

Near-Field Scanning Optical Tomography: A Nondestructive Method for Three-Dimensional Nanoscale Imaging

Jin Sun, P. Scott Carney, and John C. Schotland

(Invited Paper)

Abstract—We present the theoretical foundation for near-field scanning optical tomography, a method for three-dimensional optical imaging with subwavelength resolution. An analysis of the forward problem for both scalar and vector optical fields is described. This is followed by the construction of the pseudoinverse solution to the linearized inverse scattering problem. The results are illustrated by numerical simulations.

Index Terms—Imaging, inverse problems, inverse scattering, microscopy, near-field optics, scanning probe microscopy, tomography.

I. INTRODUCTION

NEAR-FIELD scanning optical microscopy (NSOM) [1]–[9] has attracted considerable attention as a technique to obtain images of surfaces with subwavelength resolution. This achievement is particularly important for imaging structures where spectroscopic concerns or sample handling requirements dictate the use of lower frequency fields and yet high spatial resolution is still required. Applications range from the inspection of organic and biological samples to semiconductors. Various experimental modalities are in practical use. Two prominent examples are collection-mode NSOM and illumination-mode NSOM. In illumination-mode NSOM, a tapered fiber probe with a subwavelength-size aperture serves as a source of illumination in the near zone of the sample. The scattered field intensity is then measured and recorded as a function of the probe position while the probe is scanned over the sample. In collection-mode NSOM, the fiber probe serves to detect the total field in the near zone as the sample is illuminated by a source in the far zone.

There are certain limitations of NSOM as currently practiced. Despite the fact that the sample may present a complicated

three-dimensional (3-D) structure, NSOM produces only a two-dimensional (2-D) image. Indeed, rather than being an imaging method, it is more accurate to say that NSOM maps the subwavelength structure of the optical near-field intensity in some plane above the sample. Under certain simplifying assumptions, such as homogeneity of the bulk optical properties of the sample [10]–[12], the images produced in these experiments may be related to the sample structure. However, for the more general case in which both the topography of the sample and the bulk optical properties vary, the connection between the near-field intensity and the sample structure has proven ambiguous [13].

To resolve this ambiguity, it is desirable to solve the inverse scattering problem (ISP). The ISP consists of reconstructing the 3-D object structure, in this case the spatial dependence of the dielectric susceptibility of the sample, from measurements of the scattered field. By solving the ISP we thus resolve two issues. We remove the ambiguity in connecting the sample properties and the measured data, and simultaneously obtain 3-D tomographic images of the sample.

Historically, the solution of the ISP for other scattering modalities has greatly expanded the functionality of existing methods. For instance, von Laue's analysis of X-ray diffraction made modern crystallography a reality. The work of Hounsfield and Cormack brought medical imaging out of the era of projection radiography and into the era of computed tomography (CT). In any ISP, the first step is to obtain a physically reasonable forward model for the scattering process. For instance, in CT, a geometric model of propagation, neglecting any scattering, is sufficient to describe the experiment. Likewise, as scattering becomes important, the first Born approximation is often employed. By considering the far-zone scattered field, a readily solvable ISP [14], now generally known as diffraction tomography (DT), may be obtained. The crucial step in the solution of the DT ISP is to obtain a linearized relationship between the sample properties and the scattered field or some simple function of the scattered field. This may be accomplished by making use of the first Born or Rytov approximations. Solutions for the nonlinear ISP may be obtained as well, but, in general, they present a much more computationally challenging problem.

The far-zone ISP has an inherent resolution limit imposed by the wavelength of the probe field. This limit may be traced to the fact that only the homogeneous part of the scattered field contributes to the far zone. While, in principle, a higher resolution image may be obtained by mathematical extrapolation [15], this

Manuscript received December 12, 2005; revised June 5, 2006. The work of P. S. Carney was supported in part U.S. Air Force MURI Grant F49620-03-1-0379 and in part by the National Science Foundation under CAREER Award Grant 0239265. The work of J. C. Schotland was supported in part by the Coordinated Science Laboratory and in part by the National Science Foundation under Grant DMR 0425780.

J. Sun is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA.

P. S. Carney is with the Beckman Institute for Advanced Science and Technology and the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA.

J. C. Schotland is with the Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104 USA, and also with Coordinated Science Laboratory and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA.

Digital Object Identifier 10.1109/JSTQE.2006.879567

approach is exponentially sensitive to errors in the scattering data. However, the evanescent waves which contribute to the near field carry higher spatial frequency information about the scatterer and are directly accessible experimentally. In addition, it has been demonstrated [16] that inclusion of the evanescent waves in the backpropagation algorithm of 2-D DT enhances the resolving power of that method.

To sum up the preceding discussion, the subwavelength resolution obtained in NSOM stems from the fact that direct access to the evanescent scattered waves may be gained by probing the near field. It is this part of the scattered field on which the subwavelength structure of the scattering object is encoded. The NSOM ISP is thus of great interest because it offers the possibility to obtain subwavelength-resolved 3-D reconstructions. We will refer to this method as near-field scanning optical tomography (NSOT).

NSOM instrumentation now exists that is capable of providing both the phase and amplitude of the optical field [17], [18]. It was recently demonstrated that when both the phase and amplitude of the optical near field are available, the near-field ISP may be solved, and an inversion formula was presented in [19]. Validation of the results was demonstrated by numerical simulation. The work was extended to other modalities and the analysis was modified to include a substrate [20], however simulations were limited to free space and polarization effects only considered for photon scanning tunneling microscopy (PSTM). In addition, an experimental demonstration of near-field inverse scattering was given for the closely related modality of PSTM [21]. In this paper, we expand upon and extend the results reported in our earlier letter. We discuss three experimental modalities, the illumination mode, the collection mode, and the two-probe mode NSOT and present analyses of the near-field ISP for both scalar and vector waves. A treatment of the vector case is particularly important in the near-field ISP because polarization effects are somewhat more complicated than for the far-field problem [22], and the scalar approximation may not be appropriate when the sample presents subwavelength variations in structure. The solution to the near-field ISP in all cases will be derived by a singular-value decomposition (SVD) analysis of the linearized scattering problem. The SVD is a generalized-mode decomposition that offers considerable insight into the scattering problem.

The remainder of the paper is organized as follows. In Section II, we discuss in detail the theory of near-field scattering for wave fields obeying the scalar reduced-wave equation. Specifically, we obtain expressions for the scattered field as measured in several distinct experimental modalities. In all cases, we assume that the field may be measured, i.e., the measurements are phase sensitive. The first two modalities are variations on the illumination-mode- and collection-mode NSOM experiments. The two-probe modality we discuss is conceptually novel and requires the use of two near-field probes. The expressions derived for the scattered fields are shown to be special cases of a single expression relating the measured data to the structure of the scattering object to be imaged. In Section III, we discuss the vector theory of near-field scattering. We return to both modalities earlier discussed and present the common underlying form. In Section IV-A, we take advantage of the unified form for

all modalities. We develop the SVD for the forward-scattering operator and obtain an inversion formula for the ISP. In Section IV-B, we develop the SVD for the vector-field problem. It will be seen that although the vector nature of light makes a significant difference in the physical description of the problem, the mathematical form is much like the scalar case and a similar solution may be obtained.

II. FORWARD PROBLEM: SCALAR CASE

We begin by considering the scattering of scalar waves in the near field. This will serve to introduce the necessary formalism without the mathematical complexity of the full vector theory. Note that the scalar theory is of independent interest and may prove useful in applications such as scanning acoustic microscopy [23].

Consider an experiment in which a monochromatic scalar wave is incident on a medium described by the scattering potential $\eta(\mathbf{r})$. The sample is placed on the surface of an infinite uniform medium with index of refraction n . The surface coincides with the $z = 0$ plane and the maximum height of the sample is z_m . The scalar field $U(\mathbf{r})$ satisfies the wave equation

$$\nabla^2 U(\mathbf{r}) + n^2(z)k_0^2 U(\mathbf{r}) = -4\pi k_0^2 \eta(\mathbf{r})U(\mathbf{r}) \quad (1)$$

where k_0 is the free-space wavenumber, and $n(z) = n$ for $z < 0$ and $n(z) = 1$ for $z \geq 0$. The field will be taken to consist of two parts

$$U(\mathbf{r}) = U_i(\mathbf{r}) + U_s(\mathbf{r}) \quad (2)$$

with the incident field $U_i(\mathbf{r})$ satisfying the equation

$$\nabla^2 U_i(\mathbf{r}) + n^2(z)k_0^2 U_i(\mathbf{r}) = 0 \quad (3)$$

and the scattered-field $U_s(\mathbf{r})$ satisfying the equation

$$\nabla^2 U_s(\mathbf{r}) + n^2(z)k_0^2 U_s(\mathbf{r}) = -4\pi k_0^2 \eta(\mathbf{r})U(\mathbf{r}). \quad (4)$$

Following standard procedures [24], (4) may be recast as the integral equation

$$U_s(\mathbf{r}) = k_0^2 \int d^3 r' G(\mathbf{r}, \mathbf{r}') U(\mathbf{r}') \eta(\mathbf{r}'). \quad (5)$$

Here, $G(\mathbf{r}, \mathbf{r}')$ is the half-space Green's function which obeys the equation

$$[\nabla^2 + n^2(z)k_0^2] G(\mathbf{r}, \mathbf{r}') = -4\pi \delta(\mathbf{r} - \mathbf{r}') \quad (6)$$

and satisfies the boundary conditions

$$G(\mathbf{r}, \mathbf{r}')|_{z=0+} = G(\mathbf{r}, \mathbf{r}')|_{z=0-} \quad (7)$$

$$\hat{\mathbf{z}} \cdot \nabla G(\mathbf{r}, \mathbf{r}')|_{z=0+} = \hat{\mathbf{z}} \cdot \nabla G(\mathbf{r}, \mathbf{r}')|_{z=0-}. \quad (8)$$

The scattered field is calculated perturbatively. Such an approach is only reasonable in the event that the unknown sample is small or of low contrast. When the sample is large, such as might be the case when the substrate itself is unknown, a perturbative calculation will likely yield poor results. The incident field satisfies (3) and, therefore, includes the effects of the substrate in the zeroth-order field. The Green's function also includes the substrate effects and so the first-order (in η) scattered field includes the light scattered directly from the sample as well as

the field subsequently reflected at the interface between vacuum and substrate. This result, which is known as the first Born approximation to the scattered field, is given by the expression

$$U_s(\mathbf{r}) = k_0^2 \int d^3 r' G(\mathbf{r}, \mathbf{r}') U_i(\mathbf{r}') \eta(\mathbf{r}'). \quad (9)$$

It should be noted that the incident field satisfies (3) in the domain of the sample. That is, the source of the field is assumed to lie outside the sample and is assumed to be a primary source so that it is unaffected by feedback from the scattered field or reflections from the substrate. If the incident field is generated by a secondary source, i.e., a source density dependent on the local electric field, then the incident field itself must be calculated by solving another near-field scattering problem. Such an approach is outside the scope of this paper but will be discussed elsewhere.

