Field enhancement in apertureless near-field scanning optical microscopy

John L. Bohn, D. J. Nesbitt, and A. Gallagher

JILA, National Institute of Standards and Technology, and University of Colorado, Boulder, Colorado 80309-0440

The near field of an apertureless near-field scanning optical microscopy probe is investigated with a multiple-multipole technique to obtain optical fields in the vicinity of a silicon probe tip and a glass substrate. The results demonstrate that electric field enhancements of >15 relative to the incident fields can be achieved near a silicon tip, implying intensity enhancements of several orders of magnitude. This enhancement arises both from the antenna effect of the elongated probe and from a proximity effect when the probe is near the substrate surface and its image dipoles play a role. © 2001 Optical Society of America

OCIS codes: 180.5810, 290.0290

1. INTRODUCTION

Near-field scanning optical microscopy (NSOM) opens new scientific frontiers with its promise of spectroscopically probing individual molecules. Not only will it be possible to determine *where* the molecule is, as in conventional atomic force microscopy, but also to understand *what* molecule it is and *how* it responds to its physical and chemical environment. One option for observing the spectrum of a single molecule is to disperse the molecules so sparsely on the surface that the microscope focus contains at most one molecule.[1,2] Studies of this type have yielded valuable and somewhat unexpected information on the fluorescence of single molecules as opposed to the average fluorescence of larger ensembles.[3] What is missing from this approach is the ability to probe a single molecule, or a portion of a large molecule, *in situ*, where it may be crowded in with many other molecules within the diffraction limit of visible probing light.

Conventional NSOM responds to this challenge by bringing a small-diameter aperture over the sample, usually a metal-clad optical fiber, through which either the illuminating or probe light must pass.[4] Theoretical resolution is limited to 30–50 nm in this case, however,, by the skin depth of the material cladding the probe tip,[5] while in practice the resolution is generally poorer than this. More recently a set of apertureless techniques, dubbed apertureless near-field scanning optical microscopy (ANSOM), overcomes this difficulty by replacing the optical fiber light pipe by a solid metal or semiconductor atomic force microscope probe tip.[6–9] The probe draws the electric field to its sharp tip like a lightning rod, providing a resolution of the order of nanometers, comparable to the tip's width.[6] In a further refinement, the sample may be illuminated by evanescent light generated by total internal reflection inside the glass substrate on which the sample sits. In this case virtually all the light scattered into the far field above the glass originates near the probe tip, thus significantly boosting sensitivity.[10]

To extract chemical information from ANSOM requires a spectroscopic technique that can probe a single mol-ecule's internal states. The leading proposals for extracting this information are fluorescence and Raman spectroscopies.[11–15] To make this spectroscopy work requires not only a localized source of light but a bright source of light, owing to the relative weakness of the oscillator strengths involved. Fortunately, several decades of research on surface-enhanced Raman scattering has indicated that nanometer-sized surface irregularities indeed produce enormous localized electric fields,[16] sufficient in fact to make single-molecule Raman spectroscopy feasible.[17,18]

In the context of apertureless metal tips illuminated from above the surface, large field enhancements were predicted as early as 1991 from electrostatic models[19] and verified again later in more realistic calculations.[20] The role of the frequency dependence of plasma resonances has also been investigated theoretically.[16,21] On the experimental side it has been found that the field enhancement possesses a strong dependence on the light's polarization relative to the probe's axis.[22] In addition, a scanning tunneling microscope measurement extracted a field enhancement of ~1000 for PtIr probes and graphite or gold samples.[23] In this work the enhancements were understood by using simple dipolar models of the tip. Large enhancements have also been inferred from experiments that use silicon tips.[11]

Recently another type of ANSOM experiment was reported that also measured a large field enhancement.[14] In this experiment the probe tip and sample were illuminated by an evanescent field generated by total internal reflection in a glass prism. Thus the entire probe tip was not bathed in incident illumination, as had been the case in previous studies. The probe (which was made of silicon) was scanned over various polystyrene spheres doped with dye molecules. The resulting fluorescence signal versus the tip-to-sample distance fit a simple model that replaced the probe with an effective point dipole at the center of a hemispherical tip. The only fitting parameter was the strength of this dipole, expressed as the enhancement of the electric field immediately beneath the tip relative to the evanescent electric field immediately above

the prism's surface, $\kappa \equiv E_{\text{tip}}/E_{\text{evan}}$. The results showed strikingly that the enhancement can be as large as $\kappa \approx 28$, implying an intensity enhancement of nearly one thousand.

In contrast to metal probes and samples with small finite skin depth, the silicon tip is essentially transparent to light, as is the glass substrate in the experiment of Ref. 14. The purposes of this paper are to explore the field enhancement in this novel situation and to verify that large enhancements are possible in this case as well. We do this quantitatively using numerical calculations on a model system, as well as qualitatively using a simple electrostatic model that exploits the fact that the tip is small compared with the wavelength of light. The simpler qualitative model allows us to explicitly identify the origin of the field enhancement as contributions arising from an antenna effect and a field dressing that occurs when the tip is within its own radius of curvature of the surface.

A number of problems in electromagnetic scattering theory relevant to ANSOM have been treated in the theoretical literature, yet none have quite approached the geometry of the experiment of Ref. 14. For example, prolate spheroids representing long, sharp objects have long been studied either numerically[24] or as formal extensions of Mie theory.[25] Similar solutions have had a great impact on our understanding of surface-enhanced Raman scattering.[16] Another set of solutions emphasizes scattering of light from spheres or other objects near a surface but illuminated by plane waves.[26–28] These solutions are performed by several techniques, including multiple expansions and image scatterers. These studies tend to focus on properties of the far-field scattered radiation. In addition, discrete-dipole calculations of oblate metal spheroids on surfaces have verified large near-field enhancements.[29] Specific experiments have also been modeled by assuming spherical tips, for example, Ref. 30, which models the optical contrast of the experiment.

Scattering by evanescent waves has also been considered in the case of isolated spheres[31] or small clusters of spheres representing an ANSOM tip.[32] These results have shed much light on the role of distance-dependent plasmon resonances as well as on practical issues of resolution and scattered light intensity. They have not, however, sought to address the near-field enhancement of the electric field. Nor is it completely clear how accurately results from a collection of spheres converge to a solid probe tip.

