11-15-2007

Strong Tip Effects in Near-field Scanning Optical Tomography

Jin Sun
University of Illinois

P. Scott Carney
University of Illinois

John C. Schotland
University of Pennsylvania, schotland@seas.upenn.edu

Strong tip effects in near-field scanning optical tomography

Jin Sun and P. Scott Carney[a)]
Department of Electrical and Computer Engineering, and The Coordinated Science Laboratory, University of Illinois, Urbana, Illinois 61801, USA

John C. Schotland[b)]
Department of Electrical and Computer Engineering, and The Coordinated Science Laboratory, University of Illinois, Urbana, Illinois 61801, USA
and the Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

(Received 19 July 2007; accepted 19 September 2007; published online 21 November 2007)

A model for the interaction of the scanning probe in near-field scanning optical microscopy is presented. Multiple scattering of the illuminating field with the probe is taken into account. The implications of this so-called strong tip model for the solution of the associated inverse scattering problem are studied through simulations. © *2007 American Institute of Physics*.
[DOI: 10.1063/1.2812545]

I. INTRODUCTION

Near-field scanning optical microscopy (NSOM),[1–8] as a surface-imaging method, has drawn considerable attention due to the remarkable resolution obtained that far exceeds the diffraction limit of conventional optical microscopy. Applications of NSOM range from the inspection of biological samples to semiconductor devices, where spectroscopic concerns or sample handling requirements dictate the use of lower frequency fields and yet high spatial resolution is still desired.

Two commonly practiced modalities of NSOM are the illumination mode and the collection mode. In illumination mode, a tapered optical fiber probe serves as the source of illumination and the scattered far field is recorded as the probe tip scans over the sample. In collection mode, the sample is illuminated from the far zone. The probe acts as a detector to collect the scattered field as it scans over the sample. In both modalities, the recorded field intensity as a function of the tip position is interpreted as a two-dimensional image of the sample.

One of the fundamental difficulties of NSOM is the interpretation of the images. Although the sample may present a complicated three-dimensional structure, the image nonetheless is only two-dimensional. Under certain simplifying assumptions[9–11] such as homogeneity of the sample, NSOM images may be related to surface structure. However, more generally, the effects of variations in topography and the optical properties of the sample have proven to be indistinguishable.[12] To resolve this ambiguity, it is desired to solve the NSOM inverse scattering problem (ISP) which consists of reconstructing the three-dimensional sample structure, or the spatial dependence of the dielectric susceptibility of the sample, from measurement of the scattered field. This technique is known as near-field scanning optical tomography (NSOT).[13,14]

To solve the NSOT ISP, an appropriate model of the forward problem must be constructed. The form of the forward model affects the solvability of the ISP, the computational cost of solving the problem, and the accuracy of the solution. In previous work on NSOT,[13,14] the sample was assumed to be weakly scattering and its interaction with the optical near field was treated perturbatively, within the accuracy of the first Born approximation. Such a linear forward model gives a readily solvable ISP, and a fast reconstruction algorithm based on the singular value decomposition (SVD) of the forward operator was presented. Other authors have recently considered the inverse problem beyond the linearized model where strong scattering in the sample plays an important role.[15–18]

The treatment of the tip as a passive point source or point detector in the current studies of NSOT may be unrealistic. The tip used in NSOM is typically made of a tapered, metal-coated optical fiber with a subwavelength-sized aperture. When serving as a light source, as in illumination mode, the aperture radiates as a pointlike dipole. As a detector in collection mode, the aperture collects the external field and couples it into the fiber. In both cases, the tip behaves as a pointlike scatterer.[19] Since the same fiber tip may be used in illumination mode or collection mode, the incident field will, in collection mode, first scatter from the tip and then subsequently illuminate the sample. Likewise, in illumination mode, the field will scatter from the tip after scattering from the sample. For both modalities, this extra interaction with the probe tip gives rise to a correction to the previously considered model. Using this idea, it has been shown[20] that in photon scanning tunneling microscopy (PSTM), a modality of near-field microscopy quite similar to NSOM, the scattering effect of the tip must be included in the forward model in order to obtain a correct reconstruction. The tip was modeled as a pointlike scatterer, and second-order scattering with the tip was included in the analysis.

In this paper, the effects of multiple scattering from the tip in both illumination mode and collection mode NSOT are considered. It is assumed that the sample is weakly scattering

[a)]Electronic mail: carney@uiuc.edu
[b)]Electronic mail: schotland@seas.upenn.edu

and thus a linear forward model is obtained. The ISP for this model is solved, and the results are compared with those from the ideal case by numerical simulations. Finally, the relationship between the tip strength and the solvability of the ISP is discussed.

II. FORWARD PROBLEM

Consider a near-field probe located above a dielectric half-space with constant index of refraction n. The sample, described by a dielectric susceptibility $\eta(\mathbf{r})$, is assumed to be in the near zone of the probe. The probe is taken to be described by a susceptibility $\chi(\mathbf{r})$. A monochromatic incident field $\mathbf{E}^i$ obeying the reduced vector wave equation

$$\nabla \times \nabla \times \mathbf{E}^i - k^2 \mathbf{E}^i = 0, \tag{1}$$

is scattered from the tip-sample system, where $k=k_0$ for $z \geq 0$ and $k=nk_0$ for $z<0$. The total field $\mathbf{E}=\mathbf{E}^i+\mathbf{E}^s$, where $\mathbf{E}^s$ is the scattered field, satisfies the reduced wave equation

$$\nabla \times \nabla \times \mathbf{E} - k^2 \mathbf{E} = 4\pi k_0^2(\eta + \chi)\mathbf{E}. \tag{2}$$

An integral equation for $\mathbf{E}^s$ can be obtained,

$$\mathbf{E}^s(\mathbf{r}) = k_0^2 \int d^3r' \mathbf{G}(\mathbf{r},\mathbf{r}') \cdot \mathbf{E}(\mathbf{r}')[\eta(\mathbf{r}') + \chi(\mathbf{r}')]. \tag{3}$$

Here, $\mathbf{G}$ is the half-space Green's tensor which satisfies the equation

$$\nabla \times \nabla \times \mathbf{G}(\mathbf{r},\mathbf{r}') - k^2 \mathbf{G}(\mathbf{r},\mathbf{r}') = 4\pi\delta(\mathbf{r}-\mathbf{r}')\mathbf{I}, \tag{4}$$

where $\mathbf{I}$ is identity. The Green's tensor may be expressed in the plane-wave decomposition

$$\mathbf{G}(\mathbf{r},\mathbf{r}') = \int d^2q \mathbf{g}(z,z';\mathbf{q}) \exp[i\mathbf{q}\cdot(\boldsymbol{\rho}-\boldsymbol{\rho}')], \tag{5}$$

where $\mathbf{g}(z,z';\mathbf{q})$, the plane-wave component, is given by[21]

$$g_{\alpha\beta}(z,z';\mathbf{q}) = \frac{i}{2\pi}\begin{cases} \frac{1}{k_z(\mathbf{q})}[\mathcal{D}_{\alpha\beta}(\mathbf{q})e^{ik_z(\mathbf{q})(z-z')} + \mathcal{R}_{\alpha\beta}(\mathbf{q})e^{ik_z(\mathbf{q})(z+z')}] & 0 \leq z' \leq z \\ \frac{1}{k_z(\mathbf{q})}[\tilde{\mathcal{D}}_{\alpha\beta}(\mathbf{q})e^{ik_z(\mathbf{q})(z'-z)} + \mathcal{R}_{\alpha\beta}(\mathbf{q})e^{ik_z(\mathbf{q})(z'+z)}] & 0 \leq z < z' \\ \frac{1}{k_z'(\mathbf{q})}\mathcal{T}'_{\alpha\beta}(\mathbf{q})e^{i[k_z(\mathbf{q})z - k_z'(\mathbf{q})z']} & z' < 0 \leq z \\ \frac{1}{k_z(\mathbf{q})}\mathcal{T}_{\alpha\beta}(\mathbf{q})e^{i[k_z(\mathbf{q})z' - k_z'(\mathbf{q})z]} & z < 0 \leq z' \end{cases}, \tag{6}$$

where $k_z(\mathbf{q}) \equiv \sqrt{k_0^2 - q^2}$ and $k_z'(\mathbf{q}) \equiv \sqrt{n^2k_0^2 - q^2}$. With this notation, the wave vectors in the $z \geq 0$ half-space and $z<0$ half-space may be written as $\mathbf{k}(\mathbf{q}) \equiv (\mathbf{q}, k_z(\mathbf{q}))$ and $\mathbf{k}'(\mathbf{q}) \equiv (\mathbf{q}, k_z'(\mathbf{q}))$, respectively. The polarization tensors $\mathcal{D}(\mathbf{q})$ and $\tilde{\mathcal{D}}(\mathbf{q})$, the reflection tensor $\mathcal{R}(\mathbf{q})$, and the transmission tensors $\mathcal{T}(\mathbf{q})$ and $\mathcal{T}'(\mathbf{q})$ are given in the Appendix.

