FREQUENCY DOMAIN OPTICAL TOMOGRAPHY BASED ON THE EQUATION OF RADIATIVE TRANSFER

KUI REN[*], GUILLAUME BAL [†], AND ANDREAS H. HIELSCHER [‡]

Abstract. Optical tomography consists in reconstructing the spatial distribution of absorption and scattering properties of a medium from surface measurements of transmitted light intensities. Mathematically, this problem amounts to parameter identification for the equation of radiative transfer (ERT) with diffusion-type boundary measurements. Because they are posed in the phase-space, radiative transfer equations are quite challenging to solve computationally. Most past works have considered the diffusion approximation of ERT or the steady-state ERT. In both case, substantial cross-talk has been observed in the reconstruction of the absorption and scattering properties of inclusions. In this paper, we present an optical tomographic reconstruction algorithm based on the frequency-domain ERT. The inverse problem is formulated as a regularized least-square minimization problem, in which the mismatch between forward model predictions and measurements is minimized. The ERT is discretized by using a discrete ordinates method for the directional variables and a finite volume method for the spatial variables. A limited-memory Quasi-Newton algorithm is used to minimize the least-square functional. Numerical simulations with synthetic data show that the cross-talk between the two optical parameters is significantly reduced in reconstructions based on frequency-domain data as compared to those based on steady-state data.

Key words. Optical tomography, photon density waves, equation of radiative transfer, finite volume method, discrete ordinates method, generalized minimal residual algorithm, inverse problems, regularization, numerical optimization.

AMS subject classifications. 65N21, 65Z05, 78M25, 78M50

1. Introduction. The inverse problem related to the equation of radiative transfer (ERT) has long been of interest in many fields of applied physics and mathematics. In this problem, one attempts to determine the spatial distribution of parameters (e.g., the spatial distribution of absorption and scattering coefficients) inside a medium from either transmission or reflectance data collected by illuminating the medium from many different directions. Applications include for instance astronomy [57], nuclear science [40, 41], and atmospheric science [11]; see the review [47]. Recent advances in this field have been fueled by an increased interest in medical applications. By performing optical measurements on the surface of the skin, optical tomographic techniques attempt to reconstruct the spatial distribution of absorption and scattering coefficients inside the body. These parameters can then be used for many diagnostic purposes [12].

Mathematically, optical tomography reduces to parameter identification problems (inverse problems) for the ERT, also referred to as the linear Boltzmann equation or the transport equation. Most early studies in the field deal with the plane-parallel geometry [19, 23]. Inversion procedures were also developed in many more general yet still constrained geometries; see for instance [3, 40, 41, 57] and the review papers [46, 47]. Because of their geometrical specificities, most of these works do not directly apply to biomedical applications. In the mid-nineties several groups started developing schemes and algorithms that overcame these problems in the framework of the

---

[*]Department of Applied Physics & Applied Mathematics, Columbia University, New York, NY 10027; kr2002@columbia.edu

[†]Department of Applied Physics & Applied Mathematics, Columbia University, New York, NY 10027; gb2030@columbia.edu

[‡]Department of Biomedical Engineering and Radiology, Columbia University, New York, NY 10027; ahh2004@columbia.edu

diffusion approximation to the ERT; see for instance [4, 30, 54, 59]. Many theoretical and experimental studies have shown that the diffusion approximation was only valid in specific contexts, for instance when the scattering coefficient is sufficiently large and the absorption coefficient sufficiently small [15, 16, 22, 27]. While this is often the case in biomedical applications there remain many important applications, for which reconstructions based on the ERT are needed. Examples are brain imaging, where the presence of low-absorbing and low-scattering cerebrospinal fluid leads to non-diffusive light propagation; imaging of joints, where clear synovial fluid has the same effects as cerebrospinal fluid in the brain; and small animal imaging, where small object dimensions (1-2 cm in diameter) invalidate the classical diffusion approximation [28, 29]; see however [7, 8] for ways to generalize the diffusion equation in certain situations.

Only quite recently have studies emerged that describe ERT based reconstruction codes for use in biomedical optical tomography. First, a transport-backtransport method, a nonlinear inversion method, applied to the two-dimensional time-dependent equation of radiative transfer was reported in [17]. New algorithms were developed and experimentally tested for two- and three- dimensional cases using a time-independent ERT in [36, 37, 38, 39]. While these works, which address real-life three-dimensional problems are an important step towards practical applications, they still suffer from considerable cross-talk between absorption and scattering reconstructions. What we mean by cross-talk is that purely scattering (or purely absorbing) inclusions are often reconstructed with unphysical absorption (or scattering) properties. This behavior is well-understood from the theoretical viewpoint: Different optical distributions inside the medium can lead to the same measurements collected at the surface of the medium [5, 31]. To avoid such cross-talks, which may lead to wrong diagnostics, we need different data. An experimental technique increasingly employed in recent years to obtain additional information is to use frequency domain measurements. In this case the source intensity is modulated (typically between 100-1000 MHz), leading to the propagation of so-called photon density waves. Since frequency domain measurements provide information about both the phase and the intensity of the waves (and not only the intensity as in steady-state measurements), it is expected that frequency-domain techniques will allow for better separation of absorption and scattering effects [5, 45]. Numerical reconstructions based on frequency-domain ERT, however, have not yet been developed in the literature. This is the major motivation for the present work.