A few numerical studies have also approached specific ANSOM geometries. In one, a new technique is proposed that employs illumination by a Hermite–Gaussian laser beam intended to facilitate two-photon fluorescence spectroscopy.[9] In another, a wedge-shaped tungsten probe tip was found in calculations to give intensity enhancements of several hundred when illuminated by plane waves.[20] While this result is encouraging, it is not clear how applicable it is to the transparent tip in the experiment in Ref. 14. In addition, the enhancement generated by plane waves was found to grow linearly with the volume of the probe, whereas in an evanescent wave most of the tip is not directly illuminated. Additionally, this purely numerical solution was not able to separate out contributions from the antenna and proximity effects.

It is therefore worthwhile to consider a model more closely approximating the conditions of Ref. 14. That is, we incorporate a surface, an incident evanescent wave, and an elongated probe. We model the probe by a prolate spheroid with a length that we are free to vary. In Section 2 we detail this model and its numerical solution using the multiple-multipole (MMP) method.[33] We also briefly treat this model in an electrostatic approximation that underscores the basic physics issues involved in enhancement. In Section 3 we illustrate the numerical model with useful examples that show the size of the field enhancement and its origin. These results also illustrate that resolution is predominantly limited by the size (i.e., the radius of curvature) of the probe tip.

2. THEORY

A. Computational Model

Calculating the near field is fundamentally an exercise in classical electromagnetic scattering theory. Thus the probe and the glass prism are simply represented as linear, isotropic, homogeneous dielectrics with dielectric constants ϵ and ϵ_g, respectively. We ignore the frequency dependence of these constants, since the incident light is assumed to be monochromatic. For the calculations reported below we take $\epsilon = 19$ and $\epsilon_g = 2.25$ representing visible radiation with a silicon tip and a glass prism, respectively. We have disregarded the imaginary part of ϵ, since it is small (0.63) and leads to a skin depth in the silicon of ~ 1100 nm, far larger than any tip modeled numerically in this paper.

The geometry of our model is illustrated in Fig. 1. The prism is represented as a semi-infinite dielectric slab occupying the space $z < 0$. We model the probe tip as a prolate spheroid with semimajor axis a and semiminor axis $b < a$. Using a smooth, regular shape such as a spheroid greatly simplifies numerical calculations. Notice that the detailed shape of the upper part of the probe is not extremely important, since we are interested in the near field, i.e., within several tens of nanometers near the tip, while the actual atomic force microscope probe tip in the experiment is approximately a narrowly tapered cone of several micrometers in length. We will examine the a dependence of the near field below; for the comparison between different values of a to be meaningful, the spheroids are normalized by fixing the radius of curvature, $R_c = b^2/a$, at the tip. Below we show results for two sizes of tip, $R_c = 10$ nm and $R_c = 5$ nm, which are typical of ANSOM experiments. In either case R_c remains small compared with the wavelength of the incident light. Finally, we let d denote the distance from the probe tip to the surface.

The illuminating light, with vacuum wavelength λ, is incident from within the glass prism at an angle of incidence θ_{inc} greater than the critical angle for total internal reflection. The wave vector of the incident light (indicated in Fig. 1) and that of the internally reflected light determine the scattering plane, which is taken to be the $x–y$ plane in the model. The light is polarized with its electric field in this plane, so that it will lie along rather
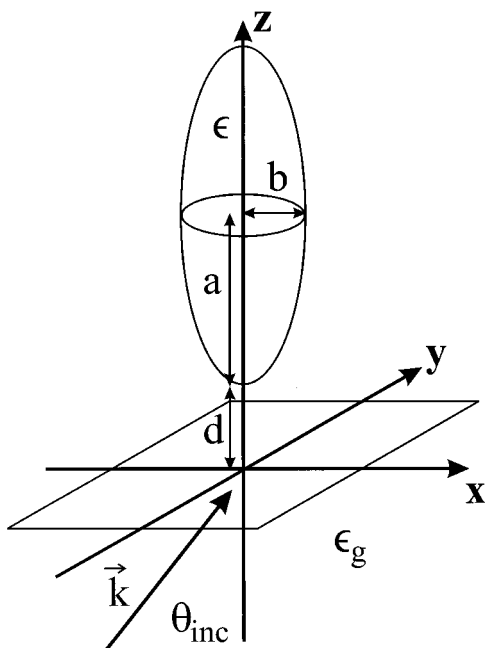
Fig. 1.   Geometry of the model.   The probe tip is modeled by a prolate spheroid with semimajor and semiminor axes a and b whose lower tip is displaced a distance d above the glass surface. The tip and glass have dielectric constants ϵ and ϵ_g, respectively. The wave vector **k** of the incident light lies in the x–z plane at an angle of incidence greater than the critical angle for total internal reflection.

than across the probe; this enhances the antenna effect.[20] In the calculations below we choose $\lambda = 543$ nm and $\theta_{\text{inc}} = 60°$; the evanescent wave's intensity then falls off exponentially in the vacuum with a $1/e$ length of $l_0 = 52$ nm.   In the absence of the probe tip the electric field can be conceptually broken into three components,

$$\mathbf{E}_0 = \mathbf{E}_{\text{inc}} + \mathbf{E}_{\text{ref}} + \mathbf{E}_{\text{evan}}, \qquad (1)$$

denoting, respectively, incident and totally internal reflected plane wave in the glass and the evanescent wave in the vacuum above.   These fields are, of course, analytically represented by elementary means.[34]

To compute the influence of the probe on the near field, we must solve Maxwell's equations and the vector Helmholtz equations for the electric and magnetic fields, subject to the usual boundary conditions on the surfaces of the prism and the probe.   In principle the fields can be expanded into a suitable set of electromagnetic multipoles[35] centered at a convenient location (e.g., the center of our spheroid), with coefficients determined to satisfy the boundary conditions on both the spheroid and the planar surface.   However, this is numerically difficult, since the boundaries are far from having the spherical symmetry to which the multipoles are adapted, and the number of coefficients required would be enormous.