The scattered field $\mathbf{E}^s$ may be calculated perturbatively from Eq. (3). It is useful to define the integral operators $\mathbf{S}$ and $\mathbf{T}_0$ such that

$$\mathbf{S}\cdot\mathbf{E}(\mathbf{r}) \equiv \int d^3r' \mathbf{S}(\mathbf{r},\mathbf{r}')\cdot\mathbf{E}(\mathbf{r}'), \tag{7}$$

and

$$\mathbf{T}_0\cdot\mathbf{E}(\mathbf{r}) \equiv \int d^3r' \mathbf{T}_0(\mathbf{r},\mathbf{r}')\cdot\mathbf{E}(\mathbf{r}). \tag{8}$$

Note that $\mathbf{S}$ and $\mathbf{T}_0$ represent a single scattering event from either the sample or the tip. The kernels of $\mathbf{S}$ and $\mathbf{T}_0$ are given by the expressions

$$\mathbf{S}(\mathbf{r},\mathbf{r}') = k_0^2 \mathbf{G}(\mathbf{r},\mathbf{r}')\eta(\mathbf{r}'), \tag{9}$$

and

$$\mathbf{T}_0(\mathbf{r},\mathbf{r}') = k_0^2 \mathbf{G}(\mathbf{r},\mathbf{r}')\chi(\mathbf{r}'). \tag{10}$$

Using the above definitions and iterating Eq. (3), we find that $\mathbf{E}^s$ is given by

$$\mathbf{E}^s = \sum_{n=1}^{\infty} (\mathbf{S}+\mathbf{T}_0)^n \cdot \mathbf{E}^i. \tag{11}$$

When the sample is weakly scattering, the series for scattered field may be truncated at first order in η to yield

$$\mathbf{E}^s = [\mathbf{T}_0 + \mathbf{T}_0^2 + \cdots + (\mathbf{I} + \mathbf{T}_0 + \mathbf{T}_0^2 + \cdots) \times \mathbf{S}(\mathbf{I} + \mathbf{T}_0 + \mathbf{T}_0^2 + \cdots)] \cdot \mathbf{E}^i. \tag{12}$$

A. The effective tip

It will prove convenient to introduce the effective tip-scattering operator $\mathbf{T} = \Sigma_{j=1}^{\infty} \mathbf{T}_0^j$. Evidently, $\mathbf{T}$ provides a nonperturbative solution of Eqs. (2) and (3) for the case that $\eta = 0$,

$$\mathbf{E}^s = \mathbf{T}\cdot\mathbf{E}^i = \mathbf{T}_0\cdot(\mathbf{E}^i + \mathbf{E}^s), \tag{13}$$

and so may be calculated in a variety of ways. In near-field optics, it is generally the case that the tip is made as sharp as possible to obtain a highly localized field. A number of authors have addressed the problem of modeling the field scattered from such tips.[22,23] For simplicity, it is assumed here that the tip is pointlike and located at $\mathbf{r}_0$ so that $\chi(\mathbf{r}) = \alpha_0\delta(\mathbf{r}-\mathbf{r}_0)$. That is, the tip is described by a point polarizability with isotropic susceptibility α_0 so that an applied electric field produces at the tip location an induced dipole with magnitude and orientation determined by the applied

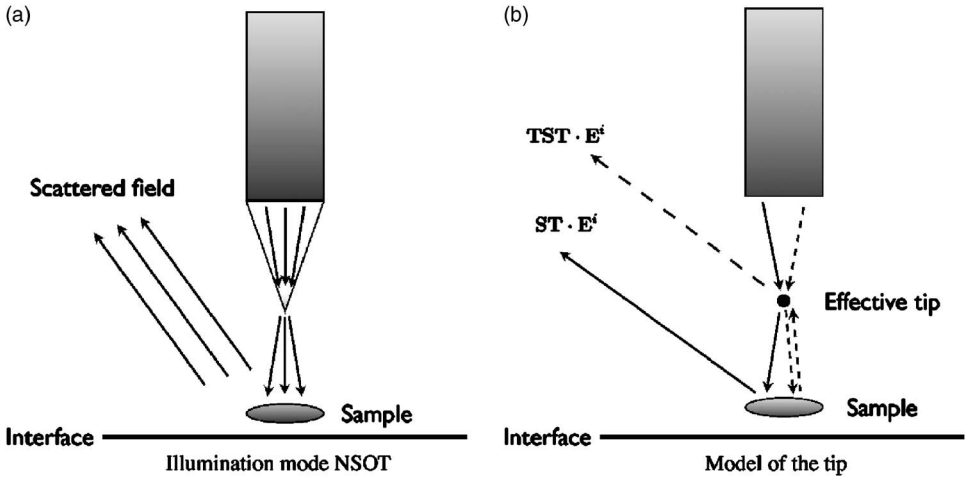


FIG. 1. The simplified model of the illumination tip.

field. This model should not be confused with a fixed dipole illumination which is insensitive to the incident field. It should be noted that other models for the tip may be substituted with straightforward changes to the following analysis. For the case of the pointlike tip, it can be seen that **T** becomes

$$\mathbf{T}(\mathbf{r},\mathbf{r}') = k_0^2\alpha_0\delta(\mathbf{r}'-\mathbf{r}_0)(\mathbf{G}(\mathbf{r},\mathbf{r}') + k_0^2\alpha_0\mathbf{G}(\mathbf{r},\mathbf{r}_0) \cdot \left\{\mathbf{I} + \sum_{j=1}^{\infty}[k_0^2\alpha\mathbf{G}(\mathbf{r}_0,\mathbf{r}_0)]^j\right\} \cdot \mathbf{G}(\mathbf{r}_0,\mathbf{r}')). \tag{14}$$

The presence of the delta function allows $\mathbf{r}'$ to be set to $\mathbf{r}_0$ and the infinite series may be formally summed to obtain

$$\mathbf{T}(\mathbf{r},\mathbf{r}') = k_0^2\alpha_0\delta(\mathbf{r}'-\mathbf{r}_0)\mathbf{G}(\mathbf{r},\mathbf{r}_0) \cdot \{\mathbf{I} + [\mathbf{I} - k_0^2\alpha_0\mathbf{G}(\mathbf{r}_0,\mathbf{r}_0)]^{-1} \cdot \mathbf{G}(\mathbf{r}_0,\mathbf{r}_0)\}. \tag{15}$$

The above expression for **T** contains a divergence which is isolated in the term $\mathbf{G}(\mathbf{r}_0,\mathbf{r}_0)$ and is the result of the assumption that the tip is pointlike. The divergence may be regularized by standard methods[24] to reintroduce a finite size scale associated with the physical tip size Λ. The divergent part of the Green's tensor can be expressed in momentum space with an appropriate cutoff,

$$G_{\alpha\beta}(\mathbf{r}_0,\mathbf{r}_0) = \frac{1}{(2\pi)^3}\int_{|\mathbf{k}|<2\pi/\Lambda} d^3k \frac{4\pi}{k^2-k_0^2}(\delta_{\alpha\beta} - k_0^{-2}k_\alpha k_\beta) + \frac{2ik_0\delta_{\alpha\beta}}{3} + G^R_{\alpha\beta}(\mathbf{r}_0,\mathbf{r}_0), \tag{16}$$

where it is assumed that the tip is located in the $z \geq 0$ half-space. Here, $\mathbf{G}^R$ is the part of the Green's tensor associated with reflection from the interface,

$$\mathbf{G}^R(\mathbf{r}_0,\mathbf{r}_0) = \int d^2q \frac{\mathcal{R}(\mathbf{q})}{k_z(\mathbf{q})} e^{2ik_z(\mathbf{q})z_0}, \tag{17}$$

which is finite and need not be regularized. The integral in Eq. (16) may be evaluated to obtain

$$\mathbf{G}(\mathbf{r}_0,\mathbf{r}_0) = \frac{2}{3}\left(\frac{1}{\pi\Lambda} - \frac{4\pi}{\Lambda^3k_0^2} + ik_0\right)\mathbf{I} + \mathbf{G}^R(\mathbf{r}_0,\mathbf{r}_0). \tag{18}$$

