SAND2018-10834J

Multiresolution Localization with Temporal Scanning for Super-Resolution Diffuse Optical Imaging of Fluorescence

Brian Z. Bentz, *Member, IEEE,* Dergan Lin, *Member, IEEE,* Justin A. Patel, *Member, IEEE,* and Kevin J. Webb, *Fellow, IEEE*

***Abstract*—A super-resolution *in vivo* optical imaging method is presented that relies on the distinct temporal information associated with the optical reporter. Optical imaging for *in vivo* applications can provide information not available with other types of imaging, but the multiply scattered light measured precludes direct observation. Notably, the imaging of fluorescence could enable the first direct imaging of neuron activity throughout the mammalian brain. However, to date, computational imaging methods based on a forward model such as the diffusion equation suffer from low spatial resolution to a degree that severely limits the application domain and prevents neuron-scale resolution. We propose a localization-based method that allows for super-resolution diffusive optical imaging in highly scattering media, such as tissue. The method relies on some degree of separation in time of the temporal response of multiple fluorescent sources. By localizing each emitter individually, a spatial resolution on the order of 10 microns through 1 cm of tissue or more is possible. This means that, in principle, the approach could be used to image all neurons in a rodent brain.**

***Index Terms*—Super-resolution, Optical imaging, turbid media, localization, matching pursuit.**

I. Introduction

Fluorescence imaging has become a standard tool in biomedical research because modulation of fluorescence intensity in space and time can provide information on biochemical processes [1]–[3]. Methods such as confocal microscopy [4] and multi-photon imaging [5] have enabled high-resolution fluorescence imaging near the surface of tissues. Fluorescence imaging in deep tissue, where the propagation direction of light becomes randomized, presents a major challenge in optical imaging. Information is lost due to scatter and absorption, which hinders image formation. In deeper tissue, cost-function-based imaging methods with a forward model can be used to for images of fluorescence. The use of a simple diffusion model to describe the photon transport has been shown to capture the mean intensity of light under suitable conditions that can be found in tissue with visible and near-infrared light. This has led to substantial interest in diffuse optical tomography (DOT, or optical diffusion tomography, ODT) [6], [7], and fluorescence diffuse optical tomography (FDOT or FODT) [8], [9]. Imaging of fluorescence resonance energy transfer (FRET) parameters, providing a spatial map of the nanometer FRET distance that is important in molecular biology, has also been shown [10], [11]. However, despite such progress, diffusive optical imaging methods suffer from low spatial resolution due to the strong attenuation of high-frequency information as light propagates through scattering media. This attenuation occurs whether the media is lossy (with absorption) or lossless (without absorption). We present a method that relies on treating fluorescent emitters whose signals vary in time as a set of point sources, and uses measurements of multiply-scattered light as a function of time to find their locations to high precision. Consequently, we present the opportunity to directly image neurons throughout the brain of a small mammal, which should provide new understanding for neuroscience. The work should also allow for imaging in other dynamic source situations in a cluttered environment.

Of particular significance is the prospect of imaging calcium (Ca^{2+}) channel activity in neurons throughout the brain. Neuron activity involves Ca^{2+} transport through channels in the cell body and then, after signal transfer along the axon, regulation of the neurotransmitter across the synaptic cleft in communication with the other neuron. Monitoring Ca^{2+} signaling thus presents a direct picture of neuron activity [12]–[14]. If it were possible to image Ca^{2+} channel activity across the brain, the resulting maps would provide new understanding about how the brain works, and also answer fundamental questions about health and disease. Fortunately, Ca^{2+} fluorescence and FRET sensors have been developed for *in vitro* studies [13], [15]. The principle is that Ca^{2+} modulates the fluorescence emission or the energy transfer from the donor to the acceptor in the case of FRET, thereby encoding concentration information in the fluorescence that can be detected. The temporal modulation of the fluorescence therefore reports on calcium channel activity. However, with optical excitation using a laser, both the excitation light and the emitted fluorescence are multiply scattered. Computational imaging of fluorescence and FRET is possible with a forward model, such as the diffusion equation, using a suitable cost function [8], [10], [11], but the spatial resolution achieved precludes being able to resolve individual neurons in deep tissue. By representing the local Ca^{2+} fluorescent reporter as a point emitter or point absorber (the donor emission can reduce due to FRET), we present a framework suitable for localization using a forward model. We show that this provides a path to

B. Z. Bentz is with Sandia National Laboratories, Albuquerque, NM, 87123 USA

D. Lin, J. A. Patel, and K. J. Webb are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, 47907 USA e-mail: webb@purdue.edu.

Manuscript received April XX, 2018; revised April XX, 2018.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

high spatial resolution, where each reporter can be imaged to high resolution, under the assumption that there is distinct temporal modulation (due to the neuronal activity).

None of the available *in vivo* methods provide direct access to neurons, i.e., they measure secondary parameters. For example, Brain imaging with fMRI has been developed using the blood-oxygen-level dependent (BOLD) contrast [16], and this has become an important tool for brain science. However, blood oxygen levels provide low resolution and a slow and remote measure of brain signaling. Optical methods such as two-photon microscopy [17] offer more direct access to a wider range of neurobiological information through optical contrast agents, but only provide information about the cortical surface of the brain visible through the use of cranial windows [18]. Our approach uses an optimization framework with a forward model for the scattering medium.

Diffuse optical imaging methods have been developed to overcome the detrimental effects of scatter, enabling deep tissue fluorescence imaging [1]. In FDOT [8], [9], [19]–[23], computational imaging allows the formation of volumetric images of the absorption, scatter, and fluorescence from boundary measurements. Fundamentally, these methods detect what have become known as diffuse photon density waves (DPDW's) emanating from excitation [24] or fluorescent sources [25]. For typical measurements and tissues, the wavelength of a DPDW is a few centimeters, and measurements are made in the near field. Considering the angular spectrum of DPDW's [26], all the spatial frequencies propagate to the detector, however they are highly attenuated, even without absorption. This causes a severe reduction in spatial resolution with depth into the medium. The dependence of the spatial resolution on depth is nonlinear, but for typical tissues, measurement geometries, and beyond a depth of about 1 cm, spatial resolutions of about depth$/2$ have been achieved [26]–[32]. We present a temporal localization method that circumvents earlier resolution limits. However, the object is modeled as a set of points, so other geometrical information is lost.

We develop a model describing the measured fluorescence intensity due to multiple inhomogeneities within a highly scattering media in Section II. We then use the model to localize fluorescent inhomogeneities separated in space and distributed in time in Section III, and the results are presented in Section IV. We compare our method to existing super-resolution imaging methods, and discuss potential applications in Section V. We conclude in Section VI. Appendix A shows how the diffusion model parameters can be scaled for other physical situations. We use a statistical detection scheme, which is described in Appendix B, and we show that this method allows super-resolution diffusive optical imaging.

II. Models

A. Physical Problem

The primary application we address is the direct *in vivo* imaging of neuron activity through optical fluorescence reporting of Ca^{2+} concentration. An optical experiment to measure Ca^{2+} signaling involves using a pulsed laser source to excite the fluorescent reporter and a time-gated camera to measure the fluorescence. With a series of deconvolved measurements, we assume that useful statistical information about neuron activity becomes available from correlations over measurements. The question we address here is how one could image each neuron using such data. Our approach is to consider the switching of the Ca^{2+} channels captured, and the fact that there will be slightly different onset times. Therefore, in principle, by treating the fluorescent Ca^{2+} reporter in each neuron as a point emitter in space, having temporal characteristics attributable to the neuron response and the fluorescence lifetime, it becomes possible to use a mathematical framework where each neuron can be modeled in space as a single emitter. This is called localization. A method for localizing a larger set of point emitters based on distinct temporal information is presented here, primarily with a view to the brain imaging application.

B. Coupled Diffusion Equations

We use the diffusion approximation to the radiative transfer equation to describe the propagation of light in a highly scattering medium such as tissue [8], [19], [33]. This is an incoherent picture that has proven useful in describing the mean optical intensity when the appropriate scattering conditions are met, which is the case for tissue having millimeter thickness or more and red or near-infrared light. The coupled diffusion model in the time domain is given by [25], [34]

$$\frac{1}{v}\frac{\partial}{\partial t}\phi_x(\mathbf{r},t) - \nabla\cdot[D_x(\mathbf{r})\nabla\phi_x(\mathbf{r},t)] + \mu_{a_x}(\mathbf{r})\phi_x(\mathbf{r},t) = S_x(\mathbf{r};t) \tag{1}$$

$$\frac{1}{v}\frac{\partial}{\partial t}\phi_m(\mathbf{r},t) - \nabla\cdot[D_m(\mathbf{r})\nabla\phi_m(\mathbf{r},t)] + \mu_{a_m}(\mathbf{r})\phi_m(\mathbf{r},t) = \phi_x(\mathbf{r},t) * S_f(\mathbf{r};t), \tag{2}$$

where $\mathbf{r}$ denotes the position, ϕ (W/mm^2) is the photon flux density, μ_a (mm^{-1}) is the absorption coefficient, $D = 1/[3(\mu_a+\mu_s')]$ (mm) is the diffusion coefficient, $\mu_s' = \mu_s(1-g)$ (mm^{-1}) is the reduced scattering coefficient, μ_s is the scattering coefficient (mm^{-1}), g is the anisotropy parameter, $v = c/n$ is the speed of light in the medium, where c is the speed of light in free space and n is the refractive index, the subscripts x and m, respectively, denote parameters at the excitation and emission wavelengths, λ_x and λ_m, S_x (W/mm^3) is the excitation source term, S_f (1/s/mm) is the fluorescence source term, and $*$ signifies a temporal convolution. With a temporal Fourier transform, resulting in the frequency domain form of (1) and (2), and considering homogeneous D and μ_a, the resulting scalar wave equations for ϕ describe the propagation of DPDW's.

Throughout this paper, we present results in terms of the mean free path, $l^* = 3D$. However, the results can be scaled to geometries of different size and amount of scatter, according to the scaling principles of Appendix A. This scaling argument allows different physical problems to be considered, and can be applied to situations as diverse as atmospheric scatter and snow.

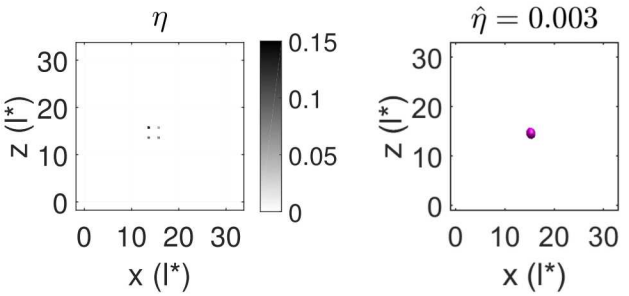


Fig. 1. 3D FDOT reconstruction of four point fluorescent inhomogeneities with 30 dB SNR. The optical properties are similar to tissue, where $\mu_s' = 2$ mm^{-1}, $\mu_a = 0.02$ mm^{-1}, and $n = 1.33$, giving a mean free path length $l^* = 3D = 0.5$ mm. The geometry is a cube of side length $32l^*$ with the point fluorescent inhomogeneities near the center. 6 sources and detectors are placed at the boundary of the cube, one source and one detector at each face. (a) Target η xz image slice, where the fluorescent inhomogeneities are located at depths of about $17l^*$, and are separated by about $2l^*$. (b) Estimated $\hat{\eta}$ reconstructed with FDOT. An isoimage with $\hat{\eta} = 0.003$ mm^{-1} threshold was formed and the same xz slice as in (a) is shown. In this case, the separation of the fluorescent inhomogeneities is much less than depth/2 and they cannot be resolved.

