Biomedical Optical Imaging[1]

Charles A. Bouman
School of Electrical and Computer Engineering
School of Biomedical Engineering
Purdue University
(765) 494-0340
bouman@purdue.edu
www.ece.purdue.edu/∼bouman

Kevin J. Webb
School of Electrical and Computer Engineering
Purdue University
(765) 494-3373
webb@purdue.edu
www.ece.purdue.edu/∼webb

Tutorial Presented at:
2004 IEEE International Symposium on Biomedical Imaging
April 15, 2004
Washington, D.C.

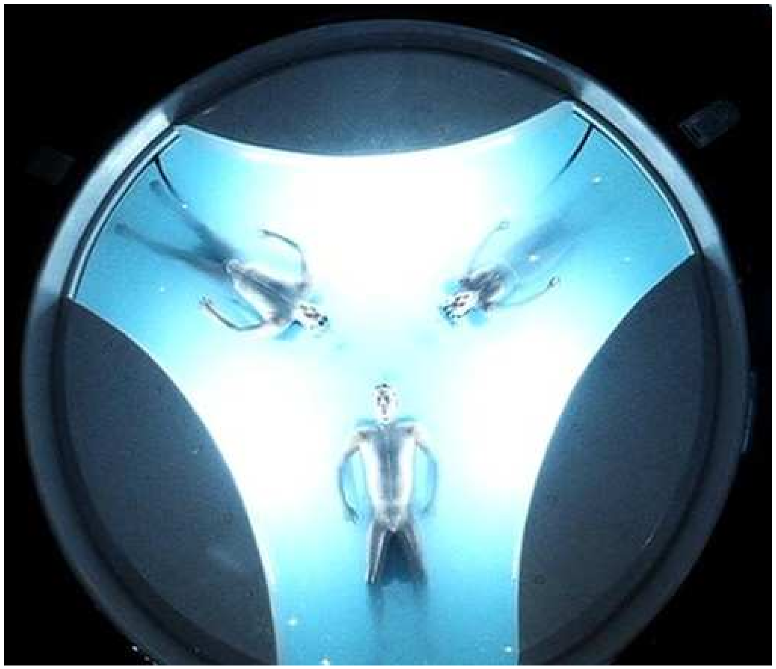
With contributions from:
Adam Milstein and Seungseok Oh
School of Electrical and Computer Engineering
Purdue University

[1]Special thanks to Dr. John Cozzens and National Science Foundation for supporting this work under contract CCR-0073357.

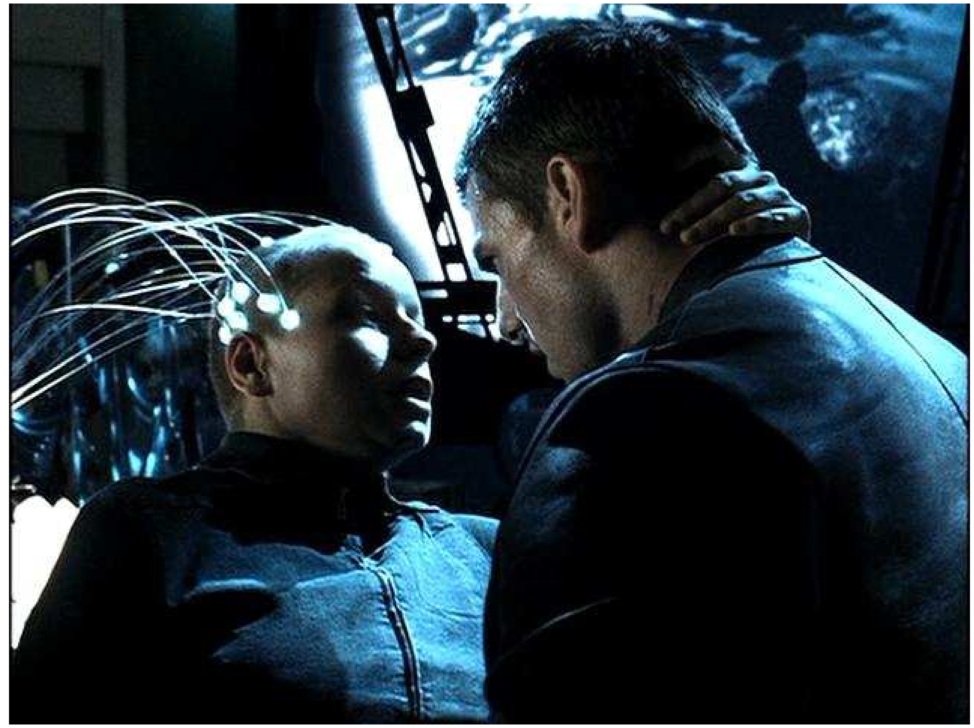
Outline

1. Sensing with Light
   (a) Fluorescence
   (b) Spectroscopy
   (c) Polarization
   (d) Optical Coherence Tomography
   (e) Optical Diffusion Tomography
   (f) ODT Systems
   (g) Applications
2. Forward Model
   (a) ODT Forward Model
   (b) Fréchet Derivative
   (c) Adjoint Differentiation
3. Inversion of Forward Model
   (a) Bayesian Framework
   (b) Optimization Methods
   (c) Linearized Approach
   (d) Source-Detector calibration
   (e) Shape-Based Reconstruction
4. Fluorescence ODT
   (a) Forward Model
   (b) Multifrequency inversion
   (c) Design Metrics
5. Molecular Imaging
   (a) Folate-Targeted Fluorescent Agents
   (b) Kinetic Imaging
6. Optical Speckle Imaging
7. Future Directions

Seeing Inside the Body with Light?

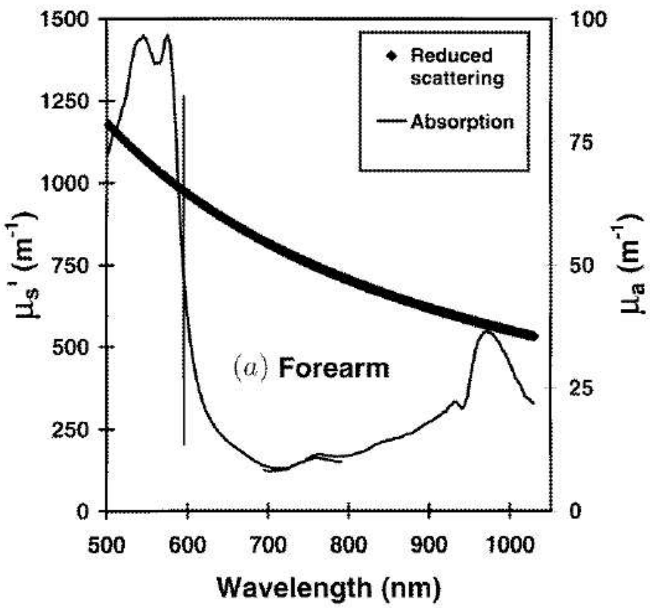


Minority Report, Twentieth Century Fox, 2002.



Wally: "We scan, by way of *optical tomography*, white light pinpoints pulse along the entire length of the headgear, and re-read after absorption through their brain tissue."

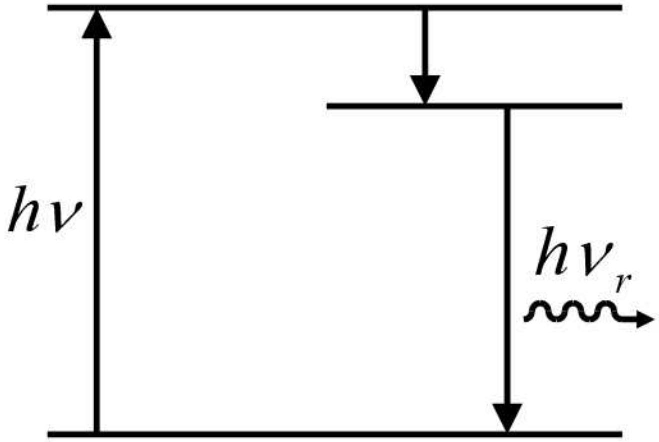
Transparency of Tissue in Near-IR Range [1]



(Reproduced from [1] R. M. P. Doornbos *et al.*, *Phys. Med. Biol.*, 1999.)

- Hemoglobin and water have relatively low absorption in near-IR
- Near-IR "window" enables optical imaging and near-infrared spectroscopy [2]

Fluorescence



- Fluorescence results from radiative decay from an excited state
- Indocyanine green (ICG): pump at 780 nm and emits at 830 nm
- Basic concept from two-level rate equation analysis [3]

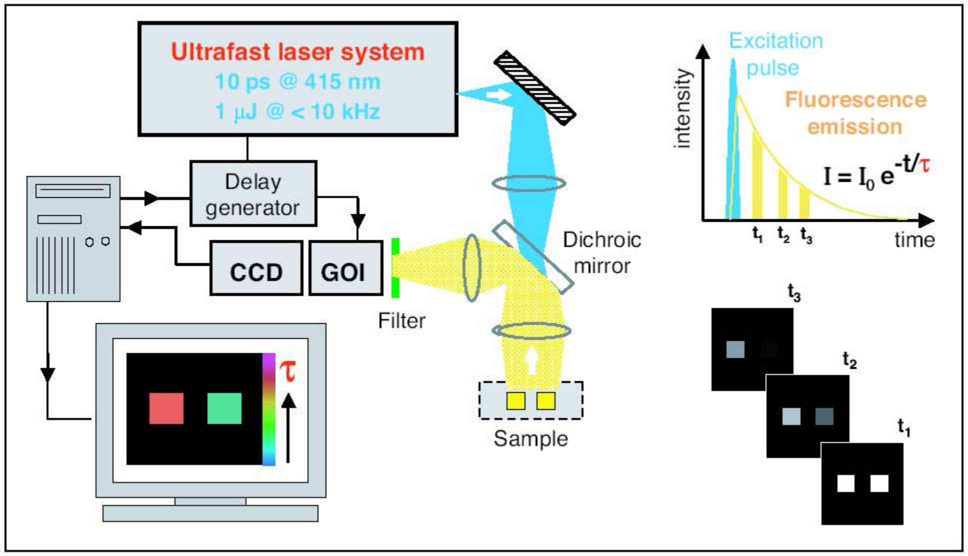
Fluorescence: Two-Level Rate Equation[3]

- Fixed total population $\rightarrow$ one rate equation
- Optical frequency transition $\rightarrow$ spontaneous emission rates dominate thermally stimulated rates

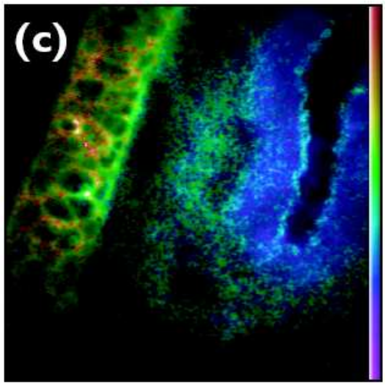
$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\Delta N(t) \;=\; -2W_{12}\Delta N(t) - \left[\frac{\Delta N(t) - \Delta N_0}{\tau_{21}}\right]}$$

- $-\left[\frac{\Delta N(t)-\Delta N_0}{\tau_{21}}\right]$ causes population difference to relax to thermal equilibrium (ΔN_0) with time constant τ_{21}
- Stimulated signal term $-2W_{12}\Delta N(t)$ acts to drive $\Delta N(t) \rightarrow 0$ (saturates the population difference)
- W_{12} proportional to strength of applied signal

Fluorescence Lifetime Imaging Microscopy (FLIM) [4, 5]



FLIM experiment



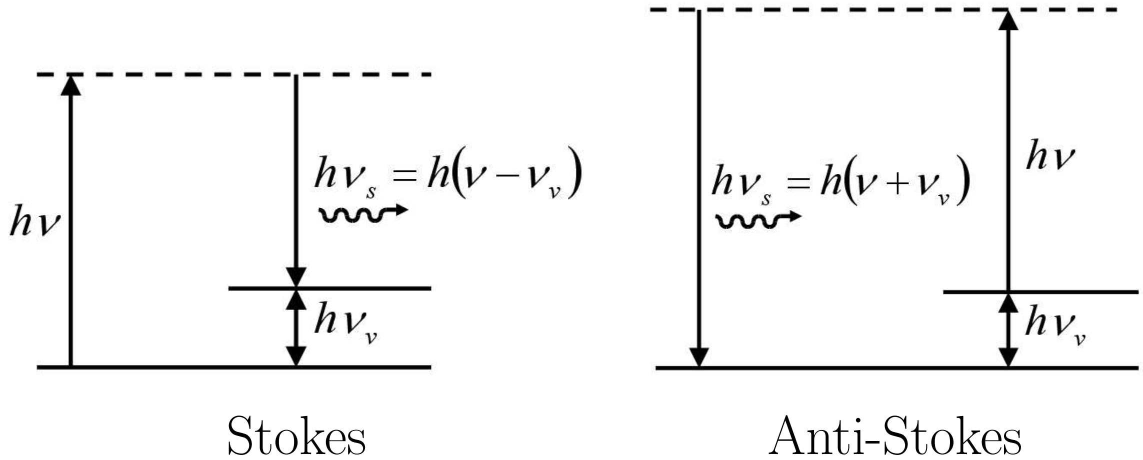
FLIM image of rat ear autofluorescence

- Time- or frequency-resolved fluorescence is recorded, and decay rate is represented as an image
- Tunable mode-locked laser and gated image intensifier can be used
- Fluorescent lifetime may provide information about tissue [6]

(Reproduced from [4] Paul French group, *Opt. Phot. News*, 2002)

Raman Spectroscopy: Basics

- Raman scatter results from coupling to molecular vibrations [7]
- Molecule and environment $\longrightarrow$ Raman signal (if any)

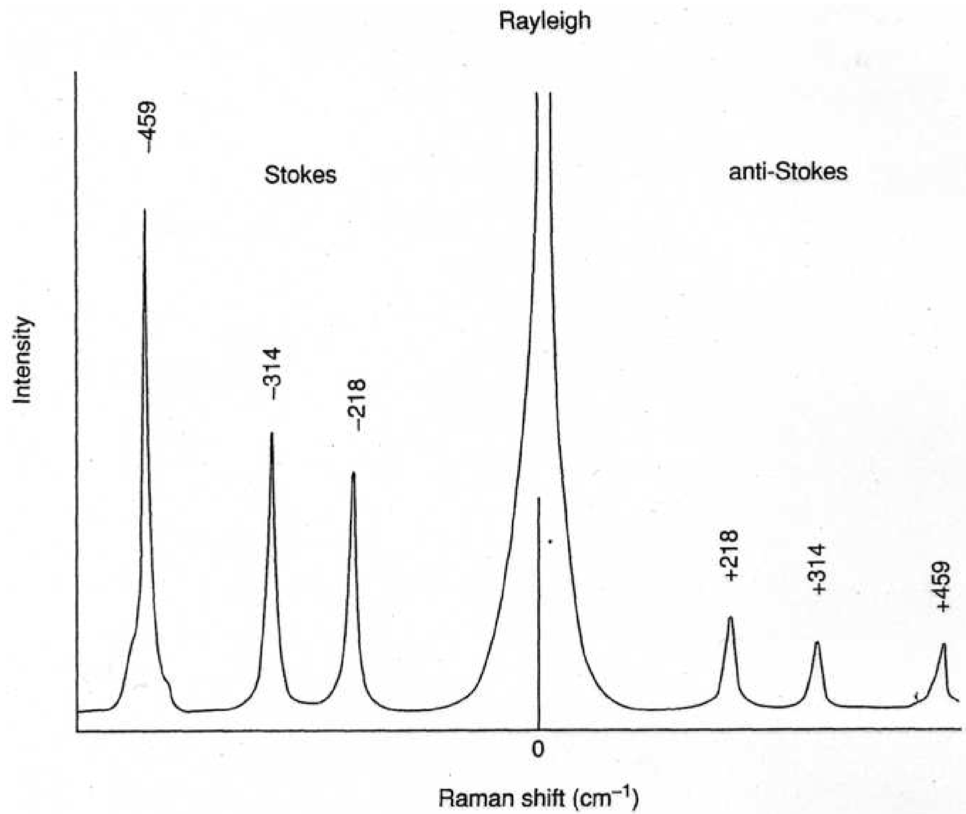


- Wave number: $\tilde{\nu} = \nu/c$
- $\lambda = 1\ \mu$m, $\Delta\tilde{\nu} = 100$ cm^{-1} [7] $\longrightarrow \Delta\lambda = 10$ nm, $\Delta\nu = 3$ THz

C. V. Raman and K. S. Krishnan, *Nature*, **121**, 501 (1928)

Raman Spectroscopy: Example

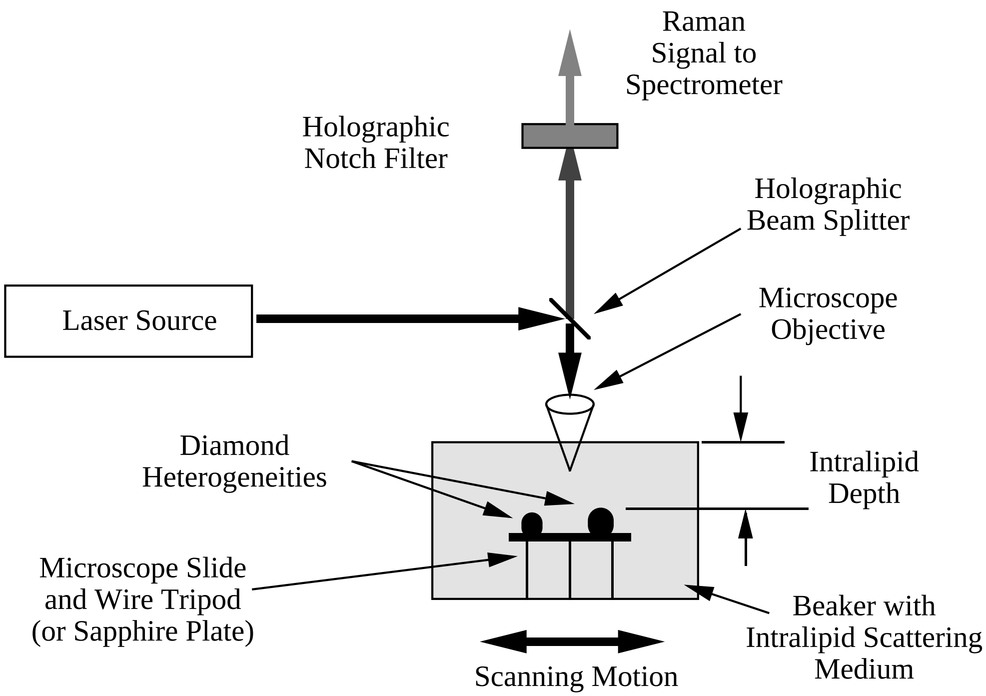
- Example Raman spectrum for CCl_4 with 488 nm excitation[8]



J. R. Ferraro, K. Nakamoto and C. W. Brown, *Introduction to Raman Spectroscopy*, Academic Press, 2003.

Raman Spectroscopy: Imaging

- Scan detector and measure counts in a specific Raman line [9]



- Glucose? [10]

Raman Spectroscopy: Surface-Enhanced Raman Scattering

- Scattering cross-section $\sigma = \Phi/W$ (Φ is scattered power and W is incident power)
- Single molecule $\sigma_{\text{Raman}} \sim 10^{-24}$ cm^2
- Single molecule fluorescence possible with $\sigma_{\text{fluor}} \sim 10^{-14} \rightarrow 10^{-10}$ cm^2 enhancement required for single molecule Raman spectroscopy
- Surface enhanced Raman scattering (SERS) may hold the promise to achieve this enhancement [11]

M. Moskovits, "Surface-enhanced spectroscopy," *Reviews of Modern Physics*, vol. 57, no. 3, pp. 783-826, July 1985.

Raman Spectroscopy: Surface-Enhanced Raman Scattering

- Excitation and Stokes field

$$E(t) = A_L e^{j\omega_L t} + A_S e^{j\omega_S t}$$

- Solution of vibrational harmonic equation $\longrightarrow$ dipole moment per Raman oscillator [12, 13]

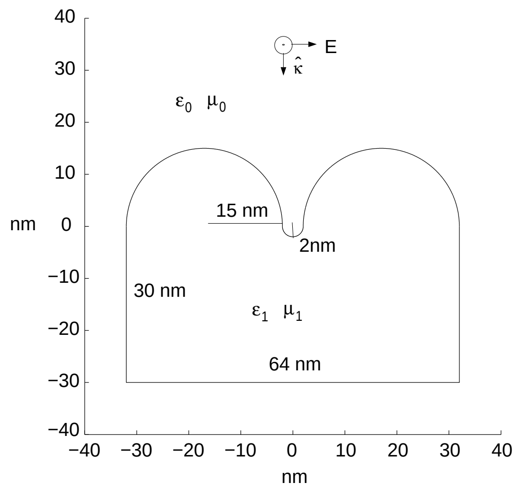
$$\boxed{\begin{aligned} \mathbf{p} &= \left(\frac{\partial\alpha}{\partial q}\right) q\mathbf{E} \\ &= \hat{e}\left(\frac{\partial\alpha}{\partial q}\right) a A_L A_S^* \cos[(\omega_L - \omega_S)t]\, \cos[\omega_L t] \end{aligned}}$$

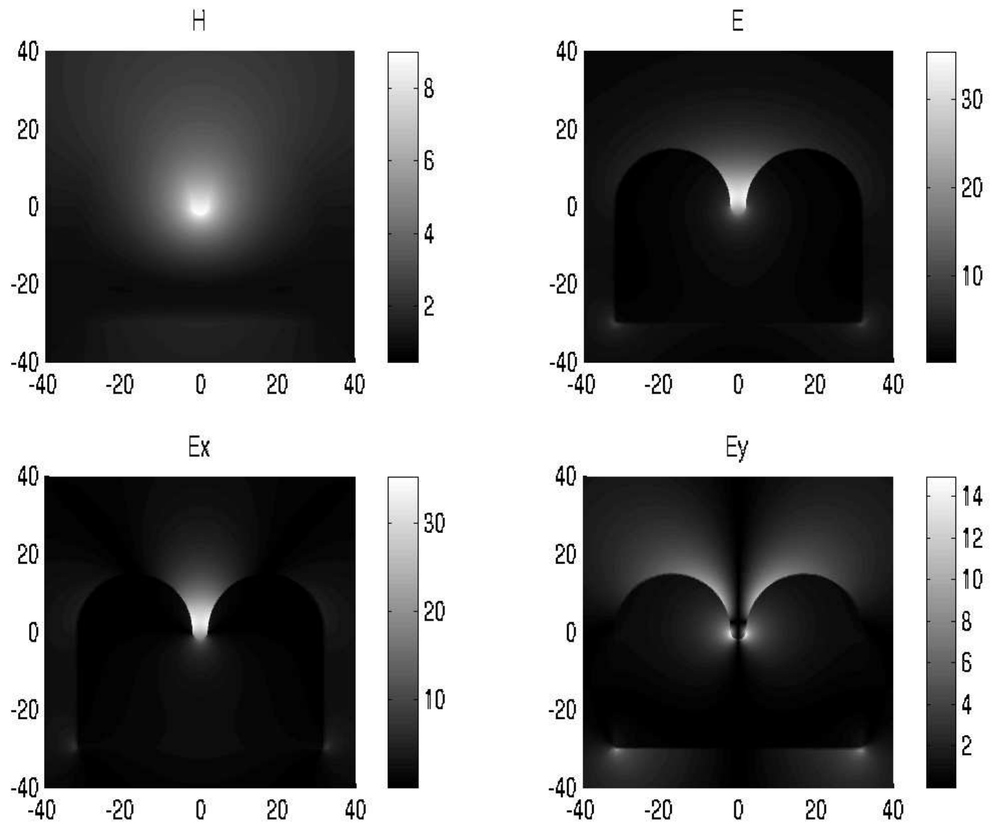
[12] D. A. Long, "Raman Spectroscopy," McGraw Hill, 1977.

[13] R. W. Boyd, "Nonlinear Optics," Academic Press, 1992.

Raman Spectroscopy: Surface-Enhanced Raman Scattering: Simulation

- Ag nano particle with $\lambda = 459.42$ nm ($\epsilon_r = -7.55 - j0.24$) [14]

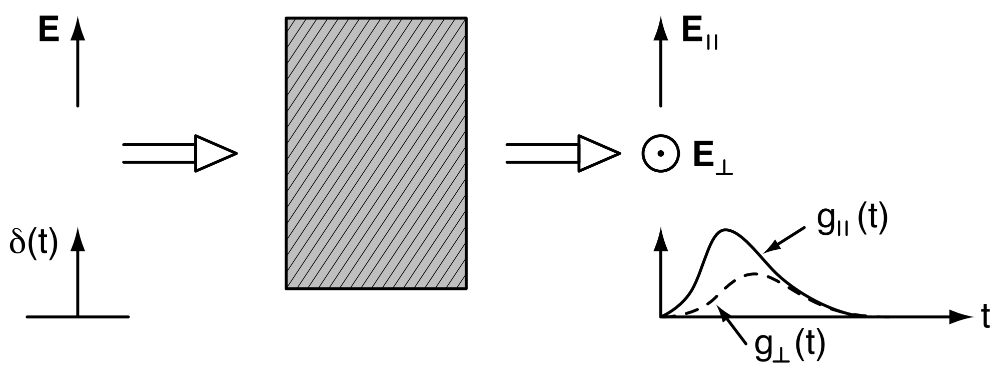


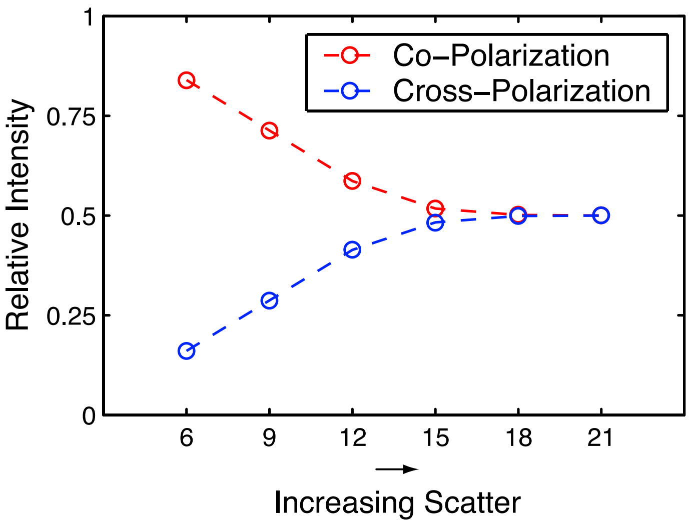


J. Li and K. J. Webb, unpublished.

Polarization

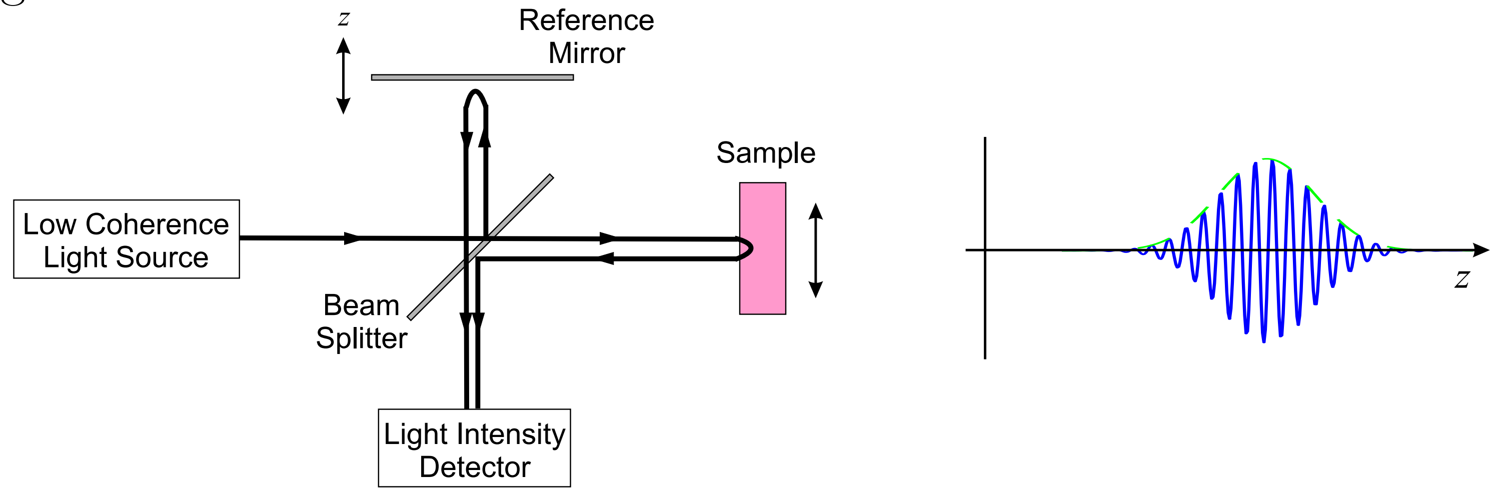
- Response for different $\mathbf{E}$-orientation $\longrightarrow$ polarization information
- As scatter increases, information in co-pol and cross-pol light becomes identical





Optical Coherence Tomography (OCT) [15]

- 2-D or 3-D image is made by using interferometric measurement of optical backreflection or backscattering from internal tissue microstructures
- Based on low coherence optical interferometry
- Negligible scatter assumed



  – z-direction moving of reference mirror ⟶ longitudinal scan
  – Beam moving on sample ⟶ transverse scan
  * Usually implemented with fiber optic

(Fujimoto Group, *Science*, vol. 254, pp. 1178-1181, 1991)

OCT: Example Image [16]

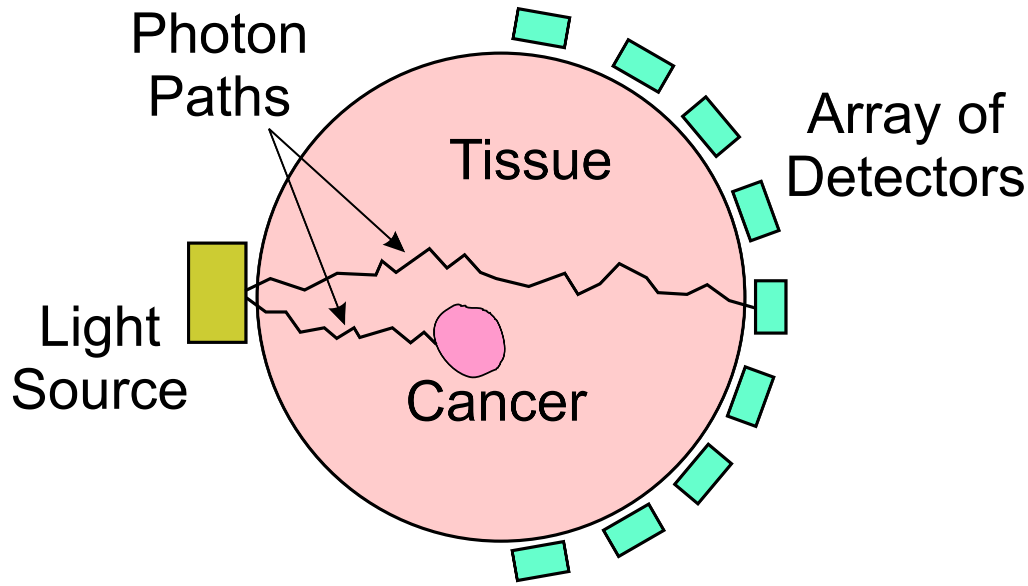
- Cellular-level image of a living African frog (Xenopus laevis) tadpole

  - *In vivo* subcellular level resolution: 1 μm $\times$ 3 μm (longitudinal $\times$ transverse)

(Reproduced from [16] W. Drexler *et al.* in Fujimoto Group, *Opt. Lett.*, vol. 24, no. 17, pp. 1221-1223, 1999)

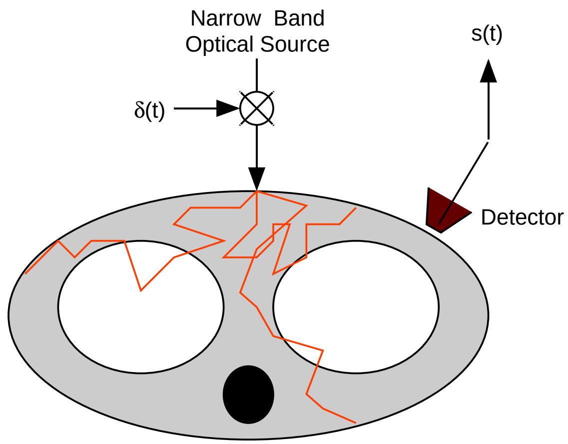
Optical Diffusion Tomography (ODT) [17]

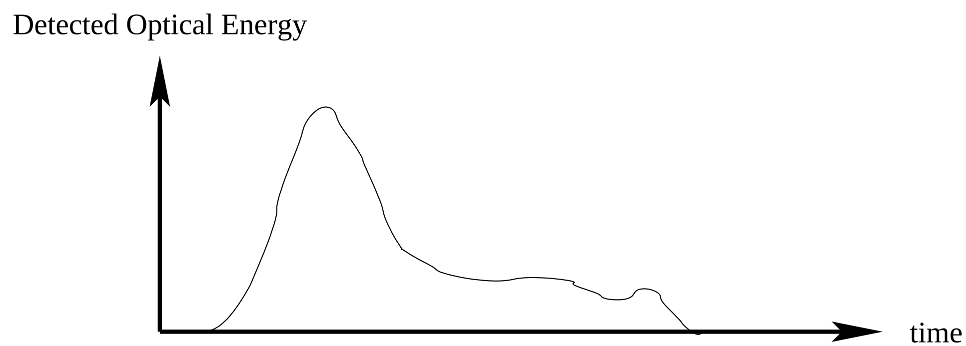
- Measure light that passes through a highly scattering medium
- Determine unknown absorption and/or diffusion cross-section of material



- With K sources and M detectors, there can be KM measurements
- Time domain: Measure delay of light pulse at detector
- Frequency domain: Measure amplitude/phase of modulated light envelope
- Also called "Diffuse Optical Tomography" and "Photon Migration"

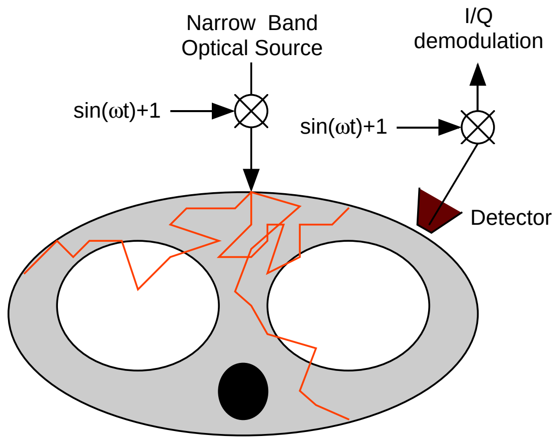
ODT: Time Domain Measurements [18, 19]

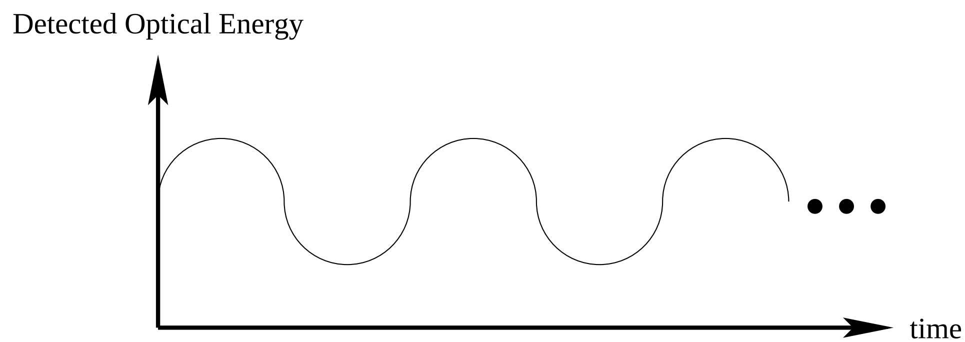




- Short pulse of light at optical source input
- Light travels to detector along different paths due to scattering
- Measure time domain response, $s(t)$, at optical detector

ODT: Frequency Domain Measurements

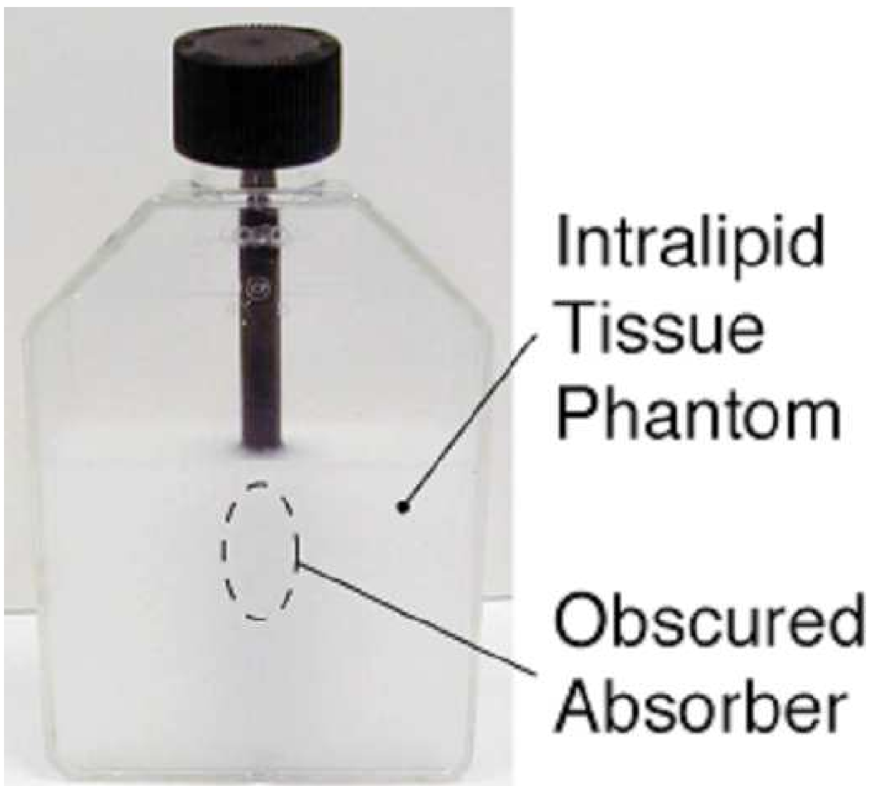




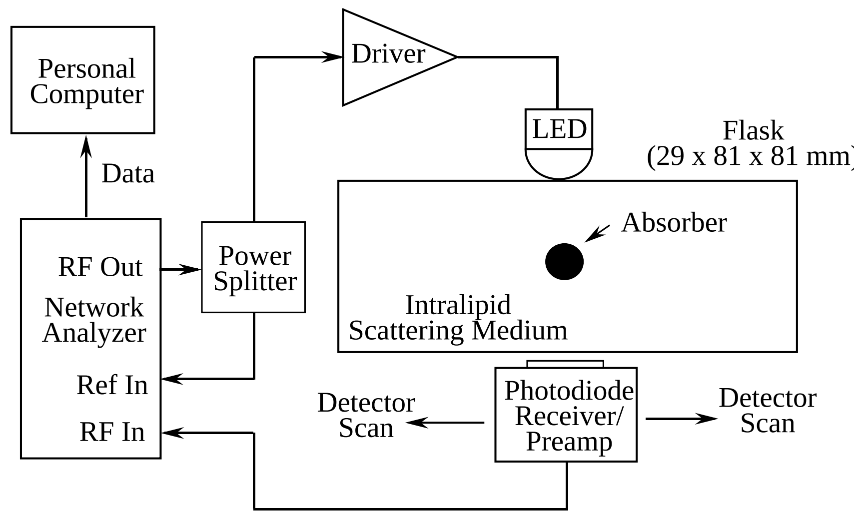
- Modulate light amplitude using RF source at frequency ω
- Measure magnitude and phase of optical detector's signal
- Scattering and absorption change magnitude and phase of detected signal
- Each measurement is a complex number
- Special case: If $\omega = 0$, this is known as continuous wave (CW) measurement

ODT Example[20, 21]: Experiment

- Intralipid solution (0.4%) solution in a 2.9 cm $\times$ 8.1 cm $\times$ 8.1 cm flask containing a 0.7 cm black plastic cylinder.