This result may be used to evaluate the expression for the tip operator, Eq. (15), from which it may be seen that

$$\mathbf{T}(\mathbf{r},\mathbf{r}') = k_0^2\delta(\mathbf{r}'-\mathbf{r}_0)\mathbf{G}(\mathbf{r},\mathbf{r}_0) \cdot \boldsymbol{\alpha}_e, \tag{19}$$

where $\boldsymbol{\alpha}_e$ is the effective tip polarizability,

$$\boldsymbol{\alpha}_e = \{\mathbf{I} + [\mathbf{I} - k_0^2\alpha_0\mathbf{G}(\mathbf{r}_0,\mathbf{r}_0)]^{-1} \cdot \mathbf{G}(\mathbf{r}_0,\mathbf{r}_0)\}\alpha_0. \tag{20}$$

Note that the effective polarizability, $\boldsymbol{\alpha}_e$, depends on the tip height above the substrate, z_0. Rather than dwell on the precise value of $\boldsymbol{\alpha}_e$ obtained from the foregoing analysis, it is perhaps of more practical importance to simply note the form of the solution obtained. In practice, near-field probe tips are neither points nor spheres, nor even cones, and so it is expected that, having established the plausibility of the form expressed in Eq. (19), the tip polarizability will need to be calibrated for each instrument.

It follows from Eq. (12) that the scattered electric field, to first order in η, is given by the sum

$$\mathbf{E}^s = (\mathbf{S} + \mathbf{T} + \mathbf{TS} + \mathbf{ST} + \mathbf{TST}) \cdot \mathbf{E}^i. \tag{21}$$

Here, the **S** and **T** terms are usually considered background and may be subtracted from the data. The **ST** and **TS** terms were the subject of earlier work.[13,14] The **TST** term was introduced in an *ad hoc* fashion in[20] and its role in the near-field ISP is the main subject of this work. By definition, in the weak-tip regime the **TST** term is neglected, in the strong tip regime it is retained.

B. Illumination mode

In illumination mode, the sample is illuminated by a small aperture at the end of a tapered fiber as illustrated in Fig. 1. The tip may be regarded as a pointlike scatterer located at the point $\mathbf{r}_1 = (\boldsymbol{\rho}_1, z_1)$. Since the incident field $\mathbf{E}^i$ emanating from the fiber must pass through the small aperture, i.e., the tip, before it interacts with the sample, the $\mathbf{S}\cdot\mathbf{E}^i$ and $\mathbf{TS}\cdot\mathbf{E}^i$ terms in Eq. (21) do not contribute. Consequently, the scattered field at the point $\mathbf{r}_2$ in the far zone, with the background subtracted, may be expressed as

$$\mathbf{E}^s = (\mathbf{S}_f\mathbf{T} + \mathbf{T}_f\mathbf{ST}) \cdot \mathbf{E}^i, \tag{22}$$

where $\mathbf{S}_f$ and $\mathbf{T}_f$ are the far-field asymptotic approximations of $\mathbf{S}$ and $\mathbf{T}$,

$$\mathbf{S}_f(\mathbf{r}_2,\mathbf{r}') = \frac{k_0^2 e^{ik_0 r_2}}{r_2} e^{-i\mathbf{k}(\mathbf{q}_2)\cdot\mathbf{r}'}[\mathcal{D}(\mathbf{q}_2) + \mathcal{R}(\mathbf{q}_2)e^{2ik_z(\mathbf{q}_2)z'}]\eta(\mathbf{r}'), \tag{23}$$

and

$$\mathbf{T}_f(\mathbf{r}_2,\mathbf{r}') = \frac{k_0^2 e^{ik_0 r_2}}{r_2} e^{-i\mathbf{k}(\mathbf{q}_2)\cdot\mathbf{r}'}[\mathcal{D}(\mathbf{q}_2) + \mathcal{R}(\mathbf{q}_2)e^{2ik_z(\mathbf{q}_2)z'}]\boldsymbol{\alpha}_e\delta(\mathbf{r}' - \mathbf{r}_0), \tag{24}$$

where $\mathbf{q}_2$ is defined so that $\mathbf{r}_2$ is in the direction of the outgoing wave vector $\mathbf{k}(\mathbf{q}_2)$ in the far zone.

The measured quantity, denoted Ψ, will be taken to be the projection of the scattered field onto a particular polarization, $\mathbf{p}_2$,

$$\Psi = \mathbf{p}_2 \cdot (\mathbf{S}_f\mathbf{T} + \mathbf{T}_f\mathbf{ST}) \cdot \mathbf{E}^i. \tag{25}$$

Making use of Eqs. (23) and (24) and taking $\mathbf{p}_1 = \mathbf{E}^i(\mathbf{r}_1)$ and $\mathbf{r}_0 = \mathbf{r}_1$, it is seen that Ψ behaves as an outgoing spherical wave which can be expressed in terms of the scattering amplitude A,

$$\Psi \sim A(\boldsymbol{\rho}_1,\mathbf{q}_2)\frac{e^{ik_0 r_2}}{r_2}, \tag{26}$$

where

$$\begin{aligned} A(\boldsymbol{\rho}_1,\mathbf{q}_2) &= \int d^3r'\,\mathbf{p}_2 \cdot \{k_0^4 e^{-i\mathbf{k}(\mathbf{q}_2)\cdot\mathbf{r}'}[\mathcal{D}(\mathbf{q}_2) + \mathcal{R}(\mathbf{q}_2)e^{2ik_z(\mathbf{q}_2)z'}] \\ &\quad + k_0^6 e^{-i\mathbf{k}(\mathbf{q}_2)\cdot\mathbf{r}_1}[\mathcal{D}(\mathbf{q}_2) + \mathcal{R}(\mathbf{q}_2)e^{2ik_z(\mathbf{q}_2)z_1}]\boldsymbol{\alpha}_e\mathbf{G}(\mathbf{r}_1,\mathbf{r}')\} \\ &\quad \times \mathbf{G}(\mathbf{r}',\mathbf{r}_1)\boldsymbol{\alpha}_e \cdot \mathbf{p}_1\eta(\mathbf{r}'). \end{aligned} \tag{27}$$

Measurements of Ψ are obtained for each position of the tip $\mathbf{r}_1$ as the tip scans over the $z = z_1$ plane with spacing h in both transverse directions. The data function $\Phi(\mathbf{q}_1,\mathbf{q}_2)$ is defined to be the lattice Fourier transform of $A(\boldsymbol{\rho}_1,\mathbf{q}_2)$ with respect to $\boldsymbol{\rho}_1$,

$$\Phi(\mathbf{q}_1,\mathbf{q}_2) = \left(\frac{h}{2\pi}\right)^2 \sum_{\boldsymbol{\rho}_1} e^{i\mathbf{q}_1\cdot\boldsymbol{\rho}_1} A(\boldsymbol{\rho}_1,\mathbf{q}_2), \tag{28}$$