In the discussion that follows, it will prove useful to express the Green's function G in the form of a plane-wave decomposition

$$G(\mathbf{r}, \mathbf{r}') = \frac{i}{2\pi} \int d^2 q g(z, z'; \mathbf{q}) \exp[i\mathbf{q} \cdot (\boldsymbol{\rho} - \boldsymbol{\rho}')] \quad (10)$$

where $\mathbf{r} = (\boldsymbol{\rho}, z)$, and $\mathbf{r}' = (\boldsymbol{\rho}', z')$. Depending on the values of z and z' , the plane-wave amplitudes $g(z, z'; \mathbf{q})$ are given by

$$g(z, z'; \mathbf{q}) = \begin{cases} \frac{1}{k_z(\mathbf{q})} [e^{ik_z(\mathbf{q})(z-z')} + R(\mathbf{q})e^{ik_z(\mathbf{q})(z+z')}], & 0 \leq z' \leq z \\ \frac{1}{k_z(\mathbf{q})} [e^{ik_z(\mathbf{q})(z'-z)} + R(\mathbf{q})e^{ik_z(\mathbf{q})(z'+z)}], & 0 \leq z < z' \\ \frac{1}{k'_z(\mathbf{q})} T'(\mathbf{q}) e^{i[k_z(\mathbf{q})z - k'_z(\mathbf{q})z']}, & z' < 0 \leq z \\ \frac{1}{k'_z(\mathbf{q})} T(\mathbf{q}) e^{i[k_z(\mathbf{q})z' - k'_z(\mathbf{q})z]}, & z < 0 \leq z' \end{cases} \quad (11)$$

where $k_z(\mathbf{q}) \equiv (k_0^2 - q^2)^{1/2}$ and $k'_z(\mathbf{q}) = \sqrt{n^2 k_0^2 - q^2}$, and the reflection coefficients $R(\mathbf{q})$, $R'(\mathbf{q})$ and the transmission coefficients $T(\mathbf{q})$ and $T'(\mathbf{q})$ are given by

$$R(\mathbf{q}) = \frac{k_z(\mathbf{q}) - k'_z(\mathbf{q})}{k_z(\mathbf{q}) + k'_z(\mathbf{q})} \quad (12)$$

$$R'(\mathbf{q}) = \frac{k'_z(\mathbf{q}) - k_z(\mathbf{q})}{k'_z(\mathbf{q}) + k_z(\mathbf{q})} \quad (13)$$

$$T(\mathbf{q}) = \frac{2k_z(\mathbf{q})}{k_z(\mathbf{q}) + k'_z(\mathbf{q})} \quad (14)$$

$$T'(\mathbf{q}) = \frac{2k'_z(\mathbf{q})}{k'_z(\mathbf{q}) + k_z(\mathbf{q})}. \quad (15)$$

The modes appearing in (10) are labeled by the transverse-wave vector \mathbf{q} . The modes for which $|\mathbf{q}| \leq k_0$ correspond to homogeneous plane waves. The modes for which $|\mathbf{q}| > k_0$ correspond to evanescent waves. For these modes, $k_z(\mathbf{q})$ is pure imaginary so that the evanescent waves decay exponentially on propagation, with a concomitant loss of high spatial-frequency components of the scattered field.

A. Illumination Mode

In illumination-mode NSOT, the sample is illuminated in the near zone by the probe. The scattered field is then measured in

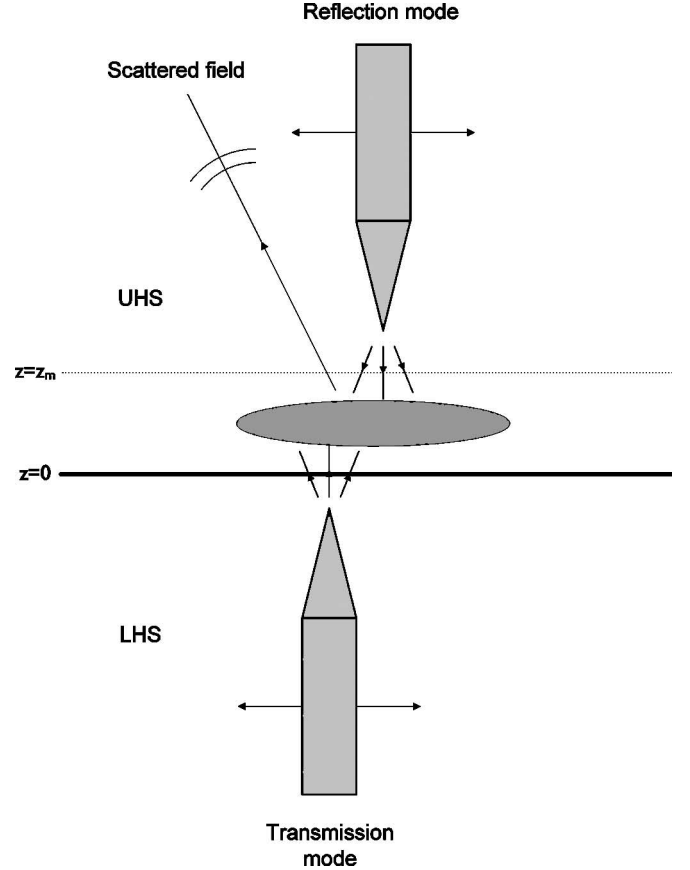


Fig. 1. Illumination-mode NSOT.

the far zone, as shown in Fig. 1. The probe serves as a point source of unit amplitude at the position \mathbf{r}_1 . The incident field is thus given by the half-space Green's function $G(\mathbf{r}, \mathbf{r}_1)$. The scattered field at a point \mathbf{r}_2 is given by the expression

$$U_s(\mathbf{r}_2) = k_0^2 \int d^3 r G(\mathbf{r}_2, \mathbf{r}) G(\mathbf{r}, \mathbf{r}_1) \eta(\mathbf{r}). \quad (16)$$

We assume that the source lies in the illumination plane $z = z_1$ so that the position of the source is $\mathbf{r}_1 = (\boldsymbol{\rho}_1, z_1)$ with $\boldsymbol{\rho}_1$ being the transverse coordinate of the source. The scatterer is assumed to lie only in the layer $0 \leq z \leq z_m$. The scattered field is measured in the upper half space. If $z_1 \leq 0$, we refer to this as *transmission mode*. If $z_1 > z_m$, we refer to this as *reflection mode*. Suppose that the observation point \mathbf{r}_2 is very far from the domain of the scatterer. Observe that for $|\mathbf{r}_2| \gg |\mathbf{r}|$, the leading term in the asymptotic expansion of the Green's function is given by

$$G(\mathbf{r}_2, \mathbf{r}) \sim \frac{e^{ik_0 r_2}}{r_2} e^{-i\mathbf{k}(\mathbf{q}_2) \cdot \mathbf{r}} [1 + R(\mathbf{q}_2) e^{2ik_z(\mathbf{q}_2)z}] \quad (17)$$

where $\mathbf{k}(\mathbf{q}_2) \equiv (\mathbf{q}_2, k_z(\mathbf{q}_2))$ is parallel to \mathbf{r}_2 . Thus, the scattered field behaves as an outgoing spherical wave which may be expressed in terms of the scattering amplitude $A(\mathbf{q}_2; \boldsymbol{\rho}_1, z_1)$ as

$$U_s(\mathbf{r}_2) \sim \frac{e^{ik_0 r_2}}{r_2} A(\mathbf{q}_2; \boldsymbol{\rho}_1, z_1) \quad (18)$$

where, using (16) and (17), it may be found that

$$A(\mathbf{q}_2; \boldsymbol{\rho}_1, z_1) = k_0^2 \int d^3r G(\mathbf{r}, \mathbf{r}_1) \times [1 + R(\mathbf{q}_2)e^{2ik_z(\mathbf{q}_2)z}]e^{-i\mathbf{k}(\mathbf{q}_2)\cdot\mathbf{r}}\eta(\mathbf{r}). \quad (19)$$

We assume that the source is scanned over a square lattice with spacing h . It will prove useful to define the data function $\Phi(\mathbf{q}_1, \mathbf{q}_2)$ by the lattice Fourier transform of $A(\mathbf{q}_2; \boldsymbol{\rho}_1, z_1)$ with respect to $\boldsymbol{\rho}_1$

$$\Phi(\mathbf{q}_1, \mathbf{q}_2) = h^2 \sum_{\boldsymbol{\rho}_1} e^{i\mathbf{q}_1\cdot\boldsymbol{\rho}_1} A(\mathbf{q}_2; \boldsymbol{\rho}_1, z_1) \quad (20)$$

where the sum over $\boldsymbol{\rho}_1$ is carried out over all lattice vectors and \mathbf{q}_1 belongs to the first Brillouin zone (FBZ) of the lattice. In this case $\text{FBZ} = [-\frac{\pi}{h}, \frac{\pi}{h}] \times [-\frac{\pi}{h}, \frac{\pi}{h}]$. Making use of (10), (19), and (20) and the identity

$$\sum_{\boldsymbol{\rho}} e^{i\mathbf{q}\cdot\boldsymbol{\rho}} = \left(\frac{2\pi}{h}\right)^2 \sum_{\mathbf{q}'} \delta(\mathbf{q} - \mathbf{q}') \quad (21)$$

where \mathbf{q} denotes a reciprocal lattice vector, we find that

$$\Phi(\mathbf{q}_1, \mathbf{q}_2) = 2\pi i k_0^2 \sum_{\mathbf{q}} \int d^3r e^{i(\mathbf{q}_1 - \mathbf{q}_2 - \mathbf{q})\cdot\boldsymbol{\rho}} \times [1 + R(\mathbf{q}_2)e^{2ik_z(\mathbf{q}_2)z}]e^{-i\mathbf{k}_z(\mathbf{q}_2)z} g(z, z_1, \mathbf{q}_1 - \mathbf{q})\eta(\mathbf{r}). \quad (22)$$