- Experimental apparatus
  - Near infrared LED (890nm) modulated at 10, 46, and 81 MHz
  - 2 light source positions (front and back) each with 25 detector locations

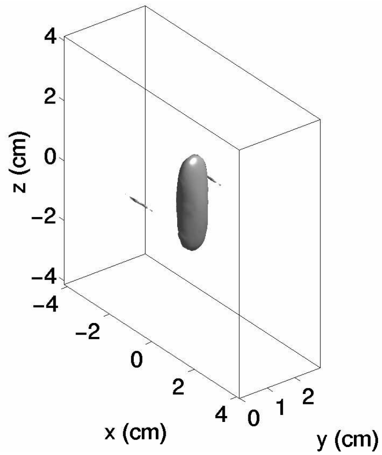


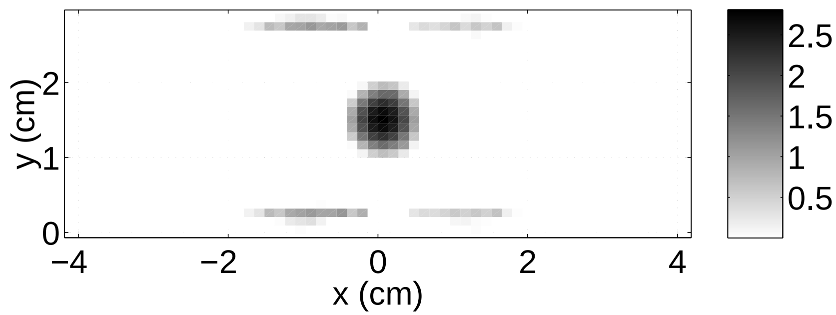
ODT: Image Reconstruction

- How do you reconstruct an Image?
- It is not ...
  - Obvious
  - Easy
  - Linear
  - Filtered back projection
- But, it can be done...

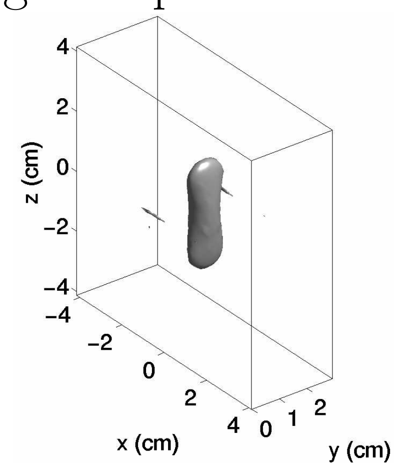
ODT Example[22]: 3-D Reconstructions

- All frequencies

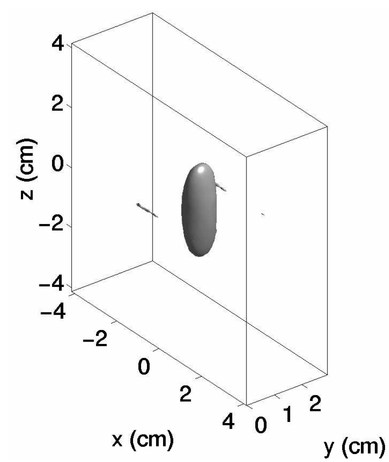




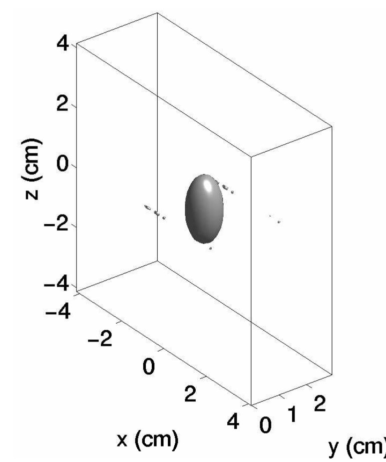
- Single frequencies



10 MHz



46 MHz

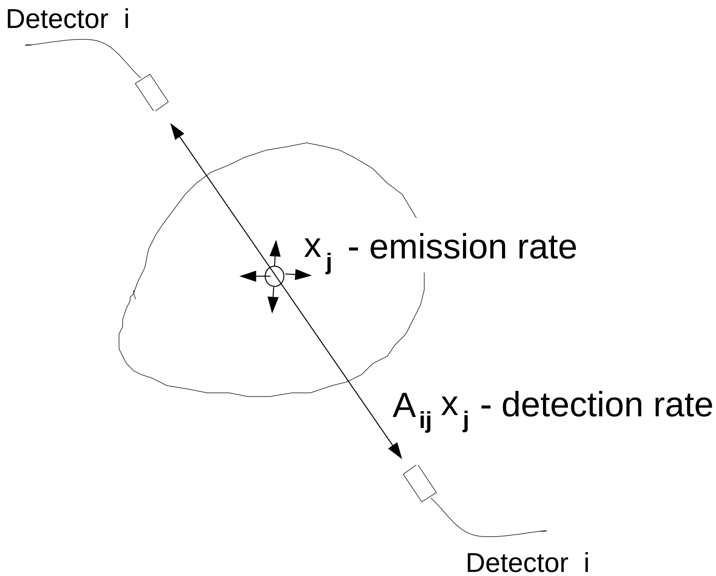


81 MHz

Milstein, Oh, Reynolds, Webb, Bouman, and Millane, *Optics Letters*, vol. 27, Jan. 2002.

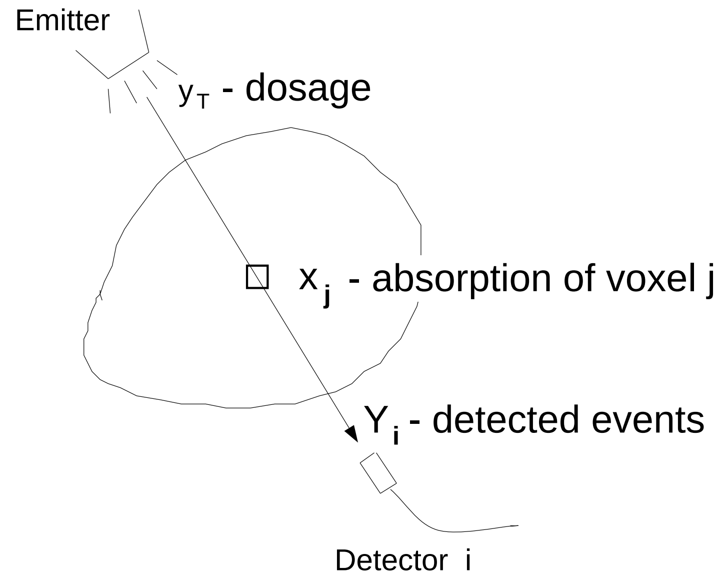
ODT: A Look at Conventional Tomography

Emission Problem



$$E[y_i] = A_{i*}x$$

Transmission Problem



$$E[y_i] = e^{-A_{i*}x} y_T$$

$$-\log\left(\frac{E[y_i]}{y_T}\right) = A_{i*}x$$

- Photons travel in a straight line
- Measurements are linearly related to unknown x
- Reconstruction is essentially matrix inversion

ODT: Contrasting ODT and Conventional Tomography

- The imaged medium determines the photons:
  - Path length distribution
  - Attenuation
  - Delay (phase)
- Forward problem:
  - Nonlinear
  - 3-dimensional
  - Modeled by partial differential equation (PDE)
  - Number of voxels >> number of measurements
- Inverse problem:
  - Filtered back projection is inappropriate
  - Nonlinear
  - Computationally expensive

ODT: Specificity

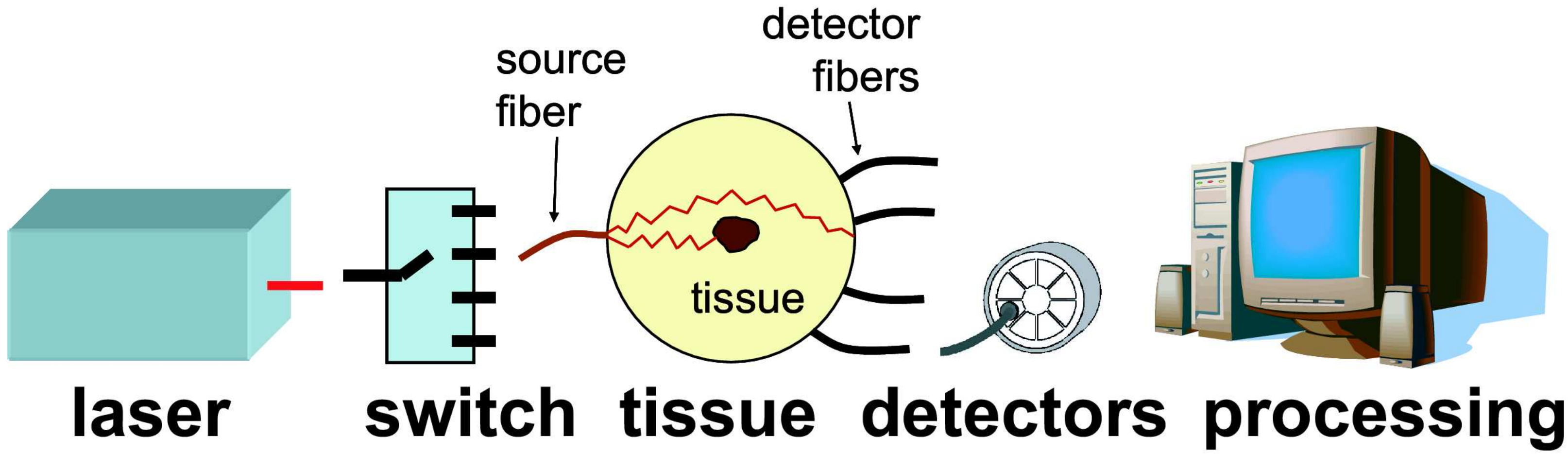
- Volumetric spectroscopy: Measure the optical properties of each voxel at each spectral wavelength
- Chemical specificity: Determine specific chemical properties through techniques such as Raman spectroscopy
- Fluorescent agents: Use near IR or visible fluorescent tracers to enhance contrast
- Fluorescent life-time imaging: Time-decay properties of fluorescent signal provide information
- Functional imaging: Measure dynamic properties of tracer uptake for applications such as pharmacokinetics
- Molecular imaging: Develop fluorescent tracers that are designed to probe specific molecular properties and target particular tissues (e.g., tumors)

ODT: Other Potential Advantages

- Safety: No radioactive exposure
- Portable: Does not require large magnet or gamma camera
- Fast: Can potentially achieve high frame rates compared to PET or MRI
- Inexpensive: Does not require coherent light

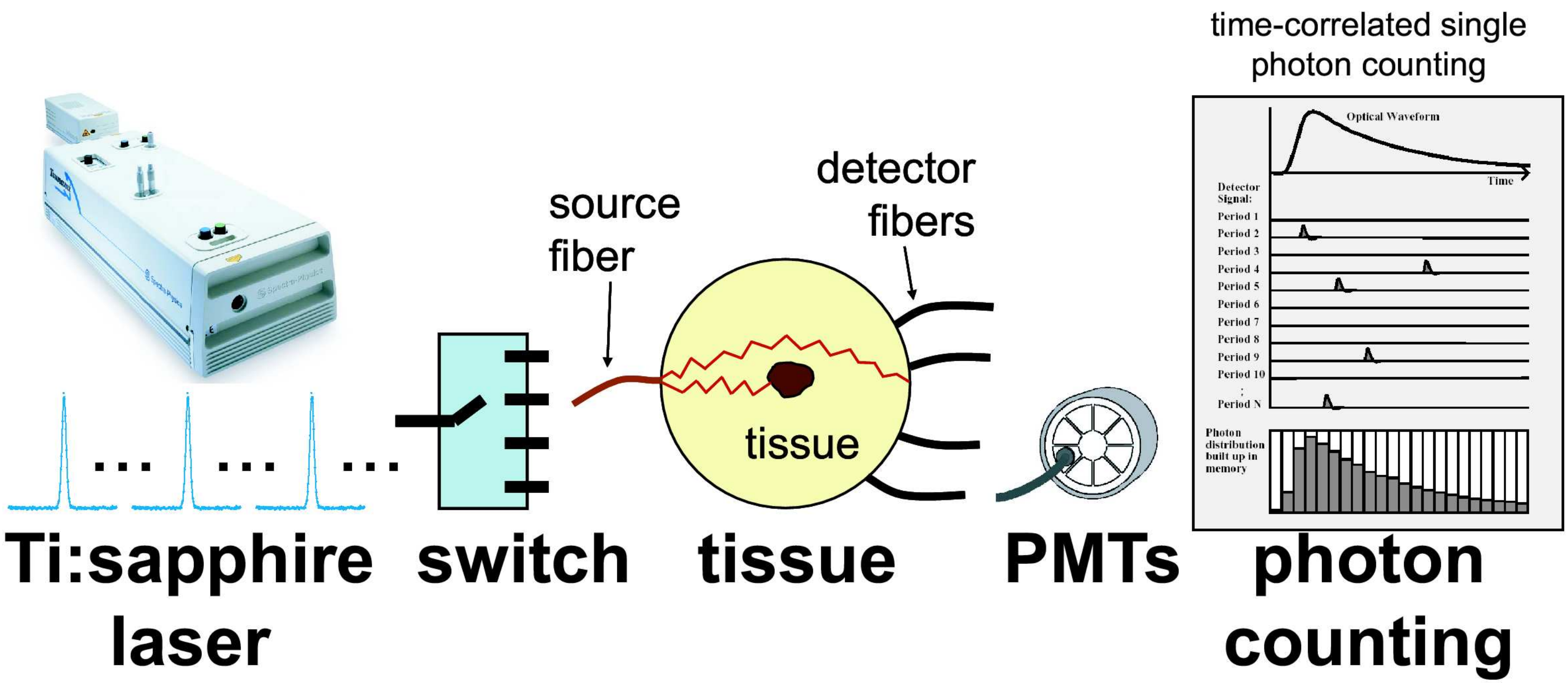
- Spatial resolution
  - Needs to be high enough to separate tissues of interest
  - Does not have resolution of CT or MRI
  - Seems reasonable to expect $\approx$1mm resolution $\approx$ 1-10 cm depth

ODT Systems: General Components



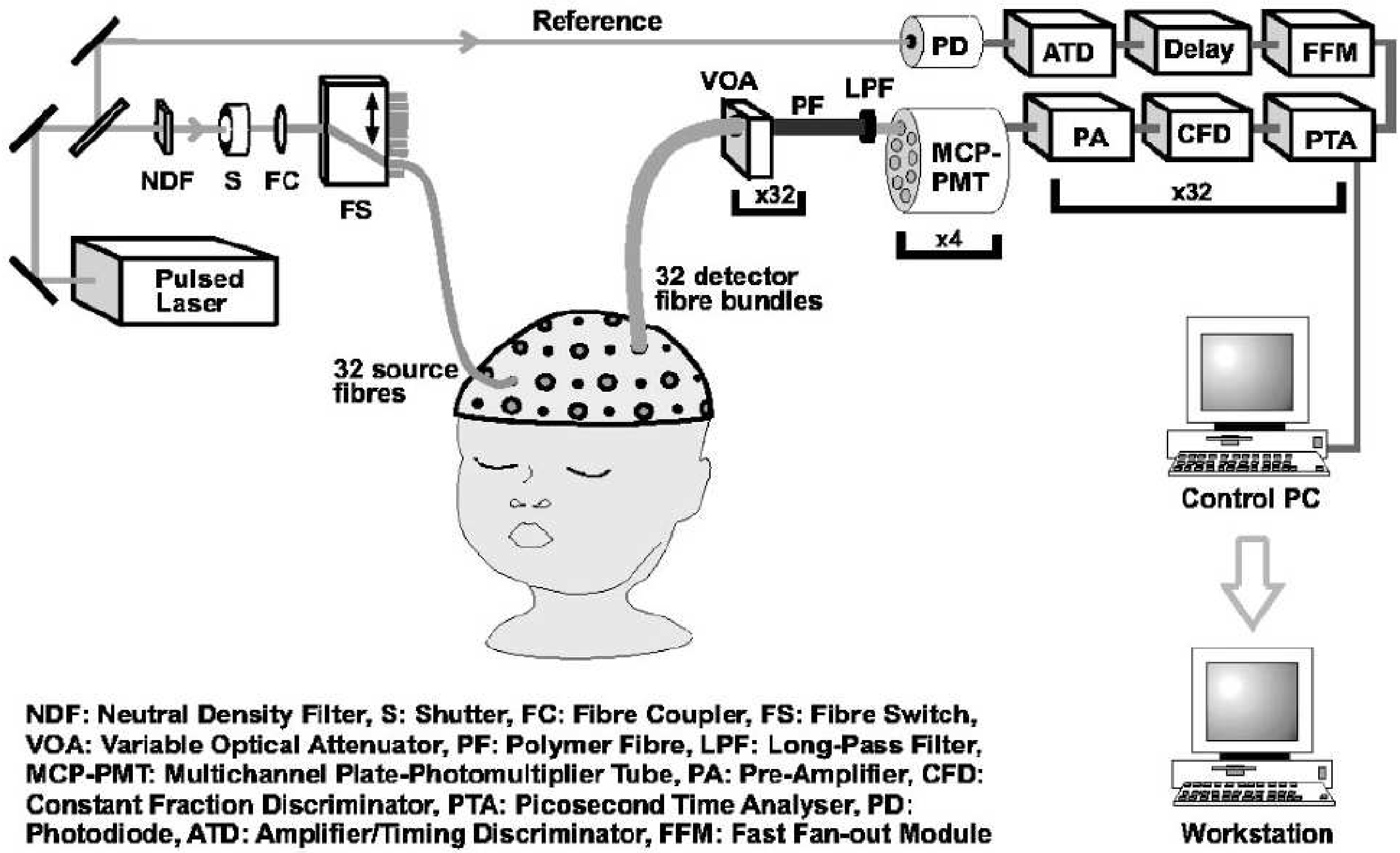
- Sources: solid state lasers, diode lasers, LEDs
- Detectors: photomultiplier tubes (PMTs), avalanche photodiodes (APDs), photodiodes (PIN), and intensified CCD cameras
- Cost and performance trade-offs (\$50 - \$250)

ODT Systems: Time Domain Overview [23, 24]



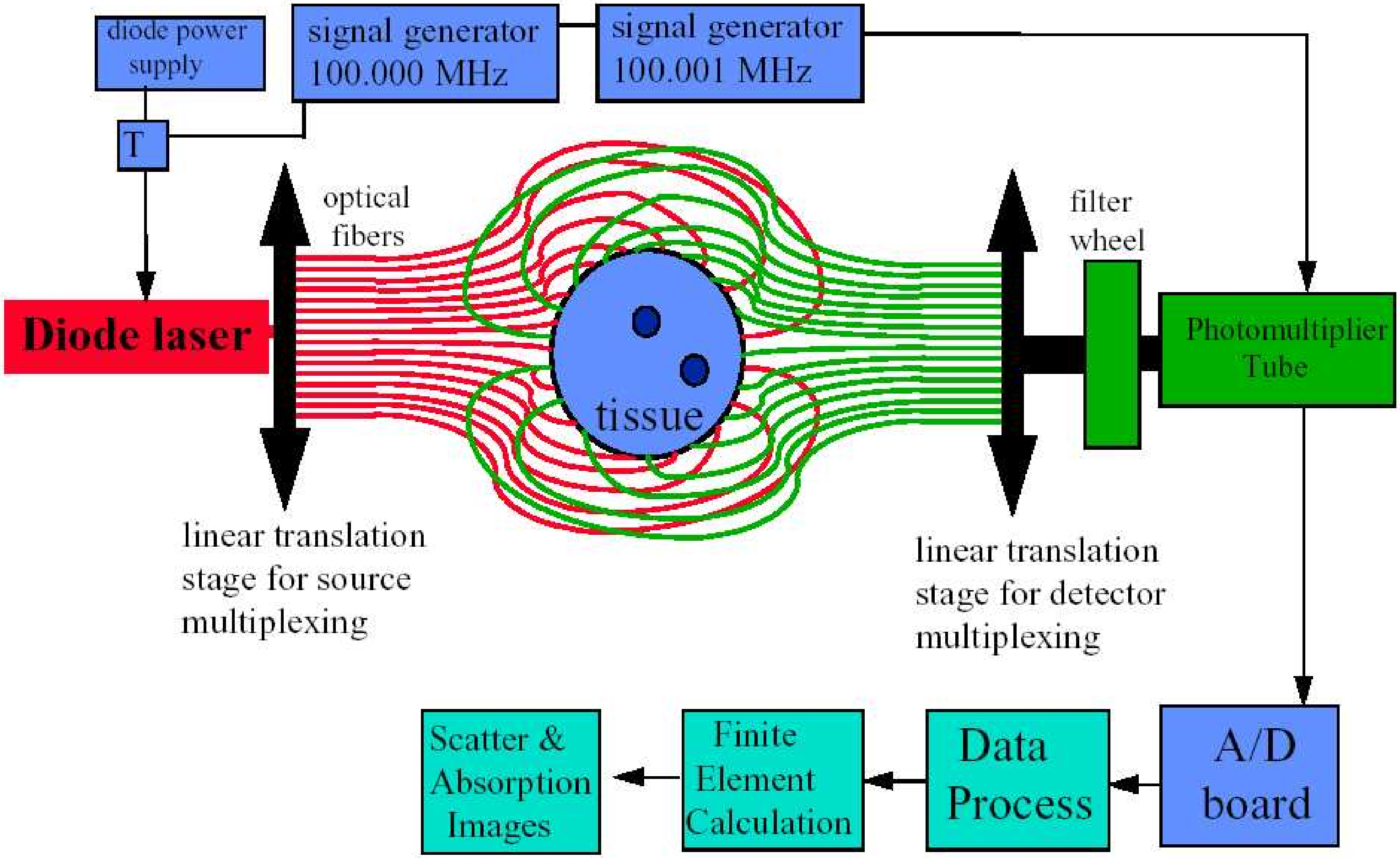
- Ti:sapphire laser pulsed with 80 MHz repetition rate
- Time-correlated single photon counting: individual photons are counted for delay times relative to trigger pulses

ODT Systems: Time Domain: UCL Imager [23]



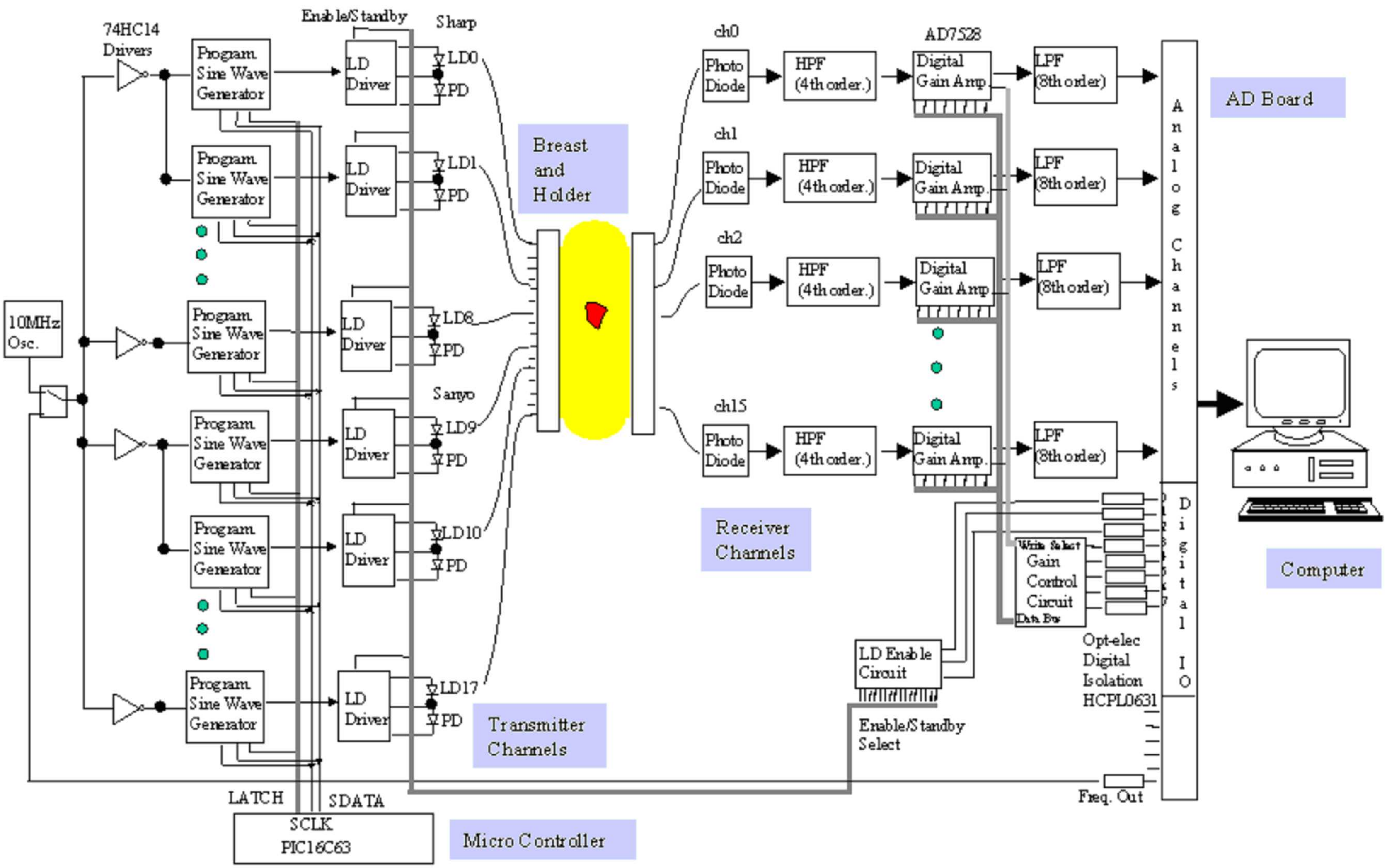
(Reproduced from Schmidt *et al.*, *Rev. Sci. Inst.*, 2000)

ODT Systems: Frequency Domain: Dartmouth Imager [25]



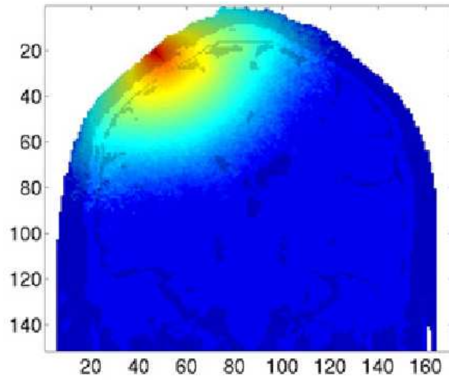
(Reproduced from Pogue *et al.*, *Opt. Exp.*, 1997)

ODT Systems: Continuous-Wave (CW): MGH Imager [26]

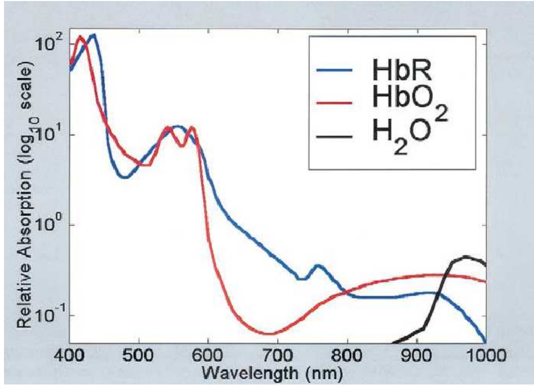


(Reproduced from Zhang *et al.*, *Proc. SPIE*, 2001)

Applications: Functional Brain Imaging [27, 28]



Simulation of light entering brain[28]

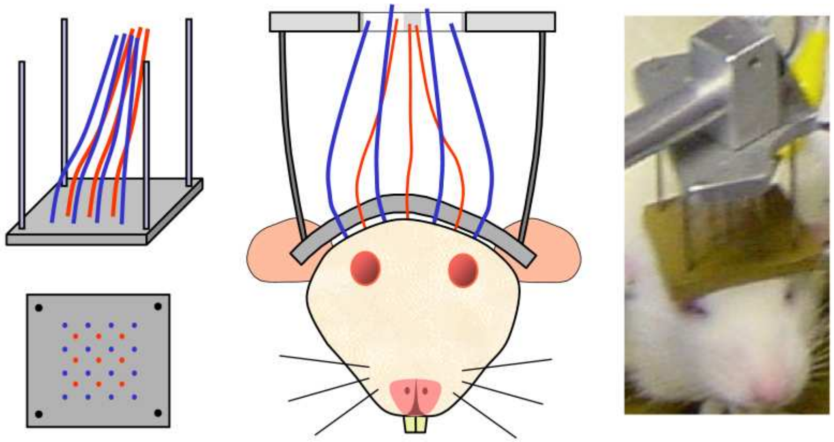


HbO_2 and HbR absorption[27]

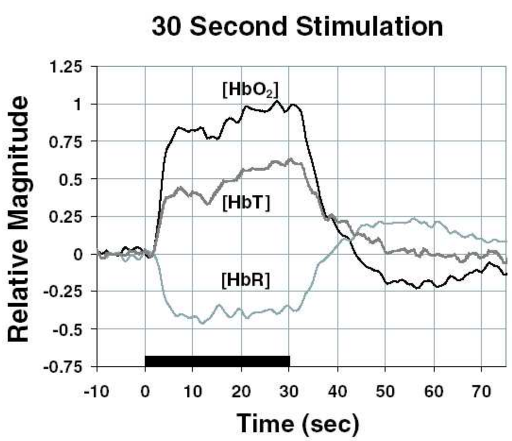
- NIR light can penetrate through the human skull into the brain
- Local changes in oxyhemoglobin ($[HbO_2]$) and deoxyhemoglobin ([HbR]) concentration indicate brain activity
- ODT can quantify $[HbO_2]$ and [HbR] in the brain due to absorption at different optical wavelengths
- Optical methods can have higher temporal resolution than BOLD fMRI, but lower spatial resolution

(Reproduced from [27] Strangman *et al.*, *Biol. Psych.*, 2002, and [28] Boas *et al.*, *Opt. Exp.*, 2002)

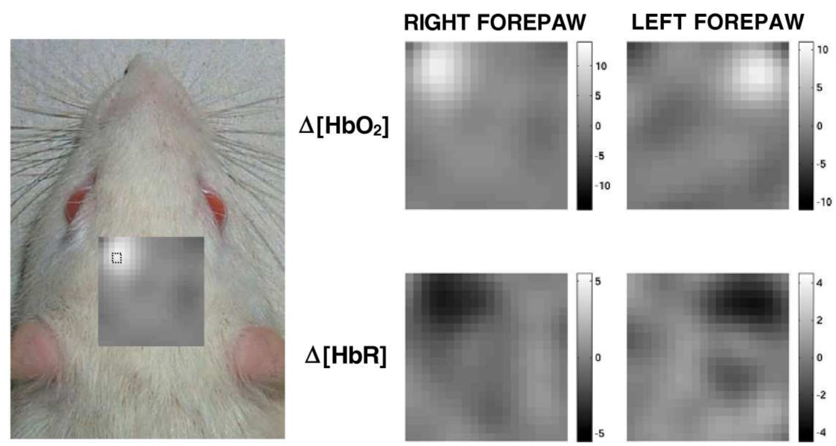
Applications: Hemodynamics in Rat Subject[29]



Fiber optic probe measures brain while forepaws electrically stimulated



Hemodynamic response

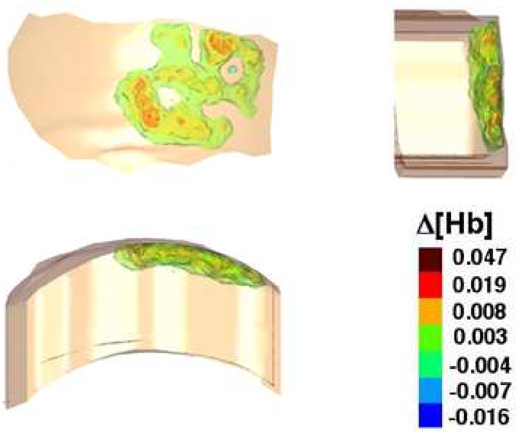


Δ[HbO$_2$] and Δ[HbR] images

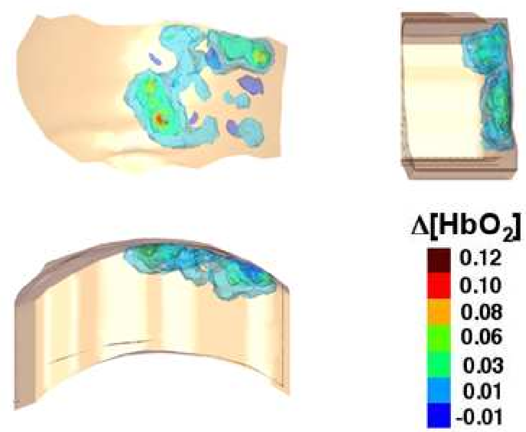
(Reproduced from Siegel *et al.*, *Phys. Med. Biol.*, 2003)

Applications: 3D Hemodynamics in Humans[30]

Commercial instrument[31, 32] measures brain during breathing exercise



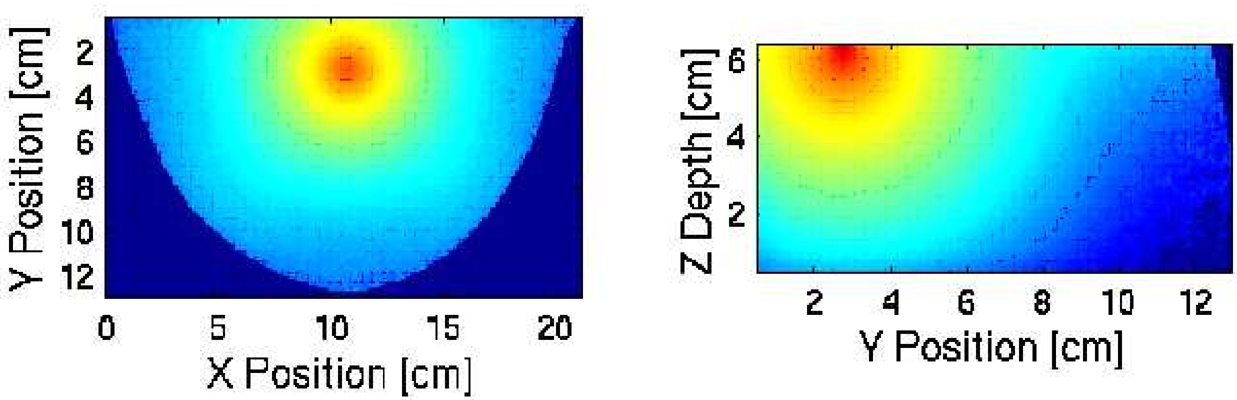
$\Delta([HbO_2]+[HbR])$ images



$\Delta[HbO_2]$ images

(Reproduced from Bluestone *et al.*, *Opt. Exp.*, 2001, and NIRx Medical Technologies website)

Applications: Breast Imaging



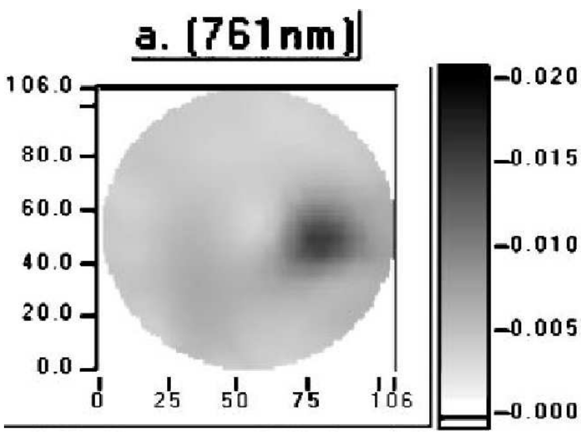
Simulation of light in breast [33]

- Currently, X-ray mammography detects structural changes in tumors compared to surrounding tissue
- Breast tumors tend to have higher absorption than surrounding tissue due to increased vascular density
- Optical methods potentially will offer earlier diagnosis by observing changes in absorption before structural changes take place

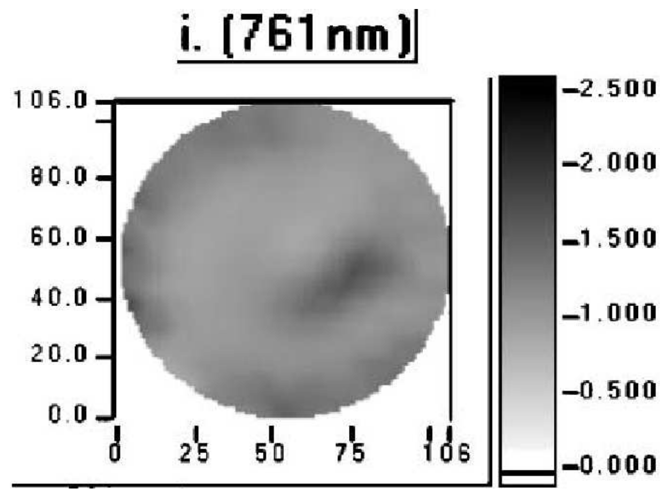
(Reproduced from Stott, MGH presentation, 2002.)

