph of experimental setup.

of different illumination pathways depending on the polarization state required. Excitation pathways for either a radial polarization state or a linearly polarized polarization state (both of which will be described in greater detail below) can be implemented through the use of removable face-plate mirrors.

3.2 AFM

The AFM is a commercial cantilever based unit from Asylum Research (MFP3D). Again, there are two essential components of this system: the AFM head itself, along with a piezo-actuated scanning stage. The AFM employed in our lab is cantilever based, meaning that the probes themselves are mounted at the end of a long (usually $\sim 200\ \mu m$) thin plank. A feedback laser reflects off the backside of the cantilever onto a segmented photodiode. The laser is weakly focused so

that the spot size is about the same size as the width of the cantilever, but as the laser bounces off the cantilever it is divergent. After reflecting off the cantilever, the beam strikes a photodiode with four segments (see Fig. 3.2); each segment puts out a voltage relative to the incident light intensity. Since the beam acts as a long lever arm, small displacements of the tip result in large displacements of the beam on the quadrant photodiode; in this manner, minuscule tip-sample interactions are measured faithfully, thus the name Atomic Force Microscope. The two main modes of operation of the AFM are contact mode, where the probe is simply dragged across the surface, and tapping mode, where the tip undergoes rapid vertical oscillations only tapping the sample surface intermittently. Tapping mode is used almost exclusively in our lab for two important reasons: it can be employed in such a manner that it is much gentler on the sample (also leading to less tip-wear), and it can lead to higher near-field contrast as will be explained in later chapters. While most of the important details of operation of the AFM can be found in the operation manual, I will discuss a few important points that are

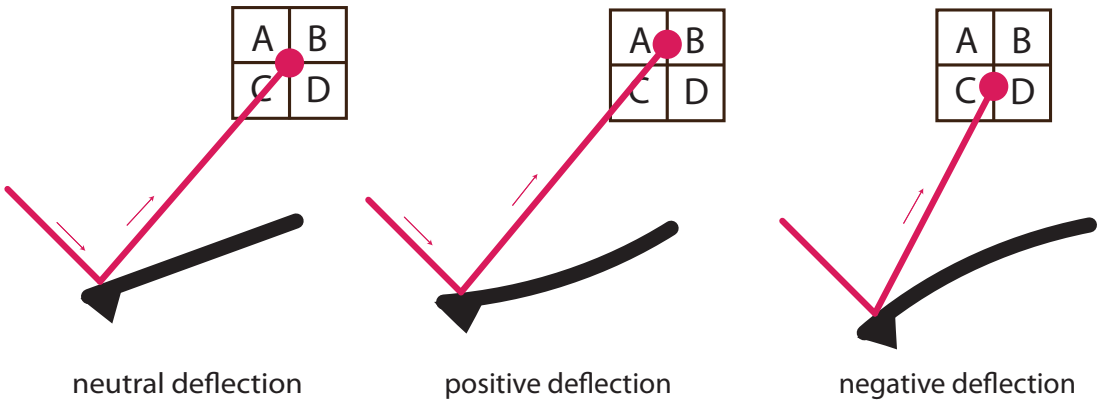

Figure 3.2. Cartoon of AFM feedback mechanism. A laser is incident on the back side of a cantilever based probe. The beam reflects off the probe onto a segmented photodiode. Each segment produces a voltage in proportion to the amount of light detected. Small changes in the cantilever deflection lead to large changes in beam displacement on the detector, thus yielding high sensitivity. Positive or negative deflections can be monitored via the photodiode output signals and converted into real displacement when a probe is properly calibrated.

important in our research.

3.2.1 Tip Calibration

The primary mode of AFM operation in our lab is tapping mode (other monikers include AC mode or intermittent contact mode). The amplitude of these tip-oscillations must be carefully calibrated: the amplitude determines the force applied when tapping, the amount of near-field contrast we can obtain, and is important for understanding the length scales of optical responses on the nanometer scale. The oscillation amplitude can be set through the software user interface and is in volts; however, an important conversion factor needed is the Inverse Optical Lever Sensitivity or InvOLS. This conversion factor can vary depending on the particular brand (geometry) of tip being used and how it is mounted. Furthermore, it depends on the sum signal from the quadrant photodiode ($V_A + V_B + V_C + V_D$); the larger the sum signal the smaller the InvOLS value. The largest difference comes from the fact that some cantilevers are bare silicon, while others are metal coated and thus much more reflective—leading to a larger sum signal. From much practice we can get a rough estimate of the InvOLS based on the sum alone, the results are summarized in Table 3.1. While Table 3.1 may be used a starting point, it must be remembered that these are only estimates and an InvOLS measurement should

Table 3.1. AFM Sum signal and corresponding approximate InvOLS values. Two classes of tips are routinely used in the lab, Si tips with no backside coating and metallic tips, or tips with a metal backside coating. These two classes typically give Sum values of 3-4 and 9 respectively, which is why no data are shown for intermediate Sum values. Note, the InvOLS values listed here are for converting the oscillation *amplitude* in volts to nm—the peak to peak amplitude being twice the values listed here. The deflection InvOLS can be calibrated in a similar way. As a rule of thumb: AmpInvOLS = 1.09·DeflInvOLS.

Sum	Amplitude InvOLS
9	70 nm/V
8	90 nm/V
5	170 nm/V
4	210 nm/V
3	240 nm/V

be performed for each tip used.

3.2.2 AFM Scanning Stage

The scanning stage is a two-layer system. Adjustment of the lower layer keeps the AFM and sample registered together but moves them both relative to an opening on the bottom allowing for optical measurements, while adjustment of the upper layer of the stage moves only the sample It is this upper layer of the stage that scans the sample during AFM operation. This upper stage is operated in closed loop mode and has a maximum scannable area of 90 $\mu m \times 90$ μm but also has hand adjustable micrometer screws for coarse positioning.

The lower stage, which adjusts the AFM head and sample in tandem, relative to the optic axis originally came with only hand adjustable micrometer screws. These have been replaced with piezo-actuated micrometers (Pico-motors from New Focus), which work by rapidly turning a drive screw then slowly relaxing back into place. The total travel of the screw can be quite large; we have purchased models with both 1/2 inch and 1 inch travel, while the minimum step of each piezo movement is ~ 20 nm. The picomotors can be actuated remotely via a computer interface or by a joystick. Tip-sample alignment can be achieved using these picomotor actuators and will be discussed in more detail later in this chapter.

3.3 The Optical Train

As mentioned, multiple excitation paths allow for easy selection of various polarization states and spatial modes for the excitation laser beam. The two options primarily employed are radial polarization and linearly polarized light; however, in each such path further changes to the spatial mode of the beam are also often made such that the beam is clipped or masked in particular ways. In order to understand the motivation behind the tailoring of the laser beams in our system, I will briefly discuss the fundamental differences between some the basic laser modes.

3.4 Transverse ElectroMagnetic Modes

In laser cavities there can exist both longitudinal modes and transverse modes of oscillation. The longitudinal modes correspond to standing waves in the cavity. For cavities with rectangular symmetry (containing Brewster windows or prisms), the transverse modes are usually called TEM_{mn} modes, TEM standing for Transverse Electric and Magnetic, while the m and n subscripts refer to number of nodal lines in the X and Y directions of the beam respectively [54]. Mathematically these beams can be described by:

$$I_{m,n}(x,y) = I_0 \left[H_m \left(\frac{\sqrt{2}x}{w} \right) exp \left(\frac{-x^2}{w^2} \right) \right]^2 \left[H_n \left(\frac{\sqrt{2}y}{w} \right) exp \left(\frac{-y^2}{w^2} \right) \right]^2 \quad (3.1)$$

where the subscripts m and n refer to the order of the Hermite polynomial H, and w is the width of the Gaussian profile. The corresponding intensity profiles can be seen in Figure 3.3.

The TEM_{00} mode is a standard laser mode produced by most commercially available lasers; it has a Gaussian profile and linear polarization state. Of special importance to our lab is the TEM_{01} mode. The superposition of the TEM_{01} mode with another TEM_{01} mode rotated at 90° to the first creates a beam with a radially shaped polarization profile as seen in Figure 3.4. This mode is often described as a "doughnut" mode as the intensity at the center of the profile is exactly zero; oftentimes it is simply referred to as a TEM_{01^*} mode. Note that the TEM_{01^*} beam is sometimes described to be a superposition of a TEM_{01} beam with a TEM_{10} beam, which is incorrect. It is also important to bear in mind that the electric field of the beam is of course oscillating. That is, while Figure 3.4 portrays a snapshot of the cross section of a collimated TEM_{10} beam at one moment in time as pointing radially outward, it is actually rapidly oscillating between radially outward and radially inward at the optical frequency.

3.5 Gaussian Illumination

The most common illumination scheme in many different types of microscope configurations is to illuminate a sample with a linearly polarized Gaussian profiled

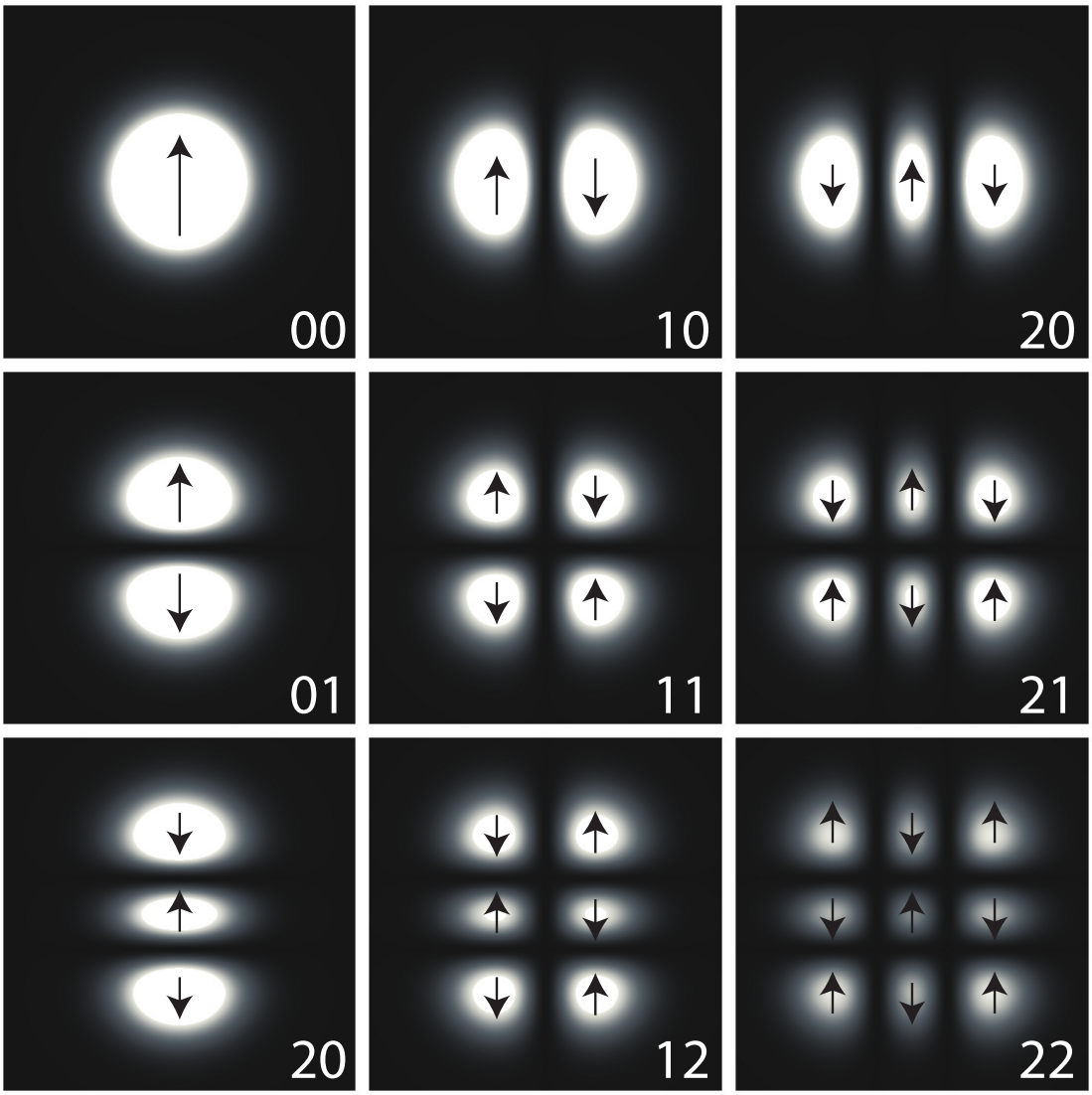

Figure 3.3. Computer generated images of TEM modes of Hermite-Gaussian beams calculated from Eq. 3.1. The subscripts *m,n* refer to the horizontal and vertical number of nodal lines respectively. Arrows representing the instantaneous polarization direction have been overlaid.

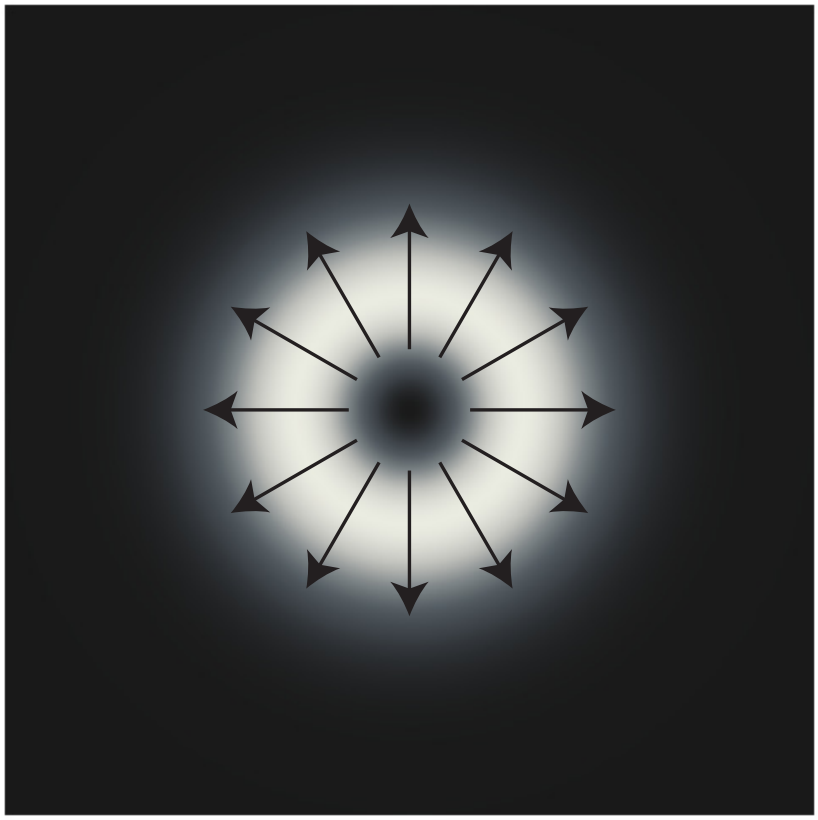

Figure 3.4. A computer generated cross section of a radially polarized beam. Arrows represent the polarization direction (the instantaneous direction of electric field vectors).

laser beam, or in other words a TEM_{00} mode beam. This type of beam is found at the output of most commercially available lasers. This mode also couples very efficiently to single mode optical fibers, whose output again is a TEM_{00} beam. The Gaussian path is the simplest to achieve in that the output of a single mode fiber is simply collimated using a short focal length lens to achieve the correct beam diameter. A commercially available out-coupler is currently used for this purpose (Optics For Research PAF-X15); however, the lens used in this commercial fiber out-coupler is extremely non-achromatic, meaning that when changing between different wavelengths, considerable adjustments must be made to reachieve a collimated beam, which consequently may change the diameter as well. Inasmuch as all commercially available out-couplers seem to suffer from

such problems, we have created a custom-built out-coupler consisting of a slightly larger focal length achromatic lens, mounted in a lens tube and screwed onto a flexure mount where the optic fiber terminates. Due to geometrical constraints focal lengths of ~ 30 mm are most appropriate, but consequently produce a large beam diameter ~ 14 mm. Due to this large beam diameter, the custom out-coupler is not currently used in the Gaussian beam path, but is used on the radial beam path.

For all paths used in our setup linearly polarized TEM_{00} laser beams are coupled into polarization maintaining optical fibers (OZ Optics) using fiberports of various focal lengths from OFR. Following the out-coupler, the next two elements in the Gaussian optical path are a half-wave plate followed by a linear polarizer. By rotating the wave plate and polarizer appropriately, the direction of the polarization can be changed to any arbitrary angle without reducing intensity.

3.6 Axial Polarization

The two basic requirements to make most near-field optical measurements a reality include first, positioning the near-field probe into the focus of a laser beam as discussed above, and second, the excitation beam must have axial polarization; That is, the electric field direction must be pointed parallel to the long axis of the probe. Several different illumination schemes have been introduced to achieve such a requirement. One such configuration is to illuminate the probe and sample from the side at a shallow angle such that a linearly polarized beam may be employed. However, most fluorescent near-field measurements utilize an inverted microscope setup such that the excitation beam is at normal incidence to the sample plane (cf. Fig. 2.1). This episcopic scheme is advantageous in that it allows for the use of high numerical aperture objectives to create tightly focused laser spots. Unfortunately, this episcopic configuration requires special beam tailoring to achieve axial polarization at the sample plane. There are principally only two configurations capable of meeting this requirement: radial polarization and evanescent illumination.

3.6.1 Radial Polarization

There are several schemes to achieve radial polarization state. One commonly employed scheme is to carefully cut and glue half-wave plates into quadrants with fast-axis orientations as indicated in Figure 3.5. A beam with linear polarization can be converted into one with quasi-radial polarization as it passes through such a device. If this beam is then passed through a spatial filter, a pure radial mode can be extracted.

One downfall of the wave plate method is that it is not a broadband device, as each wavelength employed would require its own quadrant wave plate. To circumvent this problem we have come up with an alternative scheme along these same lines, using linear polarizers rather than wave plates. Since polarizers do not actually rotate the beam itself, this can no longer be achieved by one optic component. However, by layering polarizers as diagrammed in Figure 3.6, a similar result to that of wave plates can be obtained. The trade-off of course is that while polarizers can yield a broadband device, they also result in a loss of power (50% for each layer) such that a quasi-radial beam of four quadrants would only contain ideally 12.5% of the input power (assuming no other losses through the polarizing material). Thus far we have been able to achieve a prototype of this

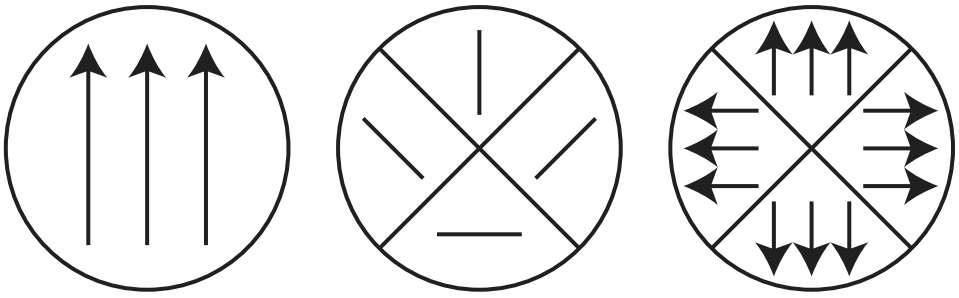

Figure 3.5. Cartoon diagramming the production of quasi-radial polarization. Linearly polarized light as shown in cross section on the left is incident on a custom half-wave plate, with the fast axis of each quadrant as shown in the middle panel. The result is a quasi-radial beam with polarization as shown on the right. True radial polarization may be obtained by spatial filtering this resultant beam.

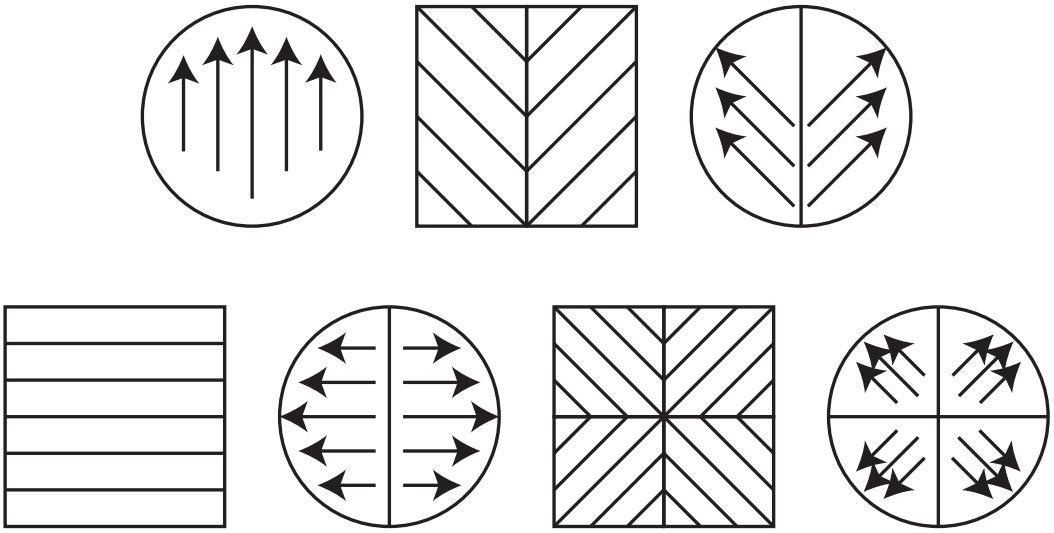

Figure 3.6. Cartoon diagramming the production of quasi-radial polarization using polarizers. Linearly polarized light as shown in cross section on the left is incident on a custom polarizer, with the transmission axis of each polarizer as shown in the square panels. The result is a quasi-radial beam with polarization as shown in cross section as the last (bottom-right) circle. True radial polarization may be obtained by spatial filtering this resultant beam.

radial polarization converter using polarizers. A patent for this device has been applied for, and is still being processed.

In a similar but more continuous fashion as the wave plate device, a commercial device using twisted nematic crystals can also produce a mode with nearly perfect radial polarization [55], which can be further refined via a spatial filter. There are two layers to this device. One layer covers only half the beam and acts as a half-wave plate to create anti-parallel field components in one half of the beam with respect to the other. The other layer rotates the beam by varying degrees in space to achieve radial polarization. This device is being used in the most current iteration of our near-field optical microscope; it also has an option that allows for producing an azimuthally polarized state. One disadvantage of this device is that it requires two input signals (square waves at two different amplitudes) for use. After a beam passes through the radial polarization converter (RPC), the now radial

beam is refined by passing through a spatial filter.

The choice of the spatial filter is crucial to obtaining good results (high power throughput and a good quality laser mode). A simple way for obtaining the spatial mode of a beam focused by a lens is simply to take the Fourier transform of the initial beam profile. It is well known that the Fourier transform of a Gaussian function yields another Gaussian function, and an inverse Fourier transform on a Gaussian yet again yields a Gaussian function. In this way as light is focused and re-collimated it maintains a Gaussian profile throughout. However, any imperfections to the mode arising from scattering off of sharp edges or dust lead to high spatial frequency components. Spatial filtering simply entails focusing light through a pinhole such that the higher frequency components of the beam become blocked by the pinhole and a clean Gaussian profile emerges. Typically the diameter of the pinhole employed in a spatial filter should match the beam waist of the focal spot in order to optimize throughput $D = F\lambda/a$, where a is the radius of the beam waist, as recommended by Newport.

The radial mode resultant from the radial polarization converter, like any beam profile, can be thought of as a superposition of high and low spatial frequency modes as would be expected in any Fourier sine or cosine series. Since there is a predominant defect line in the RPC, there also exist many high frequency components that must be removed via spatial filtering. This is made challenging due to the fact that the mathematics of Fourier transforming a radial beam are not as simple as for a Gaussian profile. Consequently, the optimal pinhole size has been determined somewhat empirically. The recommended rule of thumb for selection of pinhole size from Pascal Anger (of Lukas Novotny's lab) was to use a pinhole that is $\sim 1.5\times$ the diffraction limited spot. Our current setup utilizes an ~ 8 mm beam going into the RPC, and is focused by an 125 mm lens. Thus, calculating the diffraction limited spot, assuming a Gaussian beam and multiplying by 1.5, we should use a pinhole with a diameter of $D = 1.5 \cdot 1.22 \cdot \lambda/NA = 1.5 \cdot 1.22 \cdot \lambda/(d/FL) = 15.5\mu m$. We in fact use a $15\mu m$ pinhole.

3.6.2 The Evanescent Field

It is well known that light rays passing from a higher index of refraction material to another material of lower index of refraction will undergo total internal reflection for all rays exceeding some critical angle θ_c given by $n_1 \sin\theta_c = n_2$. Despite undergoing total internal reflection, there exist exponentially decaying components of the field into the less optically dense medium. The penetration depth of such field δ as defined as the $1/e$ depth of the field is given by [56]:

$$\delta = \frac{\lambda}{2\pi\sqrt{n_1^2 \sin^2\alpha - n_2^2}}, \tag{3.2}$$

where α is the angle of incidence of the beam, and $n_2 < n_1$.

Another important aspect to consider is the polarization state of such exponentially decaying fields. The wave propagates along the interface while decaying into the less dense medium; however the polarization depends on the incident polarization. For s-polarized light the field is polarized parallel to the interface, whose intensity at the interface is given by:

$$I_y = I_0 \frac{4\cos^2\theta}{1 - n^2}, \tag{3.3}$$

where $n = n_2/n_1$ and I_0 is the intensity incident on the interface. Alternatively, for P-polarization the evanescent field has components both parallel and perpendicular to the interface that are 90° out of phase with each other whose initial intensities at the interface are given by [57, 9, 56]:

$$I_x = I_0 \frac{4\cos^2\theta\left(\sin^2\theta - n^2\right)}{(1 - n^2)\left[(1 + n^2)\sin^2\theta - n^2\right]} \tag{3.4}$$

$$I_z = I_0 \frac{4\cos^2\theta\sin^2\theta}{(1 - n^2)\left[(1 + n^2)\sin^2\theta - n^2\right]} \tag{3.5}$$

It is this important principle that supercritical P-polarized rays create evanescent fields with strong axial components that allows for use of a total internal reflection scheme appealing for near-field microscopy. Furthermore, the relative intensities can be enhanced significantly compared to the incident intensity as shown in Figure 3.7, making such a configuration even more desirable.

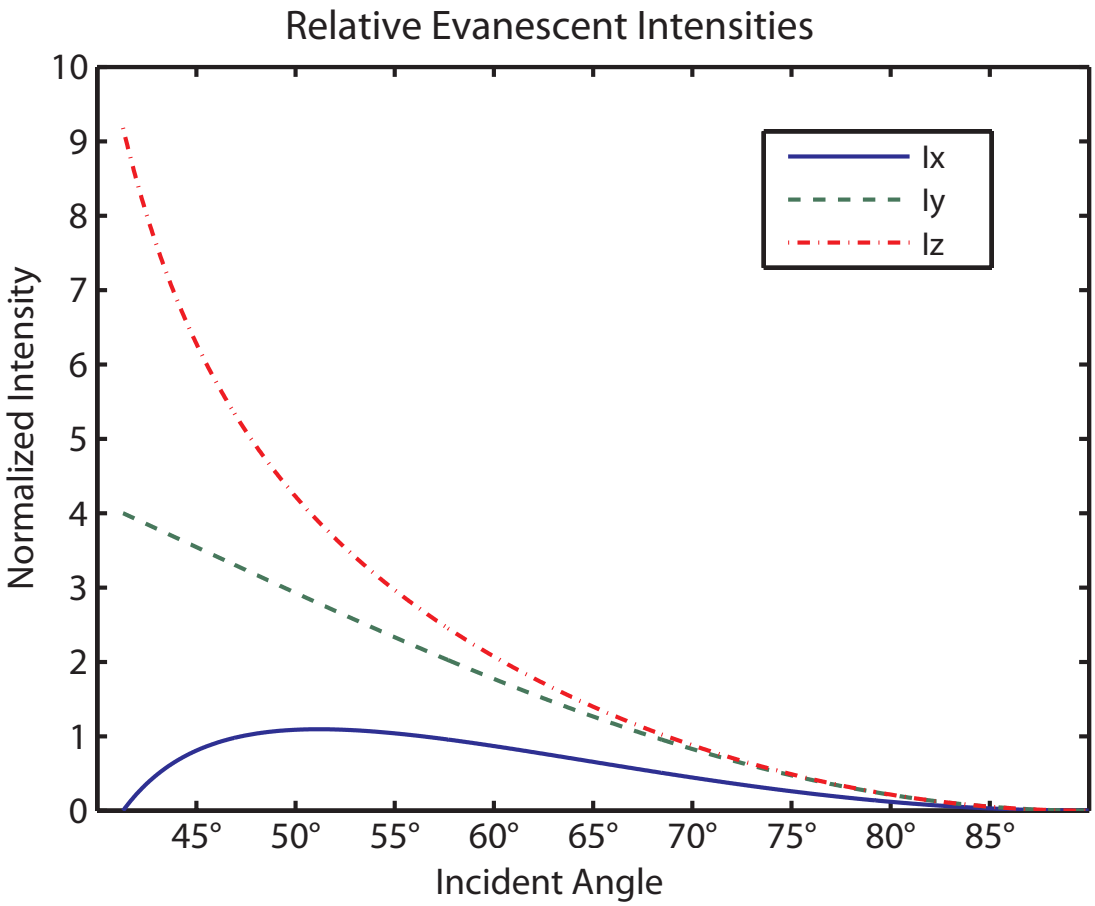

Figure 3.7. Relative evanescent intensities plotted as a function of incident angle. Eqns. (3.3), (3.4), and (3.5) are plotted as a function of incident angle.

3.6.3 Wedge Illumination

When illuminating with sufficiently large numerical apertures in a standard inverted microscope configuration as we do, some fraction of light rays incident on the back aperture of the objective will undergo total internal reflection. Since supercritical rays can lead to evanescent fields with strong axial components, we would like to block all subcritical rays that do not contribute to axial polarization.

The critical angle defining the separation between subcritical and supercritical rays is already described above. Now we wish to find how this critical angle (θ_c) corresponds to some critical diameter of the back aperture of the objective (D_c), such that rays arriving at normal incidence to the objective beyond the critical

radius ($R_c = D_c/2$) will be totally internally reflected at the glass-media interface.

The total diameter of the back aperture of a microscope objective can be calculated as follows:

$$D_{BA} = \frac{2 \cdot NA \cdot f_{TL}}{Mag}, \tag{3.6}$$

where f_{TL} is the focal length of the tube lens used by the manufacturer of the objective, and Mag is the magnification of the objective. Thus an $100\times$, $NA = 1.4$ objective designed for use with a microscope utilizing a 200 mm tube lens (e.g., Nikon) will have a back aperture of width, $D_{BA} = \frac{2 \cdot 1.4 \cdot 200mm}{100} = 5.6$ mm. Since $n \sin\theta$ is the definition of numerical aperture we can rewrite Snell's law as follows:

$$\begin{aligned} n \sin\theta_c &= n_m \\ NA_c &= n_m, \end{aligned} \tag{3.7}$$

where n is the index of refraction where the objective is working, and n_m is the index of refraction of the media the sample is in, and NA_c, is the "critical NA". Plugging this value for the "critical NA" into Equation 3.6 the critical diameter is obtained rather trivially:

$$D_{BA} = \frac{2 \cdot NA_c \cdot f_{TL}}{Mag} = \frac{2 \cdot n_m \cdot f_{TL}}{Mag} \tag{3.8}$$

Thus using the aforementioned objective with air above the coverslip, the critical diameter is readily found to be $D_c = \frac{2 \cdot 1 \cdot 200mm}{100} = 4$ mm. This result is also easily verified experimentally. Since only a small fraction of the subcritical rays become reflected at the glass air interface, the totally internally reflected rays appear at a much higher relative intensity. Placing a beamsplitter or dichroic mirror in the detection path allows for a screen to be placed to observe reflected/backscattered light from sample plane without obstructing the incident beam. Such an experiment reveals a bright halo or ring on the outermost edges of the beam, whose interior diameter corresponds to D_c.

Now, if a linearly polarized Gaussian beam incident on an objective lens is blocked by some disk with a diameter D_c such that only supercritical rays are allowed to pass, this configuration still does not yield axial polarization along the

optical axis at the focal plane for two reasons. First, as diametrically opposed rays are in phase with each other, the vector addition of any diametrically opposed beams yields a resultant vector with a polarization parallel to the sample plane (this is true for both propagating and evanescent fields). Second, portions of the beam originating from the sides of the objective orthogonal to the polarization direction (points **C** and **D** in Fig. 3.8) are S-polarized and thus have no vertical component whatsoever. Portions of the beam originating from sides of the objective along the polarization direction (points **A** and **B** in Fig. 3.8) are p-polarized, and can lead to axial components if the beam is shaped correctly. Beams originating from

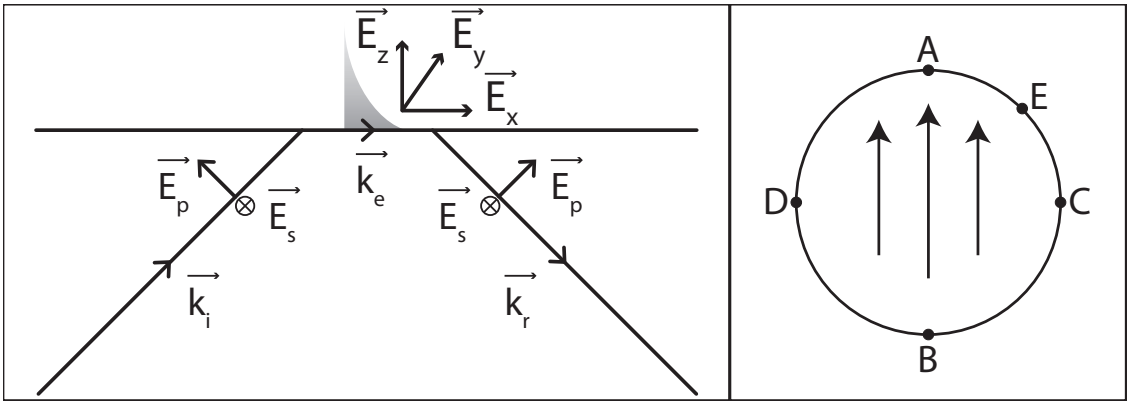

Figure 3.8. Diagram detailing the polarization state of the evanescent field. Incident and reflected values are denoted with subscripts i and r respectively, while the evanescent field quantities carry the subscript e. S-polarized light incident on an interface above the critical angle leads to an evanescent field with exponentially decaying field strength moving into the less dense media. S-polarized light yields an evanescent field polarized in the $\hat{Y}$ direction as indicated. P-polarized light incident under identical conditions again leads to an evanescent field at the interface, but is polarized primarily in the $\hat{Z}$ direction, but also has components in the $\hat{X}$ direction that are 90° out of phase. Note that the evanescent field propagates to the right as shown by $\vec{k}_e$. The shift between the incident and reflected rays is known as the Goos-Hänchen shift [9]. The right panel shows a cross section of a linearly polarized beam at the back aperture of a microscope objective. Rays originating from points A and B lead to P-polarized rays, while rays originating form C and D lead to S-polarized excitation. A ray from point E would have both S and P-polarization components. Any pair of diametrically opposed rays (A&B, C&D, or elsewhere) leads to a net longitudinal polarization at the sample surface.

intermediate positions, such as point **E**, would be comprised of both P-polarized and S-polarized components. By only permitting a wedge of light into the objective it is possible to select for essentially only super-critical rays leading to strong axially polarized evanescent field at the surface. Examining the raytracing diagram (Fig 3.9) shows the vector addition of how this is accomplished — only P-polarized rays are allowed and only from one side of the objective.

The actual implementation of the wedge illumination configuration is easily realized; a beam mask is placed in the excitation path after any polarization conditioning has been performed. Furthermore, the beam mask may be used in conjunction with either a linearly polarized beam or a radially (or even azimuthally) polarized beam. The advantage for using the beam mask with a linearly polarized beam is that the polarization of the evanescent field can easily be changed from axial to transverse through use of a half-wave plate. This has an advantage in being able to do control experiments to separate near-field effects (axial polarization) from far-field effects (transverse polarization). In principle, a similar control can be

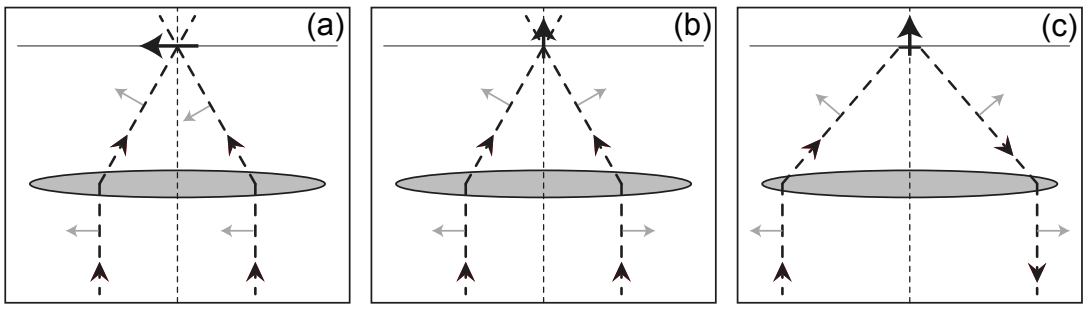

Figure 3.9. Raytracing schematic of several different epi-illumination configurations. Panel (a) represents linearly polarized light, yielding a purely horizontal resultant vector at the sample plane along the optic axis. Panel (b) demonstrates how radial polarization yields a purely vertical resultant vector at the focal plane along the optical axis. Panel (c) shows how allowing only a subsection of supercritical rays into the back-aperture of the objective for linearly polarized light can also lead to a large axial component of light at the sample plane as the beam is totally internally reflected. This is achieved by applying a wedge shaped mask to block all subcritical rays.

done using radial polarization (masked or not); however, this requires switching to azimuthal polarization which is not so easily realizable.