It is useful to identify $\tilde{\eta}(\mathbf{q}, z) = \int d^2\rho \eta(\boldsymbol{\rho}, z)e^{-i\mathbf{q}\cdot\boldsymbol{\rho}}$, the transverse Fourier transform of $\eta(\boldsymbol{\rho}, z)$. The reconstruction must be regularized. Regularization imposes an effective bandlimit on the spatial frequencies of the reconstruction. The object may thus be assumed to be bandlimited without consequence for the reconstruction. If $\eta(\boldsymbol{\rho}, z)$ is transversely bandlimited to the FBZ, that is if $\tilde{\eta}(\mathbf{q}, z) = 0$ for $\mathbf{q} \notin \text{FBZ}$, then the sum over \mathbf{q} may be truncated and only the $\mathbf{q} = \mathbf{0}$ term contributes to $\Phi(\mathbf{q}_1, \mathbf{q}_2)$. Thus, (22) may be written in the form of the integral equation

$$\Phi(\mathbf{q}_1, \mathbf{q}_2) = \int_0^{z_m} dz \kappa(\mathbf{q}_1, \mathbf{q}_2, z) \tilde{\eta}(\mathbf{q}_2 - \mathbf{q}_1, z) \quad (23)$$

where the kernel of the forward integral operator κ is given by

$$\kappa(\mathbf{q}_1, \mathbf{q}_2, z) = 2\pi i k_0^2 [1 + R(\mathbf{q}_2)] \times e^{2ik_z(\mathbf{q}_2)z} e^{-i\mathbf{k}_z(\mathbf{q}_2)z} g(z, z_1; \mathbf{q}_1). \quad (24)$$

The inversion of (23) to obtain η is discussed in Section IV-A. There we will see that it is possible to construct the transversely bandlimited minimum-norm approximation to η consistent with the lattice on which the measured scattering amplitude is sampled.

B. Collection Mode

In collection-mode NSOT, the sample is illuminated from the far zone by an incident plane wave and the scattered field is detected in the near zone by means of an idealized point detector,

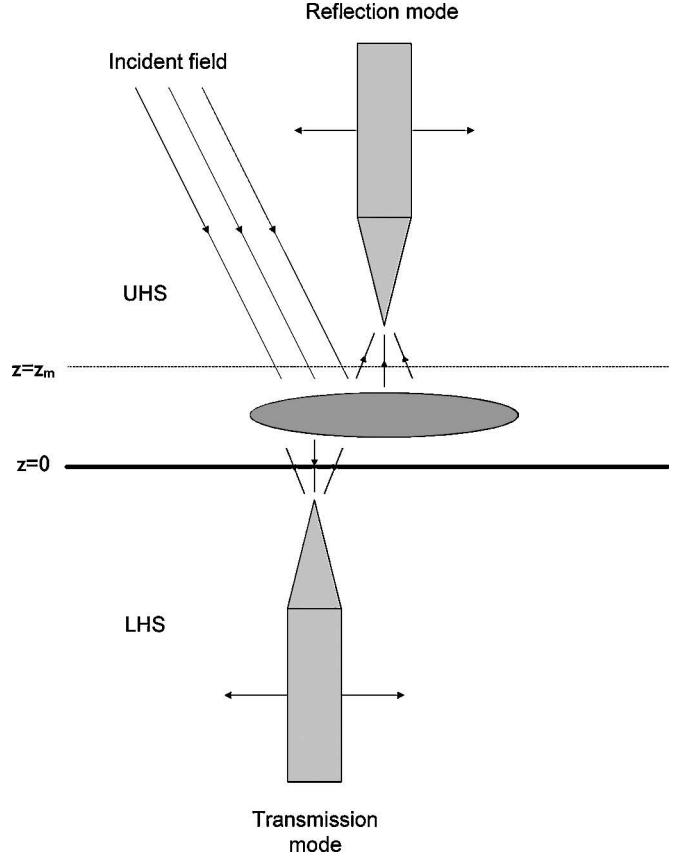


Fig. 2. Collection-mode NSOT.

as shown in Fig. 2. The incident plane wave with unit amplitude is of the form

$$U_i(\mathbf{r}) = [1 + R(\mathbf{q}_1)e^{2ik_z(\mathbf{q}_1)z}]e^{i[\mathbf{q}_1\cdot\boldsymbol{\rho} - k_z(\mathbf{q}_1)z]} \quad (25)$$

in the upper half space. It follows from (9) that the scattered field, measured in the $z = z_2$ plane at a point with coordinate $\mathbf{r}_2 = (\boldsymbol{\rho}_2, z_2)$, is given by the expression

$$U_s(\boldsymbol{\rho}_2, z_2; \mathbf{q}_1) = k_0^2 \int d^3r G(\mathbf{r}_2, \mathbf{r}) [1 + R(\mathbf{q}_1)] \times e^{2ik_z(\mathbf{q}_1)z} e^{i[\mathbf{q}_1\cdot\boldsymbol{\rho} - k_z(\mathbf{q}_1)z]}\eta(\mathbf{r}). \quad (26)$$

Note that here $z_2 > z_m$ corresponds to the reflection mode and $z_2 \leq 0$ to the transmission mode. As in illumination mode, we will define a data function $\Phi(\mathbf{q}_1, \mathbf{q}_2)$ by the lattice Fourier transform of the scattered field

$$\Phi(\mathbf{q}_1, \mathbf{q}_2) = h^2 \sum_{\boldsymbol{\rho}_2} e^{-i\mathbf{q}_2\cdot\boldsymbol{\rho}_2} U_s(\boldsymbol{\rho}_2, z_2; \mathbf{q}_1). \quad (27)$$

Making use of (10), the plane-wave representation of the Green's function, and the transverse bandlimiting of η , we find that the data function $\Phi(\mathbf{q}_1, \mathbf{q}_2)$ may be written in the form of (23) with κ given by

$$\kappa(\mathbf{q}_1, \mathbf{q}_2, z) = 2\pi i k_0^2 [1 + R(\mathbf{q}_1)e^{2ik_z(\mathbf{q}_1)z}] \times e^{-i\mathbf{k}_z(\mathbf{q}_1)z} g(z_2, z; \mathbf{q}_2). \quad (28)$$

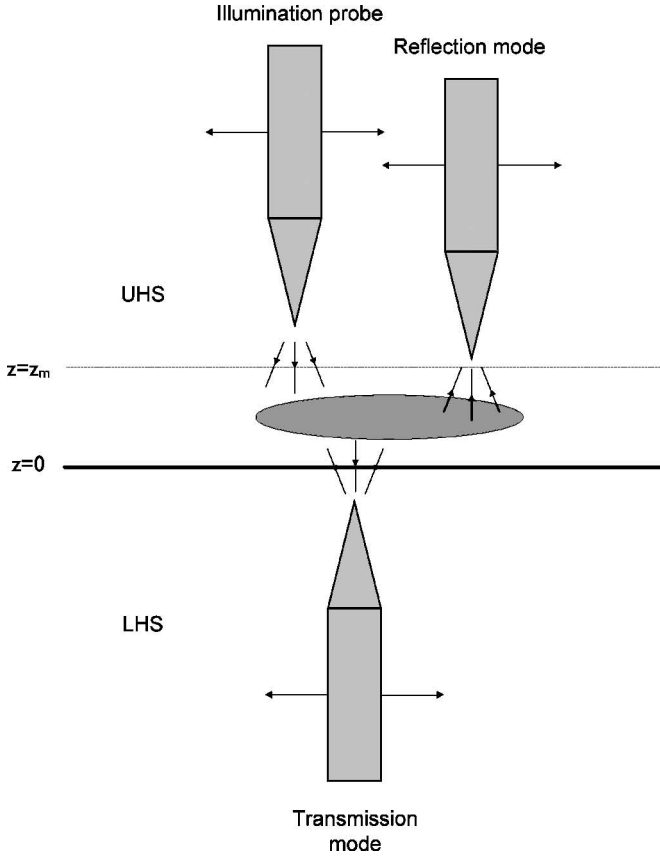


Fig. 3. Two-probe-mode NSOT.

C. Two-Probe Mode

Now, we consider the case of the two-probe mode in which both the source of illumination and the detector of the scattered field are in the near zone of the scatterer. The illumination is taken to originate from a point source of unit amplitude at $\mathbf{r}_1 = (\rho_1, z_1)$, where $z_1 > z_m$. The scattered field is measured by an idealized point detector at $\mathbf{r}_2 = (\rho_2, z_2)$, where $z_2 > z_m$ for reflection mode and $z_2 < 0$ for transmission mode, as shown in Fig. 3. As in the case of illumination mode, the incident field is given by the half-space Green's function G . Making use of (9), the scattered field is found to be given by the expression

$$U_s(\rho_2, z_2; \rho_1, z_1) = k_0^2 \int d^3r G(\mathbf{r}_2, \mathbf{r}) G(\mathbf{r}, \mathbf{r}_1) \eta(\mathbf{r}). \quad (29)$$

The data function is defined as the lattice Fourier transform of the scattered field with respect to the transverse coordinates of both the source and detector probes

$$\Phi(\mathbf{q}_1, \mathbf{q}_2) = h^4 \sum_{\rho_1, \rho_2} e^{i(\mathbf{q}_1 \cdot \rho_1 - \mathbf{q}_2 \cdot \rho_2)} U_s(\rho_2, z_2; \rho_1, z_1). \quad (30)$$

Using (10), (30), and the transverse bandlimiting of η , it may be seen that $\Phi(\mathbf{q}_1, \mathbf{q}_2)$ takes the form of (23) with κ given by

$$\kappa(\mathbf{q}_1, \mathbf{q}_2, z) = -(2\pi)^2 k_0^2 \sum_{\mathbf{q}'} g(z_2, z; \mathbf{q}_2 + \mathbf{q}') g(z, z_1; \mathbf{q}_1 + \mathbf{q}') \quad (31)$$

where the sum is taken over all reciprocal lattice vectors.