Applications: Breast Tumor Measurements [34]

Measurement table and instrument used at Dartmouth



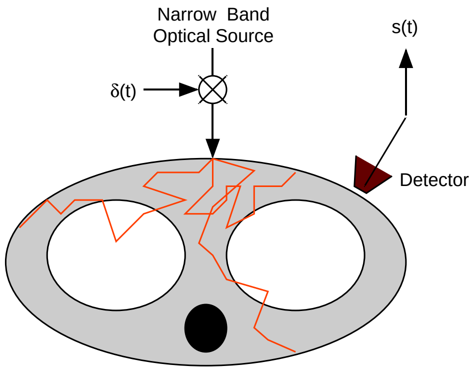
Breast absorption

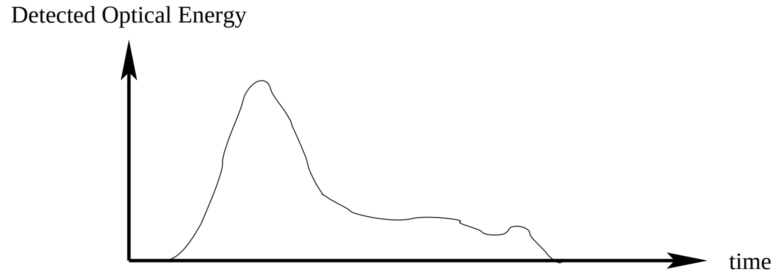


Breast scatter

(Reproduced from McBride *et al.*, *J. Biomed. Opt.*, 2002)

ODT Model: Time Domain





- The photon flux density, $\psi(r,t)$, obeys the **wave** equation

$$\frac{1}{c}\frac{\partial}{\partial t}\psi(r,t) - \nabla \cdot D(r)\nabla\psi(r,t) + \mu(r)\psi(r,t) = S(r,t)$$

$\psi(r,t)$ - photon flux density (photons per unit volume)

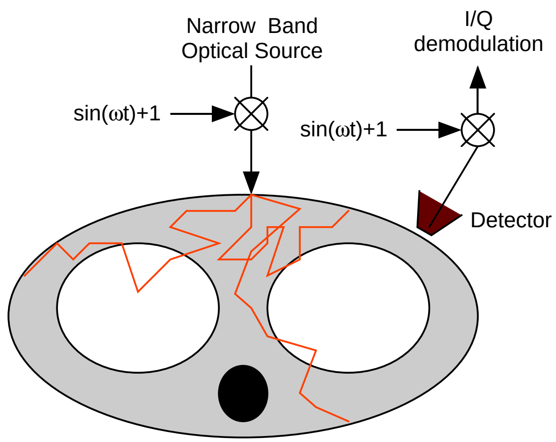
$D(r) = \frac{1}{3(\mu(r)+\mu_s'(r))}$ - Diffusion coefficient (cm)

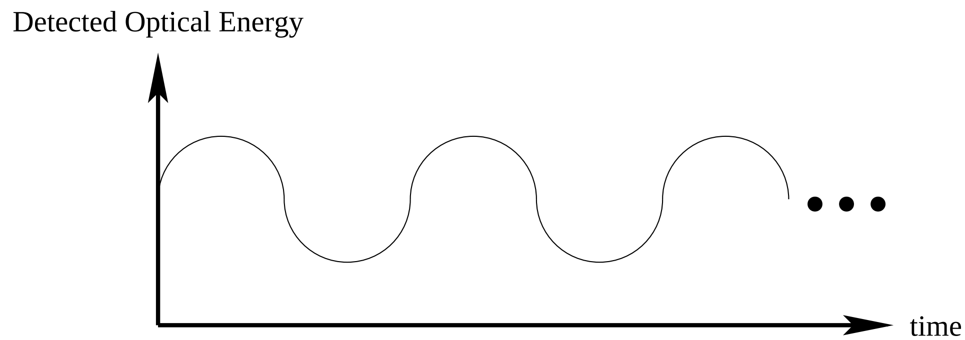
$\mu(r)$ - absorption coefficient (cm^{-1})

$S(r,t)$ - Instantaneous power density of source at location r and time t

r, t - 3-D position and time

ODT Model: Frequency Domain





- If source is modulated by $e^{j\omega t}$, then

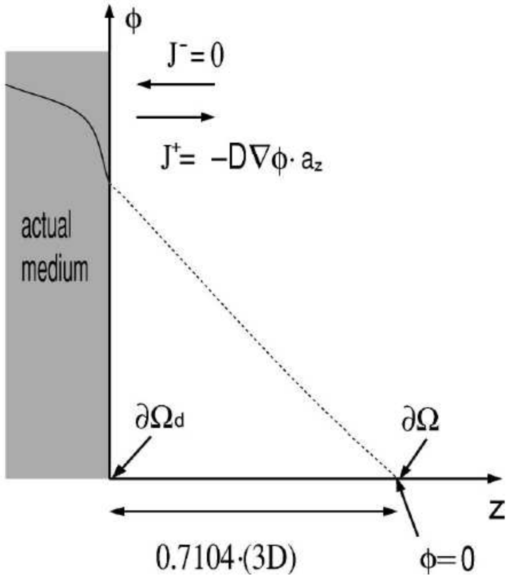
$$\begin{aligned} \psi(r,t) &= \phi(r)e^{j\omega t} \\ S(r,t) &= S(r)e^{j\omega t} \end{aligned}$$

Then the frequency modulated light, $\phi(r)$, obeys the elliptic PDE

$$\nabla \cdot D(r)\nabla\phi(r) - [\mu(r) + j\omega/c]\,\phi(r) = -S(r)$$

- Objective:
  - Measure: $\phi(r)$
  - Reconstruct: $\mu(r)$ and $D(r)$

ODT Model: Boundary Condition [35, 36]



- Basics

  $J = -D\nabla\phi \cdot n$ - photon current in direction n

  J^+ - Photon current leaving boundary

  J^- - Photon current entering boundary

  $D = \frac{1}{3(\mu+\mu_s')}$ - Reduced scattering constant μ_s'

  $l_s = 3D$ - Transport Length for randomization

- Boundary conditions

  $J^- = 0$ - No photons entering boundary

  $J^+ = -D\frac{\partial\phi}{\partial z}$ - Photons leaving boundary

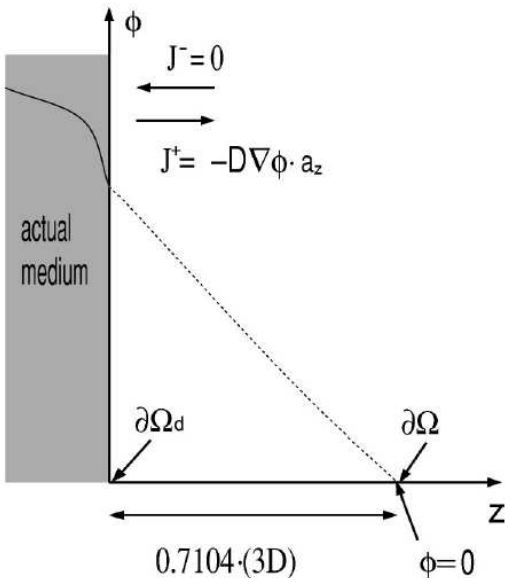
  $J = \frac{1}{3*(0.7104)}\phi$ - Photon current proportional to density ([35] p. 144)

- Boundary condition

$$\boxed{\phi = -(0.7104)3D\frac{\partial\phi}{\partial z}}$$

J. J. Duderstadt and L. J. Hamilton, *Nuclear Reactor Analysis*, Wiley, 1976.

ODT Model: Extrapolated Boundary [35]



Zero-flux condition

- Basics

  $J = -D\nabla\phi \cdot n$ - photon current in direction n

  J^+ - Photon current leaving boundary

  J^- - Photon current entering boundary

  $D = \frac{1}{3(\mu+\mu_s')}$ - Reduced scattering constant μ_s'

  $l_s = 3D$ - Transport Length for randomization

- Set boundary condition to $\phi = 0$ at extrapolated position of

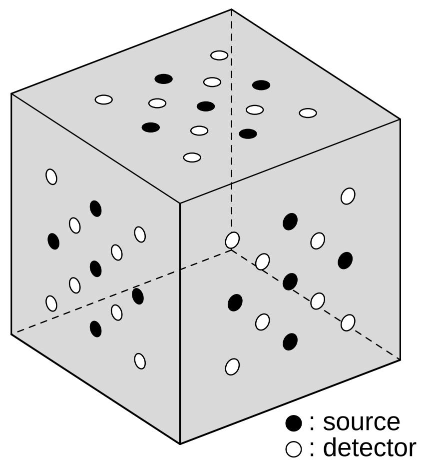
$$z_o = (0.7104)3D$$

- Approximates

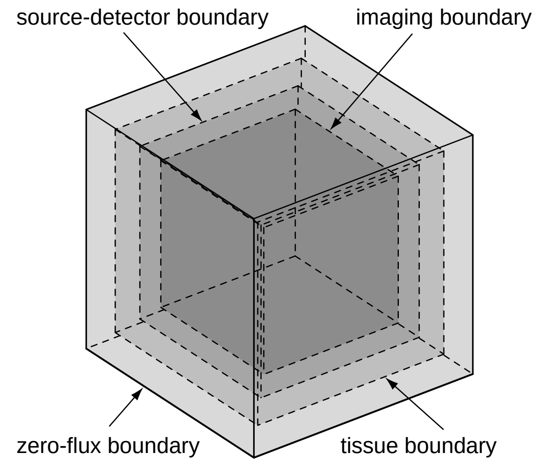
$$\phi = -z_o \frac{\partial \phi}{\partial z}$$

Converts problem to Dirichlet boundary conditions

ODT Model: Source/Detector Location [35]



Sources and detectors

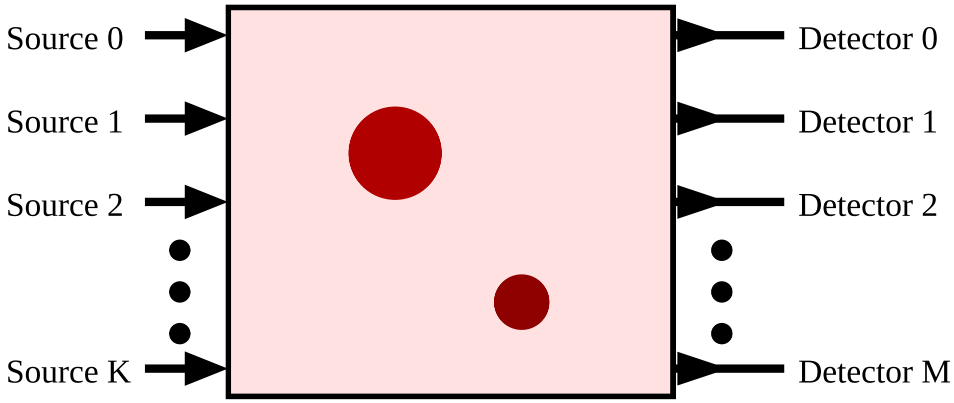


Imaging boundaries

- Computational source/detector locations should be inside tissue boundary
  - Located $l_s = 3D$ inside tissue boundary.
  - Models the distance light must travel to scatter
- Results in four boundaries
  - Zero-flux boundary - Flux set to zero to enforce boundary condition
  - Tissue boundary - Physical boundary of material being imaged
  - Source/detector boundary - Location of sources and detectors
  - Imaging boundary - Region for valid image reconstruction

ODT Model: Typical Measurement Configuration

- Typical assumptions
  - Sources 0 to $K-1$ at locations s_k
  - Detectors 0 to $M-1$ at locations d_m

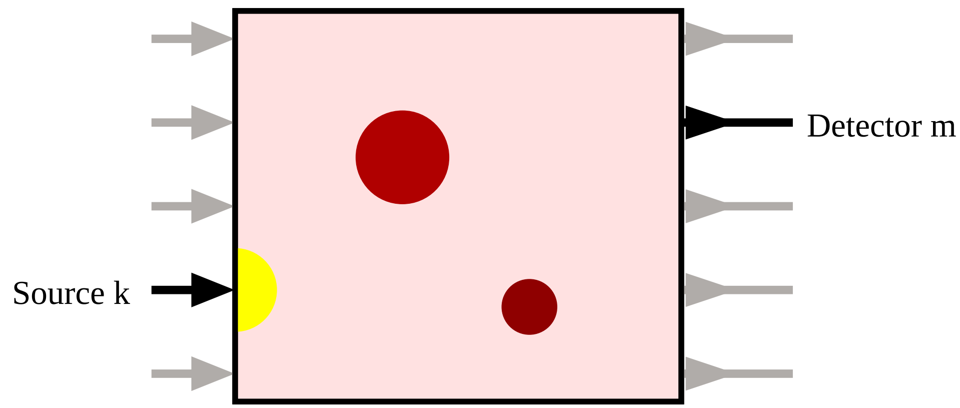


- Total of $P = MK$ measurements
  - Every combination of source and detector
  - Normally, each source must be used in sequence

ODT Model: Forward Operator

- $\phi_k(d_m)$ - The expected measurement between source k and detector m

$$\nabla \cdot D(r)\nabla\phi_k(r) - (\mu(r) + j\omega/c)\phi_k(r) = -S_k\delta(r - s_k)$$



- The full forward model is then

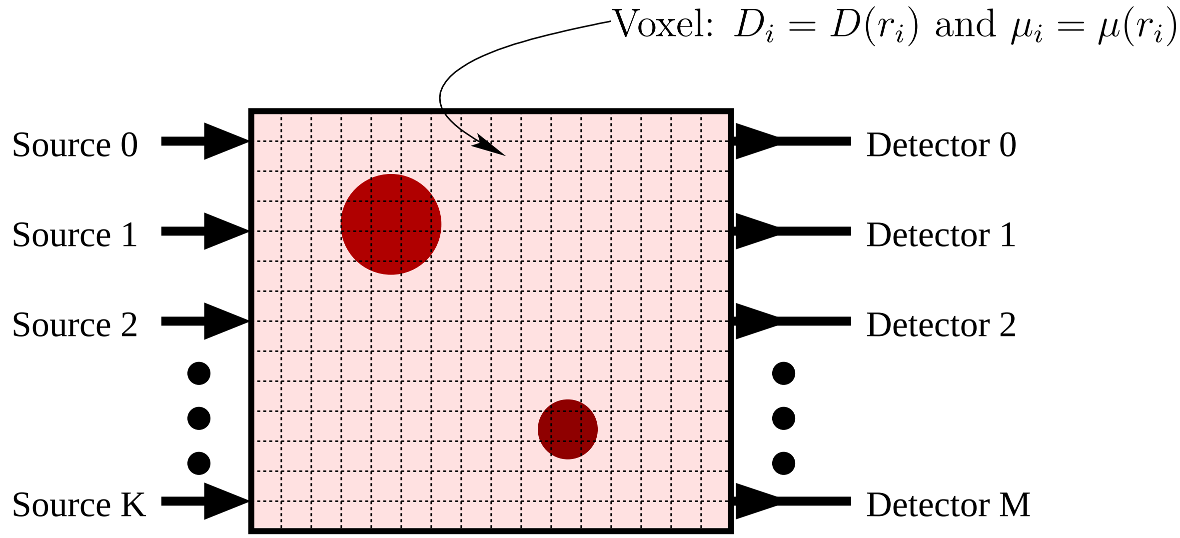
$$f(D,\mu) \triangleq [\underbrace{\phi_0(d_0),\cdots,\phi_0(d_{M-1})}_{\text{source } 0},\cdots,\underbrace{\phi_{K-1}(d_0),\cdots,\phi_{K-1}(d_{M-1})}_{\text{source } K-1}]^t$$

where

$$f_{k*M+m}(D,\mu) = \phi_k(d_m)$$

ODT Model: Discretizing Domain

- Let $r_1, r_2, \cdots, r_N$ be the voxel locations.



- Define the vectors D and μ so that

$$
\begin{aligned}
D &= [D(r_1), D(r_2), \cdots, D(r_N)] \\
\mu &= [\mu(r_1), \mu(r_2), \cdots, \mu(r_N)]
\end{aligned}
$$

- Then the forward model, $f(D, \mu)$, is a mapping from $\mathbb{R}^{2N} \rightarrow C^{KM}$

ODT Model: Computing Forward Operator

- Computing $f(D, \mu)$
  - Need to solve PDE K times (once for each source)
- Computing the gradients $\nabla_D f(D, \mu)$ and $\nabla_\mu f(D, \mu)$
  - Very important for computation of inverse
  - Both are $(MK) \times N$ matrices

$$[\nabla_D f(D, \mu)]_{k*M+m,i} = \frac{\partial \phi_k(d_m)}{\partial D_i}$$

$$[\nabla_\mu f(D, \mu)]_{k*M+m,i} = \frac{\partial \phi_k(d_m)}{\partial \mu_i}$$

  - These form the elements of a Fréchet derivative matrix operator

The Fréchet Derivative

- Question: How does ϕ depend on small changes in D and μ?

$$\begin{aligned} D(r) &\longrightarrow D(r) + \delta D(r) \\ \mu(r) &\longrightarrow \mu(r) + \delta\mu(r) \\ \phi(r) &\longrightarrow \phi(r) + \delta\phi(r) \end{aligned}$$

  Then, for small perturbations, the input/output must have the form

$$\delta\phi(r) = \int_\Omega H_D(r, r')\delta D(r')dr' + \int_\Omega H_\mu(r, r')\delta\mu(r')dr'$$

- The kernels are known as Fréchet derivatives

  $\frac{\delta\phi(r)}{\delta D(r')} = H_D(r, r')$ - Fréchet derivative of ϕ with respect to D

  $\frac{\delta\phi(r)}{\delta\mu(r')} = H_\mu(r, r')$ - Fréchet derivative of ϕ with respect to μ

- How do we compute these?

Fréchet Derivative: General PDE

- Consider a general linear PDE on the domain Ω [2]

$$\mathcal{D}_x \, \phi(r) = S(r)$$

  - $\mathcal{D}_x$ is a linear differential operator with linear coefficients x
  - Use Dirichlet boundary condition $\phi\big|_{\partial\Omega} = 0$ for all problems.

- Examples:

  If $\mathcal{D}_x = x(r)\nabla^2$, then the PDE is

$$x(r)\nabla^2\phi(r) = S(r)$$

  If $\mathcal{D}_x = \nabla \cdot x_1(r)\nabla + x_2(r)$, then the PDE is

$$\nabla \cdot x_1(r)\nabla\phi(r) + x_2(r)\phi(r) = S(r)$$

[2] This analysis can be extended to nonlinear PDE by linearizing the PDE about a solution.

Fréchet Derivative: Solution to General PDE

$$\mathcal{D}_x \, \phi(r) = S(r) \text{ with } \left.\phi\right|_{\partial\Omega} = 0$$

- Let $g_x(r, r')$ be the Green's function for this PDE
  - The Green's function is the solution to

$$\mathcal{D}_x \, g_x(r, r') = \delta(r - r')$$

  - The solution to the PDE can be expressed as a convolution

$$\phi(r) = \int_\Omega g_x(r, r') S(r') dr'$$

Fréchet Derivative: Pertubational Analysis

- By definition, we know that

$$\begin{aligned} \mathcal{D}_x \, \phi(r) &= S(r) \\ \mathcal{D}_{x+\delta x} \, (\phi(r) + \delta\phi(r)) &= S(r) \end{aligned}$$

- Subtracting these yields

$$\mathcal{D}_{x+\delta x} \, (\phi(r) + \delta\phi(r)) - \mathcal{D}_x \, \phi(r) \; = \; 0$$

  Simplifying and dropping high order terms yields

$$\mathcal{D}_x \, \delta\phi(r) = -\mathcal{D}_{\delta x} \, \phi(r)$$

- This is a PDE with forcing function $-\mathcal{D}_{\delta x} \, \phi(r)$. It can be solved using the Green's function as

$$\delta\phi(r) = -\int_\Omega g_x(r, r') \left[\mathcal{D}_{\delta x} \, \phi(r')\right] dr'$$

Fréchet Derivative: Application of Reciprocity

- If $\mathcal{D}_x$ is self-adjoint, then by reciprocity we know that

$$g_x(r, r') = g_x(r', r)$$

- For a detector at location d_m, we can write

$$\begin{aligned} g_m(r) &\triangleq g_x(d_m, r) \\ &= g_x(r, d_m) \end{aligned}$$

- So the perturbed solution has the form

$$\delta\phi(d_m) = -\int_\Omega g_m(r)\mathcal{D}_{\delta x}\,\phi(r)dr$$

where

$$\mathcal{D}_x\, g_m(r) = \delta(r - d_m)$$

Fréchet Derivative: For Diffusion Equation

- For the diffusion equation, we have

$$\mathcal{D}_{\delta x}\,\phi(r) = -\nabla \cdot \delta D \nabla \phi(r) + \delta\mu\phi(r)$$

- For source k and detector m, the perturbation is

$$\begin{aligned} \delta\phi_k(d_m) &= \int_\Omega g_m(r)\nabla \cdot \delta D \nabla \phi_k(r) dr - \int_\Omega g_m(r)\delta\mu\phi_k(r)dr \\ &= -\int_\Omega \nabla g_m(r) \cdot \nabla\phi_k(r)\,\delta D(r)dr - \int_\Omega g_m(r)\phi_k(r)\,\delta\mu(r)dr \\ &\quad + \oint_{\partial\Omega} g_m \delta D \nabla \phi_k \cdot ds \end{aligned}$$

- If $g_m(r) = 0$ and/or $\delta D(r) = 0$ on $\partial\Omega$, then we have

$$\delta\phi_k(d_m) = -\int_\Omega \underbrace{\nabla g_m(r) \cdot \nabla\phi_k(r)}_{H_D(d_m,r)}\,\delta D(r)dr - \int_\Omega \underbrace{g_m(r)\phi_k(r)}_{H_\mu(d_m,r)}\,\delta\mu(r)dr$$

Fréchet Derivative: Derivation of Integral Equality

Let u be a scalar field and v be a vector field on Ω. We know the standard vector identity

$$u\nabla \cdot v = \nabla \cdot (uv) - v \cdot \nabla u$$

Therefore using the divergence theorem, we have

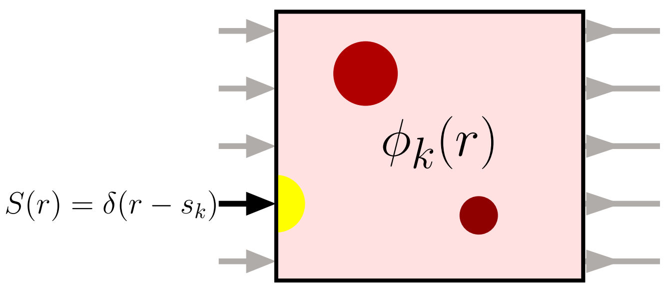
$$\begin{aligned} \int_\Omega u\nabla \cdot v dr &= \int_\Omega \nabla(uv) dr - \int_\Omega v \cdot \nabla u dr \\ &= -\int_\Omega v \cdot \nabla u dr + \oint_{\partial\Omega} uv \cdot ds \end{aligned}$$

For $u = g_m$ and $v = \delta D \nabla \phi_k$

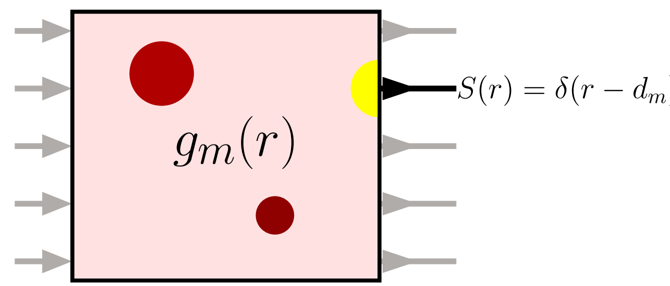
$$\int_\Omega g_m \nabla \cdot [\delta D \nabla \phi_k]\, dr = -\int_\Omega \delta D \nabla \phi_k \cdot \nabla g_m dr + \oint_{\partial\Omega} g_m \delta D \nabla \phi_k \cdot ds$$

Fréchet Derivative: Summary Form for ODT

- For each source, solve PDE for $\phi_k(r)$



- For each detector, solve PDE for $g_m(r)$



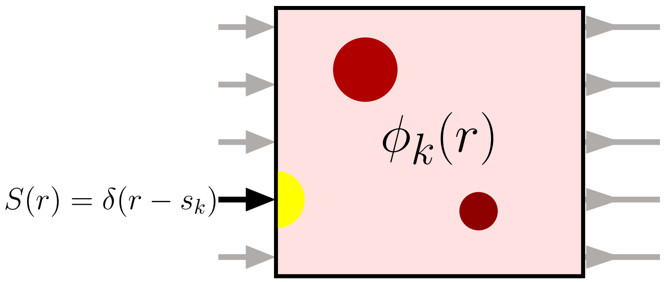
- Use solutions to form Fréchet derivative

$$\frac{\delta \phi_k(d_m)}{\delta D(r)} = -\nabla g_m(r) \cdot \nabla \phi_k(r)$$

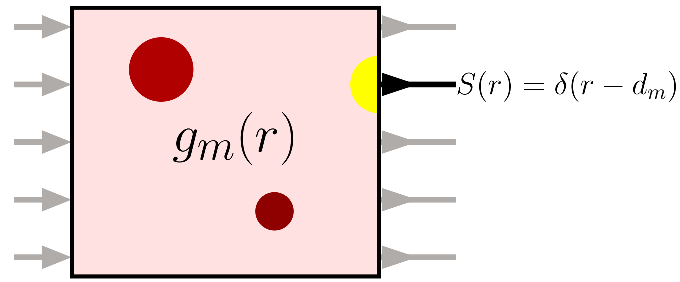
$$\frac{\delta \phi_k(d_m)}{\delta \mu(r)} = -g_m(r)\, \phi_k(r)$$

Fréchet Derivative: Computing $\nabla f(D, \mu)$

- For each source, solve PDE for $\phi_k(r)$



- For each detector, solve PDE for $g_m(r)$



- Gradients of $f(D, \mu)$ can be computed from Fréchet derivative

$$\frac{\partial \phi_k(d_m)}{\partial D_i} = -\nabla g_m(r_i) \cdot \nabla \phi_k(r_i) V$$

$$\frac{\partial \phi_k(d_m)}{\partial \mu_i} = -g_m(r_i)\, \phi_k(r_i) V$$

where V is the volume of a voxel

Fréchet Derivative: Time/Memory Complexity

$$
\begin{aligned}
\text{Time Complexity} &= \underbrace{KMN}_{\text{fill in matrix}} + \underbrace{KLN}_{\text{compute } \phi_k(r)} + \underbrace{MLN}_{\text{compute } g_m(r)} \\
&= KMN + (K+M)LN
\end{aligned}
$$

$$
\begin{aligned}
\text{Memory Complexity} &= \underbrace{KN}_{\text{store } \phi_k(r)} + \underbrace{MN}_{\text{store } g_m(r)} \\
&= (K+M)N
\end{aligned}
$$

K - number of sources

M - number of detectors

N - number of voxels

L - iterations used to solve PDE

Adjoint Differentiation: A Method for Reducing Computational Complexity [37, 38, 39]

- If $M >> L$
  - # of detectors >> # of iterations used to solve PDE
  - Filling in matrix can dominate computation
  - Matrix size = (# of measurments)×(# of voxels)
  - Matrix is fully populated
- Adjoint differentiation
  - Eliminate the explicit computation of $\nabla f(D, \mu)$
  - Based on superposition/duality
  - Applicable for methods such as conjugate gradient
  - General approach for chains of nonlinear systems

Adjoint Differentiation: Gradients of Functionals

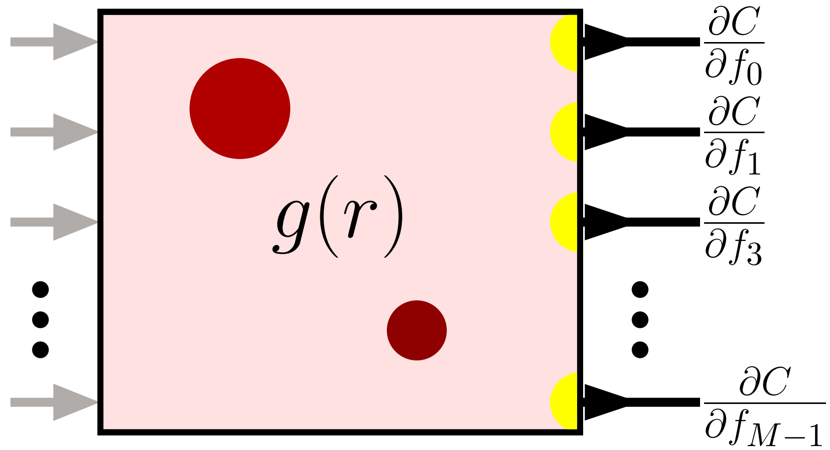
- For simplicity assume
  - $K = 1$ source
  - D known
- Consider a scalar cost functional such as $C(f(\mu)) = ||y - f(\mu)||^2$
  - Need to compute $\nabla C(f(\mu))$
  - Fréchet derivative is $\frac{\partial f_m}{\partial \mu_i} = -g_m(r_i)\,\phi(r_i)$
- By using chain rule

$$
\begin{aligned}
\frac{\partial C}{\partial \mu_i} &= \sum_{m=0}^{M-1} \frac{\partial C}{\partial f_m} \frac{\partial f_m}{\partial \mu_i} \\
&= \sum_{m=0}^{M-1} -\frac{\partial C}{\partial f_m} g_m(r_i)\,\phi(r_i)\,V \\
&= -g(r_i)\,\phi(r_i)\,V
\end{aligned}
$$

where

$$
g(r) = \sum_{m=0}^{M-1} \frac{\partial C}{\partial f_m}\, g_m(r)
$$

Adjoint Differentiation: Using Superposition



$$\frac{\partial C}{\partial \mu_i} = -g(r_i)\,\phi(r_i)\,V$$

- How to efficiently compute $g(r)$?

$$g(r) = \sum_{m=0}^{M-1} \frac{\partial C}{\partial f_m} g_m(r)$$

  – Inject signal into each detector location with amplitude

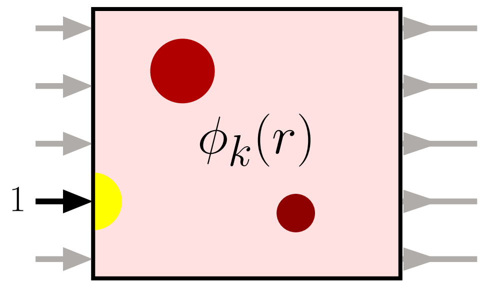
$$\frac{\partial C}{\partial f_m} = -2(y_m - f_m((\mu))$$

  – Exploits superposition of PDE

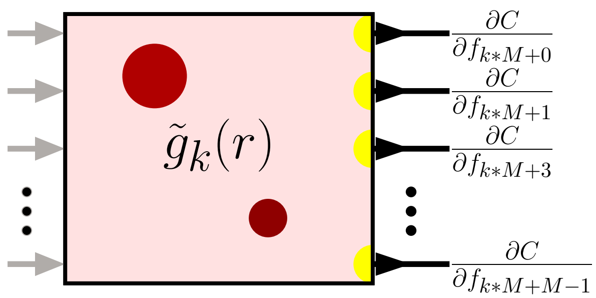
  – Reduces computation by factor of M

Adjoint Differentiation: Using Multiple Sources

- For each source, solve PDE for $\phi_k(r)$



- Solve one PDE for $g(r)$

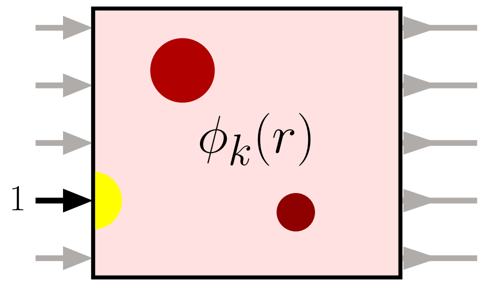


$$\frac{\partial C}{\partial \mu_i} = \sum_{k=0}^{K-1} -\tilde{g}_k(r_i)\,\phi_k(r_i)\,V$$

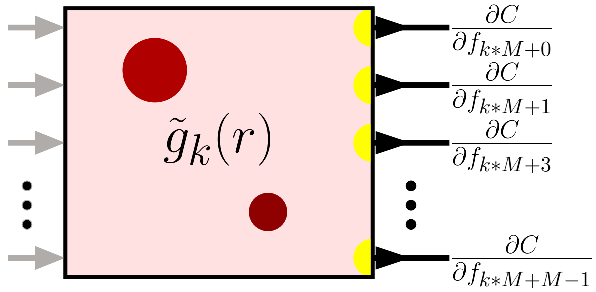
- Notice that $\tilde{g}_k$ is dependent on source index k
- Similar method possible for D

Adjoint Differentiation: Time/Memory Analysis

- For each source, solve PDE for $\phi_k(r)$



- Solve one PDE for $g(r)$



$$\text{Time Complexity} = \underbrace{KLN}_{\text{compute } \phi_k(r)} + \underbrace{KLN}_{\text{compute } \tilde{g}_k(r)}$$

$$\text{Memory Complexity} = \underbrace{KN}_{\text{store } \phi_k(r)} + \underbrace{KN}_{\text{store } \tilde{g}_k(r)}$$