The design of a beam mask is drawn in Figure 3.10. In order to predominantly allow only P-polarized light, the opening angle is 60°. Notice in Figure 3.10 that only a relatively small fraction of light is passed by this mask; in fact, the fraction of light that should pass can be easily calculated (assuming a Gaussian profile). Exactly filling the back aperture at the $1/e^2$ beam diameter, the power transmitted (P_T) by the mask is as follows:

$$P_T = \frac{P_0}{6}\left(\frac{\int_{-R_{BA}}^{R_{BA}} exp\left(-\frac{2x^2}{R_{BA}^2}\right)dx - \int_{-R_c}^{R_c} exp\left(-\frac{2x^2}{R_{BA}^2}\right)dx}{\int_{-R_{BA}}^{R_{BA}} exp\left(-\frac{2x^2}{R_{BA}^2}\right)dx}\right), \qquad (3.9)$$

where P_0 is the initial power, R_{BA} is the radius of the back aperture of the microscope objective, and R_c is the critical radius (half the critical diameter D_c as previously defined). This integration accounts for the the total power entering the back aperture of the objective minus the power blocked by a disk of diameter D_c, both normalized by the total power. The factor of (1/6) accounts for the fact

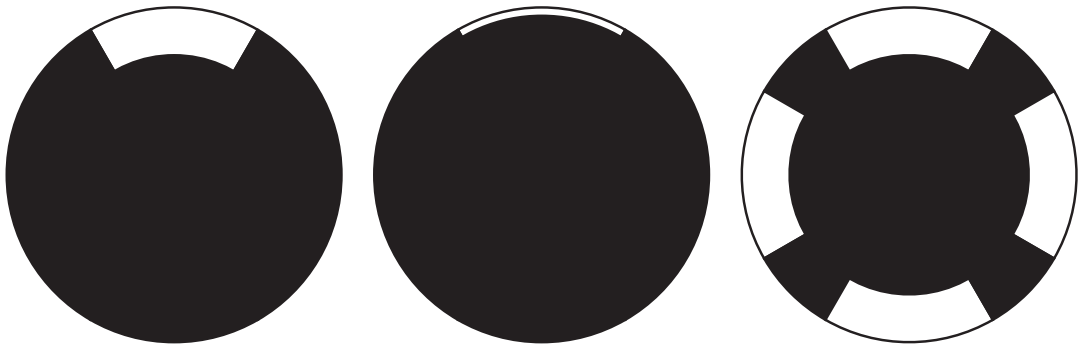

Figure 3.10. Beam masks for creating wedge illumination profiles. To scale cartoon diagrams of several different beam masks employed in the lab. The outer diameters of the mask representing the diameter of the back aperture of a 100× 1.4 NA Nikon objective. Black areas represent light that is blocked. Left mask is for use in air. Center mask is for use in water. Right mask is for use in air with radial polarization.

that only a 60° wedge of light is admitted. For a 100× 1.4 NA Nikon objective with a 5.6 mm back aperture this calculates to be about 1.9% power transmission. From Figure 3.10 it is apparent that the power transmission for water can be much lower still (0.2%). To circumvent this problem, a greatly expanded beam or one with a flat top profile can allow for higher transmission. One last consideration for mask design is that in light of Figure 3.7 it is extremely advantageous to permit transmission of all super-critical rays, especially those closest to the critical angle.

A final note of imaging with wedge illumination: the sample thickness is limited by the depth of the evanescent field δ. As TEFM is a surface technique, many samples of interest are either very thin to begin with or can be made so. Whole cells, for example, have a thickness that would not allow for use of wedge illumination. However, cells can be split open and laied flat on specially prepared coveslips, or alternatively, certain proteins of interest may be placed on lipid bilayers, Should live cell imaging, or imaging of cells closer to their native environment, wedge illumination may not be possible. Also it must be remembered that the index of refraction of the sample sitting on the interface may actually frustrate the total internal reflection of the beam, such that the beam becomes propagating, and consequently creating an illumination state with much different polarization conditions than perhaps originally intended.

The manufacture of the mask is done in a very low-tech fashion. The mask itself is made from aluminum foil. The wedge opening is made by printing a paper template from the computer, taping it to a sheet of foil, which is then cut with a razor blade. The central disk is made from a custom dye, where foil is sandwiched tightly between two plates and a metal pin is stamped through to produce a circle at the appropriate diameter with cleanly cut edges. This central circle is then glued onto the wedge opening. The entire mask is held in place with a retaining ring in an optic mount. More sophisticated fabrication methods may be used, but we have found this method to produce sufficiently acceptable results.

3.7 Beam Profiles

As TEFM is a surface technique, we are generally only concerned with the two-dimensional profile of a laser mode at the sample plane, and have little interest in the axial beam profile, only the polarization. The focal spot of a TEM_{00} laser mode has a two-dimensional Gaussian profile at the sample plane which is slightly elongated in the direction of polarization [58, 59]. The focal area of such a spot can be reduced significantly through the use of radial polarization, particularly radial polarization in combination with an annular mask [59, 60]. In fact, Dorn et al. have demonstrated a focal area of $0.16\lambda^2$ for radial polarization, compared to $0.26\lambda^2$ for Gaussian polarization, nearly a 40% improvement. We have also observed such improvements in our setup as seen in Figure 3.11.

When exciting with wedge polarization, elongated focal spot profiles are observed. This is easily explained by the fact that we are asymmetrically under-filling the back aperture of the objective such that the narrowest part of the beam becomes elongated the most, as would be expected. Also, subtle differences in the wedge profile occur between polarization states as vertical polarization tends to be slightly more elongated than horizontal polarization; however, this is difficult to quantify in that the exact focal spot profile depends sensitively on the exact placement of the wedge relative to both the incident beam and the optic axis.

3.8 Tip-Laser Alignment

To acquire near-field signals of any type, the paramount requirement is to align the tip within the focus of an excitation beam. Using radial polarization, the tip must be stringently aligned with the optic axis to obtain optimal signal enhancement. In particular, using an oil immersion objective lens with NA = 1.4, the diffraction-limited focal spot has a diameter of ~ 250 nm with green ($\lambda \sim 532$ nm) illumination, and the tip needs to be aligned into the central ~ 100 nm to achieve strong and repeatable field enhancement. Mechanical and thermal drifts on the order of several tens of nm per minute can make it difficult to maintain tip-laser alignment with radial polarization. When using metallic nanoparticle tips,

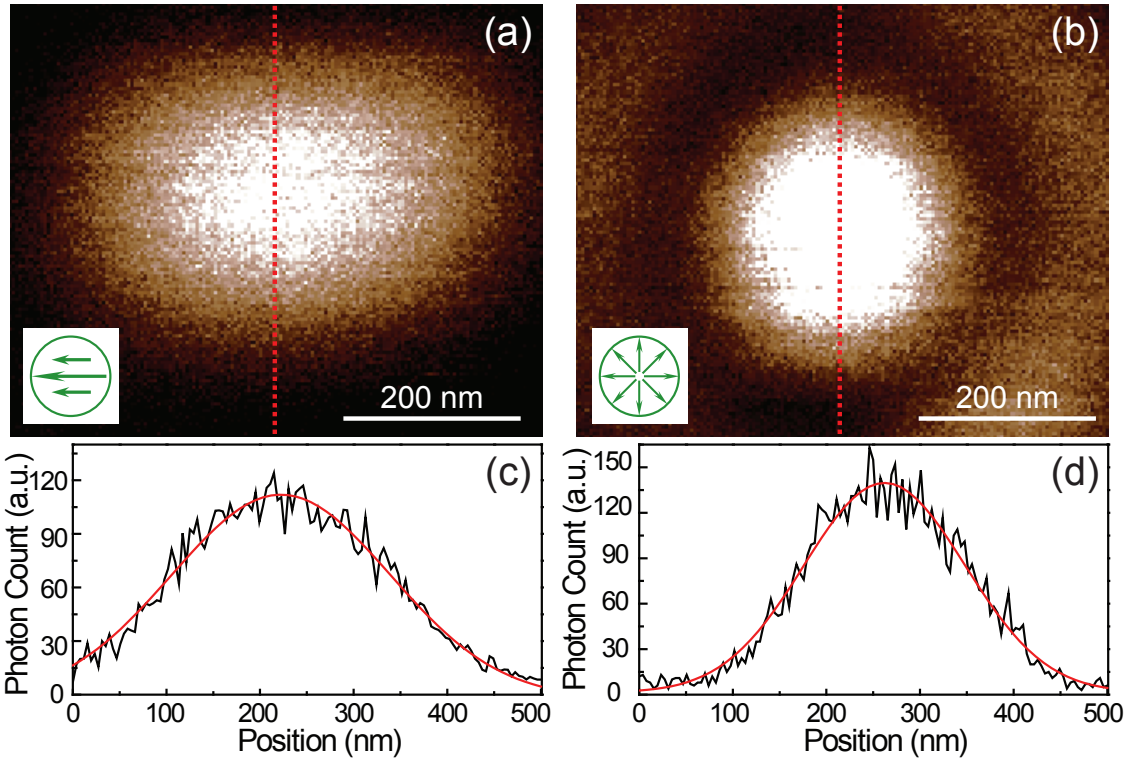

Figure 3.11. Beam profiles for Gaussian (a) and radial (b) excitation configurations. The cross-section of both profiles for each illumination condition were fit to a Gaussian function: (c)& (d). The FWHM of the Gaussian illumination profile is 234 ± 12 nm, while the FWHM of the radial profile is significantly narrower: 172 ± 5 nm. Images were taken by scanning single 20 nm dye doped latex beads through each laser profile and collecting the corresponding fluorescence signal on a single photon sensitive APD. Fluorescence excitation for both configurations was achieved by focusing a 543 nm He-Ne laser with a 100×, 1.4 NA objective. Reprinted with permission from reference [6], © 2008 IEEE.

it is possible to use the second harmonic generation (SHG) signal at the tip apex to optimize the alignment between the laser and probe. For dielectric tips (i.e., silicon) or pyramidal metal-coated tips, this is not possible, and other methods must be used, such as co-imaging onto a video camera the backscattered laser spot and a pattern cast by the tip when illuminated with a non-focused light source. The task of maintaining good alignment of the tip and laser can be a substantial technical challenge in apertureless NSOM, particularly in the episcopic configuration.

As indicated above, there are many technical challenges associated with posi-

tioning a sharp tip into a focal spot as small as $\sim$ 250 nm diameter. This task requires some indicator of alignment, as well as the precision to perform it. As our system includes a high-precision piezo-actuated scanning stage, we have the ability to scan the tip relative to a stationary laser focus. Additionally, our system includes a scanning mirror, allowing movement of the laser relative to the tip. In general, the most versatile system, particularly with regard to tip-laser alignment, would be one that has the ability to scan three elements: the sample, the tip, and the laser. Unfortunately we do not have any tip scanning abilities.

The ability to move the laser and tip relative to each other must be accompanied with the ability to measure the degree of their alignment. The top view optics in the AFM head, allowing for imaging the AFM probe and sample from above, can facilitate coarse alignment. This works best for AFM probes whose tip apexes are visible from directly above, such as ATEC tips by Nanosensors or Arrow tips by Nanoworld. In this case, top view optics can achieve alignment to within one micron with practice. Many AFM probes, however, are fabricated with the tip hidden below the cantilever, rather than protruding beyond it, limiting the alignment precision to several microns.

A more precise alignment method requires the presence of a measurable near-field optical signal embedded within the far-field background. An offset between the far-field and near-field signals then indicates the degree of misalignment, and either the tip or laser can be adjusted to compensate. This method works best for relatively simple samples where an offset between the near-field and far-field background signals is clearly discernible. In addition, it is susceptible to null results when no near-field signal is observed for a reason distinct from misalignment. This procedure also requires pre-alignment of the tip and laser at least to within the diameter of the focal area, which may or may not be possible using the top-view optics described above. A related alignment procedure requires co-registration of the optical and AFM images. This method can also work quite well, but only for samples with discernible topography.

For transparent samples, a high-precision method of alignment that is indepen-

dent of the sample topography, complexity, or the optical accessibility of the tip from above is to use the episcopic optics below the tip and sample to co-image the tip apex and laser focus backscattered from the sample. If the tip itself produces a strong scattering signal that can be distinguished from scattering off the sample surface, then good alignment can be achieved using the excitation laser alone. This is most easily accomplished if the scattering signal from the tip is either spectrally distinct from the sample scattering (e.g., probe fluorescence, second harmonic generation, white-light generation, etc.), or is modulated by tip oscillations. In any case, the small depth of focus that accompanies high NA objectives can make it difficult to obtain sufficient scattering signal from an ultra-sharp probe. In addition, light scattering from dielectric tips is generally much weaker than from metallic ones, so the excitation laser itself is unlikely to produce enough signal to facilitate alignment. To enhance sensitivity for the tip-scattered signal, it may be necessary to illuminate the tip with additional sources, possibly with spectral characteristics that are distinct from the excitation laser. Obviously, these spectral characteristics should neither interfere with AFM operation nor influence the sample response. The additional sources may be applied from the side, from below along the excitation path, or possibly even from above. This method of co-imaging a scattering signal from the tip and the laser focus backscattered from the sample, has the potential to yield tip-laser alignment with arbitrary precision, but is also the most challenging to implement.

In our system we rely on some combination of many of the techniques just described for tip-laser alignment. Typically the first step is to achieve coarse alignment using the top-view optics. Laser illumination from the side at $\lambda = 785$ nm is one option for finer alignment. This wavelength was chosen such that it is spectrally distinct from the AFM feedback laser (860 nm) and is far enough to the red to not interfere with the types of optical measurements typically performed in the lab. The side laser creates a very distinct diffraction pattern around the apex of the probe, as seen in Figure 3.12. The inverted microscope objective collects both the light scattered from the excitation laser as well as this diffraction pattern. A

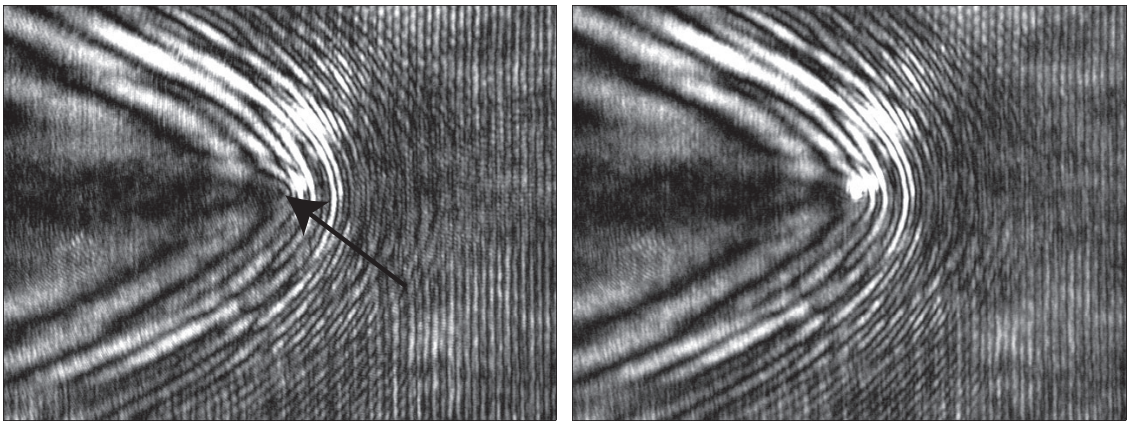

Figure 3.12. Image of a diffraction pattern around a tip under side illumination. Left panel shows the patten in the absence of an excitation laser, the arrow indicates the location of the tip apex. The right panel shows the same pattern with the excitation laser (vertical wedge illumination).

beam pick-off is inserted into the detection arm of the path such that $\sim 4\%$ of the light is directed onto a camera. The overlap of these two patterns can be used as an excellent indicator of alignment, and with practice the tip can be aligned to within several hundred nanometers of the optic axis using this indicator, and is often good enough to place the tip within a tightly focused laser beam.

Another indicator of alignment is to use defocused white light. When the AFM probe is commanded to go to the surface, it should be visible directly from the inverted microscope objective; however, since the probe is very sharp (< 10 nm) and since we want to keep it sharp (only applying small forces to the surface), it can be very difficult to image directly as the interfacial area is incredibly small. Instead of keeping the objective lens in focus on the sample plane, by focusing it above the sample plane we have a better ability to image the AFM probe. In addition to the backscattered laser excitation we also introduce an uncollimated white light source into the illumination path using a beam-splitting cube. In this way we obtain a much broader illumination spot such that after the coarse alignment step the probe is very likely to be found.

Ultimately the best indicator of alignment is the presence of a near-field signal.

If the near-field signal is strong enough in comparison to the far-field background signal present it may be observed directly. At times the near-field signal can be quite weak relative to the background, especially when first aligning the system. As a last check an approach curve can be performed, the probe is lowered directly onto the sample while monitoring the fluorescent signals; any change in the far-field signal as the tip is brought in close proximity to the sample is indicative of some tip-laser overlap.

The actual positioning of the tip relative to the laser (or vice-versa) can be accomplished in our system by moving either the lower layer of the stage or by adjustment of the scanning mirror. Moving the stage is done by the picomotors, which cause enough vibrations in the system that they can dull the tip if it is in contact with the sample. Moving the scanning mirror is totally non-perturbative to the AFM such that adjustments can be made while in contact with the surface. The scanning mirror employed in our system is a Fast-Scanning Mirror (FSM) made by Optics in Motion, which can be controlled remotely via Labview controls. This mirror has an angular resolution of < 2 μrads, which corresponds to a displacement of about 4 nm at the sample plane in our system. We have calibrated the mirror such that a when a given displacement is desired, the value can be entered into Labview where the appropriate $nm \rightarrow \theta \rightarrow V$ conversion is applied ($1\ \mu m = 0.10\ V\hat{x}, 1\ \mu m = 0.15\ V\hat{y}$).

3.9 System Specifications

3.9.1 Radial Illumination Path

The radial illumination path is the path along the top and far right side of the system (blue) shown in Figure 3.13. All lenses in the system are achromatic doublets. A beam exits a polarization maintaining fiber (OZ Optics) through a custom made achromatic out-coupler, using a 35 mm lens. The beam diameter after the out-coupler is $\sim$ 15 mm. The beam is reduced by a pair of lenses (125 mm and 75mm) and then passes through a linear polarizer (Meadowlark optics). The beam is reflected off a broadband dielectric mirror (10D20BD.1; Newport) before

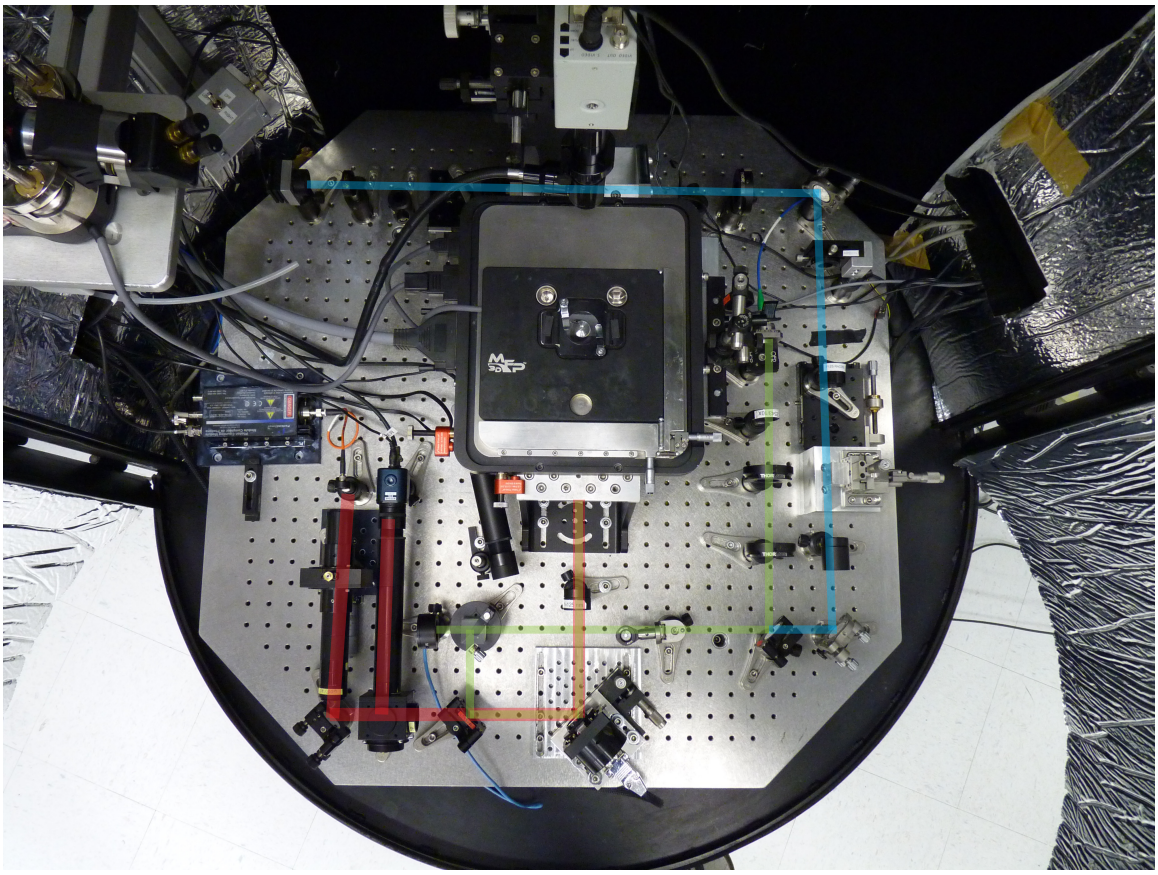

Figure 3.13. Illumination and detection paths.

passing through the radial polarization converter (Arcoptix). The beam is sized to slightly under-fill the 10 mm back aperture of the RPC. The beam is then spatial filtered by focusing with a 125 mm lens through a $15\mu m$ pinhole. The lens is on a linear translation stage, providing the Z motion of the spatial filter. The pinhole itself can be adjusted in the X-Y plane through another set of micrometers. This configuration allows for adjustment of the Z direction without need for adjustment of a collimating lens as the pinhole remains stationary along the Z axis. The beam is recollimated by a 75 mm lens before being reflected by another mirror enter into the main part of the illumination/detection path. Note that the radial polarization state is sensitive to many types of optics; for example, we have found that dielectric broadband mirrors from Newport (10D20BD.1) maintain the proper

spatial mode better than a similar model made by Thorlabs (BB1-EO2). Also, the radial polarization state can be adversely effected by some dichroic mirrors.

3.9.2 Linear/Gaussian Illumination Path

The linear polarization illumination path is to the left of the radial path (green) shown in Figure 3.13. A beam exits a polarization maintaining fiber (OZ Optics) through a commercial out-coupler (PAF; Optics for Research) that has a non-achromatic lens. Excitation filters can be placed immediately following the out-coupler, but are sometimes introduced prior to fiber coupling. The beam passes through a half-wave plate followed by a linear polarizer (Meadowlark Optics). The beam is reflected by a mirror before joining the main part of the illumination/detection path. This mirror is mounted on a removable face plate such that when in place the linear polarization path can be used, but must be removed for use with radial illumination. Alternatively, both paths may be used simultaneously if using different colors and a dichroic mirror. A beam mask is introduced into the main section of the path. It is mounted such that it can be easily removed or added by use of a flipper mount. In Figure 3.13 the mask is shown flipped down (out of the path). The remainder of the path is identical for any type of illumination.

3.9.3 Illumination Path

The main section of the illumination path is well diagrammed in Figure 2.1. The excitation beam first is incident on a non-polarizing beam splitting cube where 50% of the power goes to a power meter that is read out remotely. This beam splitter also allows for optional introduction of white light into the system, or for a backscatter detector/camera. Following the beam splitter, the beam reflects off a dichroic mirror, which is also mounted on a removable face plate for easy switching between colors. The beam next reflects off a fast scanning mirror (FSM-100; Optics in Motion), whose operation is discussed above. The beam passes through two 125 mm lenses in a 1:1 configuration, the positioning of which are important to ensure the detected light is properly de-scanned. The distance between the scanning mirror and lens is 125 mm, between the two lenses is 250 mm and between the last lens

to the back aperture of the objective is another 125 mm. A mirror is between the second lens and the microscope objective, which reflects the beam upwards. The beam is focused by a Nikon objective (usually a 100× oil objective 1.4 NA) mounted on a coarse alignment stage as well as a piezo-actuated stage for fine focusing (Piezosystem Jena).

3.9.4 Detection Path

Light is collected through the same path as just described. Since the collected light is de-scanned by the same scanning mirror, the detection path will be co-linear with the illumination path regardless of scan angle. Collected light usually includes fluorescent light from the sample, but also light from the AFM feedback laser (860 nm), the side-illumination laser (785 nm), as well as back-reflected excitation beams. The dichroic mirror reflects all wavelengths shorter than the cut-off wavelength such that the vast majority ($\sim$ 99%) of the excitation light is reflected while wavelengths longer than the cut-off frequencty are transmitted. The reflected excitation light then passes through the beam splitting cube again and can be imaged on a camera or other detector. In the detection path, after light passes through the dichroic mirror, it is filtered by an 810 short-pass filter to block the AFM laser. The remaining signal is picked off by an optical flat with an anti-reflection coating on one side, such that 4% is focused onto a camera (Watec 902H3) by a 200 mm lens (the focal length of a Nikon tube lens), thus keeping the total magnification on the camera equal to that specified on a Nikon objective. The light then reflects off a final mirror and then through several spectral filters, including the appropriate emission filters as well as 710 short-pass filters to eliminate the side laser. The final fluorescent signal is focused by a 100 mm lens into a multi-mode fiber which terminates at the face of a single photon sensitive avalanche photodiode (APD) made by Perkin Elmer (SPCM-AQR-14-FC-Si).

CHAPTER 4

IMAGE CONTRAST

While resolution is the most discussed parameter in any imaging technique, one of the less mentioned but extremely important issues in microscopy is image contrast. It has been established that TEFM is able to obtain great spatial resolution, and that it relies simply on the sharpness of the probe rather than the excitation wavelength. Less clear however, is the amount of image contrast that can be obtained with this technique. All apertureless near-field techniques struggle from a background problem, that is the near-field signal is only part of the signal collected—it rides on top of a far-field background signal. The background problem gets compounded as the density of fluorescent samples increases; a necessary requirement for being able to image most real biological samples of interest. This background problem is illustrated in Figure 4.1. The near-field signal has a resolution far superior to the far-field background signal, however, as the number

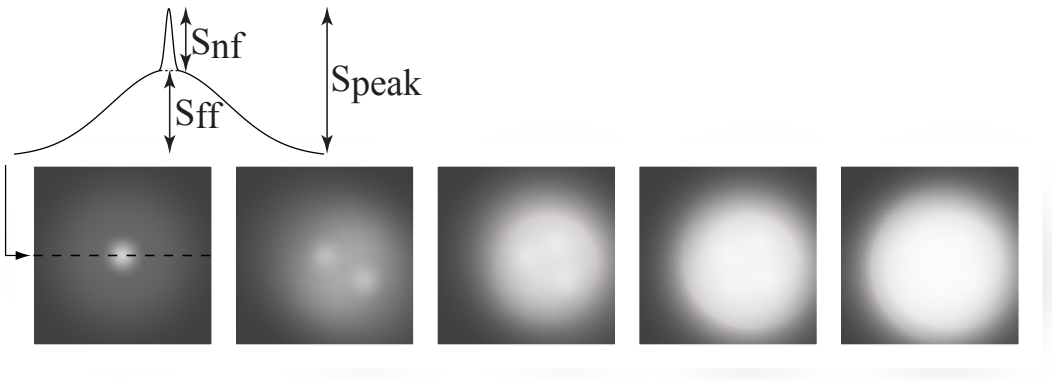

Figure 4.1. Simulated superposition of near-field and far-field signal components in apertureless NSOM. The image contrast, defined as the ratio of near-field signal to far-field signal S_{nf}/S_{ff}, decreases from left to right as the number of molecules within the far-field illumination spot increases from one to five.

of fluorophores in the focal area is increased, the near-field signal gets washed out in the overlapping far-field signals of neighboring fluorophores.

While reading *Principles of Nano-Optics*, [56] I encountered a statement that insinuated that TEFM might never provide adequate image contrast for samples of densely packed fluorophores; the argument was as follows, to paraphrase: Considering an n^{th} order optical process, and imaging a uniformly photoactive surface, the far-field background signal is proportional to $S_{ff} \sim AI_0^n$, where A is the focal area of the spot, and I_0 is the excitation intensity. On the other hand, the near-field signal S_{nf} originates from the much smaller tip-enhanced interaction area a, such that $S_{nf} \sim a(f_i I_0)^n$, where f_i is the intensity enhancement factor. Imposing that the contrast be greater than one ($S_{nf}/S_{ff} > 1$) puts a constraint on the enhancement factor: $f_i > \sqrt[n]{A/a}$. Using realistic areas; $A = (500 \text{ nm})^2$ and $a = (10 \text{ nm})^2$, we arrive at the conclusion that $f_i > \sqrt[n]{2500}$, which means for first order processes like fluorescence unrealistically large enhancement factors for dielectric tips would be required. This of course implied that, TEFM of densely packed fluorophores with dielectric tips could never achieve the image contrast necessary to be a viable imaging tool.

This worried me a great deal as this implied I was embarking on an undoubtedly fruitless endeavor to try to do the impossible: to image high density ensembles of proteins at high spatial resolution and contrast using dielectric tips. It must be remembered that this line of reasoning is only valid for the static case, where a tip is held at the position of maximum enhancement above a sample; however, it negates any improvements made possible by oscillating the tip in conjunction with demodulation techniques. At that point I decided to do a feasibility study to determine what might be the new requirements on the enhancement factor should demodulation techniques be employed. The results of this feasibility study were published in Optics Express, Volume 9 pages 6183-6193 (2008)[8].

4.1 Optics Express Paper

This chapter includes a reproduction of the Optics Express paper, but has been reformatted to fit this dissertation. Several typographical errors have been corrected here. Some clarifications and expanded derivations are made after the entire presentation of the original paper.

4.1.1 Introduction

Tip-enhanced fluorescence microscopy (TEFM) is a type of apertureless near-field scanning optical microscopy (ANSOM) that utilizes fluorescence to generate an image. By aligning the sharp tip of an atomic force microscope (AFM) probe into the focus of a laser beam with axial polarization, enhanced fields are generated at the apex of the tip [16], as shown in Fig. 4.2. This field enhancement is tightly confined to the vicinity of the tip apex and has been shown to decay rapidly as r^{-6} with distance r from the tip apex [3]. These enhanced local fields can be used to beat Abbe's diffraction limit, and various scattering processes (e.g., one- and two-photon fluorescence, Raman scattering, infrared spectroscopy, and Rayleigh scattering) have been used to image a range of samples with nanoscale resolution [3, 32, 2, 7, 61, 62, 46, 63, 64, 1, 65]. Much of the work with ANSOM to date has been on samples composed of isolated particles/molecules (e.g., fluorophores, quantum dots, nanotubes) due to the fact that ANSOM suffers from a relatively large background signal that arises from direct (nonenhanced) scattering from the laser beam. Thus, high density samples are challenging for ANSOM analysis since the background signal increases with the number of particles in the laser spot, while the tip-enhanced signal does not. This has so far prohibited the application of ANSOM to biological samples composed of a high density, heterogeneous ensemble of fluorescently-tagged biomolecules, including proteins, lipids, and nucleic acids.

Recently, a number groups have investigated various means of increasing the degree of field enhancement, including optimizing the shape of the tip to leverage plasmon and antenna resonances. These efforts have already been fruitful for increasing the enhancement, and will impact both ANSOM and sensor applications [18]. To complement these studies, it is also important to understand how much

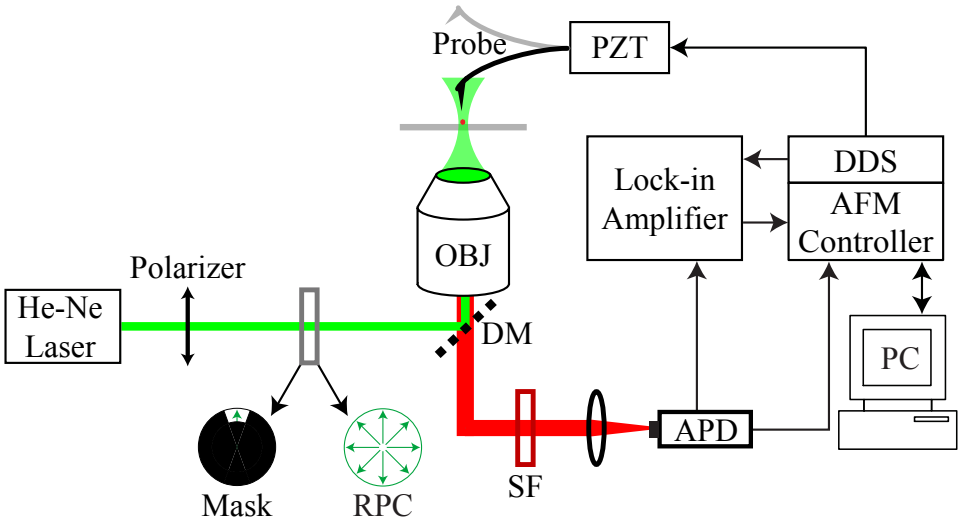

Figure 4.2. Experimental setup for TEFM. Labeled elements are as follows: He-Ne Laser — helium-neon laser ($\lambda = 543$ nm); Mask — laser-beam mask; RPC — radial polarization converter; DM — dichroic mirror; OBJ — microscope objective; Probe — AFM probe; PZT — piezoelectric transducer; SF — spectral filters; APD — avalanche photodiode; LA — lock-in amplifier; DDS — digital synthesizer; PC — personal computer. The arrows indicate the polarization of the laser beam. Axial polarization at the sample plane can be achieved either by simply focusing a radially polarized laser beam, or by placing a laser beam mask before the microscope objective such that only super-critical rays are allowed to propagate. This focused total internal reflection fluorescence (TIRF) setup is sometimes used because of its broadband capabilities and its large focal spot ($\sim$1.5 μm $\times$ 0.5 μm) lends itself to easy tip alignment, while radial polarization is preferred for smaller focal spots, $\sim(250\text{ nm})^2$.

enhancement is required to image high-density samples with sufficient contrast to resolve individual molecules within the ensemble. It has been pointed out that for dense samples, the minimum (intensity) enhancement needed to achieve sufficient image contrast ultimately depends on the n^{th} root of the ratio of the area of the illuminated spot to the area under the tip, where n is the order of the scattering process being employed [18]. Naturally for linear scattering processes such as one-photon fluorescence, larger enhancement factors are needed compared to higher-order processes, such as two-photon fluorescence or Raman spectroscopy. In this paper we specifically investigate the limits of TEFM with regard to its potential for imaging high-density samples. In particular, we use a theoretical

model based on experimental measurements to show that sufficient contrast can be obtained even for the relatively simple case of commercially available silicon tips and one-photon fluorescence.

4.1.2 Contrast in TEFM

In TEFM, the laser stimulates two distinct fluorescence signals: the far-field signal, S_{ff}, resulting from direct illumination of fluorophores within the laser focus, and the near-field signal, S_{nf}, resulting from field enhancement at the tip apex. The resolution of S_{ff} is at best diffraction limited, while S_{nf} has resolution given primarily by the sharpness of the tip [2]. Figure 4.3 shows a cartoon image composed of the superposition of S_{ff} and S_{nf} as well as a simulated profile through its center. While not shown, we also assume some noise in the far-field signal. Within this context, contrast (C) and signal-to-noise ratio (SNR) are defined as:

$$C = \frac{S_{peak} - S_{ff}}{S_{ff}} = \frac{S_{nf}}{S_{ff}} \tag{4.1}$$

$$SNR = \frac{S_{peak} - S_{ff}}{\text{Noise in } S_{ff}} = \frac{S_{nf}}{\sigma_{ff}} \tag{4.2}$$

where σ_{ff} is the standard deviation (noise) in the far-field background. The near-field signal originates from a small area on the sample surface (a_{tip}) given by the near-field interaction zone, which is determined mostly by the tip sharpness, while the far-field background originates from a much larger area (A) given by the size of the laser focus. The total fluorescence signal for a given pixel of the raster-scanned

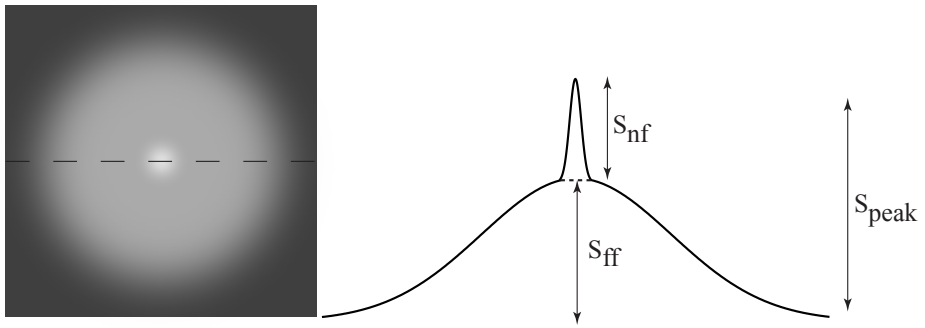

Figure 4.3. Cartoon of a fluorescent particle imaged by TEFM and the corresponding signal profile.

image, S_{peak}, is simply the sum of all photons collected during the pixel acquisition time (τ). The far-field signal S_{ff} is proportional to the number of fluorophores in the focal area of the excitation beam, N_{FA}, and also to a dimensionless parameter k that characterizes the total efficiency of the system: $S_{ff} = k\, N_{FA}$.

The probability of an illuminated fluorophore emitting a photon follows a Poisson distribution, such that the expected average number of counts in the time interval τ is simply S_{ff}. The standard deviation is given by $\sigma = \sqrt{S_{ff}} = \sqrt{k\, N_{FA}}$. In the limit of a single fluorophore in the near-field zone, $S_{nf} = f\, k$, where f characterizes the fluorescence signal enhancement induced by the tip, and is a function of several parameters related primarily to its geometry and material properties. In this limit, the peak signal is given by $S_{peak} = (f + N_{FA})k$. The overall system efficiency k is given by

$$k = I_0 \times \sigma_0 \times \tau \times Q \times CE \times \frac{\lambda}{hc} \tag{4.3}$$

where $I_0 = P_0/A$ is the intensity of the laser beam with power P_0 in a focal spot of area A; σ_0 is the absorption cross-section of the fluorophore; τ is the pixel acquisition time; Q is the quantum yield of the fluorophore; CE is the collection efficiency of the detection system; and hc/λ is the energy of a photon with wavelength λ. A green He-Ne laser ($\lambda = 543$ nm) was used for these experiments due to its low cost and the availability of fluorescent dyes and quantum dots with strong absorption at this wavelength. Although we have not done careful studies of tip-enhancement as a function of excitation wavelength, we do not expect a strong dependence since the dielectric function of silicon is fairly flat over visible wavelengths.