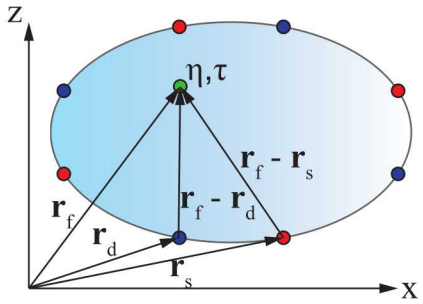


Fig. 2. Model geometry with position vector $\mathbf{r} = (x, z)$. Excitation sources at λ_x (red) are placed at known positions $\mathbf{r}_s$, point fluorescence emission locations are assumed to be $\mathbf{r}_f$ (green), and detectors at λ_m are placed at known positions $\mathbf{r}_d$ (blue).

We demonstrate the spatial resolution with an example FDOT reconstruction in Fig. 1, where we numerically reconstruct the fluorescence yield (η, as described in Sect. II-C) in 3-D space using (1) and (2), as we have described previously [8]. We assume optical properties similar to tissue at wavelengths of red and near-infrared light [35] ($\mu_s' = 2$ mm^{-1}, $\mu_a = 0.02$ mm^{-1}), and that measurements are made with a signal-to-noise ratio (SNR) of 30 dB at the detector, with Gaussian noise that is based on a shot-noise model (as described in Section II-E). The 3D geometry is a cube of side length $32l^*$, where sources and detectors were placed on each side of the cubic geometry, and the point fluorescent inhomogeneities are near the center. There are A sources and B detectors arranged (in some fashion) on each of the six faces of the cubic scattering medium. Figure 1(a) shows a 2-D slice of the target image, η, and Fig. 1(b) shows the corresponding 2-D slice isoimage of the estimated or reconstructed image, $\hat{\eta}$. It is clear from Fig. 1(b) that the point inhomogeneities in Fig. 1(a) are not resolvable with FDOT, under the conditions considered.

C. Forward Model for a Single Fluorescent Inhomogeneity

Equations (1) and (2) can be used to form a model of the detected power at λ_x and λ_m, respectively, at a set of locations around the periphery of the scattering medium. We consider the case of a set of point-emitting fluorophores, each with a differing temporal response due to the physical situation, so that the response can be attributed to a sequence of single point fluorophores. We start by describing the response due to a single fluorescent inhomogeneity, such as would be the case for a Ca^{2+} reporter, embedded in a highly scattering medium.

The problem geometry is shown in Fig. 2. The incident laser light at wavelength λ_x can be modeled as an (isotropically emitting) point source for S_x in (1) located l^* into the scattering medium (the brain). Point excitation sources represent a set of laser beam excitation locations at known positions $\mathbf{r}_s$. The excitation source term is then given by $S_x(\mathbf{r}; t) = S_o\delta(\mathbf{r} - \mathbf{r}_s, t)$, where S_o is the excitation power (W/mm^3). In this work, we let $S_o = 1$ for simplicity. Additionally, point detectors that collect light at λ_m are placed at known positions $\mathbf{r}_d$. Finally, a fluorescent inhomogeneity is assumed to be located at the unknown position $\mathbf{r}'$. Using the problem geometry in Fig. 2, we define $g_x(\mathbf{r}_s, \mathbf{r}', t)$ as the Green's function from (1) at λ_x and $g_m(\mathbf{r}', \mathbf{r}_d, t)$ as the Green's function from (2) at λ_m. The fluorescence emission photon flux density at the detector is then

$$\phi_m(\mathbf{r}_s, \mathbf{r}_d, t) = \int g_m(\mathbf{r}', \mathbf{r}_d, t) * S_f(\mathbf{r}'; t) * g_x(\mathbf{r}_s, \mathbf{r}', t) d\mathbf{r}', \tag{3}$$

where $*$ is the temporal convolution.

We assume a point fluorophore is located at $\mathbf{r}_f$, as seen in Fig. 2, giving

$$S_f(\mathbf{r}; t) = \frac{\eta_f \mu_{a_f}}{\tau_f} \exp\left(\frac{-t}{\tau_f}\right) \delta(\mathbf{r} - \mathbf{r}_f), \tag{4}$$

where η_f is the fluorophore quantum yield, μ_{a_f} is the fluorophore absorption coefficient at λ_x, τ_f is the fluorescence lifetime, and $\delta(\cdot)$ is the Dirac delta. For brevity, we define $\eta = \eta_f \mu_{a_f}$ as the fluorescence yield. For simplicity, we assume that D and μ_a are homogeneous, and that v is the same at λ_x and λ_m. Using the Green's function solution to the diffusion equation [34], we find from (3) that

$$\begin{aligned}\phi_m(\mathbf{r}_s, \mathbf{r}_f, \mathbf{r}_d, t) = & \\ & \frac{v}{(4\pi D_m v t)^{3/2}} \exp\left(\frac{-|\mathbf{r}_d - \mathbf{r}_f|^2}{4D_m v t} - \mu_{a_m} v t\right) \\ & * \frac{\eta}{\tau_f} \exp\left(\frac{-t}{\tau_f}\right) \\ & * \frac{v}{(4\pi D_x v t)^{3/2}} \exp\left(\frac{-|\mathbf{r}_f - \mathbf{r}_s|^2}{4D_x v t} - \mu_{a_x} v t\right).\end{aligned} \tag{5}$$

We also assume that $\mu_{a_x} = \mu_{a_m} = \mu_a$ and $D_x = D_m = D$. The detected power through an aperture is described by the current density $\mathbf{J}$ (W/mm^2), with a diffusion framework, and we assume pointwise measurements

$$\mathbf{J}(\mathbf{r}_s, \mathbf{r}_f, \mathbf{r}_d, t) = -D\nabla\phi_m(\mathbf{r}_s, \mathbf{r}_f, \mathbf{r}_d, t). \tag{6}$$

For laser excitation source locations $\mathbf{r}_{s_q}$ with $q \in [1, ..., Q]$ and fluorescence detector locations $\mathbf{r}_{d_m}$ with $m \in [1, ..., M]$, we write the detected fluorescent photon current density in compact form as

$$\mathbf{J}_{qm}(\mathbf{r}_f, t) = \mathbf{J}(\mathbf{r}_{s_q}, \mathbf{r}_f, \mathbf{r}_{d_m}, t), \tag{7}$$

where we emphasize the dependence on $\mathbf{r}_f$ because it will be estimated in Section III. The pointwise detected fluorescence is then

$$G_{qm}(\mathbf{r}_f, t) = |\mathbf{J}_{qm}(\mathbf{r}_f, t) \cdot \hat{\mathbf{n}}|, \tag{8}$$

where $\hat{\mathbf{n}}$ is the unit vector normal to the detector surface and G signifies the (diffusion equation) Green's function basis.

For simplicity we use the analytic solution (8) in this work, but the model data could also be generated using finite element or other numerical methods [7]. These analytical solutions to the diffusion equation subject to boundary conditions have been shown to match Monte Carlo simulations [36] and experiments of photon propagation in highly scattering media [37], [38]. It has been shown that small fluorescent inhomogeneities can be well approximated as point emitters because of the rapid attenuation of high spatial frequencies in scattering media [38], [39]. For this reason, imaging based on a point representation is possible.

D. Forward Model for Multiple Fluorescent Inhomogeneities

In the case where multiple fluorescent inhomogeneities are present, the forward model must be modified. We assume each fluorescent inhomogeneity can be represented by a point that fluoresces with a unique and increasing delay τ from $t = 0$. We let K fluorescent inhomogeneities be located at positions $\mathbf{r}_{f_k}$, where k is an index from 1 to K. Each fluorescent inhomogeneity has a different yield, η_k, and fires with different delay, τ_k. Then, the detected fluorescence for a single source-detector pair measured over the temporal support w starting at time t_o is

$$\begin{aligned}\tilde{G}_{qm}(\mathbf{R}, \boldsymbol{\tau}, t) = \text{rect}\left(\frac{t - 0.5w - t_o}{w}\right) \\ \times \sum_{k=1}^{K} \delta(t - \tau_k) * G_{qm}(\mathbf{r}_{f_k}, t),\end{aligned} \tag{9}$$

where for general x, rect(x) is 1 when $|x| < 0.5$ and zero otherwise, the vector $\boldsymbol{\tau} = [\tau_1 ..., \tau_K]^T$ corresponds to the delays τ_k, and the vector $\mathbf{R}$ contains the K positions $\mathbf{r}_{f_k}$.

We assume the detected fluorescence from (9) is sampled with sampling interval T, and we discretize the domain such that V_{vox} is the volume of a single voxel, giving the forward measurements

$$f_{qmn}(\mathbf{R}, \boldsymbol{\tau}) = V_{vox}\left[\tilde{G}_{qm}(\mathbf{R}, \boldsymbol{\tau}, t)|_{t=nT}\right] \tag{10}$$

where n is an index from 1 to N, and $t_{max} = NT$. We emphasize the dependence on $\mathbf{R}$ and $\boldsymbol{\tau}$ because these parameters will be estimated in Section III. We can now write the fluorescence data vector expected from the diffusion model as $\mathbf{f}(\mathbf{R}, \boldsymbol{\tau})$, which is

$$\begin{aligned}\mathbf{f}(\mathbf{R}, \boldsymbol{\tau}) = [f_{111}, ..., f_{11N}, f_{121}, ..., f_{12N}, ..., f_{1M1}, \\ ..., f_{1MN}, f_{211}, ..., f_{QMN}]^T.\end{aligned} \tag{11}$$

Considering (5), each f_{qmn} is linear with respect to η_k, allowing us to write

$$\mathbf{f}(\mathbf{R}, \boldsymbol{\tau}) = \mathbf{F}(\mathbf{R}, \boldsymbol{\tau})\boldsymbol{\eta}, \tag{12}$$

where $\boldsymbol{\eta} = [\eta_1 ..., \eta_K]^T$ is a vector containing the fluorescent yields (η_k) and $\mathbf{F}(\mathbf{R}, \boldsymbol{\tau})$ is a matrix of dimensions $[QMN, K]$ that contains the scaled forward measurements. The matrix $\mathbf{F}(\mathbf{R}, \boldsymbol{\tau})$ can be calculated from (9) by setting all η_k equal to 1. The vector multiplication in (12) is equivalent to the superposition of the K fluorescent responses, which, after scaling, make up the columns of $\mathbf{F}(\mathbf{R}, \boldsymbol{\tau})$.

E. Detector Noise

We use a Gaussian noise model, for which the statistics are based on a shot-noise-limited measurement system [40], [41]. We let $\mathbf{y}$ be the noisy measurement vector corresponding to $\mathbf{f}$, such that both have dimensions of $[QMN, 1]$. We assume that the noise is independent, zero mean, and Gaussian with covariance matrix $\boldsymbol{\Upsilon}$, where

$$[\Upsilon]_{ii} = \alpha|\mathbf{y}_i|, \tag{13}$$

$i = [1, ..., QMN]$ is the data index, and α is a scalar parameter that is dependent on the noise process and detector physics [40]. The form of (13) suggests a bias towards low intensity data in the cost minimization. The SNR in dB for a single source-detector pair depends on α according to

$$\text{SNR} = 10 \log\left[\frac{1}{\alpha} f_{qmn}(\mathbf{R}, \boldsymbol{\tau})\right]. \tag{14}$$

We generate simulated noisy measurements $\mathbf{y}$ for a specified SNR by calculating α from (14).