K - number of sources

N - number of voxels

M - number of detectors

L - iterations used to solve PDE

Adjoint Differentiation: Time/Memory Complexity

- If (# of Detectors) << (# of Sources)
  - Reciprocity can be used to reverse roles of sources and detectors
  - Complexity formulae replace K with M

$$\begin{aligned} \text{Time Complexity} &= \min\{K, M\}\, LN \\ \text{Memory Complexity} &= \min\{K, M\}\, N \end{aligned}$$

K - number of sources	N - number of voxels
M - number of detectors	L - iterations used to solve PDE

Direct Formulation of Inverse Problems

- Forward model $\quad x \xrightarrow{f(\cdot)} y = f(x)$

  x : image

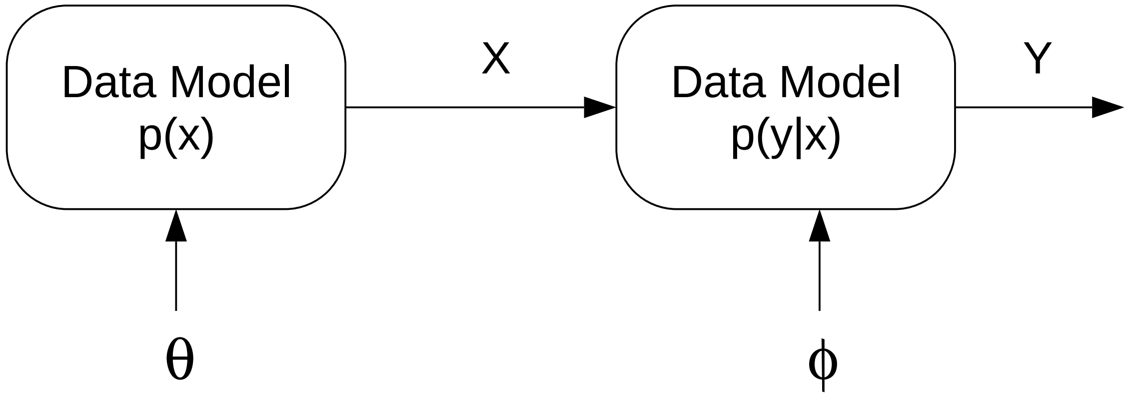
  y : measurement vector

- Inverse Problem $\quad y \xrightarrow{\text{How?}} x$

- Direct approach
  - Directly compute inverse $x = f^{-1}(y)$
  - Filtered back projection (FBP); convolution back projection (CBP); Gauss elimination
  - Computationally efficient
  - Inverse may not exist; may not be unique; may be noise sensitive
  - For non-linear problems, direct inverse may be impossible to compute

- Alternative: Search for x so that $f(x)$ approximately equals y

Bayesian Framework



- Data

  Y - Measured data: P-dimensional real random vector (or $P/2$ dimensional complex random vector)

  X - Unknown Image: N dimension random vector

- Models

  $p(y|x)$ - System model: Incorporates physical properties of system and noise

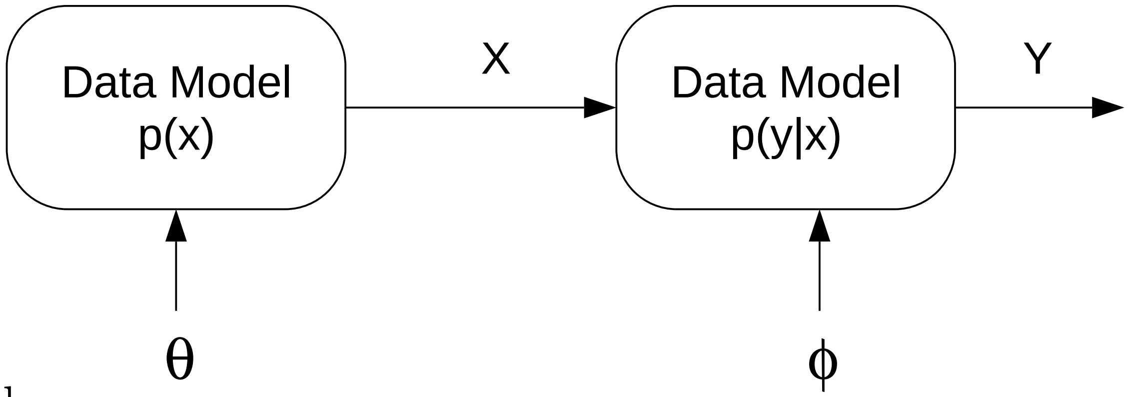
  $p(x)$ - Prior model: Incorporates assumptions of smoothness in image

- Unknown parameters

  ϕ - Unknown parameters of system

  θ - Unknown parameters of prior model

Bayesian Framework: Inversion



- Bayes Rule

$$p(x|y) = \frac{p(y|x)\ p(x)}{p(y)}$$

- Maximum *a posteriori* (MAP) estimation

$$\begin{aligned} \hat{x} &= \arg\max_x p(x|y) \\ &= \arg\max_x \{\log p(y|x) + \log p(x)\} \end{aligned}$$

- Joint MAP estimation of x, θ, and ϕ (with uniform prior)

$$\begin{aligned} \hat{x} &= \arg\max_x \max_{\theta,\phi} p(x|y,\theta,\phi) \\ &= \arg\max_x \max_{\theta,\phi} \{\log p(y|x,\phi) + \log p(x|\theta)\} \end{aligned}$$

Bayesian Framework: Forward Noise Models

- The deterministic component of forward model is $f(x)$, so
$$E[Y|X] = f(X)$$

- Gaussian model with covariance $\alpha\Lambda^{-1}$
$$p(y|x) = \frac{1}{(2\pi\alpha)^{P/2}}|\Lambda|^{1/2} \exp\left\{-\frac{1}{2\alpha}||y - f(x)||_\Lambda^2\right\}$$
where
α : measurement noise factor
$\Lambda \propto$ (measurement covariance)$^{-1}$
P : number of (real valued) dimensions to y

- Poisson model

Bayesian Framework: Generalized Gaussian Markov Random Field (GGMRF) Prior [40]

- We use the generalized GMRF (GGMRF) model

$$\log p(x) = -\frac{1}{p\sigma^p} \sum_{\{i,j\}\in \text{Neighbor}} b_{i-j}|x_i - x_j|^p + \text{constant}$$

where

$1 \le p \le 2$ controls the degree of edge smoothness

- Properties
  - $p = 2$ - Gaussian case (penalizes image edges)
  - $1.2 \approx p$ - Preserves sharp discontinuities
  - $p > 1$ Strictly convex function
  - ML estimate of σ^p

$$\hat{\sigma}^p = \frac{1}{N} \sum_{\{i,j\}\in \text{neighbor}} b_{i-j}|x_i - x_j|^p$$

Bayesian Framework: MAP Estimate with Noise Gain Estimation

$$\hat{x}_{MAP} = \arg\max_{x \geq 0} \max_{\alpha} \{\log p(y|x, \alpha) + \log p(x)\}$$

$$= \arg\max_{x \geq 0} \max_{\alpha} \left\{ -\frac{1}{2\alpha} ||y - f(x)||_{\Lambda}^2 - \frac{P}{2} \log \alpha - \frac{1}{p\sigma^p} \sum_{\{i,j\} \in \mathcal{N}} b_{i-j} |x_i - x_j|^p \right\}$$

$$= \arg\min_{x \geq 0} c(x)$$

where

$$c(x) = \frac{P}{2} \log \left(||y - f(x)||_{\Lambda}^2 \right) + \frac{1}{p\sigma^p} \sum_{\{i,j\} \in \mathcal{N}} b_{i-j} |x_i - x_j|^p$$

- Comments:
  - Automatically adjusts for unknown noise gain α
  - Estimation of α makes global convergence more robust
  - Intuition: Scaling of α preserves relative sizes of data and prior terms

Bayesian Framework: Reconstruction Strategy

- What needs to be done?
  - Formulate function $f(x)$ using PDE model of ODT
  - Specify Gaussian noise model through choice of Λ
  - Specify prior model through choice of p, σ, and b_{i-j}
  - Minimize cost functional

$$c(x) = \frac{P}{2} \log \left(||y - f(x)||^2_\Lambda \right) + \frac{1}{p\sigma^p} \sum_{\{i,j\} \in \mathcal{N}} b_{i-j} |x_i - x_j|^p$$

- Optimization strategy

$$\text{Step 1: } \hat{\alpha} = \frac{1}{P} ||y - f(\hat{x})||^2_\Lambda$$

$$\text{Step 2: } \hat{x} = \arg\min_{x \geq 0} \left\{ \frac{1}{2\hat{\alpha}} ||y - f(x)||^2_\Lambda + \frac{1}{p\sigma^p} \sum b_{i-j} |x_i - x_j|^p \right\}$$

Step 2 is difficult part

Optimization: Major Issues and Choices

- Some Choices:
  - Levenberg-Marquardt
  - Gradient descent, steepest descent
  - Conjugate gradient, preconditioned conjugate gradient
  - Iterative coordinate descent/Gauss-Seidel
  - Multiresolution, wavelet based, and multigrid
- Major Issues:
  - Per iteration computation
  - Number of iterations to convergence
  - Enforcement of convex (e.g. positivity) constraints
  - Convergence to local and (approximate) global minimum
  - Parallel implementation

Optimization: Selected Publications

- Nonlinear optimization
  - Nonlinear conjugate gradient [41, 38, 39, 42, 43, 37, 38, 44, 30]
  - Truncated Newton [45, 46, 44]
  - Levenberg-Marquardt method [47, 48, 49, 50, 51, 52]
  - Quasi-Newton [53]
- Multigrid
  - Seminal [54]
  - General [55, 56, 57, 58, 59, 60, 61]
  - Image reconstruction [62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73]
  - Matching derivative [63, 64, 74, 75]
- Linearized optimization
  - Backprojection: [76, 77, 78, 79, 80, 81]
  - Perturbation approach: Conjugate gradient [82]

Optimization: Conjugate Gradient (CG): Algorithm [83]

- To solve

$$\hat{x} = \arg\min_x c(x)$$

Apply the following algorithm

$$
\boxed{
\begin{array}{l}
\text{Initialize } x \text{ and } g = -\nabla c(x) \text{ and } h = g \\
\text{Repeat until converged } \{ \\
\qquad g_{old} \leftarrow g \ \text{/* store old gradient */} \\
\qquad g \leftarrow -\nabla c(x) \ \text{/* evaluate new gradient */} \\
\qquad \gamma \leftarrow \dfrac{(g - g_{old})^t g}{g_{old}^t \, g_{old}} \ \text{/* Polak-Ribiere */} \\
\qquad h \leftarrow g + \gamma h \ \text{/* compute conjugate */} \\
\qquad \hat{\xi} \leftarrow \arg\min_{\xi} c(x + \xi h) \ \text{/* line search */} \\
\qquad x \leftarrow x + \hat{\xi} d \\
\}
\end{array}
}
$$

Optimization: CG: Issues to Consider

- Advantages:
  - Can exploit adjoint differentiation
  - Rapid convergence for quadratic cost functionals
  - Preconditioning can speed convergence
- Disadvantages:
  - Each update requires a line search:

$$\hat{\xi} = \arg\min_{\xi} c(x + \xi\, d)$$

  - Convex constraints (positivity) requires bending or soft constraint region
  - May behave less predictably on nonquadratic cost functionals e.g. non-Gaussian priors

Optimization: CG: Time/Memory Complexity

- Complexity per iteration

$$
\begin{aligned}
\text{Time Complexity} &= \underbrace{\min\{K, M\} LN}_{\text{Adjoint Differentiation}} + \underbrace{ILN}_{\text{Line Search}} \\
&= (\min\{K, M\} + I) LN \\
\text{Memory Complexity} &= \min\{K, M\} N
\end{aligned}
$$

K - number of sources

M - number of detectors

N - number of voxels

L - iterations used to solve PDE

I - iterations per line search

Optimization: Iterative Coordinate Decent (ICD) [84, 85]

- Sequentially minimize cost functional with respect to each pixel

$$
\boxed{
\begin{array}{l}
\text{Order the voxels from 1 to } N\text{, i.e. } x_1, x_2, \cdots, x_N \\
\text{Repeat until converged } \{ \\
\quad \text{For } i = 1 \text{ to } N \ \{ \\
\qquad x_i \leftarrow \arg\min_{x_i \geq 0} c\left(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_N\right) \\
\quad \} \\
\}
\end{array}
}
$$

- Each voxel update might be for more than one parameter (e.g. μ and D)
- Algorithm for update is unspecified
- Each update monotonically reduces the cost functional
- ICD algorithm must converge to local minimum

Optimization: ICD: Efficient Update

$$c(x) = -\frac{1}{2\hat{\alpha}}||y - f(x)||^2_\Lambda - \frac{1}{p\sigma^p} \sum_{\{i,j\}\in\mathcal{N}} b_{i-j}|x_i - x_j|^p$$

Initialize x and set $e \leftarrow y - f(x)$
Repeat until converged {

$A \leftarrow \nabla f(x)$ /* Compute Fréchet derivative */

For $i = 1$ to N {

$x_{old} \leftarrow x_i$

$\theta_1 \leftarrow -2Re\left\{A^H_{*,i}\Lambda e\right\} \quad /* \; \theta_1 = \frac{\partial c(x)}{\partial x_i} \; */$

$\theta_2 \leftarrow 2A^H_{*,i}\Lambda A_{*,i} \quad /* \; \theta_2 \cong \frac{\partial^2 c(x)}{\partial x_i^2} \; */$

$x_s \leftarrow \arg\min_{\tilde{x}_i \geq 0}\left[\theta_1(\tilde{x}_i - x_i) + \frac{\theta_2}{2}(\tilde{x}_i - x_i)^2 + \frac{1}{p\sigma^p}\sum_{j\in\mathcal{N}_i}|\tilde{x}_i - x_j|^p\right]$

$e \leftarrow e + A_{*,i}(\tilde{x}_i - x_{old})$

}

}

Optimization: ICD: Update Analysis

Initialize x and set $e \leftarrow y - f(x)$
Repeat until converged {

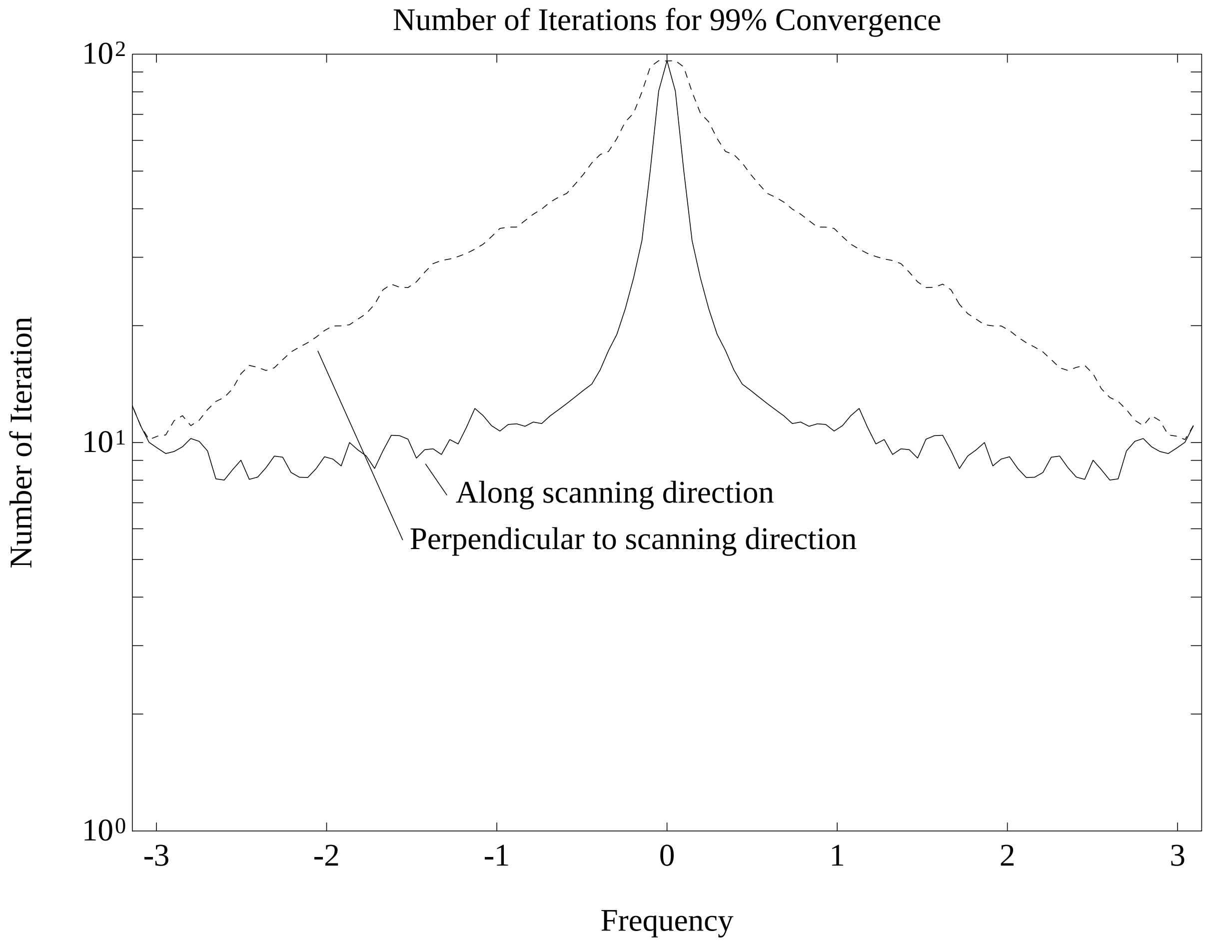
$A \leftarrow \nabla f(x)$ /* Compute Fréchet derivative */
For $i = 1$ to N {

$x_{old} \leftarrow x_i$
$\theta_1 \leftarrow -2Re\left\{A_{*,i}^H \Lambda e\right\}$
$\theta_2 \leftarrow 2A_{*,i}^H \Lambda A_{*,i}$
$x_i \leftarrow \arg\min_{\tilde{x}_i \geq 0} \left[\theta_1(\tilde{x}_i - x_i) + \frac{\theta_2}{2}(\tilde{x}_i - x_i)^2 + \frac{1}{p\sigma^p}\sum_{j\in\mathcal{N}_i} |\tilde{x}_i - x_j|^p\right]$
$e \leftarrow e + A_{*,i}(\tilde{x}_i - x_{old})$
}
}

- A is Fréchet derivative
- $A_{*,i}$ is i^{th} column of A
- $e = y - f(x)$ - forward error
- $\theta_1 = \frac{\partial c(x)}{\partial x_i}$
- $\theta_2 \cong \frac{\partial^2 c(x)}{\partial x_i^2}$

- Updates of θ_1, θ_2, and e dominate computation
- Voxel minimization can be done with simple numerical search
- Easy to implement positivity constraints
- Easy to incorporate non-Gaussian priors
- Can be easily extended to Poisson noise model [86]

Optimization: ICD: Convergence Analysis



- Example of convergence behavior for traditional tomography
- Low frequencies tend to converge more slowly
- Can be important to have good low frequency initial condition

Optimization: ICD: Time/Memory Complexity

- Complexity per iteration

$$\text{Time Complexity} = \underbrace{KMN}_{\text{voxel updates}} + \underbrace{(K+M)LN}_{\text{Fréchet derivative}}$$

$$\text{Memory Complexity} = (K+M)N$$

K - number of sources

M - number of detectors

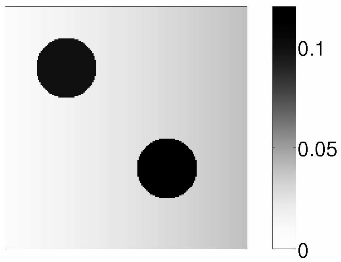
N - number of voxels

L - iterations used to solve PDE

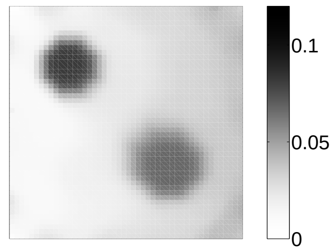
I - iterations per line search

- CG versus ICD?
  - Which is greater $(K+M) \gtrless (\min\{K, M\} + I)$?
  - Non-Gaussian priors and positivity constraints favor ICD

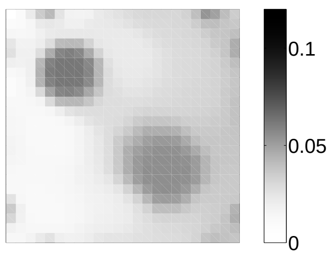
Optimization: Multigrid Inversion: The Problem



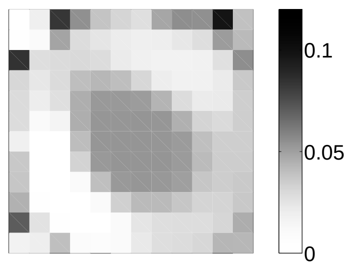
True phantom



$65 \times 65 \times 65$ Recon.



$33 \times 33 \times 33$ Recon.



$17 \times 17 \times 17$ Recon.

- Choosing a discretization resolution
  - Coarse grid $\Rightarrow$ fast
  - Fine grid $\Rightarrow$ accurate
- Approach
  - Move back-and-forth between resolutions
  - Better speed and accuracy
- Some existing multigrid methods
  - Nonlinear multigrid solvers[54, 87]
  - Tomographic reconstruction[62, 69]

- **Opportunity:**
  - Robust convergence to optimum
  - Fast and accurate
  - Use any fixed grid optimizer

- **Challenge:**
  - Fine and coarse scale cost functions are not consistent
  - Dynamically adjust cost functions for consistency

Optimization: Multigrid Inversion: Approach[64, 68, 88]

- Formulate a cost functional at scale q

$$c^{(q)}(x^{(q)}; y^{(q)}, r^{(q)}) = \frac{P}{2} \log ||y^{(q)} - f^{(q)}(x^{(q)})||_\Lambda^2 + S^{(q)}(x^{(q)}) - r^{(q)} x^{(q)}$$

$f^{(q)}(\cdot)$ - coarse scale **forward** model

$x^{(q)}$ - coarse scale solution

$S^{(q)}(\cdot)$ - coarse scale stabilizing functional

$y^{(q)}$ - coarse scale measurement

$r^{(q)}$ - adjustment factor at scale q

- **Key issues:**
  - Coarse discretization of **forward** model reduces computation
  - $y^{(q)}$ and $r^{(q)}$ can be dynamically selected at each scale
  - Novel optimization based approach [64, 74]

Optimization: Multigrid Inversion: Simplified Recursion

$x^{(q)} \leftarrow \text{MultigridV}\left[\, c^{(q)}\left(x^{(q)}; y^{(q)}, r^{(q)}\right) \,\right] \{$

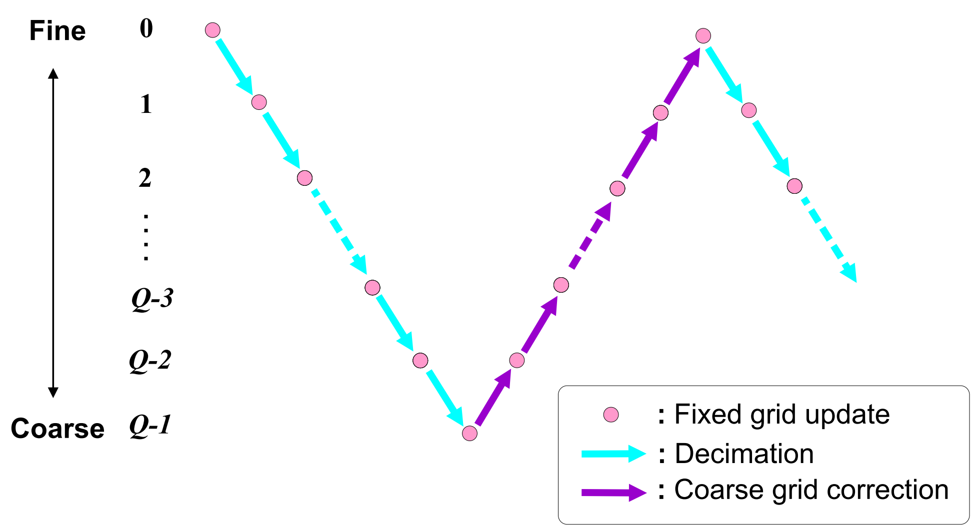
    Select $y^{(q+1)}$ and $r^{(q+1)}$ /* But, how? */

    $x^{(q+1)} \leftarrow \text{MultigridV}\left[c^{(q+1)}\left(x^{(q+1)}; y^{(q+1)}, r^{(q+1)}\right)\right]$

    Correct $x^{(q)}$ by using $x^{(q+1)}$

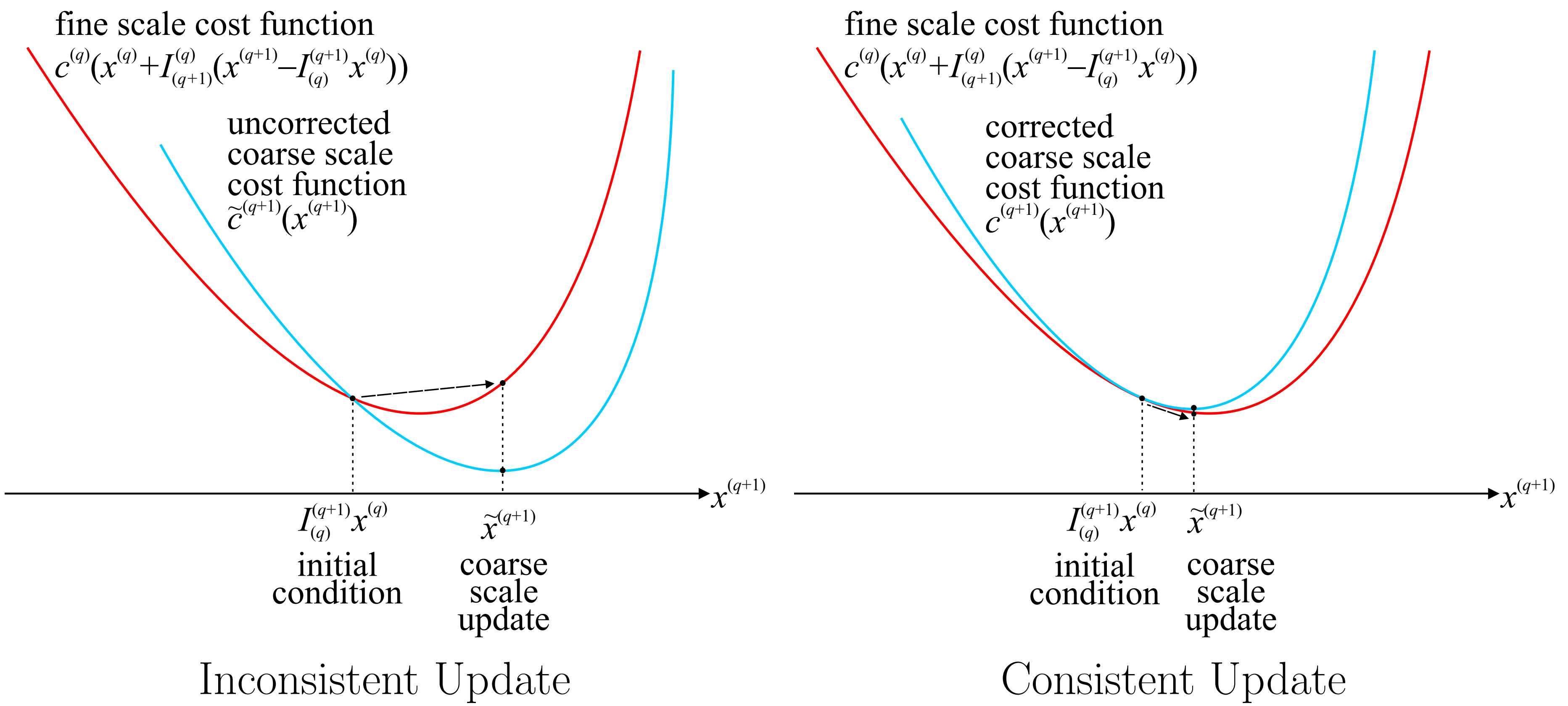
    Apply fixed grid optimizer to $x^{(q)}$

$\}$



- MultigridV recursion
  - MultigridV calls itself
  - Moves from fine to coarse to fine
- Questions
  - How are $y^{(q+1)}$ and $r^{(q+1)}$ selected?
  - How is $x^{(q)}$ corrected using $x^{(q+1)}$?

Optimization: Multigrid Inversion: Graphical Interpretation of Consistent Cost Functionals



- Choose $y^{(q+1)}$ and $r^{(q+1)}$ so that coarse scale cost function should:
  - Be tangent to fine scale cost functional at initial solution
  - Upper bound fine scale cost functional

Optimization: Multigrid Inversion: Choosing $y^{(q+1)}$ and $r^{(q+1)}$

- Match error in data term at coarse and fine scales

$$y^{(q+1)} \leftarrow y^{(q)} - \left[f^{(q)}(x^{(q)}) - f^{(q+1)}(I_{(q)}^{(q+1)} x^{(q)}) \right]$$

- Match derivatives in cost function at coarse and fine scales

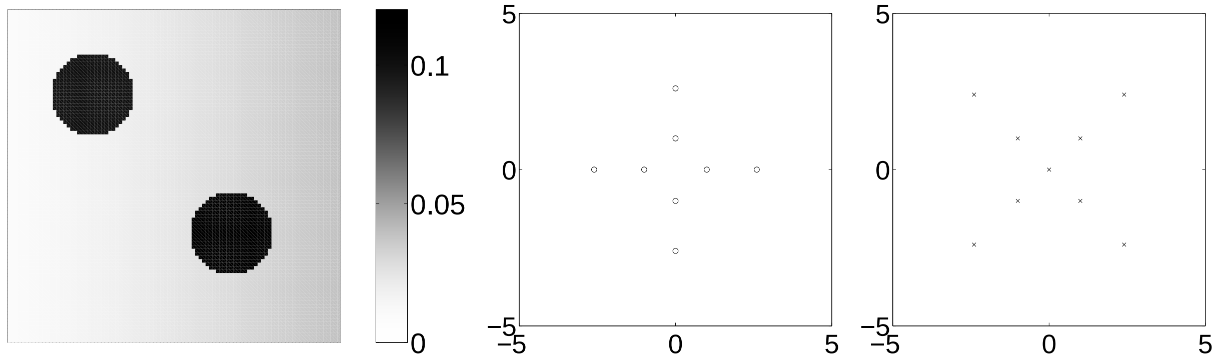
$$r^{(q+1)} \leftarrow \nabla \tilde{c}^{(q+1)}(x^{(q+1)}) \Big|_{x^{(q+1)} = I_{(q)}^{(q+1)} x^{(q)}} - \left(\nabla \tilde{c}^{(q)}(x^{(q)}) - r^{(q)} \right) I_{(q+1)}^{(q)}$$

where

$$\tilde{c}^{(q)}(x^{(q)}) \triangleq \frac{P}{2} \log ||y^{(q)} - f^{(q)}(x^{(q)})||_{\Lambda}^{2} + S^{(q)}(x^{(q)})$$

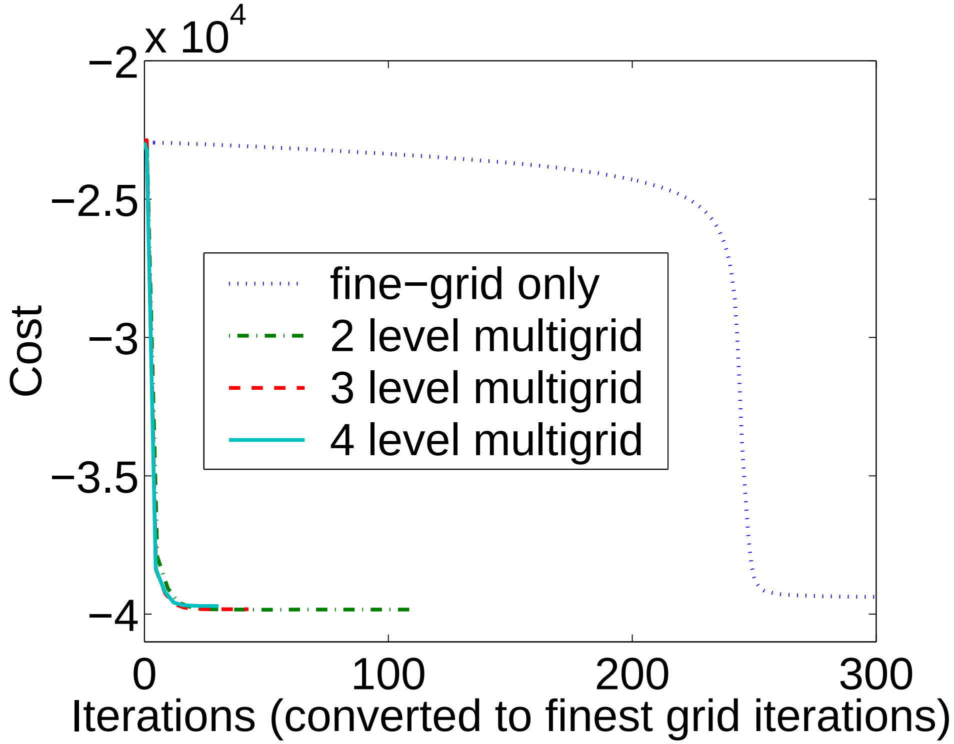
- Theorem: If the difference between cost functionals is convex, then multigrid iterations generate monotone decreasing cost.