The lower limit for detection of a near-field signal arises from the requirement that the signal-to-noise ratio (SNR) be larger than unity,

$$SNR = \frac{S_{nf}}{\sigma_{ff}} = f\sqrt{\frac{k}{N_{FA}}} > 1. \tag{4.4}$$

Below this limit, the near-field signal is indistinguishable from stochastic fluctuations of the far-field background. On the other hand, to produce an image that

can be interpreted visually dictates a more stringent requirement, namely that the contrast (C) be larger than unity,

$$C = \frac{S_{nf}}{S_{ff}} = \frac{f}{N_{FA}} > 1. \tag{4.5}$$

In this model, it is straightforward to evaluate the minimum enhancement required for sensitivity to a single fluorophore within a dense ensemble. A practical limit on density arises from the requirement that the average spacing between fluorophores be no smaller than the microscope resolution, which is given by the near-field interaction zone. In this limit, $N_{FA} = A/a_{tip}$, where A is the area of the laser focus. Using the focused-TIRF scheme described above, $A = 0.75$ μm^2 and $a_{tip} = 100$ nm^2 , which suggests a signal enhancement of $f > 7500$ is needed to achieve contrast greater than unity. Employing a radially polarized laser beam yields a smaller focus spot, $A = (250$ nm$)^2$ [60], thus reducing the required enhancement to $f \sim 600$. Silicon tips are only capable of producing an enhancement factor of $f \sim 20$ [2], well below these requirements. Simple, nonoptimized metal tips have been predicted to yield enhancement factors of $f \sim 3000$ [16] and optimized metal tips that leverage antenna resonances may yield even larger enhancement factors. Although metal tips can produce much larger field enhancements than silicon, they also strongly quench fluorescence, leading to an overall reduction in the fluorescence signal and an associated decrease in the contrast. In several previous reports, silicon tips were found to yield the largest net contrast since no quenching was observed [2, 3, 7].

At first glance, the required signal enhancements predicted above cast a shadow on the potential application of TEFM to biological systems. As discussed below, however, the contrast can be improved dramatically by oscillating the AFM probe, which induces an associated modulation in the fluorescence signal, and by the subsequent application of a phase sensitive demodulation algorithm, such as lock-in amplification. Modulation/demodulation schemes are used widely in many areas of small signal processing and have also been used before in near-field microscopy [2, 3, 7, 66, 67, 68, 69]. The analysis below demonstrates the limits of this approach for TEFM.

4.1.3 Improving Contrast via Phase Sensitive Demodulation

To calculate contrast and signal-to-noise ratio for the case of an oscillating tip, Eqs. (4.4) and (4.5) must be modified to account for the fact that the tip only intermittently contacts the sample at a particular phase of its oscillation cycle. To discuss the dependence of the near-field signal on the instantaneous height of the oscillating probe, it is useful to consider the arrival of each photon in a phase-space picture. In this scenario, each photon is assigned an angle θ_i corresponding to the instantaneous phase of the sinusoidal tip-oscillation function at the time of detection (Fig. 4.4). The photon phases can be mapped to the corresponding tip-sample separation if desired.

Since the sample remains under direct laser illumination whether the tip is oscillating or not, the far-field signal for an oscillating tip is unchanged,

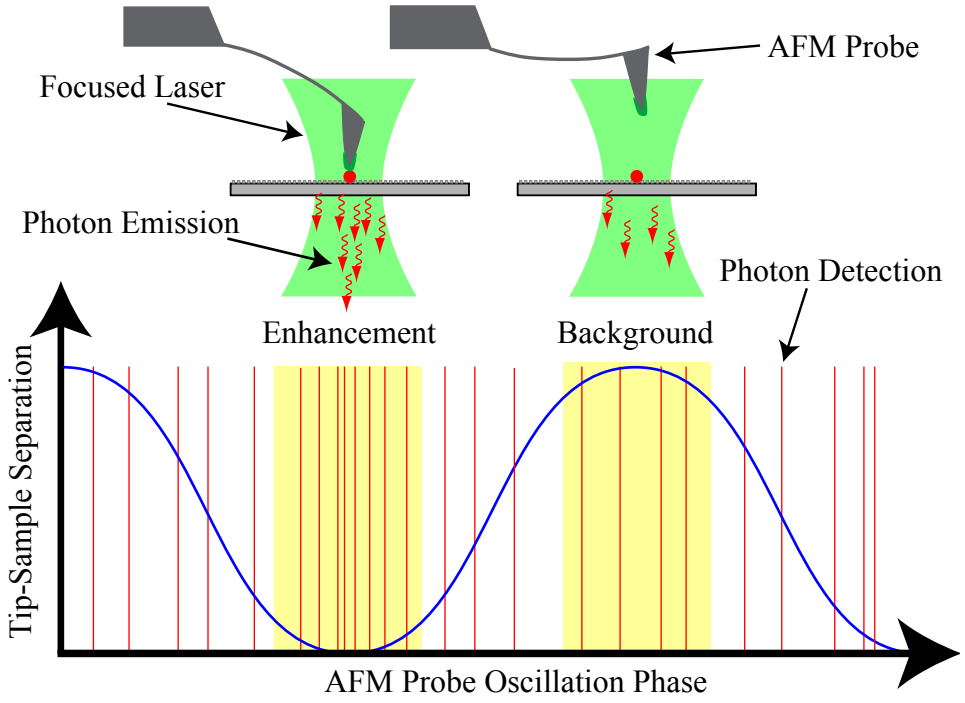

Figure 4.4. Phase-space plot showing how photon arrivals (vertical lines) are correlated to tip-oscillation phase. Squiggly arrows represent photons emitted from fluorophores within the laser focus. Higher photon count rates occur at a preferred phase θ_p corresponding to tip-sample contact, resulting in the strongest near-field signal.

$$S_{ff}^{osc} = S_{ff} = k\, N_{FA}. \tag{4.6}$$

Multiple scattering of far-field photons between the tip and sample can lead to variations in the background intensity as a function of the tip height. However, these variations have been measured to be very small ($< 5\%$) for the tip-oscillation amplitudes employed here, and are thus neglected. Therefore, we assume that the far-field signal for an oscillating tip is unchanged compared to an absent tip or one which is in constant contact with the surface.

In phase-space, the maximum near-field signal occurs at a preferred phase θ_p corresponding to tip-sample contact, and the photons are approximately Gaussian distributed around θ_p. To find the total number of near-field photons for a given pixel, S_{nf}^{osc}, the ratio γ defined as the number of photons collected in one oscillation cycle relative to the number that would have been collected had the tip been at the surface the entire time is calculated:

$$\gamma = \frac{1}{2\pi}\int_{-\pi}^{\pi} exp\left(-\frac{(\theta_i - \theta_p)^2}{2\theta_\sigma^2}\right) d\theta_i \approx \frac{\theta_\sigma}{\sqrt{2\pi}} \tag{4.7}$$

where θ_σ is the standard deviation of the photon-phase distribution, which can be obtained experimentally and is a function of oscillation amplitude. The approximation in Eq. (4.7) holds in the limit that the integration limits are extended to $\pm\infty$, or equivalently when $\theta_\sigma < \pi/3$. The near-field signal for an oscillating tip is then given by

$$S_{nf}^{osc} = S_{nf}\,\gamma = k\, f\,\gamma = k\, f\,\frac{\theta_\sigma}{\sqrt{2\pi}}. \tag{4.8}$$

Using the definitions for the oscillating signals in Eqs. (4.6) and (4.8), both the contrast and *SNR* for images produced by an oscillating tip (tapping mode TEFM) can now be calculated

$$C_{osc}^{sum} = \frac{f\,\gamma}{N_{FA}} \tag{4.9}$$

$$SNR_{osc}^{sum} = f\,\gamma\,\sqrt{\frac{k}{N_{FA}}} \tag{4.10}$$

where the subscript "sum" indicates a direct sum of the photon signals. Not surprisingly, without demodulation the contrast and *SNR* have been reduced by

a factor of γ compared to the non-oscillating scenario since the total number of near-field photons has decreased.

Lock-in amplification is a particularly powerful phase-sensitive demodulation technique that decomposes a modulated signal into real and imaginary components that are proportional to the cosine and sine projections in phase space, respectively. In TEFM, each detected fluorescence photon can be viewed as a unit vector pointing in the direction θ_i equal to the instantaneous phase of the tip oscillation at the time of detection (Fig. 4.5). In this picture, a lock-in amplifier simply performs a vector addition of the detected photons transmitted through its internal bandpass filter. If the resultant lock-in vector $\mathbf{L}$ is divided into near-field ($\mathbf{NF}$) and far-field ($\mathbf{FF}$) components, both of which are vector sums, then the lock-in signal is simply the magnitude $|\mathbf{L}| = |\mathbf{NF} + \mathbf{FF}|$.

The far-field component of the lock-in vector $\mathbf{FF}$ results from an unbiased two-dimensional random walk with unit steps, and follows the probability distribution

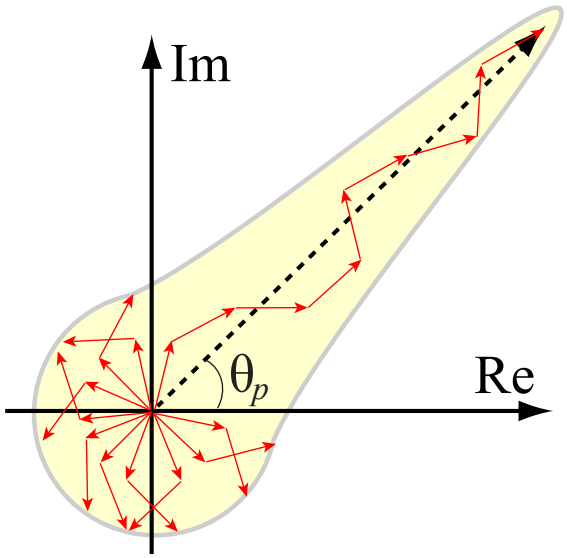

Figure 4.5. Expected phase dependency of lock-in signal. Each detected photon is considered as a unit vector with a direction corresponding to the instantaneous oscillation phase of the tip. A lock-in amplifier performs the vector addition of all such unit vectors. The near-field photon phases are Gaussian distributed around θ_p, which corresponds to tip-sample contact. Far-field background photons are detected randomly at all phases so the corresponding vector addition is simply a random walk.

originally derived by Lord Rayleigh

$$P(r) = \frac{2r}{N_{steps}} e^{-r^2/N_{steps}} \tag{4.11}$$

where r is the final end-to-end distance of the walk, and N_{steps} is the number of steps in the walk [70]. This distribution has a mean μ_r and standard deviation σ_r given by

$$\mu_r = \sqrt{\frac{\pi\ N_{steps}}{4}} \tag{4.12}$$

$$\sigma_r = \frac{1}{2}\sqrt{N_{steps}(4-\pi)}. \tag{4.13}$$

In our case, N_{steps} is given by the number of detected far-field photons that are transmitted by the lock-in bandpass filter, $N_{steps} = \beta \times S_{ff}$, where $\beta < 1$. This gives

$$|\mathbf{FF}| = \sqrt{\frac{\pi}{4}\ k\ \beta\ N_{FA}} \tag{4.14}$$

$$\sigma_{|\mathbf{FF}|} = \frac{1}{2}\sqrt{\beta\ k\ N_{FA}(4-\pi)} \tag{4.15}$$

for the average length of the far-field component $|\mathbf{FF}|$ and its uncertainty $\sigma_{|\mathbf{FF}|}$, respectively. The near-field component **NF** comes from a biased random walk about θ_p. The average value of its magnitude $|\mathbf{NF}|$ can be estimated by projecting the unit vectors corresponding to each near-field photon onto the θ_p axis and then summing the result:

$$|\mathbf{NF}| = \sum_i cos(\theta_i - \theta_p) = S_{nf}^{osc} \times \langle cos(\theta_i - \theta_p)\rangle \tag{4.16}$$

where the sum runs over all the near-field photons, $i = 1 \rightarrow S_{NF}^{osc}$. For simplification we define $\alpha = \langle cos(\theta_i - \theta_p)\rangle$. Since the phase of each photon θ_i is Gaussian distributed, the normalized expectation value is

$$\alpha = \frac{\int_{-\pi}^{\pi} cos(\theta_i - \theta_p) e^{-(\theta_i-\theta_p)^2/2\theta_\sigma^2} d\theta_i}{\int_{-\pi}^{\pi} e^{-(\theta_i-\theta_p)^2/2\theta_\sigma^2} d\theta_i} \approx e^{-\theta_\sigma^2/2}. \tag{4.17}$$

Combining this result with the definition of γ from Eq. (4.7), the average magnitude of the near-field component $|\mathbf{NF}|$ is then approximated by

$$|\mathbf{NF}| = k\, f\, \gamma\, \alpha \approx k\, f\, \frac{\theta_\sigma}{\sqrt{2\pi}}\, e^{-\theta_\sigma^2/2}. \tag{4.18}$$

When using a lock-in amplifier to demodulate the signal, an image is constructed one pixel at a time, where the value of each pixel is the magnitude of the lock-in vector, $|\mathbf{L}| = |\mathbf{NF} + \mathbf{FF}|$. The near-field component $\mathbf{NF}$ points along θ_p, but the far-field component $\mathbf{FF}$ points in a random direction. Performing the vector addition of $\mathbf{NF} + \mathbf{FF}$ and averaging over all directions for $\mathbf{FF}$, the peak lock-in signal is given by

$$|\mathbf{L}|_{peak} = \sqrt{|\mathbf{NF}|^2 + |\mathbf{FF}|^2} = \sqrt{(k\, f\, \gamma\, \alpha)^2 + \frac{\pi}{4}\, k\, N_{FA}\, \beta}. \tag{4.19}$$

The contrast C_{LI} and signal-to-noise ratio SNR_{LI} in the lock-in signal can now be found.

$$C_{LI} = \frac{|\mathbf{L}|_{peak} - |\mathbf{FF}|}{|\mathbf{FF}|} = \left[\frac{4\, k\, (f\, \gamma\, \alpha)^2}{\pi\, N_{FA}\, \beta} + 1\right]^{1/2} - 1 \tag{4.20}$$

$$SNR_{LI} = \frac{|\mathbf{L}|_{peak} - |\mathbf{FF}|}{\sigma_{|\mathbf{FF}|}} \approx 2C_{LI} \tag{4.21}$$

Equation (4.20) can be used to calculate the minimum signal enhancement factor required to achieve contrast greater than unity:

$$C_{LI} > 1 \;\Rightarrow\; f > \frac{1}{\alpha\, \gamma}\sqrt{\frac{3\pi\, N_{FA}\, \beta}{4k}}. \tag{4.22}$$

As before, we consider the case where there is only one fluorophore in the near-field zone ($\sim$10,000 fluorophores/μm^2) and the far-field illumination area is $\sim$(0.5 μm $\times$ 1.5 μm) corresponding to focused-TIRF illumination. Using typical experimental values for $k = 10$ and $\beta = 0.15$ as well as optimized values for $\gamma = 0.4$ and $\alpha = 0.6$ (see below) gives a required signal-enhancement factor of $f > 65$ to achieve a contrast greater than unity. Using radial polarization reduces the required enhancement to $f > 18$ which is very realistic for silicon tips and in fact has already been demonstrated in the case of isolated spherical quantum dots [2].

Figure 4.6 demonstrates how the lock-in demodulation scheme can be used to improve the contrast and *SNR* for samples with a high density of rod-shaped quantum dots (4 nm × 9 nm). These images were obtained using a silicon tip oscillating with an optimized amplitude of ∼30 nm peak-to-peak (see below) and focused-TIRF illumination (λ = 543 nm). Approach curve measurements where the tip is lowered onto isolated quantum dots and the fluorescence rate is measured as a function of tip-sample separation (data not shown) indicate an enhancement factor of only $f \sim 4$ for these data. The small enhancement in this case results from the fact that the elongated shape of the quantum dots leads to a somewhat small spatial overlap with the region of enhanced field at the tip apex. Furthermore, the absorption dipole for these nanorods should lie predominantly along the sample surface, while the enhanced field is strongest under the tip where it is vertically polarized. This leads to relatively weak near-field excitation of the nanorods.

Our model assumes that the fluorophores, whether quantum dots or fluorescent molecules, do not blink or photobleach. In reality, both quantum dots and molecular fluorophores blink and photobleach, which alters the contrast observed in experimental images. In particular, the background signal S_{ff} will be reduced for a blinking or photobleaching sample compared to an ideal one. Interestingly, this has the effect of increasing the contrast in experimental images in the limit of large fluorophore densities where the fluctuations in the far-field signal caused by blinking and bleaching are small compared to the total far-field signal S_{ff}. However,

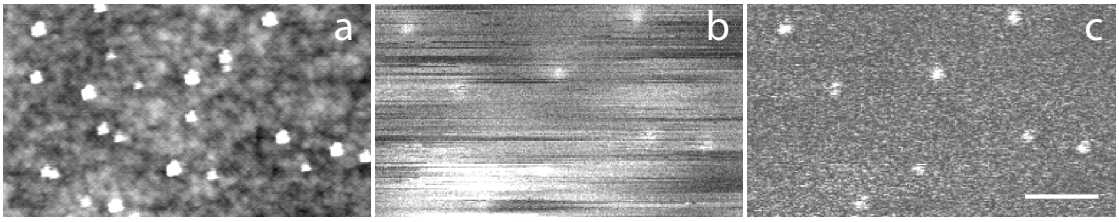

Figure 4.6. TEFM images of a high-density quantum dot sample. Panel (a) shows the AFM topography (∼50 total dots/μm^2). Panel (b) shows the scalar photon sum (∼14 bright dots/μm^2). Panel (c) shows the same image after lock-in demodulation. The scale bar is 200 nm.

the probability of the tip encountering a particular fluorophore that is "on" (i.e. not in a dark or photobleached state) is reduced by the same factor as the far-field signal S_{ff}. The consequence of this is difficult to predict without knowledge of the blinking and photobleaching rates corresponding to the particular fluorophores of interest. This issue is highlighted by Fig. 4.6, which shows the topographic image of a collection of quantum dots in addition to the undemodulated and demodulated near-field fluorescence images. The total quantum dot density as observed by the AFM topography is $\sim$50 μm^{-2}; however, many of the quantum dots do not fluoresce. The bright quantum dot density for this image is $\sim$14 μm^{-2} and there is clearly sufficient contrast to increase the density further; Eq. 4.22 predicts that a density as high as 26 bright dots/μm^2 can be achieved for $f \sim 4$.

4.1.4 Optimizing Tip Oscillation Amplitude

The lock-in contrast and signal-to-noise ratio given in Eqs. (4.20) and (4.21) are strongly influenced by the amplitude of oscillation of the AFM tip, which determines the width of the Gaussian photon-phase distribution, θ_σ, and thus the values of γ and α. Thus, to optimize the lock-in contrast, the product $\gamma \times \alpha$ must be maximized with respect to θ_σ:

$$\frac{d}{d\theta_\sigma}\left(\theta_\sigma e^{-\theta_\sigma^2/2}\right) = 0 \tag{4.23}$$

where the approximations in Eq. (4.18) have been used. Solving Eq. (4.23) for θ_σ gives an optimal value of $\theta_\sigma^{opt} = 1$ radian. The optimal oscillation amplitude, A_{opt}, can now be found using the equation of motion for the tip oscillation, $z = A(1-cos(\theta))$. To relate θ_σ^{opt} to an optimal amplitude A_{opt}, we define z_σ as the value of tip-sample separation z in an approach curve such that the integrated area under the approach curve from $0 \rightarrow z_\sigma$ contains 68% of the near-field photons. The value of z_σ depends on the sharpness of the tip and the size and shape of the fluorescent object: sharp tips and small objects yield the smallest values of z_σ. Substituting $z = z_\sigma$ and $\theta = \theta_\sigma^{opt} = 1$ into the equation of motion for the tip we obtain:

$$A_{opt} = \frac{z_\sigma}{1 - cos(1)} \approx 2.1 z_\sigma. \tag{4.24}$$

When the approximations made in Eq. (4.18) are used, a value of $A_{opt} = 2.18z_\sigma$ is obtained compared to a value of $A_{opt} = 2.11z_\sigma$ when complete numerical integrations are performed.

Experimental values for the contrast and signal-to-noise ratio as a function of the peak-to-peak oscillation amplitude of the tip are shown in Fig. 4.7, along with the theoretical predictions developed above. Isolated ($N_{FA} = 1$) CdSe/ZnS nanorods (4 nm $\times$ 9.4 nm) were imaged with different amplitudes using many different tips from the same fabrication wafer. Each data point was computed from the measured values of S_{peak}, S_{ff}, and σ_{ff}, as used in Eqs. (4.1) and (4.2) for $\sim$15 different quantum dots [61]. The values of $f = 3.7\pm1.3$, $k = 11\pm5$, $\beta = 0.15\pm0.15$, and $z_\sigma = 7.5\pm2$ nm were all obtained from a statistical analysis of image and approach curve data. Subsequently, θ_σ was computed from Eq. (4.24) using the measured value $z_\sigma = 7.5\pm2$ nm to obtain γ and α for each oscillation amplitude from Eqs. (4.7) and (4.17). Thus, the theoretical curves shown in Fig. 4.7 contain no free parameters whatsoever. The predicted peak-to-peak amplitude of 32 ± 9 nm agrees with the

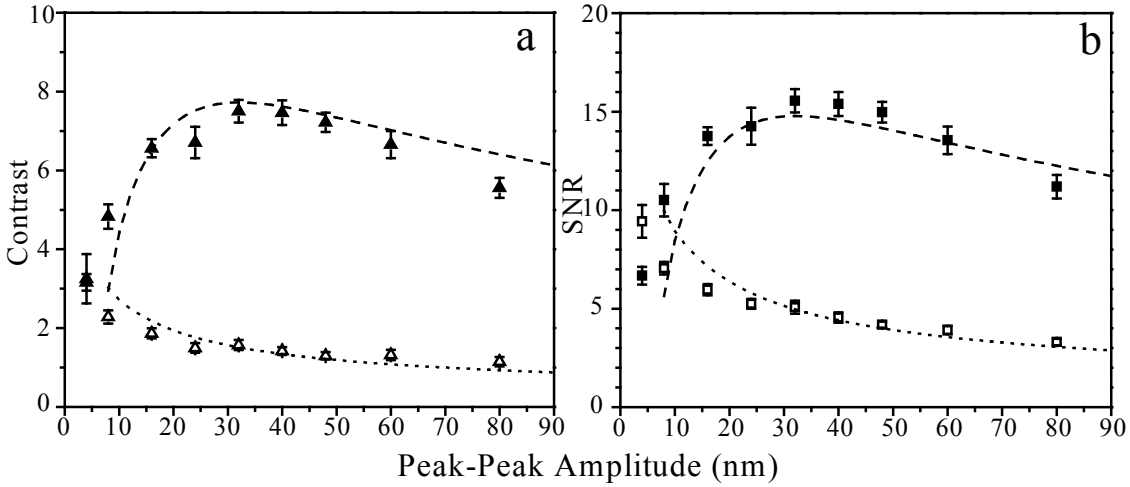

Figure 4.7. TEFM image contrast, panel (a), and signal-to-noise ratio, panel (b), for isolated quantum dots as a function of the tip oscillation amplitude. Data were obtained using BudgetSensors Multi-75 silicon tips. Data points correspond to the average value of $\sim$15 measurements for the lock-in demodulation signal (closed symbols) and the scalar sum (open symbols). Dashed and dotted lines are the corresponding theoretical predictions.

experimental value of 32 ± 4 nm. This good agreement between the predictions of this theoretical model and experimental measurements lends confidence to the calculated values of the signal enhancement factors f requisite for imaging high fluorophore densities found above.

4.1.5 Summary of Contrast Limitations

We have defined the acceptable level of near-field contrast as $C > 1$ and have calculated the amount of signal enhancement needed in TEFM to achieve such contrast for single fluorophores within high density samples. The particular density used in our calculations was 10,000 fluorophores/μm^2, which corresponds to only one fluorophore of the ensemble within the measured near-field zone (10×10 nm^2). Our model uses no free parameters, but rather extracts the values for the relevant parameters from experimental measurements. The model has been validated in part by its agreement with experimental results; mathematical optimization of the tip oscillation amplitude matches experimental measurements, as seen in Fig. 6. Using this model, we have considered the two cases of a near-field probe that is not oscillating vertically above the sample surface (contact-mode AFM imaging) and one that is (tapping mode imaging) for two experimentally-relevant illumination conditions, focused-TIRF and radial polarization. For contact-mode imaging, the requisite signal enhancement factors were calculated to be $f \sim 7500$ for focused-TIRF illumination and $f \sim 600$ for radial polarization. Both of these values are well beyond the maximum measured enhancement of $f \sim 20$ for Si tips. Tapping-mode imaging coupled with lock-in demodulation significantly increases image contrast, thus reducing the requisite signal enhancement factors to $f \sim 65$ for focused-TIRF and $f \sim 18$ for radial polarization. This last case is within the capabilities of commercially available Si AFM tips. Thus we expect that the maximum density achievable with Si tips is not limited by the enhancement factor, but rather by the requirement that each fluorophore be spatially resolved from its neighbors, in this case, at least 10 nm apart.

4.1.6 Conclusions

Determining the structure of extended biomolecular networks, and relating that structure to the physical mechanisms underlying various biological functions, are very difficult and pressing problems in molecular-scale science. Current nanoscale structural analysis tools including x-ray crystallography, electron microscopy, and atomic force microscopy, have a number of limitations that prevent their application to extended networks composed of heterogeneous mixtures of various biomolecules. Fluorescence microscopy, on the other hand, is a very powerful technique for analyzing heterogeneous molecular systems, and when combined with the spatial resolution afforded by apertureless near-field microscopy, holds great promise as a future molecular-scale structural analysis tool.

Although the potential of apertureless fluorescence microscopy in structural biology has been recognized previously, a recurring criticism has been that first-order scattering processes cannot achieve the contrast needed to resolve individual molecules within a dense ensemble. In this work, we have explicitly addressed this issue and have shown both theoretically and experimentally, that it is in fact possible to achieve the needed contrast using carefully designed modulation/demodulation schemes. The key issue discussed was the need to optimize various experimental parameters, such as the oscillation amplitude and material properties of the apertureless tip. Coupled with recent and future advances in scanning probe microscopy, such as imaging in water and fast frame imaging speeds, it may ultimately be possible to combine optical resolution approaching that of electron microscopy with the ability to image bio-molecules in physiological conditions.

4.2 Clarifications

As there have been several questions about the validity of $|\mathbf{L}|_{peak}$ from Equation 4.19, I detail the derivation here. As mentioned, the lock-in signal is the vector addition of the near-field and far-field signals, averaged over all angles. Some confusion may arise if the incorrect order of operations is performed. The proper formulation for the lock-in signal is:

$$|\mathbf{L}|_{peak} = \langle \mathbf{NF} + \mathbf{FF} \rangle \tag{4.25}$$

The incorrect way and source of confusion comes from attempting the wrong order of operations:

$$|\mathbf{L}|_{peak} \neq \langle \mathbf{NF} \rangle + \langle \mathbf{FF} \rangle. \tag{4.26}$$

The proper way to calculate $|\mathbf{L}|_{peak}$ is to first perform the vector addition and then to average over all angles. First we break both **NF** and **FF** into their standard $\hat{x}$ and $\hat{y}$ components:

$$\mathbf{NF} = |\mathbf{NF}| \cos\theta_p \, \hat{x} + |\mathbf{NF}| \sin\theta_p \, \hat{y} \tag{4.27}$$

$$\mathbf{FF} = |\mathbf{FF}| \cos\theta_r \, \hat{x} + |\mathbf{FF}| \sin\theta_r \, \hat{y}. \tag{4.28}$$

where θ_r is some random angle ranging from $0° - 360°$. Now it is straightforward to carry out the vector addition:

$$|\mathbf{L}|^2_{peak} = \left\langle (|\mathbf{NF}| \cos\theta_p + |\mathbf{FF}| \cos\theta_r)^2 + (|\mathbf{NF}| \sin\theta_p + |\mathbf{FF}| \sin\theta_r)^2 \right\rangle \tag{4.29}$$

$$= \left\langle |\mathbf{NF}|^2 + |\mathbf{FF}|^2 + 2\,|\mathbf{NF}|\,|\mathbf{FF}| \cos\theta_p \cos\theta_r + 2\,|\mathbf{NF}|\,|\mathbf{FF}| \sin\theta_p \sin\theta_r \right\rangle \tag{4.30}$$

Now since θ_r ranges over all angles, the cross terms drop out.

$$|\mathbf{L}|^2_{peak} = |\mathbf{NF}|^2 + |\mathbf{FF}|^2 \tag{4.31}$$

The final form of this equation may lead one to believe that the vectors have simply been added in quadrature, without regard to averaging over any angles. However, it must be remembered that **NF** and **FF** are not orthogonal to each other and thus to simply add in quadrature would be a naïve approach, which naïveté surprisingly has the same effect as averaging over all angles after the proper vector addition has been performed.

Taking the expectation values of the near-field and far-field terms before performing the vector addition leads to a complete removal of the far-field term such that the lock-in signal is simply equivalent to the near-field signal. By taking the limiting case it can be seen that this is not correct: assuming the near-field signal is nonexistent, this assumption would imply that the lock-in signal is also nil. This

is clearly not true as when no near-field signal is present the lock-in signal scales as the square root of the far-field signal S_{FF}, which is achieved as presented in Equation 4.19. Thus the presentation in the original paper is correct.

4.3 Negative Contrast

One other consideration is to determine the validity of the above formulations in the context that the near-field signal purely stems from quenching. Just as with enhancement, quenching is a near-field signal also capable of being used as a contrast mechanism; a decrease of signal can be as easily recognizable as an increase in signal. Contrast using quenching suffers from the exact same density issues as enhancement; as the far-field signals of multiple fluorophores begins to overlap the near-field signal, the contrast is washed out. Thus, the above formulation is also valuable for estimating the limits of quenching as a contrast mechanism. Only minor modifications need to be made to above formulation to account for the effects of quenching. First, the definition of the enhancement factor f must be expanded:

$$f > 0 \Rightarrow \quad \text{enhancement} \tag{4.32}$$

$$f = 0 \Rightarrow \quad \text{no near-field signal} \tag{4.33}$$

$$-1 \leq f < 0 \Rightarrow \quad \text{quenching.} \tag{4.34}$$

The lower bound on f in quenching comes from the definition of contrast. Again we have:

$$C = \frac{S_{peak} - S_{FF}}{S_{FF}} \tag{4.35}$$

Since negative signals are nonexistent, the peak signal in perfect quenching is $S_{peak} = 0$. This implies the greatest "enhancement" factor for quenching is $f = -1$.

Under this expanded definition for the enhancement factor, almost all quantities presented in the original paper can be used as is. One nuance is that finding the value for θ_σ must be defined in the opposite way, such that it refers to a lack of photons; this can be found by performing a histogram, fitting to an inverted Gaussian profile, and obtaining the standard deviation of the fit. Quantities derived above involving the lock-in are also valid in this expanded definition of enhancement

factor; the lock-in recognizes a lack of signal in exactly the same way it does for enhancement. In fact, when looking only at the magnitude of the lock-in signal, enhancement and quenching can be indistinguishable. Figure 4.8 demonstrates this principle showing negative contrast on a sum image, while the simultaneously acquired lock-in image shows positive near-field contrast. The only equation presented that is no longer valid for quenching is Equation 4.22, as the maximum possible contrast for quenching is $C = -1$. That said, Equations 4.23 and 4.24 may still be used to calculate the optimum oscillation amplitude for quenching.

There are several disadvantages to using quenching as a contrast mechanism that are worth discussing. The quenching scale is considerably longer than enhancement when using commercial AFM tips. This means that the obtainable resolution using quenching is worse than is possible with enhancement. However, because of this longer interaction range, metals tips can be easier to align in the center of a laser focus than dielectrics. Another disadvantage of quenching as a contrast mechanism is that the maximum attainable contrast is $C = -1$, while in principle enhancement can yield unlimited contrast. That said, an image contrast of $C = 1$ (or $C = -1$) can be rather acceptable as evidenced by Figure 4.8. Another drawback of quenching is that the maximum density that will provide adequate contrast is greatly reduced. This is in part due to the limitation on "enhancement" factors for quenching, but also the worse optical resolution that metal tips offer.

4.4 Enhancement vs. Quenching

One drawback of the equations presented in this chapter is that for either enhancement or quenching, we are able to produce a nice monotonic approach curve. As discussed in Chapter 2, it is quite often the case that a near-field signal is comprised of several competing factors. Thus, determining the expected contrast from such complicated near-field effects can be more challenging. The formulations above are amplitude dependent, remembering that f refers only to the peak enhancement value. Were it possible to measure the quenching and enhancement from the same tip in two independent measurements, this theory

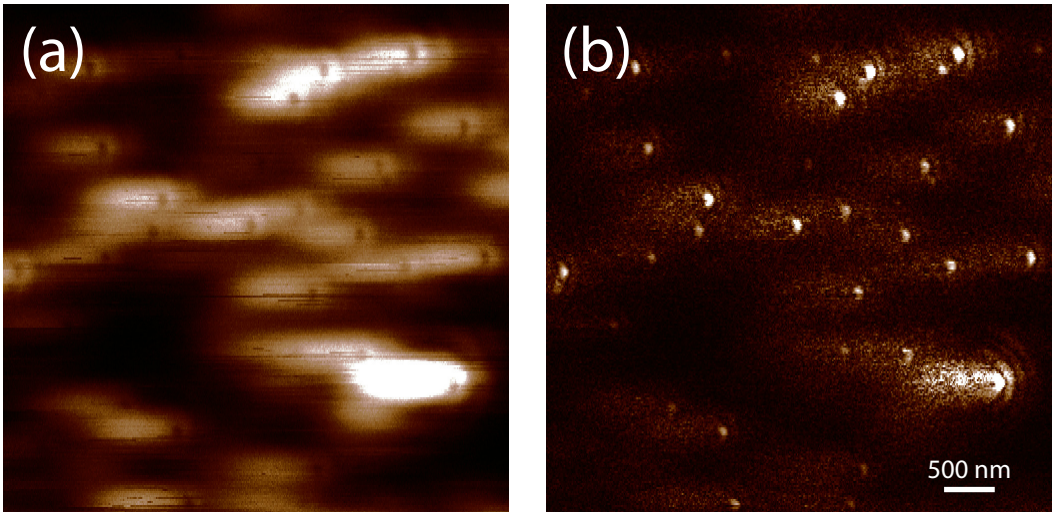

Figure 4.8. An example of negative near-field contrast. A gold coated tip was used to image 20 nm dye-doped latex beads on glass. Vertical wedge illumination was employed at 543 nm. Panel (a) shows the fluorescence sum signal; the position of the beads corresponds to the dark regions where the fluorescence is quenched. Panel (b) shows the corresponding lock-in image; the lock-in magnitude yields positive near-field contrast. Note that some longer range interference effects are clearly observable. Also notice that contrast improves with increasing intensity as some beads are brighter than others.

should also allow for calculation of the total contrast. The total enhancement factor is comprised of separate factors for enhancement, f_e, and quenching, f_q. Of course, quenching and enhancement have different functionalities in the vertical direction, so those parameters must also be treated separately; θ_σ^e for enhancement and θ_σ^q for quenching. Plugging these into Equations 4.7 and 4.17 we would obtain γ_e, γ_q, α_e, and α_q respectively. To obtain the total contrast from the lock-in (C_{LI}), the following substitutions must be made in Equation 4.20:

$$f = (f_e + 1) \cdot (f_q + 1) - 1 \tag{4.36}$$

$$\gamma = \gamma_e \cdot \gamma_q \tag{4.37}$$

$$\alpha = \alpha_e \cdot \alpha_q. \tag{4.38}$$

Of course the difficulty in all this is the ability to measure θ_σ^e and θ_σ^q on the same tip.

Again, the calculations presented in this chapter deal with the image contrast obtainable in the sample plane. Ideally one would like to have a full representation of the near-field interactions in a three-dimensional space in an X-Y-Z representation, rather than trying to collapse the net effects the various interactions onto a two-dimensional image. This is exactly the direction the research in our lab has moved to and is presented in Chapter 5.

CHAPTER 5

TOMOGRAPHY

While the previous chapter addresses the amount of attainable image contrast using lock-in demodulation techniques, there must undoubtedly be more powerful methods of suppressing the far-field background signals that plague apertureless-NSOM. When sufficient signal is present, demodulation at higher harmonics is possible, and can lead to not only better suppression of far-field signals, but also increases in resolution [71]. However, in order to be able to demodulate at higher harmonics, extremely high photon fluxes are necessary as only a fraction of the photons emitted correspond to higher harmonic frequencies. Since TEFM is a fluorescence method, every photon collected is precious; photobleaching of the sample usually limits the amount of excitation power that can be employed, and consequently total fluorescence emission rates do not allow for demodulation at higher harmonics. In order to maximize the utility of every photon collected, rather than simply counting the number of photons collected in a given time window, the precise arrival time of each photon can be recorded. This time-correlated single photon counting (TCSPC), allows for exquisite control of postprocessing of the data, and should allow for arbitrarily good image analysis algorithms to be applied. Some work towards applying various processing algorithms to TEFM has previously been demonstrated by Jordan Gerton and colleagues at Caltech [2].

-processing techniques. By time-stamping the motion of the tip, in addition to the fluorescence photons, the tip-sample separation distance can be correlated with the emission of each photon. In other words, the time of the photon emission and the corresponding tip-height at that time are both recorded. This technique allows for creation of a full three dimensional map of the near-field interactions. Since the

three dimensional map can be sliced in any arbitrary fashion, it is a tomographical technique.