The remainder of the paper is organized as follows. In section 2 we formulate the inverse problem in frequency-domain optical tomography as a classical regularized least square problem and describe the adjoint method for the computation of the gradient of the least square functional. In section 3 we present the details on the discretization of the forward and backward problems, which is based on using a discrete ordinates methods for the directional variables and a finite volume method for the spatial variables. The implementation of the minimization procedure, based on a quasi-Newton optimization algorithm, and the choice of the regularization parameter are described in section 4. Several numerical simulations illustrate the performance of the code in section 5. Reconstructions based on synthetic data indeed show a significant reduction of the cross-talk mentioned above when frequency-domain data are used. Concluding remarks are finally offered in section 6.

2. Continuous problem formulation. We now formulate the optical tomography problem. The frequency-domain equation of radiative transfer that describes the photon density in the phase space, i.e., as a function of position $\mathbf{x} \in \mathcal{D} \subset \mathbb{R}^n$ and

direction $\boldsymbol{\theta} \in S^{n-1}$ (unit sphere of $\mathbb{R}^n$) is given by [4]

$$\mathcal{T}u \equiv \left(\frac{i\omega}{v} + \boldsymbol{\theta} \cdot \nabla + \sigma_a(\mathbf{x})\right)u(\mathbf{x}, \boldsymbol{\theta}) + Q(u)(\mathbf{x}, \boldsymbol{\theta}) \;=\; 0 \qquad \text{in } \mathcal{D} \times S^{n-1}$$
$$u(\mathbf{x}, \boldsymbol{\theta}) \;=\; f(\mathbf{x}, \boldsymbol{\theta}) \quad \text{on } \Gamma_-,$$

(2.1)

where $i = \sqrt{-1}$, $n = 2, 3$ is the space dimension, $v \in \mathbb{R}^+$ is the speed of light in the medium, and ω is the modulation frequency of the boundary source $f(\mathbf{x}, \boldsymbol{\theta})$. The non-negative function $\sigma_a(\mathbf{x}) \in L^\infty(\mathcal{D})$ is the absorption coefficient. The unknown quantity, $u(\mathbf{x}, \boldsymbol{\theta})$, is the radiant power per unit solid angle per unit area perpendicular to the direction of propagation at $\mathbf{x}$ in the direction $\boldsymbol{\theta}$. Note that $u(\mathbf{x}, \boldsymbol{\theta})$ depends also on ω although, for simplicity, we do not write this dependency explicitly. The boundary sets $\Gamma_\pm$ are defined as

$$\Gamma_\pm = \{(\mathbf{x}, \boldsymbol{\theta}) \in \partial\mathcal{D} \times S^{n-1} \text{ s.t. } \pm \boldsymbol{\theta} \cdot \boldsymbol{\nu}(\mathbf{x}) > 0\},$$

with $\boldsymbol{\nu}(\mathbf{x})$ the outward unit normal to the domain at $\mathbf{x} \in \partial\mathcal{D}$. The scattering operator Q is defined as

$$Q(u)(\mathbf{x}, \boldsymbol{\theta}) = \sigma_s(\mathbf{x})\left(u(\mathbf{x}, \boldsymbol{\theta}) - \int_{S^{n-1}} k(\boldsymbol{\theta} \cdot \boldsymbol{\theta}')u(\mathbf{x}, \boldsymbol{\theta}')d\mu(\boldsymbol{\theta}')\right). \qquad (2.2)$$

Here, $\sigma_s(\mathbf{x}) \in L^\infty(\mathcal{D})$ is the scattering coefficient and $d\mu$ is the surface measure on S^{n-1} normalized so that $\int_{S^{n-1}} d\mu(\boldsymbol{\theta}) = 1$. The "collision" kernel $k(\boldsymbol{\theta} \cdot \boldsymbol{\theta}')$, which describes the probability that photons traveling in direction $\boldsymbol{\theta}'$ scatter into direction $\boldsymbol{\theta}$, is a positive function independent of $\mathbf{x}$ and satisfies the normalization condition:

$$\int_{S^{n-1}} k(\boldsymbol{\theta} \cdot \boldsymbol{\theta}')d\mu(\boldsymbol{\theta}') = 1. \qquad (2.3)$$