A stable numerical alternative is to place small sets of multipole expansions at a carefully chosen set of different expansion centers.   This technique is therefore known as the MMP method.[33] The expansion of the electric field takes the generic form

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_0(\mathbf{r}) + \sum_k c_k \mathbf{f}_k(\mathbf{r} - \mathbf{r}_k), \qquad (2)$$

with a similar expansion for the magnetic field.   Here the functions $\mathbf{f}_k$ represent analytical solutions to the Maxwell/Helmholtz equations, in our case represented by multipoles centered at the various locations $\mathbf{r}_k$.   The radial dependence of each multipole function is given by an appropriate spherical Bessel function $b_l(kr)$, where k denotes the wave number.   Boundary conditions for the multipole are specified by the type of Bessel function chosen.   If b_l is a regular spherical Bessel function j_l, then the corresponding multipole does not diverge at its origin. This kind of multipole is placed at the center of the spheroid to describe the fields inside the spheroid.   An alternative boundary condition is to set $b_l = h_l^+$, an outbound spherical Hankel function.   This kind of function is appropriate to describe, e.g., the scattered waves outside of the spheroid, but must then be placed inside the spheroid to avoid the singularity that it possesses at $r = 0$.   With the sole exception of the regular multipole describing the field inside the spheroid, all fields in any of the dielectric media are generally represented by multipole expansions centered outside that dielectric.

The coefficients c_k are determined by applying boundary conditions in a least-squares sense on a discrete set of matching points $\{\mathbf{s}_\alpha\}$ on the boundaries between dielectrics.   This procedure yields a system of linear equations for the c_k's.   Typically the number of matching points chosen is much larger than the number of unknown coefficients, so that these equations are overdetermined. The choice of matching points and multipole expansions must be made carefully, bearing in mind several issues. The multipole orders must be large enough to represent variations in the field, but not large enough that the multipoles oscillate between one matching point and the next. Moreover, the overdetermination of the system of linear equations must be kept in check to avoid numerical problems of ill conditioning.   More details can be found in Ref. 33.

A recipe that we have found useful is the following: Matching points at the prism–vacuum interface are spaced at $R_c/10$ intervals on a square grid of dimension R_c on a side.   Experience shows that beyond this distance the field rapidly reduces to the simpler solution [Eq. (1)], which we know analytically.   On the probe–vacuum interface, we use matching points regularly spaced in the spherical coordinates ϕ and θ, referred to the center of the spheroid.   The number of matching points in each coordinate varies between 20 and 80 depending on the eccentricity of the spheroid.   Typical matching points are illustrated in Fig. 2(a) for the $R_c = 10$ nm, $a = 40$ nm spheroid.

Multiple centers are placed along the z axis at various values of z, as indicated in Fig. 2(b).   To represent the fields inside the spheroidal probe tip we use three centers, at $z = d - R_c$, $d + a$, and $d + 2a + R_c$ [left side of Fig. 2(b)].   Multipoles placed at the center are of regular type, as described above.   The multipoles centered above and below the spheroid instead have outgoing boundary conditions, determined by the outgoing spherical Hankel
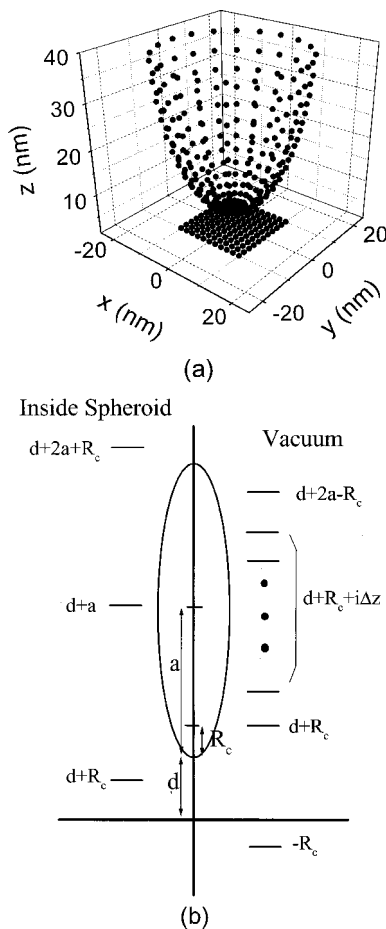
(a)

(b)

Fig. 2.   Setup of the MMP calculation.   Boundary conditions for Maxwell's equations ar enforced at a discrete set of points on the surface of the probe and the glass substrate.   A subset of these points is shown in (a) for the case of $a = 40$ nm, $b = 20$ nm, and $d = 5$ nm.   In addition the MMP calculation employs multipole expansions placed at various points along the symmetry axis of the probe and at various heights z, as indicated in (b).   The text details the locations of these expansions.

functions h_l^+.   These poles help to follow variations in the field near the comparatively sharp ends of the spheroid.

   The scattered field above the plane is represented by a string of outgoing multipoles with centers inside the probe [right side of Fig. 2(b)].   Two multipole centers are placed at the centers of curvature of the spheroid's ends, at $z = d + R_c$ and $d + 2a - R_c$.   In addition, several centers lie at equal intervals Δz (typically $2R_c$) between these centers.   Another multipole center is also placed below the surface of the prism, at $z = -R_c$.   This is a kind of "image" multipole, although experience shows that it is more useful if its position is fixed even when the probe tip moves up and down.   Finally, the field inside the glass prism is empirically well represented by a single outgoing multipole expansion placed at the center of curvature of the probe tip, $z = d + R_c$.

   With these multipoles in place the MMP method is reasonably stable for multipole orders less than $N_{max} = 6$–7.   Figure 3 illustrates the convergence versus multiple order of the intensity enhancement under the probe tip when the probe is just touching the prism.   Under these cir-

cumstances the field takes its largest value and experiences its strongest dependence on tip position.   Thus this quantity provides the most stringent test of numerical convergence.   For each of the a values in Fig. 3, two sets of enhancement versus N_{max} are shown, for different densities of matching points.   Under these circumstances the field intensities in the vicinity of the probe tip are converged to within several percent.   In addition, the difference between the fields on either side of each interface between dielectrics is monitored, and the appropriate boundary conditions are maintained at the several-percent level.   This level of convergence is adequate for the semiquantitative estimates we are aiming at here.