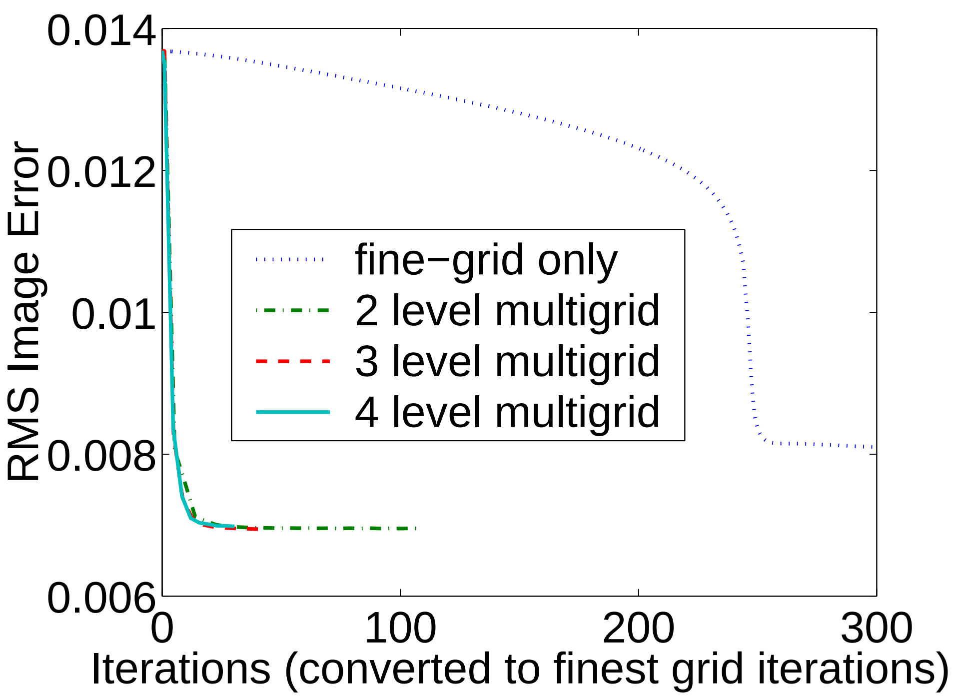
Optimization: Multigrid Inversion: Simulation



- Phantom
  - 10cm $\times$ 10cm $\times$ 10cm cube
  - Linearly varying background: $\mu = 0.01\text{cm}^{-1}$ to 0.04cm^{-1}
  - Two spherical inhomogeneities with diameters of 1.85cm with $\mu = 0.10\text{cm}^{-1}$ and $\mu = 0.12\text{cm}^{-1}$
  - Diffusion coefficient, D, is constant
- Model
  - Sources and detectors on all 6 faces using 100MHz modulation frequency and 35dB average SNR
  - GGMRF with $p = 1.2$ and $\sigma = 0.018\text{cm}^{-1}$

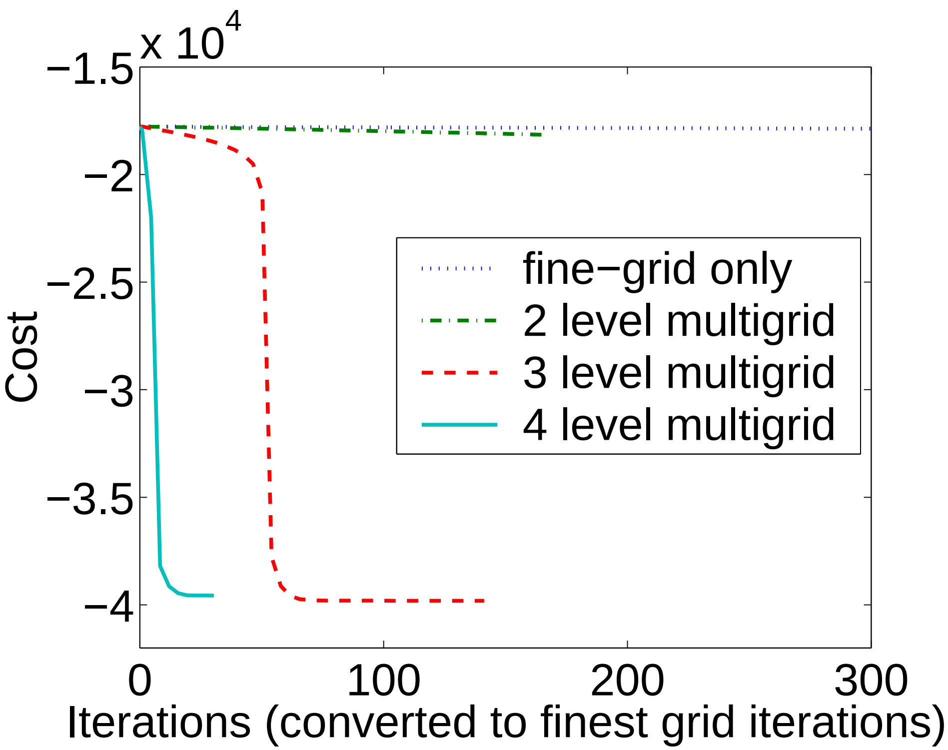
Optimization: Multigrid Inversion: Convergence Speed for $65 \times 65 \times 65$ Reconstruction with Good Initial Condition

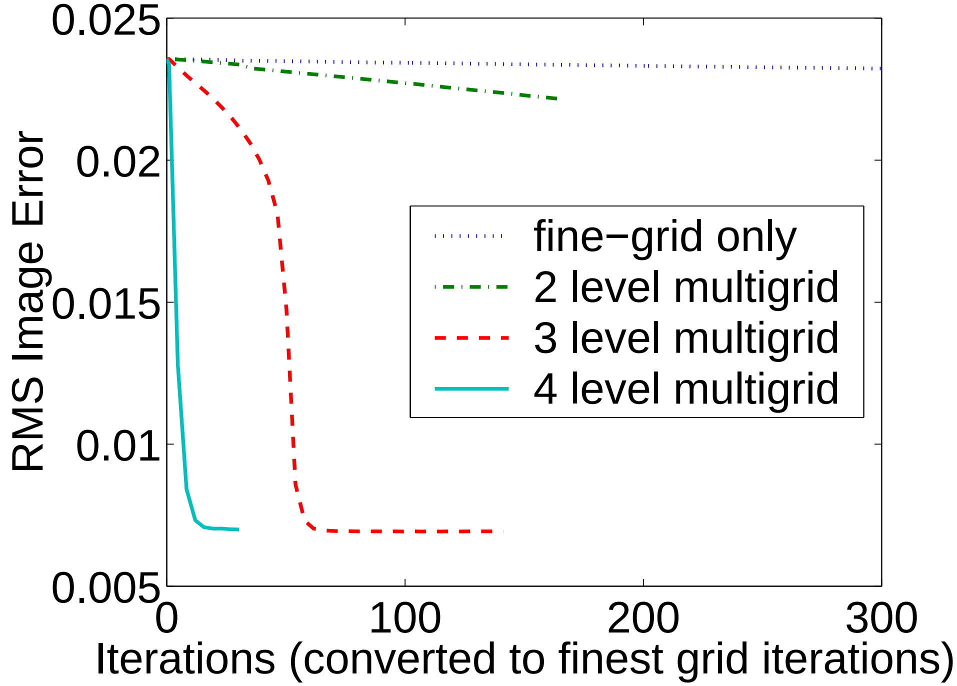




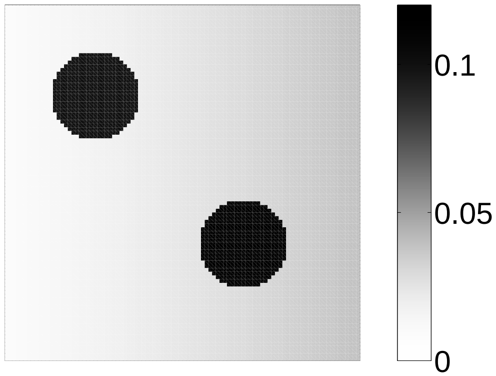
- All iterations in units of a single fixed grid iteration

Optimization: Multigrid Inversion: Convergence Speed for $65 \times 65 \times 65$ Reconstruction with Poor Initial Condition

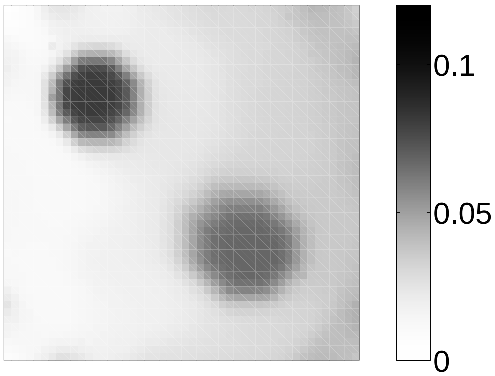




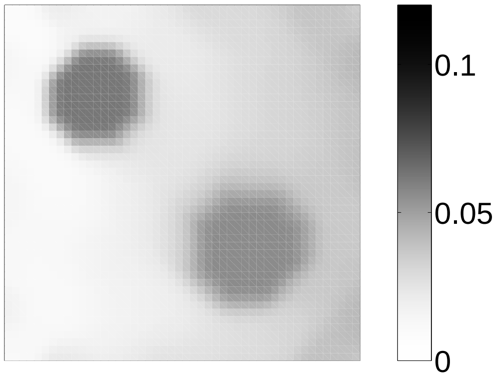
Optimization: Multigrid Inversion: Reconstruction Quality for Multigrid and Fixed Grid Algorithms



True phantom



4 level multigrid
(25.6 iter.)



Fixed-grid
(300 iter.)

Linearized Approach: Viewpoint

- Perturbation can be approximated by linearization using Fréchet derivative

$$\delta\phi_k(d_m) = \sum_{i=1}^{N} \frac{\partial\phi_k(d_m)}{\partial\mu(r_i)}\delta\mu(r_i) + \sum_{i=1}^{N} \frac{\partial\phi_k(d_m)}{\partial D(r_i)}\delta D(r_i)$$

- This can be viewed as

$$\tilde{y} \cong A \begin{bmatrix} \delta\mu \\ \delta D \end{bmatrix}$$

...or as

$$\min_{\mu,D}\left\{\frac{1}{2\alpha}\left\|\tilde{y} - A\begin{bmatrix} \delta\mu \\ \delta D \end{bmatrix}\right\|_\Lambda^2 - \frac{1}{p\sigma^p}\sum_{\{i,j\}\in\mathcal{N}} b_{i-j}|x_i - x_j|^p\right\}$$

where $\tilde{y} = y - f(\mu, D)$

- The solution is then $\mu + \delta\mu$ and $D + \delta D$

- Question: How do we choose μ and D?

Linearized Approach: Specific Form

- Perturbation can be approximated by linearization

$$\delta\phi_k(d_m) = \sum_{i=1}^{N} \frac{\partial \phi_k(d_m)}{\partial \mu(r_i)} \delta\mu(r_i) + \sum_{i=1}^{N} \frac{\partial \phi_k(d_m)}{\partial D(r_i)} \delta D(r_i)$$

- Can be formulated as a linear equation

$$\begin{bmatrix} \delta\phi_1(d_1) \\ \vdots \\ \delta\phi_1(d_M) \\ \delta\phi_2(d_1) \\ \vdots \\ \delta\phi_K(d_M) \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{\partial\phi_1(d_1)}{\partial\mu(r_1)} & \cdots & \frac{\partial\phi_1(d_1)}{\partial\mu(r_N)} & \frac{\partial\phi_1(d_1)}{\partial D(r_1)} & \cdots & \frac{\partial\phi_1(d_1)}{\partial D(r_N)} \\ \frac{\partial\phi_1(d_2)}{\partial\mu(r_1)} & \cdots & \frac{\partial\phi_1(d_2)}{\partial\mu(r_N)} & \frac{\partial\phi_1(d_2)}{\partial D(r_1)} & \cdots & \frac{\partial\phi_1(d_2)}{\partial D(r_N)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\phi_1(d_M)}{\partial\mu(r_1)} & \cdots & \frac{\partial\phi_1(d_M)}{\partial\mu(r_N)} & \frac{\partial\phi_1(d_M)}{\partial D(r_1)} & \cdots & \frac{\partial\phi_1(d_M)}{\partial D(r_N)} \\ \frac{\partial\phi_2(d_1)}{\partial\mu(r_1)} & \cdots & \frac{\partial\phi_2(d_1)}{\partial\mu(r_N)} & \frac{\partial\phi_2(d_1)}{\partial D(r_1)} & \cdots & \frac{\partial\phi_2(d_1)}{\partial D(r_N)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\phi_K(d_M)}{\partial\mu(r_1)} & \cdots & \frac{\partial\phi_K(d_M)}{\partial\mu(r_N)} & \frac{\partial\phi_K(d_M)}{\partial D(r_1)} & \cdots & \frac{\partial\phi_K(d_M)}{\partial D(r_N)} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \delta\mu(r_1) \\ \vdots \\ \delta\mu(r_N) \\ \delta D(r_1) \\ \vdots \\ \delta D(r_N) \end{bmatrix}}_{\begin{bmatrix} \mu \\ D \end{bmatrix}}$$

Linearized Approach: Reconstruction

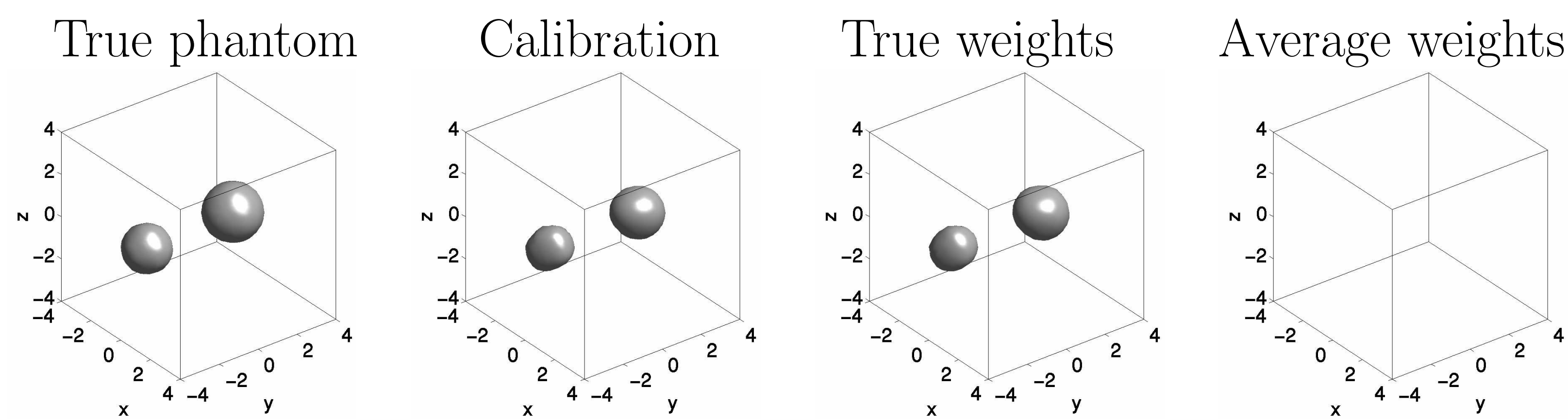
- Direct methods
  - Singular value decomposition (SVD) [89, 90, 91]
- Iterative methods
  - Algebraic reconstruction techniques (ART) and simultaneous iterative reconstruction technique (SIRT) [90]
  - Conjugate gradient (CG) [90]
  - Saves re-computation of Green's functions
- Back-projection algorithm
  - Widely used in tomographic imaging systems (e.g. CT, PET)
  - Works well for "ideal" projections with negligible scatter
- Limitations
  - Requires the selection of baseline values for μ and D
  - Results can be seriously affected by erroneous choice of μ and D
  - Quantitative imaging of spatially varying optical properties is very difficult due to linearization error [92]

Source-Detector Calibration: Coupling Variability

- Variability in practical ODT imaging systems
  - Excitation strength of sources
  - Collector efficiency of detectors
  - Effective position of sources and detectors

- Variability can be modelled with source-detector coupling parameters [93]
- Calibration of these unknown parameters is essential for effective imaging

- Importance of calibration [94]



S-D Calibration: Approaches

- No calibration: Assume known weights (e.g. use $1 + 0j$ for all weights)
- Preprocessing
  - Compare prior measurements for a homogeneous medium with the corresponding computed values [95, 96, 97, 98, 34, 99]
  - Minimize the error between the measurements for the given inhomogeneous phantom and the computed values with an assumed homogeneous medium [100]
- Simultaneous reconstruction and calibration [93, 94, 22, 101]
  - Forward model

$$f_{kM+m}(x, \beta_k, \gamma_m) = \beta_k \gamma_m \phi_k(\gamma_m; x)$$

    β_k: source excitation

    γ_m: detector collection efficiency
  - Estimate β_k's and γ_m's while reconstructing image x
  - We will focus on this approach!

S-D Calibration: Joint MAP Estimation [94, 22]

- Data likelihood function $p(y|x,\beta,\gamma,\alpha)$ involves weights β and γ

$$\log p(y|x,\beta,\gamma,\alpha) = \frac{1}{\alpha}||y - f(x,\beta,\gamma)||_\Lambda^2 + constant$$

- MAP estimate for both the image and unknown model parameters results in single cost function for both image and weights

$$\begin{aligned}(\hat{x}_{MAP},\hat{\beta},\hat{\gamma},\hat{\alpha}) &= \arg\max_{x\geq 0,\beta,\gamma,\alpha}\{\log p(y|x,\beta,\gamma,\alpha) + \log p(x)\} \\ &= \arg\max_{x\geq 0,\beta,\gamma,\alpha} c(x,\beta,\gamma,\alpha)\end{aligned}$$

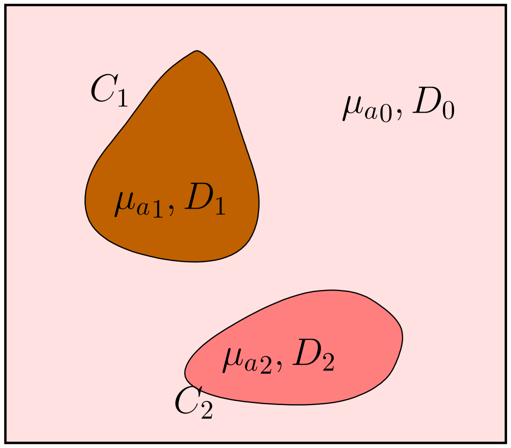
$p(x)$: image prior model

- Coordinate descent optimization of the single cost function

$$\begin{aligned}&\text{For each iteration } \{ \\ &\qquad \hat{\alpha} \leftarrow \arg\min_\alpha c(\hat{x},\hat{\beta},\hat{\gamma},\alpha) \\ &\qquad \hat{\beta} \leftarrow \arg\min_\beta c(\hat{x},\beta,\hat{\gamma},\hat{\alpha}) \\ &\qquad \hat{\gamma} \leftarrow \arg\min_\gamma c(\hat{x},\hat{\beta},\gamma,\hat{\alpha}) \\ &\qquad \hat{x} \leftarrow ICD\_update_x\ c(x,\hat{\beta},\hat{\gamma},\hat{\alpha}) \\ &\}\end{aligned}$$

Shape-Based Reconstruction

- Estimate region parameters directly from data

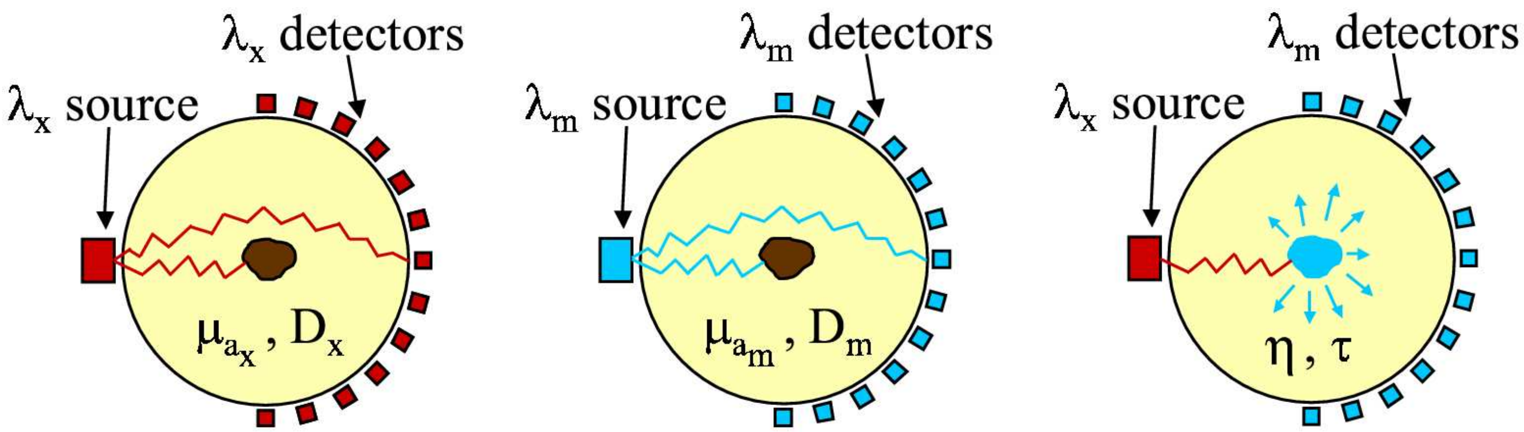


- Boundary parameters $C_0, \ldots, C_N$
- Optical parameters $\mu_0, \ldots, \mu_N$, $D_0, \ldots, D_N$

- Useful for identifying localized anomalies e.g. Breast tumor detection
- Reduced number of parameters $\Rightarrow$ higher SNR
- Regions may also vary in space or time [102]
- Regions may be parameterized by:
  – 2-D: B-spline [103], Fourier series [104, 105, 106]
  – 3-D: ellipsoid [107]
  – Level set methods [108, 109]
  – Dynamic methods [102]

Fluorescence ODT (FODT)[99]

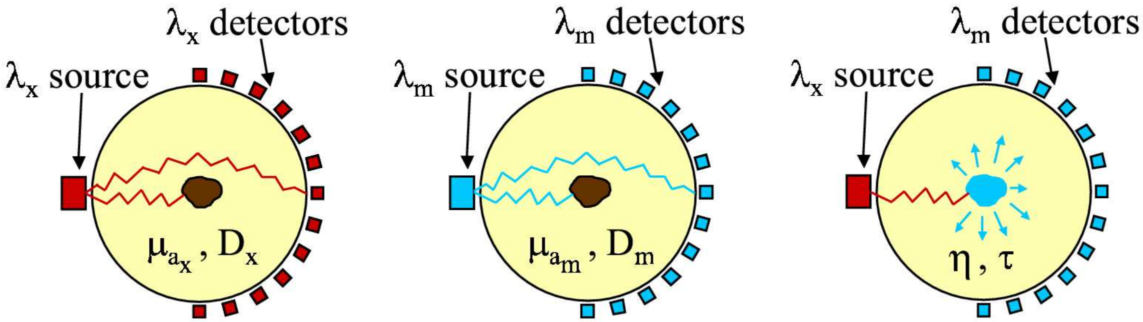
- Basic concept fluorescent tagging
  - Fluorophore provides enhanced contrast
  - With targeted delivery $\longrightarrow$ site-specific imaging (e.g. tumors)
  - Fluorophore absorbs energy at excitation wavelength, λ_x
  - Fluorophore re-emits at emission wavelength, λ_m



- Possible measurement scenarios
  - Source at λ_x and detect at $\lambda_x \Rightarrow$ ODT model
  - Source at λ_m and detect at $\lambda_m \Rightarrow$ ODT model
  - Source at λ_x and detect at $\lambda_m \Rightarrow$ Coupled ODT model

A. B. Milstein, *et al.*, "Fluorescence Optical Diffusion Tomography", *Applied Optics,* 2003.

FODT: Time Domain Model at λ_x



- The photon flux density, $\psi_x(r,t)$, obeys the diffusion equation

$$\frac{1}{c}\frac{\partial}{\partial t}\psi_x(r,t) - \nabla \cdot D_x(r)\nabla\psi_x(r,t) + \mu_x(r)\psi_x(r,t) = S_x(r,t)$$

- Quantities at excitation wavelength

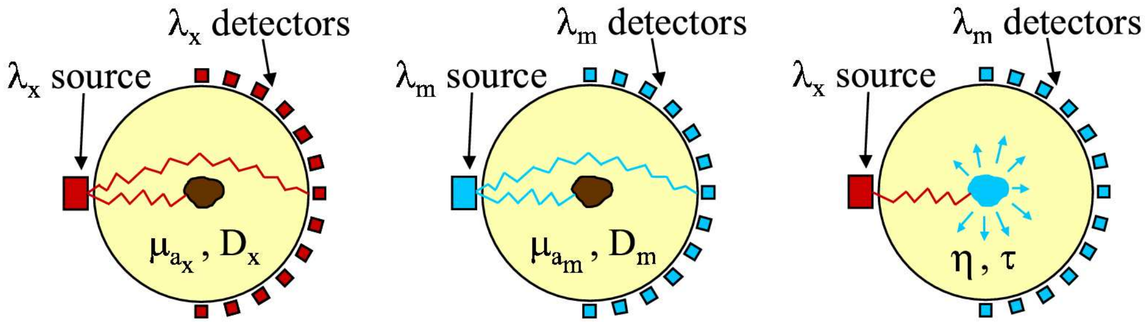
  $\psi_x(r,t)$ - photon flux density

  $D_x(r)$ - Diffusion coefficient

  $\mu_x(r)$ - absorption coefficient

  $S_x(r,t)$ - Source power at location r and time t

FODT: Time Domain Model Re-emission at λ_m



$$\frac{1}{c}\frac{\partial}{\partial t}\psi_x(r,t) - \nabla \cdot D_x(r)\nabla\psi_x(r,t) + \mu_x(r)\psi_x(r,t) = S_x(r,t)$$

- Light is re-emitted at wavelength λ_m as ($*$ denotes time-domain convolution)

$$\psi_x(r,t) * \left\{\frac{\eta(r)}{\tau(r)}e^{-t/\tau(r)}\right\}$$

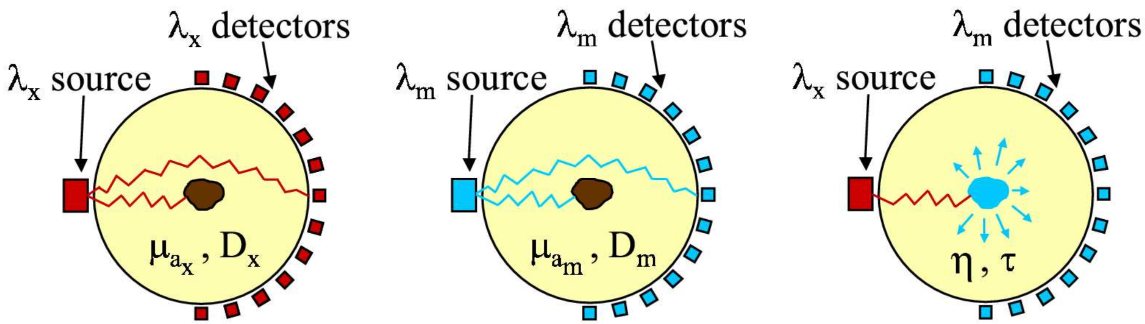
$\psi_x(r,t)$ - photon flux density at excitation wavelength

$\eta(r)$ - fluorescent yield

$\tau(r)$ - fluorescent lifetime

- Models fluorescent impulse response as exponential with time constant τ

FODT: Time Domain Model at λ_m



$$\frac{1}{c}\frac{\partial}{\partial t}\psi_m(r,t) - \nabla\cdot D_m(r)\nabla\psi_m(r,t) + \mu_m(r)\psi_m(r,t) = \underbrace{\psi_x(r,t) * \left\{\frac{\eta(r)}{\tau(r)}e^{-t/\tau(r)}\right\}}_{\text{fluorescent source}} + S_m(r,t)$$

- Quantities at emission wavelength

  $\psi_m(r,t)$ - photon flux density

  $D_m(r)$ - Diffusion coefficient

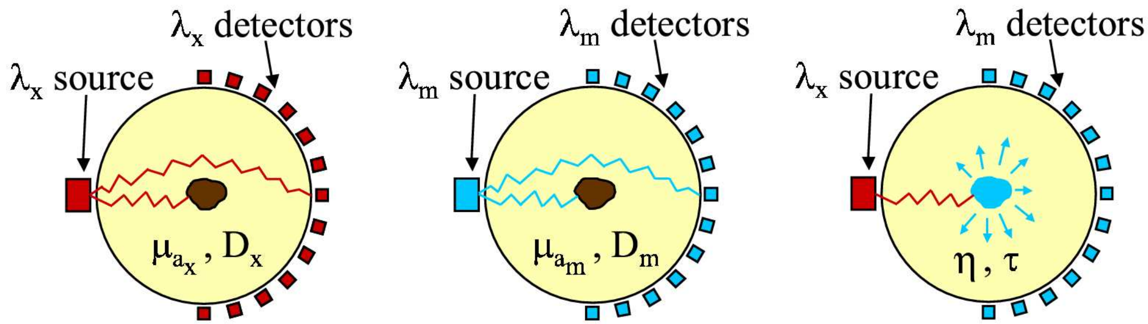
  $\mu_m(r)$ - absorption coefficient

  $S_m(r,t)$ - Source power at location r and time t

  $\eta(r)$ - fluorescent yield

  $\tau(r)$ - fluorescent lifetime

FODT: Frequency Domain Model



Source terms:

$$\begin{aligned} S_x(r,t) &= S_x(r)e^{j\omega t} \\ S_m(r,t) &= S_m(r)e^{j\omega t} \end{aligned}$$

Photon density terms

$$\begin{aligned} \psi_x(r,t) &= \phi_x(r)e^{j\omega t} \\ \psi_m(r,t) &= \phi_m(r)e^{j\omega t} \end{aligned}$$

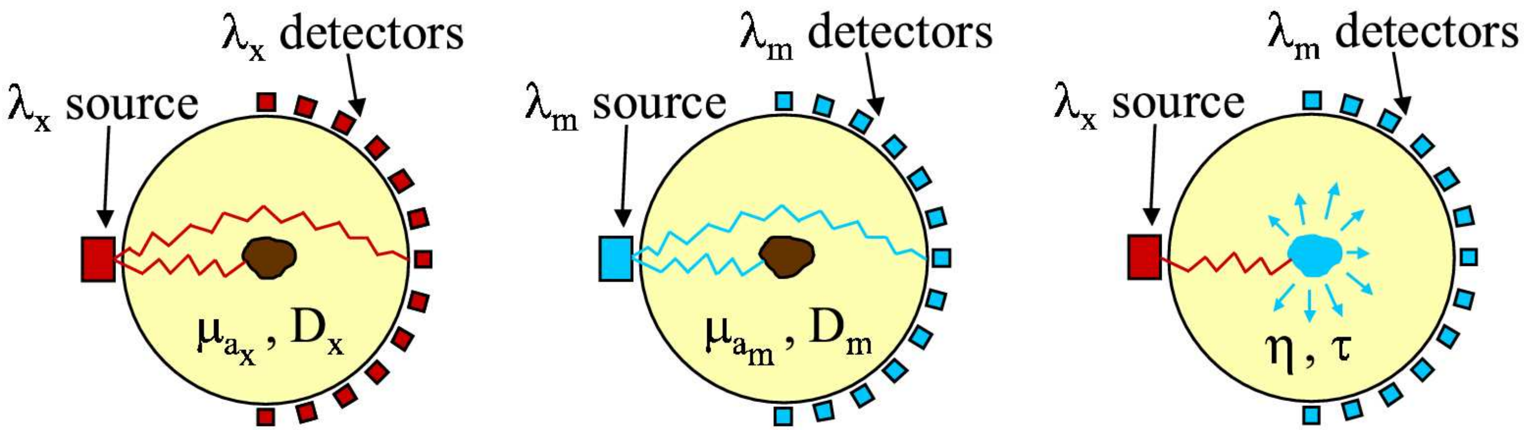
Fluorescent re-emission term

$$\psi_x(r,t) * \left\{ \frac{\eta(r)}{\tau(r)} e^{-t/\tau(r)} \right\} = \phi_x(r) \frac{\eta(r)\left[1 - j\omega\tau(r)\right]}{1 + \left[\omega\tau(r)\right]^2} e^{j\omega t}$$

- Then the frequency modulated light obeys the coupled elliptic PDEs

$$\begin{aligned} \nabla \cdot D_x(r)\nabla\phi_x(r) - \left[\mu_x(r) + j\omega/c\right]\phi_x(r) &= -S_x(r) \\ \nabla \cdot D_m(r)\nabla\phi_m(r) - \left[\mu_m(r) + j\omega/c\right]\phi_m(r) &= -\phi_x(r)\frac{\eta(r)\left[1 - j\omega\tau(r)\right]}{1 + \left[\omega\tau(r)\right]^2} - S_m(r) \end{aligned}$$

FODT: Measurement Approach



Source λ	Detector λ	Unknown Image	Observation
λ_x	λ_x	$x_x = [D_x, \mu_x]$	y_x
λ_m	λ_m	$x_m = [D_m, \mu_m]$	y_m
λ_x	λ_m	$x_f = [\eta, \tau]$	y_f

- Reconstruction strategy
  - Use y_x to determine x_x and ϕ_x
  - Use y_m to determine x_m
  - Use y_f, ϕ_x, and x_m to determine x_f

FODT: MAP Inversion

- For each image, joint MAP estimation of x and α

$$
\begin{aligned}
\hat{x}_x &= \arg \max_{x_x \geq 0, \alpha_x} \{ \, p(x_x | y_x, \alpha_x) \, \} \\
\hat{x}_m &= \arg \max_{x_m \geq 0, \alpha_m} \{ \, p(x_m | y_m, \alpha_m) \, \} \\
\hat{x}_f &= \arg \max_{x_f \geq 0, \alpha_f} \{ \, p(x_f | y_f, \alpha_f, \hat{x}_x, \hat{x}_m) \, \}
\end{aligned}
$$

- Estimation of α improves robustness of convergence
- Iterative optimization scheme:
  alternate updates with respect to α and x

$$\text{Step 1: } \hat{\alpha} = \frac{1}{P} ||y - f(\hat{x})||^2_\Lambda$$

$$\text{Step 2: } \hat{x} = \arg \min_{x \geq 0} \left\{ \frac{1}{2\hat{\alpha}} ||y - f(x)||^2_\Lambda + \frac{1}{p\sigma^p} \sum b_{i-j} |x_i - x_j|^p \right\}$$

FODT: Multifrequency: Inversion of η and τ [110]

- Problem: η and τ are nonlinearly related to fluorescent re-emission

$$\phi_x(r)\frac{\eta(r)\left[1 - j\omega\tau(r)\right]}{1 + \left[\omega\tau(r)\right]^2}$$

- Solution for single frequency case
  - Reparameterize $\{\eta, \tau\}$ as $\{\gamma, \tau\} = \{\frac{\eta}{1+[\omega\tau]^2}, \tau\}$
  - $\{\gamma, \tau\}$ are linearly related to fluorescent re-emission
- Solution for multiple frequency case
  - Parameterize using $\{\eta, \tau\}$
  - Use parametric ICD (PICD) to decouple nonlinearity
  - Compute θ_{1,ω_q} and θ_{2,ω_q} for frequencies $\omega_1, \cdots, \omega_q$
  - Perform numerical optimization of η τ at each voxel

FODT: Multifrequency: Parametric ICD (PICD) Optimization[110]

$[\omega_1, \cdots, \omega_Q]$ - the Q different frequencies used

$x_s = [\eta_s, \tau_s]$ - the unknown values at voxel s

- Define the functions

$$h(x_s) \triangleq \left[\frac{\eta_s(1 - j\omega_1\tau_s)}{1 + (\omega_1\tau_s)^2}, \cdots, \frac{\eta_s(1 - j\omega_Q\tau_s)}{1 + (\omega_Q\tau_s)^2} \right]^t$$

$$\Delta h(\tilde{x}_s, x_s) \triangleq h(\tilde{x}_s) - h(x_s)$$

- Generalize the concept of θ_1 and θ_2 to

$$\theta_1 \triangleq [\theta_{1,1}, \cdots, \theta_{1,Q}] \qquad \theta_2 \triangleq [\theta_{2,1}, \cdots, \theta_{2,Q}]$$

- Then update of voxel $x_s = [\eta_s, \tau_s]$ is given by:

$$x_s \leftarrow \arg\min_{\tilde{x}_s \geq 0} \left\{ \theta_1 \Delta h(\tilde{x}_s, x_s) + \frac{1}{2} \|\Delta h(\tilde{x}_s, x_s)\|^2_{\theta_2} + \frac{1}{p\sigma^p} \sum_{j \in \mathcal{N}_s} |\tilde{x}_s - x_j|^p \right\}$$

FODT: Multifrequency: Mutual Information Design Metric [110]

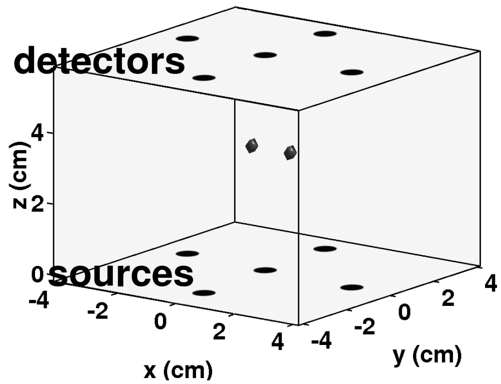
- Previously, singular value analysis has been used to evaluate measurement geometries[111]
- Information theory provides statistical framework for evaluating tomography system and reconstruction quality
- Mutual information concept: $I(X;Y)$ is the reduction in uncertainty of X due to knowledge of Y
- Consider the model

$$\begin{aligned} p_X(x) &= \frac{1}{\sqrt{(2\pi)^N |C|}} \exp\left\{-\frac{1}{2}||x||_C^2\right\} \\ p_{Y|X}(y|x) &= \frac{1}{\sqrt{(2\pi\alpha)^P |\Lambda^{-1}|}} \exp\left\{-\frac{1}{2\alpha}||y - Ax||_\Lambda^2\right\} \end{aligned}$$

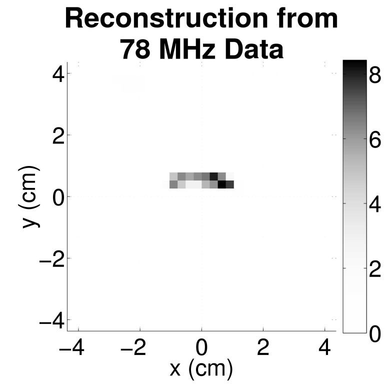
  A is the forward operator

- Then $I(X;Y) = \frac{1}{2}\log\left|I + \frac{1}{\alpha}\Lambda A C^{-1} A^H\right|$

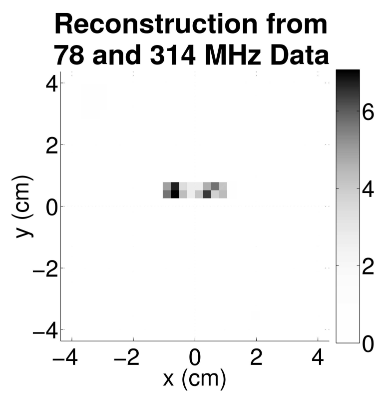
FODT: Multifrequency: Mutual Information Comparison



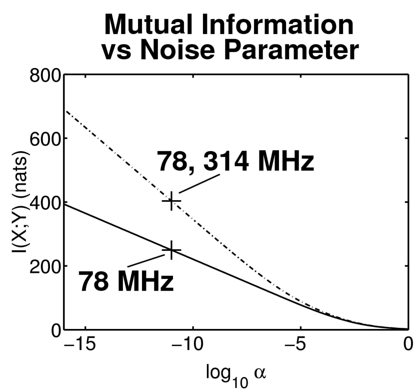
(a)



(b)



(c)



(d)

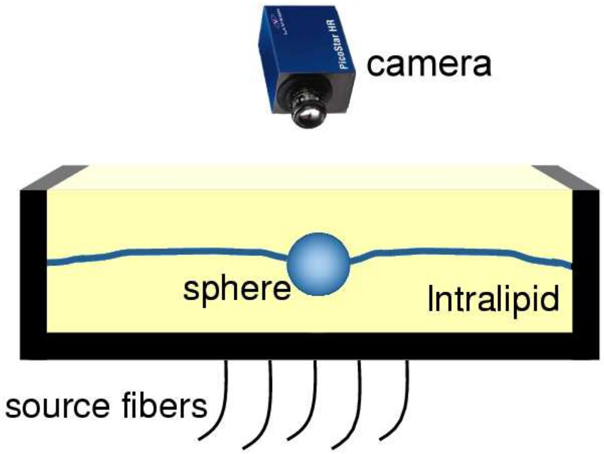
21.2dB avg. SNR at 78.4 MHz

14.7dB avg. SNR at 314 MHz

- Reconstructed 2 small, nearby objects from sparse, noisy data
- Multiple frequencies improves quality, as predicted by $I(X;Y)$