The scattering kernel for light propagation in tissues is highly peaked forward and is chosen as the Henyey-Greenstein phase function [26, 58]

$$k(\boldsymbol{\theta} \cdot \boldsymbol{\theta}') = \frac{1 - g^2}{(1 + g^2 - 2g\cos\phi)^{3/2}}, \qquad (2.4)$$

where ϕ is the angel between $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$, i.e., $\boldsymbol{\theta} \cdot \boldsymbol{\theta}' = \cos\phi$ and where $g \in [0, 1]$ is the anisotropy factor, which measures how peaked forward the phase function is. The larger g is, the more forward the scattering. The anisotropy factor is often used to define the so-called effective scattering coefficient through $\sigma_s' = (1-g)\sigma_s$. We mention that scattering kernels other than (2.4) have also been used in some situations [33] and that simplified (Fokker-Planck) models can also be used to analyze highly peaked scattering in biological tissues [34].

The optical tomography problem thus consists of reconstructing $\sigma_a(\mathbf{x})$ and $\sigma_s(\mathbf{x})$ in (2.1) from boundary current measurements; see (2.6) below. Our objective in this work is to present a numerical scheme that performs the reconstruction.

2.1. Forward problem. The absorption and scattering coefficients σ_a and σ_s cannot take negative values and have to be bounded. We thus introduce the following parameter space $\mathcal{Q}$:

$$\mathcal{Q} := \{(\sigma_a, \sigma_s) \;:\; \sigma_a \geq 0, \; \sigma_s \geq 0, \; \text{ and } (\sigma_a, \sigma_s) \in L^\infty(\mathcal{D}) \times L^\infty(\mathcal{D})\}.$$

We also introduce the functional spaces [2, 15]:

$$L^2_{\boldsymbol{\theta}\cdot\nu}(\Gamma_\pm) := \left\{ u(\mathbf{x},\boldsymbol{\theta}) : \int_{\Gamma_\pm} u^2(\mathbf{x},\boldsymbol{\theta}) |\boldsymbol{\theta}\cdot\nu(\mathbf{x})| d\sigma(\mathbf{x}) d\mu(\boldsymbol{\theta}) < +\infty \right\}$$
$$\mathcal{W}^2(\mathcal{D}\times S^{n-1}) := \left\{ u(\mathbf{x},\boldsymbol{\theta}) : u \in L^2(\mathcal{D}\times S^{n-1}) \text{ and } \boldsymbol{\theta}\cdot\nabla u \in L^2(\mathcal{D}\times S^{n-1}) \right\}.$$

Adapting well-known results [2, 15] with complex-valued absorption coefficient $\sigma_a + \frac{i\omega}{v}$ in $L^\infty(\mathcal{D})$, we have the following statement about the forward problem

PROPOSITION 2.1. *Assume that $(\sigma_a, \sigma_s) \in \mathcal{Q}$, the modulation frequency is finite $\omega < +\infty$, and $f \in L^2_{\boldsymbol{\theta}\cdot\nu}(\Gamma_-)$. Then the forward problem (2.1) is well-posed and admits a unique solution $u(\mathbf{x},\boldsymbol{\theta}) \in \mathcal{W}^2(\mathcal{D}\times S^{n-1})$.* We can then define the following *albedo* operator (as well as its adjoint) [13, 48]:

$$\Lambda : \begin{array}{l} f \longmapsto u|_{\Gamma_+} \\ L^2_{\boldsymbol{\theta}\cdot\nu}(\Gamma_-) \longmapsto L^2_{\boldsymbol{\theta}\cdot\nu}(\Gamma_+). \end{array} \tag{2.5}$$

The albedo operator Λ maps the incoming flux on the boundary into the outgoing flux and is a functional of the optical parameters σ_a and σ_s.

A major difficulty in optical tomography comes from the fact that in practice, only outgoing currents, which are angular averages of the outgoing flux and are similar to diffusion-type measurements, are available. This prevents us from using classical uniqueness and stability results in inverse transport theory [13]. In fact, the inverse problem we solve in this paper is very similar to the diffusion-based inverse problem [4], on which many more theoretical results exist. To date, we do not know of any theoretical result on the reconstruction of optical properties from outgoing currents for arbitrary geometries. This makes the development of numerical tools all the more important.