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 FBZ$=[-\pi/h,\pi/h]\times[-\pi/h,\pi/h]$. Making use of Eqs. (5) and (6) and the identity

$$\sum_{\boldsymbol{\rho}} e^{i\mathbf{q}\cdot\boldsymbol{\rho}_1} = \left(\frac{2\pi}{h}\right)^2 \sum_{\mathbf{q}'} \delta(\mathbf{q} - \mathbf{q}'), \tag{29}$$

where $\mathbf{q}'$ denotes a reciprocal lattice vector, the data function may be written as

$$\Phi(\mathbf{q}_1,\mathbf{q}_2) = \sum_{\mathbf{q}'} \int dz[K_1(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}',z) + K_2(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}',z)]\tilde{\eta}(\mathbf{q}_2 - \mathbf{q}_1 + \mathbf{q}',z), \tag{30}$$

where $\tilde{\eta}(\mathbf{q},z) = \int d^2\rho\,\eta(\mathbf{r})e^{-i\mathbf{q}\cdot\boldsymbol{\rho}}$ is the transverse Fourier transform of $\eta(\mathbf{r})$. The operators K_1 and K_2, arising from the $\mathbf{S}_f\mathbf{T}$ and $\mathbf{T}_f\mathbf{ST}$ terms, respectively, are given by

$$K_1(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}',z) = k_0^4\mathbf{p}_2 \cdot [\mathcal{D}(\mathbf{q}_2)e^{-ik_z(\mathbf{q}_2)z} + \mathcal{R}(\mathbf{q}_2)e^{ik_z(\mathbf{q}_2)z}] \times \mathbf{g}(z,z_1,\mathbf{q}_1 - \mathbf{q}')\boldsymbol{\alpha}_e \cdot \mathbf{p}_1 \tag{31}$$

and

$$\begin{aligned} K_2(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}',z) &= k_0^6\mathbf{p}_2 \cdot [\mathcal{D}(\mathbf{q}_2)e^{-ik_z(\mathbf{q}_2)z_1} + \mathcal{R}(\mathbf{q}_2)e^{ik_z(\mathbf{q}_2)z_1}] \\ &\quad \times \boldsymbol{\alpha}_e \int d^2q\,\mathbf{g}(z_1,z,\mathbf{q}) \\ &\quad \times \mathbf{g}(z,z_1,\mathbf{q} + \mathbf{q}_1 - \mathbf{q}_2 - \mathbf{q}')\boldsymbol{\alpha}_e \cdot \mathbf{p}_1. \end{aligned} \tag{32}$$

Assuming the sample function $\eta(\mathbf{r})$ is transversely bandlimited to the FBZ, that is, if $\tilde{\eta}(\mathbf{q},z) = 0$ for $\mathbf{q} \notin$ FBZ, then the sum over $\mathbf{q}'$ in Eq. (30) may be truncated and only the $\mathbf{q}' = 0$ term contributes to the data function Φ_s for $\mathbf{q}_2 - \mathbf{q}_1 \in$ FBZ. The data associated with points outside the FBZ are redundant. Thus, Eq. (30) becomes

$$\Phi(\mathbf{q}_1,\mathbf{q}_2) = \int dz[K_1(\mathbf{q}_1,\mathbf{q}_2,z) + K_2(\mathbf{q}_1,\mathbf{q}_2,z)]\tilde{\eta}(\mathbf{q}_2 - \mathbf{q}_1,z), \tag{33}$$

where $\mathbf{q}_2 - \mathbf{q}_1 \in$ FBZ, and K_1 and K_2 are evaluated by Eqs. (31) and (32) with $\mathbf{q}' = 0$, which are

$$K_1(\mathbf{q}_1,\mathbf{q}_2,z) = k_0^4\mathbf{p}_2 \cdot [\mathcal{D}(\mathbf{q}_2)e^{-ik_z(\mathbf{q}_2)z} + \mathcal{R}(\mathbf{q}_2)e^{ik_z(\mathbf{q}_2)z}]\mathbf{g}(z,z_1,\mathbf{q}_1)\boldsymbol{\alpha}_e \cdot \mathbf{p}_1 \tag{34}$$

and

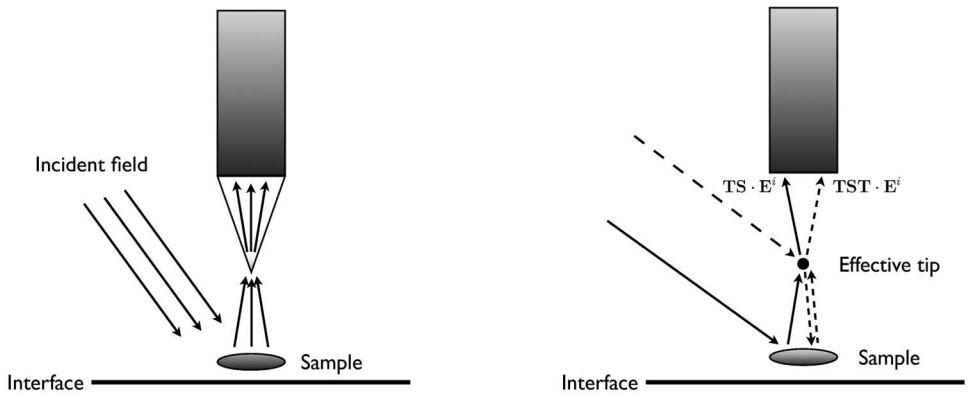
$$K_2(\mathbf{q}_1,\mathbf{q}_2,z) = k_0^6\mathbf{p}_2 \cdot [\mathcal{D}(\mathbf{q}_2)e^{-ik_z(\mathbf{q}_2)z_1} + \mathcal{R}(\mathbf{q}_2)e^{ik_z(\mathbf{q}_2)z_1}] \times \alpha_e \int d^2q\,\mathbf{g}(z_1,z,\mathbf{q})\mathbf{g}(z,z_1,\mathbf{q} + \mathbf{q}_1 - \mathbf{q}_2)\boldsymbol{\alpha}_e \cdot \mathbf{p}_1. \tag{35}$$

C. Collection mode

In collection mode, the scattered field is collected by the small aperture at the end of the fiber, which is treated as a pointlike scatterer located at the point $\mathbf{r}_2 = (\boldsymbol{\rho}_2, z_2)$ that scatters the field into the fiber face located at $\mathbf{r}'$ as illustrated in Fig. 2. Since the scattered field must pass through the small aperture before being collected by the fiber, the $\mathbf{S}$ and $\mathbf{ST}$ terms do not contribute. Subtracting background terms, Eq. (21) thus becomes

$$\mathbf{E}^s = (\mathbf{TS} + \mathbf{TST}) \cdot \mathbf{E}^i. \tag{36}$$

As in illumination mode, a scalar measured quantity representing the projection of the scattered field onto a certain polarization direction $\mathbf{p}_2$ is considered,



(a) Collection mode NSOT (b) Model of the tip

FIG. 2. The simplified model of the collection tip.

$$\Psi = \mathbf{p}_2 \cdot (\mathbf{TS} + \mathbf{TST}) \cdot \mathbf{E}^i. \tag{37}$$

The illuminating field is taken to be a plane wave of polarization $\mathbf{p}_1$ in the direction of the wave vector $\mathbf{k}(\mathbf{q}_1)$,

$$\mathbf{E}^i(\mathbf{r}) = e^{i\mathbf{q}_1\cdot\boldsymbol{\rho} - ik_z(\mathbf{q}_1)z}[\mathbf{I} + \mathcal{R}(\mathbf{q}_1)e^{2ik_z(\mathbf{q}_1)z}] \cdot \mathbf{p}_1. \tag{38}$$

It may be seen that

$$\begin{aligned}\Psi(\mathbf{q}_1,\boldsymbol{\rho}_2) = \int d^3r' \mathbf{p}_2 \cdot \mathbf{G}(\mathbf{r}_2',\mathbf{r}_2)\boldsymbol{\alpha}_e\mathbf{G}(\mathbf{r}_2,\mathbf{r}') \\ \times\{k_0^4 e^{i\mathbf{q}_1\cdot\boldsymbol{\rho}' - ik_z(\mathbf{q}_1)z'}[\mathbf{I} + \mathcal{R}(\mathbf{q}_1)e^{2ik_z(\mathbf{q}_1)z'}] \\ + k_0^6 e^{i\mathbf{q}_1\cdot\boldsymbol{\rho}_2 - ik_z(\mathbf{q}_1)z_2}\mathbf{G}(\mathbf{r}',\mathbf{r}_2) \\ \times\boldsymbol{\alpha}_e[\mathbf{I} + \mathcal{R}(\mathbf{q}_1)e^{2ik_z(\mathbf{q}_1)z_2}]\} \cdot \mathbf{p}_1\eta(\mathbf{r}').\end{aligned} \tag{39}$$