The tomography technique itself, including a few relevant examples, was published in Nano Letters, volume 9, pages 3440-3446 (2009). This chapter includes a reproduction of the Nano Letters paper in its entirety, but has been reformatted to fit this dissertation. Additionally, section headings not present in the original have been included here. Furthermore, material originally presented as supplementary material available online has also been included, with the exception of a movie that could not be included in this print document. This chapter then necessarily includes some redundancies with previous chapters, which is important for the flow and understanding of the concepts to be presented.

5.1 Introduction

Light-matter interactions are of fundamental importance for a host of nanoscale phenomena and emerging nanotechnology applications. Such interactions can be altered by proximate surfaces, particularly near nanostructures with sharp facets such as lightning rods [33, 11], optical antennae [72, 47, 14], and roughened surfaces [73, 74, 75]. For example, in surface-enhanced Raman spectroscopy (SERS) the optical scattering cross-section is enhanced for molecules near "hot spots" on a nanostructured surface, leading to an increase in the molecular detection efficiency [76]. The enhanced near-field at the end of a sharp tip can also be used for high contrast optical microscopy/spectroscopy of molecules on a surface with spatial resolution down to the nanometer scale [32, 41, 3]. Recently, a large amount of effort has been dedicated to optimizing the size and shape of optical nano-antennae to enhance light-matter interactions, thereby altering the direction and rate of a dipole's radiated emission [72, 47, 14, 56].

Within the context of apertureless near-field scanning optical microscopy (ANSOM), the optical signal is generally proportional to a dipole's excitation rate, which in the linear regime is in turn proportional to the local light intensity [56]. In ANSOM, the tip can enhance the optical intensity via both the lightning-rod

effect [77, 11] and plasmon resonances [72, 47, 14, 56], and can also induce an interference pattern from the superposition of the excitation light and the light scattered off the tip [39, 40, 41, 26]. These two effects act independently to create a three-dimensional intensity pattern that contains fluctuations on a number of length scales ranging from nanometers to roughly the wavelength. Additionally, a proximate tip can also suppress a dipole's spontaneous emission rate via non-radiative energy transfer (fluorescence quenching) [23, 1, 25, 22, 12, 78, 27]. The quenching efficiency depends sensitively on the probe material, being highest for metals. All three tip-induced effects, field enhancement, optical interference, and fluorescence quenching, are exquisitely sensitive to tip geometry. The interplay of the complex intensity pattern and the quenching distribution yields an optical image with rich, three-dimensional structure whose precise shape is due to the particular tip-sample coupling mechanisms at play for the given experimental conditions [18]. In this letter, we present the first method that can produce a three-dimensional, sub nanometer precision map of the optical signal resulting from these tip-sample interactions.

5.2 Methods

A three-dimensional theoretical model of the interplay between the complex tip-modified intensity pattern and the fluorescence quenching distribution is quite difficult to generate, and is left to the future. Rather, in this work we use the three-dimensional structure of the resulting optical signal to demonstrate the capabilities of our tomographical reconstruction technique. In particular, we reveal this intricate pattern by scanning 20-nm diameter dye-doped latex spheres through the region near the apex of an illuminated atomic force microscope (AFM) tip. These spheres contain several tens of individual dye molecules (up to $\sim$200) with random orientations, thus mitigating effects related to the direction of the emission dipole moment. Using these spheres, the topographical signal is easily and accurately registered with the optical image. The resulting composite images reveal three-dimensional features of the tip-sample coupling mechanisms in extraordinary detail.

This work differs from previous reports in that it reveals the full three-dimensional structure of the net tip-sample coupling map, rather than simple one-dimensional approach curves [2, 12, 26, 78] or two-dimensional maps that obscure structure in the third (vertical) dimension [25, 41, 3, 7, 26, 78, 79].

Our method is based on scanning near-field optical microscopy (SNOM), or more particularly, tip-enhanced fluorescence microscopy (TEFM) [7, 6, 8]. In this technique, a focused laser beam illuminates a nm-scale AFM tip composed of either metal (typically gold) or silicon, which is brought into close proximity with the sample of interest (Fig. 5.1). The fluorescence rate is monitored as the sample is scanned laterally relative to the tip position. Significantly, large amplitude vertical oscillations in the tip-sample distance can be obtained when the cantilever is resonantly excited via a piezoelectric transducer. As described below, these oscillations provide the means to map the tip-sample coupling in the vertical (Z) direction, i.e. transverse to the sample plane. Any method of tip-sample distance control that can be made to induce significant vertical oscillations of the

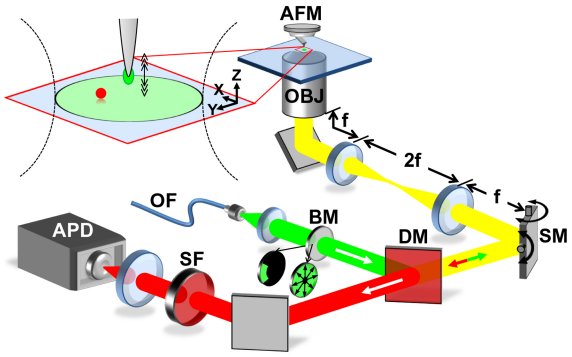

Figure 5.1. Schematic of experiment. A laser beam is directed through a beam mask (BM), producing either a radially polarized laser beam or a 60° section (wedge) of an annular beam. A microscope objective (OBJ: NA = 1.4) focuses the laser beam and collects emitted fluorescence, which is focused onto an avalanche photodiode (APD). The laser focus is positioned onto an AFM tip (oscillation amplitude: 10 - 500 nm) using a scanning mirror (SM). The sample is raster-scanned laterally, where by convention the X-axis corresponds to the fast-scan direction. The inset shows the tip-sample interaction region in more detail. Other important components include an optical fiber (OF), a dichroic mirror (DM), and a spectral filter (SF).

tip, including all types of both cantilever and tuning-fork based SNOM techniques [79], can be employed to produce the desired three-dimensional map.

Our TEFM system utilizes a commercial AFM (Asylum Research) sitting atop a custom optical setup. A green He-Ne laser ($\lambda = 543$ nm) is used for excitation and the laser beam is directed through either a radial polarization converter (Arcoptix) or through a focused-TIRF (total internal reflection fluorescence) arrangement. The radially polarized beam produces longitudinal polarization (along the tip axis) at the sample interface. In focused-TIRF, all but a small wedge of supercritical rays are blocked in the infinity space behind the microscope objective, producing an evanescent field above the interface within a near diffraction limited spot (≈ 1.5 μm $\times$ 0.5 μm). The polarization of the annular beam can be adjusted to produce longitudinal or transverse polarization at the interface. The laser beam is focused through a high numerical aperture objective (NA = 1.4) and fluorescence signals are collected through the same objective and detected by an APD (Perkin Elmer). The laser focus is aligned onto an AFM probe; when the excitation polarization is longitudinal, an enhancement in the optical intensity is expected at the tip apex due to a nonresonant lightning-rod effect (i.e., far from any plasmon resonance). Topographical and photon sum signals are recorded simultaneously via the AFM controller and displayed in real time. The APD signal can also be analyzed in real time using a commercial lock-in amplifier (Stanford Research Systems) [7, 8]. These real time signals are primarily used for aligning the tip and laser and to produce one and two dimensional maps of the near-field coupling between the tip and sample.

To obtain three dimensional maps, a pair of data acquisition cards (NI PCI-6251, NI USB-6210) is used to record photon arrival times, tip oscillations, and AFM line markers at 80 MHz sampling rate. The tip oscillations are obtained by transforming the AFM deflection signal into a square wave that triggers a time measurement at a particular phase of each oscillation cycle. After acquiring these data channels, the tip oscillation phase corresponding to each photon arrival is computed and recorded to a computer disk. If the cantilever oscillation is harmonic, the photon phase delays can be correlated with the height of the tip above the sample after

calibrating the tip oscillation amplitude. Alternatively, if the cantilever oscillations are anharmonic, the AFM deflection signal can be recorded by an analog-to-digital card at a predefined sample rate, enabling real time measurement of the precise cantilever trajectory. In either case, the fluorescence rate can be correlated with the lateral position of the probe using the AFM line markers.

As the sample is rastered, histograms of photon phase delays are accumulated and parsed into the assigned lateral $(x-y)$ pixels. This results in a two-dimensional array of vertical approach curves, which represents a three-dimensional map of the near-field coupling between the tip and sample. The photon sum for each three-dimensional pixel (voxel) is normalized to its corresponding acquisition time. Thus, the value of a given voxel within the three-dimensional image space corresponds to the fluorescence count *rate* at that particular coordinate (x, y, z). Since the photon count rates are stored as a three-dimensional array, the data can be sectioned arbitrarily. This method for producing three-dimensional tomographical reconstructions relies upon the phase-correlation of single photons, and we therefore call it single photon near-field tomography, or SP-NFT.

Figure 5.2 shows an X-Z slice generated by graphically displaying a linear array of vertical approach curves at a particular value of y (the slow scan axis). Here, the near-field coupling between a 20-nm diameter fluorescent sphere and a gold tip is measured along a Y-section that cuts through the topographical center of the sphere. When the tip is sufficiently far from the sample ($z \sim \infty$) such that no tip-sample coupling occurs, voxel values reflect the laser-induced far-field fluorescence rate, $S_{z\sim\infty} \approx S_{ff}$. Thus, a precise measure of the *local* far-field contribution to the fluorescence rate is obtained at each lateral position by averaging the values of a number of voxels corresponding to large tip heights. At small tip-sample separations, the fluorescence signal contains contributions from both far-field and near-field interactions $S_{z\sim0} \approx S_{nf} + S_{ff}$. The tip-induced modification to the fluorescence rate is then easily isolated by subtracting the local far-field rate, as shown in Figure 5.2(e). This subtraction dynamically accounts for spatial variations and temporal fluctuations in the laser intensity and also for single-molecule blinking

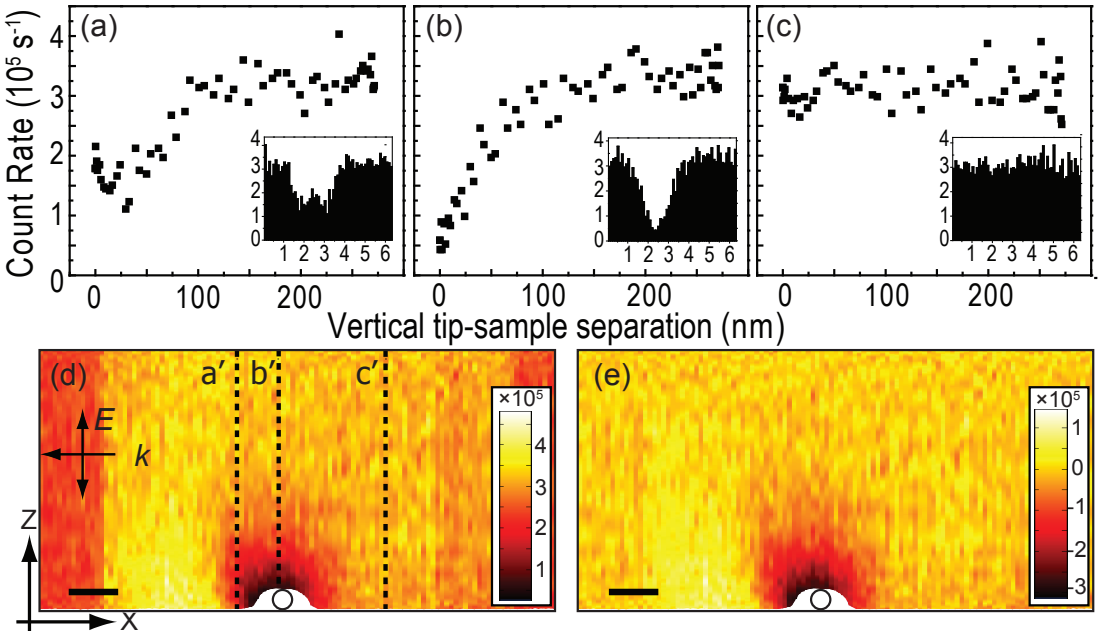

Figure 5.2. Tomographical reconstruction of a 20-nm diameter fluorescent sphere. The vertical approach curves and associated phase histograms shown in panels (a) - (c) correspond to the lateral positions indicated by the dashed lines a′, b′, and c′ in panel (d). The lower image panels are total (d) and far-field subtracted (e) X-Z sections where each column of pixels corresponds to a vertical approach curve. In all panels, the detected photon counts have been normalized to the relevant acquisition time to generate photon count *rates*. For these data, focused-TIRF illumination was used (Fig. 5.1) to produce an evanescent field with polarization parallel to the long axis of a gold-coated tip. The scale bars correspond to 50 nm.

or slow photobleaching of a multi-chromophoric sample. To further elucidate three-dimensional variations of the tip-sample coupling, the fluorescence data can be superimposed onto the topographical data. This is shown in Figure 5.2(d) and (e) as the white cutout region at the bottom of the image panels; the actual size of the fluorescent sphere is indicated by the circular outline in that region. Duplicity in the fluorescence signal acquired with the data acquisition card and the AFM controller accounts for any electronic delays between the two, ensuring that the *topography* signal is properly registered with the SP-NFT image.

5.3 Tomographical Slices

Figure 5.2 reveals a number of interesting effects that would be difficult to capture accurately with either a two-dimensional lateral image or a one-dimensional approach curve. First, there is a three-dimensional "halo" of reduced fluorescence signal extending ∼100 nm from the surface of the sphere. In this region there is no observable enhancement in the fluorescence signal even though the axial polarization of the excitation field should lead to an enhancement in the optical intensity at the tip. For metal tips with a closed geometry, such as a sphere or prolate spheroid, the lightning rod effect and surface plasmon resonances can augment each other leading to exceedingly large field enhancement factors near the plasmon resonance frequency. However, for metal tips with an open geometry, such as commercial AFM probes, surface plasmons will be strongly dissipated [77, 18, 26, 78]. Furthermore, elongated tip geometries lead to a large red shift in the plasmon resonance spectrum. These two factors act to suppress plasmon-based field enhancement in this experiment, which used commercial gold-coated AFM tips and a green excitation wavelength (543 nm).

Ohmic dissipation mechanisms are responsible for strong quenching of fluorescence by metal tips, and reduce the optical signal in direct competition with field enhancement [1, 25]. This tug-of-war has been observed in recent one-dimensional approach curve measurements using spherical gold tips [12, 27]. In those studies, the closed spherical geometry led to clear field enhancement, which was mitigated by quenching only at very short range. For tips with open geometries, the balance between field enhancement and fluorescence quenching is pushed toward weaker enhancement and stronger quenching [13]. In this work, commercial gold-coated AFM tips were used, and the quenching evidently overwhelms field enhancement even at very short tip-sample separation distances. The complete lack of observable signal enhancement for metal-coated tips with pyramidal geometry has also been observed in previous work [1, 25, 26, 78].

Note that the approach curve shown in Figure 5.2(a) corresponding to axis a′ to the left side of the sphere exhibits partial recovery in the fluorescence signal.

If Figure 5.2(a) had been obtained in isolation, it would be tempting to attribute the partial recovery to a more favorable balance between field enhancement and fluorescence quenching. However, after several tens of measurements at this illumination wavelength (543 nm), such a partial recovery has never been observed when the approach curve is measured along an axis directly above the geometric center of the sphere, e.g., panel (b) corresponding to axis b′. Thus, we attribute the partial recovery in this case to the particular location with respect to the detailed shape of the tip-sample coupling map at which this approach curve was acquired. Specifically, at this location the sample is quenched to a lesser extent so that the fluorescence signal begins approaching the far-field background rate as the tip is lowered to the glass surface. On the other hand, Figure 5.3 (originally in the supplementary material) shows approach curve data obtained using an excitation wavelength of 633 nm. In that case, the partial recovery in the fluorescence rate

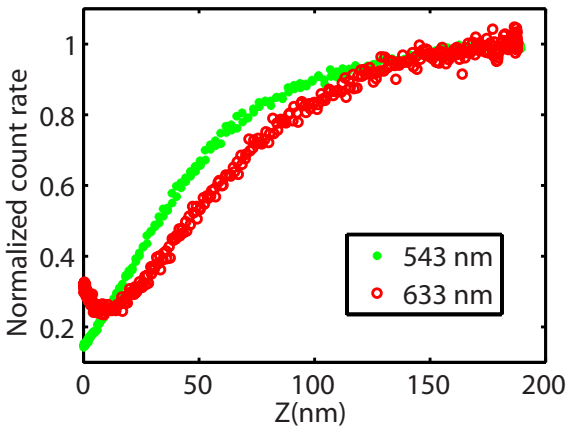

Figure 5.3. Comparison of approach curves at two different wavelengths, 543 nm (solid green circles) and 633 nm (open red circles). Both curves were obtained using gold-coated AFM tips on dye-doped latex spheres of nominally the same size (20 nm) with vertically-polarized evanescent illumination. Although these data were acquired by freezing the lateral motion of the tip (to improve photon statistics), the shape of each curve agrees with those extracted from a full three-dimensional data set. Clearly the competition between field enhancement and fluorescence quenching tilts more toward enhancement at 633 nm. However, the monotonic decrease in observed fluorescence rate at 543 nm for decreasing tip-sample separations does not indicate a complete lack of field enhancement, rather that quenching overwhelms enhancement at all length scales.

is due to a more favorable balance between enhancement and quenching. This more favorable balance is expected at longer wavelengths given the elongated tip geometry, as discussed above.

The optical interference pattern generated by the superposition of the direct excitation light and that scattered from from the tip also contributes to the three-dimensional structure of the image. In Figure 5.2, this interference pattern manifests itself as the bright region to the left of the sphere position, which arises from the constructive interference between the evanescent wave and the light scattered off the tip when it is downstream of the sphere. For these data, focused-TIRF illumination was used (Fig. 5.1 - annular wedge beam mask) yielding a vertically polarized evanescent field at the glass-air interface traveling from right to left. Such interference features have been observed in previous work when similar illumination conditions were used [39, 40, 26, 41]. As shown Figure 5.4 (originally in the supplementary material), when the polarization of the evanescent field is horizontal (perpendicular to the tip axis) these long-range features are strongly suppressed and only the primary dark halo is visible [40]. This is reasonable since the effective dipole moment of the tip should be predominantly along the tip axis, so light scattering should be weaker for horizontal polarization.

Note that the dark halo shown in Figure 5.2 is not spherically symmetric and is skewed to the left. This asymmetry is independent of both the propagation direction of the evanescent field and the sample scan direction (see Fig. 5.4), but does depend somewhat on the particular tip used. Thus, the asymmetry of the halo is evidently a tip-specific effect. Although a detailed explanation of this effect would require more extensive investigation, it is likely that the redirection of fluorescence emission caused by an antenna-like coupling between the tip and the fluorophores in the sphere plays a role. This type of coupling has been demonstrated previously [14], and would contribute to the observed reduction of the fluorescence signal by biasing the emission pattern toward the tip, and thus away from the collection solid angle of the microscope objective. Within this context, the observed skew in the dark halo is likely caused by a tilt in the effective antenna axis of the gold tip or by

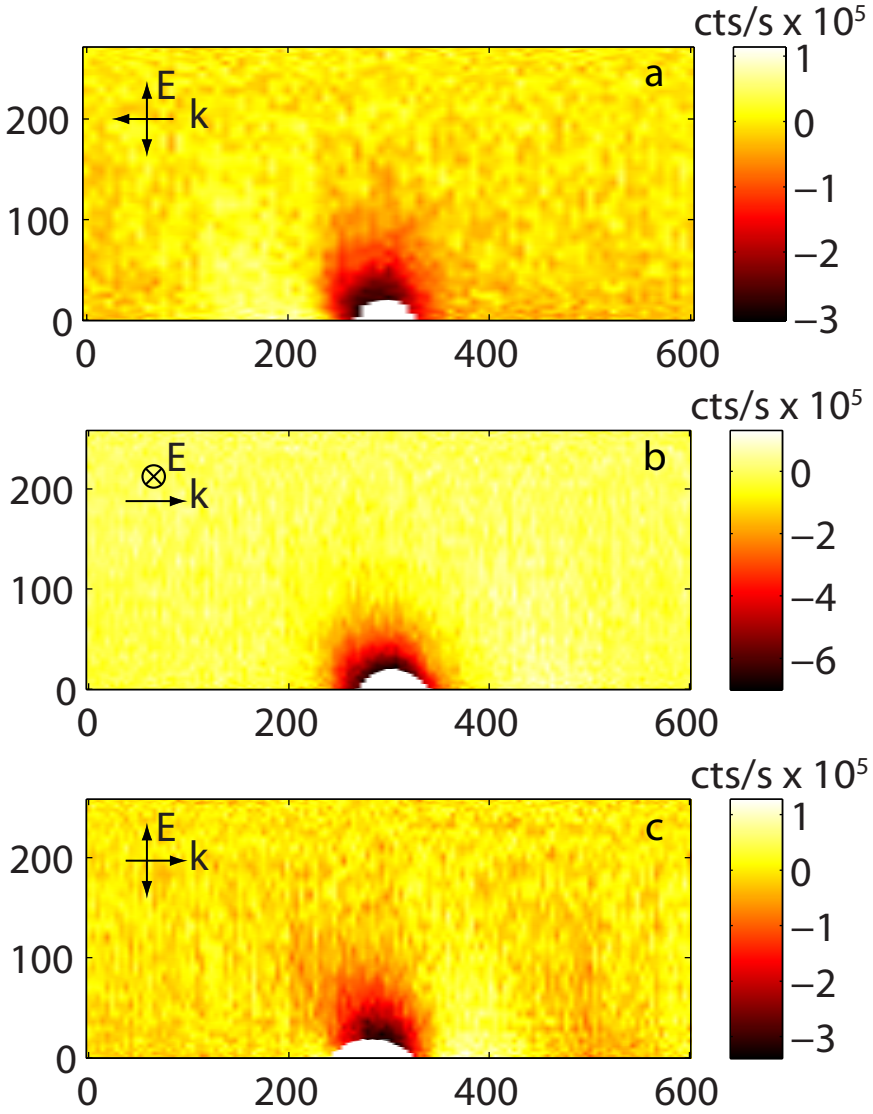

Figure 5.4. Comparison of different illumination and scan conditions from Figure 5.2. All images are obtained with evanescent illumination at 543 nm excitation and gold-coated AFM tips and have been far-field subtracted. Distances on both vertical and horizontal axis are in nm. The AFM topography can be seen in white. Notice the leftward skew of the central dark halo in all images. Retrace of 5.2 is shown in (a). Panel (b) shows that the bright spot due to interference is drastically reduced for s-polarized wedge illumination. Panel (c) corresponds to vertically polarized light incident from the left instead of the right. The same tip was used for both (b) and (c).

an asymmetric shape. In this case, this effect is subtle since the sample is composed of a large number of fluorescent molecules with random dipole orientations, and only due to the extremely high precision of our technique is it visible at all.

SP-NFT can generate two-dimensional planar sections with arbitrary orientation and position. To illustrate this ability, Figure 5.5 shows the intersection of an X-Z and a Y-Z plane on top of an X-Y plane ($z \sim 0$) for a 20-nm diameter fluorescent sphere. In this image, strong enhancement in the fluorescence signal is clearly visible only when the tip is directly above the sphere and within ∼10 nm of its surface. These data were taken using a silicon tip with radial polarization, and the fluorescence enhancement is attributed to the lightning-rod effect in agreement with previous observations [11, 2, 41, 7, 3, 8]. The *signal* enhancement factor $(S_{z\sim 0}/S_{z\sim\infty} - 1)$ under these conditions has been measured to be as large as ∼8. Since the enhanced-field volume is much smaller than volume of the fluorescent sphere, however, the *field* enhancement factor is significantly larger [2, 8].

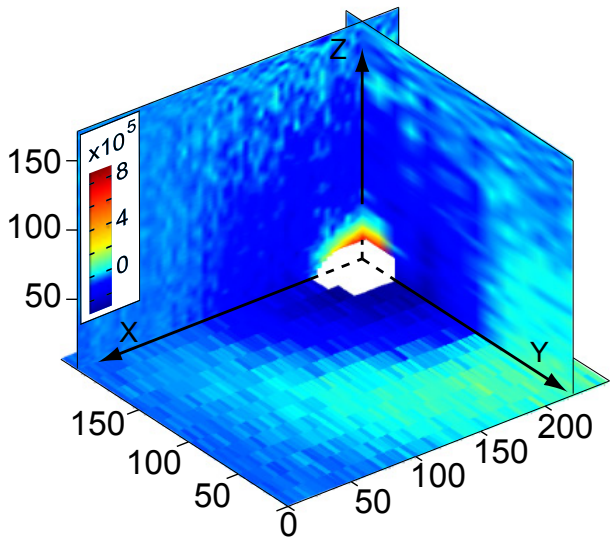

Figure 5.5. Three-dimensional tomographic reconstruction of a 20-nm diameter fluorescent sphere using a silicon tip. The intersection of two vertical planes, X-Z and Y-Z, that cut through the topographical center of the sphere is shown on top of a $z \sim 0$ plane. The local far-field fluorescence rate has been subtracted from these data as described above. Note, the colorbar has been chosen to emphasize the spatial extent of the enhancement and the surrounding dark halo.

The interference pattern generated by the superposition of the excitation light and that scattered from the tip is clearly visible in Figure 5.5 as the three-dimensional halo of fluorescence suppression that surrounds the sphere and a bright halo beyond that. In this case, the radial polarization beam mask was used, so the interference pattern is fairly symmetric around the sphere. The three-dimensional symmetry of the halo would be very difficult to deduce from one- or two-dimensional measurements. Evidently, short tip-sample separation distances correspond to destructive interference, presumably due to a 180° phase shift upon scattering from the tip, irrespective of whether the tip is silicon or gold. The resulting dark halo surrounding the tightly confined enhanced near-field region extends ∼120 nm away from the tip. These three-dimensional renderings clearly show that the length scale for tip-induced interference is much longer than field enhancement, but comparable to quenching. For silicon tips the interference pattern is not obscured by quenching at short tip-sample separations. For gold tips, however, it is difficult to fully separate the contributions of these two effects without performing lifetime measurements.

5.4 Image Contrast

Another important aspect of the single-photon analysis method described here is the potential to increase contrast in tip-enhanced images. This will be particularly important for complex samples composed of closely spaced fluorophores, such as biological samples, where the background fluorescence signal will be large [8]. One way to increase contrast is to demodulate the fluorescence signal at the tip oscillation frequency [2, 3, 7, 8, 79], which can be done by directing the output pulses from the APD into a commercial lock-in amplifier, along with a reference wave from the cantilever oscillation signal [7, 8] (Fig. 5.1). In the context of our single photon counting technique, each photon can be considered a unit vector in phase space whose angle is given by its phase delay relative to the preceding tip-oscillation timestamp. Computationally, a lock-in algorithm is essentially equivalent to performing a *vector* rather than a scalar sum of the photon signal. Thus, photons arising from tip-induced field enhancement cluster around the particular phase corresponding to

tip-sample contact, while background photons are distributed uniformly in phase space. By the same token, tip-induced fluorescence suppression (e.g., quenching), results in a relative lack of photons at the tip-sample contact phase, and thus the net lock-in phase tends toward the value corresponding to maximum tip-sample separation. Note that the offline lock-in analysis described here generally yields better performance than a commercial lock-in amplifier because the fidelity of each photon is preserved perfectly. In addition, averaging windows and spectral filters can be made arbitrarily sharp.

Figure 5.6 shows a detailed comparison between near-field images of a 20-nm diameter fluorescent sphere generated by sectioning the data into X-Y planes (a - c) and the offline lock-in algorithm described above (e - g). These data were obtained using radial polarization and a silicon tip. Although difficult to display on a two-dimensional page, SP-NFT can easily generate a stack of two-dimensional images to elucidate the full three-dimensional symmetry of the tip-sample interaction (see supplementary material). Panels (a) and (b) of Figure 5.6 show two such X-Y sections corresponding to two different ranges of tip height, $z \geq 145$ nm (a) and $z \leq 1$ nm (b). Panel (c) is the difference between the two panels (b - a), and thus provides a graphical representation of the tip-sample coupling at $z \sim 0$. In panel (e), the value of each lateral pixel is the magnitude of the vector photon sum normalized to the pixel acquisition time, while panel (g) shows the phase angle of this resultant vector. Panel (f) shows the vector photon sum after the application of a phase filter that eliminates photons which do not contribute to the near-field signal, as described below.

Note that it is not possible to determine if a particular commercial lock-in signal corresponds to an increase or decrease in the fluorescence signal at tip-sample contact. However, since the offline lock-in analysis is derived from the same dataset used to generate the X-Y sections, it is trivial to perform this correlation with our system. Furthermore, a lock-in amplifier essentially reports the *net* change in fluorescence signal as the tip-sample distance is modulated during an oscillation cycle. As such, a lock-in analysis cannot reveal nonmonotonic variations in the

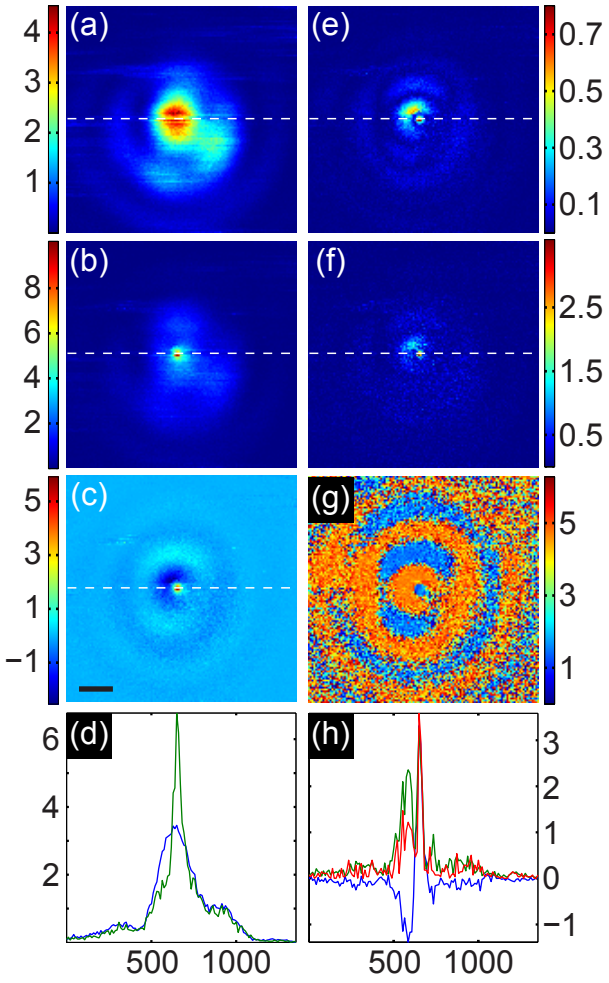

Figure 5.6. Comparison of tomographical and lock-in analyses, all colorbars represent count rates of 10^5 counts per second. Panel (a) shows an X-Y section corresponding to the *scalar* count rate of photons when the tip is far from the sample ($z \geq 145$ nm), panel (b) corresponds to $z \leq 1$ nm, and panel (c) is their difference (b - a). The signal profiles in panel (d) are taken along the dotted line, and include the count rate from panel (a) in blue, and the count rate from panel (b) in green. Panel (e) shows the magnitude of the *vector* sum of all detected photons within each pixel, which is essentially equivalent to lock-in amplification. Panel (f) shows a phase-filtered lock-in algorithm generated from the vector sum of photons within narrow phase windows. Panel (g) shows the phase of the resultant vector from (e). The signal profiles in panel (h) include the count rate from panel (e) in green, the count rate from panel (f) in red, and count rate from panel (c) in blue. The data in panels (e) and (f) and their corresponding signal profiles have been normalized to the peak value in panel (c).

fluorescence signal as a function of tip-sample separation distance. In contrast, such variations are preserved through SP-NFT. For instance, Figure 5.5 shows both a decrease and increase in the fluorescence signal for decreasing tip-sample distance for lateral positions directly above the sphere. These non-monotonic signal variations are also present in the data corresponding to Figure 5.6, but are not visible in the lock-in image. These observations demonstrate another way in which SP-NFT outperforms a commercial lock-in amplifier when demodulating near-field images. Furthermore, this performance is achieved at a fraction of the cost of a commercial instrument [7, 8].

The lock-in image pattern in panel (e) of Figure 5.6 is similar to the far-field subtracted image in panel (c), with a sharp central maximum due to tip-induced field enhancement, surrounded by a diffuse ring corresponding to the interference effect discussed above. For imaging purposes, the interference pattern is artifactual, as it will reduce contrast particularly for dense samples. One way to suppress such artifacts is to apply a phase filter to the data before performing the vector sum, as shown in Figure 5.6(f). Here the phase filter was chosen to eliminate all photons from the data set except those within a narrow phase window centered at tip-sample contact ($z \sim 0$). To help offset the far-field contribution within this window, photons within an equal-width phase window centered 180° from tip-sample contact ($z \sim \infty$) are also included in the vector sum [2, 3]. This algorithm eliminates photons corresponding to features with long length scales in the vertical (z) direction, without affecting those photons that contribute directly to the near-field signal [8]. This algorithm is globally applied photon by photon, and the signal cross-sections in Figure 5.6(h) clearly demonstrate that the halo-like imaging artifact is suppressed, although not completely eliminated. It is crucial to note that there is no limit to the number and variety of analysis algorithms that can be implemented to optimize image contrast for any particular situation.

The tip-induced interference pattern is visible as the dark halo in Figure 5.6(c), but is most clear in the lock-in phase image shown in panel (g). Under these conditions, the lock-in phase primarily adopts two values in the tip oscillation trajectory

corresponding to tip-sample contact and maximum tip-sample separation, which very sensitively report an increase or decrease, respectively, in the fluorescence rate relative to the background. Note that this interference pattern should not be confused to be a far-field Airy pattern, as evidenced by Figure 5.7. This binary distribution of the lock-in phase holds fairly rigorously even when the tip is relatively far from the fluorescent sphere. When this is the case, the sphere is illuminated by the periphery of the laser focal spot (e.g., Fig. 5.6(a)) where the intensity is quite low, while the tip is positioned near the center of the focus. Despite this, the lock-in phase is not strictly random, which clearly highlights the spatial extent of the tip-induced optical interference pattern. A number of measurements using different tips, both silicon and gold, demonstrate that the symmetry of the interference pattern depends on *both* the detailed intensity pattern within the far-field focal spot *and* the particular tip geometry. For clarity, we have chosen a dataset with minimal asymmetry, but the pattern can become distorted when using tips with different geometries, or if the illumination conditions are altered significantly (e.g., focused-TIRF vs. radial).

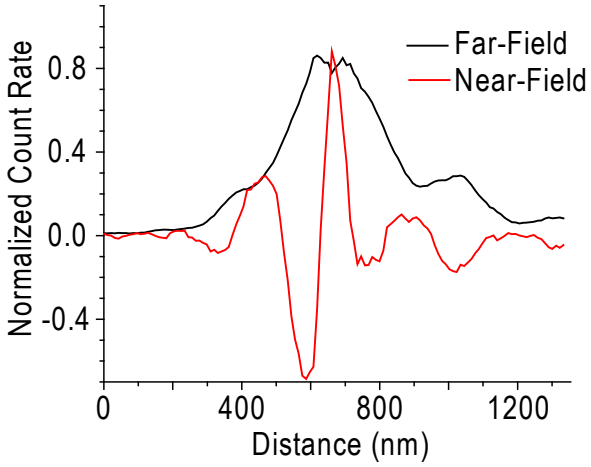

Figure 5.7. A comparison of cross sections over the fluorescent bead taken from images found in Figure 5.6 corresponding to panel (a) far-field and panel (c) near-field. The patterns shown indicate that the interference pattern from the tip (red) is distinct from the Airy pattern visible in the far-field (black).

5.5 Summary

In summary, we presented a single-photon counting technique for generating three-dimensional tomographical reconstructions of tip-sample interactions in tip-enhanced fluorescence microscopy. Our technique has a number of advantages compared to previous near-field tomographical measurements [80, 81, 31]. Most importantly, all three spatial dimensions of the fluorescence map are acquired simultaneously in a single measurement along with standard AFM channels (topography, oscillation amplitude, and phase). Thus, fluctuations in the fluorescence signal unrelated to tip-sample coupling can be easily detected and removed. Furthermore, uncertainties in the tip position relative to the sample are largely eliminated. The versatility of the technique was demonstrated by revealing a number of interesting three-dimensional features when imaging 20-nm diameter fluorescent spheres. In particular, the extremely high sensitivity and precision of this technique enabled the first measurement of the three-dimensional image pattern resulting from the complex interplay between field enhancement, fluorescence quenching, and tip-induced optical interference.

The ability of this technique to tomographically map the full three-dimensional dependence of the tip-sample interactions makes it quite powerful for developing and testing near-field models and we anticipate its extensive use in the future to aid in the study of a variety of nanophotonic phenomena and in the development of materials with increased functionality. It may be particularly useful in the design and development of novel nano-antenna geometries based on plasmon resonances, a topic which has gained considerable attention in recent years. Furthermore, this technique can be applied to study the near-field interactions between any two particles, provided one of them can be attached to an AFM tip. For example, it is now being used to study energy transfer between various fluorophore species (quantum dots, organic polymers, etc.) and carbon nanotubes attached to the tip, with the ultimate goal of designing nanotube-based photovoltaic materials. Finally, the technique can be used to improve image contrast in tip-enhanced fluorescence microscopy, as demonstrated by applying a number of phase-sensitive

analysis algorithms.

5.6 Simulations

While the interference patterns reported here, have been reported previously, the literature does not entirely agree on the origin of this phenomenon. In order to more fully prove that the source of these patterns stems from interference in the way we claim, some numerical modeling using COMSOL, has been done. As reported above, we hypothesized that the interference pattern results from the direct excitation laser incident on the sample interfering with excitation light scattering from the tip. This is supported by several, but not all previously published reports on the matter. By performing a full electrodynamic simulation, we have verified that this hypothesis was correct. The results are seen below in Figure 5.8. In this image a two-dimensional calculation was performed such that the results are only qualitative. The model includes a Si tip with realistic dimensions (10 nm radius of curvature and a 4:1 aspect ratio) placed 10 nm above an air-glass interface. An optical field was incident from below at $\theta_i = 61°$, well above the critical angle ($\theta_c = 41°$), and typical of rays stemming from the particular objective and illumination conditions used in the experiments reported above. The wavelength of the simulation was 532 nm. As can be seen in Figure 5.8, the regions of higher field (bright) are similar to those observed in the experiment.