III. Localization for Super-Resolution Imaging

Compared to diffusive imaging with FDOT, fluorescence localization is a simpler problem [38], [42]–[45]. Biomedical applications of localization have included determining the location of a fluorescing tumor in a mouse [46] or the location of a targeted fluorophore embedded in the tongue of a live mouse [47]. Here, we describe the localization of multiple fluorescent inhomogeneities for the formation of super-resolution images. Our method assumes there is a significant temporal separation τ_f such that the fluorescent impulse responses can be separated, and we explore this assumption in Section III-C. By exploiting this information as prior knowledge, higher-resolution imaging compared to FDOT is achieved.

A. Localization of Multiple Fluorescent Inhomogeneities

In order to localize K fluorescent inhomogeneities inside a highly scattering medium, the positions $\mathbf{r}_{f_k}$ must be estimated. To further characterize the fluorescent inhomogeneities, the yields η_k, which are proportional to the fluorophore concentration, can also be estimated. We form the maximum likelihood (ML) estimation as [38], [39]

$$\hat{\theta} = \arg\max_{\theta} p_{1,\theta}(\mathbf{y}), \tag{15}$$

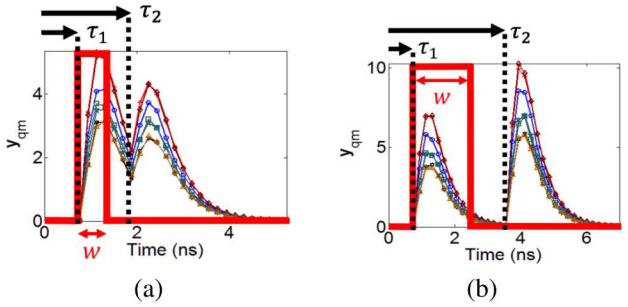


Fig. 3. Typical fluorescence temporal responses for one source and seven detectors ($Q = 1$, $M = 7$) using the same optical properties as Fig. 1. The 7 different symbols and corresponding colors represent different source-detector measurement pairs. The time axis is a discrete set of points $t_1, ..., t_N$, with T between sample points. (a) The delay τ_2 is short, causing substantial overlap due to superposition. (b) The delay τ_2 is long such that the detected fluorescence decays substantially before the next fluorescence response. We show that localization of the fluorescence inhomogeneities is possible in both cases.

with $p_{1,\theta}(\mathbf{y})$ is the probability distribution given by

$$p_{1,\theta}(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^P |\boldsymbol{\Upsilon}|}} \exp\left(-\frac{1}{2}||\mathbf{y} - \mathbf{F}(\mathbf{R},\boldsymbol{\tau})\boldsymbol{\eta}||^2_{\boldsymbol{\Upsilon}^{-1}}\right), \tag{16}$$

where $P = QMN$ is the number of measurements, for an arbitrary vector $\mathbf{u}$, $||\mathbf{u}||^2_{\mathbf{V}} = \mathbf{u}^H\mathbf{V}\mathbf{u}$, H being the Hermitian transpose, and θ corresponds to the parameters of interest for localization, which are $\mathbf{R}$ and $\boldsymbol{\eta}$. After taking the logarithm, (15) is equivalent to minimizing

$$c(\mathbf{R},\boldsymbol{\tau}) = \min_{\boldsymbol{\eta}} ||\mathbf{y} - \mathbf{F}(\mathbf{R},\boldsymbol{\tau})\boldsymbol{\eta}||^2_{\boldsymbol{\Upsilon}^{-1}}. \tag{17}$$

Minimizing (17) could be accomplished using matching pursuit [48], [49], where a dictionary of basis functions is formed using (9) (with $K = 1$) for all possible $\mathbf{r}_{f_k}$ and τ_k. The weight of each basis function is then estimated by finding the maximum inner products between the measured data $\mathbf{y}$ and the dictionary basis functions. For (17), the basis function weight would be η_k. However, if high temporal and spatial resolution is desired, as is the case in this work, forming the dictionary and calculating the inner products becomes computationally difficult, which we show in Section IV-B.

We propose a method to minimize (17) that takes advantage of the causality of the problem. The approach allows us to avoid calculation of the basis functions at each τ_k and their inner products, alleviating the computational burden. We assume that each τ_k is unique, allowing temporal separation of the fluorescent impulse responses, as shown in Fig. 3 for the case of two emitters. In Fig. 3(a) the temporal responses from two point fluorophore locations in a scattering domain are measured to be displaced in time but overlapping, whereas in Fig. 3(b) they are distinct. We consider the general case where the delays τ_k are not large enough for the fluorescence to decay to noise before the start of the next response, as shown in Fig. 3(a). The discretized measurement data leads to a temporal sampling period T and we consider a temporal window w, illustrated in Fig. 3, consisting of an integer number of samples separated in time by T.

Considering (9), we see that η_k and $\mathbf{r}_{f_k}$ could be estimated using data within a temporal window starting at $t_o = \tau_k$ with $w < \tau_{k+1} - \tau_k$. Thus, we assume that the data within the temporal window w comes from a single fluorophore at $\mathbf{r}_{f_k}$. A simplified cost function based on (17) can then be written as

$$c_k(\mathbf{r}_{f_k}, \tau_k) = \min_{\eta_k} ||\mathbf{y}_k - \eta_k \mathbf{f}_k(\mathbf{r}_{f_k})||^2_{\boldsymbol{\Upsilon}_k^{-1}}, \tag{18}$$

where $\mathbf{f}_k(\mathbf{r}_{f_k})$ is derived from $\mathbf{f}(\mathbf{R},\boldsymbol{\tau})$ in (11) with $t_o = \tau_k$ and $w < \tau_{k+1} - \tau_k$, $\mathbf{y}_k$ is the corresponding windowed measurement vector, and $\boldsymbol{\Upsilon}_k$ is the noise covariance matrix for $\mathbf{y}_k$. In this description, $\mathbf{y}_k$ contains the subset of measurements from $\mathbf{y}$ at all detectors that are within w. Note that w is being used here as a model parameter to identify the response of a single point fluorophore, and that the measured data $\mathbf{y}$ has a temporal support that is much larger than w.

Forming (18) requires that τ_k is known. We estimate τ_k using the generalized likelihood ratio test (GLRT) [38], which is described in Appendix B. The GLRT uses a binary hypothesis to calculate a threshold for determining whether a point fluorophore is detectable or "localizable." We apply the GLRT at each time index starting at $t = 0$ and progressing in positive time in increments of T. The time index where fluorescence is first detected is then our estimate for τ_k. Once detectability is established, (18) can be minimized using the estimated τ_k and a preselected w. To detect the next emitter, in general the estimated forward solution, $\hat{\mathbf{y}}_k = \hat{\eta_k}\mathbf{f}(\hat{\mathbf{r}}_{f_k})$, must be subtracted from $\mathbf{y}$ before moving the temporal window w and calculating the GLRT at the next time point. This removes the influence of one reporter's response from the responses following it, and is valid by superposition. A similar procedure is used in matching pursuit [48], [49]. The robustness of the method can be improved by adding positive noise to $\hat{\mathbf{y}}_k$ before subtraction, where values that are negative after subtraction are set to zero.

We also consider the special case where w can be large compared to the decay time of the fluorescence responses. This is shown in Fig. 3(b), demonstrating that the subtraction step is not needed if the time scan is resumed at $t = \tau_k + w$. However, we use the general case which allows extraction of the response of each emitter from the signals in Fig. 3(a) or (b).

To minimize (18), we use the method developed by Milstein *et al.* [38], which is a two-step procedure described by

$$\tilde{\eta_k}(\mathbf{r}_{f_k}) = \frac{\mathbf{f}_k^T(\mathbf{r}_{f_k})\boldsymbol{\Upsilon}_k^{-1}\mathbf{y}_k}{\mathbf{f}_k^T(\mathbf{r}_{f_k})\boldsymbol{\Upsilon}_k^{-1}\mathbf{f}_k(\mathbf{r}_{f_k})} \tag{19}$$

$$c_k(\mathbf{r}_{f_k}, \tau_k) = ||\mathbf{y}_k - \tilde{\eta_k}(\mathbf{r}_{f_k})\mathbf{f}_k(\mathbf{r}_{f_k})||^2_{\boldsymbol{\Upsilon}_k^{-1}}. \tag{20}$$

Equation (19) was found by setting the derivative of $||\mathbf{y}_k - \eta_k\mathbf{f}_k(\mathbf{r}_k)||^2_{\boldsymbol{\Upsilon}_k^{-1}}$ with respect to η_k equal to zero. Equations (19) and (20) are then solved at positions $\mathbf{r}_{f_k}$ of interest (within the time window w), and the kth fluorescent emitter's position and yield are estimated by

$$\hat{\mathbf{r}}_{f_k} = \arg\min_{\mathbf{r}_{f_k}} c_k(\mathbf{r}_{f_k}, \tau_k) \tag{21}$$

$$\hat{\eta_k} = \tilde{\eta_k}(\hat{\mathbf{r}}_{f_k}). \tag{22}$$

In an experiment, where calibration scale factors are unknown, $\hat{\eta_k}$ can be considered a nuisance parameter [38] where its estimate is no longer quantitative. Interestingly, if $\boldsymbol{\Upsilon}^{-1}$ was

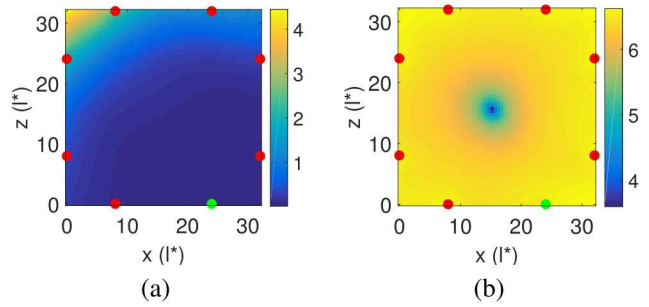


Fig. 4. Localization of a single fluorescent inhomogeneity ($K = 1$) using an (x, z) coordinate system. 1 source (green) and 7 detectors (red) are placed at the boundary of a square of side length $32l^*$. The optical properties are the same as in Fig. 1, where $l^* = 0.5$ mm, and we assume an SNR of 30 dB. A fluorescent inhomogeneity with $\eta_1 = 0.1$ mm^{-1} was placed at $x = 15.26 \times l^*$ and $z = 15.56 \times l^*$. The simulated noisy data is the same as the first set of curves at τ_1 shown in Fig. 3(b). (a) Yield $\tilde{\eta}_k(\mathbf{r}_{f_k})$ from (19) plotted over the region of interest. (b) Cost $c_k(\mathbf{r}_{f_k}, \tau_k)$ from (20) plotted over the region of interest. The position with lowest cost in (b) is $\hat{\mathbf{r}}_{f_k}$, and the value of $\tilde{\eta}_k$ at $\hat{\mathbf{r}}_{f_k}$ in (a) is $\hat{\eta}_k$. Here, $\hat{\mathbf{r}}_{f_k} = (x, z) = (15.24, 15.75) \times l^*$ and $\hat{\eta}_k = 0.0999$ mm^{-1}. The percent errors in the estimated x and z positions are 0.143% and 1.196%, respectively. The discretization of the region of interest is a primary contributor to the estimation error.

an identity matrix, (19) would be equivalent to calculating the basis function weight in matching pursuit. Because of this, the localization method presented here could be considered a constrained version of matching pursuit that incorporates a noise model.