III. FORWARD PROBLEM: VECTOR CASE

We now turn our attention to the vector theory of light. The mathematical treatment of the vector formulation of near-field tomography is quite similar to the scalar case. Nevertheless, the full vector theory is necessary for our purposes since the scalar approximation to the scattering of electromagnetic waves breaks down when the dielectric susceptibility varies on subwavelength scales.

We will restrict our attention to nonmagnetic materials and again consider the half-space geometry. The spatial part of the monochromatic electric field satisfies the reduced-wave equation

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) - n^2(z) k_0^2 \mathbf{E}(\mathbf{r}) = 4\pi k_0^2 \mathbf{E}(\mathbf{r}) \eta(\mathbf{r}) \quad (32)$$

where $\eta(\mathbf{r})$ is the dielectric susceptibility related to the dielectric permittivity $\varepsilon(\mathbf{r})$ by $\varepsilon(\mathbf{r}) = 1 + 4\pi\eta(\mathbf{r})$ and k_0 is the free-space wavenumber. The field is taken to consist of two parts

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^i(\mathbf{r}) + \mathbf{E}^s(\mathbf{r}), \quad (33)$$

where $\mathbf{E}^i(\mathbf{r})$ is the incident field satisfying the equation

$$\nabla \times \nabla \times \mathbf{E}^i(\mathbf{r}) - n^2(z) k_0^2 \mathbf{E}^i(\mathbf{r}) = 0 \quad (34)$$

and $\mathbf{E}^s(\mathbf{r})$ is the scattered field obeying the equation

$$\nabla \times \nabla \times \mathbf{E}^s(\mathbf{r}) - n^2(z) k_0^2 \mathbf{E}^s(\mathbf{r}) = 4\pi k_0^2 \mathbf{E}(\mathbf{r}) \eta(\mathbf{r}). \quad (35)$$

Equation (35) may be reformulated as the integral equation

$$\mathbf{E}_\alpha^s(\mathbf{r}) = k_0^2 \int d^3r' G_{\alpha\beta}(\mathbf{r}, \mathbf{r}') E_\beta^i(\mathbf{r}') \eta(\mathbf{r}') \quad (36)$$

where $G_{\alpha\beta}$ is the Green's tensor which satisfies the equation

$$\nabla \times \nabla \times G_{\alpha\beta}(\mathbf{r}) - n^2(z) k_0^2 G_{\alpha\beta}(\mathbf{r}, \mathbf{r}') = 4\pi \delta(\mathbf{r} - \mathbf{r}') \delta_{\alpha\beta}. \quad (37)$$

The Green's tensor also obeys the boundary conditions

$$\hat{\mathbf{z}} \times \mathbf{G}(\mathbf{r}, \mathbf{r}')|_{z=0+} = \hat{\mathbf{z}} \times \mathbf{G}(\mathbf{r}, \mathbf{r}')|_{z=0-} \quad (38)$$

$$\hat{\mathbf{z}} \times \nabla \times \mathbf{G}(\mathbf{r}, \mathbf{r}')|_{z=0+} = \hat{\mathbf{z}} \times \nabla \times \mathbf{G}(\mathbf{r}, \mathbf{r}')|_{z=0-}. \quad (39)$$

Throughout this paper, Greek subscripts indicate vector components and the summation convention applies to repeated indices. Within the accuracy of the first Born approximation the scattered field is given by

$$\mathbf{E}_\alpha^s(\mathbf{r}) = k_0^2 \int d^3r' G_{\alpha\beta}(\mathbf{r}, \mathbf{r}') E_\beta^i(\mathbf{r}') \eta(\mathbf{r}'). \quad (40)$$

As in the formulation of the forward problem for the scalar case, it may be noted that the first Born approximation is appropriate for small samples on a known substrate and that the substrate is included nonperturbatively. Moreover, it is again implicit that the source of the field is external and primary so that though the incident field is consistent with the presence of the substrate and the attendant half-space boundary conditions, the source density itself is not dependent on the field and so no probe-substrate resonances are directly taken into account.

The Green's tensor may be expressed in the plane-wave decomposition

$$G_{\alpha\beta}(\mathbf{r}, \mathbf{r}') = \frac{i}{2\pi} \int d^2q g_{\alpha\beta}(z, z'; \mathbf{q}) \exp[i\mathbf{q} \cdot (\boldsymbol{\rho} - \boldsymbol{\rho}')] \quad (41)$$

where the plane-wave amplitudes $g_{\alpha\beta}(z, z'; \mathbf{q})$ are given by (42), shown at the bottom of the page, and $\Pi_{\alpha\beta}(\mathbf{q})$ is given by the expression

$$\Pi_{\alpha\beta}(\mathbf{q}) = \begin{cases} \delta_{\alpha\beta} - k_0^{-2} k_\alpha(\mathbf{q}) k_\beta(\mathbf{q}), & z - z' > 0 \\ \delta_{\alpha\beta} - k_0^{-2} \tilde{k}_\alpha(\mathbf{q}) \tilde{k}_\beta(\mathbf{q}), & z - z' < 0 \end{cases} \quad (43)$$

with $\tilde{\mathbf{k}}(\mathbf{q}) \equiv (\mathbf{q}, -k_z(\mathbf{q}))$. The reflection tensor $R_{\alpha\beta}(\mathbf{q})$ and the transmission tensors $T_{\alpha\beta}(\mathbf{q})$ and $T'_{\alpha\beta}(\mathbf{q})$ are defined in the Appendix.

A. Illumination Mode

We now return to the modality considered in Section II-A in which the sample is illuminated by a point source at the position $\mathbf{r}_1 = (\boldsymbol{\rho}_1, z_1)$, as shown in Fig. 1. The incident field is given by

$$E_\alpha^i(\mathbf{r}) = G_{\alpha\beta}(\mathbf{r}, \mathbf{r}_1) e_\beta \quad (44)$$

where $G_{\alpha\beta}$ is the half-space Green's tensor and e_β is the dipole moment of the source of the incident field. Using (40), the scattered field is seen to be of the form

$$E_\alpha^s(\mathbf{r}_2) = k_0^2 \int d^3r G_{\alpha\beta}(\mathbf{r}_2, \mathbf{r}) G_{\beta\gamma}(\mathbf{r}, \mathbf{r}_1) e_\gamma \eta(\mathbf{r}). \quad (45)$$

In the far zone of the scatterer, for $r_2 \gg r$, the asymptotic form of the Green's tensor is given by

$$G_{\alpha\beta}(\mathbf{r}_2, \mathbf{r}) \sim \frac{e^{ik_0 r_2}}{r_2} e^{-i\mathbf{k}(\mathbf{q}_2) \cdot \mathbf{r}} [\Pi_{\alpha\beta}(\mathbf{q}_2) + R_{\alpha\beta}(\mathbf{q}_2) e^{2ik_z(\mathbf{q}_2)z}] \quad (46)$$

where $\mathbf{k}(\mathbf{q}_2) \equiv (\mathbf{q}_2, k_z(\mathbf{q}_2))$ is in the direction of \mathbf{r}_2 . We see that the scattered field may be expressed in terms of the scattering amplitude $A_{\alpha\beta}(\mathbf{q}_2; \boldsymbol{\rho}_1, z_1)$

$$E_\alpha^s(\mathbf{r}_2) \sim A_{\alpha\beta}(\mathbf{q}_2; \boldsymbol{\rho}_1, z_1) e_\beta \frac{e^{ik_0 r_2}}{r_2}. \quad (47)$$

Making use of (45) and (46) it may be found that

$$A_{\alpha\beta}(\mathbf{q}_2; \boldsymbol{\rho}_1, z_1) = k_0^2 \int d^3r e^{-i\mathbf{k}(\mathbf{q}_2) \cdot \mathbf{r}} \times [\Pi_{\alpha\gamma}(\mathbf{q}_2) + R_{\alpha\gamma}(\mathbf{q}_2) e^{2ik_z(\mathbf{q}_2)z}] G_{\gamma\beta}(\mathbf{r}, \mathbf{r}_1) \eta(\mathbf{r}). \quad (48)$$

As before, we define a data function through a lattice Fourier transform of the angular part of the far-zone scattered field

$$\Phi_\alpha(\mathbf{q}_1, \mathbf{q}_2) = h^2 \sum_{\boldsymbol{\rho}_1} e^{i\mathbf{q}_1 \cdot \boldsymbol{\rho}_1} A_{\alpha\beta}(\mathbf{q}_2; \boldsymbol{\rho}_1, z_1) e_\beta. \quad (49)$$

Making use of the plane-wave decomposition (41) as well as the transverse bandlimiting of η , we find that

$$\Phi_\alpha(\mathbf{q}_1, \mathbf{q}_2) = \int dz \kappa_\alpha(\mathbf{q}_1, \mathbf{q}_2, z) \tilde{\eta}(\mathbf{q}_2 - \mathbf{q}_1, z) \quad (50)$$

where the vector integral operator κ_α is given by

$$\kappa_\alpha(\mathbf{q}_1, \mathbf{q}_2, z) = 2\pi i k_0^2 \left[\Pi_{\alpha\gamma}(\mathbf{q}_2) + R_{\alpha\gamma}(\mathbf{q}_2) e^{2ik_z(\mathbf{q}_2)z} \right] \times e^{-ik_z(\mathbf{q}_2)z} g_{\gamma\beta}(z, z_1; \mathbf{q}_1) e_\beta. \quad (51)$$