For each direction of illumination $\mathbf{q}_1$, the tip scans over the $z=z_2'$ plane with spacing h to sample the scattered field. This enables us to define the data function $\Phi(\mathbf{q}_1,\mathbf{q}_2)$ as the lattice Fourier transform of $\Psi(\mathbf{q}_1,\boldsymbol{\rho}_2)$ with respect to $\boldsymbol{\rho}_2$,

$$\Phi(\mathbf{q}_1,\mathbf{q}_2) = \left(\frac{h}{2\pi}\right)^2 \sum_{\boldsymbol{\rho}_2} e^{-i\mathbf{q}_2\cdot\boldsymbol{\rho}_2}\Psi(\mathbf{q}_1,\boldsymbol{\rho}_2), \tag{40}$$

where the sum over $\boldsymbol{\rho}_2$ is carried out over all lattice vectors and $\mathbf{q}_2$ belongs to the FBZ of the lattice. Making use of the plane-wave representation of the Green's tensor and proceeding much as in the case of illumination mode, the data function can be seen to be of the form of Eq. (33), with

$$\begin{aligned}K_1(\mathbf{q}_1,\mathbf{q}_2,z) = k_0^4\mathbf{p}_2 \cdot \mathbf{G}(\mathbf{r}_2',\mathbf{r}_2)\boldsymbol{\alpha}_e\mathbf{g}(z_2,z,\mathbf{q}_2) \\ \times[\mathbf{I}e^{-ik_z(\mathbf{q}_1)z} + \mathcal{R}(\mathbf{q}_1)e^{ik_z(\mathbf{q}_1)z}] \cdot \mathbf{p}_1\end{aligned} \tag{41}$$

and

$$\begin{aligned}K_2(\mathbf{q}_1,\mathbf{q}_2,z) \\ = k_0^6\mathbf{p}_2 \cdot \mathbf{G}(\mathbf{r}_2',\mathbf{r}_2)\boldsymbol{\alpha}_e \int d^2q\mathbf{g}(z_2,z,\mathbf{q})\mathbf{g}(z,z_2,\mathbf{q}+\mathbf{q}_1-\mathbf{q}_2) \\ \times\boldsymbol{\alpha}_e[\mathbf{I}e^{-ik_z(\mathbf{q}_1)z_2} + \mathcal{R}(\mathbf{q}_1)e^{ik_z(\mathbf{q}_1)z_2}] \cdot \mathbf{p}_1,\end{aligned} \tag{42}$$

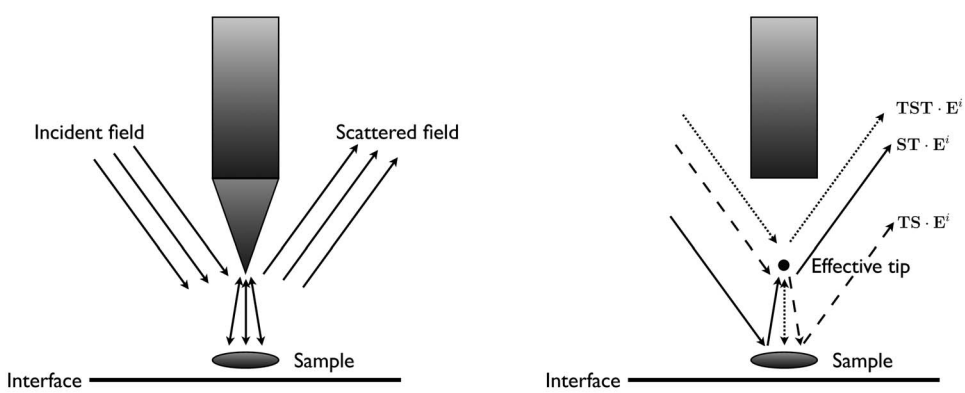
where K_1 and K_2 represent the **TS** and **TST** terms in Eq. (37), respectively.

D. Apertureless mode

In apertureless mode, a sharp metal tip located at $\mathbf{r}_t = (\boldsymbol{\rho}_t, z_t)$ is excited by an external incident field, serving as a pointlike secondary source to illuminate the sample as illustrated in Fig. 3. The scattered near-field interacts with the tip and is converted into propagating modes in the far field. The data collected consist of all terms in Eq. (21), with the exclusion of the background terms **S** and **T**,

$$\mathbf{E}^s = (\mathbf{T}_f\mathbf{S} + \mathbf{S}_f\mathbf{T} + \mathbf{T}_f\mathbf{ST}) \cdot \mathbf{E}^i. \tag{43}$$

As in the illumination mode, the scattered field is measured in the far zone. A scalar measured quantity representing the



(a) Apertureless NSOT (b) Model of the apertureless tip

FIG. 3. The simplified model of the apertureless tip.

projection of the scattered field onto a certain polarization direction $\mathbf{p}_2$ is considered,

$$\Psi = \mathbf{p}_2 \cdot (\mathbf{T}_f\mathbf{S} + \mathbf{S}_f\mathbf{T} + \mathbf{T}_f\mathbf{ST}) \cdot \mathbf{E}^i. \tag{44}$$

The illuminating field is taken to be the same as in the collection mode, and is given by Eq. (38). It may be seen that Ψ behaves as an outgoing spherical wave with scattering amplitude $A = A_1 + A_2$, where

$$\begin{aligned} A_1(\mathbf{q}_1,\mathbf{q}_2,\boldsymbol{\rho}_t) &= k_0^4 \int d^3r' \mathbf{p}_2 \cdot \{e^{-i\mathbf{k}(\mathbf{q}_2)\cdot\mathbf{r}_t}[\mathcal{D}(\mathbf{q}_2) + \mathcal{R}(\mathbf{q}_2)e^{2ik_z(\mathbf{q}_2)z_t}] \\ &\times \boldsymbol{\alpha}_e \mathbf{G}(\mathbf{r}_t,\mathbf{r}')[\mathbf{I} + \mathcal{R}(\mathbf{q}_1)e^{2ik_z(\mathbf{q}_1)z'}] \\ &\times e^{i\mathbf{q}_1\cdot\boldsymbol{\rho}' - ik_z(\mathbf{q}_1)z'} + e^{-i\mathbf{k}(\mathbf{q}_2)\cdot\mathbf{r}'}[\mathcal{D}(\mathbf{q}_2) + \mathcal{R}(\mathbf{q}_2)e^{2ik_z(\mathbf{q}_2)z'}] \\ &\times \mathbf{G}(\mathbf{r}',\mathbf{r}_t)\boldsymbol{\alpha}_e[\mathbf{I} + \mathcal{R}(\mathbf{q}_1)e^{2ik_z(\mathbf{q}_1)z_t}] \\ &\times e^{i\mathbf{q}_1\cdot\boldsymbol{\rho}_t - ik_z(\mathbf{q}_1)z_t}\} \cdot \mathbf{p}_1 \eta(\mathbf{r}') \end{aligned} \tag{45}$$

corresponds to the **TS** and **ST** terms, and

$$\begin{aligned} A_2(\mathbf{q}_1,\mathbf{q}_2,\boldsymbol{\rho}_t) &= k_0^6 \int d^3r' \mathbf{p}_2 \cdot \{e^{-i\mathbf{k}(\mathbf{q}_2)\cdot\mathbf{r}_t}[\mathcal{D}(\mathbf{q}_2) + \mathcal{R}(\mathbf{q}_2)e^{2ik_z(\mathbf{q}_2)z_t}] \\ &\times \boldsymbol{\alpha}_e \mathbf{G}(\mathbf{r}_t,\mathbf{r}')\mathbf{G}(\mathbf{r}',\mathbf{r}_t)\boldsymbol{\alpha}_e[\mathbf{I} + \mathcal{R}(\mathbf{q}_1)e^{2ik_z(\mathbf{q}_1)z_t}] \\ &\times e^{i\mathbf{q}_1\cdot\boldsymbol{\rho}_t - ik_z(\mathbf{q}_1)z_t}\} \cdot \mathbf{p}_1 \eta(\mathbf{r}') \end{aligned} \tag{46}$$

corresponds to the **TST** term.