   Another issue must be kept in mind when this model is used.   To realistically model the tip as a macroscopic object in the z direction requires extending our model to the $a \rightarrow \infty$ limit.   However, as a becomes large, this calculation becomes problematic, since the spheroid is bounded and ultimately exhibits scattering resonances, which alter the field distributions.   Typically a resonance places field nodes at the surface of the spheroid, thus effectively reducing the apparent field near the probe tip.   We can estimate the value of a for which these resonances appear in the spheroidal model.   Regarding the electromagnetic field inside the spheroid as analogous to a quantum mechanical particle in a box, we expect the spheroid to resonate when its smallest dimension $2b = 2\sqrt{R_c a}$ becomes comparable to 1/2 the wavelength of the light in the material of which the probe is constructed, $\lambda_{probe} = \lambda/\sqrt{\epsilon}$.   Equating the two dimensions suggests that the lowest resonance appears for silicon tips when b becomes larger than ≈ 30 nm.

   In the case of a real probe tip these resonances are not expected to be a significant issue, for two reasons: (i) When the tip is very large compared with λ_{probe}, the resonances should blend together into a quasicontinuum, whereby only an averaged, nonresonant field is observed; (ii) in any real material there are dispersive losses, mean-
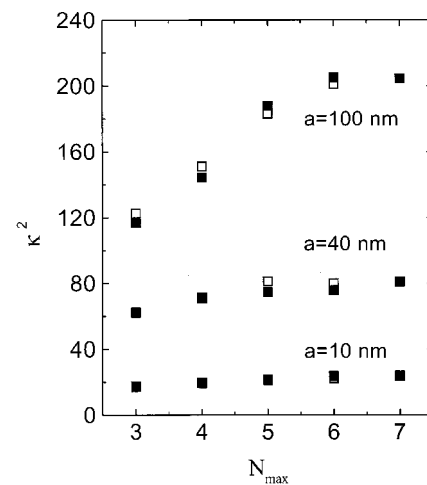

Fig. 3.   Convergence of the MMP calculation.   The intensity enhancement factor κ^2 at $d = 0$ is plotted versus N_{max}, the maximum order of multipoles used in the expansions.   Results for three different values of the probe's semimajor axis a are shown.   In each case the difference between open and solid symbols represents an increase in the number of matching points on the probe tip by ~20%.

ing that the standing waves in the tip responsible for resonances may not even form in the first place. For the silicon tips of interest here, the skin depth of visible light is ~1100 nm, far smaller than the size of the tip. Thus damping of such resonances would be significant in a real tip, even though we neglect it in the numerical calculations here.

B. Electrostatic Approximation

As illustrated in the next section, much can be learned about the electric field distribution underneath the tip from numerical calculations for finite values of a, specifically in the nonresonant region $a \leq 100$ nm. Nevertheless, it is worth a digression here to construct a simpler, analytic model that can suggest qualitatively how large the fields can become in the $a \rightarrow \infty$ limit. To do this we assemble an electrostatic model that is easy to implement, is analytic, and approximates the correct shape of the field. These results are presented alongside those of the numerical calculations in Section 3.

The approximate model replaces the probe tip by a point dipole located at the tip's radius of curvature, $z = d + R_c$.[14,23,30] The fields due to this dipole and its image dipole in the glass prism, located at $z = -d - R_c$, then approximate the near field. We determine the size of this dipole using a self-consistent electrostatic approximation. We represent the evanescent field by a uniform field of magnitude $\mathbf{E}_e$ pointing in the z direction to mimic the vertical component of the incident light. In this field the tip acquires an induced dipole moment of magnitude

$$p_{\text{ind}} = \alpha(E_e + E_{\text{im}}), \qquad (3)$$

where

$$E_{\text{im}} = \frac{2p_{\text{im}}}{(2d + 2R_c)^3} \qquad (4)$$

is the field at the sphere due to the image dipole of magnitude p_{im}. In the case of a small spherical tip of radius R_c, α is the familiar polarizability given by

$$\alpha^{\text{sphere}} = R_c^3 \left(\frac{\epsilon - 1}{\epsilon + 2} \right). \qquad (5)$$

More generally, for an extremely elongated ellipsoid, we must replace this value by the effective field enhancement near the ellipsoid's tip.

$$\alpha = \frac{(\epsilon - 1)}{2} R_c^3, \qquad (6)$$

as described in Appendix A.

The electrostatic potentials above ($z > 0$) and below ($z < 0$) the prism are given in cylindrical coordinates by

$$\Phi_{\text{above}} = -E_e z + \frac{p_{\text{ind}}(z - R_c - d)}{[\rho^2 + (z - R_c - d)^2]^{3/2}}$$

$$+ \frac{p_{\text{im}}(z + R_c + d)}{[\rho^2 + (z + R_c + d)^2]^{3/2}}, \qquad (7)$$

and

$$\Phi_{\text{below}} = -\frac{E_e}{\epsilon_g} z + \frac{p_{\text{prism}}(z - R_c - d)}{[\rho^2 + (z - R_c - d)^2]^{3/2}} \qquad (8)$$

where ϵ_g stands for the dielectric constant of the glass from which the prism is made, and p_{prism} is an effective dipole moment to be determined. By applying the appropriate boundary conditions to the electric field $\mathbf{E} = -\nabla\Phi$ at the prism–vacuum interface, we arrive at the values of the remaining dipoles:

$$p_{\text{im}} = \left(\frac{\epsilon_g - 1}{\epsilon_g + 1} \right) p_{\text{ind}} \equiv \beta p_{\text{ind}}, \qquad (9)$$

$$p_{\text{prism}} = \frac{2}{\epsilon_g + 1} p_{\text{ind}}. \qquad (10)$$

Inserting Eq. (4) and (9) to substitute for E_{im} in Eq. (3) yields a self-consistent equation for the induced dipole moment

$$p_{\text{ind}} = \alpha E_e \left[1 - \frac{\alpha\beta}{4(d + R_c)^3} \right]^{-1}. \qquad (11)$$

With these dipole moments determined, the gradients of Eqs. (7) and (8) yield the electric fields in our model.