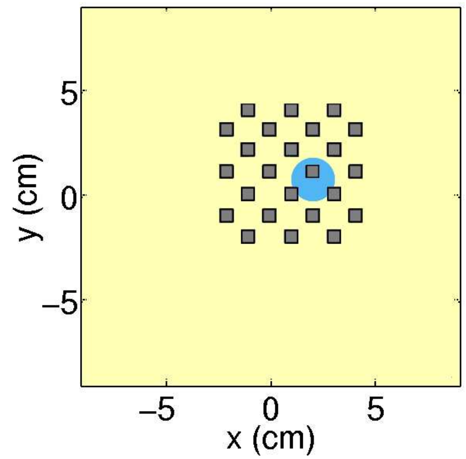
FODT: Experiment: Fluorophore in Lipid Suspension

- 0.125 μM indocyanine green (ICG) in 1% Intralipid suspension
- MaiTai tunable, ultrafast Ti:sapphire source, and PicoStar intensified CCD detector
- $\lambda_x = 780$ nm, $\lambda_m = 830$ nm



Instrument

Measurement Box



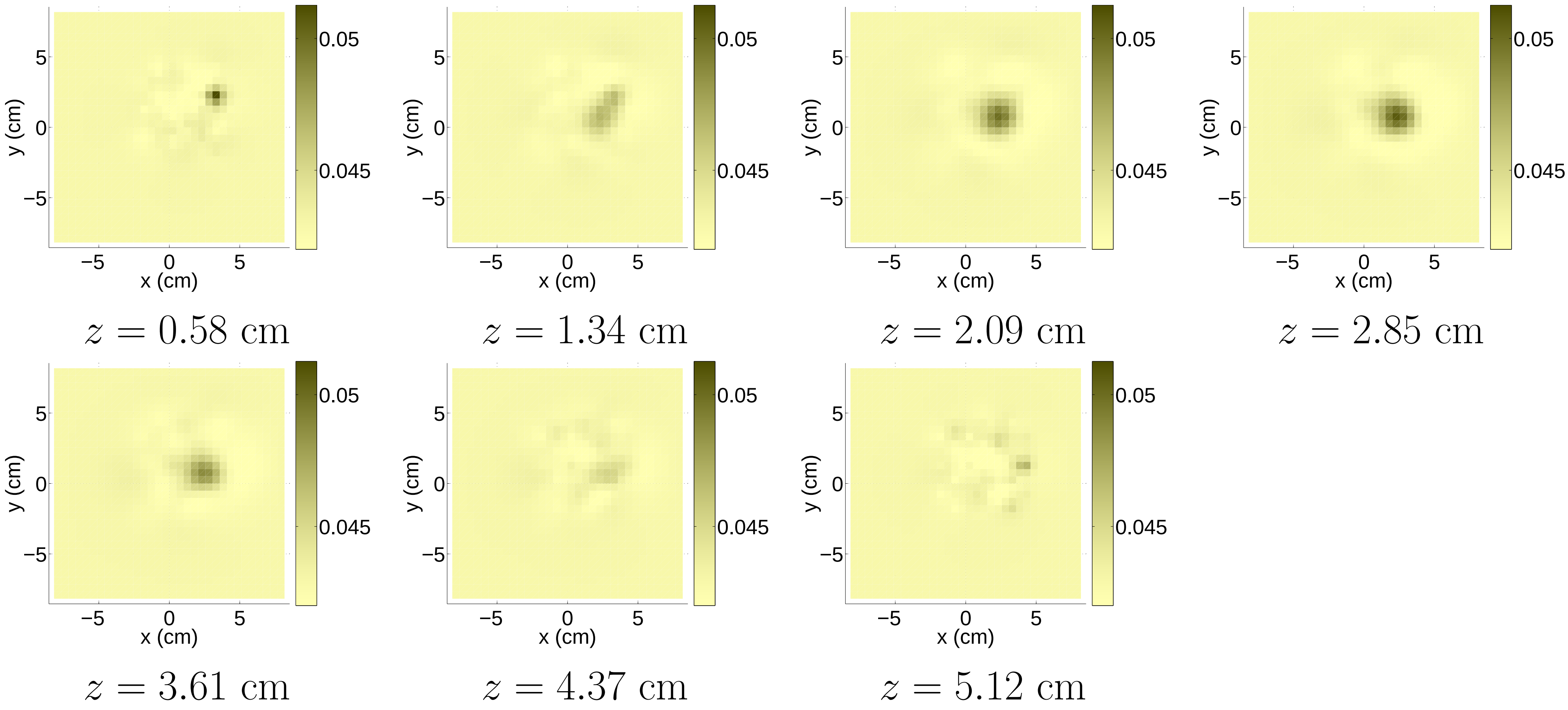
Source/Detection Positions

FODT: Experiment: Parameters

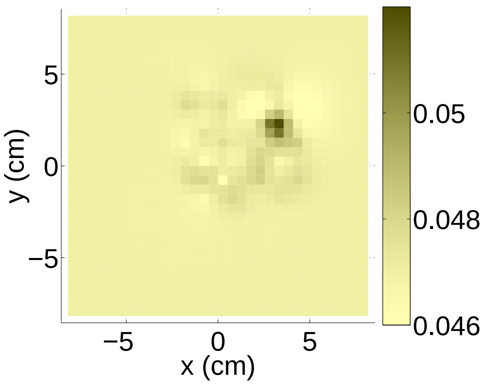
- Impulse responses from 24 source fibers, 24 detection positions
- $\mu_{a_{x,m}}$ data: 250 ps time step, 75 ms acquisition time/measurement, lower intensifier gain
- η, τ data: 500 ps time step, 1 s acquisition time/measurement, higher intensifier gain
- From FFT, 78, 314 MHz, and 627 MHz data used
- 16 cm $\times$ 16 cm $\times$ 5.7 cm domain modeled
- $p = 2.0$ for all reconstructions

σ used	Parameter
0.005 cm^{-1}	μ_a
0.1 ns	τ
0.00005 AU	η

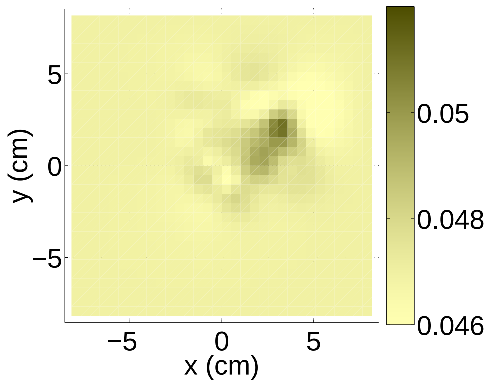
FODT: Experiment: Results - μ_x (cm^{-1})



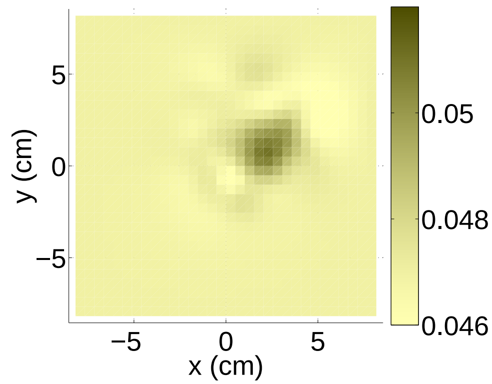
FODT: Experiment: Results - μ_m ($\mathbf{cm}^{-1}$)



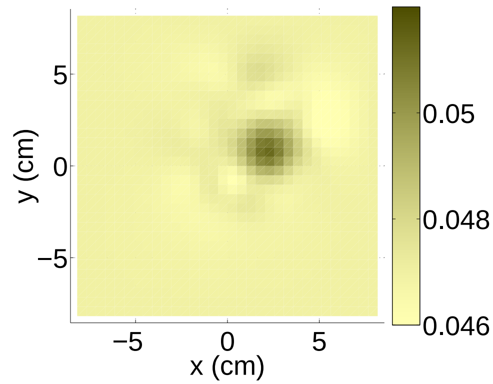
$z = 0.58$ cm



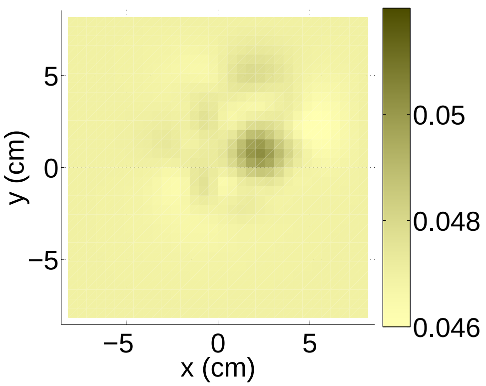
$z = 1.34$ cm



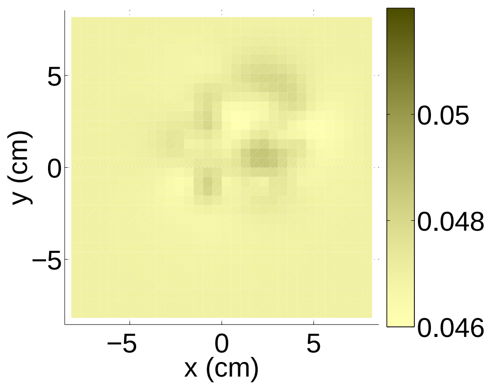
$z = 2.09$ cm



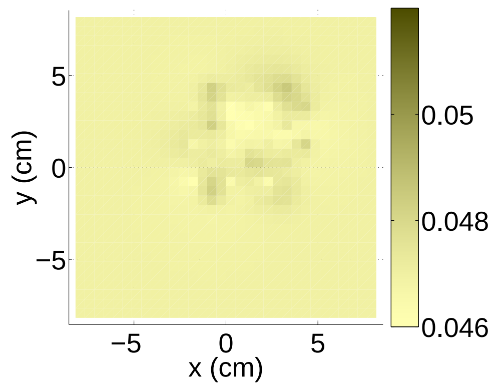
$z = 2.85$ cm



$z = 3.61$ cm

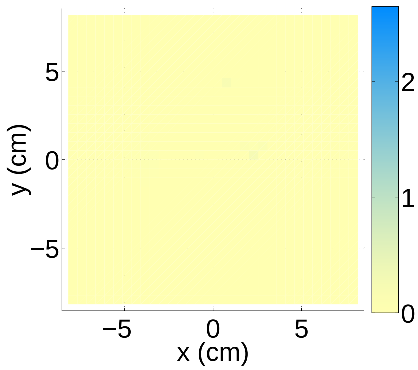


$z = 4.37$ cm

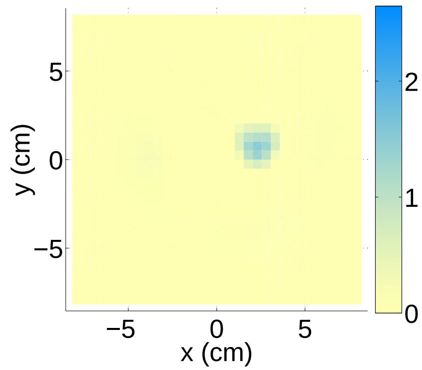


$z = 5.12$ cm

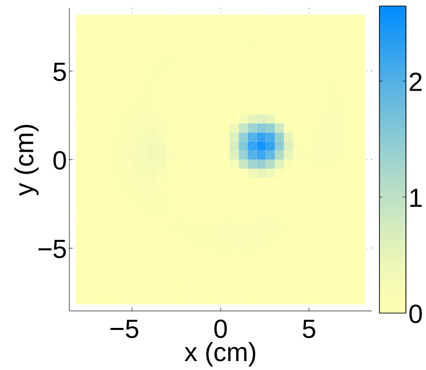
FODT: Experiment: Results - η (AU)



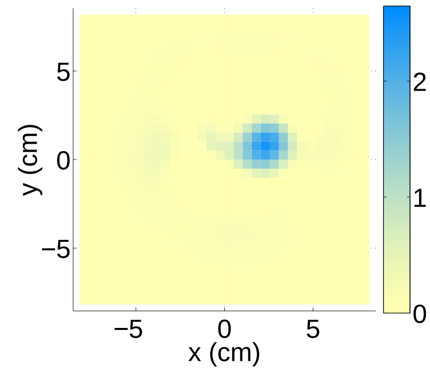
$z = 0.58$ cm



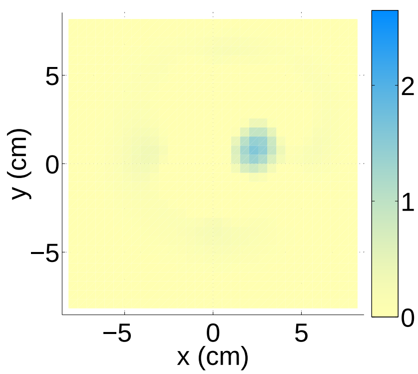
$z = 1.34$ cm



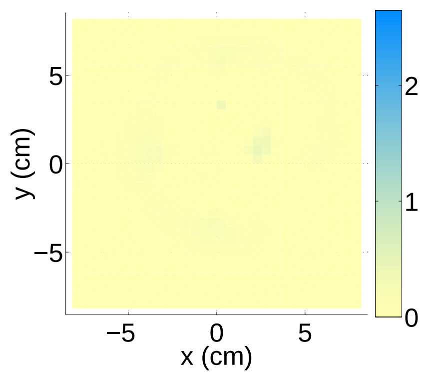
$z = 2.09$ cm



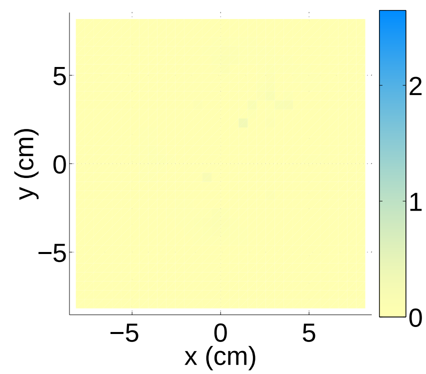
$z = 2.85$ cm



$z = 3.61$ cm

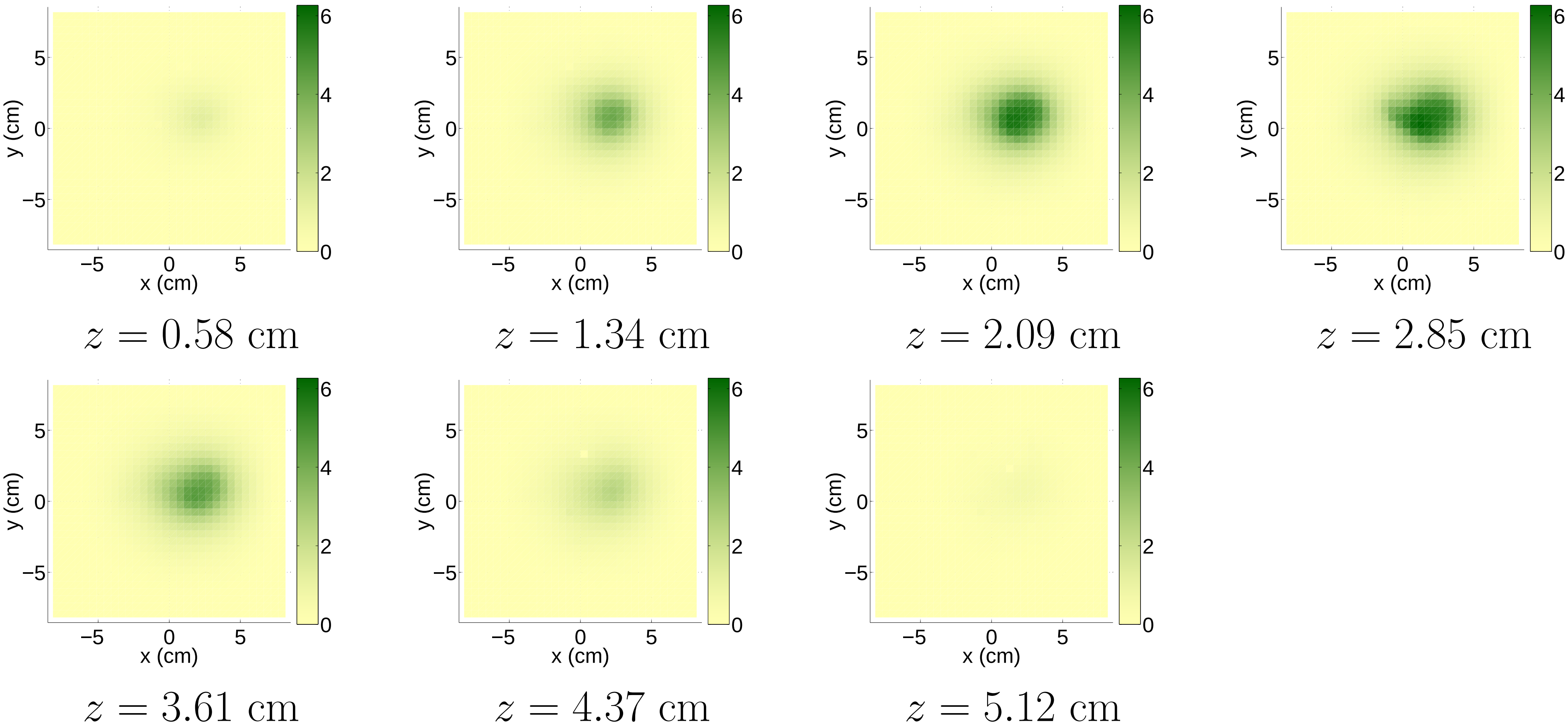


$z = 4.37$ cm



$z = 5.12$ cm

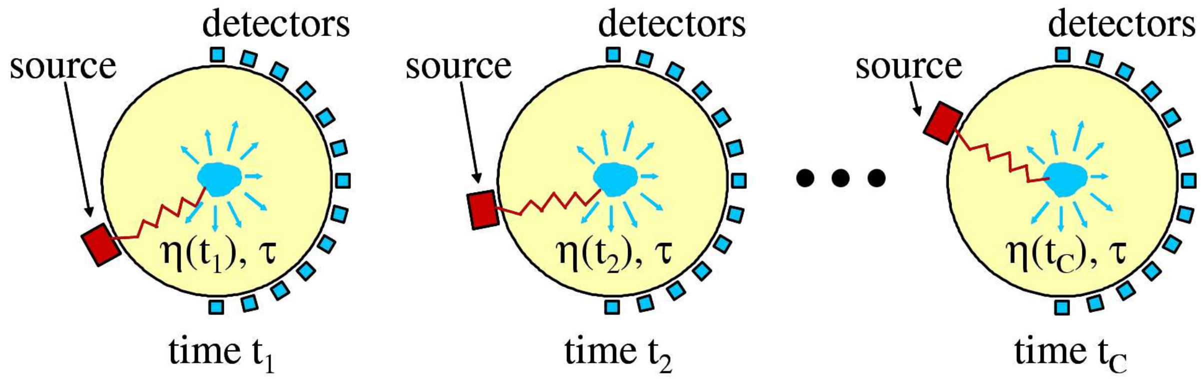
FODT: Experiment: Results - τ $(1 \times 10^{-10}$ s)



Kinetic Parameter Models in FODT

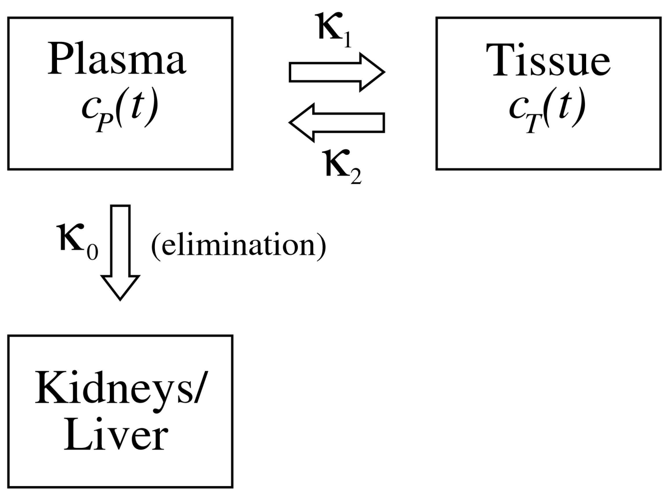
- Recent advances in targeted fluorescent probes have paved the way for molecular imaging
- Several groups have measured dynamic behavior of optical contrast agents in human or animal subjects [112, 113]
- Uptake rate of targeted dyes may form a basis for distinguishing diseased and healthy tissue

Kinetic FODT: Approach [114, 115]



- Problem
  - Would like to extract kinetic parameters for each voxel
  - Source position is typically moving
  - Insufficient data to reconstruct at each time
- Solution: Directly reconstruct kinetic parameters for each voxel
- Advantages
  - Only reconstruct one image of parameters
  - Account for time varying illumination
  - Reduced number of parameters
  - Increased SNR

Kinetic FODT: Compartment Model for Each Voxel



$$\begin{aligned} c_P &= A\exp(-\kappa_3 t) \\ \frac{dc_T}{dt} &= \kappa_1 c_P - \kappa_2 c_T \end{aligned}$$

c_P: concentration in plasma	c_T: concentration in tissue
κ_1: rate to enter tissue	κ_2: rate to leave tissue
κ_3: exponential parameter for c_P	κ_O: rate of elimination into kidneys

- $\eta(r,t)$ is proportional to a weighted sum of c_P and c_T, with the form:

$$\eta(r,t) = \gamma_1(r)\exp[-\gamma_4(r)t] - \gamma_2(r)\exp[-\gamma_3(r)t].$$

- Parameters for voxel are

$$x_s = [\gamma_{1,s}, \gamma_{2,s}, \gamma_{3,s}, \gamma_{4,s}]$$

Kinetic FODT: Bayesian Approach with PICD Optimization

- At each time t_c, forward operator is linear with respect to $h_{t_c}(x) = \frac{\eta(t_c)}{1+j\omega\tau}$,
- Optimization of x may be written:

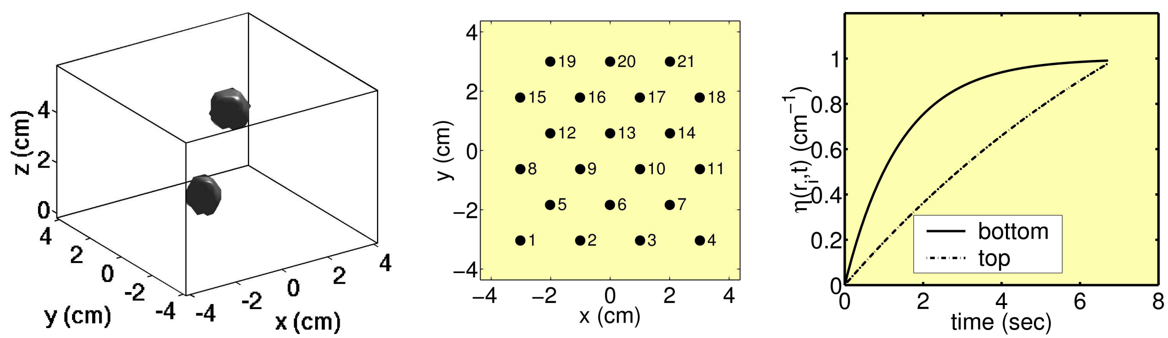
$$\hat{x} = \arg\min_x \left\{ \frac{1}{2\hat{\alpha}} \sum_{c=1}^{C} ||y_{t_c} - J_{t_c} h_{t_c}(x))||^2_{\Lambda_{t_c}} - \log p(x) \right\}$$

- We update voxels of τ and $\gamma_1 \cdots \gamma_4$ using nonlinear line search, with constraints $x \geq 0$, $\gamma_1 \geq \gamma_2$, and $\gamma_3 \geq \gamma_4$:

$$\hat{x}_i \leftarrow \arg\min_{x_i} \left\{ \sum_{c=1}^{C} \left(\theta_{1,t_c}[h_{t_c}(x_i)] + \frac{1}{2}\theta_{2,t_c}[h_{t_c}(x_i)]^2 \right) - \log p(x_i) \right\}$$

Simulation - Two Objects with Different Uptake

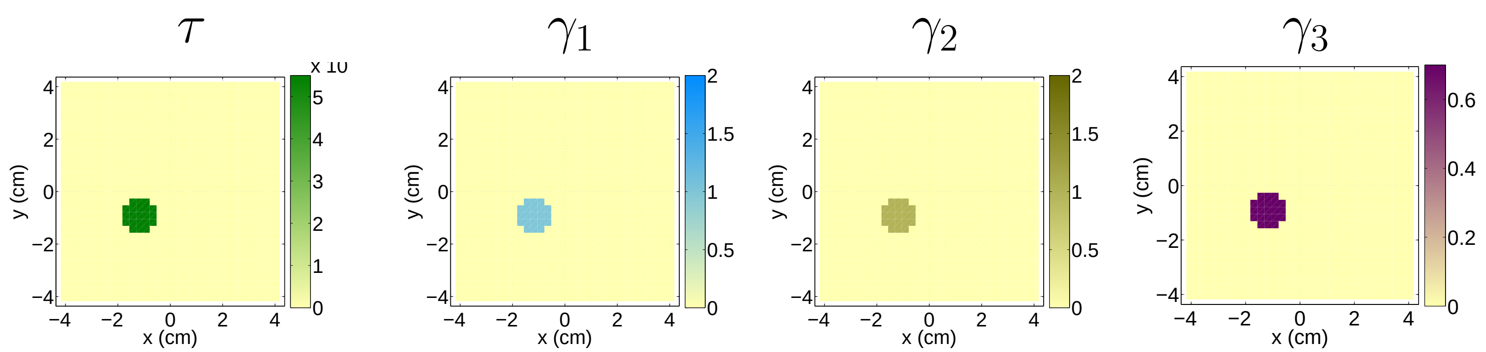
- Synthetic time series of data generated from cube-shaped phantom with two heterogenieties
- Objects have same τ, but different γ_1, γ_2, and γ_3.
- Sources illuminated one at a time in raster order, with each source used once
- Average signal/noise: 28 dB



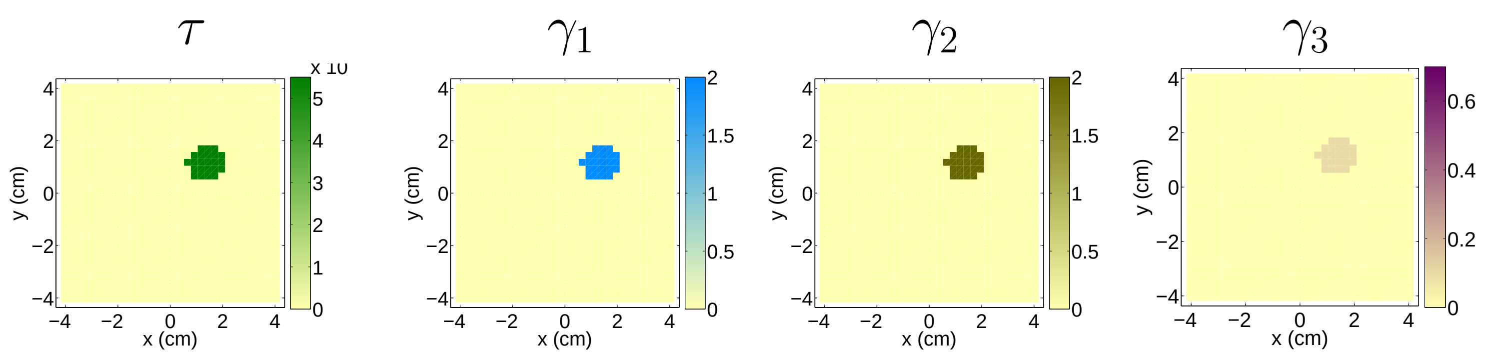
Two Objects — Sources — η vs. Time

Kinetic FODT: Simulation - True Phantom

Bottom: $z = 2.1$ cm

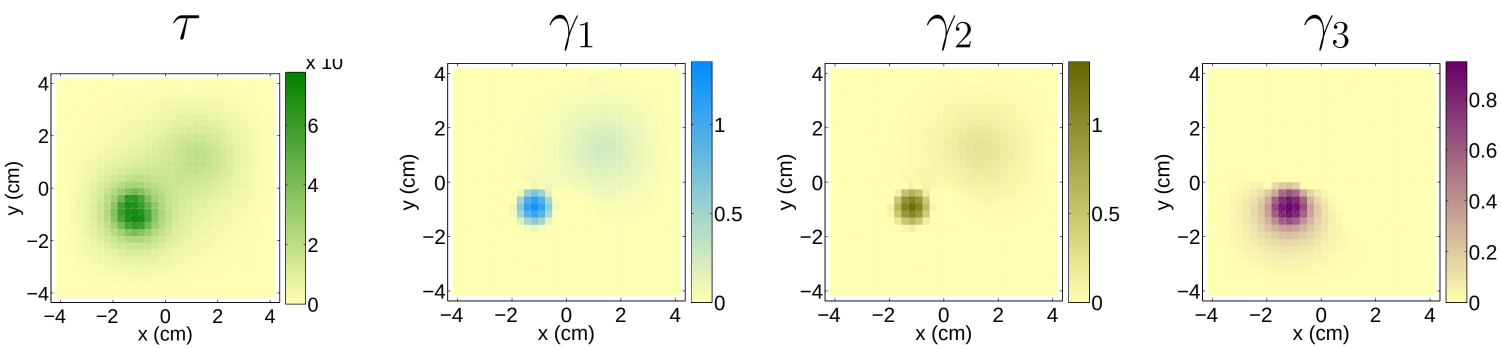


Top: $z = 4.0$ cm

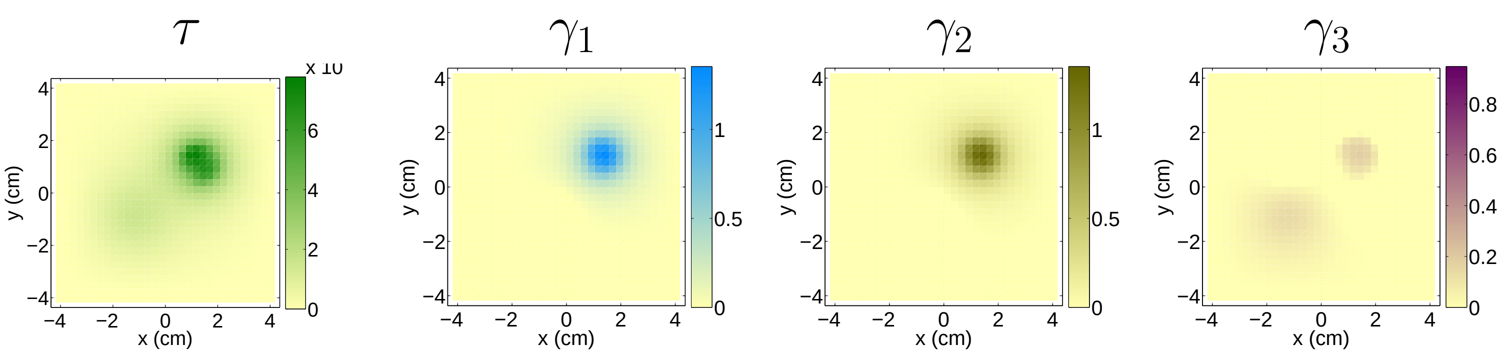


Kinetic FODT: Simulation - Reconstruction

Bottom:
$z = 2.1$ cm



Top:
$z = 4.0$ cm

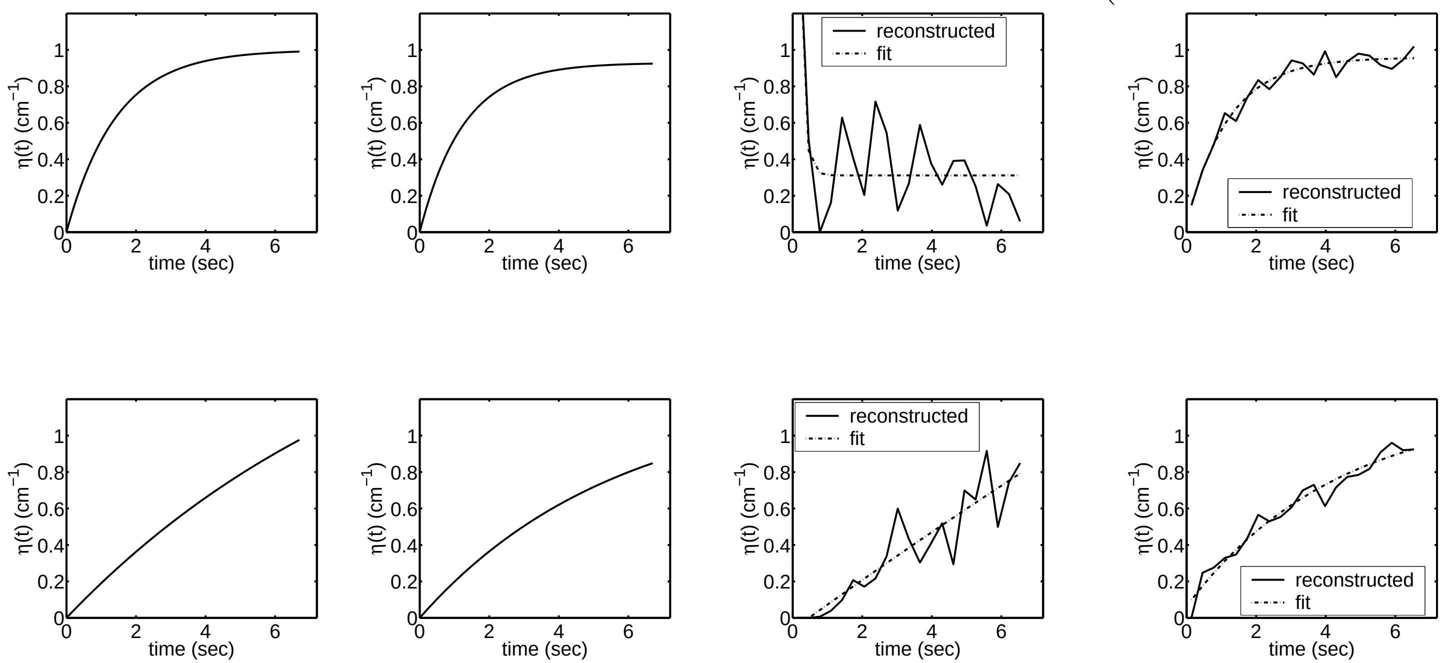


Kinetic FODT: Simulation - Plots of $\eta(t)$ and Reconstructions

True | Reconstruction | Frame-by-frame | Frame-by-frame ($21\times$ more data)

Bottom: $z = 2.1$ cm

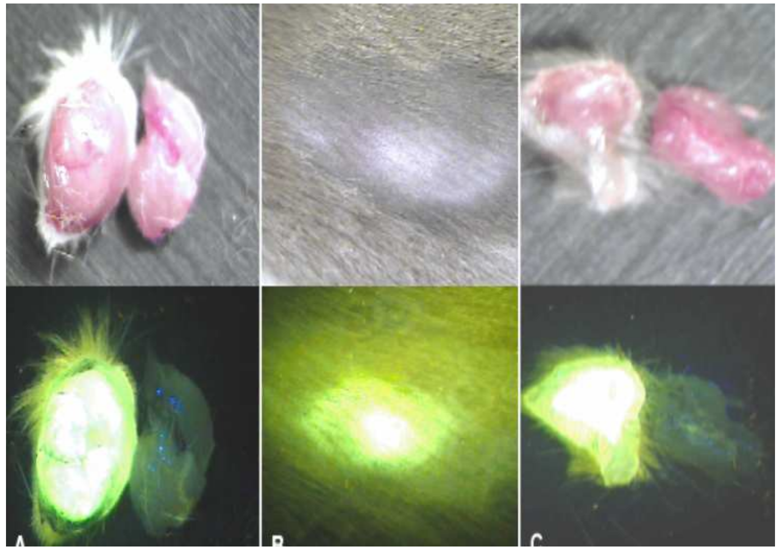
Top: $z = 4.0$ cm



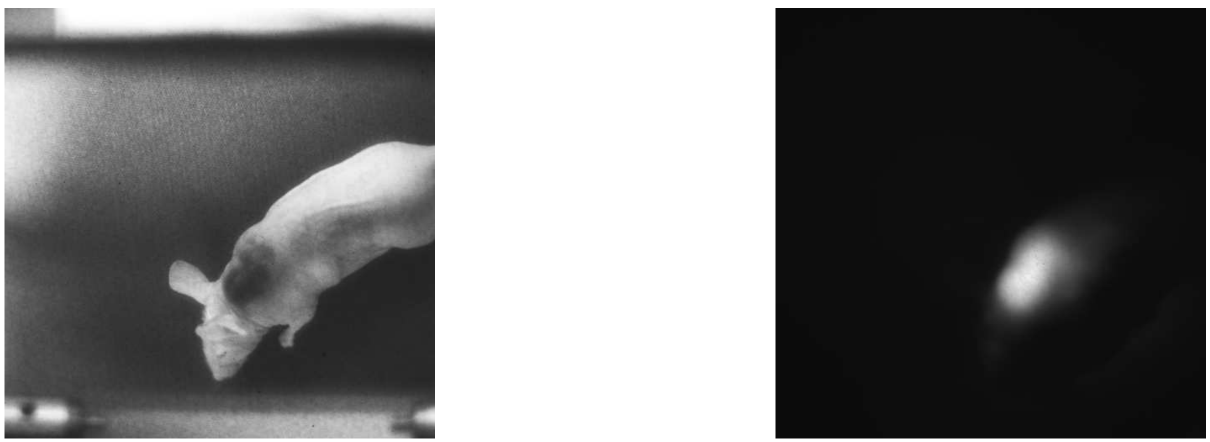
Targeted Contrast Agents

- Clinical diagnostic imaging often relies on different uptake behavior between tumors and the surrounding tissue
- Non-targeted dyes may accumulate in tumor due to increased vascular density or capillary permeability [116]
- Some imaging agents specifically target certain receptors which are overexpressed in malignant cells
- Examples of targeting ligands for delivery of diagnostic imaging agents include antibodies[117], hormones [118], small peptides [119], and folic acid [120]