For each pair of $\mathbf{q}_1$ and $\mathbf{q}_2$, the tip scans over the $z = z_t$ plane with spacing h to sample the scattered field. We thus define the data function $\Phi(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_t)$ as the lattice Fourier transform of $A(\mathbf{q}_1,\mathbf{q}_2,\boldsymbol{\rho}_t)$ with respect to $\boldsymbol{\rho}_t$,

$$\Phi(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_t) = \left(\frac{h}{2\pi}\right)^2 \sum_{\boldsymbol{\rho}_t} e^{-i\mathbf{q}_t\cdot\boldsymbol{\rho}_t} A(\mathbf{q}_1,\mathbf{q}_2,\boldsymbol{\rho}_t), \tag{47}$$

where the sum over $\boldsymbol{\rho}_t$ is carried out over all lattice vectors and $\mathbf{q}_t$ belongs to the FBZ of the lattice. Following similar procedures as in the previous sections, the data function may be found to be

$$\begin{aligned} \Phi(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_t) = \int dz[&K_1(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_t,z) \\ &+ K_2(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_t,z)]\tilde{\eta}(\mathbf{q}_2 - \mathbf{q}_1 + \mathbf{q}_t,z), \end{aligned} \tag{48}$$

where

$$\begin{aligned} K_1(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_t,z) = k_0^4 \mathbf{p}_2 \cdot \{&[\mathcal{D}(\mathbf{q}_2)e^{-ik_z(\mathbf{q}_2)z_t} \\ &+ \mathcal{R}(\mathbf{q}_2)e^{ik_z(\mathbf{q}_2)z_t}]\boldsymbol{\alpha}_e \mathbf{g}(\mathbf{q}_2 + \mathbf{q}_t,z_t,z) \\ &\times [\mathbf{I}e^{-ik_z(\mathbf{q}_1)z} + \mathcal{R}(\mathbf{q}_1)e^{ik_z(\mathbf{q}_1)z}] \\ &+ [\mathcal{D}(\mathbf{q}_2)e^{-ik_z(\mathbf{q}_2)z} + \mathcal{R}(\mathbf{q}_2)e^{ik_z(\mathbf{q}_2)z}] \\ &\times \mathbf{g}(\mathbf{q}_1 - \mathbf{q}_t,z,z_t)\boldsymbol{\alpha}_e[\mathbf{I}e^{-ik_z(\mathbf{q}_1)z_t} \\ &+ \mathcal{R}(\mathbf{q}_1)e^{ik_z(\mathbf{q}_1)z_t}]\} \cdot \mathbf{p}_1 \end{aligned} \tag{49}$$

and

$$\begin{aligned} K_2(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_t,z) &= k_0^6 \mathbf{p}_2 \cdot [\mathcal{D}(\mathbf{q}_2)e^{-ik_z(\mathbf{q}_2)z_t} + \mathcal{R}(\mathbf{q}_2)e^{ik_z(\mathbf{q}_2)z_t}]\boldsymbol{\alpha}_e \\ &\times \int d^2q \mathbf{g}(\mathbf{q},z_t,z)\mathbf{g}(\mathbf{q} + \mathbf{q}_1 - \mathbf{q}_2 - \mathbf{q}_t,z,z_t)\boldsymbol{\alpha}_e[\mathbf{I}e^{-ik_z(\mathbf{q}_1)z_t} \\ &+ \mathcal{R}(\mathbf{q}_1)e^{ik_z(\mathbf{q}_1)z_t}] \cdot \mathbf{p}_1, \end{aligned} \tag{50}$$

where K_1 and K_2 represent the **TS**+**ST** and **TST** terms in Eq. (44), respectively.

III. INVERSE PROBLEM

The inverse problem may be solved in a manner similar to that employed in earlier works.[13,25,26] It is convenient to work with a subset of the data. By fixing $\mathbf{q}_2 - \mathbf{q}_1 \equiv \mathbf{q}_0$ as a constant, the integral Eq. (33) can be rewritten as

$$\Phi(\mathbf{q}_1,\mathbf{q}_1 + \mathbf{q}_0) = \int dz[K_1(\mathbf{q}_1,\mathbf{q}_1 + \mathbf{q}_0,z) + K_2(\mathbf{q}_1,\mathbf{q}_1 + \mathbf{q}_0,z)]\tilde{\eta}(\mathbf{q}_0,z). \tag{51}$$

Similar results may be obtained for Eq. (48).

Solving Eq. (51) consists of finding the inverse of the forward operator $K \equiv K_1 + K_2$ that maps the data function Φ, now a function of $\mathbf{q}_1$ only, to the $\mathbf{q}_0$ Fourier component of the sample structure, $\tilde{\eta}$, as a function of z. Since Φ_s and $\tilde{\eta}$ belong to different Hilbert spaces, the inverse of the forward operator K is not defined. As an alternative, we seek the pseudoinverse solution $\tilde{\eta}^+$, which is defined as the minimizer of $\|K\eta - \Phi\|$, with minimum L_2 norm. This pseudoinverse solution is uniquely defined. It can be obtained through the SVD of the forward operator, and is given by[13,25,26]

$$\tilde{\eta}^+(\mathbf{q}_0,z) = \sum_{\mathbf{q}_1',\mathbf{q}_1} K^*(\mathbf{q}_1',\mathbf{q}_1' + \mathbf{q}_0,z)M^+(\mathbf{q}_1',\mathbf{q}_1;\mathbf{q}_0)\Phi(\mathbf{q}_1,\mathbf{q}_1 + \mathbf{q}_0), \tag{52}$$

where $K^*(\mathbf{q}_1',\mathbf{q}_1' + \mathbf{q}_0,z)$ is the complex conjugate of $K(\mathbf{q}_1',\mathbf{q}_1' + \mathbf{q}_0,z) \equiv K_1(\mathbf{q}_1',\mathbf{q}_1' + \mathbf{q}_0,z) + K_2(\mathbf{q}_1',\mathbf{q}_1' + \mathbf{q}_0,z)$, and $M^+(\mathbf{q}_1',\mathbf{q}_1;\mathbf{q}_0)$ is the element of the regularized pseudoinverse matrix of the self-adjoint matrix M, defined by its matrix elements

$$M(\mathbf{q}_1,\mathbf{q}_1';\mathbf{q}_0) = \int dz K(\mathbf{q}_1,\mathbf{q}_1 + \mathbf{q}_0,z)K^*(\mathbf{q}_1',\mathbf{q}_1' + \mathbf{q}_0,z). \tag{53}$$

The linearized scattering operator is the sum of two parts, K_1 and K_2. When the incident field is changed, K_2 picks up a change in an overall multiplicative factor, but the variables in the integration are not involved. The independent variables of the contribution to the data associated with K_2 are thus restricted to being two-dimensional and so K_2 is not invertible. It can further be observed that K_1 is first order in $\boldsymbol{\alpha}_e$, the effective tip strength, whereas K_2 is second order in $\boldsymbol{\alpha}_e$. As a consequence, when the tip is weakly scattering, i.e., $k_0^3\alpha_0 \ll 1$, the contribution of K_1 is generally greater than K_2 and therefore is the major contribution to the data function. In the limiting case that the multiple tip-scattering effect is

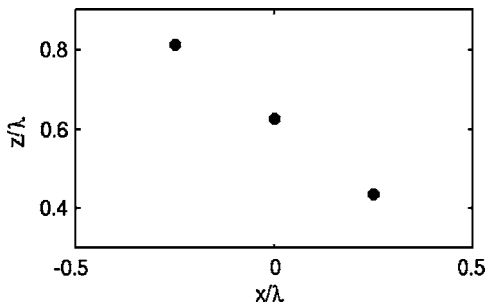


FIG. 4. Point scatterers in the $y=0$ plane.

negligible, the problem reduces to the ideal case assumed in Ref. 13. On the other hand, when K_2 becomes more significant and therefore cannot be neglected, the inverse problem becomes more ill-conditioned. In the limiting case of a very strong tip, the K_2 operator becomes dominant and meaningful solutions to the inverse problem cannot be obtained for the measurement schemes considered here.