Thus the magnitude of the near field can be obtained in our electrostatic picture, and in particular the value of κ. Recalling that κ is the ratio of the field's magnitude to that of the evanescent wave, we must divide the resulting field by E_e. The resulting expression for the d-dependent tip enhancement for a spherical probe is

$$\kappa = 1 + \frac{2\alpha[1/R_c^3 + \beta/(2d + R_c)^3]}{1 - [\alpha\beta/4(d + R_c)^3]}. \qquad (12)$$

The first term in this expression is just the applied field itself. The second term is proportional to the polarizability of the tip and contains terms in its numerator arising from the tip's induced dipole as well as terms due to the image dipole. In addition, this term contains a resonance denominator that arises from the self-consistency of the total electric field. This denominator is largely responsible for additional enhancement that is the proximity of the tip to the surface.

The dependence of the enhancement on α highlights the role of the antenna effect. For example, for a spherical silicon tip ($\epsilon = 19$), we find in this approximation $\kappa = 2.95$, which is not particularly large. However, for an elongated tip where Eq. (6) is the appropriate polarizability, we find $\kappa = 187$, far larger than the computed MMP value obtained below. The reason for this overestimate is clear: We are assuming an arbitrarily long probe tip uniformly illuminated by an electric field whose strength is constant in z, in contrast to the actual illumination provided by the evanescent field. In addition, the dipole moments excited in the real tip by a time-dependent field only add their contributions to the field coherently within a portion of the tip of order $\lambda/4$.

3. NEAR FIELD: SIZE, SHAPE, AND STRENGTH

In this section we explore the ANSOM near field using both the detailed MMP calculation and the electrostatic

approximation described above. We consider tips of two different representative radii of curvature $R_c = 10$ nm and $R_c = 5$ nm, to help illustrate the importance of R_c in determining the properties of the near field.

A. Dependence on Probe Position: Lateral Resolution

An important requirement for successful ANSOM imaging is that the resolution be determined by the size of the probe tip, and not the macroscopic length of the probe. This has been explicitly verified in various calculations of the scattered light[32] and in field enhancement under plane-wave illumination.[20] Here we also verify sharp resolution for the near fields in the evanescently excited silicon tip, showing also the dependence of resolution on the tip's height above the surface.

Figure 4 plots the intensity of the field in the x–y plane immediately above the surface of the glass prism, where the molecules being probed would sit, for a probe tip of radius $R_c = 10$ nm hovering $d = 2$ nm above the surface. In Fig. 4(a) is shown the result of a full MMP calculation for a spheroid of semimajor axis $a = 100$ nm, while in Fig. 4(b) is shown the corresponding field distribution for the electrostatic approximation in the $a \rightarrow \infty$ limit. In both cases the intensity is normalized to the same maximum value in order to emphasize the overall shape of the intensity profile. This normalization accounts for the apparently different background intensity levels in the two cases, since the electrostatic model yields a far higher peak intensity, as discussed above.

The shapes of the two peaks are qualitatively similar, speaking to the general success of the electrostatic model in describing the field shape. Note in particular that the electrostatic results is by construction cylindrically symmetric about the probe's axis and that the full MMP solu-
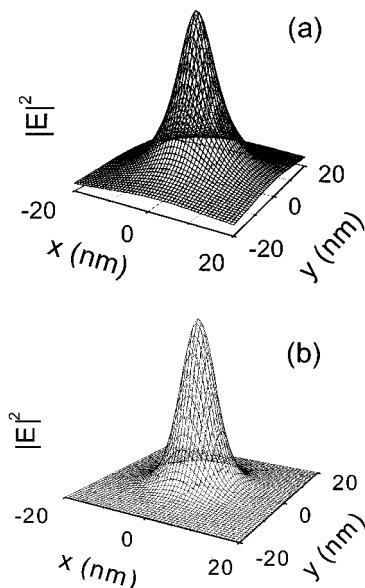

Fig. 5. ANSOM resolution: The FWHM of the near field intensity in the x direction is plotted versus the probe tip–surface distance d. Both quantities are normalized by R_c, the tip's radius of curvature. (a) and (b) show the results for the $R_c = 10$ nm and $R_c = 5$ nm probe tips, respectively. The values of the semimajor axis a are indicated in the legend. In both panels the result of the electrostatic approximation is shown as a heavy line.

tion is nearly so. Closer examination reveals a small asymmetry in the MMP result, leading to a slightly higher intensity on the downstream (i.e., positive-x) side of the probe. One essential feature apparent in the intensity profiles in Fig. 3 is that their FWHM is of order $\sim$10 nm, nearly the same as the radius of curvature of the tip.

To emphasize the relation between resolution and the radius of curvature R_c, we plot in Fig. 5 the FWHM of the intensity peaks at the surface versus the probe tip–surface distance d, with both quantities normalized by R_c. Figure 5(a) shows the result for the $R_c = 10$ nm tip, while 5(b) shows the result for the $R_c = 5$ nm tip. In both cases the data are calculated for various values of semimajor axis a to illustrate the rapid convergence as a grows. Moreover, the converged curves are very nearly the same for the two tip radii, especially when the tip is close to the surface.

This good agreement underscores the importance of R_c in determining the resolution width. Thus when the probe is in contact with the surface ($d = 0$), the FWHM is roughly $0.6R_c$. It then grows with increasing d, nearly doubling in width by the time $d/R_c \approx 1$. The heavy solid curves in Fig. 5 show the same result calculated within the electrostatic model. This model reproduces the width quite well in the small-d limit yet underestimates the growth in width as the probe tip is lifted from the surface. Still, this simple qualitative model also confirms that R_c is the scale of lateral resolution.