B. Collection Mode

We now return to the modality discussed in Section II-B in which an incident plane wave illuminates the sample and the scattered field is measured in the $z = z_2$ plane, as shown in Fig. 2. The incident field may be expressed as

$$E_\alpha^i(\mathbf{r}) = [\delta_{\alpha\beta} + R_{\alpha\beta}(\mathbf{q}_1) e^{2ik_z(\mathbf{q}_1)z}] e_\beta e^{i[\mathbf{q}_1 \cdot \boldsymbol{\rho} - k_z(\mathbf{q}_1)z]} \quad (52)$$

where e_β is the polarization of the incoming plane wave in the upper half space. The scattered field is measured at the point $\mathbf{r}_2 = (\boldsymbol{\rho}_2, z_2)$, where $z_2 > z_m$ for reflection mode and $z_2 < 0$ for transmission mode. Making use of (40) and (52), it may be seen that the scattered field is given by the expression

$$E_\alpha^s(\boldsymbol{\rho}_2, z_2; \mathbf{q}_1) = k_0^2 \int d^3r G_{\alpha\beta}(\mathbf{r}_2, \mathbf{r}) [\delta_{\beta\gamma} + R_{\beta\gamma}(\mathbf{q}_1) \times e^{2ik_z(\mathbf{q}_1)z}] p_\gamma e^{i[\mathbf{q}_1 \cdot \boldsymbol{\rho} - k_z(\mathbf{q}_1)z]} \eta(\mathbf{r}). \quad (53)$$

We now define a data function as the lattice Fourier transform of the scattered field

$$\Phi_\alpha(\mathbf{q}_1, \mathbf{q}_2) = h^2 \sum_{\boldsymbol{\rho}_2} e^{-i\mathbf{q}_2 \cdot \boldsymbol{\rho}_2} E_\alpha^s(\boldsymbol{\rho}_2, 0; \mathbf{q}_1). \quad (54)$$

Making use of the plane-wave decomposition (41) along with the assumed transverse bandlimiting of η , the data function is given by (50) with the operator κ_α of the form

$$\kappa_\alpha(\mathbf{q}_1, \mathbf{q}_2, z) = 2\pi i k_0^2 g_{\alpha\beta}(z_2, z; \mathbf{q}_2) \times [\delta_{\beta\gamma} + R_{\beta\gamma}(\mathbf{q}_1) e^{2ik_z(\mathbf{q}_1)z}] e_\gamma e^{-ik_z(\mathbf{q}_1)z}. \quad (55)$$

C. Two-Probe Mode

We now return to the case considered in Section II-C in which the illumination is provided by a point source of unit amplitude and polarization e_β located at $\mathbf{r}_1 = (\boldsymbol{\rho}_1, z_1)$, where $z_1 > z_m$. The scattered field is measured at $\mathbf{r}_2 = (\boldsymbol{\rho}_2, z_2)$, where $z_2 > z_m$ for reflection mode and $z_2 < 0$ for transmission mode. The incident field is given by (44) and the scattered field is therefore found by (40) to be of the form

$$E_\alpha^s(\boldsymbol{\rho}_2, z_2; \boldsymbol{\rho}_1, z_1) = k_0^2 \int d^3r G_{\alpha\gamma}(\mathbf{r}_2, \mathbf{r}) G_{\gamma\beta}(\mathbf{r}, \mathbf{r}_1) e_\beta \eta(\mathbf{r}). \quad (56)$$

$$g_{\alpha\beta}(z, z'; \mathbf{q}) = \begin{cases} \frac{1}{k_z(\mathbf{q})} [\Pi_{\alpha\beta}(\mathbf{q}) e^{ik_z(\mathbf{q})(z-z')} + R_{\alpha\beta}(\mathbf{q}) e^{ik_z(\mathbf{q})(z+z')}], & 0 \leq z' \leq z \\ \frac{1}{k_z(\mathbf{q})} [\Pi_{\alpha\beta}(\mathbf{q}) e^{ik_z(\mathbf{q})(z'-z)} + R_{\alpha\beta}(\mathbf{q}) e^{ik_z(\mathbf{q})(z'+z)}], & 0 \leq z < z' \\ \frac{1}{k'_z(\mathbf{q})} T'_{\alpha\beta}(\mathbf{q}) e^{i[k_z(\mathbf{q})z - k'_z(\mathbf{q})z']}, & z' < 0 \leq z \\ \frac{1}{k'_z(\mathbf{q})} T_{\alpha\beta}(\mathbf{q}) e^{i[k_z(\mathbf{q})z' - k'_z(\mathbf{q})z]}, & z < 0 \leq z' \end{cases} \quad (42)$$

The data function is defined as a lattice Fourier transform of the scattered field

$$\Phi_\alpha(\mathbf{q}_1, \mathbf{q}_2) = h^4 \sum_{\boldsymbol{\rho}_1, \boldsymbol{\rho}_2} e^{i(\mathbf{q}_1 \cdot \boldsymbol{\rho}_1 - \mathbf{q}_2 \cdot \boldsymbol{\rho}_2)} E_\alpha^s(\boldsymbol{\rho}_2, z_2; \boldsymbol{\rho}_1, z_1). \quad (57)$$

Using the plane-wave decomposition (41) and the transverse bandlimiting of η , the data function is found to be of the form (50) where κ_α is given by

$$\kappa_\alpha(\mathbf{q}_1, \mathbf{q}_2, z) = -(2\pi)^2 k_0^2 \sum_{\mathbf{q}'} g_{\alpha\gamma}(z_2, z; \mathbf{q}_2 + \mathbf{q}') \times g_{\gamma\beta}(z, z_1; \mathbf{q}_1 + \mathbf{q}') \quad (58)$$

with the sum taken over reciprocal lattice vectors.

IV. NEAR-FIELD INVERSE SCATTERING PROBLEM

In the near-field ISP, we wish to reconstruct the scattering potential $\eta(\mathbf{r})$ from the data function. To this end, the pseudoinverse solution to the integral equations (22) and (50) will be constructed. We begin with a brief review of the SVD of linear operators on Hilbert spaces [25].

Let K denote a linear operator with kernel $K(x, y)$ which maps the Hilbert space \mathcal{H}_1 into the Hilbert space \mathcal{H}_2 . The SVD of K is a representation of the form

$$K(x, y) = \sum_n \sigma_n g_n(x) f_n^*(y) \quad (59)$$

where σ_n is the singular value associated with the singular functions f_n and g_n . The $\{f_n\}$ and $\{g_n\}$ are orthonormal bases of \mathcal{H}_1 and \mathcal{H}_2 , respectively, and are eigenfunctions with eigenvalues σ_n^2 of the nonnegative self-adjoint operators K^*K and KK^*

$$K^*K f_n = \sigma_n^2 f_n \quad (60)$$

$$KK^* g_n = \sigma_n^2 g_n. \quad (61)$$

In addition, f_n and g_n are related by

$$K f_n = \sigma_n g_n \quad (62)$$

$$K^* g_n = \sigma_n f_n. \quad (63)$$

The pseudoinverse solution to the equation $Kf = g$ is defined to be the minimizer of $\|Kf - g\|$ with the smallest norm. This well-defined element $f^+ \in N(K)^\perp$ is unique and may be shown to be of the form $f^+ = K^+g$, where the pseudoinverse operator A^+ is given by $K^+ = K^*(KK^*)^{-1}$ and $N(K)^\perp$ is the orthogonal complement of the null space of K . The SVD of K may be used to express K^+ as

$$K^+(x, y) = \sum_n \frac{1}{\sigma_n} f_n(x) g_n^*(y). \quad (64)$$

The SVD approach is next applied to the inverse problem for scalar waves.

A. Scalar Case

The integral equation (22) may be rewritten in the form

$$\Phi(\mathbf{q}_1, \mathbf{q}_2) = \int d^3r K(\mathbf{q}_1, \mathbf{q}_2; \mathbf{r}) \eta(\mathbf{r}) \quad (65)$$

where the scattering operator $K(\mathbf{q}_1, \mathbf{q}_2; \mathbf{r})$ is given by

$$K(\mathbf{q}_1, \mathbf{q}_2; \mathbf{r}) = \exp[i(\mathbf{q}_1 - \mathbf{q}_2) \cdot \boldsymbol{\rho}] \kappa(\mathbf{q}_1, \mathbf{q}_2, z) \quad (66)$$

with the form of κ chosen appropriately for each of the experimental modalities we have considered.

In order to obtain the SVD of $K(\mathbf{q}_1, \mathbf{q}_2; \mathbf{r})$, it will prove useful to introduce the following identity:

$$K(\mathbf{q}_1, \mathbf{q}_2; \mathbf{r}) = \sum_{\mathbf{Q}} \exp(i\mathbf{Q} \cdot \boldsymbol{\rho}) \times \delta(\mathbf{Q} + \mathbf{q}_2 - \mathbf{q}_1) \kappa(\mathbf{Q} + \mathbf{q}_2, \mathbf{q}_2, z) \quad (67)$$

where δ denotes the Kronecker delta. Using this result, the matrix elements of the operator KK^* are found to be given by

$$KK^*(\mathbf{q}_1, \mathbf{q}_2; \mathbf{q}'_1, \mathbf{q}'_2) = \sum_{\mathbf{Q}} M(\mathbf{q}_2, \mathbf{q}'_2; \mathbf{Q}) \delta(\mathbf{Q} + \mathbf{q}_2 - \mathbf{q}_1) \times \delta(\mathbf{Q} + \mathbf{q}'_2 - \mathbf{q}'_1) \quad (68)$$

where

$$M(\mathbf{q}_2, \mathbf{q}'_2; \mathbf{Q}) = \int_0^{z_m} dz \kappa(\mathbf{Q} + \mathbf{q}_2, \mathbf{q}_2, z) \times \kappa^*(\mathbf{Q} + \mathbf{q}'_2, \mathbf{q}'_2, z). \quad (69)$$

To find the singular vectors $g_{\mathbf{Q}\mathbf{Q}'}$ of K which satisfy

$$KK^* g_{\mathbf{Q}\mathbf{Q}'} = \sigma_{\mathbf{Q}\mathbf{Q}'}^2 g_{\mathbf{Q}\mathbf{Q}'} \quad (70)$$

it will be useful to make the ansatz

$$g_{\mathbf{Q}\mathbf{Q}'}(\mathbf{q}_1, \mathbf{q}_2) = C_{\mathbf{Q}'}(\mathbf{q}_2; \mathbf{Q}) \delta(\mathbf{Q} + \mathbf{q}_2 - \mathbf{q}_1). \quad (71)$$