To be consistent with existing measurement technologies, we define the following "measurement operator":

$$\begin{array}{rcl} \mathcal{G}u|_{\Gamma_+} & := & \displaystyle\int_{S^{n-1}_+} \boldsymbol{\theta}\cdot\nu(\mathbf{x}) u|_{\Gamma_+} d\mu(\boldsymbol{\theta}) \equiv z(\mathbf{x}) \\ \mathcal{G} & : & L^2_{\boldsymbol{\theta}\cdot\nu}(\Gamma_+) \longmapsto L^2(\partial\mathcal{D}) \equiv \mathcal{Z} \end{array} \tag{2.6}$$

with $S^{n-1}_+ := \{\boldsymbol{\theta} : \boldsymbol{\theta} \in S^{n-1} \text{ s.t. } \boldsymbol{\theta}\cdot\nu(\mathbf{x}) > 0\}$. We will call $\mathcal{Z}$ the "measurement space". Now the composite operator $\mathcal{G}\Lambda : f \mapsto z$ maps the incoming flux into the tomographic measurements. The adjoint operator $\mathcal{G}^*$ of $\mathcal{G}$ is defined via the identity

$$\langle \mathcal{G}^* g_1, g_2 \rangle_{L^2_{\boldsymbol{\theta}\cdot\nu}(\Gamma_+)} = \langle g_1, \mathcal{G}g_2 \rangle_{\mathcal{Z}}, \tag{2.7}$$

for all $g_1 \in \mathcal{Z}$ and $g_2 \in L^2_{\boldsymbol{\theta}\cdot\nu}(\Gamma_+)$, where $\langle \cdot, \cdot \rangle_X$ is the usual inner product in a Hilbert space X. Observe that $\mathcal{G}^*$ is nothing but the operation of multiplication by $\boldsymbol{\theta}\cdot\nu(\mathbf{x})$.

2.2. Least square formulation. The inverse problem of optical tomography can be formulated as follows: determine $(\sigma_a, \sigma_s) \in \mathcal{Q}$ such that

$$\mathcal{G}\Lambda f = z, \tag{2.8}$$

where $z \in \mathcal{Z}$ is the measured data. This problem is in general severely ill-posed (assuming that uniqueness of reconstruction holds as in diffusion theory [4]) in the sense that when no regularization is applied, noise contained in the data z is more amplified during the inversion procedure than what would results from an arbitrary

number of differentiations [20]. Another practical difficulty in solving (2.8) lies in the fact that the amount of available data may be quite limited [44]. After discretizing (2.8) on a reasonable mesh, we will end up with a very under-determined nonlinear system. A classical way to resolve the lack of measurements is to turn to the following least square formulation: find (σ_a, σ_s) solving:

$$\min \frac{1}{2} \|\mathcal{G}\Lambda f - z\|_{\mathcal{Z}}^2. \tag{2.9}$$

For the reasons we have mentioned above, the least square problem (2.9) is usually not stable [4]. To stabilize the problem, we add additional smoothness restrictions on the coefficients we wish to reconstruct. We thus require that the coefficients lie in the following space of *admissible parameters*:

$$\mathcal{Q}_{ad} := \{(\sigma_a, \sigma_s) \ : \ (\sigma_a, \sigma_s) \in \mathcal{Q}, \text{ and } (\sigma_a, \sigma_s) \in \mathcal{H}^1(\mathcal{D}) \times \mathcal{H}^1(\mathcal{D})\},$$

where $\mathcal{H}^1(\mathcal{D})$ is the usual Hilbert space of L^2 functions with first-order partial derivatives in $L^2(\mathcal{D})$. It is known that $\mathcal{Q}_{ad}$ is a closed and convex subset of $\mathcal{H}^1(\mathcal{D})$. We can thus introduce the following regularized least square functional:

$$\begin{aligned}
\mathcal{F}_\alpha(\sigma_a, \sigma_s) &:= \frac{1}{2}\|\mathcal{G}\Lambda f - z\|_{\mathcal{Z}}^2 + \frac{\alpha}{2} \sum_{p=\{a,s\}} \|\sigma_p\|_{\mathcal{H}^1(\mathcal{D})}^2, \\
\mathcal{F}_\alpha &: \qquad \mathcal{H}^1(\mathcal{D}) \times \mathcal{H}^1(\mathcal{D}) \longmapsto \mathbb{R}^+
\end{aligned} \tag{2.10}$$

where the last term is called a regularization term and α is the regularization parameter [20]. The method for choosing a good α will be described in section 4.3. The $\mathcal{H}^1$ norm is defined as usual via:

$$\|\mathscr{Y}\|_{\mathcal{H}^1(\mathcal{D})}^2 := \|\mathscr{Y}\|_{L^2(\mathcal{D})}^2 + \|\nabla\mathscr{Y}\|_{L^2(\mathcal{D})}^2, \quad \text{for} \quad \mathscr{Y} \in \mathcal{H}^1(\mathcal{D}). \tag{2.11}$$

We thus formulate the optical tomography problem as the following regularized least square (RLS) problem:

(**RLS**) Find a pair (σ_a, σ_s) that minimizes the least square functional $\mathcal{F}_\alpha(\sigma_a, \sigma_s)$ given in (2.10).