We demonstrate localization of a single fluorescent inhomogeneity using optical properties similar to tissue in Fig. 4. For simplicity, we assume a two-dimensional geometry and a region of interest of length $32 \times l^*$ along x and of length $32 \times l^*$ along z, where the mean free path length $l^* = 3D$. Extension to three dimensions is unnecessary to demonstrate the method, but it is straightforward. We place a single source and seven detectors along the boundary of the region of interest, as seen in Fig. 4. We use a discretized grid with $N_x = 64$ points in the x dimension and $N_z = 64$ points in the z dimension. The position with lowest cost in Fig. 4(b) is $\hat{\mathbf{r}}_f$, and the value of $\tilde{\eta}_k$ at $\hat{\mathbf{r}}_{f_k}$ in Fig. 4(a) is $\hat{\eta}_k$. Here, $\hat{\mathbf{r}}_{f_k} = (x, z) = (15.24, 15.75) \times l^*$ and $\hat{\eta} = 0.0999$ mm^{-1}, which are close to the true values given in the Fig. 4 caption. We calculated the localization error as $[(\mathbf{r}_{f_k} - \hat{\mathbf{r}}_{f_k})/\mathbf{r}_{f_k} \times 100]\%$. The results can be extrapolated to three dimensions, and the problem is scalable in the amount of scatter.

B. Localization with Multiresolution Analysis

We introduce a method to simultaneously reduce the computation time and improve the accuracy of the localization by incorporating a hierarchy of discretization grids into the localization. We refer to this as a multiresolution analysis (MRA) method [41], [50], and it is demonstrated in Fig. 5. First, $c_k(\mathbf{r}_{f_k}, \tau_k)$ is calculated using (19) and (20) on a coarse grid over the entire region of interest. Then, $c_k(\mathbf{r}_{f_k}, \tau_k)$ is iteratively calculated on successively smaller and finer grids, such that each new grid contains the point of minimum cost of the current iteration. This procedure is repeated until convergence, which, in our case, was two successive grids

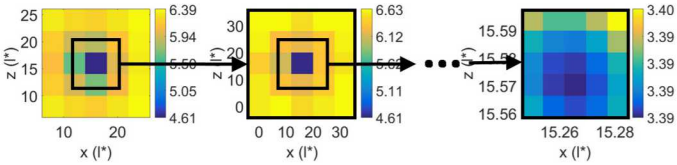


Fig. 5. Localization with MRA of the single fluorescent inhomogeneity in Fig. 4. The cost is calculated using (19) and (20) on progressively finer grids, where each new grid contains the region of smallest cost. Here, $\hat{\mathbf{r}}_{f_k} = (x, z) = (15.27, 15.57) \times l^*$ and $\hat{\eta}_k = 0.1003$ mm^{-1}. The percent errors in the estimated x and z positions are 0.037% and 0.066%, respectively. The discretization error has been reduced, especially for the z coordinate. The number of positions where the cost must be calculated has also been reduced, decreasing the computation time. The reduction is even greater when extrapolated to 3D.

$k \leftarrow 0$
$t_o \leftarrow 0$
while $t_o <= t_{max} - 3 \times T$ **do**
  $w \leftarrow T$
  $Grid \leftarrow coarse\ Grid$
  $\hat{\theta}_{\text{temp}} \leftarrow \min_{\eta_k} ||\mathbf{y}_k - \eta_k \mathbf{f}_k(\mathbf{r}_{f_k})||^2_{\mathbf{\Upsilon}_k^{-1}}$
  $\tilde{q} \leftarrow \mathbf{h}^T(\hat{\theta}_{\text{temp}})\mathbf{y}_k$
  if $\tilde{q} > k_{P_F}(\hat{\theta}_{\text{temp}})$ **then**
    $k \leftarrow k + 1$
    $w \leftarrow 3 \times T$
    for I iterations **do**
      $\hat{\theta}_k \leftarrow \min_{\eta_k} ||\mathbf{y}_k - \eta_k \mathbf{f}_k(\mathbf{r}_{f_k})||^2_{\mathbf{\Upsilon}_k^{-1}}$
      $Grid \leftarrow$ make finer
    end for
    $\mathbf{y} \leftarrow \mathbf{y} - \hat{\mathbf{y}}_k$
  end if
  $t_o \leftarrow t_o + T$
end while

return $\hat{\theta}_k$

Fig. 6. Pseudocode for the temporal localization algorithm using MRA and a window length of $w = 3T$. Note that the detection step can be very fast since $\tilde{q}$ is calculated using only the first coarse grid. For the definitions of $\tilde{q}$, k_{P_F}, and $\mathbf{h}$, see Appendix B.

where the change in the minimum cost was less than 1%, but not equal to zero. Importantly, not converging when the change in cost is zero avoids the situation where the same point happens to have the lowest cost on two grids.

Pseudocode demonstrating the localization of multiple inhomogeneities is shown in Fig. 6 for a window length of $w = 3T$. The MRA method is used to calculate $\hat{\theta}_k$, where the number of multiresolution iterations is I. We found that the full MRA was not needed to calculate $\hat{\theta}_{\text{temp}}$ in order to determine $\tilde{q}$; calculating $\hat{\theta}_{\text{temp}}$ on only the first coarse grid was sufficient, saving significant computation time.

C. Probability of Successful Localization

Our method assumes that there is significant temporal separation τ_f between fluorescence signals. Here, we explore

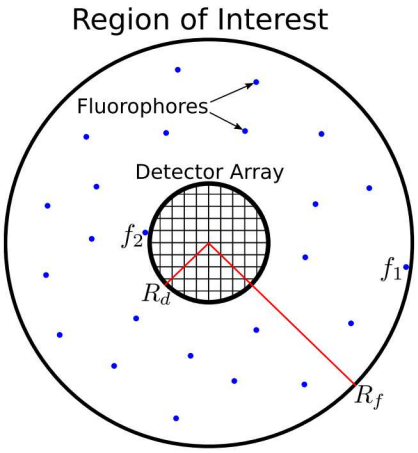


Fig. 7. Geometry used for our stochastic analysis of the requirements on the temporal separation τ_f. We will assume a circular region of interest, with radius R_f. We also assume that all of the detectors are gathered in the center of the region of interest, in a circular array of radius R_d. Two example fluorophores are labeled: f_1, which is far from the detector array, and f_2, which is close to it. Let us denote the number of total fluorophores in the region of interest as N_f. That allows us to define the average fluorophore density ρ_f with the equation $N_f = \rho_f \pi R_f^2$.

that assumption and develop a stochastic model to estimate the probability that it holds.

In order to be able to distinguish between the fluorescence impulse responses from two fluorophores, their peaks must be separated by a significant enough amount of time τ_f. We develop a framework to estimate a conservative value of τ_f such that, if all impulse responses are separated by this amount, the algorithm will function correctly. For now, we assume a geometry described in Fig. 7. As depicted there, R_f is the farthest a fluorophore (for example, f_1 in the figure) can be from the center of the detector array, and $R_f + R_d$ is the longest possible geometrical distance between a fluorophore in the region of interest and a detector in the array. If photons travel at an average speed of $\langle v \rangle$ through the scattering medium, then it would take, on average, $\frac{1}{\langle v \rangle}(R_f + R_d)$ seconds to guarantee that a fluorophore's impulse response reaches all detectors.

However, if another impulse response was to begin after this amount of time, for example fluorophore f_2 in Fig. 7, the peaks of the two impulse responses might still coincide. If we were to wait an additional amount of time equal to the full-width-at-half-max (FWHM) of the widest-possible impulse response, then the peaks would be sufficiently separated in time that our algorithm could work properly, no matter the locations of the two fluorophores in question. Let us denote this additional amount of time as $\Delta t_{\text{FWHM,max}}$ This gives us the conservative value of

$$\tau_f = \frac{1}{\langle v \rangle}(R_f + R_d) + \Delta t_{\text{FWHM,max}}. \tag{23}$$

Next, we will develop a model for estimating the probability of achieving this temporal separation τ_f. Let us model the fluorophore emission times as independent and identically distributed exponential random variables, with rate parameter λ_f emissions per second. This means the fluorophore emissions can be modeled as independent Poisson processes, which is consistent with the shot-noise-limited measurement model described in Section II-E. If emissions are measured over a time interval Δt, then the number of emissions that any given fluorophore gives off is distributed as a Poisson random variable, with rate parameter $\lambda_f \Delta t$. The probability mass function for the number of emissions in this time interval is then

$$\text{Poisson}(\lambda_f \Delta t) \sim \frac{(\lambda_f \Delta t)^n}{n!} e^{-(\lambda_f \Delta t)} \tag{24}$$

Now, suppose there are N_f fluorophores in the region of interest, as described in Fig. 7. By merging these N_f independent and identically distributed Poisson processes, we have that the number of total fluorophore emissions during some time interval Δt is distributed as a Poisson random variable, with rate parameter $N_f \lambda_f \Delta t$. Let us denote this total number of emissions as $F(\Delta t)$. Then, if we denote as F_1 the event that no more than one fluorophore emission occurs in the time interval Δt, its probability is

$$\begin{aligned} P\{F_1\} &= P\{F(\Delta t) = 0 \cup F(\Delta t) = 1\} \\ &= (1 + N_f \lambda_f \Delta t)\, e^{-(N_f \lambda_f \Delta t)}. \end{aligned} \tag{25}$$

The value given in (23) for τ_f is a conservative one; having fluorophore emissions separated by this amount of time is a sufficient condition for the algorithm to function correctly, but not a necessary condition. Still, suppose we require enough temporal separation of fluorophore emissions such that each measured time interval contains at most one event. That is, let us conservatively assign $\Delta t = \tau_f$. If S is the event that the algorithm is able to function correctly, then we can bound the probability of this event as

$$P\{S\} \geq (1 + \lambda_F)\, e^{-\lambda_F}, \tag{26}$$

where

$$\begin{aligned} \lambda_F &= N_f \lambda_f \Delta t \\ &= N_f \lambda_f \tau_f \\ &= \rho_f \pi R_f^2 \lambda_f \left(\frac{R_f + R_d}{\langle v \rangle} + \Delta t_{\text{FWHM,max}} \right). \end{aligned} \tag{27}$$

Equation (26) shows that we can increase the probability of successful localization by decreasing λ_F, as defined in (27). Of course, we are free to choose a measurement interval $\Delta t < \tau_f$, which would increase $P\{S\}$ further by decreasing λ_F. Indeed, the lower bound on Δt is limited only by one's detection system and the noise involved, thus linking $P\{S\}$ to the SNR that one is able to achieve. Additionally, if decreasing the average fluorophore emission rate λ_f is too difficult for one's application, it may be easier to decrease the emitter density ρ_f instead. The average speed $\langle v \rangle$ is material- and geometry-dependent, so that may be harder to affect. Importantly, λ_F depends cubically on the radius R_f of the region of interest, but this is often simple to adjust. While (26) and (27) only provide an approximation for the probability of success of our localization algorithm, they are nonetheless useful for analyzing it and how it relates to the experimental setup.