Equation (68) now implies that

$$\sum_{\mathbf{q}'} M(\mathbf{q}, \mathbf{q}'; \mathbf{Q}) C_{\mathbf{Q}'}(\mathbf{q}'; \mathbf{Q}) = \sigma_{\mathbf{Q}\mathbf{Q}'}^2 C_{\mathbf{Q}'}(\mathbf{Q}; \mathbf{Q}). \quad (72)$$

Thus, $C_{\mathbf{Q}'}(\mathbf{q}_2; \mathbf{Q})$ is an eigenvector of $M(\mathbf{Q})$ labeled by \mathbf{Q}' with eigenvalue $\sigma_{\mathbf{Q}\mathbf{Q}'}^2$. Since $M(\mathbf{Q})$ is self-adjoint, the $C_{\mathbf{Q}'}(\mathbf{Q}_2; \mathbf{q})$ may be taken to be orthonormal. Next, the $f_{\mathbf{Q}\mathbf{Q}'}$ may be found from the relation $K^* g_{\mathbf{Q}\mathbf{Q}'} = \sigma_{\mathbf{Q}\mathbf{Q}'} f_{\mathbf{Q}\mathbf{Q}'}$ and are given by

$$f_{\mathbf{Q}\mathbf{Q}'}(\mathbf{r}) = \frac{1}{\sigma_{\mathbf{Q}\mathbf{Q}'}} \sum_{\mathbf{q}} \exp(-i\mathbf{Q} \cdot \boldsymbol{\rho}) \times \kappa^*(\mathbf{Q} + \mathbf{q}, \mathbf{q}; z) C_{\mathbf{Q}'}^*(\mathbf{q}; \mathbf{Q}). \quad (73)$$

It follows that the SVD of $K(\mathbf{q}_1, \mathbf{q}_2; \mathbf{r})$ is given by the expression

$$K(\mathbf{q}_1, \mathbf{q}_2; \mathbf{r}) = \sum_{\mathbf{Q}, \mathbf{Q}'} \sigma_{\mathbf{Q}\mathbf{Q}'} f_{\mathbf{Q}\mathbf{Q}'}^*(\mathbf{r}) g_{\mathbf{Q}\mathbf{Q}'}(\mathbf{q}_1, \mathbf{q}_2). \quad (74)$$

The SVD (74) may now be used to obtain the pseudoinverse solution to the integral equation (65)

$$\eta^+(\mathbf{r}) = \sum_{\mathbf{q}_1, \mathbf{q}_2} K^+(\mathbf{r}; \mathbf{q}_1, \mathbf{q}_2) \Phi(\mathbf{q}_1, \mathbf{q}_2) \quad (75)$$

where $K^+(\mathbf{r}; \mathbf{q}_1, \mathbf{q}_2)$ is the pseudoinverse of $K(\mathbf{q}_1, \mathbf{q}_2; \mathbf{r})$. Using the result (64), the pseudoinverse K^+ may be seen to be

given by

$$K^+(r; q_1, q_2) = \sum_{Q, Q'} \frac{1}{\sigma_{QQ'}} f_{QQ'}(r) g_{QQ'}^*(q_1, q_2). \quad (76)$$

Substituting (71) and (73) into (76) and using the spectral decomposition

$$\sum_{Q'} \frac{1}{\sigma_{QQ'}} C_{Q'}(q; Q) C_{Q'}^*(q'; Q) = M^{-1}(q, q'; Q) \quad (77)$$

where $M^{-1}(q, q'; Q)$ is the qq' matrix element of $M^{-1}(Q)$ one obtains

$$\begin{aligned} \eta^+(r) &= \sum_{q_1, q_2, q'_2} \sum_Q \exp(-iQ \cdot \rho) \\ &\times \delta(Q + q_2 - q_1) M^{-1}(q_2, q'_2; Q) \\ &\times \kappa^*(Q + q'_2, q'_2, z) \Phi(q_1, q_2) \end{aligned} \quad (78)$$

which is the inversion formula for scalar NSOT.

B. Vector Case

The integral equation (50) may be rewritten in the form

$$\Phi_\alpha(q_1, q_2) = \int d^3r K_\alpha(q_1, q_2; r) \eta(r) \quad (79)$$

where

$$K_\alpha(q_1, q_2; r) = \exp[i(q_1 - q_2) \cdot \rho] \kappa_\alpha(q_1, q_2, z). \quad (80)$$

The SVD for the vector case is now obtained. Following the previous development, it may be found that the SVD of $K_\alpha(q_1, q_2; r)$ is of the form

$$K_\alpha(q_1, q_2; r) = \sum_{Q, Q'} \sigma_{QQ'} f_{QQ'}(r) g_{QQ'}^\alpha(q_1, q_2). \quad (81)$$

Here the singular functions are given by

$$g_{QQ'}^\alpha(q_1, q_2) = C_{Q'}^\alpha(q_2; Q) \delta(Q + q_2 - q_1) \quad (82)$$

$$\begin{aligned} f_{QQ'}(r) &= \frac{1}{\sigma_{QQ'}} \sum_{q \in \Lambda} \exp(-iQ \cdot \rho) \\ &\times \kappa^*(Q + q, q; z) C_{Q'}^*(q; Q). \end{aligned} \quad (83)$$

The $C_{Q'}^\alpha(q_2; Q)$ are the α, q_2 elements of eigenfunctions, $C_{Q'}(Q)$, labeled by Q' of the operator $M(Q)$ with matrix elements $M_\alpha^{\alpha'}(q_2, q'_2; Q)$ with eigenvalues $\sigma_{qq'}^2$. This means

$$\sum_{q'} M_\alpha^{\alpha'}(q, q'; Q) C_{Q'}^{\alpha'}(q'; Q) = \sigma_{QQ'}^2 C_{Q'}^\alpha(q; Q) \quad (84)$$

where

$$\begin{aligned} M_\alpha^{\alpha'}(q_2, q'_2; Q) &= \int_0^{z_m} dz \kappa_\alpha(Q + q_2, q_2, z) \kappa_{\alpha'}^* \\ &\times (Q + q'_2, q'_2, z). \end{aligned} \quad (85)$$

The pseudoinverse solution to the integral equation (86) is given by

$$\eta^+(r) = \sum_{q_1, q_2} K_\alpha^+(r; q_1, q_2) \Phi_\alpha(q_1, q_2) \quad (86)$$

where

$$K_\alpha^+(r; q_1, q_2) = \sum_{Q, Q'} \frac{1}{\sigma_{QQ'}} f_{QQ'}(r) g_{QQ'}^{\alpha*}(q_1, q_2). \quad (87)$$

More explicitly,

$$\begin{aligned} \eta^+(r) &= \sum_{q_1, q_2, q'_2} \sum_Q \exp(-iQ \cdot \rho) \delta(Q + q_2 - q_1) \\ &\times [M^{-1}(Q)]_\alpha^{\alpha'}(q_2, q'_2) \\ &\times \kappa_{\alpha'}^*(Q + q'_2, q'_2, z) \Phi_\alpha(q_1, q_2) \end{aligned} \quad (88)$$

which is the inversion formula for vector NSOT.

C. Regularization

It was demonstrated in [19] that the singular values of K may be very small and thus the ISP is ill posed. Accordingly, the kernel K^+ is highly singular and must be regularized to obtain a stable image reconstruction algorithm. Any of the standard approaches, such as Wiener filtering or the Tikhonov method may be used. For the SVD-derived pseudoinverse, these methods result in a modification of the singular values via the introduction of a suitable regularizer $R(\sigma)$ which replaces σ^{-1} in (76) and (87). The simplest choice for $R(\sigma)$ leads to truncation of the small singular values where we take

$$R(\sigma_Q) = \begin{cases} \sigma_Q^{-1}, & \sigma_Q \geq \sigma_{\min} \\ 0, & \sigma_Q < \sigma_{\min}. \end{cases} \quad (89)$$

for some σ_{\min} . As a variant of this approach, the regularizer

$$R(\sigma_Q) = \begin{cases} \sigma_Q^{-1}, & \sigma_Q \geq \varepsilon \max_{q'}(\sigma_Q) \\ 0, & \sigma_Q < \varepsilon \max_{q'}(\sigma_Q). \end{cases} \quad (90)$$

may be employed, where ε is the regularization parameter and $\max_{q'}(\sigma_Q)$ is the largest singular value for a fixed q' . This has the effect of conditioning each of the matrices $M(q')$ individually. If Tikhonov regularization is used

$$R(\sigma_Q) = \frac{\sigma_Q}{\lambda + \sigma_Q^2} \quad (91)$$

where λ is the regularization parameter. This choice leads to smoothing of $\tilde{\eta}^+(z, q')$ by penalizing functions with large L^2 norm. The regularization parameter λ must be chosen commensurate with the noise. Of course, under the assumption that the noise is Gaussian, this approach, that is least-squared error with additive Gaussian noise, is equivalent to maximum likelihood (ML). While photon-limited optical intensity signals exhibit Poissonian statistics, the data for this problem must be acquired interferometrically. The noise is then set or dominated by the reference field. Without entanglement, the noise is thus Gaussian for a reference field much larger than the scattered field [26].

V. NUMERICAL SIMULATIONS

To study the achievable resolution and stability of NSOT, we have performed a series of numerical simulations. We consider the collection-mode NSOT. The object was taken to consist of

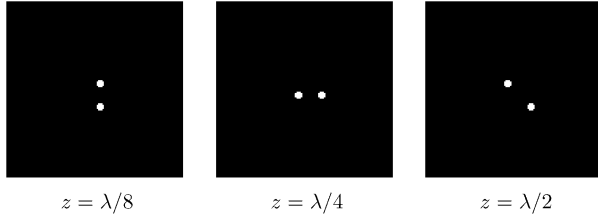


Fig. 4. Positions of point scatterers of the scattering object in a $\lambda \times \lambda$ window centered at the origin.