B. Vertical Resolution: The Proximity Effect

A second crucial characteristic of the optical near field is its extremely rapid falloff as the tip is removed from the surface. Figure 6 illustrates this effect, plotting field intensity at the surface versus d/R_c. Here all fields are

Fig. 4. Distribution of electric field intensity at the glass surface for a probe with radius of curvature $R_c = 10$ nm and a probe tip–surface distance $d = 2$ nm. (a) shows the result of the full MMP calculation for an $a = 100$ nm probe, while (b) shows the approximate result obtained in the electrostatic approximation with $a \rightarrow \infty$. The two calculations are normalized to the same peak intensity.
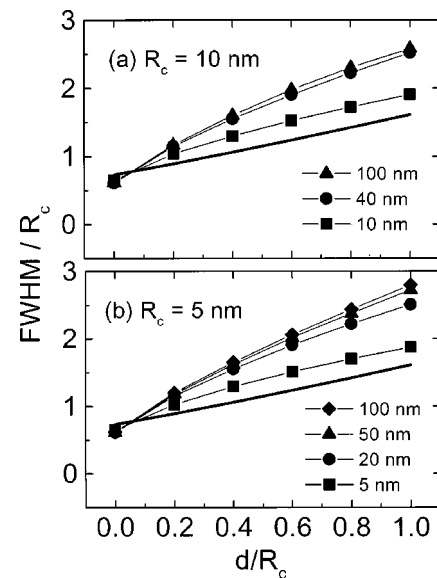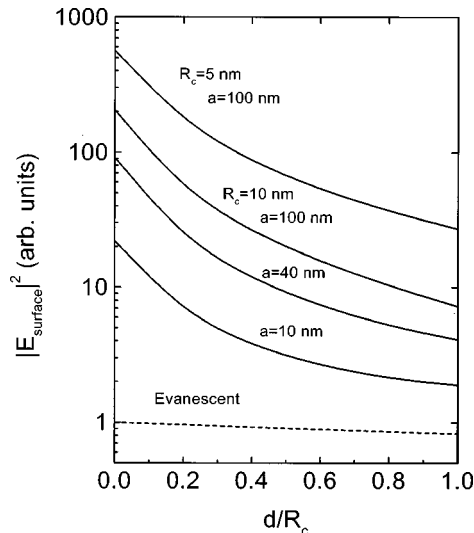
Fig. 6. Proximity effect: The intensity on the glass surface immediately beneath the probe tip is plotted as a function of the tip–surface distance, d, in units of R_c. For comparison the dashed curve shows the rate at which the evanescent field falls off with distance in the absence of a probe tip. MMP calculations are shown for three different probe lengths a for the $R_c = 10$ nm tip. Also shown is the result for the $R_c = 5$ nm, $a = 100$ nm tip.

normalized by the intensity of the evanescent field just above the prism's surface. For comparison, the dashed curve plots $\exp(-d/l_0)$, i.e., the rate at which the evanescent field itself would fall off if the probe had no effect. Recall that for our calculations the $1/e$ length l_0 of the evanescent intensity is 52 nm. For comparison, the calculation for the $R_c = 5$ nm, $a = 100$ nm tip is also shown.

In all cases, two features are immediately apparent: The intensity is far higher, by two orders of magnitude, than the ambient evanescent intensity, and it falls off far more rapidly. Indeed, the intensity falls to roughly half its $d = 0$ value by $d/R_c \sim 0.1$ (note the logarithmic axis in the figure). The vertical resolution is thus even sharper than the lateral resolution and in fact is basically independent of the probe length a. This sharp vertical resolution originates in the proximity effect and falls off rapidly as a result of the $1/d^3$ falloff of the induced and images dipole, as suggested by the electrostatic approximation.

C. Dependence on Probe Length: The Antenna Effect

Finally, we address explicitly the dependence of the field intensity itself on the probe length, i.e., the antenna effect. To demonstrate this effect we show in Fig. 7 the intensity enhancement κ^2 at the surface as a function of the semimajor axis a, again scaled by R_c. When $a/R_c = 1$ the probe is spherical, and when $a/R_c > 1$ the probe is prolate with radius of curvature maintained at R_c nm at the tip. The case shown is for the tip just touching the surface, $d = 0$.

As was the case for the resolution width, here the agreement between the $R_c = 10$ nm and $R_c = 5$ nm tips is quite good, especially for small values of a. For a spherical tip we understand that the strength of the induced dipole is $\sim R_c^3$, while the near field a distance R_c

from the dipole is proportional to $1/R_c^3$. The two factors cancel, leaving the absolute intensities nearly equal in the limit where the sphere is far smaller than the wavelength of light. As a/R_c grows, this trend changes somewhat, and the results for the two tips begin to diverge. Moreover, at larger values of a/R_c, both curves begin to show saturation behavior as less of the tip is directly illuminated by the evanescent wave. For the $R_c = 10$ nm tip, the intensity in fact drops abruptly when $a/R_c \sim 12$, i.e., when $b \sim 30$ nm, signaling the onset of internal resonances as anticipated in subsection 2 A. When the spheroid becomes large enough to support resonances, the fields inside the probe (and hence in the near field outside the probe) begin to change qualitatively in character, destroying the growing trend in κ^2. By contrast, the $R_c = 5$ nm tip does not experience this falloff in the range shown, since even for the largest a calculated, the semiminor axis only gets as large as $b \approx 22$ nm, too narrow to support these resonant states.

Figure 7 illustrates that the intensity enhancement is $\geqslant 250$, implying a field enhancement $\kappa \geqslant 15$, already half the observed value. How much larger it would go in the limit of large a remains uncertain, owing to the fundamental resonance limit of the ellipsoidal model. However, we can attempt an estimate based on the Padé approximant technique.[36] Specifically, we fit the rising part of κ^2 versus a/R_c to the approximant

$$\kappa^2 = \frac{A(a/R_c)}{1 + B(a/R_c)}. \tag{13}$$

This particular rational function correctly expresses both the observed linear behavior of κ^2 at small a/R_c and the fact that the intensity must saturate at large a/R_c.