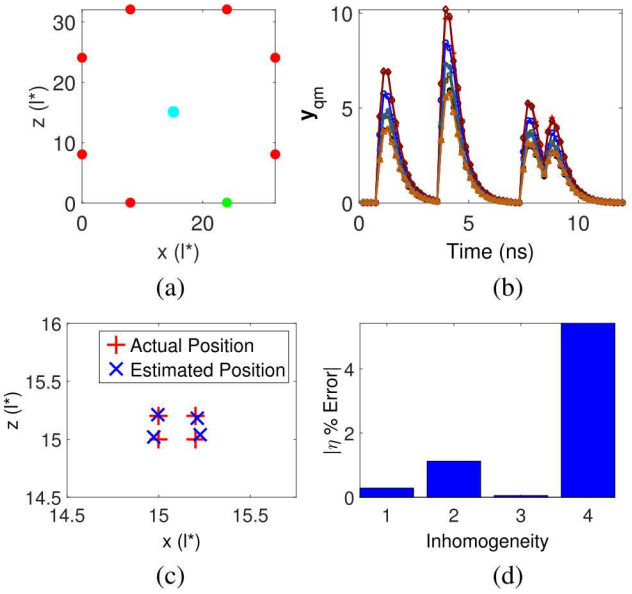


Fig. 8. Localization of four fluorescent inhomogeneities at different positions with different delays and yields using the same optical parameters as in Fig. 1 and the same geometry as in Fig. 4. The true parameters describing the inhomogeneities are $(x, z, \eta, \tau) = (15l^*, 15l^*, 0.1, 5T)$, $(15l^*, 15.2l^*, 0.15, 20T)$, $(15.2l^*, 15.0l^*, 0.075, 40T)$, and $(15.2l^*, 15.2l^*, 0.05, 46T)$, where $l^* = 0.5$ mm and $T = 0.19$ ns. All of these parameters are estimated by the algorithm. We assume the fluorescence lifetime τ_f is known. (a) Problem geometry as in Fig. 4, where the positions of the excitation source (green), detectors (red), and fluorescent inhomogeneities (cyan) are plotted. (b) Detected fluorescence temporal profile. One source ($Q = 1$) and seven detectors ($M = 7$) give 7 measurements $\mathbf{y}_{qm}$. The 7 different symbols and their corresponding colors represent different source-detector measurement pairs. The data was generated using the true parameters with simulated noise for an SNR of 30 dB. (c) True positions of the inhomogeneities $\mathbf{r}_k$ and the estimated positions $\hat{\mathbf{r}}_{f_k}$ determined by the localization algorithm. Note the accuracy of the estimated positions. (d) Yield η_k experimental errors. Labels one to four correspond to delays from shortest to longest. Each fluorescent inhomogeneity was successfully localized, even for the case when there is overlap between the temporal signals.

IV. Results

We show localization of four different fluorescent inhomogeneities separated by varying delays. We use the same geometry and optical parameters as in Fig. 4, except we place four inhomogeneities within the medium instead of one.

Figure 8(a) shows the problem geometry as in Fig. 4, where the positions of the excitation source (green), detectors (red), and fluorophores (cyan) are shown. Figure 8(b) shows the simulated noisy data from the four inhomogeneities with delays of 0.9, 3.75, 7.5, and 8.6 ns. The first two inhomogeneities are temporally separated, while the second two have a significant temporal overlap. Figure 8(c) plots the actual and predicted locations of the inhomogeneities, which are separated by $0.2 \times l^*$. Figure 8(d) shows the yield η (proportional to concentration) experimental errors, calculated as $[(\eta_k - \hat{\eta}k)/\eta_k \times 100]$. For the multiresolution operation, we used $N_x = N_z = 5$ equally spaced points for each grid, the same as what was used in Fig. 5. The first coarse grid was defined over the entire square region of interest from Fig. 4, and each finer square grid was centered at (21) and extended a distance slightly greater than the previous grid spacing along the x and z directions (similar to Fig. 5). Both the position and η were predicted with high accuracy, even for the case when there is overlap between the temporal signals. The accuracy of the localization is the focus of Section IV-A.

A. Localization for High Spatial Resolution

We assume insignificant background signal and that the fluorophores do not diffuse or change positions significantly during the integration time. In this case, the localization uncertainty is dominated by measurement noise and each $\hat{\mathbf{r}}_{k_f}$ coordinate x and z has a Gaussian probability distribution characterized by the standard deviations σ_x and σ_z [51], [52]. These standard deviations are commonly used to characterize the spatial precision of localization [53], [54]. Here, we show the capability of localization to extract high-spatial-resolution information by generating these statistics from numerical calculations.

We iteratively localize the single emitter in Fig. 4 using MRA, and the results are shown in Fig. 9. During each iteration, random noise was added to the forward calculation in order to generate independent simulated measurements. Both σ_x and σ_z were then calculated from the localized positions and used to plot ellipses with major and minor axis equal to $4\sigma_x$ or $4\sigma_z$. The MRA method described in Section III-B was important in reducing the time required to compute these statistics, as discussed in Sect. IV-B.

Figures 9(a) and (b) show localization uncertainly statistics for different SNR, where the noise was added as described by (13) [40]. Even for a low SNR of 10 dB, the location of the fluorescent inhomogeneity can be estimated with much higher accuracy compared to traditional diffusive imaging methods, as seen for example in Fig. 1. Figure 9(c) and (d) show localization uncertainly statistics for different numbers of detectors M. The detectors were distributed as in Fig. 4, and the statistics depend on the spatial support. Compared to volumetric image reconstruction with FDOT, little information, or fewer detectors, are needed to localize the fluorophore.

Considering Fig. 6, it is not clear how or if the localization uncertainty will depend on the window length w. To study this, we build up localization uncertainly statistics for different w with a constant SNR of 30 dB and $M = 7$. The results are shown in Fig. 9(e) and (f). For all results in this figure (and also in Fig. 8), the time axis starts at $t = 0$ and continues in increments of $T = 0.19$ ns for 64 increments, giving $t_{max} = 12.16$ ns. As can be seen in Fig. 9(e) and (f), the localization statistics change slightly with window size, but they are still highly accurate even for the shortest window of length $2T$.

The localization uncertainty of the inhomogeneities in Fig. 8(c) is described by the blue ellipse in Fig. 9(b). The extension to imaging is straightforward. If we consider an image which is made up of voxels, and each voxel can be individually localized, then Fig. 9 describes the resolution limit of the image. Comparing to Fig. 1 where the resolution limit is less than about $8.5l^*$, the resolution limit from Fig. 9(b), or the minimum distance between two inhomogeneities such that they can both be accurately localized, is about $0.1l^*$ or 50 μm. This is a remarkable spatial resolution for deep tissue optical imaging.

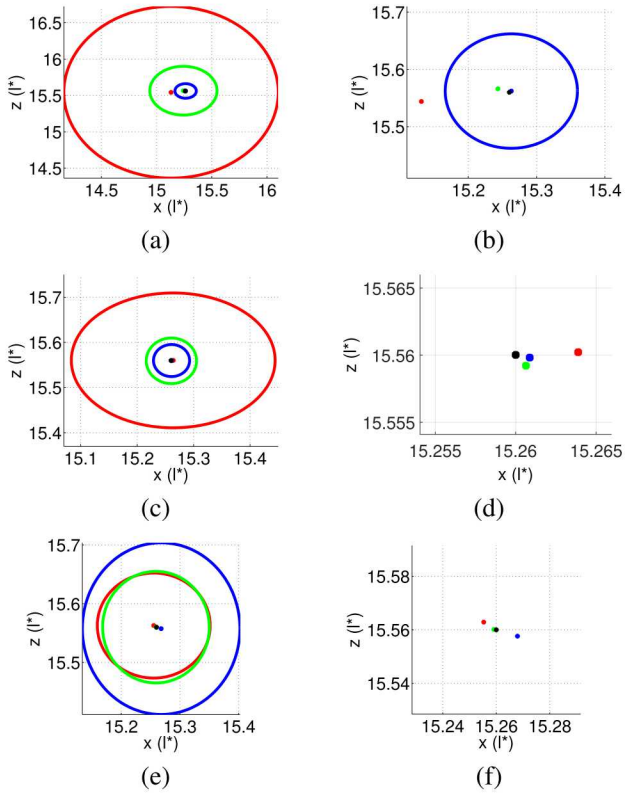


Fig. 9. Localization uncertainty of a single fluorescent inhomogeneity using the same optical parameters as in Fig. 1 and the same geometry as in Fig. 4 with $Q = 1$. The fluorescent inhomogeneity location was estimated 150 times using noisy simulated independent data sets. The true location is the black point. The ellipses have major and minor axes of length $4\sigma_x$ or $4\sigma_z$, such that they contain 95% of the $\hat{x}$ and $\hat{z}$ positions. The center points of the ellipses are the mean of the $\hat{x}$ and $\hat{z}$ positions. (a) Localization uncertainty for different SNR with $M = 7$ and $w = 3T$. Blue, green, and red correspond to SNR of 30, 20, and 10 dB, respectively. (b) Zoomed version of (a) to show the mean values. (c) Localization uncertainty for different numbers of detectors M with an SNR of 30 dB and $w = 3T$. Red, green, and blue correspond to $M = 7$, $M = 31$, and $M = 50$. (d) Zoomed version of (c) to show the mean values. (e) Localization uncertainty for different window lengths w, 30 dB SNR, and $M = 7$. Red, green, and blue correspond to windows $w = 32T$, $17T$, and $2T$, where $T = 0.19$ ns and $t_{max} = 64T$. (f) Zoomed version of (e) to show the mean values. The ellipses are not circles because the fluorescent inhomogeneity is not located at the center of the medium and equidistant to all detectors. Note that the fluorescent inhomogeneity can be accurately localized even with low SNR, few detectors, and a short window w.

B. Run Time Analysis

We investigate the computational advantage of the MRA localization algorithm shown in Fig. 6 compared to matching pursuit, an established method for sparse approximation. The run time of matching pursuit is (number of iterations) $\times$ (time per iteration) [48], [49], where on average there are K iterations. The time per iteration is (number of inner products) $\times$ (time per inner product). We let $N_{x_{MP}}$ and $N_{z_{MP}}$ be the number of grid points along the x and z dimensions for matching pursuit (MP), where $N_{x_{MP}}$ and $N_{z_{MP}}$ must be very large in order to achieve high spatial resolution. To minimize (17), matching pursuit must calculate $\mathcal{O}(N_{x_{MP}} N_{z_{MP}} N)$ inner products since $t_{max} = NT$, and each inner product involves $\mathcal{O}(QMN)$ multiply-add operations. Therefore the expected run time of matching pursuit is $\mathcal{O}(QMN_{x_{MP}} N_{z_{MP}} N^2 K)$. The run time of the Fig. 6 algorithm is (time for the threshold tests) + (time for localizations). Both the threshold tests and the localizations require repeated evaluation of grids, each representing progressively finer resolution. Each of these grids has the same number of points as the first (lowest-resolution) one. Suppose this grid has N_{MG} points. Since each of these grid points could potentially be the one visited in a second iteration of MRA, there are N_{MG}^2 points represented in a two-iteration MRA algorithm. Since we perform I iterations and eventually reach $N_{x_{MP}} N_{z_{MP}}$ points, it must be that $N_{MG}^I = N_{x_{MP}} N_{z_{MP}}$, or $N_{MG} = (N_{x_{MP}} N_{z_{MP}})^{(1/I)}$. So, the time for the scanning threshold tests is $\mathcal{O}[QM(N_{x_{MP}} N_{z_{MP}})^{(1/I)} N]$, where the number of multiresolution iterations I to achieve the same discretization as matching pursuit is typically small (< 10). We expect there to be K localizations on average, so the time spent for localization is $\mathcal{O}[KQM(N_{x_{MP}} N_{z_{MP}})^{(1/I)} I]$, using similar analysis as before for the number of points in each multiresolution iteration. Therefore the expected run time of the Fig. 6 algorithm is $\mathcal{O}[QM(N_{x_{MP}} N_{z_{MP}})^{(1/I)}(N + KI)]$. We summarize these results in Table I.