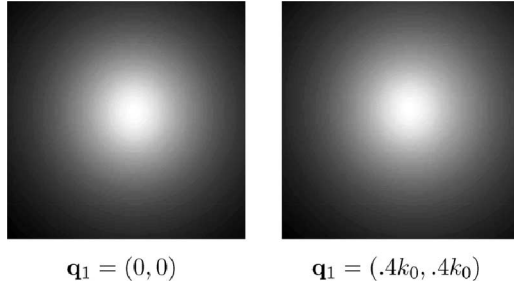


Fig. 5. Scalar case scattered field intensity distributions in a $\lambda \times \lambda$ window centered at the origin in the measurement plane ($z = 0$) with different incident transverse wave vectors. The intensities are normalized to unity at peak value.

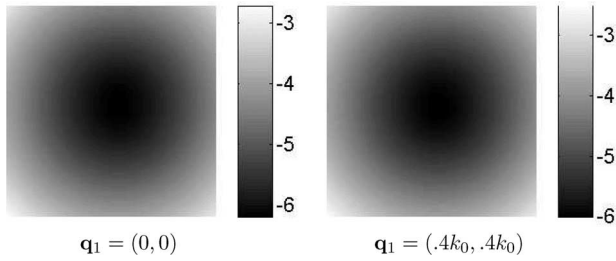


Fig. 6. Scalar case scattered field phase distributions in a $\lambda \times \lambda$ window centered at the origin in the measurement plane ($z = 0$) with different incident transverse wave vectors.

six point-like scatterers, shown in Fig. 4, located in three horizontal layers of different heights, in a cubic box of volume $\lambda \times \lambda \times \lambda$ just above the $z = 0$ plane, centered at the origin, where $\lambda = 2\pi/k_0$ is the free-space wavelength of the illumination field.

The data were simulated for 36 different illuminating plane waves with transverse-wave vectors \mathbf{q}_1 evenly sampled within the square $[0, k_0) \times [0, k_0)$. For the vector case, the incident plane waves are taken to have TE polarization, i.e., the polarization is normal to the wave vector and the surface normal of the interface. Raw NSOM data are shown in Figs. 5 and 6 for three different illuminating fields. The scattered field was computed in an $8\lambda \times 8\lambda$ window in the measurement plane $z = 0$.

Data were simulated to be consistent with a point-like probe taking measurements at a spacing $\lambda/16$ in a window $32\lambda \times 32\lambda$. From the data, the transverse 2-D Fourier components of the scatterer structure were computed in the square $[-8k_0, 7k_0] \times [-8k_0, 7k_0]$ with spacing k_0 in the frequency domain. The reconstruction is realized by truncated SVD, with singular functions computed in the interval $z \in [0, \lambda]$, within which the entire scatterer exists. The truncation thresholds were

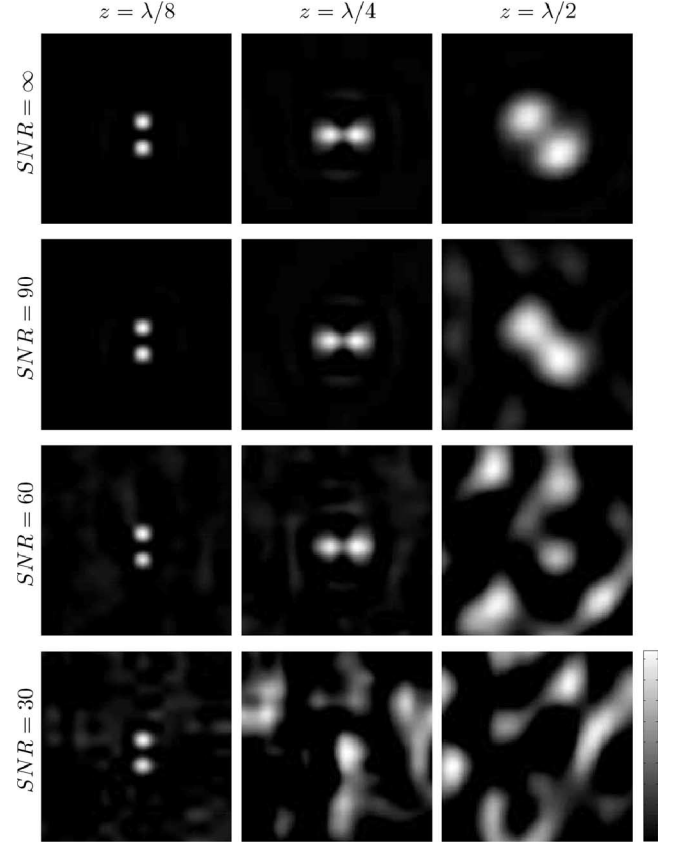


Fig. 7. Scalar case reconstructions from data of different noise levels in different layers in a $\lambda \times \lambda$ window centered at the origin. The images have been normalized to the maximum value in each frame and plotted on the linear scale shown in the lower right corner.

decided by numerical experiments. Evidently, these settings are consistent with our assumption that the scatterer is within the cube mentioned above and no other priors are included. For the scalar case, the reconstructed structure is shown in Fig. 7. The vector-case reconstruction, which was reconstructed from only one component of the vector-scattered fields, is shown in Fig. 8. It can be seen that the reconstruction is better near the measurement plane than away from it. The obtainable resolution is at least equal to the distance from the measurement plane to the reconstruction plane, which is the distance over which high-resolution information is decaying to a barely retrievable level.

In Figs. 7 and 8, reconstructions from data contaminated with randomly generated Gaussian noise of different levels of signal-to-noise ratio (SNR) are also shown. It can be seen that the quality of reconstruction is more susceptible to noise away from the measurement plane than close to it. Again, this is due to the exponential decay of the evanescent field with distance. Regularization was varied from plane to plane, though the general trend of worsening resolution with distance from the measurement is observed even when the regularization is the same for the entire reconstruction. For both sets of reconstruction when $\text{SNR} = \infty$ or 90, for the closest layer, the SVD was truncated at 10^{-8} of the largest singular value, and for the midlayer at 10^{-10} , and for the farthest layer at 10^{-12} . When $\text{SNR} = 60$, the SVD was truncated at 10^{-7} of the largest singular value for the closest

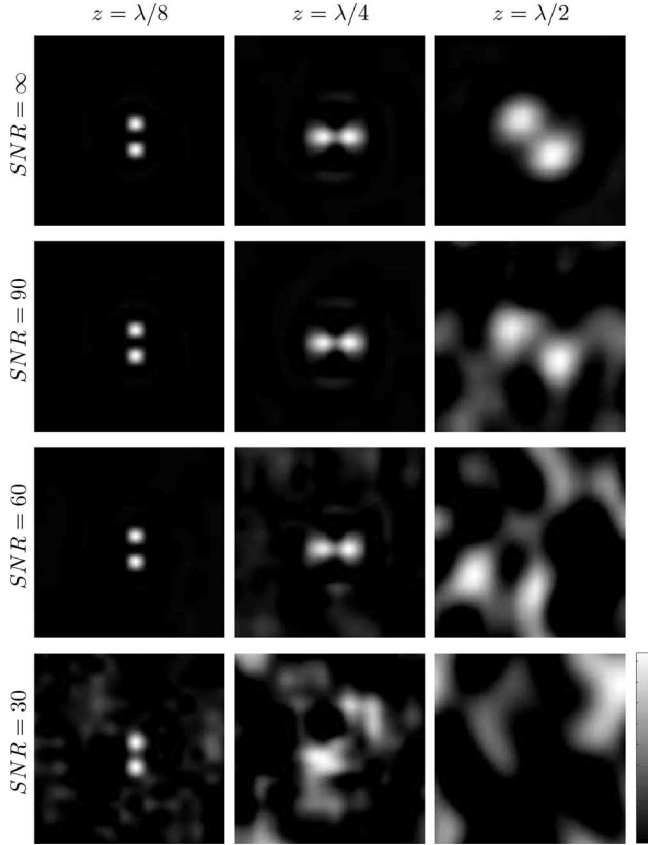


Fig. 8. Vector case reconstructions from data of different noise levels in different layers in a $\lambda \times \lambda$ window centered at the origin. The images have been normalized to the maximum value in each frame and plotted on the linear scale shown in the lower right corner.

layer, at 10^{-8} for the midlayer, and at 10^{-9} for the farthest layer. Finally, for $\text{SNR} = 30$, the SVD was truncated at 10^{-6} , 10^{-7} , and 10^{-8} of the largest singular value for increasingly deeper planes. It may be noted that some of the reconstructions look underregularized as they simply appear noisy. However, past a certain point, increasing regularization only serves to cause the entire reconstruction to be dominated by the features of the top layer. Thus, we observe that at increased levels of noise, deep features are simply not recoverable.

VI. DISCUSSION

We have presented an analysis of the inverse scattering problem for different variations of the NSOM. The analysis expands upon our earlier work presented in [19], extending those results to new modalities and taking account of polarization effects and boundaries.

The inversion formulae obtained here make possible two new capabilities in the NSOM. First, we can reconstruct the 3-D structure of a sample on subwavelength scales. Second, we can unambiguously map the measured optical field to the susceptibility of the sample. This connection had previously been unclear even for effectively 2-D samples.

There remain in this area several open problems. We are investigating the means by which to reconstruct strongly scat-

tering objects, better models for the tip-sample interaction, and various methods to include any available prior knowledge about the object.