Folate-Targeted Fluorescent Agents[120, 121]

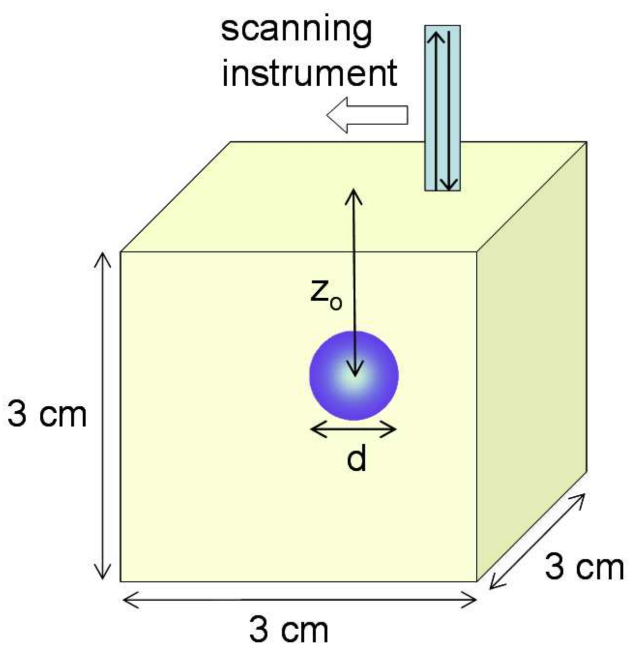


Mouse tumors containing folate-fluorescein



Mouse tumors containing folate-indocyanine

Folate Targeting: Simulation Geometry

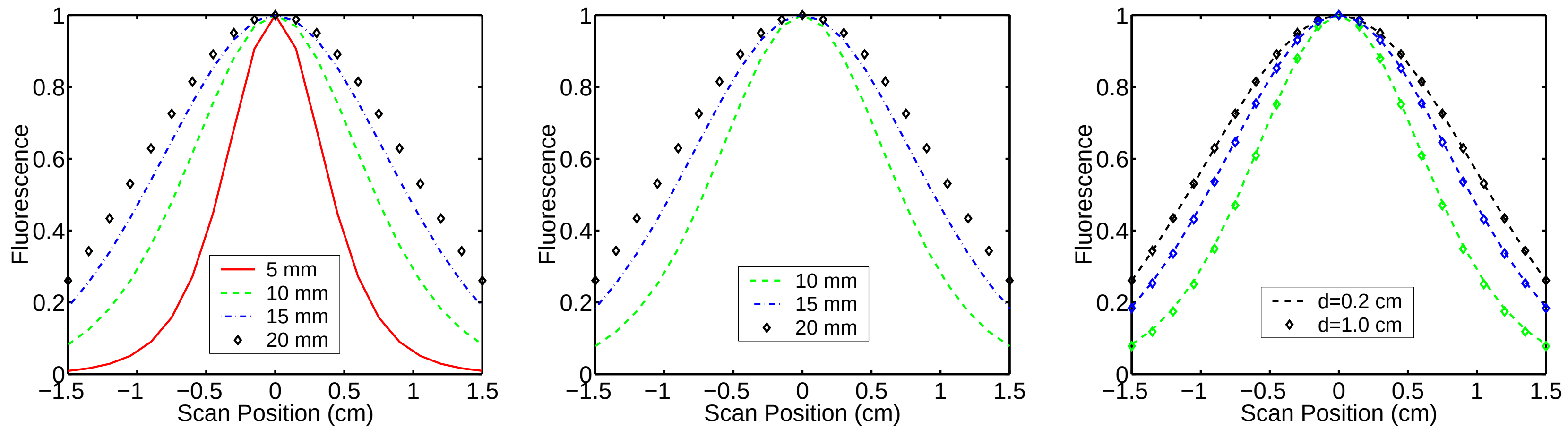


- Concept: a fiber optic probe incorporated into a surgical instrument
- Patient injected with targeted fluorescent contrast agent
- Two fibers for fluorescent measurement: excitation and collection
- Scan measurement and mathematical model allow reconstruction of tumor position

Folate Targeting: Simulation: Effect of Tumor Depth and Size

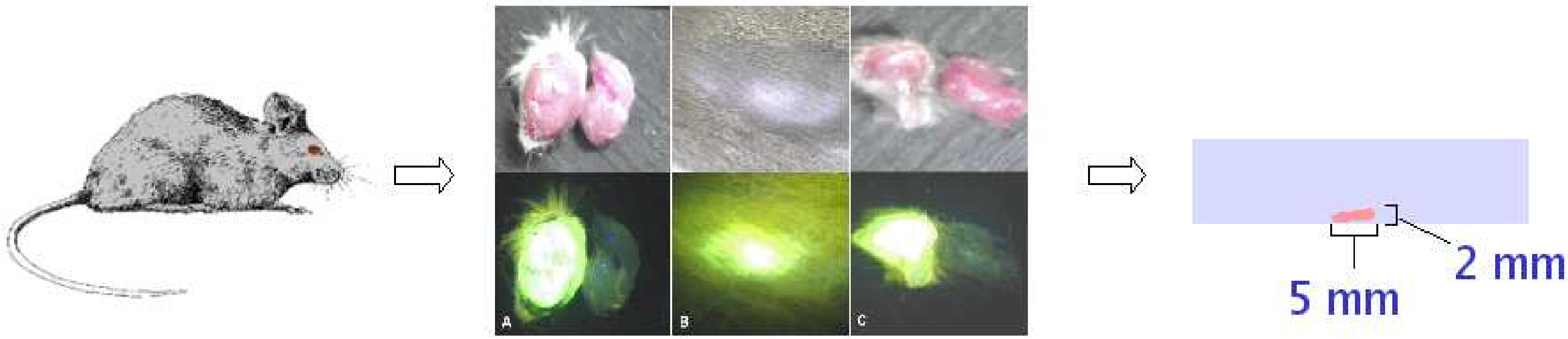
Normalized Fluorescence Scans for Several Tumor Depths

2mm Diameter   1 cm Diameter   2 mm, 1 cm Comparison



- Simulated tissue: $\mu_a = 0.05$ cm^{-1}, $\mu_s' = 1/3D = 9$ cm $^{-1}$
- Source and detector separated by 2 mm scanned over surface
- Result: normalized scan data depend strongly on depth, and weakly on tumor size

Folate Targeting: Mouse Tumors Grown[121]



- Mouse induced to grow lung tumor and injected with 10 nmols of folate-fluorescein
- Portion of tumor excised and frozen for later experimental use
- Before measurement, tumor thawed and glued to Petri dish

Folate Targeting: Tumor Localization: Fit to Model

- Tumor position estimated by fitting measurements to diffusion model
- Computational domain with $65 \times 65 \times 65$ volume elements simulates volume of size 6 cm $\times$ 6 cm $\times$ 4 cm
- Assumption: $\mu_a = 0 \text{ cm}^{-1}$, $\mu_s' = 1/3D = 15 \text{ cm}^{-1}$

$$
\begin{aligned}
\mathbf{r}^* &= \arg\min_{\mathbf{r}} c(\mathbf{r}) \\
c(\mathbf{r}) &= \min_{w} \left\{ \sum_{k=1}^{K} \frac{[y_k - w f_k(\mathbf{r})]^2}{y_k} \right\},
\end{aligned}
$$

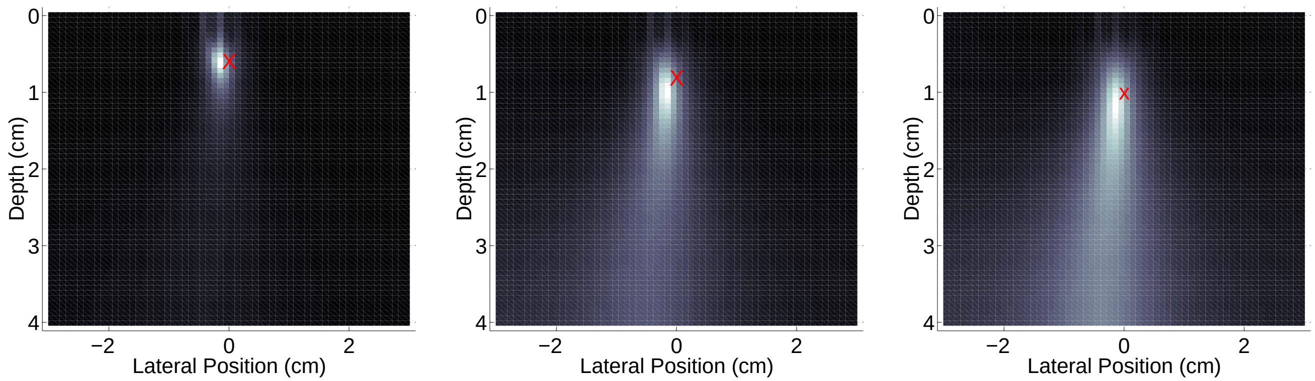
$\mathbf{r}^*$: estimated tumor position

$f_k(\mathbf{r})$: computed data for k^{th} scan position and point source tumor at $\mathbf{r}$

y_k: k^{th} measurement in scan

w: normalizing weight

Folate Targeting: Results



- $\log c(\mathbf{r})$ plots show estimated tumor position
- Uncertainty in estimate increases with depth
- Despite limited data, horizontal and vertical positions recovered accurately within 2.1 mm

Folate Targeting: Summary of Tumor Localization Experiment

- While preliminary, results demonstrate that folate imaging agents coupled with a scan measurement can accurately localize an obscured tumor
- In principle, a diagnostic instrument could perform a similar measurement
- We aim to develop technology that will allow imaging of tumor loci in folate receptor-expressing tumors

Optical Speckle

- Scattering of coherent light creates speckle $\rightarrow$ information on scattering medium [122, 123]
- The temporal response is an important parameter for characterizing a random medium
- The temporal response for co-polarized light can be extracted from third order speckle intensity correlation measurements [124]
- Intensity correlations for cross-polarized light can provide additional information for weakly scattering media

[122] Thompson, Webb, and Weiner, *Applied Optics,* vol. 36, June 1997.

[123] McKinney, Webster, Webb and Weiner, *Optics Letters,* vol. 25, no. 1, Jan. 2000.

[124] Webster, Webb, and Weiner, *Phys. Rev. Letters*, vol. 88, no. 3, Jan. 2002.

Optical Speckle: Diffusion Approximation

- The diffusion equation is

$$\frac{1}{c}\frac{\partial \psi}{\partial t}(\vec{r},t) - \nabla \cdot D \nabla \psi(\vec{r},t) + \mu_a \psi(\vec{r},t) = S(\vec{r},t)$$

with $D = \dfrac{1}{3(\mu_s' + \mu_a)}$, where μ_s' - scattering coefficient
μ_a - absorption coefficient

- Measurable quantity is the photon current density

$$J_n(\vec{r},t) = -D\nabla\psi(\vec{r},t) \cdot \ \hat{\mathbf{n}}$$

with $\hat{\mathbf{n}}$ = unit normal at measurement surface

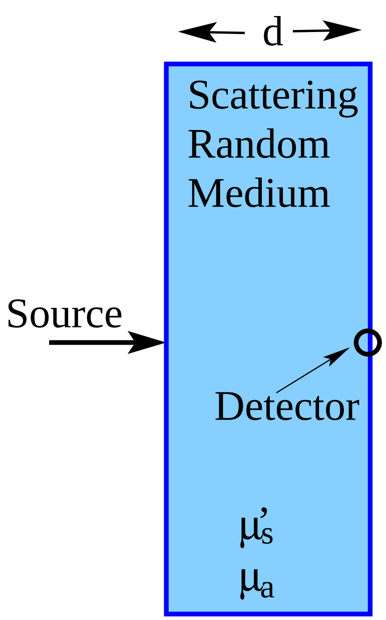
- For a homogeneous region, with $S(\vec{r},t) = \delta(\vec{r},t)$

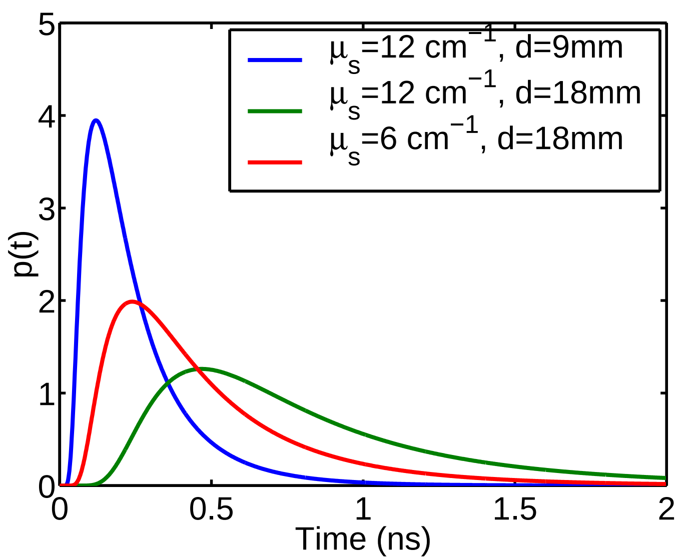
$$\psi_{\text{homo}}(\vec{r},t) = \frac{c}{(4\pi D c\, t)^{m/2}} \exp(-\mu_a c\, t) \exp\left[\frac{-r^2}{4Dct}\right]$$

Optical Speckle: Photon Transit-Time Distribution

- $J_n(\vec{r},t)$ = measured photon current density at detector
- For a source of $\delta(\vec{r},t)$ the photon transit-time distribution is

$$\boxed{p(\vec{r},t) \equiv \frac{J_n(\vec{r},t)}{\int_0^\infty dt\, J_n(\vec{r},t)}}$$





Optical Speckle: Laser Speckle from Diffuse Media

- $S(\nu)$ - normalized power spectral density of the laser source
- $p(t)$ - photon transit-time distribution of diffuse media
- Speckle intensity correlation

$$\frac{\langle I(\nu)I(\nu+\Delta\nu)\rangle}{\langle I(\nu)\rangle^2} = 1 + |P(\Delta\nu)|^2$$

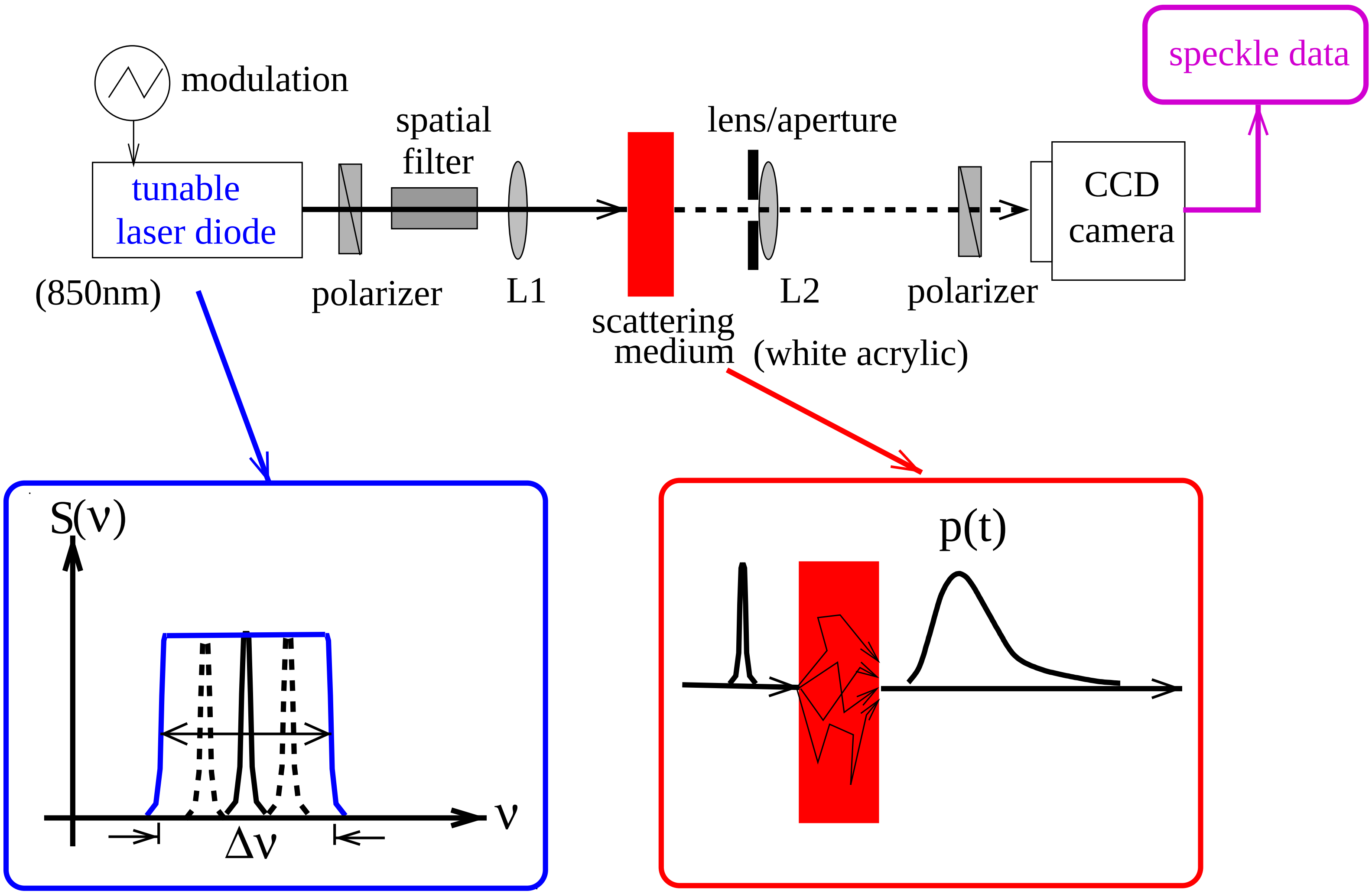
where

$$P(\Delta\nu) = \int_0^\infty dt\, p(t) \exp(-j2\pi\Delta\nu t)$$

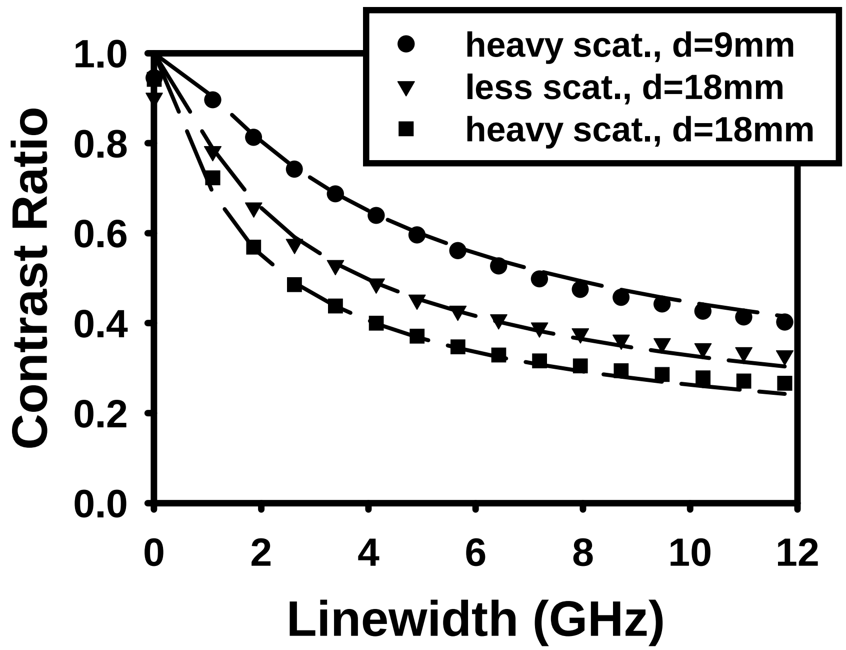
- Speckle contrast ratio

$$\frac{\sigma_I}{\mu_I} = \left\{\int_0^\infty d\nu_1 \int_0^\infty d\nu_2\, S(\nu_1)S(\nu_2)|P(\nu_1-\nu_2)|^2\right\}^{1/2}$$

Optical Speckle: Experimental Setup



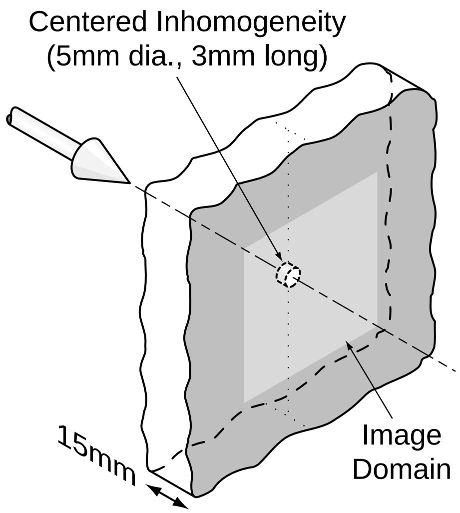
Optical Speckle: Measured Results for Homogeneous Slab



- Analytic $p(t)$ from slab model with μ_s' as variable ($\mu_a = 0$)
- d = slab thickness
- Speckle data acquired from a $1\,\mathrm{mm} \times 1\,\mathrm{mm}$ imaging area

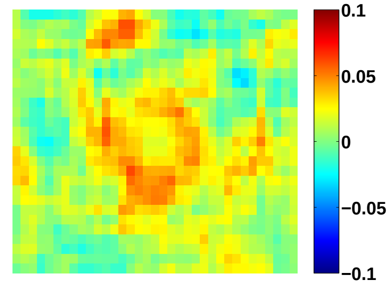
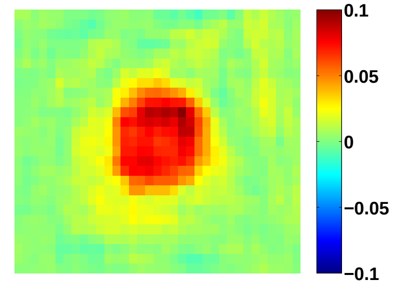
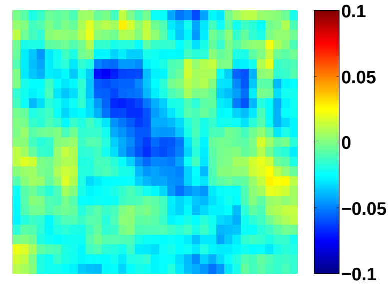
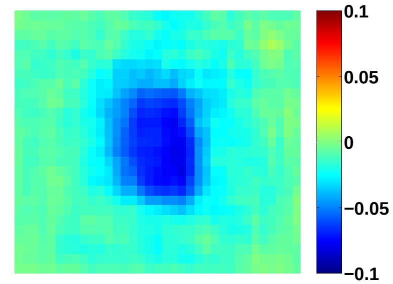
Optical Speckle: Inhomogeneous Scattering Medium Experiments

Sample geometry



Imaging area of
$45\,\mathrm{mm} \times 36\,\mathrm{mm}$

Contrast ratio difference

	$(\Delta\nu = 5\,\mathrm{MHz})$	$(\Delta\nu = 17\,\mathrm{GHz})$
void		
dense		

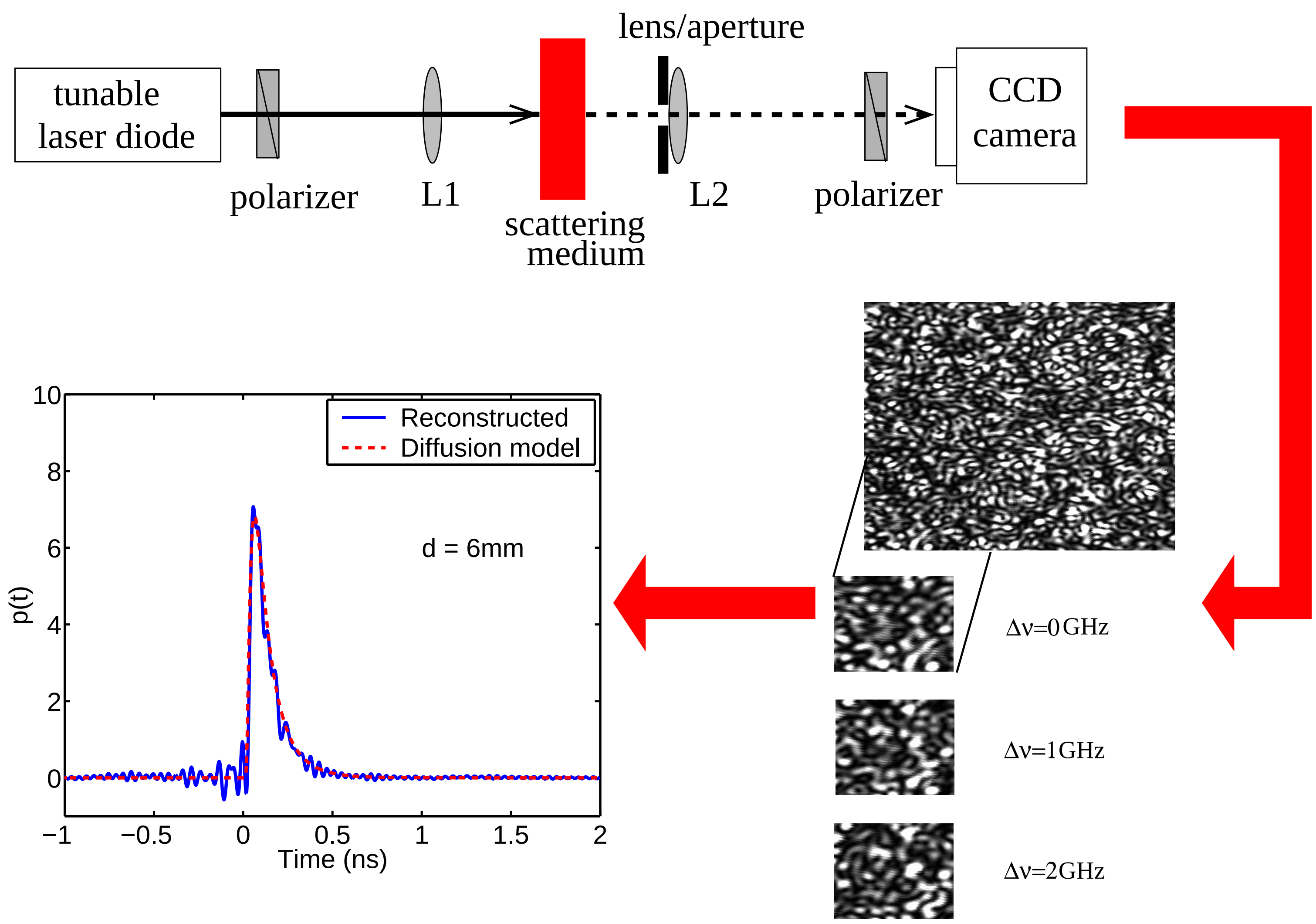
Optical Speckle: Using Speckle Data For Optical Diffusion Imaging

- Optical diffusion imaging algorithms require $J_n(\vec{r}, t)$ or $\tilde{J}_n(\vec{r}, \omega)$
- Recalling

$$p(\vec{r}, t) \equiv \frac{J_n(\vec{r}, t)}{\int_0^\infty dt \; J_n(\vec{r}, t)} \quad \Longrightarrow \quad J_n(\vec{r}, t) = \mu_I(\vec{r}) p(\vec{r}, t)$$

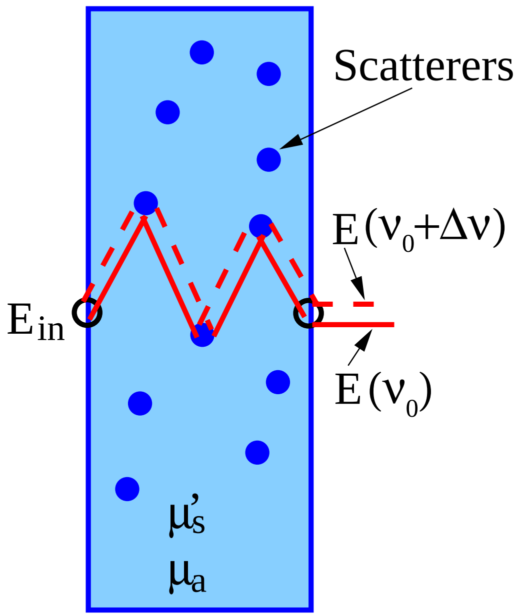
- Task is to determine $p(\vec{r}, t)$ from measured speckle data

Optical Speckle: Third Order Frequency Correlation



Optical Speckle: Circular Gaussian Fields

$$\boxed{\begin{aligned} \langle E(\nu_0+\Delta\nu)E^*(\nu_0)\rangle &= \langle I(\nu_0)\rangle \int_0^\infty dt\, p(t) \exp\left(-j2\pi\Delta\nu t\right) \\ &= \langle I(\nu_0)\rangle P(\Delta\nu) \end{aligned}}$$



- Fourier relation between *field* correlation and $p(t)$ [125]
- $\langle E(\nu_0 + \Delta\nu)E^*(\nu_0)\rangle$ is difficult to measure at optical frequencies
- $\langle I(\nu_0 + \Delta\nu)I(\nu_0)\rangle$ is convenient to measure

Optical Speckle: Speckle Intensity Correlations

- Intensity measurements at different frequencies

$$\boxed{I_1 = I(\nu_0), \quad I_2 = I(\nu_0 + \Delta\nu_1), \quad I_3 = I(\nu_0 + \Delta\nu_1 + \Delta\nu_2)}$$

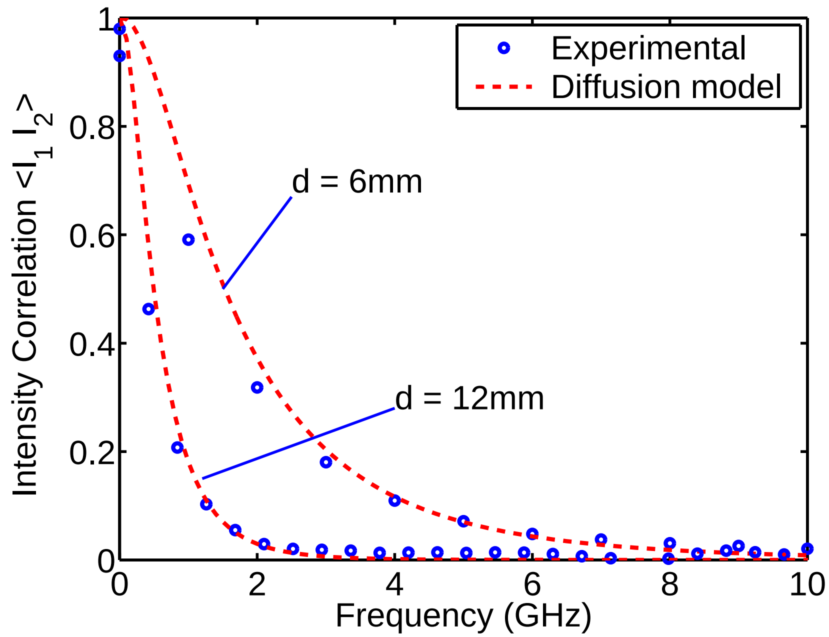
- Define normalized intensities $\boxed{\tilde{I}_j = \dfrac{I_j - \langle I \rangle}{\langle I \rangle}, \quad j = 1, 2, 3}$

- Second and third order intensity correlations

$$\boxed{\begin{aligned} \langle \tilde{I}_1 \tilde{I}_2 \rangle &= |P(\Delta\nu_1)|^2 \\ \langle \tilde{I}_1 \tilde{I}_2 \tilde{I}_3 \rangle &= 2\,\mathrm{Re}\{P(\Delta\nu_1) P(\Delta\nu_2) P^*(\Delta\nu_1 + \Delta\nu_2)\} \end{aligned}}$$

Optical Speckle: Second Order Correlations

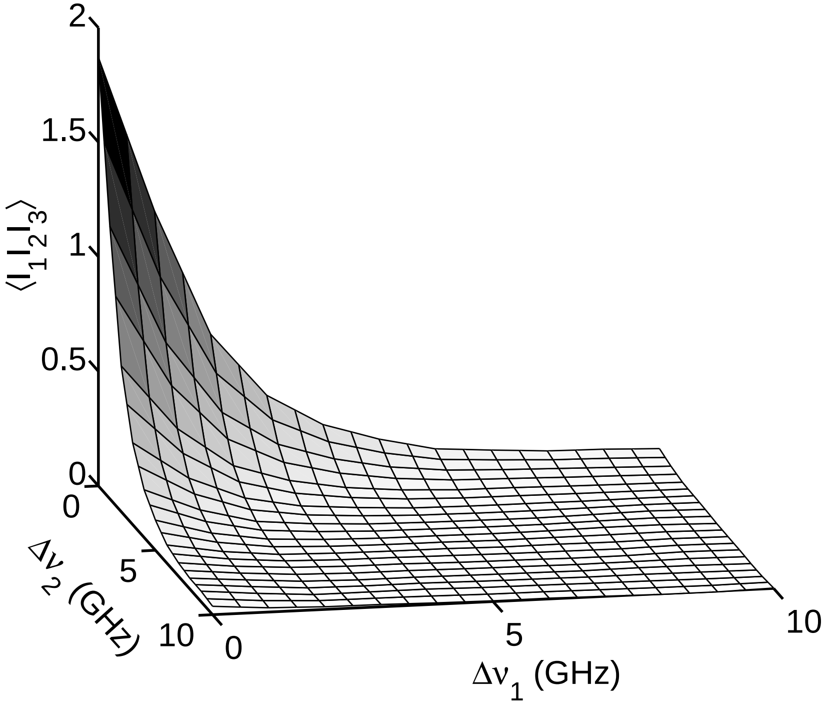
- Gives only magnitude of $P(\Delta\nu)$ (phase information is lost)
- Cannot obtain $p(t)$ without using *a priori* information



Second order correlations for homogeneous slabs with $\mu_s' = 15\,\mathrm{cm}^{-1}$

Optical Speckle: Third Order Correlation

$$\boxed{\langle \tilde{I}_1 \tilde{I}_2 \tilde{I}_3 \rangle = 2\,\mathrm{Re}\{P(\Delta\nu_1)P(\Delta\nu_2)P^*(\Delta\nu_1 + \Delta\nu_2)\}}$$