TABLE I
RUN TIME COMPARISON WITH MATCHING PURSUIT

	Run Time
Matching Pursuit	$\mathcal{O}(QMN_{x_{MP}} N_{z_{MP}} N^2 K)$
Fig. 6 Algorithm	$\mathcal{O}[QM(N_{x_{MP}} N_{z_{MP}})^{(1/I)}(N + KI)]$

There are some clear advantages to the MRA algorithm presented in Fig. 6. First, the run time is linear in N instead of quadratic. Second, the grid size ($(N_{x_{MP}} N_{z_{MP}})^{(1/I)}$) is much smaller because of the MRA method. In situations where a high-resolution image is desired, $N_{x_{MP}}$, $N_{z_{MP}}$, and N must all be very large, severely limiting the practicality of matching pursuit. We show example run times in Fig. 10 using a 3.47 GHz Intel X5690 with 96 GB RAM. It is clear from the plot that for large N, matching pursuit is quadratic in N while the new localization method remains linear. For matching pursuit, the amount of memory needed to store the large dictionary is $8 \times (QMN_{x_{MP}} N_{z_{MP}} N^2)$ bytes, and using the parameters given in Fig. 10, this gives an ideal maximum value of $N = 855$. In practice, more than the dictionary must be saved in RAM, and in Fig. 10 the maximum value of N possible for matching pursuit was found to be 481.

V. Discussion

We have shown that localization is a powerful tool for retrieving high spatial and temporal resolution information of fluorescent inhomogeneities embedded in scattering media. To accomplish this, we have assumed as prior information that the data from each fluorescent inhomogeneity is separated by a varying and increasing delay.

Other super-resolution methods have been developed for improving the spatial resolution beyond the diffraction limit in microscopy. Fundamentally, imaging methods can surpass resolution limits with the addition of prior information to compensate for the information that is lost due to attenuation or randomization of the signal. For example, structured

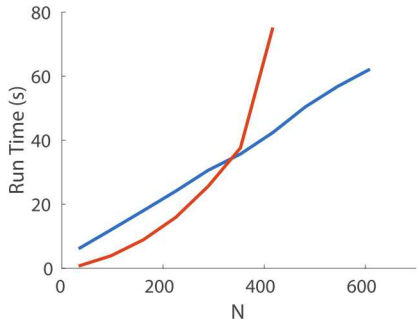


Fig. 10. Run times of matching pursuit (red) and the Fig. 6 algorithm (blue) for varying N, where $t_{max} = NT$. We use the same optical properties as Fig. 1, and the same geometry as Fig. 4. The computation time required to generate and load the dictionary by matching pursuit is not included. Including this almost triples the run times shown in the plot.

illumination microscopy (SIM) [55] breaks the resolution limit through spatial modulation of the excitation light, and can be further improved by leveraging the nonlinear dependence of the fluorescent emission rate in the saturation regime [56]. Stimulated emission depletion (STED) [57] forms a smaller effective point spread function (PSF) by saturating fluorophores at the periphery of the focal point. Other techniques such as photoactivated localization microscopy (PALM) [58], fluorescence PALM (fPALM) [59], and stochastic optical reconstruction microscopy (STORM) [60] are able to localize switchable fluorescent molecules by distinguishing the emission between their fluorescent and non-fluorescent states. These are all super-resolution microscopy methods that achieve position information beyond that directly represented by the point spread function. The shared use of temporal switching of fluorescence states in fluorescent molecules makes our method analogous to PALM, fPALM, and STORM. However, our incorporation of a forward model, such as the coupled diffusion model for scattered light we use here, allows imaging through scatter. Moreover, our super-resolution imaging method can be applied to a wider variety of applications that employ forward models, such as photoacoustic tomography [58], electrical impedance tomography [59], seismic waveform tomography [60], and microwave imaging [61].

One important potential application of our method is brain imaging, because the response of fluorescent contrast agents due to neurons or groups of neurons firing may follow a model similar to in (9) [62]–[66]. If this is found to be true, and a reasonable SNR can be achieved, then it would be possible to form images of the whole brain or brain surface with high spatial and temporal resolution. These images could be used to form correlation maps of brain activity, useful for diagnosing and studying neurological diseases such as Parkinson's or Alzheimer's disease, and developing treatments [67]. Considering Fig. 4, only one source and a few detectors are required, simplifying the experiment setup. The experiment could be performed in the time domain with fast and sensitive detectors, or with an integration time equal to w if the delays τ_k are sufficiently large. If an intensified CCD camera is used, gated measurements of pacemaking neurons, which fire periodically at 2-3 Hz [15], could be performed, assuming the jitter is insignificant. Alternatively, data captured with an intensified CCD over an integration time of w could be used to separate temporally overlapping neuron responses in time and in space due to the large number of detectors. Another potential application is the localization of blinking quantum dots [68]. There are likely additional biomedical applications or physical problems where fast dynamics or kinetics could be studied.

Semiconductor quantum dots (QDs) are popular in fluorescence imaging due to their convenient absorption and emission spectrum properties and greater brightness compared to most organic dyes. However, they commonly exhibit an inherent intermittence in their emission (blinking), a property that is often considered a nuisance, and work has been done to suppress it. [69], [70]. Still, some believe that this blinking could be exploited, and it has been shown that this blinking can be affected by variables such as the excitation power, the shell material about the QD, and environment conditions such as the temperature and surrounding medium [71]. We believe this inherent blinking of QDs could allow our localization method to be used in many application where the required temporal separation of fluorescent emissions would not otherwise occur.

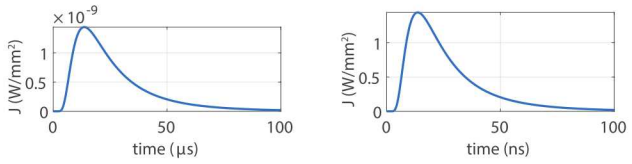
Outside of fluorescence imaging altogether, our method may be applied to, for example, acoustic waves generated by multiple speakers in a room. A suitable forward model can easily be obtained if the shape of the room is known in advance, and the required intermittent generation of sound may arise in normal conversation between people. Our method could then be used to localize these speakers, and once that is achieved, we envision the location-specific amplification or attenuation of certain speakers, thus allowing a measure of control over the acoustic environment.

VI. Conclusion

We have developed an approach for the fast localization of multiple fluorescent inhomogeneities within a highly scattering medium that takes advantage of variations in temporal delays between responses. The method allows formation of super-resolution optical images from highly scattered light. We demonstrate through simulations that MRA and temporal scanning can significantly reduce the computational burden compared to matching pursuit. The geometries can be scaled and other forward models can be used, allowing a broad range of applications.

Appendix A
Scaling the Diffusion Equation

In this Appendix we show scaling of the spatial dimensions of the diffusion equation. Considering (1) and (2) along with (5), it is clear that the dimensions of the diffusion problem can be scaled [72]. We let $\mu_a = 1/l_a$, where l_a is the distance between absorption events, and $\mu_s = 1/l_s$ where l_s is the distance between scattering events. Then these parameters can be scaled by a parameter α_s, resulting in $\tilde{\mathbf{r}} = \alpha_s \mathbf{r}$, $\tilde{\mu_a} = \mu_a/\alpha_s$, $\tilde{\mu_s} = \mu_s/\alpha_s$, $\tilde{g} = g$, and $\tilde{c} = \alpha_s c$, giving $\tilde{D} = \alpha_s D$. This scaling allows a material with widely spaced scatterers to be modeled by a more densely packed material or vice versa.



(a) (b)

Fig. 11. Scaling of the current density from (6) detected through cloud. In (a), the cloud is 1 km thick and $\mu_a = 0.1$ km^{-1}, $\mu_s' = 10$ km^{-1}, and $n = 1.5$, which are typical measured values [73]. In (b), $\alpha_s = 1/1000$ scaling is applied with the result that $\tilde{\mu_a} = 0.1$ m^{-1}, $\tilde{\mu_s}' = 10$ m^{-1}, a thickness of 1 m, and we keep $n = 1.5$. Note that the time axis is scaled by α_s and the current density J is scaled by α_s^{-3}. If we had scaled the refractive index as $\tilde{n} = n/\alpha_s$, then the current density J would be scaled by α_s^{-2} and the time axis would not be scaled.

Significant scaling of the velocity or refractive index is not realistic, however in the case of unmodulated light ($\omega = 0$ in the frequency domain) this is not a problem. Alternatively, the time axis can be scaled as $\tilde{t} = \alpha_s t$ instead of scaling the velocity. Figure 11 shows scaling of (6) with $\alpha_s = 1/1000$.

APPENDIX B
DETECTION

In this Appendix we develop the method used to detect a single fluorescent inhomogeneity. The model is used to determine the threshold needed for a decision to localize an inhomogeneity during the temporal scan. Following Milstein *et al.* [38], a fluorescent inhomogeneity is detected for a specified false-alarm rate P_F. We let hypothesis H_0 correspond to the absence of a fluorescent inhomogeneity, and we let $H_{1,\boldsymbol{\theta}_k}$ correspond to the presence of a fluorescent inhomogeneity that is parameterized by the vector $\theta_k = [\ \mathbf{r}_{f_k} \quad \eta_k\ ]^T$. Then the densities for measurement vector $\mathbf{y}_k$ of length P under both hypotheses are given by

$$p_0(\mathbf{y}_k) = \frac{1}{\sqrt{(2\pi)^P |\boldsymbol{\Upsilon}_k|}} \exp\left(-\frac{1}{2}||\mathbf{y}_k||^2_{\boldsymbol{\Upsilon}_k^{-1}}\right) \tag{28}$$

$$p_{1,\theta_k}(\mathbf{y}_k) = \frac{1}{\sqrt{(2\pi)^P |\boldsymbol{\Upsilon}_k|}} \exp\left(-\frac{1}{2}||\mathbf{y}_k - \eta_k \mathbf{f}_k(\mathbf{r}_{f_k})||^2_{\boldsymbol{\Upsilon}_k^{-1}}\right). \tag{29}$$

Then, we form the log likelihood ratio

$$L(\mathbf{y}_k, \theta_k) = \ln \frac{p_{1,\theta_k}(\mathbf{y}_k)}{p_0(\mathbf{y}_k)} = \mathbf{h}^T(\theta_k)\mathbf{y}_k - c(\theta_k), \tag{30}$$

where $\mathbf{h}^T(\theta_k) = \eta_k \mathbf{f}_k^T(\mathbf{r}_{f_k})\boldsymbol{\Upsilon}_k^{-1}$ and $c(\theta_k) = \frac{1}{2}\eta_k^2 \mathbf{f}_k^T(\mathbf{r}_{f_k})\boldsymbol{\Upsilon}_k^{-1}\mathbf{f}_k(\mathbf{r}_{f_k})$. By the Neyman-Pearson lemma, the likelihood ratio test (LRT) produces the highest probability of detection for a specified false alarm rate of P_F. For this paper, we used $P_F = 0.03$ to generate all related results. We declare an inhomogeneity present when

$$L(\mathbf{y}_k, \theta_k) > \tilde{k}_{P_F}(\theta_k). \tag{31}$$

where $\tilde{k}_{P_F}(\theta_k)$ is a threshold determined by P_F. The LRT suggests that if the correlation between the data $\mathbf{y}_k$ and the model $\mathbf{f}_k(\mathbf{r}_{f_k})$ is above a threshold, an inhomogeneity exists.