We first observe that problem (**RLS**) has at least one solution. That is, the functional $\mathcal{F}_\alpha(\sigma_a, \sigma_s)$ has at least one minimizer. This existence result is classical and follows from the weak lower semicontinuity and coercivity of $\mathcal{F}_\alpha(\sigma_a, \sigma_s)$ [43, 56]. However, we cannot show that $\mathcal{F}_\alpha(\sigma_a, \sigma_s)$ is strictly convex and thus cannot conclude that the minimizer is unique [56].

Our implementation of the inverse problem of optical tomography is a gradient-based minimization approach. We thus need to compute the Fréchet derivative of the least square functional $\mathcal{F}_\alpha(\sigma_a, \sigma_s)$. Direct estimates of the Fréchet derivatives being quite costly because the optical parameters are (at least at the continuous level) infinite dimensional objects, we adopt the adjoint state (or co-state) approach [56] to estimate the derivatives. We have the following result:

THEOREM 2.2 (Fréchet derivatives). *The functional $\mathcal{F}_\alpha(\sigma_a, \sigma_s)$ is Fréchet differentiable with respective to σ_a and σ_s. The derivative at (σ_a, σ_s) in the direction (h_a, h_s) is given by*

$$\begin{pmatrix} \mathcal{F}_\alpha'(\sigma_a, \sigma_s)h_a \\ \mathcal{F}_\alpha'(\sigma_a, \sigma_s)h_s \end{pmatrix} = \begin{pmatrix} \left\langle \varphi, (\frac{\partial \mathcal{T}}{\partial \sigma_a} h_a)u \right\rangle_{L^2(\mathcal{D} \times S^{n-1})} + \alpha \langle \sigma_a, h_a \rangle_{\mathcal{H}^1(\mathcal{D})} \\ \left\langle \varphi, (\frac{\partial \mathcal{T}}{\partial \sigma_s} h_s)u \right\rangle_{L^2(\mathcal{D} \times S^{n-1})} + \alpha \langle \sigma_s, h_s \rangle_{\mathcal{H}^1(\mathcal{D})} \end{pmatrix}, \tag{2.12}$$

5

where φ is the solution of the adjoint problem (2.15) below and $\mathcal{T}$ is defined in (2.1).

Proof. Let us denote by r the residual $\mathcal{G}\Lambda f - z = \mathcal{G}u|_{\Gamma_+} - z$. According to [17, 18], r is Fréchet differentiable with respect to both σ_a and σ_s. The L^2-norm is Fréchet differentiable as shown in [43]. By the chain rule, we conclude that $\mathcal{F}_\alpha$ is differentiable with respect to σ_a and σ_s.

We now compute these Fréchet derivatives. Let us compute the derivative with respect to σ_a:

$$
\begin{aligned}
\mathcal{F}'_\alpha(\sigma_a, \sigma_s)h_a &= \left\langle r, \mathcal{G}(\frac{\partial u|_{\Gamma_+}}{\partial \sigma_a} h_a) \right\rangle_{\mathcal{Z}} + \alpha \left\langle \sigma_a, h_a \right\rangle_{\mathcal{H}^1(\mathcal{D})} \\
&= \left\langle \mathcal{G}^* r, \frac{\partial u|_{\Gamma_+}}{\partial \sigma_a} h_a \right\rangle_{L^2_{\boldsymbol{\theta}\cdot\nu}(\Gamma_+)} + \alpha \left\langle \sigma_a, h_a \right\rangle_{\mathcal{H}^1(\mathcal{D})}
\end{aligned}
\tag{2.13}
$$

where we have used the properties of the adjoint operator (2.7). On the other hand, differentiating the transport equation (2.1) gives:

$$
\begin{aligned}
\mathcal{T}\phi + (\frac{\partial \mathcal{T}}{\partial \sigma_a} h_a)u &= 0 \quad \text{in } \mathcal{D} \times S^{n-1} \\
\phi &= 0 \quad \text{on } \Gamma_-,
\end{aligned}
\tag{2.14}
$$

where $\phi \equiv \dfrac{\partial u}{\partial \sigma_a} h_a$, and $\mathcal{T}$ is the transport operator defined in (2.1). We need also to introduce an adjoint variable φ which is the solution of the following adjoint transport equation:

$$
\begin{aligned}
\mathcal{T}^*\varphi \equiv \left(\frac{i\omega}{v} - \boldsymbol{\theta}\cdot\nabla + \sigma_a(\mathbf{x})\right)\varphi(\mathbf{x}, \boldsymbol{\theta}) + Q(\varphi)(\mathbf{x}, \boldsymbol{\theta}) &= 0 \quad \text{in } \mathcal{D} \times S^{n-1} \\
\varphi(\mathbf{x}, \boldsymbol{\theta}) &= -\mathcal{G}^* r \quad \text{on } \Gamma_+.
\end{aligned}
\tag{2.15}
$$