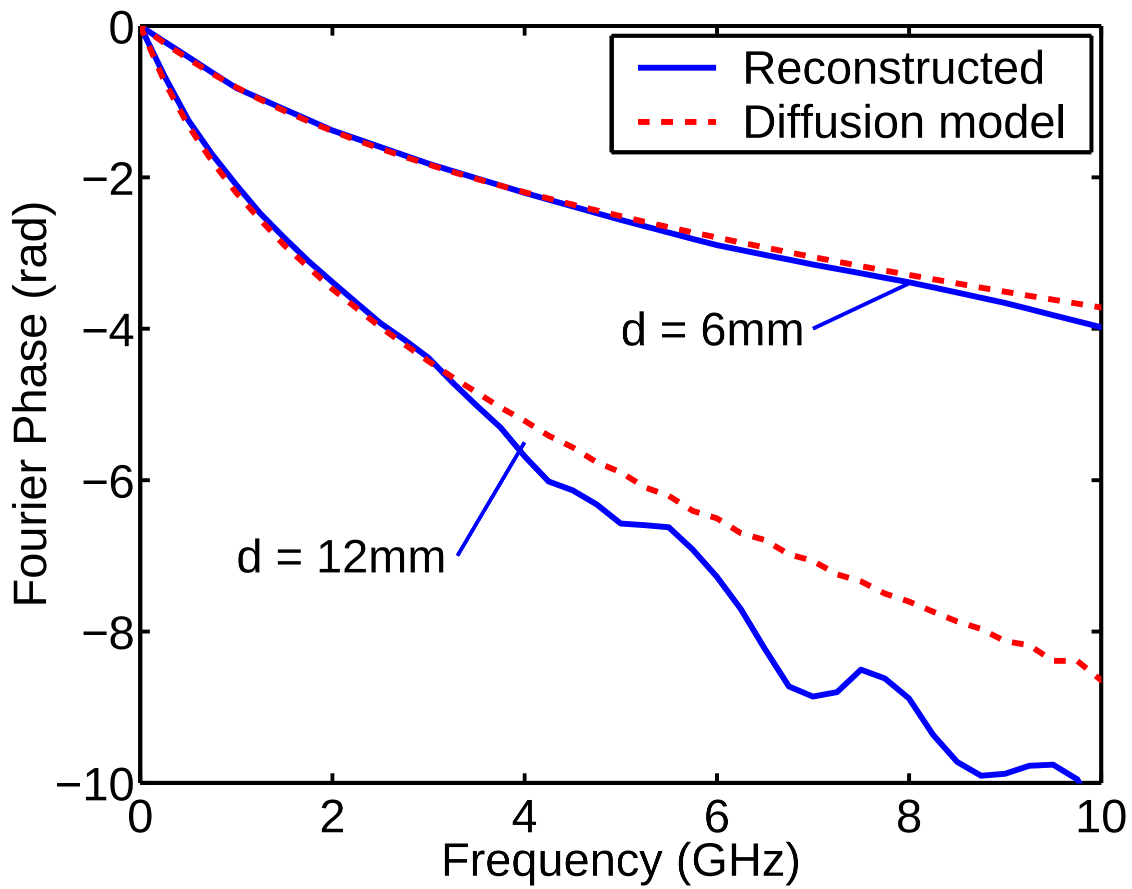


- Real part of bispectrum [126] of $p(t)$
- Contains Fourier phase information
- Half of the data in the plot is unique (from symmetry)
- Correlations averaged over all possible frequency combinations from experimental data set

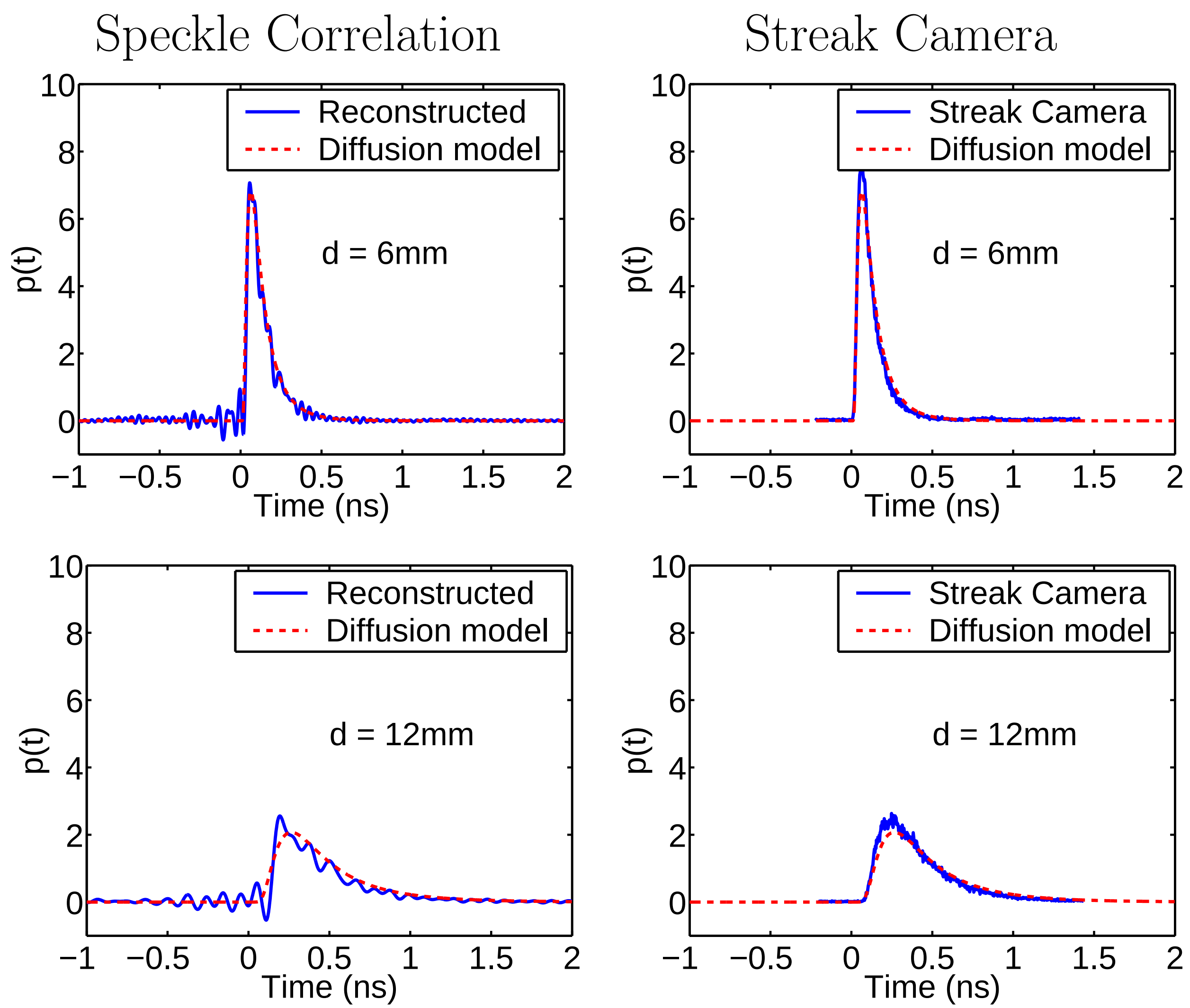
- For slab with thickness $d = 6\,\mathrm{mm}$ and $\mu_s' = 15\,\mathrm{cm}^{-1}$

Optical Speckle: Fourier Phase Reconstruction

- Bispectral techniques [127] allow for simple reconstruction of the Fourier phase
- Invariant to linear Fourier phase (arbitrary time-offset)



Optical Speckle: Co-Polarized Temporal Response [128]



Frontiers

- Spectral and temporal information $\longrightarrow$ better functional imaging, and early tumor detection
- Molecular imaging $\longrightarrow$ targeting with contrast agent
- Kinetic imaging $\longrightarrow$ pharmacokinetics
- In vivo imaging $\longrightarrow$ clinical developments (oxygen monitor is there)
- What will be the "Killer App"?

Thank You to Collaborators

Massachusetts General Hospital, Harvard Medical School

Prof. David Boas

Dr. Q. Zhang

Dr. J. Stott

Purdue University, Department of Chemistry

Prof. Philip S. Low

Dr. Michael Kennedy

University of Canterbury

Prof. Rick P. Millane

References

[1] R. M. P. Doornbos, R. Lang, M. C. Aalders, F. W. Cross, and H. J. C. M. Sterenborg, "The determination of *in vivo* human tissue optical properties and absolute chromophore concentrations using spatially resolved steady-state diffuse reflectance spectroscopy," *Phys. Med. Biol.*, vol. 44, pp. 967–981, 1999.

[2] F. F. Jobsis, "Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters," *Science*, vol. 198, pp. 1264–1267, 1977.

[3] A. E. Siegman, *Lasers.* University Science Books, 1986.

[4] D. Elson, S. Webb, J. Siegel, K. Suhling, D. Davis, J. Lever, D. Philips, A. Wallace, and P. French, "Biomedical applications of fluorescence lifetime imaging," *Opt. Photon. News*, pp. 26–32, Nov. 2002.

[5] K. Dowling, M. J. Dayel, M. J. Lever, P. M. W. French, J. D. Hares, and A. K. L. Dymoke-Bradshaw, "Fluorescence lifetime imaging with picosecond resolution for biomedical applications," *Optics Letters*, vol. 23, no. 10, pp. 810–812, May 15 1998.

[6] S. B. Bambot, J. R. Lakowicz, and G. Rao, "Potential applications of lifetime-based, phase-modulation fluorimetry in bioprocess and clinical monitoring," *Trends Biotechnol.*, vol. 13, pp. 106–115, March 1995.

[7] C. V. Raman and K. S. Krishnan, "A new type of secondary radiation," *Nature*, vol. 121, no. 3048, p. 501, March 1928.

[8] J. R. Ferraro, K. Nakamoto, and C. W. Brown, *Introduction to Raman Spectroscopy.* Academic Press, 2003.

[9] C. A. Thompson, J. S. R. K. J. Webb, F. P. LaPlant, and D. Ben-Amotz, "Raman spectroscopic studies of diamond in Intralipid," *Optics Letters*, vol. 20, no. 10, pp. 1195–1197, May 15 1995.

[10] A. J. Berger, T.-W. Koo, I. Itzkan, G. Horowitz, and M. S. Feld, "Multicomponent blood analysis by near-infrared raman spectroscopy," *Applied Optics*, vol. 38, no. 13, p. 2916, May 1999.

[11] M. Moskovits, "Surface-enhanced spectroscopy," *Reviews of Modern Physics*, vol. 57, no. 3, pp. 783–826, July 1985.

[12] D. A. Long, *Raman Spectroscopy.* McGraw Hill, 1977.

[13] R. W. Boyd, *Nonlinear Optics.* Academic Press, 1992.

[14] F. J. Garcia-Vidal and J. B. Pendry, "Collective theory for surface enhanced Raman scattering," *Physical Review Letters*, vol. 77, no. 6, pp. 1163–1166, August 1996.

[15] D. Huang, E. Swanson, C. Lin, J. Schuman, W. Stinson, W. Chang, M. Hee, T. Flotte, D. Gregory, C. Puliafito, and J. Fujimoto, “Optical coherence tomography,” *Science*, vol. 254, pp. 1178–1181, Nov 1991.

[16] W. Drexler, U. Morgner, F. X. Kartner, C. Pitris, S. A. Boppart, X. D. Li, E. P. Ippen, and J. G. Fujimoto, “*In vivo* ultrahigh-resolution optical coherence tomography,” *Opt. Lett.*, vol. 24, no. 17, pp. 1221–1223, Sep. 1999.

[17] S. R. Arridge, “Optical tomography in medical imaging,” *Inverse Problems*, vol. 15, pp. R41–R93, 1999.

[18] D. A. Benaron and D. K. Stevenson, “Optical time-of-flight and absorbance imaging of biologic media,” *Science*, vol. 259, no. 5100, pp. 1463–1466, 1993.

[19] B. Chance, J. S. Leigh, H. Miyake, D. S. Smith, S. Nioka, R. Greenfeld, M. Finander, K. Kaufmann, W. Levy, and *et al.*. M. Young, “Comparison of time-resolved and -unresolved measurements of deoxyhemoglobin in brain,” *Proc Natl Acad Sci U S A.*, vol. 85, no. 14, pp. 4971–4975, July 1988.

[20] J. S. Reynolds, A. Przadka, S. Yeung, and K. J. Webb, “Optical diffusion imaging: a comparative numerical and experimental study,” *Applied Optics*, vol. 35, no. 19, pp. 3671–3679, July 1996.

[21] J. S. Reynolds, C. A. Thompson, K. J. Webb, F. P. LaPlant, and D. Ben-Amotz, “Frequency domain modeling of reradiation in highly scattering media,” *Applied Optics*, vol. 36, pp. 2252–2259, April 1997.

[22] A. B. Milstein, S. Oh, J. S. Reynolds, K. J. Webb, C. A. Bouman, and R. P. Millane, “Three-dimensional Bayesian optical diffusion tomography using experimental data,” *Optics Letters*, vol. 27, pp. 95–97, January 2002.

[23] F. E. W. Schmidt, M. E. Fry, E. M. C. Hillman, J. C. Hebden, and D. T. Delpy, “A 32-channel time-resolved instrument for medical optical tomography,” *Rev. Sci. Inst.*, vol. 71, no. 1, pp. 256–265, Jan. 2000.

[24] W. Becker, A. Bergmann, G. Biscotti, and A. Ruck, “Advanced time-correlated single photon counting techniques for spectroscopy and imaging in biomedical systems,” *Proceedings of the SPIE 5340: Commercial and Biomedical Applications of Ultrafast Lasers*, vol. 5340, January 2004, San Jose, CA.

[25] B. W. Pogue, M. Testorf, T. McBride, U. Østerberg, and K. Paulsen, “Instrumentation and design of a frequency-domain diffuse optical tomography imager for breast cancer detection,” *Opt. Express*, vol. 1, no. 13, pp. 391–403, Dec. 22 1997.

[26] Q. Zhang, T. J. Brukilacchio, T. Gaudett, L. Wang, A. Li, and D. A. Boas, “Experimental comparison of using continuous-wave and frequency-domain diffuse optical imaging systems to detect heterogeneities,” *Proc. SPIE*, vol. 4250, June 2001, pp. 219–238.

[27] G. Strangman, D. A. Boas, and J. P. Sutton, "Non-invasive neuroimaging using near-infrared light," *Biol. Psychiatry*, vol. 52, pp. 679–693, 2002.

[28] D. A. Boas, J. P. Culver, J. J. Stott, and A. K. Dunn, "Three dimensional Monte Carlo code for photon migration through complex heterogeneous media including the adult human head," *Opt. Express*, vol. 10, no. 3, pp. 159–170, Feb. 11 2002.

[29] A. M. Siegel, J. P. Culver, J. B. Mandeville, and D. A. Boas, "Temporal comparison of functional brain imaging with diffuse optical tomography and fMRI during rat forepaw stimulation," *Phys. Med. Biol.*, vol. 48, no. 10, pp. 1391–1403, May 21 2003.

[30] A. Y. Bluestone, G. Abdoulaev, C. H. Schmitz, R. L. Barbour, and A. H. Hielscher, "Three-dimensional optical tomography of hemodynamics in the human head," *Opt. Express*, vol. 9, no. 6, pp. 272–286, Sept. 10 2001.

[31] "NIRx Medical Technologies, LLC." http://www.nirx.net/products/dynot.html, 2004.

[32] C. H. Schmitz, H. L. Graber, H. Luo, I. Arif, J. Hira, Y. Pei, A. Bluestone, S. Zhong, R. Andronica, I. Soller, N. Ramirez, S. S. Barbour, and R. L. Barbour, "Instrumentation and calibration protocol for imaging dynamic features in dense-scattering media by optical tomography," *Appl. Opt.*, vol. 39, no. 34, pp. 6466–6486, December 1 2000.

[33] J. J. Stott, "Introduction to optical breast imaging." Presentation at MGH, available at http://www.nmr.mgh.harvard.edu/˜jstott/, 2002.

[34] T. O. McBride, B. W. Pogue, S. Poplack, S. Soho, W. A. Wells, S. Jiang, U. L. Østerberg, and K. D. Paulsen, "Multispectral near-infrared tomography: a case study in compensating for water and lipid content in hemoglobin imaging of the breast," *J. Biomed. Opt.*, vol. 7, no. 1, pp. 72–79, January 2002.

[35] J. J. Duderstadt and L. J. Hamilton, *Nuclear Reactor Analysis*. New York: Wiley, 1976.

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

[37] S. S. Saquib, K. M. Hanson, and G. S. Cunningham, "Model-based image reconstruction from time-resolved diffusion data," *Proc. of SPIE Conf. on Medical Imaging: Image Processing*, vol. 3034, February 25-28 1997, Newport Beach, CA, pp. 369–380.

[38] S. R. Arridge and M. Schweiger, "A gradient-based optimisation scheme for optical tomography," *Optics Express*, vol. 2, no. 6, pp. 213–226, March 1998.

[39] A. H. Hielscher, A. D. Klose, and K. M. Hanson, "Gradient-based iterative image reconstruction scheme for time-resolved optical tomography," *IEEE Trans. on Medical Imaging*, vol. 18, no. 3, March 1999.

[40] C. A. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. on Image Processing*, vol. 2, no. 3, pp. 296–310, July 1993.

[41] R. Roy, *Image reconstruction from light measurement on biological tissue.* Ph.D. dissertation, Universit of Hertfordshire, Harfield, England, 1996.

[42] A. D. Klose and A. H. Hielscher, "Iterative reconstruction scheme for optical tomography based on the equation of radiative transfer," *Medical Physics*, vol. 26, pp. 1698–1707, 1999.

[43] M. V. Klibanov, T. R. Lucas, and R. M. Frank, "A fast and accurate imaging algorithm in optical diffusion tomography," *Inverse Problems*, vol. 13, pp. 1341–1361, 1997.

[44] R. Roy and E. Sevick-Muraca, "A numerical study of gradient-based nonlinear optimization methods for contrast enhanced optical tomography," *Opt. Express*, vol. 9, no. 1, pp. 49–65, July 2001.

[45] R. Roy and E. Sevick-Muraca, "Truncated Newton's optimization scheme for absorption and fluorescence optical tomography: Part I theory and formulation," *Opt. Express*, vol. 4, pp. 353–371, 1999.

[46] R. Roy and E. Sevick-Muraca, "Truncated Newton's optimization scheme for absorption and fluorescence optical tomography: Part II reconstruction from synthetic measurements," *Opt. Express*, vol. 4, pp. 372–382, 1999.

[47] D. Paithankar, A. Chen, B. Pogue, M. Patterson, and E. Sevick-Muraca, "Imaging of fluorescent yield and lifetime from multiply scattered light reemitted from random media," *Applied Optics*, vol. 36, no. 10, pp. 2260–2272, April 1997.

[48] K. D. Paulsen and H. Jiang, "Enhanced frequency-domain optical image reconstruction in tissues through total-variation minimization," *Applied Optics*, vol. 35, no. 19, pp. 3447–3458, July 1996.

[49] H. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson, "Optical image reconstruction using frequency-domain data: simulation and experiment," *J. Optical Society America A*, vol. 13, no. 2, pp. 253–266, February 1996.

[50] M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh, "Experimental images of heterogeneous turbid media by frequency-domain diffusing-photon tomography," *Optics Letters*, vol. 20, no. 5, pp. 426–428, March 1995.

[51] S. R. Arridge, M. Schweiger, M. Hiraoka, and D. T. Delpy, "Performance of an iterative reconstruction algorithm for near infrared absorption and scattering imaging," *Proc. of SPIE Conf. on Photon Migration and Imaging in Random Media and Tissues* (G. Muller, B. Chance, R. Alfano, S. Arridge, J. Beuthan, E. Gratton, M. Kaschke, B. Masters, S. Svanberg, and P. van der Zee, eds.), vol. 1888, May 1993, Los Angeles, CA., pp. 360–371.

[52] J. C. Ye, K. J. Webb, R. P. Millane, and T. J. Downar, "Modified distorted Born iterative method with an approximate Fréchet derivative for optical diffusion tomography," *J. Optical Society America A*, vol. 16, no. 7, pp. 1814–1826, July 1999.

[53] A. D. Klose and A. H. Hielsher, "Quasi-Newton methods in optical tomographic image reconstruction," *Inverse Problems*, vol. 19, pp. 387–409, April 2003.

[54] A. Brandt, "Multi-level adaptive solutions to boundary value problems," *Mathematics of Computation*, vol. 31, no. 138, pp. 333–390, April 1977.

[55] U. Trottenberg, C. Oosterlee, and A. Schueller, *Multigrid*. London: Academic Press, 2000.

[56] W. L. Briggs, V. E. Henson, and S. F. McCormick, *A Multigrid Tutorial, 2nd Ed.* Philadelphia: Society for Industrial and Applied Mathematics, 2000.

[57] S. McCormick, ed., *Multigrid Methods*. Philadelphia: Society for Industrial and Applied Mathematics, 1987.

[58] W. Hackbusch, *Multigrid Methods and Applications*. Springer Series in Computational Mathematics, Berlin: Springer-Verlag, 1985.

[59] P. Wesseling, *An Introduction to Multigrid Methods*. Chichester: John Wiley & Sons, 1992.

[60] A. Brandt, "Multiscale and multiresolution methods: Theory and applications," in *Multiscale Scientific Computation: Review 2001* (T. J. Barth, T. F. Chan, , and R. Haimes, eds.), pp. 3–96, Heidelberg: Springer Verlag, 2001.

[61] A. Brandt and D. Ron, "Multigrid solvers and multilevel optimization strategies," in *Multilevel Optimization and VLSI-CAD* (J. Cong and J. Shinnerl, eds.), pp. 1–69, Boston: Kluwer Academic Publishers, 2002.

[62] C. A. Bouman and K. Sauer, "Nonlinear multigrid methods of optimization in Bayesian tomographic image reconstruction," *Proc. of SPIE Conf. on Neural and Stochastic Methods in Image and Signal Processing*, vol. 1766, July 19-24 1992, San Diego, CA, pp. 296–306.

[63] J. C. Ye, C. A. Bouman, R. P. Millane, and K. J. Webb, "Nonlinear multigrid optimization for Bayesian diffusion tomography," *Proc. of IEEE Int'l Conf. on Image Proc.*, October 25-28 1999, Kobe, Japan.

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

[65] S. Oh, A. B. Milstein, C. A. Bouman, and K. J. Webb, "Multigrid inversion algorithms with applications to optical diffusion tomography," *Proc. of 36th Asilomar Conference on Signals, Systems, and Computers*, Nov. 2002, Monterey, CA, pp. 901–905.

[66] S. Oh, A. B. Milstein, C. A. Bouman, and K. J. Webb, "Multigrid algorithms for optimizations and inverse problems," *Proc. the SPIE/IS&T Conference on Computational Imaging 2003*, Jan. 20-25 2003, Santa Clara, CA, USA, pp. 59–70.

[67] S. Oh, A. B. Milstein, C. A. Bouman, and K. J. Webb, "Adaptive nonlinear multigrid inversion with applications to Bayesian optical diffusion tomography," *Proc. IEEE Workshop on Statistical Signal Processing*, Sep. 2003, St. Louis, MO, USA.

[68] S. Oh, A. B. Milstein, C. A. Bouman, and K. J. Webb, "Nonlinear multigrid inversion," *Proc. of IEEE Int'l Conf. on Image Proc.*, Sep. 2003, Barcelona, Spain.

[69] S. F. McCormick and J. G. Wade, "Multigrid solution of a linearized, regularized least-squares problem in electrical impedance tomography," *Inverse Problems*, vol. 9, pp. 697–713, 1993.

[70] L. Borcea, "Nonlinear multigrid for imaging electrical conductivity and permittivity at low frequency," *Inverse Problems*, vol. 17, pp. 329–359, April 2001.

[71] R. Gandlin and A. Brandt, "Two multigrid algorithms for inverse problem in electrical impedance tomography," *Proc. 2003 Copper Mountain Conf. Multigrid Methods*, March 30-April 4 2003, Copper Mountain, CO, USA.

[72] C. R. Johnson, M. Mohr, U. Ruede, A. Samsonov, and K. Zyp, "Multilevel methods for inverse bioelelectric field problems," *Lecture Notes in Computational Science and Engineering - Multiscale and Multiresolution Methods: Theory and Applications, Eds. T.J. Barth, T.F. Chan, R. Haimes.*, vol. 20, Oct 2001, Springer-Verlag Publishing, Heidelberg.

[73] M. Guven, B. Yazici, Z. Intes, and B. Chance, "An adaptive multigrid algorithm for region of interest diffuse optical tomography," *Proc. of IEEE Int'l Conf. on Image Proc.*, Sep. 2003, Barcelona, Spain.

[74] S. G. Nash, "A multigrid approach to discretized optimization problems," *J. of Optimization methods and software*, vol. 14, pp. 99–116, 2000.

[75] R. M. Lewis and S. G. Nash, "A multigrid approach to the optimization of systems governed by differential equations," *8-th AIAA/USAF/ISSMO Symp. Multidisciplinary Analysis and Optimization*, 2000, Long Beach, CA.

[76] J. C. Hebden and K. S. Wong, "Time-resolved optical tomography," *Appl. Opt.*, vol. 32, no. 4, pp. 372–380, Feb. 1993.

[77] Y. Yamada, Y. Hasegawa, and Y. Yamashita, "Simulation of fan-beam-type optical computed-tomography imaging of strongly scattering and weakly absorbing media," *Appl. Opt.*, vol. 32, no. 25, pp. 4808–4814, Sep. 1993.

[78] I. Oda, H. Eda, Y. Tsunazawa, M. Takada, Y. Yamada, G. Nishimura, and M. Tamura, "Optical tomography by the temporally extrapolated absorbance method," *Appl. Opt.*, vol. 35, no. 1, pp. 169–175, Jan. 1996.

[79] S. B. Colak, D. G. Papaioannou, G. W. 't Hooft, M. B. van der Mark, H. Schomberg, J. C. J. Paasschens, J. B. M. Mellissen, and N. A. A. J. van Asten, "Tomographic image reconstruction from optical projections in light-diffusing media," *Applied Optics*, vol. 36, no. 1, pp. 180–213, January 1997.

[80] M. A. Franceschini, V. Toronov, M. E. Filiaci, E. Gratton, and S. Fantini, "On-line optical imaging of he human brain with 160-ms temporal resolution," *Opt. Express*, vol. 6, no. 3, pp. 49–57, Jan. 2000.

[81] A. B. Konovalov, V. V. Lyubimov, I. I. Kutuzov, O. V. Kravtsenyuk, A. G. Murzin, G. B. Mordvinov, L. N. Soms, and L. M. Yavorskaya, "Application of transform algorithms to high-resolution image reconstruction in optical diffusion tomography of strongly scattering media," *Journal of Electronic Imaging*, vol. 12, no. 4, pp. 601–612, Oct. 2003.

[82] W. Zhu, Y. Wang, Y. Yao, J. Chang, H. L. Graber, and R. L. Barbour, "Iterative total least squres image reconstruction algorithm for optical tomography by the conjugate gradient method," *J. Opt. Soc. Am. A*, vol. 14, no. 4, pp. 799–807, April 1997.

[83] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge: Cambridge University Press, 1992.

[84] K. Sauer and C. A. Bouman, "A local update strategy for iterative reconstruction from projections," *IEEE Trans. on Signal Processing*, vol. 41, no. 2, pp. 534–548, February 1993.

[85] J. C. Ye, K. J. Webb, C. A. Bouman, and R. P. Millane, "Optical diffusion tomography using iterative coordinate descent optimization in a Bayesian framework," *J. Optical Society America A*, vol. 16, no. 10, pp. 2400–2412, October 1999.

[86] C. A. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Trans. on Image Processing*, vol. 5, no. 3, pp. 480–492, March 1996.

[87] W. Briggs, *A Multigrid Tutorial*. Philadelphia: Society for Industrial and Applied Mathematics, 1987.

[88] S. Oh, A. B. Milstein, C. A. Bouman, and K. J. Webb, "A general framework for nonlinear multigrid inversion," *IEEE Trans. Image Process.*

[89] C. P. Gonatas, M. Ishii, J. S. Leigh, and J. C. Schotland, "Optical diffusion imaging using a direct inversion method," *Physical Review E*, vol. 52, no. 4, pp. 4361–4365, Oct. 1995.

[90] R. J. Gaudette, D. H. Brooks, C. A. DiMarzio, M. E. Kilmer, E. L. Miller, T. Gaudette, and D. A. Boas, "A comparison study of linear reconstruction techniques for diffuse optical tomographic imaging of absorption coefficient," *Phys. Med. Biol.*, vol. 45, pp. 1051–1070, 2000.

[91] V. A. Markel, V. Mital, and J. C. Schotland, "Inverse problem in optical diffusion tomography. iii. inversion formulas and singular-value decomposition," *J. Opt. Soc. Am. A*, vol. 20, no. 5, pp. 890–902, May 2003.

[92] D. A. Boas, "A fundamental limitation of linearized algorithms for diffuse optical tomography," *Opt. Express*, vol. 1, no. 13, pp. 404–413, Dec. 1997.

[93] D. Boas, T. Gaudette, and S. Arridge, "Simultaneous imaging and optode calibration with diffuse optical tomography," *Opt. Express*, vol. 8, no. 5, pp. 263–270, February 2001.

[94] S. Oh, A. B. Milstein, R. P. Millane, C. A. Bouman, and K. J. Webb, "Source-detector calibration in three-dimensional Bayesian optical diffusion tomography," *J. Optical Society America A*, vol. 19, no. 10, pp. 1983–1993, Oct. 2002.

[95] H. Jiang, K. Paulsen, and U. Osterberg, "Optical image reconstruction using dc data: simulations and experiments," *Phys. Med. Biol.*, vol. 41, no. 8, pp. 1483–1498, August 1996.

[96] H. Jiang, K. Paulsen, U. Osterberg, and M. Patterson, "Improved continuous light diffusion imaging in single- and multi-target tissue-like phantoms," *Phys. Med. Biol.*, vol. 43, no. 3, pp. 675–693, March 1998.

[97] B. W. Pogue, S. P. Poplack, T. O. McBride, W. A. Wells, K. S. Osterman, U. L. Østerberg, and K. D. Paulsen, "Quantitative hemoglobin tomography with diffuse near-infrared spectroscopy: pilot results in the breast," *Radiology*, vol. 218, no. 1, pp. 261–266, January 2001.

[98] B. W. Pogue, C. Willscher, T. O. McBride, U. L. Østerberg, and K. D. Paulsen, "Contrast-detail analysis for detection and characterization with near-infrared diffuse tomography," *Med. Phys.*, vol. 27, no. 12, pp. 2693–2700, December 2000.

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

[100] N. Iftimia and H. Jiang, "Quantitative optical image reconstructions of turbid media by use of direct-current measurements," *Applied Optics*, vol. 39, no. 28, pp. 5256–5261, October 2000.

[101] J. J. Stott, J. P. Culver, S. R. Arridge, and D. A. Boas, "Optode positional calibration in diffuse optical tomography," *Appl. Opt.*, vol. 42, no. 16, pp. 3154–3162, June 2003.

[102] Y. Shi and W. C. Karl, "Tomographic reconstruction of dynamic objects," *Proceedings of the SPIE/IS&T Conference on Computational Imaging*, vol. 5016, January 23-24 2003, Santa Clara, CA, pp. 151–160.

[103] M. E. Kilmer, E. L. Miller, D. Boas, and D. Brooks, "A shape-based reconstruction technique for DPDW data," *Opt. Express*, vol. 7, no. 13, pp. 481–491, Dec. 2000.

[104] V. Kolehmainen, S. R. Arridge, W. R. B. Lionheart, M. Vauhkonen, and J. P. Kaipio, "Recovery of region boundaries of piecewise constant coefficients of an elliptic PDE from boundary data," *Inverse Problems*, vol. 15, pp. 1375–1391, 1999.

[105] V. Kolehmainen, M. Vauhkonen, and J. P. Kaipio, "Recovery of piecewise constant coefficients in optical diffusion tomography," *Opt. Express*, vol. 7, no. 13, pp. 468–480, Dec. 2000.

[106] V. Kolehmainen, S. R. Arridge, M. Vauhkonen, and J. P. Kaipio, "Simultaneous reconstruction of internal tissue region boundaries and coefficients in optical diffusion tomography," *Phys. Med. Biol.*, vol. 45, pp. 3267–3283, 2000 2000.

[107] M. E. Kilmer, E. L. Miller, A. Barbaro, and D. Boas, "Three-dimensional shape-based imaging of absorption perturbation for diffuse optical tomography," *Appl. Opt.*, vol. 42, no. 16, pp. 3129–3144, June 2003.

[108] J. C. Ye, Y. Bresler, and P. Moulin, "Cramer-rao bounds for parametric shape estimation in inverse problems," *IEEE Trans. on Image Processing*, vol. 12, no. 1, pp. 71–84, January 2003.

[109] H. Feng, W. C. Karl, and D. A. Castanon, "A curve evolution approach to object-based tomographic reconstruction," *IEEE Trans. on Image Processing*, vol. 12, no. 1, pp. 44–57, January 2003.

[110] A. B. Milstein, J. J. Stott, S. Oh, D. A. Boas, R. P. Millane, C. A. Bouman, and K. J. Webb, "Fluorescence optical diffusion tomography using multiple-frequency data," *J. Opt. Soc. Am. A*, (to appear).

[111] J. P. Culver, V. Ntziachristos, M. J. Holbrooke, and A. G. Yodh, "Optimization of optode arrangements for diffuse optical tomography: A singular-value analysis," *Opt. Lett.*, vol. 26, no. 10, pp. 701–703, May 15 2001.

[112] M. Gurfinkel, A. B. Thompson, W. Ralston, T. L. Troy, A. L. Moore, T. A. Moore, J. D. Gust, D. Tatman, J. S. Reynolds, B. Muggenburg, K. Nikula, R. Pandey, R. H. Mayer, D. J. Hawrysz, and E. M. Sevick-Muraca, "Pharmacokinetics of ICG and HPPH-car for the detection of normal and tumor tissue using fluorescence, near-infrared reflectance imaging: a case study," *Photochem. Photobiol.*, vol. 72, no. 1, pp. 94–102, July 2000.

[113] D. J. Cuccia, F. Bevilacqua, A. J. Durkin, S. Merritt, B. J. Tromberg, G. Gulsen, H. Yu, J. Wang, and O. Nalcioglu, "*In vivo* quantification of optical contrast agent dynamics in rat tumors by use of diffuse optical spectroscopy with magnetic resonance imaging coregistration," *Appl. Opt.*, vol. 42, no. 16, pp. 2940–2950, June 2003.

[114] A. B. Milstein, S. Oh, K. J. Webb, and C. A. Bouman, "Direct reconstruction of kinetic parameter images in fluorescence optical diffusion tomography," *IEEE Workshop on Biomedical Imaging*, April 15 2004, Washington, D.C.

[115] A. B. Milstein, S. Oh, K. J. Webb, and C. A. Bouman, "Estimation of kinetic model parameters in optical diffusion tomography," *Computational Imaging II*, vol. 5299, Jan. 19-20 2004, San Jose, CA.

[116] M. V. Knopp, E. Weiss, H. P. Sinn, J. Mattern, H. Junkermann, J. Radeleff, A. Magener, G. Brix, S. D. S, I. Zuna, and G. van Kaick, "Pathophysiologic basis of contrast enhancement in breast tumors," *Journal of Magnetic Resonance Imaging*, vol. 10, no. 3, pp. 260–266, 1999.

[117] P. J. Hudson, "Recombinant antibodies: a novel approach to cancer diagnosis and therapy," *Expert Opinion on Investigational Drugs*, vol. 9, no. 6, pp. 1231–1242, 2000.

[118] S. D. Jonson and M. J. Welch, "PET imaging of breast cancer with fluorine-18 radiolabeled estrogens and progestins," *Quarterly Journal of Nuclear Medicine*, vol. 42, no. 1, pp. 8–17, 1998.

[119] A. Heppeler, S. Froidevaux, A. N. Eberle, and H. R. Maecke, "Receptor targeting for tumor localisation and therapy with radiopeptides," *Current Medicinal Chemistry*, vol. 7, no. 9, pp. 971–994, 2000.

[120] J. A. Reddy and P. S. Low, "Folate-mediated targeting of therapeutic and imaging agents to cancers," *Critical Reviews in Therapeutic Drug Carrier Systems*, vol. 15, no. 6, pp. 587–627, 1998.

[121] K. J. Webb, A. B. Milstein, M. D. Kennedy, K. N. Jallad, C. A. Bouman, D. Ben-Amotz, and P. S. Low, "Folate conjugate fluorescence labeling for tumor localization," *Third Inter-Institute Workshop on Diagnostic Optical Imaging and Spectroscopy: The Clinical Adventure. Bethesda, MD, 2002, vol Poster Presentation*, September 2002, Bethesda, MD.

[122] C. A. Thompson, K. J. Webb, and A. M. Weiner, "Diffusive media characterization using laser speckle," *Applied Optics*, vol. 36, pp. 3726–3734, June 1997.

[123] J. D. McKinney, M. A. Webster, K. J. Webb, and A. M. Weiner, "Characterization and imaging in optically scattering media by use of laser speckle and a variable-coherence source," *Optics Letters*, vol. 25, no. 1, p. 4, January 2000.

[124] M. A. Webster, K. J. Webb, and A. M. Weiner, "Temporal response of a random medium from third-order laser speckle frequency correlations," *Physical Review Letters*, vol. 88, no. 3, pp. 033901/1–4, January 2002.

[125] A. Z. Genack and J. M. Drake, "Relationship between optical intensity, fluctuations and pulse propagation in random media," *Europhysics Letters*, vol. 11, no. 4, pp. 331–336, February 1990.

[126] A. W. Lohmann and B. Wirnitzer, "Triple correlations [optical computing]," *Proc. of the IEEE*, vol. 72, no. 7, pp. 889–901, July 1984.

[127] H. Bartelt, A. W. Lohmann, and B. Wirnitzer, "Phase and amplitude recovery from bispectra," *Applied Optics*, vol. 23, no. 18, pp. 3121–3129, September 1984.

[128] M. A. Webster, K. J. Webb, A. M. Weiner, J. Xu, and H. Cao, "Temporal response of a random medium from speckle intensity frequency correlations," *J. Optical Society America A*, vol. 20, no. 11, pp. 2057–2070, November 2003.