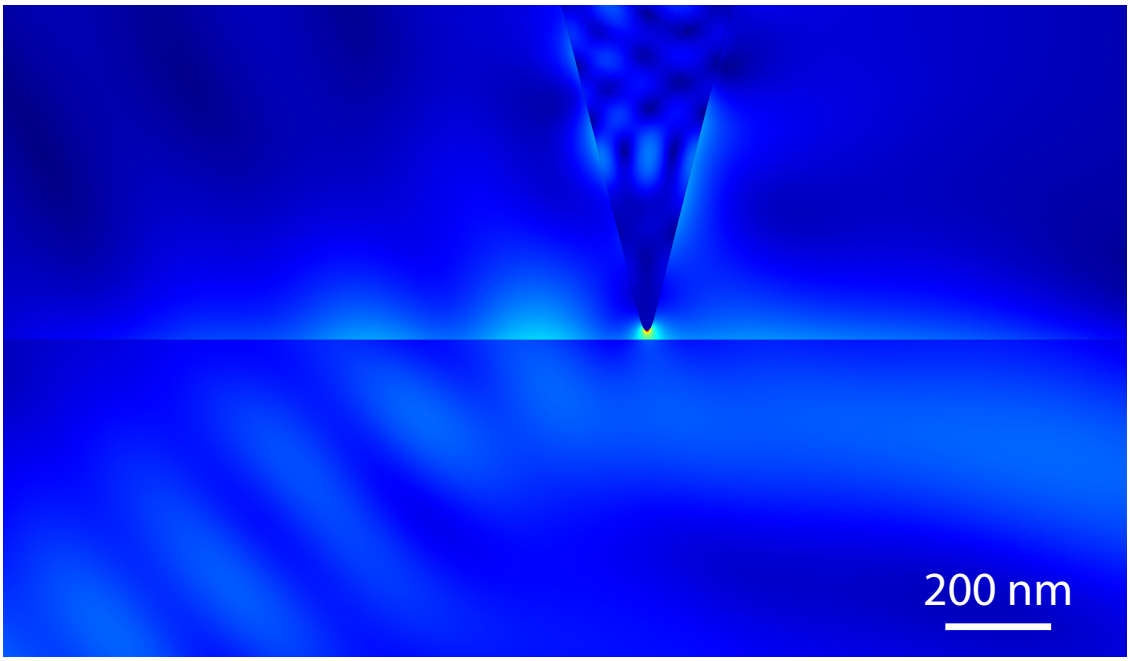

Figure 5.8. A wedge illumination simulation. A Si tip with realistic dimensions (10 nm radius of curvature and a 4:1 aspect ratio) was placed 10 nm above an air-glass interface. A field was incident from below (left to right) at $\theta_i = 61°$, well above the critical angle ($\theta_c = 41°$), and typical of rays stemming from the particular objective and illumination conditions used in the experiments reported above. The wavelength of the simulation was 532 nm. Regions of high field are represented by brighter colors.

CHAPTER 6

AQUEOUS IMAGING

Paramount in using TEFM to image biological specimens is the ability to operate TEFM in aqueous environments, particularly buffer solutions that mimic physiological conditions. Currently, TEFM can provide adequate resolution to begin to probe serious biological questions, achieving sub 10 nm resolution [3], and thus should hopefully allow for exploration of the realm of single proteins. The original goal of TEFM was to determine the structure of complex biological networks by identifying the relative position and role of proteins within these complicated systems. With willing collaborators in biology, we have had very specific aims: to elucidate the relative positions of protein complexes responsible for the formation of multivesicular bodies (MVBs).

The lab of Markus Babst in the biology department here at the University of Utah researches a system of protein complexes responsible for the formation of MVBs that seemed well suited to analysis with our TEFM system. An important part of the regular function of a eukaryotic cell is to recycle transmembrane proteins that are nonfunctional. These proteins are removed from the membrane via endocytosis where they are trafficked to lysosomes to be recycled. This process is reliant on the proper function of MVBs. The Babst lab is interested in understanding the molecular mechanisms of MVB formation, in part because certain types of viruses, such as HIV, exploit the MVB pathway for entry into a cell. Several protein complexes known as ESCRT (Endosomal Sorting Complex Required for Transport) are responsible for tagging and sorting transmembrane proteins for degradation and regulate MVB formation.

Since this system involves transmembrane proteins, key elements of the ESCRT protein complexes could be purified and reconstituted on lipid bilayers in vitro. Such a scheme is ideal for imaging with TEFM. The ultimate goal of this project was to be able to selectively label key elements within the complex with various colors of fluorescent proteins. An artist's rendition of what we hoped to discover is shown in Figure 6.1.

This goal has not yet been realized, due to the fact that tip-enhancement has not been achieved in aqueous conditions. Many efforts were made to image fluorescent test samples with TEFM; however, none have been successful. This has lead to an in-depth investigation as to the physical limits of TEFM, especially more fully

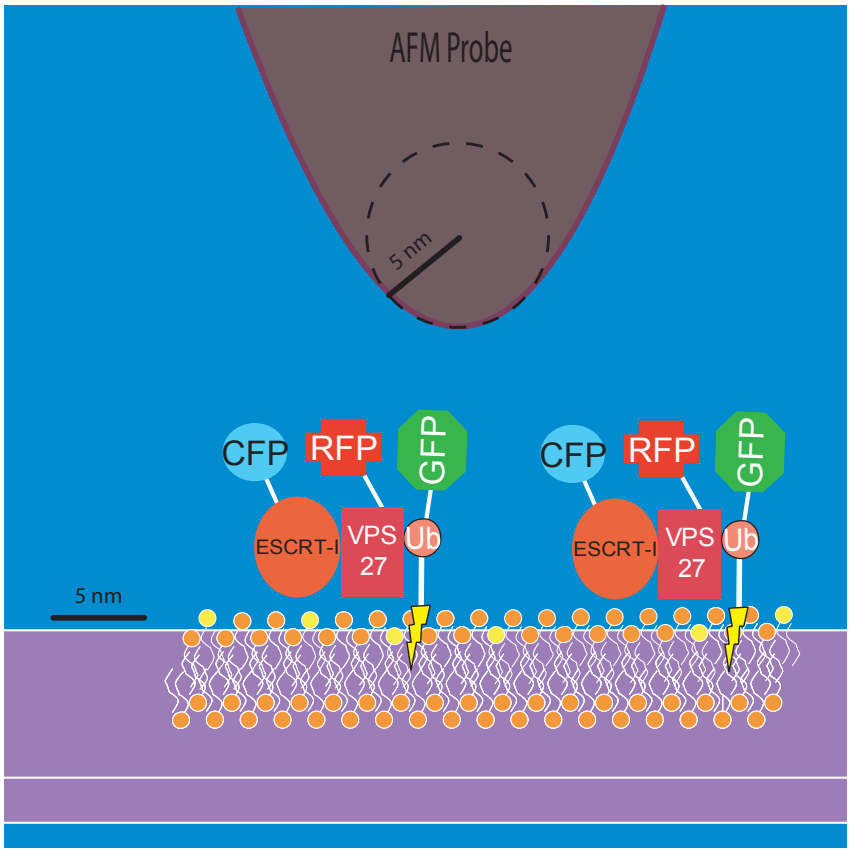

Figure 6.1. Cartoon diagram of imaging ESCRT protein complexes with TEFM. Several various elements are fluorescently labeled with different colored fluorescent proteins (FP): cyan (CFP), green (GFP), and yellow (YFP). The near-field probe would have hopefully been able to elucidate the relative positions of various proteins within the complex.

understanding the role of various illumination schemes and how they influence the amount of expected field enhancement. Furthermore, the role of oxide layers on silicon tips has also been investigated. This chapter will also discuss the problems associated with imaging in water and several variations of TEFM imaging that have also been attempted: embedding biological samples in plastics, and far-field fluorescence correlation with AFM topography.

6.1 Sample Preparation

One considerable technical challenge of imaging in a fluid environment is the problem of keeping samples attached to the surface. When imaging in a dry environment, samples can be simply dried down on a surface, where due to electrostatic interactions, they remain firmly in place. Typical test samples include dye-doped latex beads, nominally 20 nm in size, as well as quantum dots. Both are commercially available and come in a variety of moieties. Fluorescent beads of this size contain several hundred dye molecules and can thus be quite bright and stable; however, due to their extended size, they show less net fluorescence enhancement due to reduced intensity-volume overlap. On the other hand, quantum dots are favorable as they act as single emitters. Quantum dots can be much smaller and thus exhibit better intensity-volume overlap (IVO), resulting in higher net signal enhancement factors. However, they are not as bright as beads and also suffer from blinking. For working in aqueous solutions, quantum dots must first be solubilized. This process includes adding several layers of ligand groups to the quantum dot, making it essentially the same size as latex beads, thus negating any intensity-volume overlap benefits. In this regard, to avoid blinking events, 20-nm dye-doped latex beads have been used almost exclusively as test samples for water.

There has been a considerable amount of effort devoted to robustly attaching bead samples to glass surfaces to allow for fluorescence imaging in a fluid environment. Common protocols include biotin-streptavidin conjugation techniques, as well as amine-carboxyl chemistries. In recruiting outside help to tackle this issue, our collaborators in biology and chemistry alike insist that this is an easily solvable

problem, as they routinely attach fluorescent samples to glass surfaces. This has not proved quite so easy in that we have the added complexity of needing to do AFM imaging.

Several reoccurring problems plague sample preparation issues: samples becoming detached from the surface, samples becoming attached to the tip, and samples bleaching when the tip scans over them. When performing only far-field fluorescence imaging (in the absence of the tip), the bead samples remain bright and well attached for extended periods of time. As the AFM probe is brought in close proximity to the glass surface, beads may easily become detached. This is thought to arise from the sonicator-like action of the tip. One standard method for cleaning glass is by sonication, where the sample is put in water that is vibrated at high frequencies. These high frequency pressure waves cause small particles to become released from the surface in a process called cavitation as micro bubbles expand and collapse near any surface present. The operating frequency of commercial bath sonicators is several tens of kHz, precisely the operating frequency of some AFM probes in water. Samples becoming detached from the surface has largely been overcome through refinement of attachment protocols (see Appendix E) in conjunction with careful choice of tip-oscillation frequencies (lower being better).

Another related problem is that of beads becoming attached to AFM probes. One standard attachment protocol recommended to us by biologists is that of biotin-streptavidin conjugation. Biotin and streptavidin are biomolecules with a very high affinity for one another. Both quantum dots and latex beads can be purchased with biotin coatings. By coating the glass coverslip with streptavidin, biotin coated fluorophores can be readily attached. Further treatment of the surface with biotin can block any further unwanted reactions. This has been attempted in our lab with poor results, the frequency of bead attachment being quite high. It is hypothesized that as the tip scans over the sample it can build up layers of biotin and streptavidin, which when subsequently contacting a bead can cause it to be readily attached to the tip. This bead pick-up is easily observed in Figure 6.2. As the AFM probe contacts the sample the background fluorescence count rate

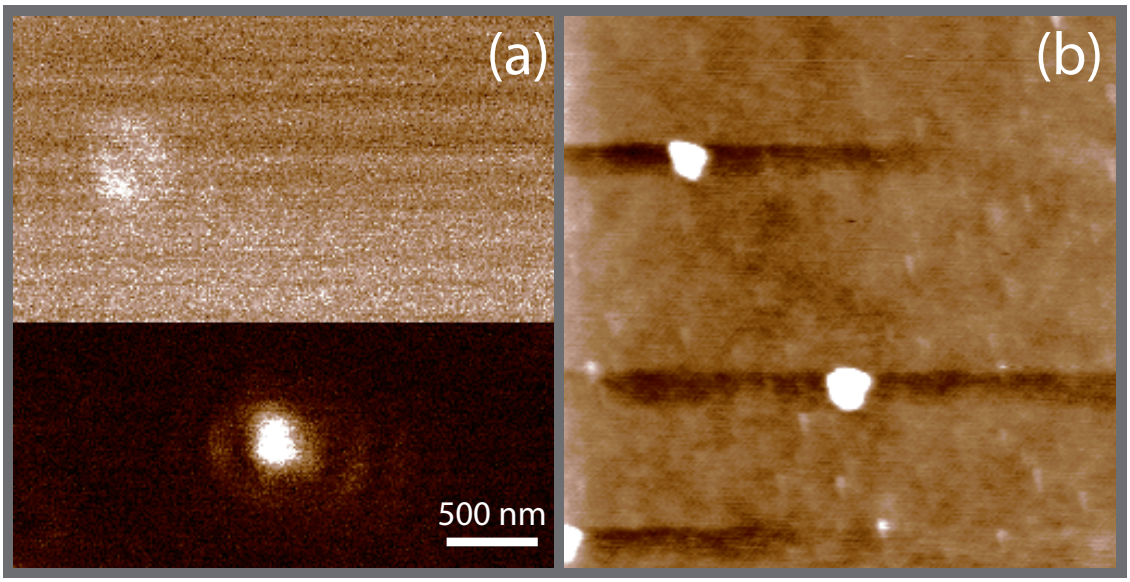

Figure 6.2. An example of a fluorescent sample becoming attached to the AFM tip in water. In this image, a silicon tip is scanning over a sample of fluorescent latex beads attached to a glass coverslip. The scan direction is from left-to-right and from bottom-to-top. Panel (a) shows the fluorescence signal, while panel (b) shows the simultaneously recorded height trace. The tip and excitation laser are misaligned. As the tip passes over the lower bead, it becomes attached to the AFM tip; as evidenced by the sudden increase in the background fluorescence count rate.

immediately increases, indicating the sample has become attached to the tip, which is maintained in alignment with the excitation beam. We have found that this is mitigated to some extent using amine-carboxy chemistries for bead attachment to glass coverslips.

Finally, one last persistent problem is that vven when fluorescent beads remain firmly attached to glass, when scanned over by the AFM tip they often become dark. As the bead enters the far-field illumination spot the fluorescent signal is recorded and the outer portions of the focal spot pattern begin to be observable. As the tip scans over the bead, as indicated by the height trace (Fig. 6.2), the fluorescence vanishes. Subsequent AFM scans confirm that the sample is in place, but permanently dark. Oftentimes, no topological change to the sample can be seen, but for a still unknown reason, some samples become irreversibly damaged in this way. This is not necessarily categorized as photo-bleaching either, as AFM scans

done in the absence of an excitation source can lead to the same result. We thus hypothesize that some sort of mechanical damage to the sample may be occurring, but this has not been rigorously confirmed. Small tip oscillations amplitudes can alleviate this problem to some extent but not completely remove it. Furthermore, small oscillation amplitudes are also associated with higher occurrence of bead samples becoming attached to the tip.

6.2 Early Efforts

These challenges in preparing a suitable sample aside, we naively expected that achieving tip-enhancement in water would not be significantly different from air. To our surprise, no enhancement has been observable. Figure 6.3 shows several attempts at obtaining signal enhancement from a Si tip. We verified that enhancement is indeed attainable with a similar setup in air. Panel (a) demonstrates that under similar illumination conditions in air, with similar tips, good enhancement can be obtained. Silicon tips were freshly etched in a buffered hydrofluoric acid solution (BOE) and used to image dye-doped latex beads on glass coverslips. Fluorescent count rates vs. tip-oscillation phase is plotted in a histogram as a tip oscillates on top of a bead. Panels (b)-(f), were obtained one week later, with a tip from the same batch as (a), which was also freshly etched. No clear enhancement is observable in these panels, nor has it ever been, in any measurements performed in water (or saline buffer solution) for any tip-illumination condition.

After several early attempts with no apparent tip-enhancement, I spent some time trying to understand the source of the absence of field enhancement. No one reason jumps out as the culprit, but a combination of several effects acting in concert could be responsible for the lack of field enhancement. One of these factors stems from permittivity considerations; the mismatch between the tip and the surrounding media is smaller in water compared to air, leading to a reduced lightning-rod effect. Other contributing factors include the formation of oxide layers on the tip, as well as reduced efficiency of creating axial fields in water.

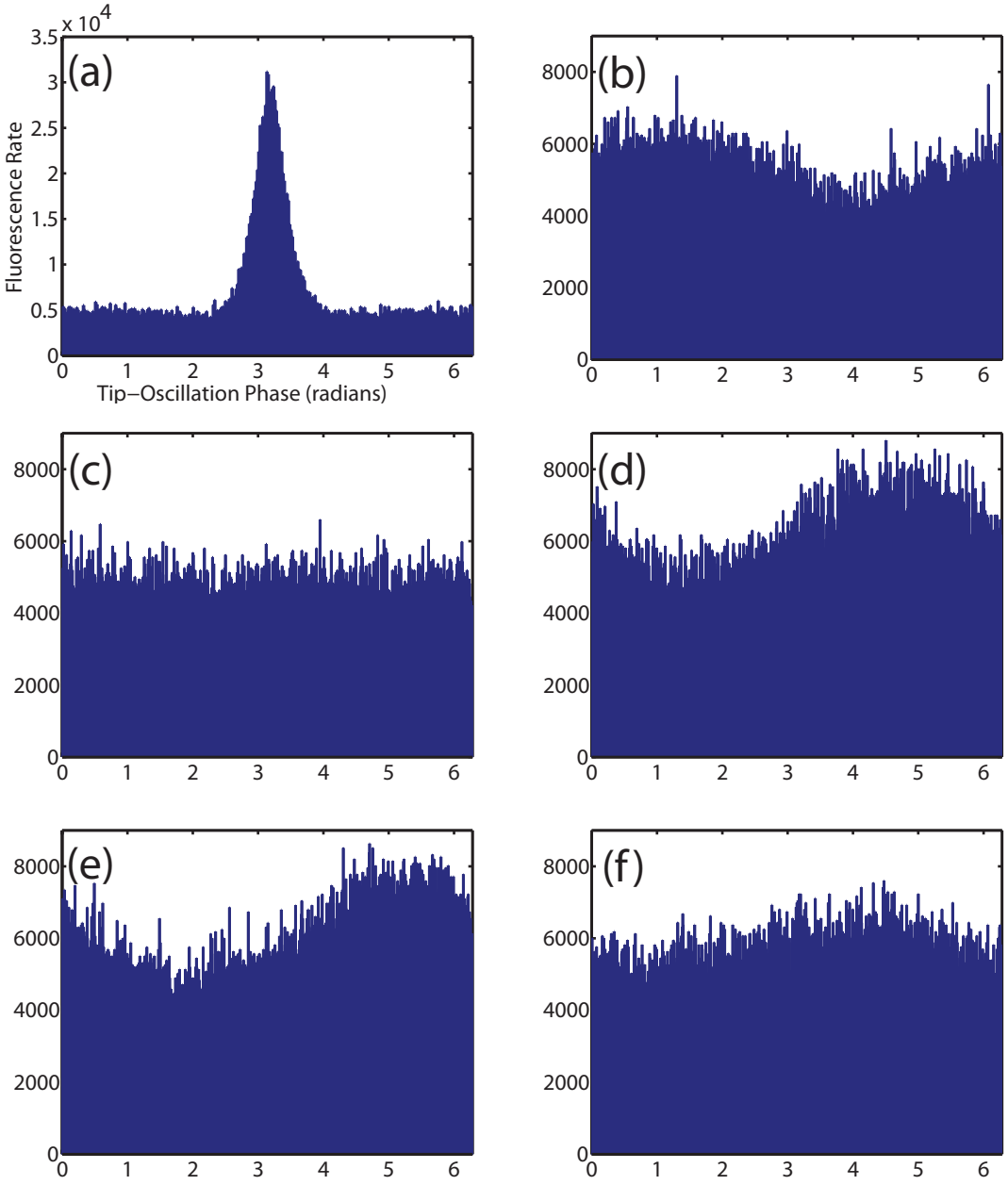

Figure 6.3. Histograms of TEFM in water. Panel (a) shows a histogram of a freshly etched Si tip on a latex bead in air. Panels (b)-(f) show histograms of a freshly etched Si tip on different latex beads in water. All panels were taken with vertical wedge illumination at nominally the same laser power (actual intensities may vary somewhat as the wedge profile is different in water as compared to air). Panels (b)-(f) have the same tip-oscillation amplitude; the amplitude of oscillation in (a) is 10% larger. Sinusoidal patterns in some histograms are due to interference effects, larger amplitudes being an indicator of better alignment.

6.3 Permittivity Considerations

One major concern of imaging in water is that of relative permittivity. As we have only had the capacity to perform numeric analyses recently, a good first order approximation is to check what might be expected from using the analytic model for a sphere. Repeating the solution for Equation 2.1:

$$\vec{E_{out}}(r,\theta) = \vec{E_0} \cdot \cos\theta \left[1 + 2\frac{\epsilon_r - 1}{\epsilon_r + 2}\frac{R^3}{r^3}\right]\hat{r} + \vec{E_0} \cdot \sin\theta \left[-1 + \frac{\epsilon_r - 1}{\epsilon_r + 2}\frac{R^3}{r^3}\right]\hat{\theta}, \quad (6.1)$$

It can be seen that the enhancement of a sphere is set by the relative permittivity of the sphere compared to the surrounding media $\epsilon_r = \epsilon_s/\epsilon_{media}$, thus as the permittivity of the media is increased, the enhancement is expected to go down. Inserting the permittivity of water at optical wavelengths $\epsilon'_{water} = n^2_{water} = 1.33^2 = 1.77$, we expect that the peak intensity enhancement from a sphere should decrease by a factor of $\sim 15\%$. Of course we are also interested in how the intensity-volume overlap might change. The IVO was calculated for 20 nm spherical regions adjacent to Si spheres of different radii. As seen in Table 6.1, the actual drop in the anticipated enhancement factor for latex beads is actually much less than 15%, especially for sharp probes.

Again, keep in mind that the permittivity is complex valued. Presumably screening of ions in the water can be responsible for reduced local fields at the tip apex. Typical indices of refraction for water in the visible spectrum are well known, as well as absorption coefficients. Absorption coefficients are directly related

Table 6.1. Intensity-Volume overlap integrals for spheres of varying sizes from a 3D electrostatic calculation using COMSOL . The intensity-volume overlap was calculated in the region of 20-nm spheres directly below Si spheres of various radii placed in an otherwise uniform electric field. This table plots the intensity-volume overlap for each size sphere normalized to the corresponding result performed in air.

Radius	Overlap Water/Air
40	0.89
30	0.90
20	0.92
10	0.95

to the imaginary part of the permittivity in a straightforward way:

$$k = \frac{\lambda_0 \alpha_{abs}}{4\pi} \Rightarrow \epsilon'' = 2nk = \frac{n\lambda_0 \alpha_{abs}}{2\pi}, \tag{6.2}$$

where k is the imaginary part of the index of refraction, n is the real part, and λ_0 is the free space wavelength. Water absorbs strongly at microwave frequencies, which is how food heats up in a microwave oven. At visible frequencies, the absorption of water is very small, $\sim 5 \times 10^{-3}$ cm^{-1} [82, 83]. Equating this to the imaginary part of the index of refraction we get: $\epsilon'' = 3 \times 10^{-5}$.

Arguably, this permittivity value is for ultrapure water, and different results may be expected when imaging in biologically relevant buffers. Unfortunately, data for absorption coefficients of common buffer solutions is difficult to obtain, as there is no real standard for which type of salts should be included and at which concentration. However, some data are available for seawater [84], which shows only slight variations in absorption coefficient and thus ϵ''. With such a small imaginary part of the permittivity, even in saline buffer solutions, screening from any ionic effects is minimal. Furthermore, any extra nonradiative decay due to ϵ'' of the water is also small, as we know fluorescence samples are routinely imaged in aqueous solutions for extended periods of time under far-field illumination.

6.4 Oxide Layers

The results of these models show that a reduction in relative permittivity coupled with ionic screening effects are not able to account for the complete lack in observable enhancement. I next hypothesized that a drastic reduction in near-field signal might be due to effects of oxide layers forming on the tip. It was previously known to us that as silicon tips oxidize, the net signal enhancement is decreased. Just as the dielectric sphere in a uniform field has an analytic solution (Appendix A), so does a dielectric sphere surrounded by a shell of another dielectric material; the solution is provided in Appendix B. To first estimate the effects of oxide layer formation, the peak enhancement of Si spheres with a layer of SiO_2 can be calculated for both air and water. These peak values indicated that the peak enhancement

values are reduced by fairly modest amounts with ever thickening oxide layers as seen in Figure 6.4.

Of greater interest than the peak enhancement factors, it is valuable to understand how the IVO of enhanced fields changes with increasing oxide formation on Si tips. The IVO values as a function of oxide thickness, IVO(t), are plotted for Si spheres in both air and water, Figure 6.5. The IVO(t) values in the left panel demonstrate the relative effect of oxide layers for a given medium, as IVO values are normalized to their peak values found in the corresponding medium: $\text{IVO}(t)_{air}/\text{IVO}(t=0)_{air}$, and $\text{IVO}(t)_{water}/\text{IVO}(t=0)_{water}$. The right panel compares the relative effect of the oxide values in water to the results otherwise obtainable in air by plotting the following values: $\text{IVO}(t)_{air}/\text{IVO}(t=0)_{air}$,

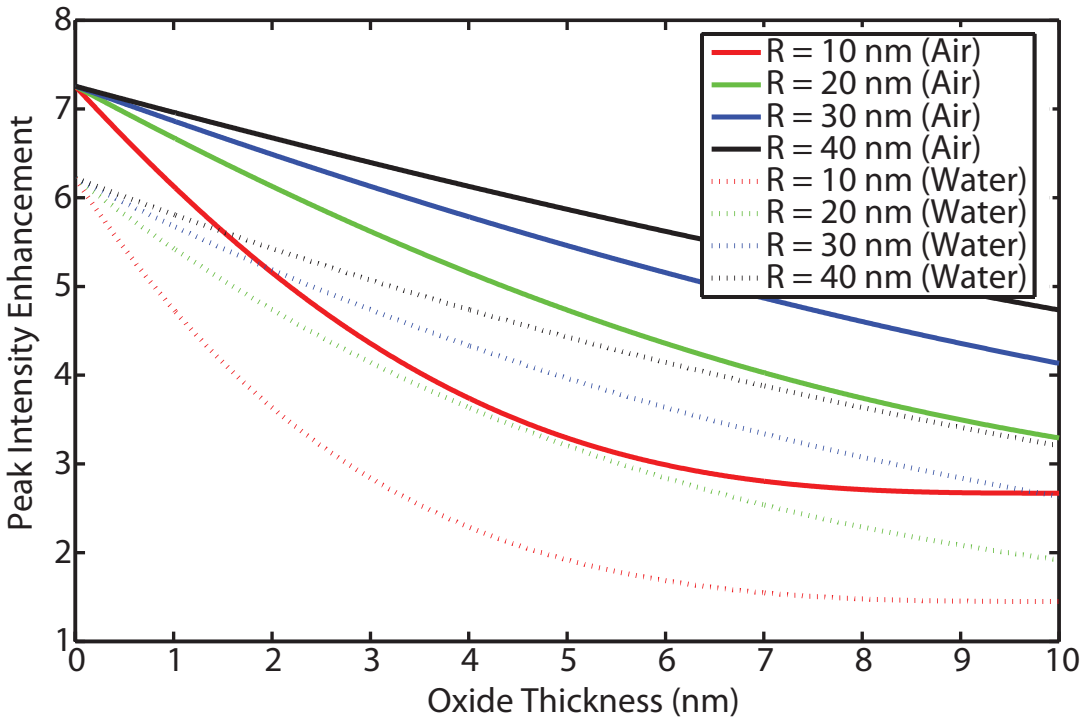

Figure 6.4. Analytic electrostatic peak intensity enhancement values computed for Si spheres surrounded by both air and water with varying oxide thickness. Permittivities used correspond to excitation at 543 nm and are as follows: $\epsilon_{Si} = 17.6 + 0.12i, \epsilon_{air} = 1.0, \epsilon_{water} = 1.78, \epsilon_{SiO_2} = 2.39$.

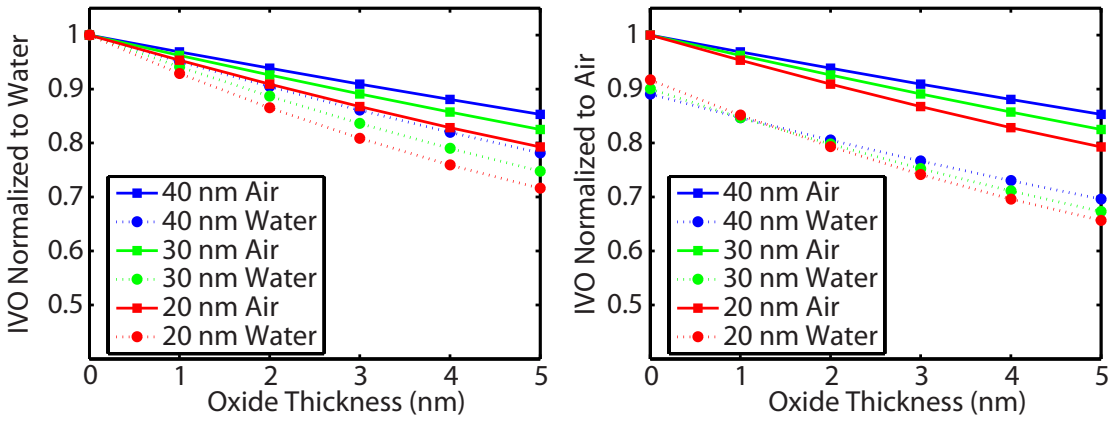

Figure 6.5. IVO values for Si spheres surrounded by both air and water with varying oxide thickness. Permittivities used correspond to excitation at 543 nm and are as follows: $\epsilon_{Si} = 17.6 + 0.12i, \epsilon_{air} = 1.0, \epsilon_{water} = 1.78, \epsilon_{SiO_2} = 2.39$. Diameter of IVO integration is 20 nm.

$\text{IVO}(t)_{water}/\text{IVO}(t=0)_{air}$.

A key to more fully understanding the effects of oxide layer formation is to use more realistic tip geometries. Using again the same tip geometry as before, especially the same outer measurements of the tip, an interior layer of SiO_2 of variable thickness was added to the simulation. Figure 6.6 illustrates the geometry of the calculation while Figure 6.7 shows the numeric results. Figure 6.6 shows the solution for a 3 nm thick oxide layer having formed along the boundaries of the tip, the IVO was calculated for a sphere of 20 nm in diameter located directly beneath the tip as indicated by the dark circle. The effects of the oxide layer can be immediately seen; at very least it acts as a buffer zone, limiting the closest approach of the tip to the sample. The oxide layer has a permittivity of $\epsilon' = 2.39$, a value between that of Si and water. This intermediate step in permittivity lowers the effective mismatch between the tip and water, further reducing the magnitude of the lightning rod effect.

As seen in Figure 6.7, the results of calculating the IVO values was also very striking. The values here are normalized to the IVO obtained with the tip in air and no oxide layer present; again, the total volume of the tip + oxide layer is

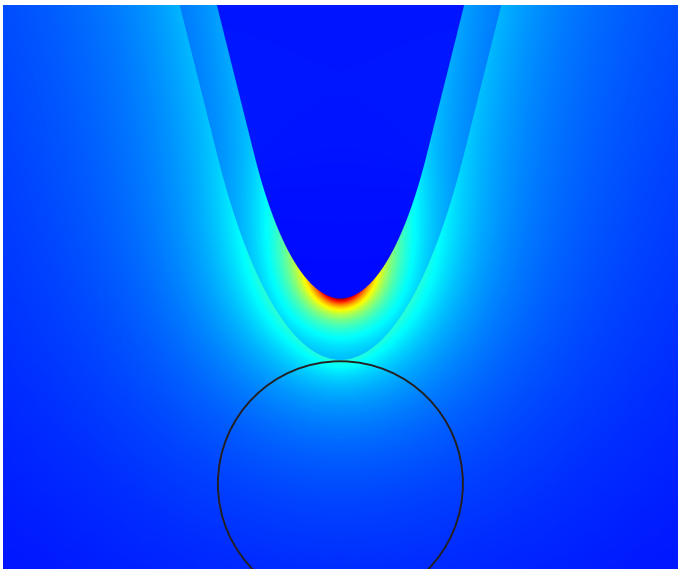

Figure 6.6. 3D electrostatic COMSOL simulation of an Si tip with a 3 nm layer of SiO_2. IVO values were calculated in a 20 nm diameter spherical region directly below the tip for varying oxide thicknesses. Brighter (more red) regions indicate higher intensity values.

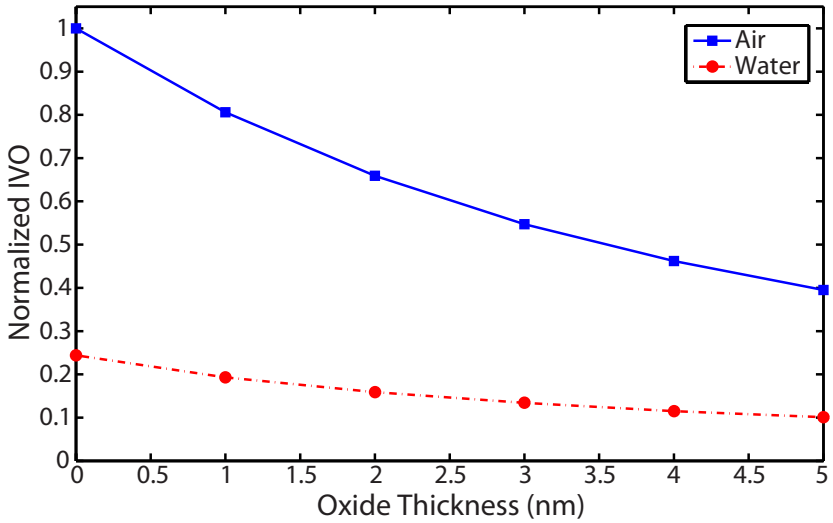

Figure 6.7. An electrostatic COMSOL calculation of IVO values calculated for a 20 nm diameter sphere located beneath a Si tip with varying oxide thickness. The calculations were performed with the tip in air and water. IVO values are normalized to that of an SI tip with no oxide layer in air. Permittivities used correspond to excitation at 543 nm and are as follows: $\epsilon_{Si} = 17.6 + 0.12i, \epsilon_{air} = 1.0, \epsilon_{water} = 1.78, \epsilon_{SiO_2} = 2.39$.

kept constant. From permittivity considerations alone, with no oxide layer, the lightning rod-effect in water was much lower than we anticipated from our previous calculations with spheres. Adding oxide layers only compounds the problem even more; however, the effects of the oxide layer in air are more significant.

6.5 Diminished Evanescent Fields

One major concern in imaging in water is finding the appropriate illumination conditions. As discussed in Chapter 3, wedge illumination is often favorable, in that it allows for relatively easy alignment between tip and laser. For imaging in aqueous environments, the critical angle for total internal reflection is increased from $\theta_c = 41^\circ$ in air to $\theta_c = 62^\circ$ for water. This means that using the same microscope objective a prohibitively small amount of power may be transmitted (0.2%, see Chapter 3) when using the appropriately sized beam mask. Thus, for imaging in water, it may seem preferable to use a radially polarized illumination beam.

Another factor further reducing any expected enhancement is the intensity and polarization state of the evanescent fields in the water. In air, p-polarized beams can have evanescent intensities that are axially polarized and enhanced by a factor of nine compared to the incoming beam. As shown in Figure 6.8, when working in water, the magnitude of the axial field components is not so large, due to the increase in the critical angle. The intensities of these axial components are reduced considerably—to a little over five—when working in water. When using a wedge illumination scheme, by design all rays lead to evanescent fields. When illuminating with radial polarization; in air, a significant portion of the radially polarized beam leads to axially polarized evanescent fields, where the resulting axial components are much stronger compared to their propagating counterparts. In fact, the majority of the axial field components created via radial illumination stem from the totally internally reflected components of the beam. When illuminating with radial polarization in water, as has been attempted many times, the total field enhancement factor is expected to be reduced as there is a diminished amount of

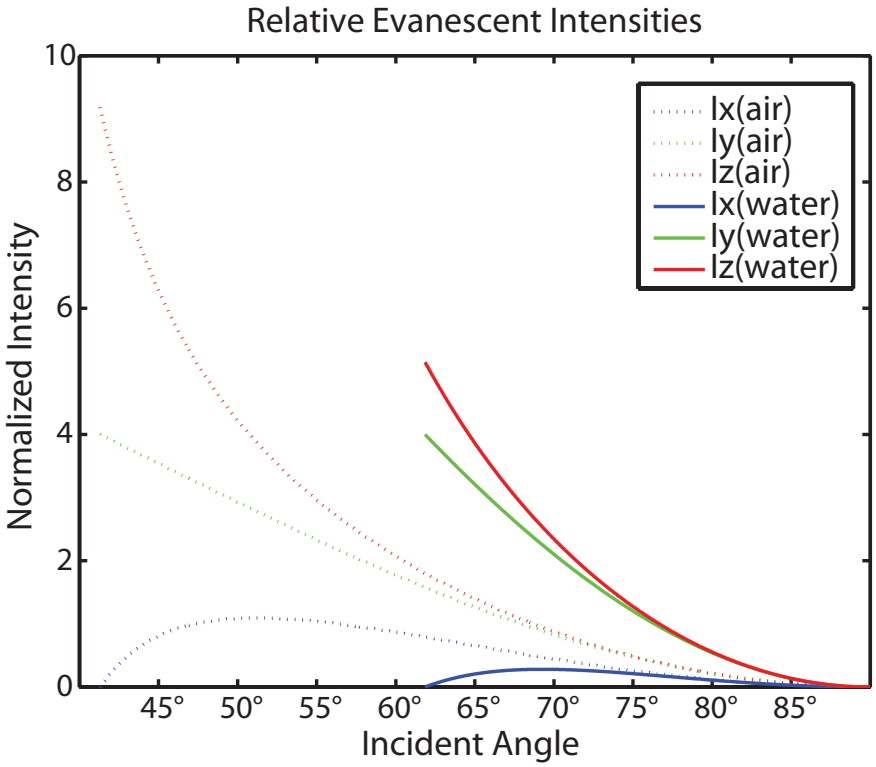

Figure 6.8. Axial components of evanescent intensities plotted as a function of incident angle for both air and water. Calculations are from Eq. 3.5.

supercritical rays.