Here we have used that $Q^* = Q$, which follows from the definition (2.2). Multiplying (2.14) by φ and (2.15) by ϕ, then integrating over $\mathcal{D} \times S^{n-1}$, we obtain

$$
\langle \mathcal{G}^* r, \varphi \rangle_{L^2_{\boldsymbol{\theta}\cdot\nu}(\Gamma_+)} = \left\langle \varphi, (\frac{\partial \mathcal{T}}{\partial \sigma_a} h_a)u \right\rangle_{L^2(\mathcal{D} \times S^{n-1})},
\tag{2.16}
$$

which leads to

$$
\mathcal{F}'_\alpha(\sigma_a, \sigma_s)h_a = \left\langle \varphi, (\frac{\partial \mathcal{T}}{\partial \sigma_a} h_a)u \right\rangle_{L^2(\mathcal{D} \times S^{n-1})} + \alpha \left\langle \sigma_a, h_a \right\rangle_{\mathcal{H}^1(\mathcal{D})}.
\tag{2.17}
$$

The derivative with respect to σ_s can be computed similarly. $\square$

This result shows that in order to compute the Fréchet derivative of the objective functional $\mathcal{F}_\alpha(\sigma_a, \sigma_s)$, we need to solve one forward transport problem (2.1) and one adjoint transport problem (2.15).

3. Discretization methods. There is a vast literature on the discretization of radiative transfer equations; see for instance [1, 24, 42]. In this paper, we have chosen to use the discrete ordinates method to discretize the directional variables and the finite volume method [21] to discretize the spatial variables.

3.1. The discrete ordinates formulation. In the discrete ordinates method [1, 42], we approximate the total scalar flux, defined as the integral of $u(\mathbf{x}, \boldsymbol{\theta})$ over

S^{n-1}, by the following quadrature rule

$$\int_{S^{n-1}} u(\mathbf{x}, \boldsymbol{\theta}) d\mu(\boldsymbol{\theta}) \approx \sum_{j=1}^{J} \eta_j u(\mathbf{x}, \boldsymbol{\theta}_j), \tag{3.1}$$

where $\boldsymbol{\theta}_j$ is the jth direction and η_j the associated weight, for $1 \leq j \leq J$. Details on how to choose the set of directions $\{\boldsymbol{\theta}_j\}_{j=1}^{J}$ and the corresponding weights $\{\eta_j\}_{j=1}^{J}$ can be found in [42]. To ensure particle conservation, we impose that

$$\sum_{j=1}^{J} \eta_j = 1. \tag{3.2}$$

The equation of radiative transfer is now decomposed as a discrete set of J coupled differential equations that describe the photon flux field along J directions:

$$\nabla \cdot (\boldsymbol{\theta}_j u) + (\sigma_t + \frac{i\omega}{v}) u(\mathbf{x}, \boldsymbol{\theta}_j) = \sigma_s(\mathbf{x}) \sum_{j'=1}^{J} \eta_{j'} k_{jj'} u(\mathbf{x}, \boldsymbol{\theta}_{j'}), \tag{3.3}$$

for $j = 1, 2, ..., J$, where $k_{jj'} = k(\boldsymbol{\theta}_j \cdot \boldsymbol{\theta}_{j'})$, and where $\sigma_t = \sigma_a + \sigma_s$. We impose

$$\sum_{j=1}^{J} \eta_j k_{jj'} = 1, \qquad 1 \leq j' \leq J, \tag{3.4}$$

so that the number of photons in the system is preserved by the scattering process.

3.2. Spatial discretization. We use a finite volume method to perform the spatial discretization. Finite volume methods [21] ensure the conservation of mass (or momentum, energy) in a discrete sense, which is important in transport calculations. They also have the advantage of easily handling complicated geometries by arbitrary triangulations, which we need in tomographic applications.

We implement a cell-centered version of the finite volume methods. Consider a mesh of $\mathbb{R}^n$, $\mathcal{M}$, consisting of polyhedral bounded convex subsets of $\mathbb{R}^n$ which covers our computational domain $\mathcal{D}$. Let $\mathcal{C} \in \mathcal{M}$ be a *control cell*, that is an element of the mesh $\mathcal{M}$, $\partial\mathcal{C}$ its boundary, and $V_\mathcal{C}$ its Lebesgue measure. We assume that the unknown quantity, for example $u(\mathbf{x}, \theta_j)$, is constant in $\mathcal{C}$ and denote the value of $u(\mathbf{x}, \boldsymbol{\theta}_j) \equiv u_j$ on $\mathcal{C}$ by $u_j^{\mathcal{C}}$.