We define $\tilde{q} = \mathbf{h}^T(\theta_k)\mathbf{y}_k$ as the decision statistic, and we find that it has a normal distribution under each of the two hypotheses, such that $(\tilde{q}|H_0) = N(0, \sigma_{\tilde{q}}^2)$ and $(\tilde{q}|H_{1,\boldsymbol{\theta}_k}) = N(\mu_{\tilde{q}}, \sigma_{\tilde{q}}^2)$, where the mean $\mu_{\tilde{q}} = \eta_k^2 \mathbf{f}_k^T(\mathbf{r}_{f_k})\boldsymbol{\Upsilon}_k^{-1}\mathbf{f}_k(\mathbf{r}_{f_k})$ and the variance $\sigma_{\tilde{q}}^2 = \mathbf{h}^T(\theta_k)\boldsymbol{\Upsilon}_k\mathbf{h}(\theta_k)$. Equation (31) is then equivalent to $\tilde{q} > k_{P_F}(\theta_k)$, where $k_{P_F}(\theta_k)$ is a modified threshold that we calculate below.

By definition, the false alarm rate P_F is given by

$$P_F = \int_{k_{P_F}(\theta_k)}^{\infty} p_0(\tilde{q})d\tilde{q} = \frac{1}{2}\text{erfc}\left(\frac{k_{P_F}(\theta_k)}{\sigma_{\tilde{q}}}\right), \tag{32}$$

where we used the substitution $u = 1/\left(\sqrt{2}\sigma_{\tilde{q}}\right)$ and

$$\text{erfc}(r) = \int_r^{\infty} \frac{2}{\sqrt{\pi}} \exp\left(-u^2\right) du, \tag{33}$$

is the complementary error function. The threshold $k_{P_F}(\theta_k)$ can then be calculated as

$$k_{P_F}(\theta_k) = \sqrt{2}\sigma_{\tilde{q}}\text{erfc}^{-1}(2P_F), \tag{34}$$

where $\text{erfc}^{-1}(\cdot)$ is the inverse complementary error function. Using this threshold, the probability of detection P_D can be computed as

$$P_D = \int_{k_{P_F}(\theta_k)}^{\infty} p_{1,\theta_k}(\tilde{q})d\tilde{q} = \frac{1}{2}\text{erfc}\left(\frac{k_{P_F}(\theta_k) - \mu_{\tilde{q}}}{\sqrt{2}\sigma_{\tilde{q}}}\right). \tag{35}$$

The LRT assumes θ_k is known, but in practice it must be estimated. We use the generalized likelihood ratio test (GLRT), where θ_k in the LRT is replaced by the ML estimate $\hat{\theta}_k = \arg\max_{\theta_k} p_{1,\theta_k}(\mathbf{y})$. Calculating $\hat{\theta}_k$ constitutes localization.

ACKNOWLEDGMENT

The authors would like to acknowledge funding through the National Science Foundation (NSF) (CBET-0854249, CISE-1218909) and the National Institutes of Health (NIH) (1R21CA182235-01A1). Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

REFERENCES

[1] V. Ntziachristos, C. Bremer, and R. Weissleder, "Fluorescence imaging with near-infrared light: new technological advances that enable *in vivo* molecular imaging," *Eur. Radiol.*, vol. 13, no. 1, pp. 195–208, 2003.

[2] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*. Springer, 2009.

[3] S. A. Hilderbrand and R. Weissleder, "Near-infrared fluorescence: application to *in vivo* molecular imaging," *Curr. Opin. Chem. Biol.*, vol. 14, no. 1, pp. 71–79, 2010.

[4] J. White, W. Amos, and M. Fordham, "An evaluation of confocal versus conventional imaging of biological structures by fluorescence light microscopy." *J. Cell. Biol.*, vol. 105, no. 1, pp. 41–48, 1987.

[5] W. Denk, J. H. Strickler, and W. W. Webb, "Two-photon laser scanning fluorescence microscopy," *Science*, vol. 248, no. 4951, pp. 73–76, 1990.

[6] A. B. Milstein, S. Oh, J. S. Reynolds, K. J. Webb, C. A. Bouman, and R. P. Millane, "Three-dimensional Bayesian optical diffusion tomography with experimental data," *Opt. Lett.*, vol. 27, no. 2, pp. 95–97, January 15 2002.

[7] M. Schweiger and S. Arridge, "The Toast++ software suite for forward and inverse modeling in optical tomography," *J. Biomed. Opt.*, vol. 19, no. 4, pp. 040 801–040 801, 2014.

[8] A. B. Milstein, S. Oh, K. J. Webb, C. A. Bouman, Q. Zhang, D. A. Boas, and R. P. Millane, "Fluorescence optical diffusion tomography," *Appl. Opt.*, vol. 42, no. 16, pp. 3081–3094, Jun. 2003.

[9] V. Y. Soloviev, C. D'Andrea, G. Valentini, R. Cubeddu, and S. R. Arridge, "Combined reconstruction of fluorescent and optical parameters using time-resolved data," *Appl. Opt.*, vol. 48, no. 1, pp. 28–36, 2009.

[10] V. Gaind, K. J. Webb, S. Kularatne, and C. A. Bouman, "Towards *in vivo* imaging of intramolecular fluorescence resonance energy transfer parameters," *J. Opt. Soc. Am. A*, vol. 26, pp. 1805–1813, 2009.

[11] V. Gaind, S. Kularatne, P. S. Low, and K. J. Webb, "Deep tissue imaging of intramolecular fluorescence resonance energy transfer parameters," *Opt. Lett.*, vol. 35, pp. 1314–1316, 2010.

[12] K. Chung, J. Wallace, S.-Y. Kim, S. Kalyanasundaram, A. S. Andalman, T. J. Davidson, J. J. Mirzabekov, K. A. Zalocusky, J. Mattis, A. K. Denisin, S. Pak, H. Bernstein, C. Ramakrishnan, L. Grosenick, V. Gradinaru, and K. Deisseroth, "Structural and molecular interrogation of intact biological systems," *Nature*, vol. 497, no. 7449, pp. 332–337, 2013.

[13] T. Thestrup, J. Litzlbauer, I. Bartholomäus, M. Mues, L. Russo, H. Dana, Y. Kovalchuk, Y. Liang, G. Kalamakis, Y. Laukat, S. Becker, G. Witte, A. Geiger, T. Allen, L. C. Rome, T.-W. Chen, D. S. Kim, O. Garaschuk, C. Griesinger, and O. Griesbeck, "Optimized ratiometric calcium sensors for functional *in vivo* imaging of neurons and t lymphocytes," *Nat. Meth.*, vol. 11, no. 2, pp. 175–182, 2014.

[14] J. T. Vogelstein, B. O. Watson, A. M. Packer, R. Yuste, B. Jedynak, and L. Paninski, "Spike inference from calcium imaging using sequential monte carlo methods," *Biophys. J.*, vol. 97, no. 2, pp. 636 – 655, 2009.

[15] J. N. Guzman, J. Sánchez-Padilla, C. S. Chan, and D. J. Surmeier, "Robust pacemaking in substantia nigra dopaminergic neurons," *J. Neurosci.*, vol. 29, no. 35, pp. 11 011–11 019, 2009.

[16] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation," *Proc. Nation. Acad. Sci.*, vol. 87, no. 24, pp. 9868–9872, 1990.

[17] D. Kobat, N. G. Horton, and C. Xu, "In vivo two-photon microscopy to 1.6-mm depth in mouse cortex," *J. Biomed. Opt.*, vol. 16, no. 10, pp. 106 014–106 014, 2011.

[18] G. J. Goldey, D. K. Roumis, L. L. Glickfeld, A. M. Kerlin, R. C. Reid, V. Bonin, D. P. Schafer, and M. L. Andermann, "Removable cranial windows for long-term imaging in awake mice," *Nat. Protoc.*, vol. 9, no. 11, pp. 2515–2538, 2014.

[19] S. R. Arridge, "Optical tomography in medical imaging," *Inverse Probl.*, vol. 15, no. 2, p. R41, 1999.

[20] V. Ntziachristos, C.-H. Tung, C. Bremer, and R. Weissleder, "Fluorescence molecular tomography resolves protease activity *in vivo*," *Nat. Med.*, vol. 8, no. 7, pp. 757–761, 2002.

[21] V. Ntziachristos and R. Weissleder, "Charge-coupled-device based scanner for tomography of fluorescent near-infrared probes in turbid media," *Med. Phys.*, vol. 29, no. 5, pp. 803–809, 2002.

[22] D. Hall, G. Ma, F. Lesage, and Y. Wang, "Simple time-domain optical method for estimating the depth and concentration of a fluorescent inclusion in a turbid medium," *Opt. Lett.*, vol. 29, no. 19, pp. 2258–2260, 2004.

[23] N. Ducros, A. Da Silva, J.-M. Dinten, C. S. Seelamantula, M. Unser, and F. Peyrin, "A time-domain wavelet-based approach for fluorescence diffuse optical tomography," *Med. Phys.*, vol. 37, no. 6, pp. 2890–2900, 2010.

[24] D. A. Boas, M. OLeary, B. Chance, and A. Yodh, "Detection and characterization of optical inhomogeneities with diffuse photon density waves: a signal-to-noise analysis," *Appl. Opt.*, vol. 36, no. 1, pp. 75–92, 1997.

[25] X. Li, M. O'Leary, D. Boas, B. Chance, and A. Yodh, "Fluorescent diffuse photon density waves in homogeneous and heterogeneous turbid media: analytic solutions and applications," *Appl. Opt.*, vol. 35, no. 19, pp. 3746–3758, 1996.

[26] J. Ripoll, M. Nieto-Vesperinas, and R. Carminati, "Spatial resolution of diffuse photon density waves," *J. Opt. Soc. Am. A*, vol. 16, no. 6, pp. 1466–1476, 1999.

[27] A. Gandjbakhche, R. Nossal, and R. Bonner, "Resolution limits for optical transillumination of abnormalities deeply embedded in tissues," *Med. Phys.*, vol. 21, no. 2, pp. 185–191, 1994.

[28] B. W. Pogue, T. O. McBride, U. L. Osterberg, and K. D. Paulsen, "Comparison of imaging geometries for diffuse optical tomography of tissue," *Opt. Exp.*, vol. 4, no. 8, pp. 270–286, 1999.