For example, a subcritical p-polarized ray incident at a given angle would have a normalized axial component of its intensity given by the Fresnel transmission coefficient multiplied by the vertical component of the field, which can be obtained via Snell's law. Given a ray at a fixed angle of incidence (θ), the final amount of axial intensity per unit of p-polarized incident intensity can be readily calculated:

$$\begin{aligned} \frac{I_z}{I_0} &= (\text{Fresnel Transmission Coefficient}) \times \text{Axial Component}_{Snell} \\ &= \left(1 - \left(\frac{n_m^2\cos(\theta) - n\sqrt{n_m^2 - n^2\sin(\theta)^2}}{n_m^2\cos(\theta) + n\sqrt{n_m^2 - n^2\sin(\theta)^2}}\right)^2\right) \times \frac{n\sin\theta}{n_m}, \end{aligned} \tag{6.3}$$

where n_m is the index of refraction of the media of the sample, and n is index of refraction of the media of the objective/coverslip. Correspondingly, the axial component of a super-critical ray has already been given in Equation 3.5:

$$I_z = I_0 \frac{4\cos^2\theta\sin^2\theta}{(1-n_m/n^2)\left[(1+n_m/n^2)\sin^2\theta - n_m/n^2\right]} \tag{6.4}$$

Now the total strength of axial field components stemming from both subcritical and super-critical rays can be found for a given intensity profile. In particular, radial polarization is the easiest TEM mode to solve for because of its symmetry: all rays are p-polarized.

$$2\pi \int_0^{\theta_c} TEM_{m,n}(\theta)\,(\text{Fresnel Coefficient})\,(\theta)\,\text{Axial Component}_{Snell}(\theta)\,d\theta + 2\pi \int_{\theta_c}^{\theta_{BA}} TEM_{m,n}(\theta)\,I_z(\theta)\,d\theta \tag{6.5}$$

Here, the integrals correspond to subcritical and super-critical components with the appropriate limits of integration to account for all rays the objective can produce. The factors $TEM_{m,n}$ account for the intensity distribution of the incoming beam at the back aperture of the objective. One nuance then is ensuring that the radial intensity profile has been mapped appropriately to the correct incident angle. This mapping is done through a rearrangement of Equations 3.6 and 3.8:

$$r = 2n\sin(\theta). \tag{6.6}$$

As a reminder, the critical angle, θ_c, and maximum angle corresponding to the back aperture of the objective lens, θ_{BA}, are given by:

$$\theta_c = \sin^{-1}\left(\frac{n_m}{n}\right) \tag{6.7}$$

$$\theta_{BA} = \sin^{-1}\left(\frac{D_{BA}\cdot Mag}{2f_{TL}\cdot n}\right). \tag{6.8}$$

We can now solve for the amount of axial field expected given our typical experimental setup for imaging in both air and water. Using the values for a $100\times$ $NA = 1.4$ oil immersion objective and assuming a perfect radially polarized beam whose $1/e^2$ width is equal to the diameter of the back aperture, we find the amount of axial field is given by:

$$2\pi \int_0^{\sin^{-1}(n_m/n)} \left[H_1\left(\frac{\sqrt{2}\cdot 2n\sin(\theta)}{2.8/\sqrt{2}} \right) exp\left(\frac{-(2n\sin(\theta))^2}{(2.8/\sqrt{2})^2} \right) \right]^2 \times \left[1 - \left(\frac{n_m^2\cos(\theta) - n\sqrt{n_m^2 - n^2\sin(\theta)^2}}{n_m^2\cos(\theta) + n\sqrt{n_m^2 - n^2\sin(\theta)^2}} \right)^2 \right] \times [n\sin(\theta)]\, d\theta +$$
$$2\pi \int_{\sin^{-1}(n_m/n)}^{\sin^{-1}\left(\frac{D_{BA}\cdot Mag}{2f_{TL}\cdot n}\right)} \left[H_1\left(\frac{\sqrt{2}\cdot 2n\sin(\theta)}{2.8/\sqrt{2}} \right) exp\left(\frac{-(2n\sin(\theta))^2}{(2.8/\sqrt{2})^2} \right) \right]^2 \times \left[\frac{4\cos^2\theta\sin^2\theta}{(1-(n_m/n)^2)\left[(1+(n_m/n)^2)\sin^2\theta - (n_m/n)^2\right]} \right] d\theta, \qquad (6.9)$$

where n is the index of refraction of the glass $n = 1.515$, and n_m is the index of refraction of the media n_{air}, or n_{water}. The numeric result is that imaging in water will yield only 42% the axial field as would be found in air. To be sure this is not an exact solution, here we assumed a ray diagram-like treatment, assuming all rays converge at the focal point fully negating any diffraction effects. Also implicit in this calculation is that tip was on the surface, i.e., that the evanescent intensities were maximum, and that all evanescent fields did not decay as they propagated towards the center of the focal spot—a reasonable assumption given the length of the Goos-Hänchen shift [9] in this configuration.[4] Nevertheless, this should give a reasonable approximation. Combining this result with that of the COMSOL simulations just discussed, we arrive at this conclusion: for a Si tip in water, even without any oxide layers we only expect only $0.42 \cdot 0.24 = 0.10$ the enhancement we would otherwise expect in air when imaging 20-nm fluorescent beads.

These results give further support to the notion that evanescent illumination can lead to large axial field components, and thus should be the preferred choice of excitation wherever possible. For water, while the net power may be low, this gives the highest possible enhancement. Some effort was directed to this end, by making beam masks shown in Figure 3.10. Still, no signal enhancement was observed.

[4]The Goos-Hänchen shift is dependent on the incident angle as well as index of refraction of the media. For water the smallest shift is 770 nm occurring at the maximum angle of the objective. For air, the smallest shift is 155 nm.

6.6 Success in Other Groups

Despite the difficulties associated with using TEFM in aqueous environments, there has been two published reports of achieving some success to this end. Using a rather different experimental setup than we employ, Höppener and co-workers have imaged fluorescently labeled calcium channels in an aqueous environment [85] as seen in Figure 6.9. In this report no indication of actual resolution nor enhancement factors were reported for imaging in aqueous environments. However, the contrast and resolution was sufficient to determine that the sample of interest has an average protein-protein spacing of 90 nm, a result that is completely unattainable with traditional far-field techniques. To the best of my knowledge, no other reports of imaging in liquids have been made.

There are several major differences between our imaging system and the one Höppener used. First, this experiment was performed using probes made of 80-nm gold nanospheres attached to dielectric tips. The resulting resolution of course was limited by this large tip size. Second, the system used was not a cantilever based

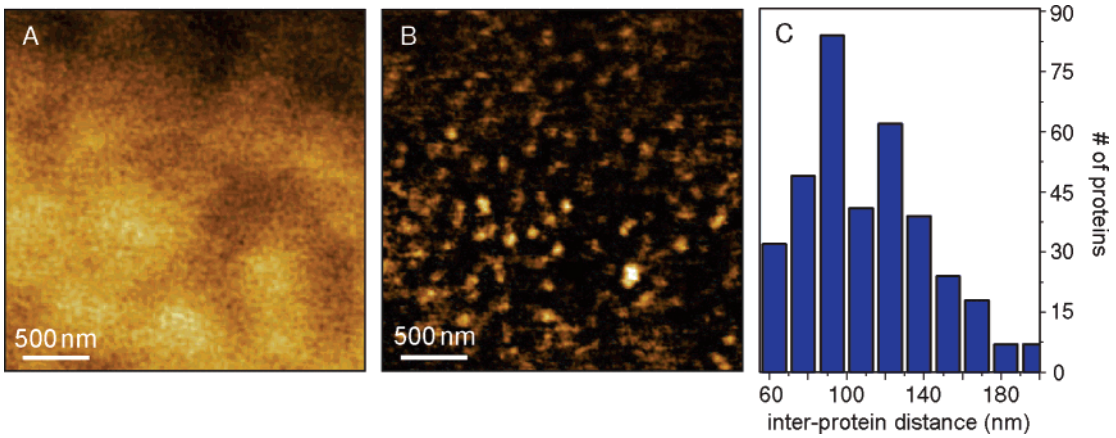

Figure 6.9. TEFM imaging of Calcium channels. Plasma membrane-bound Ca2+ ATPases (PMCAs) were labeled with Alexa 633 via immunolabeling techniques. (A) Confocal fluorescence image of an erythrocyte plasma membrane immersed in water. (B) Corresponding near-field fluorescence image showing individually resolved PMCA4 proteins. The image was acquired using an 80 nm gold nanoparticle antenna. The confocal background has been subtracted using high-pass filtering. (C) Distribution of nearest-neighbor interprotein distances, revealing an average protein-protein separation of 90 nm. Reprinted with permission from [85].

one, but rather relies on a shear-force feedback mechanism. Lastly, the sample images were dye molecules attached to calcium channels.

These differences might be key to understanding why they were able to accomplish this goal. Metal nanosphere probes can have large enhancement factors, as they can support localized surface plasmon resonances as well as polarization effects. Looking at Equation 2.2 again we find that for metal spheres, as the relative permittivity ϵ_r approaches $\epsilon_r \to -2$ that a resonance condition is met.

$$E_{max} = E_0 \left(1 + 2\frac{\epsilon_r - 1}{\epsilon_r + 2}\right), \tag{6.10}$$

Gold has a relative permittivity in air of $\epsilon_r^{650}(Au) = -13.6 + 1.0i$ at 650 nm excitation [86], which equates to a peak intensity enhancement of 12.3. In water, the relative permittivity becomes $\epsilon_r^{650}(Au) = -7.62 + 0.56i$, and consequently moves closer to the resonant condition. Thus, the peak intensity enhancement for a gold sphere becomes 16.3—over a 30% increase, a marked difference between the decreases seen for Si.

Aside from permittivity considerations, the AFM feedback system might also play an important role in understanding the different results for water. Shear-force feedback based AFM systems oscillate the probe laterally, while cantilever based systems as the one in our lab oscillate vertically. We have measured anharmonic oscillations in our system in water, due to unwanted pressure waves at the nanoscale (turbulence) that is more pronounced as the tip approaches the sample. This can prevent the tip from coming in close enough contact to the sample to allow for the full extent of the near-field signal to interact with the sample. By comparison, lateral oscillations of shear-force AFM are much smaller ($\sim$ 1 nm) than those employed in cantilever based AFM systems. Thus, less turbulence might be expected in an aqueous environment when compared to the much larger vertical oscillations (10 nm) used in a cantilever based system. Furthermore, since the oscillations are lateral, there is nothing precluding the tip getting arbitrarily close to the sample. Small samples (dye molecules) mean that the intensity-volume overlap can be appropriately optimized. Imaging with metal nanospheres yields quenching at

very short range, but good enhancement at a tip-sample gap of several nanometers [12].

More recently, one other group using a similar experimental setup to our own has also had some success using a TEFM system in water [87]. As seen in Figure 6.10, Frey et al. were able to obtain TEFM images of single molecules attached to glass coverslips with Si tips in an aqueous environment. They used a focused wedge illumination similar to ours but employed a higher numerical aperture objective, NA = 1.45. In this experiment two photon counters were gated such that they

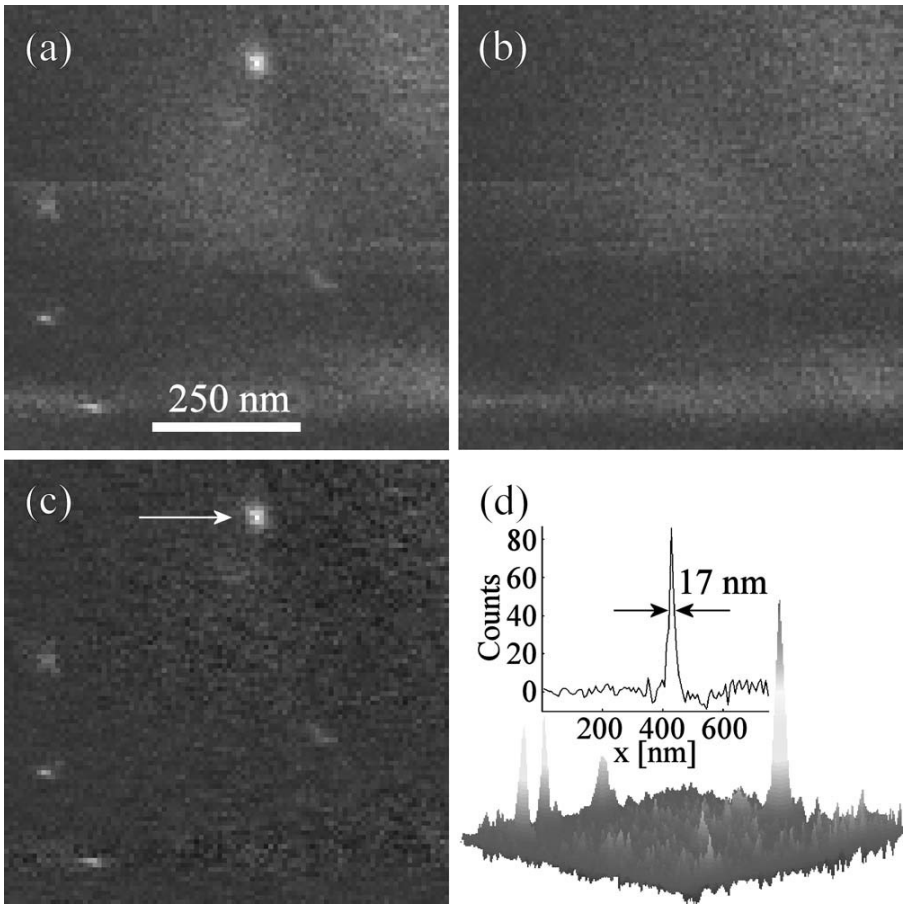

Figure 6.10. TEFM images of single ATTO-740 molecules in water. (a) Raw fluorescence image including near-field and far-field signals. (b) Far-field fluorescence image. (c) Near-field fluorescence image obtained by the difference of the previous two. (d) 3D image and cross section of the the single molecule indicated in (c). Reprinted with permission from [87].

collected photons at the top and bottom of the tips trajectory respectively. This allows for isolation of near-field signals in a similar fashion that was shown in Figure 5.6(c).

Importantly the single molecule samples used, ATTO-740, had a relatively low quantum yield, $q = 0.1$. As the TEFM tip approaches a fluorescent sample both the radiative and nonradiative decay rates can change [88, 89]. The quantum yield in the absence of a tip is defined as:

$$q_0 = \frac{\gamma_r^0}{\gamma_r^0 + \gamma_{nr}^0}, \tag{6.11}$$

where γ_r^0 and γ_{nr}^0 are the radiative and nonradiative rates in the absence of the tip. As the tip approaches both of these rates can be changed in some distance dependent fashion. These rates in the presence of the tip become $\gamma_r = A(z)\gamma_r^0$ and $\gamma_r = B(z)\gamma_{nr}^0$, where $A(z)$ and $B(z)$ represent the exact distance dependent changes to these rates respectively. The apparent quantum yield in the presence of a tip can then be written as follows:

$$q = \frac{A(z)\gamma_r^0}{A(z)\gamma_r^0 + B(z)\gamma_{nr}^0}. \tag{6.12}$$

Looking at the case of closest approach, $A(z = 0) = A$ and $B(z = 0) = B$, the apparent quantum yield can be easily compared to the initial quantum yield:

$$\frac{q}{q_0} = \frac{1 + \dfrac{\gamma_{nr}^0}{\gamma_r^0}}{1 + \dfrac{B\gamma_{nr}^0}{A\gamma_r^0}}. \tag{6.13}$$

Plotting Equation 6.13 as a function of initial quantum yield reveals that for lower initial values of the quantum yield, increasing the radiative rate relative to the nonradiative rate yields larger observed signals even in the absence of field enhancement, see Figure 6.11. This means that for low initial quantum yields, larger total enhancement factors are possible.

This report from Frey et al. indicates that with a judicious choice of sample, that imaging in water is indeed possible with silicon tips. In this experiment there are several other factors of significance worth mentioning. First, the emission

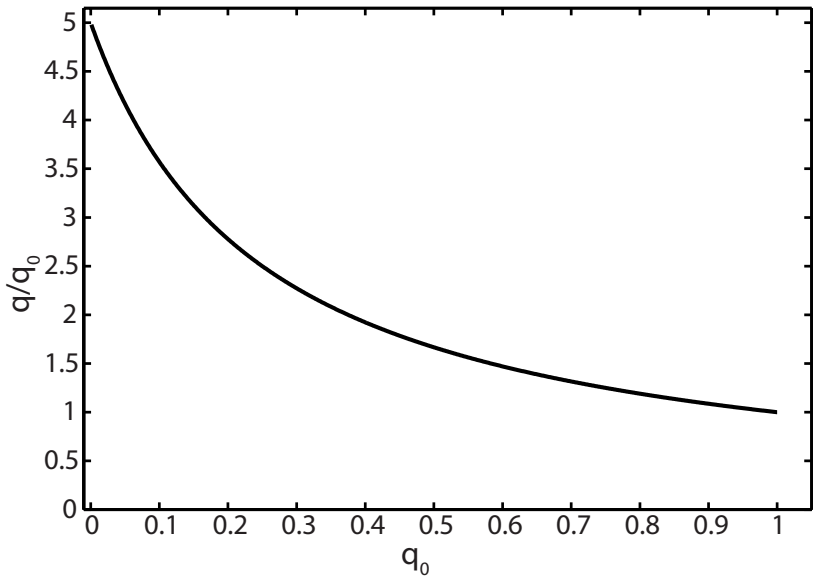

Figure 6.11. Apparent quantum yield plotted as a function of q_0. The apparent quantum yield, q, is normalized to the initial quantum yield q_0 as defined in Eq. 6.13. Here the values for the terms A and B in Eq. 6.13 have been set such that A = 5B, or equivalently, the radiative rate is increased five times as much as the nonradiative rate in the presence of a tip. Note that for low initial quantum yields, the apparent quantum yield can be increased significantly.

wavelength of the fluorophore employed ($\lambda = 740$ nm) is further to the red than we use. This means that the imaginary part of the permittivity of Si at this wavelength is smaller than what we encounter, which should reduce any quenching effects of the Si. Second, the NA of the objective used here ($NA = 1.45$) is somewhat higher than what we employ, meaning that more evanescent (and consequently axial) field components are included. Using higher numerical aperture objectives also improves the total collection efficiency of the system. Finally the excitation laser was pulsed in synchrony with the oscillation cycle of the tip. By only allowing the excitation laser to be on when the tip is at the peak and bottom of its trajectory, unnecessary photobleaching of the sample is avoided.

6.7 Quenching

In light of all the difficulties encountered in attempting TEFM in water, these can be alleviated to a great extent if quenching can be used as a contrast mechanism.

As discussed previously, quenching is limited in both the maximum attainable resolution as well as contrast. However, for certain types of samples that must be imaged in water, it at least offers an easily implementable solution if not at ideal conditions. Figures 6.12 and 6.13 demonstrate the efficacy of quenching as a contrast mechanism in water.

While a combination of reduced axial fields as well as reduced lightning-rod effect make enhancement a difficult image contrast mechanism in water, quenching does not seem to be much affected. Some difficulties associated with quenching as a contrast mechanism in water still exist to be sure: difficulties in sample preparation and performing high fidelity AFM scans in fluids can be persistent issues, but generally not ones that cannot be overcome. While no thorough study has been performed on the subject, quenching efficiencies appear to be slightly reduced compared to those found in air. Selected images from a full tomographical data set of TEFM of 20-nm fluorescent latex beads in a buffer solution is shown in

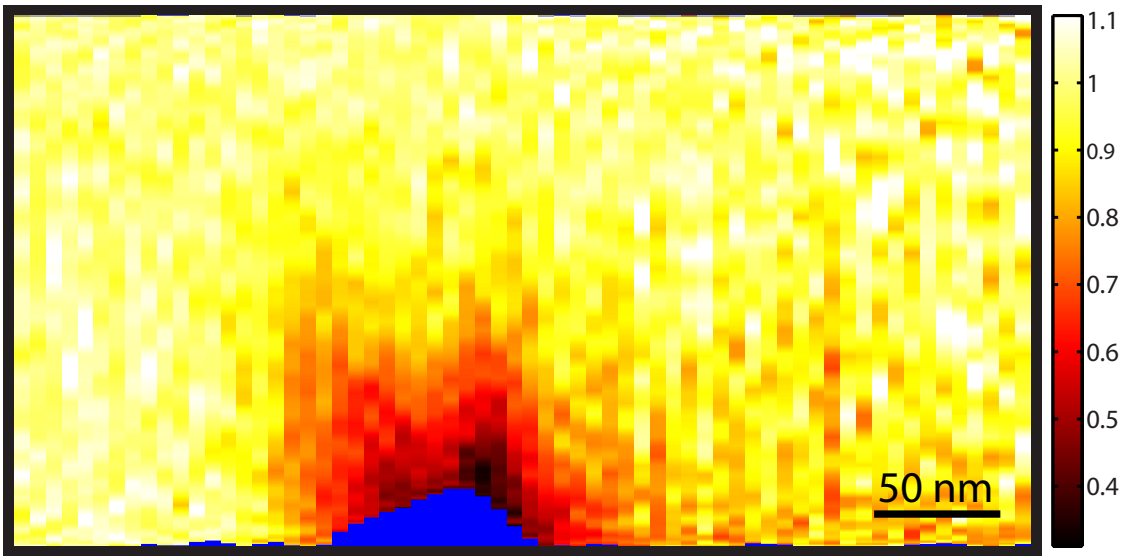

Figure 6.12. X-Z tomographical slice demonstrating quenching in water. Gold coated silicon tips were scanned over 20 nm fluorescent beads in an aqueous environment. The normalized fluorescence rate is plotted as a function of the X-Z position of the tip relative to the sample. The fluorescence rate has been divided by the far-field fluorescence rate ($z > 168$nm).

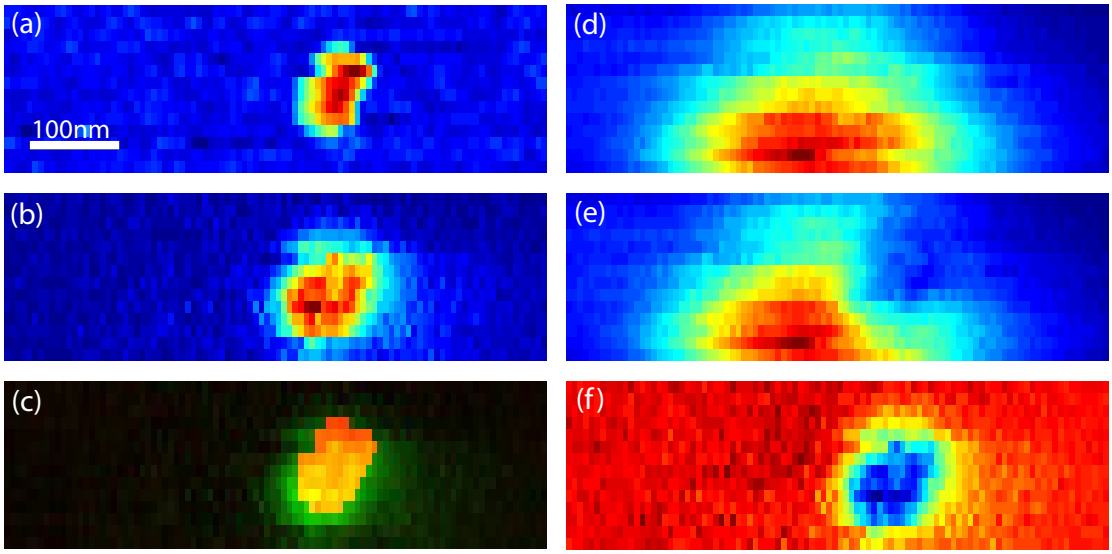

Figure 6.13. X-Y slices of a full tomographical data set taken as a Au tip scans over a fluorescent latex bead. (a)—AFM height trace. (b)—Lock-in trace. (c)—Height data in green overlaid with lock-in data in red, overlapping regions yield yellow. (d)—Far-field fluorescence signal ($z > 168nm$). (e)—Combined near and far-field signal ($z < 21nm$). (f)—The difference between images (e) and (f).

Figures 6.12 and 6.13. Both X-Z slices and various X-Y cuts in these two figures demonstrate that good contrast of at least single fluorescent beads can be obtained in aqueous solutions.

6.8 Embedded Worms

As it became clear that TEFM imaging in aqueous environments was at least very difficult, another potential workaround was to image biological samples embedded in plastics. Erik Jorgensen's group at the University of Utah has some expertise in embedding *C. elegans* worms in plastics for electron microscopy studies. As mentioned in the introduction, electron microscopy (EM) has excellent resolution but low chemical specificity. Image contrast, and to some extent chemical specificity, can be enhanced by labeling parts of the cell with electron dense metals such as Osmium. Such electron microscopy preparations generally do not leave fluorophores in tact; however, the Jorgensen group has developed a gentler treatment that allows

for imaging with EM but still leaves fluorescence labels in tact. Worms embedded in plastic can then be physically sectioned using a microtome to very thin slices and imaged under fluorescence microscopy as well as EM. The correlation of the two resulting images can lead to a better understanding of cell architecture.

It was our hope that one of these sections could be imaged with TEFM, as reportedly the microtome device could cut sections as few of several tens of nanometers thick. In this way we could use Si tips in air to image real biological specimens. We knew that our near fields would only extend a fraction of the way into the plastic, and thus the fluorescence background rate would be a more significant problem. Many attempts were made to obtain a TEFM image of *C. elegans* fluorescently labeled with RFP. Many of the initial trials failed due to inadequate fluorescence signal as a result of the EM prep. After some refinement of this step by Shigeki Watanabe, strong fluorescence signals were obtained as shown in Figure 6.14.

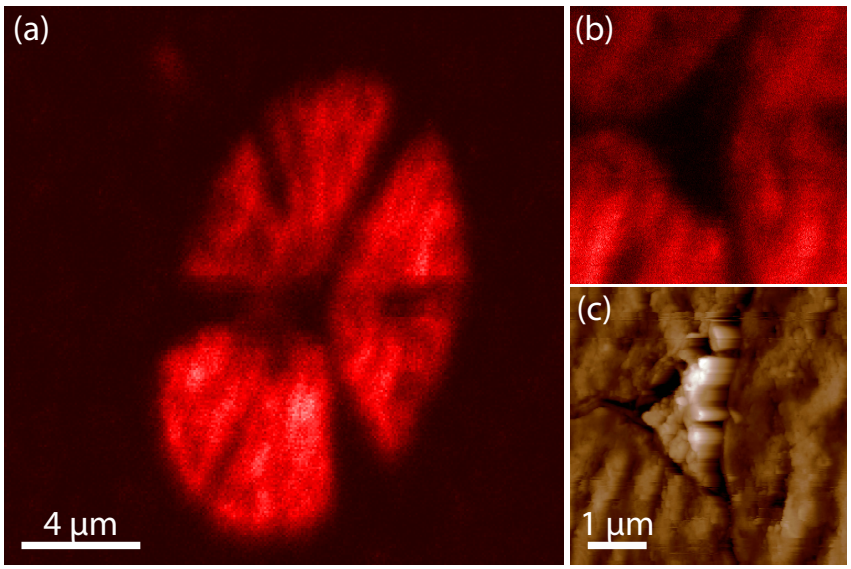

Figure 6.14. Fluorescence image of a *C. elegans* section in plastic. A *C.elegans* sample with a fluorescently labeled pharynx was fixed, embedded in plastic, and sectioned to a nominal thickness of 50 nm. A far-field image is shown in (a). Panels (b) and (c) show zoomed in fluorescence and simultaneously acquired AFM height images respectively—no near-field signal was observed. Ridges in the bottom left corner of panel (c) have a peak-to-peak height of $\sim$ 100 nm, indicating a true sample thickness much larger than 50 nm. Correlation between fluorescence intensity (b) and sample thickness (c) is also observed.

Once strongly fluorescing samples were available it also became clear that we were unable to successfully do TEFM. This failure was due to thick sample sections. The nominal section thickness shown here was 50 nm as reported by the microtome. Figure 6.14(c) shows the AFM height trace of the section, where the peak-to-peak height of the ridges observable in the bottom left corner is $\sim$ 100 nm. With this much variability in just one section, the total thickness is more than 100 nm. Approach curves were performed to check for any near-field signal at all, but none was found. This is most likely due to photobleaching of fluorophores near the surface, or at least these might be dying first, with the strong fluorescence originating from deeper within the sample. There is some evidence for this as there is some correlation between fluorescence signal and sample thickness in this image.

6.9 Fluorescence Correlation of VSV

Not giving up hope yet on imaging biological samples with TEFM, our most recent biological collaboration is with Saveez Saffarian, who studies the Vesicular Stomatitis Virus (VSV). VSV is a negative sense RNA virus, meaning that its genetic material is coded in a single strand of RNA that cannot be read by the host cell because the RNA reading direction is backwards. This means that VSV must package its own polymerase into the capsid. Since a VSV virion is asymmetric, it might be expected that the location of the polymerase might also be directional.

Our goal then is to use TEFM to identify the location of a fluorescently labeled polymerase within the virion. From EM images it is known that each virion is on average 180 nm long and 80 nm across, in a bullet shape. This means that either enhancement in an air (inert gas) environment, that fluorescence enhancement might be obtainable. Other options for obtaining the desired result would be to use fluorescence quenching with metal tips. Especially when combined with centration techniques this could also yield the relative position of the polymerase with high precision.

This work is still currently underway, attempts thus far have yielded very few images of virions resembling bullet shapes, one such image is shown in Figure 6.15.

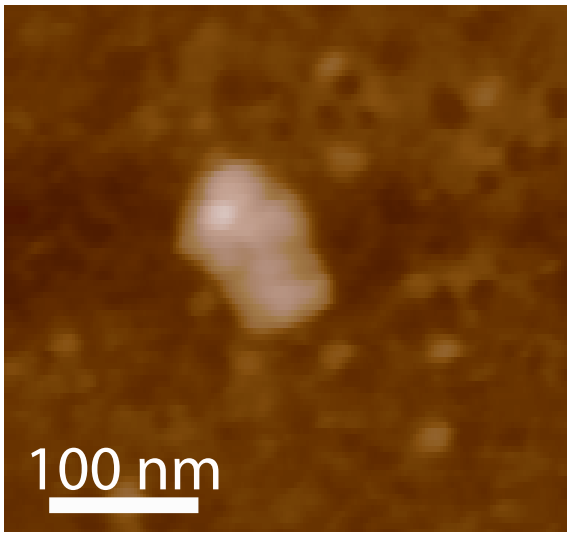

Figure 6.15. AFM image of a VSV virion, note the definitive bullet shape. The maximum height of the virion is 30 nm.

This is in part to due low virion densities such that much hunting is required to find one. Also, the virions imaged thus far seem to be flattened in height, which might be expected as they are simply attached to glass coverslips via van der Waals interactions. Further refinement of sample preparation techniques will undoubtedly solve these issues.

A more troubling issue is that as the tip scans over a virion, it causes it to go dark. This has been seen before in latex beads, especially when imaged in water. To circumvent this problem it is also possible to do a fluorescence-AFM correlation experiment. This will require attaching both small fluorescent beads as well as the virions to a glass coverslip. We can then perform a far-field only optical image followed by an AFM trace. We plan to use the 20-nm dye-doped beads as fiducial marks to allow for calibration of the optical images, correcting for any offset between the optical and height images. By centrating the far-field fluorescence spots we can correlate the fluorescent centers with AFM topography. This should give us adequate resolution to also determine the location of the fluorescent centers within the virion.

6.10 Conclusions

While TEFM imaging in water has been unsuccessful in our particular experimental setup, it has been shown to in fact be possible. The results of this work indicate that silicon, while yielding excellent results in air, is not the best suited material for obtaining field enhancement in water. By attaching metal nanoparticles to the end of dielectric tips, field enhancement should be observable in our system as well. Utilizing the ideas discussed here about maximizing axial field components, it is recommended to use a radial polarization in conjunction with a wedge beam mask to optimize axial fields. Furthermore, the mask should be constructed with great care such that *all* supercritical rays are utilized, especially those closest to the critical angle. Lastly, low-quantum yield fluorophores are desired such that maximal signal enhancement factors may be obtained. For low density samples such as the virus experiment just detailed, correlation between far-field fluorescence signals and AFM topography seems well suited for dealing with delicate samples.

Another interesting avenue for obtaining near-field signals is to use fluorescence lifetime imaging. Several groups have demonstrated that fluorescence lifetime is a viable near-field contrast mechanism. Furthermore, it has been shown that the resolution of near-field lifetime measurements can yield resolution on the same scale as the AFM topography [90]. This makes another compelling reason for implementing the use of metal nanoparticles as probes for aqueous imaging with TEFM.

CHAPTER 7

CARBON NANOTUBES

Besides working towards extending the abilities of TEFM into aqueous imaging, several other major projects in the lab have been pursued. One of these projects is the use of carbon nanotubes (CNTs) as near-field optical probes. I have been involved in all aspects of this project: from the growth of the CNTs themselves, to attachment of the CNTs to AFM probes, and finally in taking optical data with these ultrasharp probes. As I have been involved in the growth of CNTs, I have a working knowledge of many of the important aspects for successful growth; however, as this project was headed by other students, I give only a sketch of the growth procedure here. I have been much more intimately involved in actually using CNTs as optical probes, so this chapter will focus primarily on imaging.

7.1 Growth

Nanotubes are grown on silicon substrates using a chemical vapor deposition (CVD) method. This involves spreading catalyst particles on ultraclean silcon wafers, and then flowing some form of carbon-containing gas over the catalyst particles at high temperatures. The particular catalyst particles used are iron. Mixtures of argon, methane, and hydrogen are input into a tube oven at temperatures greater than 900 $^{\circ}C$. The protocol used in our lab was developed by Jason Hafner [91], later adapted by the Pat Collier group, and has been shown to produce predominately single-walled nanotubes with average diameters between 3-5 nm [92, 93]. Furthermore, as is typical in most CVD growth methods, it is expected that roughly 2/3 of the nanotubes are semiconducting.

The specific catalyst deposition procedure as well as flow rates and other technical details can be found in our growth protocol in Appendix F. Nanotubes were

routinely grown in our lab for quite some time, until the procedure suddenly stopped working. After troubleshooting for several months, we have determined that an important factor for successful growth is the presence of an oxide layer on the Si substrates.

As there was a substantial period of time when we were unable to grow our own nanotubes, we acquired a nanotube wafer from Yoshie Narui from Pat Collier's group at Caltech. This wafer was made from the same protocol we also use. Most of the data presented in this chapter were taken using nanotubes from Caltech that were attached to AFM probes in our lab.

7.2 Pickup and Shortening

To use CNTs as optical probes we attach them to the apex of standard AFM probes. This is done in a relatively simple fashion: the AFM tip scans over a substrate containing CNTs until one becomes attached to the probe. This pick-up method is not original to us, it was first developed by Hafner et al. [91]. The attachment of the CNT to the tip occurs via van der Waals interactions, thus it is thought that primarily vertically oriented CNTs become attached. In this regard, gold-coated AFM probes have been the preferred tip for CNT pick-up, as CNTs become attached more readily and stay attached for longer periods of time as compared to silicon AFM tips.

Attachment of a CNT can be determined from several different AFM channels. One such indicator is a sudden change in resolution. As CNTs have much smaller diameters than a Si or Au AFM probe, attachment manifests itself as an increase in resolution as shown in Figure 7.1. This is not always the case, as extremely long CNTs can also become attached. These long CNTs lack the lateral stiffness of their shorter counterparts and can actually cause low fidelity scans and/or decreased resolution [91, 93]. Another indicator of CNT pick-up is the AFM phase trace. Attachment of a CNT leads to large differences in interfacial contact areas and thus tip-sample forces can change dramatically, as exhibited in Figure 7.1(b). To avoid further CNT pick-up, care is taken to only scan several more lines, ensuring

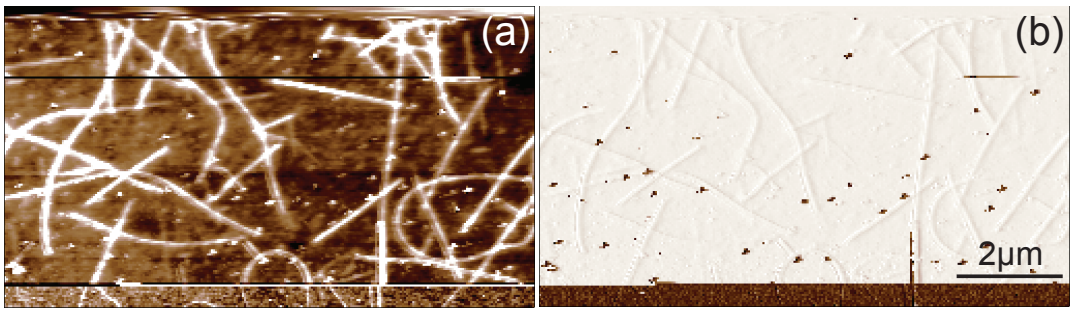

Figure 7.1. AFM topography (a) and phase (b) images obtained while AFM is scanned from top to bottom over a nanotube-covered silicon wafer. Resolution increased dramatically when a CNT was picked up, while the phase suddenly changed. Reprinted with permission from reference [6], © 2008 IEEE.

the CNT has become robustly attached.

Once a CNT is attached to the tip, it is still generally too long to be used for imaging. As a general rule, a CNT probe will produce a good AFM image if it protrudes from the apex of the tip less than about 170 nm. Excessively long CNTs can exhibit large thermal vibrations [91]. Also, as mentioned, lateral stiffness is also a function of length and diameter—wider CNTs being stiffer, while longer CNTs are less rigid.

Figure 7.2 shows an SEM image of a nanotube that has been picked up on a silicon tip using the method just described. The distal end of the tip appears blurry due to vibrations, illustrating how long CNT protrusion lengths can be bad for AFM imaging. Also seen are several other nanotubes attached closer to the apex of the silicon pyramid, exemplifying how multiple CNTs may become attached, possibly leading to image artifacts—multiple CNTs attached to a tip could cause multiply peaked features.