Integrating the above discrete ordinates equations (3.3) over cell $\mathcal{C}$ and using the divergence theorem on the first term, we obtain the following equations

$$\int_{\partial\mathcal{C}} \boldsymbol{\theta}_j \cdot \mathbf{n}_{\mathcal{C}}(\mathbf{x}) u_j d\gamma(\mathbf{x}) + (\sigma_t^{\mathcal{C}} + \frac{i\omega}{v}) V_\mathcal{C} u_j^{\mathcal{C}} = V_\mathcal{C} \sigma_s^{\mathcal{C}} \sum_{j'=1}^{J} \eta_{j'} k_{jj'} u_{j'}^{\mathcal{C}}, \tag{3.5}$$

for $1 \leq j \leq J$, where, $\mathbf{n}_{\mathcal{C}}(\mathbf{x})$ denotes the outward normal to $\partial\mathcal{C}$ at point $\mathbf{x} \in \partial\mathcal{C}$, $d\gamma(\mathbf{x})$ denotes the surface Lebesgue measure on $\partial\mathcal{C}$ and $\sigma_s^{\mathcal{C}}$ ($\sigma_t^{\mathcal{C}}$) is the value of σ_s (σ_t) on cell $\mathcal{C}$.

Now we have to approximate the flux through the boundary of $\mathcal{C}$, i.e., the first integral term in equation (3.5). Let $\{\mathcal{C}_i\}_{i=1}^{I}$ be the set of neighboring cells of $\mathcal{C}$. We denote by $S_{\mathcal{C},i}$ the common edge of cell $\mathcal{C}$ and $\mathcal{C}_i$, i.e., $S_{\mathcal{C},i} = \partial\mathcal{C} \cap \partial\mathcal{C}_i$. We then have

$$\int_{\partial\mathcal{C}} \boldsymbol{\theta}_j \cdot \mathbf{n}_{\mathcal{C}}(\mathbf{x}) u_j d\gamma(\mathbf{x}) = \sum_i \int_{S_{\mathcal{C},i}} \boldsymbol{\theta}_j \cdot \mathbf{n}_{\mathcal{C}}(\mathbf{x}) u_j d\gamma(\mathbf{x}). \tag{3.6}$$

The flux $\int_{S_{\mathcal{C},i}} \boldsymbol{\theta}_j \cdot \mathbf{n}_{\mathcal{C}}(\mathbf{x})u_j d\gamma(\mathbf{x})$ can be approximated by various numerical schemes. In this work, we take a first-order upwind scheme:

$$F_{j,i}^{\mathcal{C}} := \int_{S_{\mathcal{C},i}} \boldsymbol{\theta}_j \cdot \mathbf{n}_{\mathcal{C}}(\mathbf{x})u_j d\gamma(\mathbf{x}) = \begin{cases} \boldsymbol{\theta}_j \cdot \mathbf{n}_{\mathcal{C}}|S_{\mathcal{C},i}|u_j^{\mathcal{C}} & \text{if } \boldsymbol{\theta}_j \cdot \mathbf{n}_{\mathcal{C}} \geq 0 \\ \boldsymbol{\theta}_j \cdot \mathbf{n}_{\mathcal{C}}|S_{\mathcal{C},i}|u_j^{\mathcal{C}_i} & \text{if } \boldsymbol{\theta}_j \cdot \mathbf{n}_{\mathcal{C}} < 0, \end{cases} \quad (3.7)$$

where $|S_{\mathcal{C},i}|$ is the Lebesgue measure of $S_{\mathcal{C},i}$. We then obtain a full discretization of the discrete ordinates equations

$$\sum_i F_{j,i}^{\mathcal{C}} + (\sigma_t^{\mathcal{C}} + \frac{i\omega}{v})V_{\mathcal{C}}u_j^{\mathcal{C}} = V_{\mathcal{C}}\sigma_s^{\mathcal{C}} \sum_{j'=1}^{J} \eta_{j'}k_{jj'}u_{j'}^{\mathcal{C}}, \quad (3.8)$$

for $j = 1, 2, ..., J$. Let N denote the total number of control cells. After collecting the discretized transport equation (3.8) on all control cells, we arrive at the following system of complex-valued algebraic equations

$$\mathbf{AU} = \mathbf{SU} + \mathbf{G} \quad (3.9)$$

where $\mathbf{A} \in \mathbb{C}^{NJ \times NJ}$ and $\mathbf{S} \in \mathbb{C}^{NJ \times NJ}$ are the discretized streaming-collision and scattering operators, respectively. The boundary source $f(\mathbf{x}, \boldsymbol{\theta})$, which comes into the discretized system via the flux approximation (3.7) is denoted by $\mathbf{G}$. The vector $\mathbf{U} \in \mathbb{C}^{NJ \times 1}$, which contains the values of $u(\mathbf{x}, \boldsymbol{\theta})$ on the cell $\mathcal{C}$ in the direction $\boldsymbol{\theta}_j$ is organized as

$$\mathbf{U} = (u_1^1, ..., u_1^N, u_2^1, ..., u_2^N, ..., ..., u_J^1, ..., u_J^N)^T.$$

We refer to our earlier work [51] for some numerical tests on the finite volume discretization of the transport equation.