The approximants are shown in Fig. 7; the dashed curve is for the $R_c = 10$ nm tip, and the solid curve is for the $R_c = 5$ nm tip. The fits are easily extrapolated to tips of infinite length, yielding $\kappa = 30 \pm 5$ for the $R_c = 10$ nm tip and $\kappa = 22 \pm 2$ for the $R_c = 5$ nm tip, with uncertainties arising from the scatter in the computed
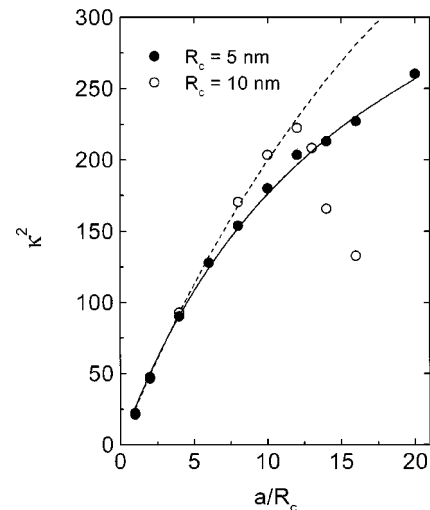

Fig. 7. Antenna effect: The intensity enhancement factor κ^2 is plotted versus normalized probe length a/R_c. The solid and dashed curves represent Padé approximants to the numerical data.

points. Without taking these values too seriously, we nevertheless note that they are of approximately the same magnitude as the experimentally inferred value $\kappa = 28$. Moreover, the smaller tip yields a somewhat smaller enhancement, since it presents a smaller volume to the evanescent field.

4. CONCLUSIONS

To reiterate, we have theoretically verified several aspects of the near field generated by apertureless NSOM: (i) The *resolution* in both the horizontal and vertical directions is determined primarily by the sharpness of the probe tip; (ii) The *intensity* delivered to the sample arises both from the macroscopic length of the probe and from a significant proximity effect as the tip nears the surface; (iii) based on numerical simulations for $\epsilon = 19$, the magnitude of the field enhancement at the probe tip is $\kappa \geq 15$. Additionally, the electrostatic approximation notes that κ grows linearly with ϵ, suggesting that further enhancement can be obtained by crafting the probe from a suitable material. This large enhancement occurs even in the absence of a plasmon resonance, which is of course expected to enhance the field even further.

APPENDIX A: ELECTROSTATIC FIELD ENHANCEMENT

In this appendix we determine the field enhancement at the tip of a prolate spheroidal dielectric in a uniform static field directed along its long axis. For this purpose we adapt the general result derived in Section 5.3 of Ref. 37 but with the long axis of the spheroid along the z, rather than the x, axis. Reference 37 casts the problem in the prolate spheroidal coordinates (ξ, η, ζ).

An applied potential $\Phi_e = -E_e z$ induces a potential outside the spheroid that is a function of ξ only:

$$\Phi_{\rm ind} = \frac{\Phi_e(1 - \epsilon)\left(\dfrac{ab^2}{2}\right)\displaystyle\int_\xi^\infty \dfrac{{\rm d}q}{{\rm D}(q)}}{1 + (\epsilon - 1)L_3}, \qquad ({\rm A1})$$

where L_3 is given by the integral

$$L_3 = \frac{ab^2}{2}\int_0^\infty \frac{{\rm d}q}{{\rm D}(q)} \qquad ({\rm A2})$$

and the denominator in the integrand is

$$ {\rm D}(q) = (q + a^2)^{3/2}(q + b^2). \qquad ({\rm A3})$$

We are interested in the limit where $a \gg b$, in which case $L_3 \to 0$ (Ref. 37) and the potential reduces to

$$\Phi_{\rm ind} = -E_e z(1-\epsilon)\frac{ab^2}{2}\int_\xi^\infty \frac{{\rm d}q}{{\rm D}(q)}. \qquad ({\rm A4})$$

To evaluate the induced electric field right at the tip, we need to compute $-\partial\Phi_{\rm ind}/\partial z$ at the Cartesian coordinates $(x,y,z) = (0,0,a)$. Setting $z = a+\delta$ in the vicinity of the tip, we find that $\xi \approx 2a\,\delta$, and that

$$\Phi_{\rm ind} = -E_e(1 - \epsilon)\frac{ab^2}{2}[(a + \delta)g(\delta)], \qquad ({\rm A5})$$

where the integral now takes the form

$$g(\delta) = \int_{2a\,\delta}^\infty \frac{{\rm d}q}{{\rm D}(q)}. \qquad ({\rm A6})$$

The gradient of $\Phi_{\rm ind}$ with respect to z is now equivalent to its gradient with respect to δ. A simple derivative yields

$$\frac{\partial}{\partial\delta}[(a+\delta)g(\delta)] \xrightarrow[\delta \to 0]{} \frac{\partial g}{\partial\delta}(\delta = 0) + g(0) = \frac{-2}{ab^2} + 0. \qquad ({\rm A7})$$

In this last step we have used the fact that in the limit of small δ, $g(\delta)$ approaches L_3, which vanishes for large a. The magnitude of the induced field becomes, after substitution of the appropriate coefficients from Eq. (A5),

$$E_{\rm ind} = E_e(\epsilon - 1). \qquad ({\rm A8})$$

This enhancement factor, $(\epsilon - 1)$, replaces the factor $2(\epsilon - 1)/(\epsilon + 2)$ that arises from a dielectric sphere in a field.

ACKNOWLEDGMENTS

John L. Bohn's e-mail address is bohn @murphy.colorado.edu.