The protrusion length of the CNT can be measured via force curves as demonstrated in Figure 7.3; as the tip contacts the surface it begins to buckle, further pushing reveals the location of the AFM tip pyramid. The distance between these points indicates the length. CNTs are shortened by applying short voltage pulses between the tip and a gold coated glass slide. These pulses are generally 10 μs long

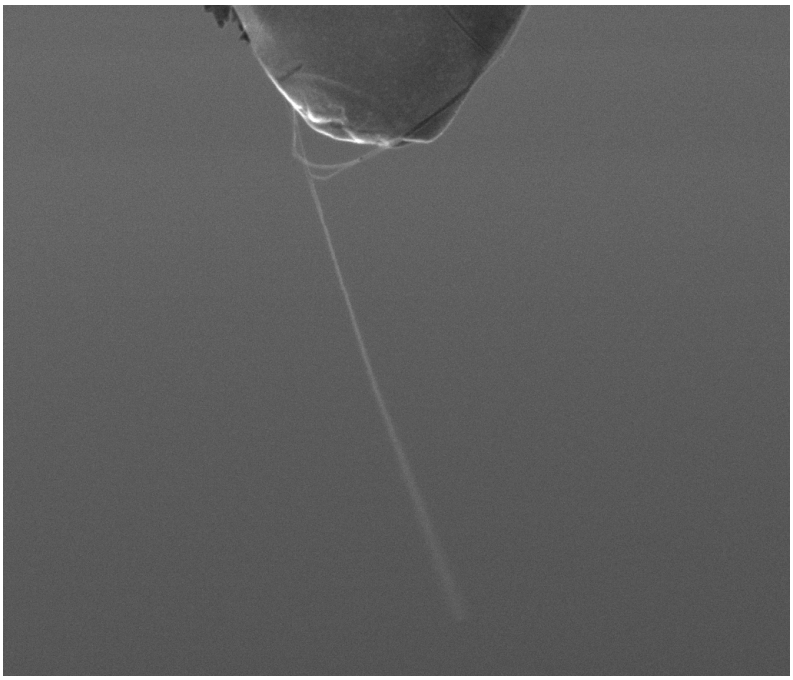

Figure 7.2. SEM image of a CNT probe. Multiple CNTs became attached to this Si tip using the pick-up method. The length of the protruding CNT is too long for good AFM imaging as thermal vibrations at the tip lead to broadening of the image. Reprinted with permission from reference [6], © 2008 IEEE.

and have an amplitude of 10-20 V. It should be noted that not all CNTs withstand the shortening procedure: we usually find a success rate of about one in three for Au tips and one in five to ten for Si tips.

7.3 Fluorescence Imaging

The first report of using CNTs as fluorescence probes was made by our lab in 2008 [6]. After testing many CNT probes, we found that they all strongly quench fluorescence. Initially we were surprised at this fact, as we conjectured that some fraction of the attached CNTs would be semiconducting and might lead to enhancement. Unfortunately we lack the equipment necessary to determine the particular type of CNT we attached to an AFM probe. Using the recently developed near-field tomography system in our lab, we have used CNTs to probe quantum dots (QDs) quite extensively. All QD studies presented here consist of a CdSe core and ZnS shell and have nominal dimensions of 4 nm $\times$ 9 nm (Invitrogen).

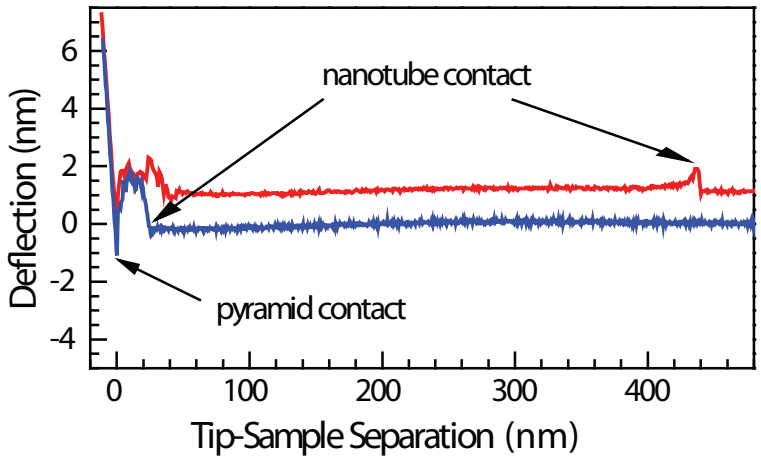

Figure 7.3. Approach curves of nanotube shortening. AFM contact-mode approach curves of the same CNT probe: before shortening (top curve - red), and after shortening (bottom curve - blue). There is a kink in the deflection curve when the nanotube touches the surface, as indicated in the figure. The distance between these kinks and where the pyramid contacts the surface gives the length of the nanotube: ~ 440 nm before shortening and ~ 30 nm after shortening in this case. Reprinted with permission from reference [6], © 2008 IEEE.

Figure 7.4 demonstrates a typical quenching profile as a CNT tip scans over a QD. Averaging over many CNTs and many more QDs, typical quenching efficiencies of $\sim 95\%$ are usually seen.

Approach curves constructed from photons corresponding to the QD "on" state have a very different functionality compared to those that were emitted during the QD "off" state, as shown in Figure 7.5. The QD "off" state is a low quantum yield state. The quantum yield (q) represents the number of photons emitted per the number absorbed, and is thus always less than one. It can also be defined in terms of radiative and nonradiative rates: the total fluorescence decay rate arises from the addition of contributions from the radiative decay rate γ_r, the intrinsic nonradiative decay rate γ_i associated with intramolecular energy dissipation, and an extra nonradiative decay rate term γ_{nr} accounting for energy losses to the local environment [13].

$$\gamma = \gamma_r + \gamma_{nr} + \gamma_i \tag{7.1}$$

The apparent quantum yield is then given by:

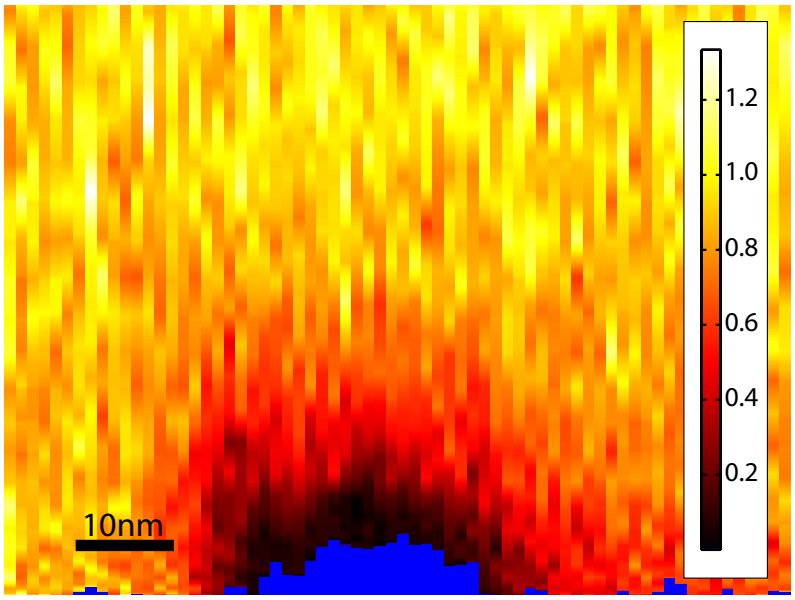

Figure 7.4. X-Z tomographical slice of CNT tip scanning over a quantum dot. The fluorescence count rate has been normalized by the far-field count rate ($z > 48$ nm) to emphasize the extent of quenching.

$$q = \frac{\gamma_r}{\gamma} = \frac{\gamma_r}{\gamma_r + \gamma_{nr} + \gamma_i}. \tag{7.2}$$

In this picture the intrinsic nonradiative decay rate, γ_i, represents the decay rate associated with intrinsic loss mechanisms in the absence of any perturbation (free space). From reference [13], we know that the quantum yield can play a significant role on the particular balance between enhancement/quenching of a fluorophore. High quantum yield states ($\gamma_r >> \gamma_i$) have nowhere to go but down; as a metal nanostructure approaches, the apparent quantum yield decreases as nonradiative channels become available (γ_{nr} increases) and start to compete with the radiative decay rates. On the other hand, low quantum yield states already have high intrinsic nonradiative rates, thus as a metallic nanostructure approaches, the extrinsic nonradiative rate (γ_{nr}) must also be very large to compete. It would thus be expected that "off" states are influenced to a lesser extent by the addition of nonradiative decay channels as a CNT approaches, as seen in Figure 7.5.

Note the short length scale of of the QD-CNT interaction as seen in Figure 7.5; quenching begins at a tip-sample separation of about 20-30 nm. This is quite

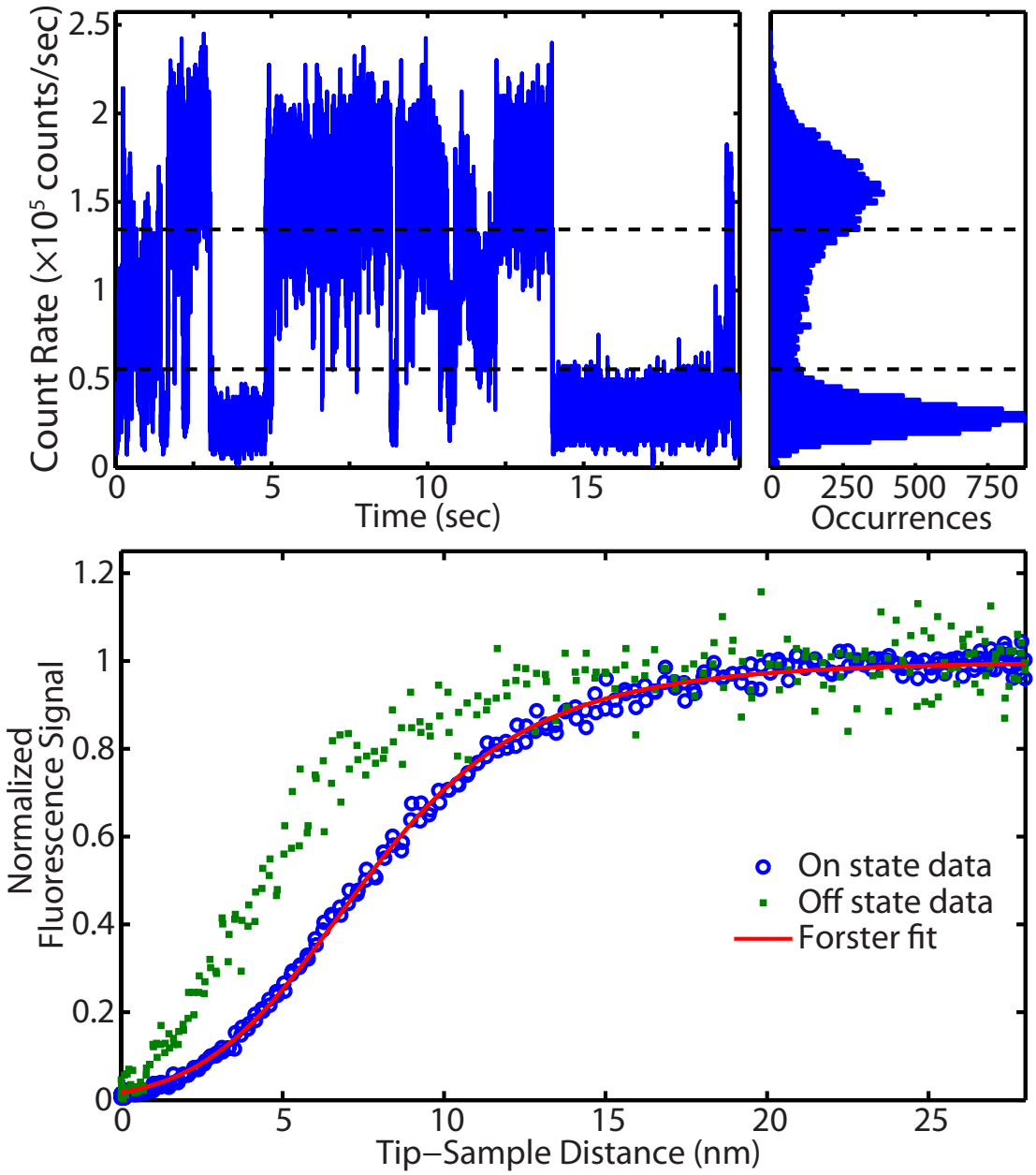

Figure 7.5. Blinking analyses of a single CdSe/ZnS QD interacting with a CNT probe. A 20 second time trace of the fluorescence count rate while a CNT probe intermittently taps the surface highlights the "on-off" blinking behavior of the QD. "On" states are determined to be those above the upper dashed line, while "off" states are those below the lower dashed line. Thresholds were set from a histogram of the count rates (upper right panel). The tip-sample functionality is plotted for both states (lower panel).

different than the long-range quenching observed with gold coated AFM probes (cf. Fig. 5.3) that starts to occur at length scales greater than 150 nm. This lack of long-range quenching is due to a reduced ability of the QD to couple to surface plasmon polariton modes, relying instead on only shorter ranged Förster-like energy transfer as described in Chapter 2. The exact functionality of this energy transfer is currently an extremely active area of research in our lab.

A Förster model of energy can be made to fit this data extremely well as indicated in Figure 7.5. This model assumes energy transfer between a pair of dipoles. The basic Förster model for energy transfer is given below [56]:

$$E = \frac{1}{1 - \left(\frac{z}{R_0}\right)^6}, \tag{7.3}$$

where z is the tip-sample separation distance (or dipole-dipole distance), and E is the energy transfer (100% at $z = 0$). The value R_0 is known as the Förster radius; experimentally it is measured as the separation distance at which 50% energy transfer is achieved. Theoretically it represents the relative dipole-dipole coupling strength with larger R_0 values meaning stronger interaction. Fundamentally R_0 depends on coupling parameters such as the overlap integral between the spectrum of the energy donor with the absorption spectrum of the acceptor, $f_D(\omega)$. Other parameters folded into R_0 include a term that accounts for the dipoles' relative orientations, κ , as well as the absorption cross section of the acceptor σ_A [56]:

$$R_0^6 = \frac{9c^4\kappa^2}{8\pi} \int_0^\infty \frac{f_D(\omega)\sigma_A(\omega)}{n^4(\omega)\omega^4} d\omega. \tag{7.4}$$

In order to fit the data in Figure 7.5, the standard Förster model has been adjusted to include an offset term, z_0:

$$E = \frac{1}{1 - \left(\frac{z - z_0}{R_0}\right)^6}. \tag{7.5}$$

Initially this was done to account for the fact that the CNT could never come to a zero distance of the center of the QD. As seen in Figure 7.5, fitting the data in

this way leads to excellent agreement between the data and theory. One issue we are still addressing is that the offset value in such a fit is generally much larger than originally thought; rather than being a few nm (the width of a QD), the offset values for z_0 are more typically in the range of 10-20 nm. Through further careful analysis of the data, it was discovered that there is a strong correlation between the coupling strength R_0 and this offset term z_0: typical R_0 values are found between 15-35 nm, where longer values of R_0 correspond to larger values of z_0 for a given CNT-QD measurement. This is the very issue that is currently being addressed—to determine the specific relationship between these two parameters, and consequently the most appropriate interpretation of z_0.

7.4 Asymmetric Histograms

Some CNT probes, regardless of length, behave quite differently than most. Asymmetric histograms of tip-oscillation phase versus fluorescence rate are sometimes observed for certain CNTs. It is important to remember approach curves are simply calculated from histograms, such that asymmetry in the histogram consequently will translate to asymmetric approach curves. An important step in this procedure is being able to identify the phase that corresponds to tip-sample contact. With symmetric histograms this is not problematic, as Gaussian fits provide excellent results. Asymmetric histograms are also often marked by a broad flat bottom of the quenching profile (Fig. 7.6 - inset). This makes determining the precise phase corresponding to tip-sample contact nearly impossible; only reasonable estimates may be made.

Since the histogram is asymmetric it is difficult to assign a "good" or "bad" edge to the curve from a single data sample. However, from many data samples there is at least an "average" looking histogram that is expected. The falling edge of Figure 7.6 corresponds quite well to this idea of an average histogram, while the rising edge exhibits a much steeper functionality. To interpret, the CNT behaves more or less as would be expected as the CNT approaches the sample; upon retraction of the probe however, there is some irregularity.

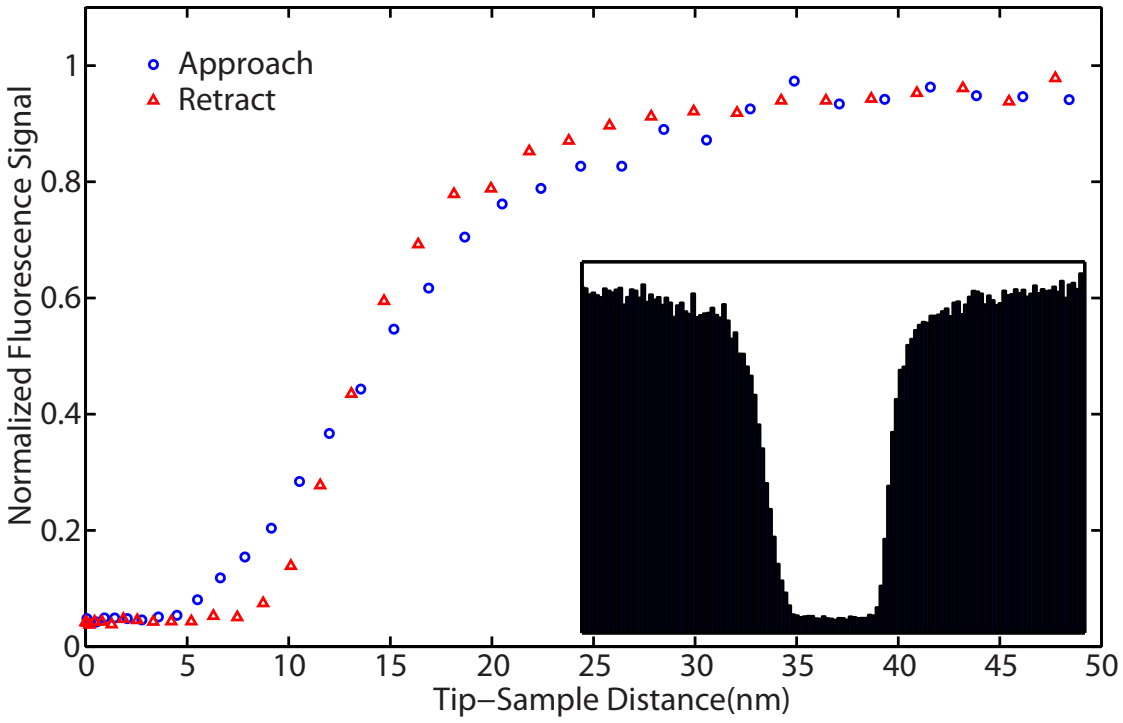

Figure 7.6. Approach curve and corresponding phase histogram (inset) of a CNT buckling event on a QD. The falling edge of the histogram corresponds to the approach of the tip, which is more gradual than the retraction (rising edge).

The source of this irregularity is still debatable. It can possibly be attributed to a CNT buckling event, where as the CNT pushes into the sample surface, the CNT buckles due to a lack of axial stiffness. Buckling would cause a larger portion of the CNT to interact with the surface, thus a greater lift-off force would be required to overcome van der Waals interactions with the sample surface. During this time the buckled nanotube could presumably be in extremely close proximity (if not touching) to the QD, leading to an extended duration of quenching. This buckling hypothesis is supported by some evidence of poor AFM height traces (exhibited by image broadening and low-fidelity scans) being associated with CNT probes that yield asymmetric approach curves. This still is somewhat of a simple observation, and a more thorough data analysis is needed to confirm this. Should this be the case, this might be evidence of extremely small diameter CNTs, as even at short

protrusion lengths (< 100 nm), they might lack axial rigidity.

Another possible explanation is the possibility of charge transfer between the QD and CNT. Some previous evidence has shown that as a CNT scans over a QD that it can remain completely quenched in the region corresponding to the AFM height trace [6]. Since the tip is oscillating, even when the CNT is directly probing the QD it would be expected that there still be some fluorescence signal from the QD—at least when the tip is at higher points in its trajectory. It could be the case that asymmetric histograms are caused by charge transfer from the QD to the NT. The hypothesis is that at when the CNT contacts the QD there is some charge transfer, causing the QD to go into a dark state, thus requiring some recovery time (longer than an oscillation cycle) before the QD becomes uncharged and returns to a higher emitting state again. If the recovery time is less than one oscillation period, this charge transfer hypothesis is quite testable. After some dark period, should the QD suddenly become bright, it might be expected that the rising edge of the histogram might have a very narrow linewidth (or at least for some portion of the rise time - depending on the tip-height before it went bright).

Further investigation of the source of this asymmetry is still needed. Performing approach curves at different oscillation amplitudes and set-point levels may confirm that when pushing hard enough even average CNT probes can be made to buckle. Alternatively or additionally, attaching CNT to AFM probes of a wide range of resonance frequencies might confirm a characteristic time scale for fluorescence recovery, which would be a further indicator of charge transfer. The definitive experiment would be to perform this measurement in a vacuum environment while monitoring any current across the tip, thus confirming any possible charge transfer.

7.5 Outlook

These studies on CNT-QD interactions highlight the extreme power of the TEFM system that has been developed in our lab. Clearly TEFM is far more than simply an imaging technique; it allows for mapping of complex interactions on the shortest of length scales. Another implication beyond the study of CNT probes

for energy/charge transfer studies is that of imaging.

It has been shown that CNT AFM probes can be further stabilized by coating them with layers of silicon oxide [94]. These stabilized tips are able to survive immersion in water. This implies that such probes may also be used for near-field studies in water. Quenching has already been demonstrated as a contrast mechanism in water, however, at low resolution. The short range quenching of CNTs may also provide increased quenching resolution in that application.

Quenching is not expected to lend itself well to imaging dense ensembles of fluorophores as it lacks contrast. However, fluorescence lifetime imaging using stabilized CNTs may well prove advantages in terms of both resolution and contrast in water. This is due to the fact that fluorescence lifetime profiles have been shown to be much narrower than quenching [90]. This is definitely an important experiment worth attempting for fluorescence near-field measurements in water.

CHAPTER 8

CONCLUSIONS

When I first joined the lab, the original goal of my project was to be writing a dissertation about the many biological questions we have been able to answer using TEFM. Clearly this has not been the case. TEFM had indeed established itself as a world-class imaging system in terms of resolution; however, as I embarked on this project, it became clear that there were still many unknowns about the field enhancement mechanism, most appropriate illumination conditions, the role of quenching, image contrast etc. Certainly this is no fault of the previous work done by my advisor at Caltech; there are still many unknowns in near-field optics as it is still a developing field, and answering some of these questions has been an entire dissertation project. Here I summarize some of the more important contributions this dissertation has made toward the advancement of TEFM towards our original goal.

8.1 Enhancement vs. Contrast

Some of the basic theories presented in Chapter 2 have required a great deal of research to bring them to one coherent picture of the roles of enhancement and quenching in near-field optics. Even at the most recent edition of the premier conference in our field (10th International conference on near-field optics, nanophotonics and related techniques in Sept. 2008), leading experts struggled to understand all the nuances and complexities associated with tip-sample interactions of various sorts, but more particularly quenching. The stark differences between the long-range quenching behavior observed in our lab with metal coated AFM probes versus the short range quenching observed by metal nanoparticle probes stymied even the most expert researchers in the field. In this regard, the theoretical back-

ground found in Chapter 2 represents a great deal of advancement in understanding some fundamental principles at work in the particular embodiment in our near-field system.

8.2 Image Contrast

A considerable contribution to understanding the limitations of TEFM has been the semianalytic theory of image contrast presented in Chapter 4. Image contrast is a vital component in achieving the goal of imaging high density fluorescent samples. This treatment allows for a measure of expected image contrast for any density sample when the characteristics of the tip-sample interactions are known. No other treatments of near-field image contrast achievable when utilizing demodulation techniques is known to exist. Furthermore, this model for image contrast is able to predict the optimal oscillation amplitude—a very important experimental parameter that might otherwise take a great deal of experimentation to optimize empirically. Lastly, this model has been extended to account for using fluorescence quenching as a contrast mechanism.

8.3 Tomography

Chapter 5 presents a powerful new technique for mapping three-dimensional tip-sample interactions. This single-photon near-field tomography technique allows for recording the X-Y-Z position of the tip for each photon emitted by the sample. Collecting the data in this fashion allows for arbitrary sectioning and post processing of the data. Post processing has been used to completely isolate near-field signals. Such treatment allows for further gains in image contrast.

8.4 Aqueous Imaging

Extending TEFM into the realm of aqueous imaging has been fraught with many difficulties. Difficulties and challenges in sample preparation and tip-alignment aside, it has been shown that aqueous imaging with dielectric tips is expected to yield enhancement factors only $\sim 10\%$ as large as might be expected in air. Unfortunately this has resulted in a complete inability to achieve any enhanced

near-field images in an aqueous environment using silicon probes. On the positive side, many valuable lessons have been learned from these studies. The optimal illumination mode for achieving field enhancement for air or water appears to be wedge illumination in conjunction with P-polarized light. Furthermore, several ways forward have been identified: using metal nanoparticle probes for field enhancement, or using fluorescence lifetime imaging as a contrast mechanism, either separately or in conjunction with metal nanoparticle probes. Near-field fluorescence lifetime measurements using sharp silicon probes may also lead to sufficient image contrast at unprecedented resolution. This concept needs a proof of principle experiment. As Si tips with s-polarized light exhibit weak quenching, lifetime changes may occur on the shortest of length scales, due to the small imaginary part of the permittivity.

8.5 Carbon Nanotubes

The use of carbon nanotubes as near-field probes has been demonstrated with high precision. Extremely efficient energy transfer between CNTs and QDs has been observed at length scales much shorter than gold tips, but much longer than typical dipole-dipole interactions. Furthermore, several experiments involving CNTs for aqueous imaging have also been proposed.

APPENDIX A

SOLUTION OF A DIELECTRIC SPHERE IN AN OTHERWISE UNIFORM ELECTRIC FIELD

A.1 Field-Enhancement

When a dielectric material is placed in a uniform electric field, the applied field polarizes the material as the electrons and ions migrate toward their respective sides of the dielectric. This charge separation creates an induced electric field, and the total field near the surface of the material is the superposition of both the applied field and the induced field. For a dielectric sphere, the total resulting field is found analytically. The solution is detailed in this appendix, closely following the derivation by Griffiths [95].

We set out to find the potential, V, both inside and outside the sphere by solving the Laplacian in spherical coordinates:

$$\nabla^2 V = 0$$

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial V}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 V}{\partial\phi^2} = 0, \tag{A.1}$$

Assuming azimuthal symmetry (V independent of ϕ) this reduces to:

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial V}{\partial\theta}\right) = 0. \tag{A.2}$$

As usual, we try separation of variables: $V(r,\theta) = R(r)\Theta(\theta)$ this simplifies to an ordinary differential equation:

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{1}{\Theta\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) = 0. \tag{A.3}$$

In the usual fashion, both the first and second halves of the equation must independently equal some constant. With a lot of foresight these constants can be written as follows:

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) = l\,(l+1) \tag{A.4}$$

$$\frac{1}{\Theta\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) = -l\,(l+1) \tag{A.5}$$

$R(r)$ has the general solution:

$$R\,(r) = Ar^l + \frac{B}{r^{l+1}}, \tag{A.6}$$

where A and B are yet to be determined constants. The solutions for $\Theta(\theta)$ are the Legendre Polynomials:

$$\Theta\,(\theta) = P_l\,(\cos\theta)\,. \tag{A.7}$$

Thus the most general solution for Laplace's equation in spherical coordinates with azimuthal symmetry is then:

$$V\,(r,\theta) = \sum_{l=0}^{\infty}\left(A_l r^l + \frac{B_l}{r^{l+1}}\right)P_l\,(\cos\theta)\,. \tag{A.8}$$

Now the specific task at hand is to solve Laplace's equation with a few boundary conditions. The potential is continuous across the interface:

$$V_{in} = V_{out} \qquad \text{at } r = R, \tag{A.9}$$

The normal component of the electric displacement ($D = \epsilon E$) is also continuous across the interface:$D_\perp^{in} = D_\perp^{out}$ at $r = R$. We use ϵ to denote the permittivity of the sphere and ϵ_m to denote the permittivity of the medium surrounding the sphere. It is also useful to define the quantity $\epsilon_r = \epsilon/\epsilon_m$, where the subscript r refers to *relative*.

$$\epsilon\frac{\partial V_{in}}{\partial r} = \epsilon_{medium}\frac{\partial V_{out}}{\partial r} \qquad \text{at } r = R, \tag{A.10}$$

Finally, the potential at infinity must be equal to the applied potential:

$$V_{out} \to -E_0 r\cos\theta \qquad \text{for } r >> R. \tag{A.11}$$

From these boundary conditions we find the potential inside the sphere must have the form:

$$V_{in}(r,\theta) = \sum_{l=0}^{\infty} A_l r^l P_l(\cos\theta). \tag{A.12}$$

This is due to the fact that as $r \to 0$ the term $\frac{B_l}{r^{l+1}}$ would tend to infinity. Outside the sphere, by similar arguments we get:

$$V_{out}(r,\theta) = -E_0 r\cos\theta + \sum_{l=0}^{\infty} \frac{B_l}{r^{l+1}} P_l(\cos\theta). \tag{A.13}$$

Now using the first boundary condition we obtain:

$$\sum_{l=0}^{\infty} A_l R^l P_l(\cos\theta) = -E_0 R\cos\theta + \sum_{l=0}^{\infty} \frac{B_l}{R^{l+1}} P_l(\cos\theta) \tag{A.14}$$

This leads to:

$$\begin{aligned} A_l R^l &= \frac{B^l}{R^{l+1}} \\ A_1 R &= -E_0 R + \frac{B_1}{R^2} \end{aligned} \tag{A.15}$$

The second boundary condition requires:

$$\epsilon_r \sum_{l=0}^{\infty} l A_l R^{l-1} P_l(\cos\theta) = -E_0 R\cos\theta - \sum_{l=0}^{\infty} \frac{(l+1)B_l}{R^{l+2}} P_l(\cos\theta) \tag{A.16}$$

$$\begin{aligned} \epsilon_r l A_l R^{l-1} &= -\frac{(l+1)B^l}{R^{l+2}} \\ \epsilon_r A_1 &= -E_0 - \frac{2B_1}{R^3} \end{aligned} \tag{A.17}$$

Solving now for the constants A and B:

$$\begin{aligned} A_l = B_l = 0 \quad &\text{for } l \neq 1 \\ A_1 &= -\frac{3}{\epsilon_r + 2} E_0, \end{aligned}$$

$$B_1 = \frac{\epsilon_r - 1}{\epsilon_r + 2} R^3 E_0 \tag{A.18}$$

Plugging these in we get the final solutions for the potentials:

$$V_{in}(r,\theta) = -\frac{3}{\epsilon_r + 2} E_0 r\cos\theta \tag{A.19}$$

$$V_{out}(r,\theta) = -E_0 r\cos\theta + \frac{\epsilon_r - 1}{\epsilon_r + 2}\frac{R^3}{r^2}E_0\cos\theta \tag{A.20}$$

We can now solve for the final form of the field, both inside and outside of the sphere:

$$\vec{E} = -\nabla V \Rightarrow \tag{A.21}$$

$$\vec{E}_{in}(r,\theta) = -\frac{3}{\epsilon_r + 2}\vec{E}_0 \tag{A.22}$$

$$\vec{E}_{out}(r,\theta) = \vec{E}_0 \cdot \cos\theta\left[1 + 2\frac{\epsilon_r - 1}{\epsilon_r + 2}\frac{R^3}{r^3}\right]\hat{r} + \vec{E}_0 \cdot \sin\theta\left[-1 + \frac{\epsilon_r - 1}{\epsilon_r + 2}\frac{R^3}{r^3}\right]\hat{\theta} \tag{A.23}$$

Another useful quantity is the maximum electric field at the poles of the sphere, which is given by:

$$E_{max} = E_0\left(1 + 2\frac{\epsilon_r - 1}{\epsilon_r + 2}\right), \tag{A.24}$$

where E_0 is the applied field.

APPENDIX B

SPHERICAL SHELLS

The analytic solution of a dielectric sphere with concentric layers of spherical shells in much the same way as outlined in Appendix A. As before, we seek a solution to the Laplacian in spherical coordinates, assuming azimuthal symmetry. The only real change now the addition of extra boundary conditions.

Consider a sphere of radius a, and permittivity ϵ_a, surrounded by a uniform shell of radius b and permittivity ϵ_b. The surrounding media has a permittivity ϵ_m and potential V_m. The boundary conditions now are as follows:

$$V_a = V_b \text{ at } r = a \tag{B.1}$$

$$\epsilon_a \frac{\delta V_a}{\delta r} = \epsilon_b \frac{\delta V_b}{\delta r} \text{ at } r = a \tag{B.2}$$

$$V_b = V_c \text{ at } r = b \tag{B.3}$$

$$\epsilon_b \frac{\delta V_b}{\delta r} = \epsilon_m \frac{\delta V_c}{\delta r} \text{ at } r = b \tag{B.4}$$

As before we seek solutions of the form:

$$V(r,\theta) = \sum_{l=0}^{\infty} \left(A_l r^l + \frac{B_l}{r^{l+1}} \right) P_l(\cos\theta). \tag{B.5}$$

Inside the sphere we find that V_a must have the form:

$$V_a(r,\theta) = \sum_{l=0}^{\infty} A_l r^l P_l(\cos\theta). \tag{B.6}$$

This is due to the fact that as $r \to 0$ the term $\frac{B_l}{r^{l+1}}$ would tend to infinity. We will find that the anything other than $l = 1$ leads to trivial solutions. Making this substitution immediately:

$$V_a(r,\theta) = Ar\cos\theta. \tag{B.7}$$

In the region of the shell we obtain:

$$V_b(r,\theta) = Br\cos\theta + \frac{C\cos\theta}{r^2}. \tag{B.8}$$

Outside the sphere, we must regain the applied field as $r \to \infty$, thus by similar arguments as found in Appendix A we get:

$$V_m(r,\theta) = -E_0 r\cos\theta + \frac{D\cos\theta}{r^2}. \tag{B.9}$$

The challenge now is to find the coefficients A-D. This is a straightforward task in linear algebra as there are four equations and four unknowns. The equations to be solved from the boundary conditions are:

$$A = B + \frac{C}{a^3} \tag{B.10}$$

$$B + \frac{C}{b^3} = -E_0 + \frac{D}{b^3} \tag{B.11}$$

$$\epsilon_a A = \epsilon_b B - \epsilon_b \frac{2C}{a^3} \tag{B.12}$$

$$\epsilon_b B - \epsilon_b \frac{2C}{b^3} = -\epsilon_m E_0 - \epsilon_m \frac{2D}{b^3} \tag{B.13}$$

This system of equations are most easily solved using a reduced-row echelon command using a computer.

The coefficients are found to be (barring any transcription errors):

$$A = \frac{9\epsilon_m\epsilon_b b^3}{(-2b^3\epsilon_b^2 - 4\epsilon_m b^3\epsilon_b - 2a^3\epsilon_b\epsilon_a + 2a^3\epsilon_b^2 + 2a^3\epsilon_m\epsilon_a - 2a^3\epsilon_m\epsilon_b - b^3\epsilon_b\epsilon_a - 2b^3\epsilon_m\epsilon_a)}$$

$$B = \frac{3\epsilon_m b^3(2\epsilon_b + \epsilon_a)}{(-2b^3\epsilon_b^2 - 4\epsilon_m b^3\epsilon_b - 2a^3\epsilon_b\epsilon_a + 2a^3\epsilon_b^2 + 2a^3\epsilon_m\epsilon_a - 2a^3\epsilon_m\epsilon_b - b^3\epsilon_b\epsilon_a - 2b^3\epsilon_m\epsilon_a)}$$

$$C = \frac{3a^3b^3\epsilon_m(-\epsilon_a + \epsilon_b)}{(-2b^3\epsilon_b^2 - 4\epsilon_m b^3\epsilon_b - 2a^3\epsilon_b\epsilon_a + 2a^3\epsilon_b^2 + 2a^3\epsilon_m\epsilon_a - 2a^3\epsilon_m\epsilon_b - b^3\epsilon_b\epsilon_a - 2b^3\epsilon_m\epsilon_a)}$$

$$D = \frac{b^3(-2b^3\epsilon_b^2 + 2\epsilon_m b^3\epsilon_b - 2a^3\epsilon_b\epsilon_a + 2a^3\epsilon_b^2 - a^3\epsilon_m\epsilon_a + a^3\epsilon_m\epsilon_b - b^3\epsilon_b\epsilon_a + b^3\epsilon_m\epsilon_a)}{(-2b^3\epsilon_b^2 - 4\epsilon_m b^3\epsilon_b - 2a^3\epsilon_b\epsilon_a + 2a^3\epsilon_b^2 + 2a^3\epsilon_m\epsilon_a - 2a^3\epsilon_m\epsilon_b - b^3\epsilon_b\epsilon_a - 2b^3\epsilon_m\epsilon_a)}$$

We can now solve for the final form of the field:

$$\vec{E} = -\nabla V \Rightarrow \tag{B.14}$$

$$\vec{E}_a\left(r,\theta\right) = -\vec{E_0}\cdot\left(A\cos\theta\right)\hat{r} + \vec{E_0}\cdot\left(A\sin\theta\right)\hat{\theta} \;\text{ for } r < a \tag{B.15}$$

$$\vec{E}_b\left(r,\theta\right) = \vec{E_0}\cdot\cos\theta\left[-B + \frac{2C}{r^3}\right]\hat{r} + \vec{E_0}\cdot\sin\theta\left[B + \frac{C}{r^3}\right]\hat{\theta} \;\text{ for } a < r < b \tag{B.16}$$

$$\vec{E}_m\left(r,\theta\right) = \vec{E_0}\cdot\cos\theta\left[1 + \frac{2D}{r^3}\right]\hat{r} + \vec{E_0}\cdot\sin\theta\left[-1 + \frac{D}{r^3}\right]\hat{\theta} \;\text{ for } r > b \tag{B.17}$$

APPENDIX C

TIME STAMPING CIRCUIT

The AFM tip oscillation signal is output via the AFM control box. The actual signal monitored is the cantilever deflection signal. This has been deliberately chosen (as opposed to the drive signal) to account for any deviations in the tip-oscillation trajectory. Under most imaging conditions in air, the actual probe oscillation is extremely sinusoidal, and consequently so is the deflection output signal. For non sinusoidal operating conditions, tip oscillations can be time stamped by continuously recording the deflection signal with a ADC.