IV. NUMERICAL SIMULATIONS

To study the effect of the strong tip correction to the NSOT ISP, a series of numerical simulations was performed. Collection mode NSOT is considered. As shown in Fig. 4, the object was taken to consist of three pointlike scatterers in vacuum, located at different heights in the $y=0$ plane, in a cubic box of size $\lambda\times\lambda\times\lambda$ just above the $z=0$ plane (that is, the interface with the substrate), and centered at the origin, where $\lambda=2\pi/k_0$ is the free-space wavelength of the illuminating field. The point scatterers of the sample were taken to have a polarizability of λ^3 and the effective polarizability of the tip was taken to be $\lambda^3/100$. The substrate was taken to have a real index of refraction equal to 1.5.

The data were simulated for 64 different illuminating plane waves with transverse wave vectors, $\mathbf{q}_1$, uniformly sampled within the square $[0,k_0)\times[0,k_0)$. The incident plane waves are taken to have TM polarization, that is, the electric field is polarized in the plane defined by the wave vector and the normal direction of the interface. The scattered field was simulated with both the weak tip model and the strong tip corrected model. Data were simulated at a discrete set of points with spacing $\lambda/16$ in a window $32\lambda\times32\lambda$ in the $z=\lambda$ measurement plane. Since the measurement plane is at a fixed position the effective polarizability is constant. From the data, samples of the transverse two-dimensional Fourier components of the scatterer structure were computed. 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 reconstructions of the scatterer from the simulated data are shown in Fig. 5. These reconstructions are computed in a $\lambda\times\lambda$ window in three horizontal layers in which the three pointlike scatterers reside. Reconstructions in the vertical $y=0$ cross section are also shown with certain degrees of contrast enhancement. In column (a), the forward data were simulated under the weak tip assumption and the reconstructions were made with only the week tip assumption. In column (b) the strong tip term, that is the **TST** term, is taken into account in both the forward simulation and the solution of the inverse problem. It may be noticed that the strong tip reconstruction has a wider point-spread function than the

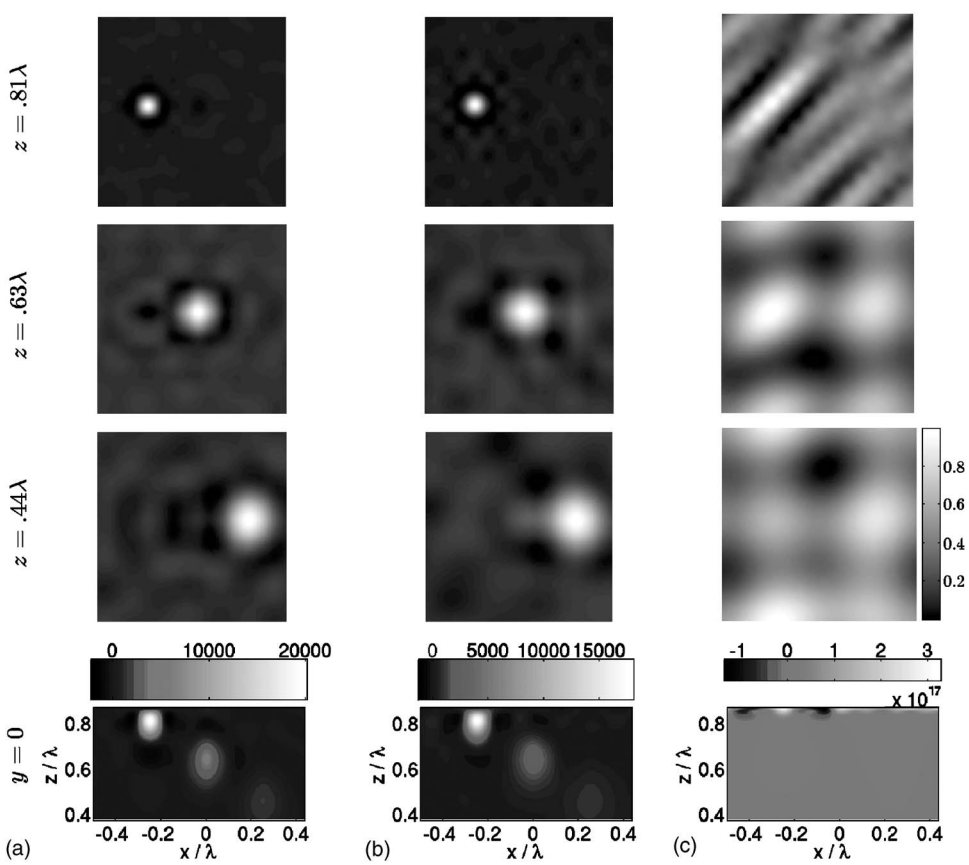


FIG. 5. Comparison of reconstructions from weak tip and strong tip models. The effective tip polarizability is $\lambda^3/100$; column (a) shows the weak tip reconstruction from weak tip simulated data, column (b) is the strong tip reconstruction with strong tip simulated data, and column (c) is the reconstruction with the weak tip model using strong tip simulated data. Images in the first three rows are linearly normalized.

weak tip model, especially when the reconstruction layer is further away from the measurement plane. This is consistent with our conclusion that the strong tip ISP is more poorly conditioned than the weak tip ISP. In column (c) data were simulated including the strong tip term, but reconstructions were computed only under the weak tip assumption. It is seen that the reconstruction barely resembles the original sample. These simulations demonstrate the significance of strong tip effect in the NSOT image reconstruction. Moreover, the results are consistent with the experiences of the authors in developing the inversion methods of Ref. 20 where a weak-tip model initially yielded poor results like those seen in column (c) while the inclusion of a strong tip term yielded results in good agreement with known sample structure.

V. DISCUSSION

Multiple-scattering effects from the tip in both illumination mode and collection mode NSOT have been analyzed. For the inverse problem, inclusion of the strong tip correction is seen, through numerical simulations, to be essential for a valid reconstruction of the sample.

It is shown qualitatively that the strong tip effect increases the ill-posedness of the inverse problem for these modalities. Under the extreme circumstance that the tip-scattering effect becomes dominant, no meaningful reconstruction can be obtained and therefore alternative approaches to NSOT experiment and data acquisition might be advised.

ACKNOWLEDGMENTS

This work was supported in part by U.S. Air Force MURI Grant F49620-03-10379 and in part by the National Science Foundation under CAREER Award Grant 0239265 and a University of Illinois Urbana-Champaign Critical Initiatives in Research and Scholarship grant. J.C.S. acknowledges support under Grant Nos. NSF DMR 0425780 and USAFOSR FA9550-07-1- 0096.

APPENDIX

The polarization tensors $\mathcal{D}(\mathbf{q})$ and $\tilde{\mathcal{D}}(\mathbf{q})$ in Eq. (6) are given by the expressions

$$\mathcal{D}_{\alpha\beta}(\mathbf{q}) = \delta_{\alpha\beta} - k_0^{-2} k_\alpha(\mathbf{q}) k_\beta(\mathbf{q}), \tag{A1}$$

$$\tilde{\mathcal{D}}(\mathbf{q})_{\alpha\beta} = \delta_{\alpha\beta} - k_0^{-2} \tilde{k}_\alpha(\mathbf{q}) \tilde{k}_\beta(\mathbf{q}), \tag{A2}$$

where $\tilde{\mathbf{k}}(\mathbf{q}) \equiv [\mathbf{q}, -k_z(\mathbf{q})]$ is the down-going wave vector in the $z \geq 0$ half-space. The reflection and transmission tensors for the vector field in the $z \geq 0$ half-space and the $z < 0$ half-space 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 $\mathcal{R}$ and $\mathcal{R}'$ reflection in the $z \geq 0$ half-space and the $z < 0$ half-space, respectively, and we denote by $\mathcal{T}$ and $\mathcal{T}'$ transmission from the $z \geq 0$ half-space into the $z < 0$ half-space and transmission from the $z < 0$ half-space into the $z \geq 0$ half-space respectively. $\mathbf{P}$ and $\mathbf{P}'$ are projection operators onto the TE/TM basis in the $z \geq 0$ half-space and the $z < 0$ half-space, respectively, and are discussed below. We find for the reflection and transmission tensors the expressions