[29] E. E. Graves, J. Ripoll, R. Weissleder, and V. Ntziachristos, "A submillimeter resolution fluorescence molecular imaging system for small animal imaging," *Med. Phys.*, vol. 30, no. 5, pp. 901–911, 2003.

[30] D. Boas, K. Chen, D. Grebert, and M. Franceschini, "Improving the diffuse optical imaging spatial resolution of the cerebral hemodynamic response to brain activation in humans," *Opt. Lett.*, vol. 29, no. 13, pp. 1506–1508, 2004.

[31] L. Zhao, V. K. Lee, S.-S. Yoo, G. Dai, and X. Intes, "The integration of 3-D cell printing and mesoscopic fluorescence molecular tomography of vascular constructs within thick hydrogel scaffolds," *Biomaterials*, vol. 33, no. 21, pp. 5325–5332, 2012.

[32] M. S. Ozturk, V. K. Lee, L. Zhao, G. Dai, and X. Intes, "Mesoscopic fluorescence molecular tomography of reporter genes in bioprinted thick tissue," *J. Biomed. Opt.*, vol. 18, no. 10, pp. 100 501–100 501, 2013.

[33] H. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson, "Optical image reconstruction using frequency domain data: simulations and experiments," *J. Opt. Soc. Am. A*, vol. 13, no. 2, pp. 253–266, Feb 1996.

[34] M. S. Patterson, B. Chance, and B. C. Wilson, "Time resolved reflectance and transmittance for the non-invasive measurement of tissue optical properties," *Appl. Opt.*, vol. 28, no. 12, pp. 2331–2336, 1989.

[35] V. G. Peters, D. Wyman, M. Patterson, and G. Frank, "Optical properties of normal and diseased human breast tissues in the visible and near infrared," *Phys. Med. Biol.*, vol. 35, no. 9, p. 1317, 1990.

[36] J. B. Fishkin and E. Gratton, "Propagation of photon-density waves in strongly scattering media containing an absorbing semi-infinite plane bounded by a straight edge," *J. Opt. Soc. Am. A*, vol. 10, no. 1, pp. 127–140, 1993.

[37] R. C. Haskell, L. O. Svaasand, T.-T. Tsay, T.-C. Feng, B. J. Tromberg, and M. S. McAdams, "Boundary conditions for the diffusion equation in radiative transfer," *J. Opt. Soc. Am. A*, vol. 11, no. 10, pp. 2727–2741, 1994.

[38] A. B. Milstein, M. D. Kennedy, P. S. Low, C. A. Bouman, and K. J. Webb, "Statistical approach for detection and localization of a fluorescing mouse tumor in Intralipid," *Appl. Opt.*, vol. 44, no. 12, pp. 2300–2310, 2005.

[39] G. Cao, V. Gaind, C. A. Bouman, and K. J. Webb, "Localization of an absorbing inhomogeneity in a scattering medium in a statistical framework," *Opt. Lett.*, vol. 32, no. 20, pp. 3026–3028, 2007.

[40] J. C. Ye, K. J. Webb, C. A. Bouman, and R. P. Millane, "Optical diffusion tomography by iterative-coordinate-descent optimization in a Bayesian framework," *J. Opt. Soc. Am. A*, vol. 16, no. 10, pp. 2400–2412, October 1999.

[41] J. C. Ye, C. A. Bouman, K. J. Webb, and R. P. Millane, "Nonlinear multigrid algorithms for Bayesian optical diffusion tomography," *IEEE Trans. Image Process.*, vol. 10, no. 6, pp. 909–922, 2001.

[42] H. Schau and A. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Trans. Acoust. Speech*, vol. 35, no. 8, pp. 1223–1225, 1987.

[43] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust. Speech*, vol. 36, no. 10, pp. 1553–1560, 1988.

[44] E. L. Hull, M. G. Nichols, and T. H. Foster, "Localization of luminescent inhomogeneities in turbid media with spatially resolved measurements of CW diffuse luminescence emittance," *Appl. Opt.*, vol. 37, no. 13, pp. 2755–2765, 1998.

[45] M. Pfister and B. Scholz, "Localization of fluorescence spots with space-space music for mammographylike measurement systems," *J. Biomed. Opt.*, vol. 9, no. 3, pp. 481–487, 2004.

[46] Y. Chen, G. Zheng, Z. H. Zhang, D. Blessington, M. Zhang, H. Li, Q. Liu, L. Zhou, X. Intes, and S. Achilefu, "Metabolism-enhanced tumor localization by fluorescence imaging: *in vivo* animal studies," *Opt. Lett.*, vol. 28, no. 21, pp. 2070–2072, 2003.

[47] I. Gannot, A. Garashi, G. Gannot, V. Chernomordik, and A. Gandjbakhche, "*In vivo* quantitative three-dimensional localization of tumor labeled with exogenous specific fluorescence markers," *Appl. Opt.*, vol. 42, no. 16, pp. 3073–3080, 2003.

[48] Y. C. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*. IEEE, 1993, pp. 40–44.

[49] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[50] A. Brandt, *Multigrid Techniques: 1984 Guide, with Applications to Fluid Dynamics*. Sankt Augustin, Germany: GMD-Studien, 1984.

[51] R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise nanometer localization analysis for individual fluorescent probes," *Biophys. J.*, vol. 82, no. 5, pp. 2775–2783, 2002.
[52] X. Michalet, "Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium," *Phys. Rev. E*, vol. 82, no. 4, p. 041914, 2010.
[53] T. D. Lacoste, X. Michalet, F. Pinaud, D. S. Chemla, A. P. Alivisatos, and S. Weiss, "Ultrahigh-resolution multicolor colocalization of single fluorescent probes," *Proc. Natl. Acad. Sci.*, vol. 97, no. 17, pp. 9461–9466, 2000.
[54] F. Balzarotti, Y. Eilers, K. C. Gwosch, A. H. Gynnå, V. Westphal, F. D. Stefani, J. Elf, and S. W. Hell, "Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes," *Science*, vol. 355, no. 6325, pp. 606–612, 2017.
[55] M. G. L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microsc.*, vol. 198, no. 2, pp. 82–87, 2000.
[56] M. G. L. Gustafsson, "Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution," *Proc. Natl. Acad. Sci. USA*, vol. 102, no. 37, pp. 13 081–13 086, 2005.
[57] S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Opt. Lett.*, vol. 19, no. 11, pp. 780–782, 1994.
[58] M. Xu and L. V. Wang, "Universal back-projection algorithm for photoacoustic computed tomography," *Phys. Rev. E*, vol. 71, p. 016706, 2005.
[59] M. Cheney, D. Isaacson, and J. Newell, "Electrical impedance tomography," *SIAM Rev*, vol. 41, no. 1, pp. 85–101, 1999.
[60] R. G. Pratt, "Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model," *GEOPHYSICS*, vol. 64, no. 3, pp. 888–901, 1999.
[61] T. Rubk, P. M. Meaney, P. Meincke, and K. D. Paulsen, "Nonlinear microwave imaging for breast-cancer screening using gaussnewton's method and the cgls inversion algorithm," *IEEE Trans. Antennas Propag.*, vol. 55, no. 8, pp. 2320–2331, 2007.
[62] M. Scherg, "Functional imaging and localization of electromagnetic brain activity," *Brain Topogr.*, vol. 5, no. 2, pp. 103–111, 1992.
[63] R. Yasuda, E. A. Nimchinsky, V. Scheuss, T. A. Pologruto, T. G. Oertner, B. L. Sabatini, and K. Svoboda, "Imaging calcium concentration dynamics in small neuronal compartments," *Sci. STKE*, vol. 2004, no. 219, p. pl5, 2004.
[64] R. Prevedel, Y.-G. Yoon, M. Hoffmann, N. Pak, G. Wetzstein, S. Kato, T. Schrödel, R. Raskar, M. Zimmer, E. S. Boyden, and A. Vaziri, "Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy," *Nat. Meth.*, vol. 11, no. 7, pp. 727–730, 2014.
[65] A. T. Eggebrecht, S. L. Ferradal, A. Robichaux-Viehoever, M. S. Hassanpour, H. Dehghani, A. Z. Snyder, T. Hershey, and J. P. Culver, "Mapping distributed brain function and networks with diffuse optical tomography," *Nat. Photon.*, vol. 8, no. 6, pp. 448–454, 2014.
[66] M. L. Castanares, V. Gautam, J. Drury, H. Bachor, and V. R. Daria, "Efficient multi-site two-photon functional imaging of neuronal circuits," *Biomed. Opt. Express*, vol. 7, no. 12, pp. 5325–5334, 2016.
[67] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.
[68] K. T. Shimizu, R. G. Neuhauser, C. A. Leatherdale, S. A. Empedocles, W. Woo, and M. G. Bawendi, "Blinking statistics in single semiconductor nanocrystal quantum dots," *Phys. Rev. B*, vol. 63, no. 20, p. 205316, 2001.
[69] S. Hohng and T. Ha, "Near-complete suppression of quantum dot blinking in ambient conditions," *J. Am. Chem. Soc*, vol. 126, no. 5, pp. 1324–1325, 2004.
[70] W. C. W. Chan and S. Nie, "Quantum dot bioconjugates for ultrasensitive nonisotopic detection," *Science*, vol. 281, no. 5385, pp. 2016–2018, 1998.
[71] R. Khatchadourian, A. Bachir, S. J Clarke, C. D Heyes, P. Wiseman, and J. Nadeau, "Fluorescence intensity and intermittency as tools for following dopamine bioconjugate processing in living cells," *J Biomed Biotechnol.*, vol. 2007, p. 10, 02 2007.
[72] F. Martelli, S. Del Bianco, A. Ismaelli, and G. Zaccanti, *Light Propagation through Biological Tissue and Other Diffusive Media: Theory*. SPIE Press, Bellingham, 2010.
[73] I. Mel'Nikova and V. Mikhaylov, "Spectral scattering and absorption coefficients in strati derived from aircraft measurements," *J. Atmos. Sci.*, vol. 51, no. 7, pp. 925–931, 1994.

PLACE PHOTO HERE

Brian Z. Bentz received his B.S. and Ph.D. in 2011 and 2017, respectively, from Purdue University, West Lafayette, Indiana, both in electrical and computer engineering. He was a Research Assistant with the School of Electrical and Computer Engineering at Purdue University from 2012 to 2017. In 2017, he joined Sandia National Laboratories, Albuquerque, New Mexico.

PLACE PHOTO HERE

Dergan Lin received his M.S. in 2012 from Purdue University in electrical engineering and and is currently working towards a Ph.D. at Purdue University.

PLACE PHOTO HERE

Justin A. Patel received his B.S. in 2016 from the University of California, Berkeley in electrical engineering and computer science. He currently studies at Purdue University in West Lafayette, Indiana, where he is working towards a Ph.D. in electrical and computer engineering, and works there as a Research Assistant with the School of Electrical and Computer Engineering.

PLACE PHOTO HERE

Kevin J. Webb is a professor in the School of Electrical an Computer Engineering at Purdue University, West Lafayette, Indiana, where he has been on the faculty since 1990. Prior to this, he was on the faculty of the University of Maryland, College Park, following completion of his Ph.D. at the University of Illinois, Urbana-Champaign, in 1984. He is a Fellow of the IEEE, the OSA, and APS.