A homebuilt circuit is used as part of time-stamping the tip-oscillation signal. The sine wave deflection signal is converted into a 0-5 V square wave in order to be compatible with a data acquisition card (NI USB-6210). First the sine wave is amplified in three stages using AD8002 op amps. This is done to account for the large variations in sine-wave amplitude that can be used in the system: the output is saturated at 5 volts regardless of the input signal. The resulting signal from the amplification stages is then put through a fast discriminator (AD8561), to yield the desired output. The final output must be buffered so that it is not pulled down, this is done with a $25-\Omega$ octal buffer/driver (SN6425244NT). The complete circuit diagram is shown in Figure C.1

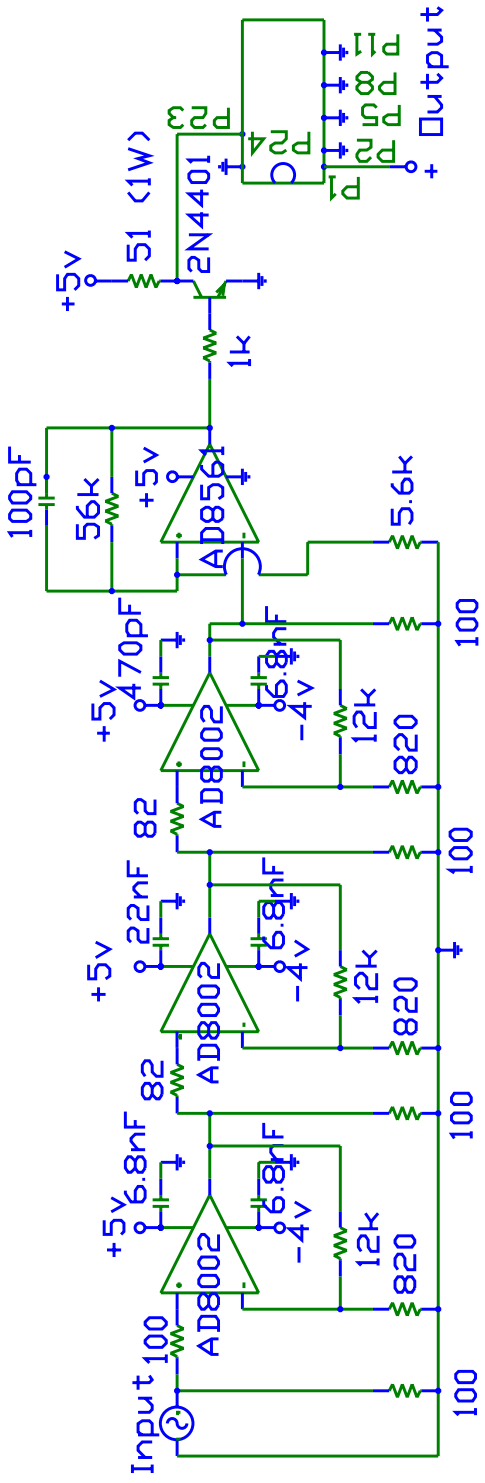

Figure C.1. Time-stamping circuit diagram for tip-oscillations: it converts a sinusoidal deflection signal into a 0-5 V square wave. Note: place 1nF capacitors between +5V supplies and ground.

APPENDIX D

DETAILED SMOOTH COVERSLIP CLEANING PROCEDURE

Last Modified June 5, 2009

1. Handling with ethanol-cleaned tweezers, rinse rack thoroughly with ultra pure water. Placing rack on rinsed lid, fill with round glass coverslips and lower gently into rinsed dish. If using closely spaced rack, place slips in every other slot.

2. Cover rack and slips with 2% Micro-90 solution, sonicate for 30 minutes and rinse thoroughly in ultra pure water.

    - To rinse thoroughly remove teflon racks and place them on rinsed lid, rinse crystallization dish, place teflon racks back into crystallization dish, fill dish (with racks) with ultra pure water. Repeat $\sim$ 3 times.

3. Sonicate in NaOH: fill crystallization dish with 30% (w/w) NaOH.

    - To create 30% NaOH solution from pellets:
        - Weigh 100 g NaOH in each of 2 tared weigh boats.
        - Mix 200 g NaOH with 467 mL ultra pure water. Place container with water on stir plate, add mixing bar, and then slowly add pellets while stirring, keeping heat reaction in check.
        - Make sure the tops of the coverslips are covered - if not add a little more ultra pure water until they are.
        - Don't fill the dish too full - be sure not to spill any NaOH into the sonicator.

4. Sonicate at room temperature for 1 hour.

5. Rinse well in ultra pure water. Always handle rack with clean tweezers and hold rack on clean dish lid while rinsing rack and dish separately.

6. Sonicate in HCl. Repeat steps 4-5 using 1 N HCl instead of NaOH.

    - To create 1 N HCl from stock 12.1 N (8.26% solution): Add 190 mL HCl to 2109 mL ultra pure water (in that order).

7. Sonicate in ultra pure water. Repeat steps 4-5 using only ultra pure water.

8. After final rinsing, store rack with cleaned slips in ultra pure water and cover with lid until ready for use. Store in laminar flow hood.

APPENDIX E

CONJUGATING LATEX BEADS TO AMINE COATED SLIDES

Last updated May 13, 2008

1. Coating slides with an Amine group:

   - Make a mixture of 2 mM APTES in dry n-Hexane: Add 15 mL of dry n-Hexane and 7 μL of 3-aminopropyltriethoxisilane (APTES) to a 15 mL tube. Seal tightly and mix by inverting for a couple of minutes. For larger volumes use 11.75 μL APTES in 25 mL Hexane, or 23.5 μL APTES in 50 mL Hexane.
   - With the previously cleaned slides laying flat in a cover dish, add the Hexane/APTES solution until the slides are completely submerged. Keep at room temperature for $\sim$ 2 hrs. (Don't wait too long - 2 hours is the upper limit, 30 minutes is probably ok).
   - Rinse the slides in toluene once, then with methanol. Place in a glass rack and bake (while still covered) in oven (JFB 326) at 150° C for more than 30 minutes (overnight is ok). After removing the coverslips from the oven allow them to cool and store in methanol for up to one week.
   - Air dry in the laminar flow hood immediately before use.

2. Making the latex bead dilution:

   - Dilute the latex beads (Nile Red) to 10-5 (not absolute concentration but relative to how they come) with ultra pure water. Add 1 mL of the latex bead dilution to 9 mL of MES buffer (50 mM, pH $\sim$ 6.1) so the

concentration of the beads is now 10-6. Add 20 mg of EDAC to the latex bead solution and vortex until mixed well.

3. Applying the beads to the slides:

   - The amine coated slides from Step 2 should be placed laying flat in a cover dish. Transfer $\sim$ 500 μL of the latex bead solution from Step 3 on the top surface of each slide. Cover the dish and leave at room temperature for more than 4 hours (overnight is ok). Gently rinse with ultra pure water 3 times by dipping the slide in the water and gently waving it. Air dry in the laminar flow hood.

APPENDIX F

GROWTH OF NANOTUBES

Notes on Preparation:

- Use latex or equivalent gloves for all activities involving the handling, removal or storage of wafers. Also use gloves to handle furnace tube and sample holder.
- Entire procedure should be done under the cleanest conditions possible. Use laminar flow hood for cutting, cleaning and catalyst procedures.

Catalyst bottle cleaning:

1. After emptying both the 25 mL and 50 mL screwtop bottles of their old catalyst solution, rinse thoroughly first with pure water and then again with ultra pure water (making sure to rinse the outsides, rims, and caps of both bottles).
2. Fill the 25 mL and 50 mL bottles with ultra pure water and sonicate for at least 15 minutes.
3. Empty both bottles and fill with electronic-grape isopropyl alcohol (IPA). Sonicate for another 15 minutes.

Cutting:

1. While catalyst bottles are sonicating, cut desired number of wafers using a razor blade and ruler. Wafers that will stand vertically in the holder must be at least 12.5 mm $\times$ 12.5 mm. Wafers that will lie flat in holder must be 10 mm $\times$ 10 mm.
2. Label each wafer with diamond scribe.

Cleaning:

1. For all steps of the cleaning process, make sure to thoroughly rinse all glassware with ultra pure water before and after each use, specifically when switching between two different chemicals.

2. Before loading wafers into tray/holder, blow each wafer with nitrogen to remove any large dust particles.

3. Rinse beaker/dish with a small amount of toluene and discard into non-halogenous waste container. Fill beaker/dish with toluene to an appropriate level (enough to completely cover wafers when submerged).

4. Rinse wafers (and holder) with methanol and place in toluene solution. Sonicate wafers in toluene for 15 minutes.

5. Rinse another beaker/dish with a small amount of acetone and discard into non-halogenous waste container. Fill beaker/dish with acetone (enough to completely cover wafers when submerged).

6. Remove wafers from toluene solution and immediately rinse with acetone. Place wafers in acetone solution and sonicate for another 15 minutes.

7. Rinse another beaker/dish with a small amount of (IPA) and discard into non-halogenous waste container. Fill beaker/dish with IPA (enough to completely cover wafers when submerged).

8. Remove wafers from acetone solution and immediately rinse with IPA. Place wafers in IPA solution and sonicate for another 15 minutes.

9. Clean tweezers with ultra pure water.

Catalyst:

1. During wafer cleaning, empty the two catalyst bottles of the sonicated electronic-grade IPA into non-halogenous waste container.

2. Fill both bottles with 25 mL of electronic-grade IPA.

3. Weigh out approximately 250 mg of ferric nitrate on weigh paper (avoiding big crystals), making sure to clean spatula before and after use with ultra pure water and an IPA rinse.

4. Deposit ferric nitrate into 50 mL bottle and swirl (instead of shake) well until no more crystals are visible in solution.

5. Dilute solution by extracting 25 μL with pipette from 50 mL bottle and inserting into 25 mL bottle. Swirl thoroughly to mix.

Spinning:

1. Before catalyst application, obtain and thoroughly clean a clear, plastic box, lining the bottom with Kodak paper.

2. Clean platform with acetone and kimwipe before use.

3. Cover surrounding surface with large kimwipe, tearing a hole in the center for spin platform.

4. Without anything on the platform, press "Power, Vacuum, Start".

5. Adjust dial to 3.00 krpm and timer to 900+ seconds.

6. Press "Stop, Vacuum" when finished adjusting.

Coating:

- Note: if one of the wafers will be lying flat during growth, remove from IPA solution, allow to air dry, deposit enough catalyst solution to completely cover wafer surface, and place directly into box. Do not spin dry.

1. Place first wafer on platform.

2. Press "Vacuum, Start".

3. While spinning, deliver 4 × 50 μL of catalyst per wafer with pipette. This is equivalent to approximately 4 × 4 drops, with intervals of 5+ seconds. Wait until wafer has dried before applying another drop of catalyst.

4. When catalyst application is complete, press "Stop, Vacuum," remove wafer from platform and place in box.

5. Repeat application process for all other wafers.

6. Clean tweezers with ultra pure water.

Baking:

1. Prep:

    - Turn on argon, hydrogen, and methane gases.
    - Assemble gas-flow system and run argon, hydrogen, and methane for 5 minutes, setting to desired flow rates (on flow meter, argon should read "100", hydrogen "50", and methane "4". These numbers correspond to 440 mL/min, 125 mL/min, and 1080 mL/min respectively).
    - Turn off hydrogen and methane at flow meter.
    - Load wafers into holder in number order upstream to downstream.
    - Gently slide the holder into quartz tube, noting where holder is positioned in furnace.
    - Reassemble gas-flow system and insert thermocouple, noting where the thermocouple is positioned within the furnace.

2. Flush argon at 440 mL/min through system for 15 minutes.

3. While argon is running, stuff furnace ends with Fiberfrax insulation.

4. Add hydrogen at 125 mL/min.

5. Turn on and direct fans toward two o-rings. Keep on throughout entire process to protect o-rings from melting.

6. Heat furnace to 900° C using "Set Point" dial.

   Note alt. temp:

   Note time from room temp to desired temp:

7. Hold at desired temp for 15 minutes with same flow rates.

   Note control TC temp:

   Note read TC temp:

8. Turn off hydrogen and flush with argon only for 5 minutes.

9. Turn on methane and hydrogen simultaneously and then immediately turn off argon flow. Methane flow rate should be 1080 mL/min. Growth time can vary between 30 seconds and 5 minutes.

   Note growth time used:

   Note temps of read and control TCs after methane was turned on: read TC: control TC:

10. Turn on argon and hydrogen and then immediately turn off methane. Flush system with argon and hydrogen for 5 minutes (holding temp constant).

Cooldown:

1. Dial down "Set Point" dial to zero. Unplug controller.

2. Remove insulation from both ends of furnace.

3. Blow cool air from right end of furnace until both thermocouples read $< 150°$ C.

4. Turn off hydrogen.

5. Turn off two fans.

Removal:

1. Remove thermocouple from quartz tube.
2. Remove holder and wafers from tube.
3. Place wafers back in box.

Shutdown:

1. Run argon through system for 5 minutes.
2. Remove left tubing from quartz end and replace with rubber stopper.
3. Remove right tubing from the other quartz end and replace with rubber stopper.
4. Turn off argon gas at flow meter.
5. Turn off argon, hydrogen, and methane gases at tanks.
6. Vacuum any spilled insulation and clean workstation.

APPENDIX G

ALIGNING THE TEFM SYSTEM

Fine alignment of our custom built optical set-up is required for good results. This appendix describes the fine alignment of the system, which should be performed frequently—ideally before each use, but definitely after switching between excitation paths. It is assumed that the user is familiar with standard best practices for handling and adjusting optical components. This appendix also does not speak to the course alignment procedure required for building the system.

The primary concern is that the excitation laser be centered along the optical axis, as defined by the microscope objective, meaning that the excitation beam should be both at normal incidence and centered on the back aperture of the objective. We have developed a standardized procedure in the lab for ensuring that this condition is met.

The primary component of this alignment protocol is a custom built adapter that screws into the microscope objective seat. This adapter has male Nikon threads and female SM1 threads compatible with one inch diameter lens tube made by Thorlabs. Bulls-eye style targets are printed on transparency film such that they fit snugly into the lens tube. The objective is removed and the adapter with a length of lens tube is mounted in its place.

First a short section of lens tube (3") is used with a target, to verify that the excitation beam is striking the center. Should this not be the case, a mirror is adjusted (#1 or #2 as indicated in Figure G.1 depending on the path being utilized) until it is centered. Next, a much longer section of lens tube is added (9") with the target being placed at the top end. Again, it is verified that the beam strikes the center of the target. If not, a different mirror (beamsplitter) is adjusted (#3) until the beam is in the center. The adjustment of the mirrors is done such

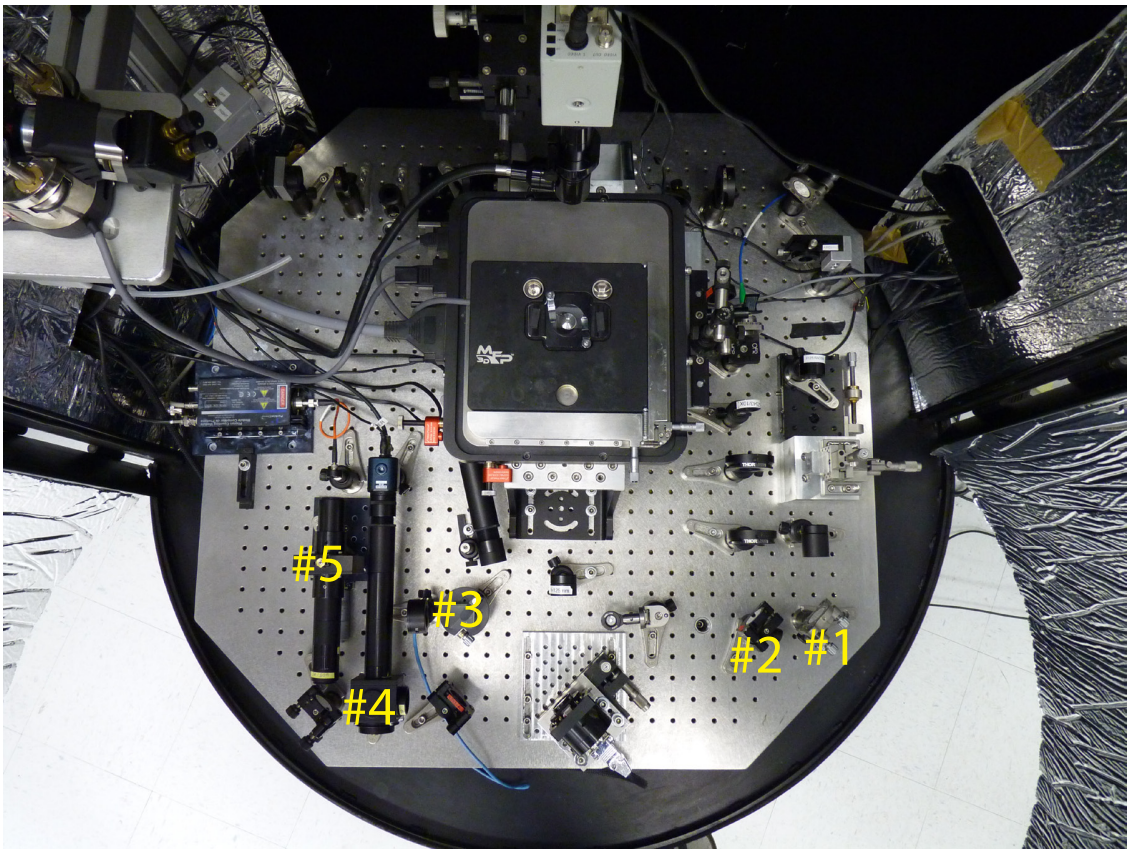

Figure G.1. Numbered photo of TEFM system

that when using the short lens tube the far mirror is adjusted, and when using the long lens tube, the closer mirror is adjusted (near, and far referring to the distance from the objective mount). This process is repeated several times until no further adjustments are necessary by either mirror. By verifying that the excitation beam is centered at two locations ensures that it will also enter the objective at normal incidence.

As described in Chapter 3, a beam pickoff sends 4% of collected light in the emission path to a camera. This camera is used to monitor the focus of the excitation beam on the sample. Once the bulls-eye procedure is complete, the objective is put in place and a clean glass coverslip is mounted. Minor adjustments to the beam pickoff are made so that the resulting image of the focused excitation beam in centered within the field of view. By defocusing the excitation beam,

the quality of the excitation mode can be examined. Should there be some light asymmetries, mirrors #1/#2 and #3 are readjusted to correct for this, while keeping the focal spot carefully centered within the field of view of the camera.

The final element to be adjusted is the detector. The detector used is a single pixel avalanche photodiode, the particular version in use is fibercoupled with a multimode fiber. The free end of the fiber is fixed to the table with a standard optics mount. The lens that focuses onto the fiber is mounted on a three dimensional translation stage (#5).

Importantly, the detector is adjusted with the appropriate spectral filters in place. A fluorescent marker is used to draw on the clean coverslip, the objective is focused. A custom built Labview program displays a trace of the real-time count rate of the APD as adjustments are made. The fine position of the final lens is adjusted until the count-rate is maximized along all three axes. After each axis is maximized, the process is repeated until no further adjustments are necessary. The system is now aligned and ready for use.

REFERENCES

[1] T. J. Yang, G. A. Lessard, and S. R. Quake. An apertureless near-field microscope for fluorescence imaging. *Applied Physics Letters*, 76(3):378–380, 2000.

[2] J. M. Gerton, L. A. Wade, G. A. Lessard, Z. Ma, and S. R. Quake. Tip-enhanced fluorescence microscopy at 10 nanometer resolution. *Physical Review Letters*, 93(18), 2004.

[3] Z. Y. Ma, J. M. Gerton, L. A. Wade, and S. R. Quake. Fluorescence near-field microscopy of dna at sub-10 nm resolution. *Physical Review Letters*, 97(26), 2006.

[4] L. D. Landau and E. M. Lifshitz. *The Classical Theory of Fields (Course of Theoretical Physics; v.2)*. Elsevier, Burlington, fourth revised english edition edition, 2004.

[5] S. W. Hell. Far-field optical nanoscopy. *Science*, 316:1153–1158, 2007.

[6] C. Mu, B. D. Mangum, C. Xie, and J. M. Gerton. Nanoscale fluorescence microscopy using carbon nanotubes. *IEEE Journal of Selected Topics in Quantum Electronics*, 14(1):206–216, 2008.

[7] C. A. Xie, C. Mu, J. R. Cox, and J. M. Gerton. Tip-enhanced fluorescence microscopy of high-density samples. *Applied Physics Letters*, 89(14), 2006.

[8] B. D. Mangum, C. Mu, and J. M. Gerton. Resolving single fluorophores within dense ensembles: Contrast limits of tip-enhanced fluorescence microscopy. *Optics Express*, 16(9):6183–6193, 2008.

[9] J. D. Jackson. *Classical Electrodynamics*. John Wiley & Sons, Hoboken, third edition, 1999.

[10] L. Novotny. Near-field optical characterization of nanocomposite materials. *J. Am. Ceram. Soc.*, 85(5):1057–1060, 2002.

[11] J. L. Bohn, D. J. Nesbitt, and A. Gallagher. Field enhancement in apertureless near-field scanning optical microscopy. *Journal of the Optical Society of America A*, 18(12):2998–3006, 2001.

[12] P. Anger, P. Bharadwaj, and L. Novotny. Enhancement and quenching of single-molecule fluorescence. *Physical Review Letters*, 96(11), 2006.

[13] N. A. Issa and R. Guckenberger. Fluorescence near metal tips: The roles of energy transfer and surface plasmon polaritons. *Optics Express*, 15:12131–12144, 2007.

[14] T. H. Taminiau, F. D. Stefani, F. B. Segerink, and N. F. Van Hulst. Optical antennas direct single-molecule emission. *Nature Photonics*, 2(4):234–237, 2008.

[15] K. L. Kelly, E. Coronado, L. L. Zhao, and G. C. Schatz. The optical properties of metal nanoparticles: The influence of size, shape, and dielectric environment. *Journal of Physical Chemistry B*, 107(3):668–677, 2003.

[16] L. Novotny, R. X. Bian, and X. S. Xie. Theory of nanometric optical tweezers. *Physical Review Letters*, 79(4):645–648, 1997.

[17] C. Kittel. *Introduction to Solid State Physics*. John Wiley Sons, Hoboken, N.J., eigth edition, 2005.

[18] L. Novotny and S. J. Stranick. Near-field optical microscopy and spectroscopy with pointed probes. *Annual Review of Physical Chemistry*, 57:303–331, 2006.

[19] W. L. Barnes, A. Dereux, and T. W. Ebbesen. Surface plasmon subwavelength optics. *Nature*, 424(6950):824–830, 2003.

[20] S. A. Maier and H. A. Atwater. Plasmonics: Localization and guiding of electromagnetic energy in metal/dielectric structures. *Journal of Applied Physics*, 98(1), 2005.

[21] C. Sonnichsen, T. Franzl, T. Wilk, G. von Plessen, J. Feldmann, O. Wilson, and P. Mulvaney. Drastic reduction of plasmon damping in gold nanorods. *Physical Review Letters*, 88(7), 2002.

[22] C.C. Neacsu, G.A. Steudle, and M.B. Raschke. Plasmonic light scattering from nanoscopic metal tips. *Applied Physics B*, 80(3):295–300, 2005.

[23] S. Link and M. A. El-Sayed. Size and temperature dependence of the plasmon absorption of colloidal gold nanoparticles. *Journal of Physical Chemistry B*, 103(21):4212–4217, 1999.

[24] B.D. Mangum, E. Shafran, C. Mu, and J.M. Gerton. Three-dimensional mapping of near-field interactions via single-photon tomography. *Nano Letters*, 9(10):3440–3446, 2009.

[25] W. Trabesinger, A. Kramer, M. Kreiter, B. Hecht, and U. P. Wild. Single-molecule near-field optical energy transfer microscopy. *Applied Physics Letters*, 81(11):2118–2120, 2002.

[26] S. Kuhn, U. Hakanson, L. Rogobete, and V. Sandoghdar. Enhancement of single-molecule fluorescence using a gold nanoparticle as an optical nanoantenna. *Physical Review Letters*, 97(1), 2006.

[27] P. Bharadwaj and L. Novotny. Spectral dependence of single molecule fluorescence enhancement. *Optics Express*, 15(21):14266–14274, 2007.

[28] T. H. Taminiau, R. J. Moerland, F. B. Segerink, L. Kuipers, and N. F. van Hulst. Lambda/4 resonance of an optical monopole antenna probed by single molecule fluorescence. *Nano Letters*, 7(1):28–33, 2007.

[29] L. Novotny. Effective wavelength scaling for optical antennas. *Physical Review Letters*, 98(26), 2007.

[30] G. W. Bryant, F. J. G. De Abajo, and J. Aizpurua. Mapping the plasmon resonances of metallic nanoantennas. *Nano Letters*, 8(2):631–636, 2008.

[31] H. G. Frey, F. Keilmann, A. Kriele, and R. Guckenberger. Enhancing the resolution of scanning near-field optical microscopy by a metal tip grown on an aperture probe. *Applied Physics Letters*, 81(26):5030–5032, 2002.

[32] H. G. Frey, S. Witt, K. Felderer, and R. Guckenberger. High-resolution imaging of single fluorescent molecules with the optical near-field of a metal tip. *Physical Review Letters*, 93(20), 2004.

[33] L. Novotny. Single molecule fluorescence in inhomogeneous environments. *Applied Physics Letters*, 69(25):3806–3808, 1996.

[34] R. Carminati, J. J. Greffet, C. Henkel, and J. M. Vigoureux. Radiative and non-radiative decay of a single molecule close to a metallic nanoparticle. *Optics Communications*, 261:368–375, 2006.

[35] S. Vukovic, S. Corni, and B. Mennucci. Fluorescence enhancement of chromophores close to metal nanoparticles. optimal setup revealed by the polarizable continuum model. *The Journal of Physical Chemistry C*, 113(1):121–133, 2008.

[36] D. Hu, M. Micic, N. Klymyshyn, Y. Suh, and H. P. Lu. Correlated topographic and spectroscopic imaging beyond diffraction limit by atomic force microscopy metallic tip-enhanced near-field fluorescence lifetime microscopy. *Review of Scientific Instruments*, 74(7):3347–3355, 2003.

[37] B. Knoll and F. Keilmann. Near-field probing of vibrational absorption for chemical microscopy. *Nature*, 399:134–137, 1999.

[38] H. Gersen, M. F. García-Parajó, L. Novotny, J. A. Veerman, L. Kuipers, and N. F. van Hulst. Influencing the angular emission of a single molecule. *Physical Review Letters*, 85(25):5312, 2000.

[39] C. Girard, O.J.F. Martin, and A. Dereux. Molecular lifetime changes induced by nanometer scale optical fields. *Physical Review Letters*, 75(17):3098–3101, 1995.

[40] F. H'Dhili, R. Bachelot, A. Rumyantseva, G. Lerondel, and P. Royer. Nanopatterning photosensitive polymers using local field enhancement at the end of apertureless snom tips. *Journal of Microscopy*, 209(3):214–222, 2003.

[41] V. V. Protasenko, A. Gallagher, and D. J. Nesbitt. Factors that influence confocal apertureless near-field scanning optical microscopy. *Optics Communications*, 233(1-3):45–56, 2004.

[42] S. Kühn and V. Sandoghdar. Modification of single molecule fluorescence by a scanning probe. *Applied Physics B*, 84:211–217, 2006.

[43] T. Kalkbrenner, M. Ramstein, J. Mlynek, and V. Sandoghdar. A single gold particle as a probe for apertureless scanning near-field optical microscopy. *Journal of Microscopy-Oxford*, 202:72–76, 2001. Part 1.

[44] L. Eligal, F. Culfaz, V. McCaughan, N. I. Cade, and D. Richards. Etching gold tips suitable for tip-enhanced near-field optical microscopy. *Review of Scientific Instruments*, 80(3):3, 2009.

[45] S. Adachi. Model dielectric constants of si and ge. *Physical Review B*, 38(18):12966, 1988.

[46] V. V. Protasenko, M. Kuno, A. Gallagher, and D. J. Nesbitt. Fluorescence of single zns overcoated cdse quantum dots studied by apertureless near-field scanning optical microscopy. *Optics Communications*, 210(1-2):11–23, 2002.

[47] J. N. Farahani, D. W. Pohl, H. J. Eisler, and B. Hecht. Single quantum dot coupled to a scanning optical antenna: A tunable superemitter. *Physical Review Letters*, 95(1):4, 2005.

[48] E. D. Palik. *Handbook of Optical Constants of Solids*. Academic Press, Orlando, 1985.

[49] D. F. Edwards and E. Ochoa. Infrared refractive index of diamond. *J. Opt. Soc. Am.*, 71(5):607–608, 1981.

[50] L. G. Schulz. The optical constants of silver, gold, copper, and aluminum. i. the absorption coefficient k. *J. Opt. Soc. Am.*, 44(5):357–362, 1954.

[51] D. Davazoglou, G. Pallis, V. Psycharis, M. Gioti, and S. Logothetidis. Structure and optical properties of tungsten thin films deposited by pyrolysis of w(co)[sub 6] at various temperatures. *Journal of Applied Physics*, 77(11):6070–6072, 1995.

[52] S. Wolf and R.N. Taubner. *Silicon Processing for the Vlsi Era*. Process Technology. Lattice Press, Sunset Beach, 1986.

[53] P. Bharadwaj, P. Anger, and L. Novotny. Nanoplasmonic enhancement of single-molecule fluorescence. *Nanotechnology*, 18(4), 2007.

[54] E. Hecht. *Optics*. Addison Wesley, San Francisco, fourth edition, 2002.

[55] M. Stalder and M. Schadt. Linearly polarized light with axial symmetry generated by liquid-crystal polarization converters. *Optics Letters*, 21(23):1948–1950, 1996.

[56] L. Novotny and B. Hecht. *Principles of Nano-Optics*. Cambridge University Press, Cambridge, 2006.

[57] M. W. Davidson. Evanescent field polarization and intensity profiles, February 2010.

[58] M. Born and E. Wolf. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, New York, seventh edition, 2002.

[59] S. Quabis, R. Dorn, M. Eberler, O. Glöckl, and G. Leuchs. Focusing light to a tighter spot. *Optics Communications*, 179(1-6):1–7, 2000.

[60] R. Dorn, S. Quabis, and G. Leuchs. Sharper focus for a radially polarized light beam. *Physical Review Letters*, 91(23):233901, 2003.

[61] H.F. Hamann, M. Kuno, A. Gallagher, and D. J. Nesbitt. Molecular fluorescence in the vicinity of a nanoscopic probe. *Journal of Chemical Physics*, 114(19):8596–8609, 2001.

[62] A. Hartschuh, E.J. Sanchez, X. S. Xie, and L. Novotny. High-resolution near-field raman microscopy of single-walled carbon nanotubes. *Physical Review Letters*, 90(9):095503–1 – 095503–4, 2003.

[63] E. J. Sanchez, L. Novotny, G. R. Holtom, and X. S. Xie. Room-temperature fluorescence imaging and spectroscopy of single molecules by two-photon excitation. *Journal of Physical Chemistry A*, 101(38):7019–7023, 1997.

[64] E. J. Sanchez, L. Novotny, and X. S. Xie. Near-field fluorescence microscopy based on two-photon excitation with metal tips. *Physical Review Letters*, 82(20):4014–4017, 1999.

[65] R. Hillenbrand, F. Keilmann, P. Hanarp, D. S. Sutherland, and J. Aizpurua. Coherent imaging of nanoscale plasmon patterns with a carbon nanotube optical probe. *Applied Physics Letters*, 83(2):368–370, 2003.

[66] B. Knoll and F. Keilmann. Enhanced dielectric contrast in scattering-type scanning near-field optical microscopy. *Optics Communications*, 182(4-6):321–328, 2000.

[67] R. Hillenbrand and F. Keilmann. Complex optical constants on a subwavelength scale. *Physical Review Letters*, 85(14):3029, 2000.

[68] P. G. Gucciardi and G. Bachelier. Far-field background suppression in tip-modulated apertureless near-field optical microscopy. *Journal of Applied Physics*, 99(12):9, 2006.

[69] F. Keilmann. Scattering-type near-field optical microscopy. *Journal of Electron Microscopy*, 53(2):187–192, 2004.

[70] J. W. Strutt. On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. *Philosophical Magazine*, X:73–78, 1880.

[71] R. Hillenbrand and F. Keilmann. Material-specific mapping of metal/semiconductor/dielectric nanosystems at 10 nm resolution by backscattering near-field optical microscopy. *Applied Physics Letters*, 80(1):25–27, 2002.

[72] P. Mühlschlegel, H. J. Eisler, O. J. F. Martin, B. Hecht, and D. W. Pohl. Resonant optical antennas. *Science*, 308:1607–1609, 2005.

[73] M. Fleischmann, P. J. Hendra, and A. J. McQuillan. Raman spectra of pyridine adsorbed at a silver electrode. *Chemical Physics Letters*, 26(2):163–166, 1974.

[74] D. L. Jeanmaire and R. P. VanDuyne. Surface raman spectroelectrochemistry: Part i. heterocyclic, aromatic, and aliphatic amines adsorbed on the anodized silver electrode. *Journal of Electroanalytical Chemistry*, 84(1):1–20, 1977.

[75] M. G. Albrecht and J. A. Creighton. Anomalously intense raman spectra of pyridine at a silver electrode. *Journal of the American Chemical Society*, 99(15):5215–5217, 1977.

[76] K. Kneipp, Y. Wang, H. Kneipp, L. T. Perelman, I. Itzkan, R. R. Dasari, and M.S. Feld. Single molecule detection using surface-enhanced raman scattering (sers). *Physical Review Letters*, 78(9):1667–1670, 1997.

[77] J. T. Krug, E. J. Sanchez, and X. S. Xie. Design of near-field optical probes with optimal field enhancement by finite difference time domain electromagnetic simulation. *Journal of Chemical Physics*, 116(24):10895–10901, 2002.

[78] S. Kühn, U. Håkanson, L. Rogobete, and V. Sandoghdar. Enhancement of single-molecule fluorescence using a gold nanoparticleas an optical nanoantenna. *Physical Review Letters*, 97(1):017402, 2006.

[79] C. Höppener, R. Beams, and L. Novotny. Background suppression in near-field optical imaging. *Nano Letters*, 9(2):903–908, 2009.

[80] T. Grosges and D. Barchiesi. Tomography of the near-field optical signal. *Optics Letters*, 31(23):3435–3437, 2006.

[81] D. Barchiesi. Scanning near-field optical data contrast measurement: A tomographylike near-field reconstruction. *Applied Optics*, 45(29):7597–7601, 2006.

[82] H. Buiteveld, J. M. H. Hakvoort, and M. Donze. The optical properties of pure water. In J. S. Jaffe, editor, *SPIE Proceedings on Ocean Optics XII*, volume 2258, pages 174–183, 1994.

[83] R. M. Pope and E. S. Fry. Absorption spectrum (380–700 nm) of pure water. II. integrating cavity measurements. *Applied Optics*, 36:8710–8723, 1997.

[84] K.S. Shifrin. *Physical Optics of Ocean Water*. American Institute of Physics, New York, english edition, 1988.

[85] C. Höppener and L. Novotny. Antenna-based optical imaging of single ca2+ transmembrane proteins in liquids. *Nano Letters*, 8(2):642–646, 2008.

[86] P. B. Johnson and R. W. Christy. Optical constants of the noble metals. *Phys. Rev. B*, 6(12):4370–4379, Dec 1972.

[87] Heinrich G. Frey, Jan Paskarbeit, and Dario Anselmetti. Tip-enhanced single molecule fluorescence near-field microscopy in aqueous environment. *Applied Physics Letters*, 94(24):241116, 2009.

[88] Joseph R. Lakowicz. Radiative decay engineering: Biophysical and biomedical applications. *Analytical Biochemistry*, 298(1):1 – 24, 2001.

[89] E. Dulkeith, A. C. Morteani, T. Niedereichholz, T. A. Klar, J. Feldmann, S. A. Levi, F. C. J. M. van Veggel, D. N. Reinhoudt, M. Möller, and D. I. Gittins. Fluorescence quenching of dye molecules near gold nanoparticles: Radiative and nonradiative effects. *Phys. Rev. Lett.*, 89(20):203002, 2002.

[90] E. Yoskovitz, D. Oron, I. Shweky, and U. Banin. Apertureless near-field distance-dependent lifetime imaging and spectroscopy of semiconductor nanocrystals. *The Journal of Physical Chemistry C*, 112(42):16306–16311, 2008.

[91] J. H. Hafner, C. Cheung, T. H. Oosterkamp, and C. M. Lieber. High-yield assembly of individual single-walled carbon nanotube tips for scanning probe microscopies. *The Journal of Physical Chemistry B*, 105:743–746, 2001.

[92] L. Wade, I. Shapiro, Z. Ma, S. Quake, and C. P. Collier. Single-molecule fluorescence and force microscopy employing carbon nanotubes. *Nanotechnology*, 3:317–320, 2003.

[93] L. Wade, I. Shapiro, Z. Ma, S. Quake, and C. P. Collier. Correlating afm probe morphology to image resolution for single-wall carbon nanotube tips. *Nano Letters*, 4(4):725–731, 2004.

[94] Y. Narui, D. M. Ceres, J. Chen, K. P. Giapis, and C. P. Collier. High aspect ratio silicon dioxide-coated single-walled carbon nanotube scanning probe nanoelectrodes. *Journal of Physical Chemistry C*, 113(16):6815–6820, 2009.

[95] D. Griffiths. *Introduction to Electrodynamics*. Prentice-Hall, Upper Saddle River, third edition, 1999.