$$\mathcal{R}_{\alpha\beta}(\mathbf{q}) = P_{\gamma\alpha}(\mathbf{k}) r_{\gamma\delta}(\mathbf{k},\mathbf{k}') P_{\delta\beta}(\tilde{\mathbf{k}}), \tag{A3}$$

$$\mathcal{R}'_{\alpha\beta}(\mathbf{q}) = P'_{\gamma\alpha}(\tilde{\mathbf{k}}') r'_{\gamma\delta}(\mathbf{k},\mathbf{k}') P'_{\delta\beta}(\mathbf{k}'), \tag{A4}$$

$$\mathcal{T}_{\alpha\beta}(\mathbf{q}) = P'_{\gamma\alpha}(\tilde{\mathbf{k}}') t_{\gamma\delta}(\mathbf{k},\mathbf{k}') P_{\delta\beta}(\tilde{\mathbf{k}}), \tag{A5}$$

$$\mathcal{T}'_{\alpha\beta}(\mathbf{q}) = P_{\gamma\alpha}(\mathbf{k}) t'_{\gamma\delta}(\mathbf{k},\mathbf{k}') P'_{\delta\beta}(\mathbf{k}'). \tag{A6}$$

where the TEM reflection and transmission coefficients r and t in the $z \geq 0$ half-space are given by

$$r = \begin{pmatrix} \dfrac{n^2 k_z - k_z'}{n^2 k_z + k_z'} & 0 \\ 0 & \dfrac{k_z - k_z'}{k_z + k_z'} \end{pmatrix}, \tag{A7}$$

and

$$t = \begin{pmatrix} \dfrac{2nk_z}{n^2 k_z + k_z'} & 0 \\ 0 & \dfrac{2k_z}{k_z + k_z'} \end{pmatrix}, \tag{A8}$$

and their $z < 0$ half-space counterparts are given by $r' = -r$ and $t' = k_z' t / k_z$. The projection operator onto the TE/TM basis in the $z \geq 0$ half-space is given by

$$\mathbf{P}(\mathbf{k}) = \frac{1}{\sqrt{k_x^2 + k_y^2} k_0} \begin{pmatrix} -k_x k_z & -k_y k_z & k_x^2 + k_y^2 \\ -k_y k_0 & k_x k_0 & 0 \end{pmatrix}, \tag{A9}$$

and in the $z < 0$ half-space

$$\mathbf{P}'(\mathbf{k}') = \frac{1}{\sqrt{k_x^2 + k_y^2} n k_0} \begin{pmatrix} -k_x k_z' & -k_y k_z' & k_x^2 + k_y^2 \\ -k_y n k_0 & k_x n k_0 & 0 \end{pmatrix}. \tag{A10}$$

The following identities are useful in derivations in the text:

$$P_{\gamma\alpha}(\mathbf{k}) P_{\gamma\beta}(\mathbf{k}) = \mathcal{D}_{\alpha\beta}(\mathbf{q}), \tag{A11}$$

$$P_{\gamma\alpha}(\tilde{\mathbf{k}}) P_{\gamma\beta}(\tilde{\mathbf{k}}) = \tilde{\mathcal{D}}_{\alpha\beta}(\mathbf{q}), \tag{A12}$$

$$P_{\alpha\gamma}(\mathbf{k}) P_{\beta\gamma}(\mathbf{k}) = P_{\alpha\gamma}(\tilde{\mathbf{k}}) P_{\beta\gamma}(\tilde{\mathbf{k}}) = \delta_{\alpha\beta}, \tag{A13}$$

$$\mathcal{R}(\mathbf{q})_{\alpha\beta} = \mathcal{R}(-\mathbf{q})_{\beta\alpha}, \tag{A14}$$

$$\mathcal{R}'(\mathbf{q})_{\alpha\beta} = \mathcal{R}'(-\mathbf{q})_{\beta\alpha}, \tag{A15}$$

$$\mathcal{T}'_{\gamma\alpha}(\mathbf{q}) \mathcal{R}_{\gamma\beta}(\mathbf{q}) = \frac{-k_z'}{k_z} \mathcal{R}'_{\gamma\alpha}(\mathbf{q}) \mathcal{T}_{\gamma\beta}(\mathbf{q}), \tag{A16}$$

$$\mathcal{R}_{\mu\alpha}(\mathbf{q}) \mathcal{R}_{\mu\beta}(\mathbf{q}) + \frac{k_z'}{k_z} \mathcal{T}_{\mu\alpha}(\mathbf{q}) \mathcal{T}_{\mu\beta}(\mathbf{q}) = \delta_{\alpha\beta} - \frac{\tilde{k}_\alpha \tilde{k}_\beta}{k_0^2}. \tag{A17}$$

Equations (A11)–(A13) simply reflect the fact the $\mathbf{P}$ is a projection operator with the usual properties that $\mathbf{P}^2 = \mathbf{P}$ and

P is the identity on the subspace into which it projects. Equation (A16) may be understood to be a statement of Stokes reciprocity.

[1]C. Girard and A. Dereux, Rep. Prog. Phys. **59**, 657 (1996).
[2]D. Courjon, K. Sarayeddine, and M. Spajer, Opt. Commun. **71**, 23 (1989).
[3]E. Synge, Philos. Mag. **6**, 356 (1928).
[4]R. C. Reddick, R. J. Warmack, and T. L. Ferrell, Phys. Rev. B **39**, 767 (1989).
[5]E. Ash and G. Nicholls, Nature (London) **237**, 510 (1972).
[6]A. Lewis, M. Isaacson, A. Harootunian, and A. Muray, Ultramicroscopy **13**, 227 (1984).
[7]D. W. Pohl, W. Denk, and M. Lanz, Appl. Phys. Lett. **44**, 651 (1984).
[8]E. Betzig and J. K. Trautman, Science **257**, 189 (1992).
[9]N. Garcia and M. Nieto-Vesperinas, Opt. Lett. **18**, 2090 (1993).
[10]N. Garcia and M. Nieto-Vesperinas, Opt. Lett. **20**, 949 (1995).
[11]R. Carminati, J.-J. Greffet, N. Garcia, and M. Nieto-Vesperinas, Opt. Lett. **21**, 501 (1996).
[12]R. Carminati and J.-J. Greffet, Opt. Commun. **116**, 316 (1995).
[13]P. S. Carney and J. C. Schotland, Appl. Phys. Lett. **77**, 2798 (2000).
[14]J. Sun, J. C. Schotland, and P. S. Carney, IEEE J. Sel. Top. Quantum Electron. **12**, 1072 (2006).
[15]K. Belkebir, P. Chaumet, and A. Sentenac, J. Opt. Soc. Am. A **22**, 1889 (2005).
[16]G. Bao and L. Peijun, SIAM J. Appl. Math. **65**, 2049 (2005).
[17]G. Panasyuk, V. Markel, P. Carney, and J. Schotland, Appl. Phys. Lett. **89**, 221116 (2006).
[18]P. C. Chaumet, K. Belkebir, and A. Sentenac, Opt. Lett. **29**, 2740 (2004).
[19]S. I. Bozhevolnyi, V. A. Markel, V. Coello, W. Kim, and V. M. Shalaev, Phys. Rev. B **58**, 11441 (1998).
[20]P. S. Carney, R. A. Frazin, S. I. Bozhevolnyi, V. S. Volkov, A. Boltasseva, and J. C. Schotland, Phys. Rev. Lett. **92**, 163903 (2004).
[21]C. Tai, *Dyadic Green Functions in Electromagnetic Theory*, 2nd ed. (IEEE Press, New York, 1994).
[22]L. Novotny and B. Hecht, *Principles of Nano-Optics*, 1st ed. (Cambridge University Press, Cambridge, 2006).
[23]A. Cvitkovic, N. Ocelic, and R. Hillenbrand, Opt. Express **15**, 8550 (2007).
[24]P. de Vries, D. V. van Coevorden, and A. Lagendijk, Rev. Mod. Phys. **70**, 447 (1998).
[25]P. S. Carney and J. C. Schotland, J. Opt. Pure Appl. Opt. **4**, S140 (2002).
[26]P. S. Carney and J. C. Schotland, J. Opt. Soc. Am. A **20**, 542 (2003).