APPENDIX

The reflection and transmission coefficients for the vector field in the upper and lower half spaces may be obtained by first projecting the field onto the TE/TM basis, then multiplying by the appropriate Fresnel coefficients, and projecting back onto the original basis. We denote by R and R' the reflection in the upper and lower half spaces, respectively, and we denote by T and T' the transmission from upper half space into the lower half space and transmission from the lower half space into the upper half space, respectively. P and P' are projection operators onto the TE/TM basis in the upper and lower half spaces, respectively and are shortly. The wave vectors \mathbf{k} , \mathbf{k}' , $\tilde{\mathbf{k}}$, and $\tilde{\mathbf{k}}'$ are defined the same way as in the previous sections, where the prime symbol ($'$) denotes quantities in lower half space and the tilde symbol (\sim) denotes wave vectors with negative z component. They are functions of the transverse-wave vector \mathbf{q} . We find for the reflection and transmission tensors, the expressions

$$R_{\alpha\beta}(\mathbf{q}) = P_{\gamma\alpha}(\mathbf{k})r_{\gamma\delta}(\mathbf{k}, \mathbf{k}')P_{\delta\beta}(\tilde{\mathbf{k}}) \quad (92)$$

$$R'_{\alpha\beta}(\mathbf{q}) = P'_{\gamma\alpha}(\tilde{\mathbf{k}}')r'_{\gamma\delta}(\mathbf{k}, \mathbf{k}')P'_{\delta\beta}(\mathbf{k}') \quad (93)$$

$$T_{\alpha\beta}(\mathbf{q}) = P'_{\gamma\alpha}(\tilde{\mathbf{k}}')t_{\gamma\delta}(\mathbf{k}, \mathbf{k}')P_{\delta\beta}(\tilde{\mathbf{k}}) \quad (94)$$

$$T'_{\alpha\beta}(\mathbf{q}) = P_{\gamma\alpha}(\mathbf{k})t'_{\gamma\delta}(\mathbf{k}, \mathbf{k}')P'_{\delta\beta}(\mathbf{k}') \quad (95)$$

where r and t are given by

$$r = \begin{pmatrix} \frac{n^2 k_z - k'_z}{n^2 k_z + k'_z} & 0 \\ 0 & \frac{k_z - k'_z}{k_z + k'_z} \end{pmatrix} \quad (96)$$

and

$$t = \begin{pmatrix} \frac{2n k_z}{n^2 k_z + k'_z} & 0 \\ 0 & \frac{2k_z}{k_z + k'_z} \end{pmatrix} \quad (97)$$

and $r' = -r$, $t' = k'_z t / k_z$. The projection operator onto the TE/TM basis in the upper half space is given by

$$P(\mathbf{k}) = \frac{1}{\sqrt{k_x^2 + k_y^2 k_0}} \times \begin{pmatrix} -k_x k_z & -k_y k_z & k_x^2 + k_y^2 \\ -k_y k_0 & k_x k_0 & 0 \end{pmatrix} \quad (98)$$

and in the lower half space

$$P'(\mathbf{k}') = \frac{1}{\sqrt{k_x^2 + k_y^2 n k_0}} \times \begin{pmatrix} -k_x k'_z & -k_y k'_z & k_x^2 + k_y^2 \\ -k_y n k_0 & k_x n k_0 & 0 \end{pmatrix}. \quad (99)$$

The following identities are useful in derivations in the text:

$$P_{\gamma\alpha}(\mathbf{k})P_{\beta\gamma}(\mathbf{k}) = \delta_{\alpha\beta} - \frac{k_\alpha k_\beta}{k_0^2} \quad (100)$$

$$P_{\alpha\gamma}(\mathbf{k})P_{\beta\gamma}(\mathbf{k}) = \delta_{\alpha\beta} \quad (101)$$

$$T'_{\gamma\alpha}(\mathbf{q})R_{\gamma\beta}(\mathbf{q}) = \frac{-k'_z}{k_z}R'_{\gamma\alpha}(\mathbf{q})T_{\gamma\beta}(\mathbf{q}) \quad (102)$$

$$R_{\mu\alpha}(\mathbf{q})R_{\mu\beta}(\mathbf{q}) + \frac{k'_z}{k_z}T_{\mu\alpha}(\mathbf{q})T_{\mu\beta}(\mathbf{q}) = \delta_{\alpha\beta} - \frac{\tilde{k}_\alpha\tilde{k}_\beta}{k_0^2}. \quad (103)$$

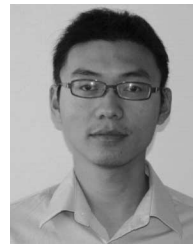
Equations (100) and (101) simply reflect the fact that P is a projection operator with the usual properties that $P^2 = P$ and P is the identity on the subspace into which it projects. Equation (102) may be understood to be a statement of Stokes reciprocity.

ACKNOWLEDGMENT

This work was carried out in part during the sabbatical of J. C. Schotland at the University of Illinois at Urbana-Champaign.

REFERENCES

- [1] E. Synge, "A suggested method for extending microscopic resolution into the ultra-microscopic region," *Philos. Mag.*, vol. 6, p. 356, 1928.
- [2] R. C. Reddick, R. J. Warmack, and T. L. Ferrell, "A new form of scanning optical microscopy," *Phys. Rev. B*, vol. 39, p. 767, 1989.
- [3] D. Courjon and C. Bainier, "Near field microscopy and near field optics," *Rep. Prog. Phys.*, vol. 57, p. 989, 1994.
- [4] C. Girard and A. Dereux, "Near-field optics theories," *Rep. Prog. Phys.*, vol. 59, p. 657, 1996.
- [5] E. Ash and G. Nicholls, "Super-resolution aperture scanning microscope," *Nature*, vol. 237, p. 510, 1972.
- [6] A. Lewis, M. Isaacson, A. Harootunian, and A. Muray, "Development of a 500Å spatial resolution light microscope: I. Light is efficiently transmitted through $\lambda/16$ diameter apertures," *Ultramicroscopy*, vol. 13, p. 227, 1984.
- [7] D. W. Pohl, W. Denk, and M. Lanz, "Optical stethoscopy: Image recording with resolution $\lambda/20$," *Appl. Phys. Lett.*, vol. 44, p. 651, 1984.
- [8] E. Betzig and J. K. Trautman, "Near-field optics: Microscopy, spectroscopy, and surface modification beyond the diffraction limit," *Science*, vol. 257, p. 189, 1992.
- [9] R. Dickson, D. Norris, Y. L. Tzeng, and W. Moerner, "Three-dimensional imaging of single molecules solvated in pores of poly(acrylamide) gels," *Science*, vol. 274, p. 966, 1996.
- [10] N. Garcia and M. Nieto Vesperinas, "Near-field optics inverse-scattering reconstruction of reflective surfaces," *Opt. Lett.*, vol. 18, p. 2090, 1993.
- [11] —, "Direct solution to the inverse scattering problem for surfaces from near-field intensities without phase retrieval," *Opt. Lett.*, vol. 20, p. 949, 1995.
- [12] R. Carminati, J.-J. Greffet, N. Garcia, and M. Nieto-Vesperinas, "Direct reconstruction of surfaces from near-field intensity under spatially incoherent illumination," *Opt. Lett.*, vol. 21, p. 501, 1996.
- [13] J.-J. Greffet, A. Sentenac, and R. Carminati, "Two-dimensional numerical simulation of the photon scanning tunneling microscope: Concept of transfer function," *Opt. Commun.*, vol. 116, p. 20, 1995.
- [14] E. Wolf, "Principles and development of diffraction tomography," in *Trends in Optics*, A. Consortini, Ed. San Diego, CA: Academic, 1996, pp. 83–110.
- [15] T. Habashy and E. Wolf, "Reconstruction of scattering potentials from incomplete data," *J. Mod. Opt.*, vol. 41, p. 1679, 1994.
- [16] A. Schatzberg and A. J. Devaney, "Super-resolution in diffraction tomography," *Inverse Probl.*, vol. 8, p. 149, 1992.
- [17] R. Hillenbrand and F. Keilmann, "Complex optical constants on a sub-wavelength scale," *Phys. Rev. Lett.*, vol. 85, p. 3029, 2000.
- [18] R. Hillenbrand, B. Knoll, and F. Keilmann, "Pure optical contrast in scattering-type scanning near-field microscopy," *J. Opt. Microsc.*, vol. 202, pp. 77–83, 2001.
- [19] P. S. Carney and J. C. Schotland, "Inverse scattering for near-field microscopy," *Appl. Phys. Lett.*, vol. 77, p. 2798, 2000.
- [20] —, "Near-field tomography," in *Inside Out*, G. Uhlmann, Ed. Cambridge, U.K.: Cambridge Univ. Press, 2003, pp. 133–168.
- [21] P. S. Carney, R. A. Frazin, S. I. Bozhevolnyi, V. S. Volkov, A. Boltasseva, and J. C. Schotland, "Computational lens for the near field," *Phys. Rev. Lett.*, vol. 92, p. 163903, 2004.
- [22] J.-J. Greffet and R. Carminati, "Image formation in near-field optics," *Prog. Surf. Sci.*, vol. 56, p. 133, 1997.
- [23] Z. Yu and S. Boseck, "Scanning acoustic microscopy and its applications to material characterization," *Rev. Mod. Phys.*, vol. 67, p. 863, 1995.
- [24] M. Born and E. Wolf, *Principles of Optics*, 7th ed. Cambridge, U.K.: Cambridge Univ. Press, 1999.
- [25] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
- [26] R. A. Frazin, D. G. Fischer, and P. S. Carney, "Information content of the near-field: Two-dimensional samples," *J. Opt. Soc. Amer. A*, vol. 21, pp. 1050–1057, 2004.



Jin Sun received the M.S. degree in electrical engineering in 2005 from the University of Illinois at Urbana-Champaign, Urbana, where he is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering.

His current research interests include the theoretical modeling of near-field optical microscopy and tomography and associated inverse scattering problems.



P. Scott Carney received the B.S. degree in engineering physics from the University of Illinois at Urbana-Champaign, Urbana, in 1994, and the Ph.D. degree from the University of Rochester, Rochester, NY, in 1999.

He was a Postdoctoral Associate at Washington University, St. Louis, MO, from 1999 to 2001 when he joined the faculty of the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign. He is a theorist and his current research interests include inverse problems, imaging,

coherence theory, and other branches of optical physics.



John C. Schotland received the M.D. and Ph.D. degrees from the University of Pennsylvania, Philadelphia.

He is an Associate Professor of Bioengineering and Electrical Engineering at the University of Pennsylvania. His research is focused on theoretical optical physics with applications to biomedical imaging and nano-optics. His current research interests include optical tomography, optical imaging of nanoscale systems, and the use of quantum states of light for optical imaging. Inverse problems, particularly inverse scattering problems, are a unifying theme that connects these areas.