REFERENCES

1. L. Kador, D. E. Horne, and H. W. Moerner, "Optical detection and probing of single dopant molecules of pentacene in a p-terphenyl host crystal by means of absorption spectroscopy," J. Phys. Chem. **94**, 1237–1248 (1990).
2. S. Nie, D. T. Chiu, and R. N. Zare, "Probing individual molecules with confocal fluorescence microscopy," Science **266**, 1018–1021 (1994).
3. T. Ha, Th. Enderle, D. S. Chemla, P. R. Selvin, and S. Weiss, "Quantum jumps of single molecules at room temperature," Chem. Phys. Lett. **271**, 1–5 (1997).
4. E. Betzig and J. K. Trautman, "Near-field optics: microscopy, spectroscopy, and surface modification beyond the diffraction limit," Science **257**, 189–195 (1992).
5. B. Hecht, H. Bielefeldt, Y. Inouye, and D. W. Pohl, "Facts and artifacts in near-field optical microscopy," J. Appl. Phys. **81**, 2492–2498 (1997).
6. F. Zenhausern, M. P. O'Boyle, and H. K. Wickramsinge, "Apertureless near-field optical microscope," Appl. Phys. Lett. **65**, 1623–1625 (1994).
7. Y. Inouye and S. Kawata, "Near-field scanning optical microscope with a metallic probe tip," Opt. Lett. **19**, 159–161 (1994).
8. P. Gleyzes, A. C. Boccara, and R. Bachelot, "Near-field optical microscopy using a metallic vibrating tip," Ultramicroscopy **57**, 318–322 (1995).
9. L. Novotny, E. J. Sánchez, and X. S. Xie, "Near-field optical imaging using metal tips illuminated by higher-order Hermite–Gaussian beams," Ultramicroscopy **71**, 21–29 (1998).
10. H. F. Hamann, A. Gallagher, and D. J. Nesbitt, "Enhanced sensitivity near-field scanning optical microscopy at high spatial resolution," Appl. Phys. Lett. **73**, 1469–1471 (1998).

11. J. Azoulay, A. Débarre, A. Richard, and P. Tcénio, "Field enhancement and apertureless near-field optical spectroscopy of single molecules," J. Microsc. **194**, 486–490 (1999).

12. E. J. Sánchez, L. Novoty, and X. S. Xie, "Near-field fluorescence microscopy based on two-photon excitation with metal tips," Phys. Rev. Lett. **82**, 4014–4017 (1999).

13. T. J. Yang, G. A. Lessard, and S. R. Quake, "An apertureless near-field microscope for fluorescence imaging," Appl. Phys. Lett. **76**, 378–380 (2000).

14. H. F. Hamann, A. Gallagher, and D. J. Nesbitt, "Near-field fluorescence imaging by localized field enhancement near a sharp probe tip," Appl. Phys. Lett. **76**, 1953–1955 (2000).

15. B. Knoll and F. Keilmann, "Near-field probing of vibrational absorption for chemical microscopy," Nature **399**, 134–137 (1999).

16. J. Gersten and A. Nitzan, "Electromagnetic theory of enhanced Raman scattering by molecules adsorbed on rough surfaces," J. Chem. Phys. **73**, 3023–3037 (1980).

17. S. Nie and S. R. Emory, "Probing single molecules and single nanoparticles by surface-enhanced Raman scattering," Science **275**, 1102–1106 (1997).

18. 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)," Phys. Rev. Lett. **78**, 1667–1670 (1997).

19. W. Denk and D. W. Pohl, "Near-field optics: Microscopy with nanometer-size fields," J. Vac. Sci. Technol. B **9**, 510–513 (1991).

20. O. J. F. Martin and C. Girard, "Controlling and tuning strong optical field gradients at a local probe microscope tip apex," Appl. Phys. Lett. **70**, 705–707 (1997).

21. A. V. Zayats, "Electromagnetic field enhancement in the context of apertureless near-field microscopy," Opt. Commun. **161**, 156–162 (1999).

22. L. Aigouy, A. Lahrech, S. Grsillon, H. Cory, A. C. Boccara, and J. C. Rivoal, "Polarization effects in apertureless scanning near-field optical microscopy: an experimental study," Opt. Lett. **24**, 187–189 (1999).

23. A. V. Bragas, S. M. Landi, and O. E. Martinez, "Laser field enhancement at the scanning tunneling microscope junction measured by optical rectification," Appl. Phys. Lett. **72**, 2075–2077 (1998).

24. H. M. Al-Rizzo and J. M. Tranquilla, "Electromagnetic wave scattering by highly elongated and geometrically composite objects of large size parameters: the generalized multipole technique," Appl. Opt. **34**, 3502–3521 (1995).

25. S. Asano and G. Yamamoto, "Light scattering by a spheroidal particle," Appl. Opt. **14**, 29–49 (1975).

26. G. Videen, M. G. Turner, V. J. Iafelice, W. S. Bickel, and W. L. Wolfe, "Scattering from a small sphere near a surface," J. Opt. Soc. Am. A **10**, 118–126 (1993).

27. B. R. Johnson, "Calculation of light scattering from a spherical particle on a surface by the multipole expansion method," J. Opt. Soc. Am. A **13**, 326–337 (1996).

28. J. C. Chao, F. J. Rizzo, I. Elshafiey, Y. J. Liu, L. Upda, and P. A. Martin, "General formulation for light scattering by a dielectric body near a perfectly conducting surface," J. Opt. Soc. Am. A **13**, 338–344 (1996).

29. W.-H. Yang, G. C. Schatz, and R. P. Van Duyne, "Discrete dipole approximation for calculating extinction and Raman intensities for small particles with arbitrary shapes," J. Chem. Phys. **103**, 869–875 (1995).

30. C. J. Hill, P. M. Bridger, G. S. Picus, and T. C. McGill, "Scanning apertureless microscopy below the diffraction limit: Comparison between theory and experiment," Appl. Phys. Lett. **75**, 4022–4024 (1999).

31. R. Wannemacher, A. Pack, and M. Quinten, "Resonant absorption and scattering in evanescent fields," Appl. Phys. B **68**, 225–232 (1999).

32. M. Quinten, "Evanescent wave scattering by aggregates of clusters—application to optical near-field microscopy," Appl. Phys. B **70**, 579–586 (2000).

33. Ch. Hafner and L. Bomholt, *The 3D Electrodynamic Wave Simulator* (Wiley, Chichester, UK, 1993).

34. R. K. Wangsness, *Electromagnetic Fields* (Wiley, New York, 1979).

35. J. D. Jackson, *Classical Electrodynamics*, 2nd ed. (Wiley, New York, 1975), Chap. 16.

36. G. A. Baker, Jr., and P. Graves-Morris, *Padé Approximants, Part I: Basic Theory*, Vol. 13 of Encyclopedia of Mathematics and Its Applications, G. Rota, ed. (Addison-Wesley, Reading, MA, 1981).

37. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1998), Section 5.3.