3.3. Discrete adjoint problem. We present in this section the numerical method we have employed to compute the gradient of discrete objective function with respect to the optical properties on each cell.

To simplify the notation, we denote from now on by $\sigma_a \in \mathbb{R}^{N \times 1}$ the absorption coefficient vector $(\sigma_a^1, ..., \sigma_a^{\mathcal{C}}, ..., \sigma_a^N)^T$ and $\sigma_s \in \mathbb{R}^{N \times 1}$ the scattering coefficient vector $(\sigma_s^1, ..., \sigma_s^{\mathcal{C}}, ..., \sigma_s^N)^T$.

The discretized objective function that we want to minimize takes the form

$$\mathcal{F}_\alpha(\sigma_a, \sigma_s) = \frac{1}{2} \sum_{d=1}^{N_d} (\mathcal{P}_d \mathbf{U} - z_d^\delta)^2 + \frac{\alpha}{2} \sum_{\mathcal{C}=1}^{N} \sum_{p=\{a,s\}} \left(\sum_{\kappa=\{x,y,z\}} (\mathcal{D}_\kappa^{\mathcal{C}} \sigma_p)^2 + (\sigma_p^{\mathcal{C}})^2 \right) \quad (3.10)$$

which is equivalent to (2.10) up to a multiplying constant when we have enough detectors. The superscript δ is used to denote the level of noise contained in the measurements. Here $\mathcal{D}_\kappa^{\mathcal{C}} \in \mathbb{R}^{1 \times N}$ denotes the discretized partial differential operator at cell $\mathcal{C}$ in the κ $(= x, y, z)$ direction. $\mathcal{P}_d \in \mathbb{R}^{1 \times N}$ is a discretized version of the operation that takes the outgoing flux and averages over S_+^{n-1}.

One can check that

$$\frac{\partial \mathcal{F}_\alpha}{\partial \sigma_a^{\mathcal{C}}} = \sum_{d=1}^{N_d} (\mathcal{P}_d \mathbf{U} - z_d^\delta)\mathcal{P}_d \frac{\partial \mathbf{U}}{\partial \sigma_a^{\mathcal{C}}} + \alpha \sum_{\mathcal{C}=1}^{N} \left(\sum_{\kappa=\{x,y,z\}} (\mathcal{D}_\kappa^{\mathcal{C}} \sigma_a)(\mathcal{D}_\kappa^{\mathcal{C}} \mathcal{I}_{\mathcal{C}}) + \sigma_a^{\mathcal{C}} \right), \quad (3.11)$$

8

with the unit direction vector $\mathcal{I}_\mathcal{C} \in \mathbb{R}^{N \times 1}$ a vector whose $\mathcal{C}$-th element is 1 and all other components are zero.

At the same time, we notice from (3.9) that

$$\frac{\partial \mathbf{A}}{\partial \sigma_a^\mathcal{C}} \mathbf{U} + \mathbf{A} \frac{\partial \mathbf{U}}{\partial \sigma_a^\mathcal{C}} = \frac{\partial \mathbf{S}}{\partial \sigma_a^\mathcal{C}} \mathbf{U} + \mathbf{S} \frac{\partial \mathbf{U}}{\partial \sigma_a^\mathcal{C}}, \tag{3.12}$$

which is equivalent to saying that

$$\frac{\partial \mathbf{U}}{\partial \sigma_a^\mathcal{C}} = -(\mathbf{A} - \mathbf{S})^{-1} \frac{\partial (\mathbf{A} - \mathbf{S})}{\partial \sigma_a^\mathcal{C}} \mathbf{U}, \tag{3.13}$$

provided that $\mathbf{A} - \mathbf{S}$ is invertible. We thus have

$$\frac{\partial \mathcal{F}_\alpha}{\partial \sigma_a^\mathcal{C}} = -\sum_{d=1}^{N_d} (\mathcal{P}_d \mathbf{U} - z_d^\delta) \mathcal{P}_d (\mathbf{A} - \mathbf{S})^{-1} \frac{\partial (\mathbf{A} - \mathbf{S})}{\partial \sigma_a^\mathcal{C}} \mathbf{U}$$

$$+ \alpha \sum_{\mathcal{C}=1}^{N} \left(\sum_{\kappa=\{x,y,z\}} (\mathcal{D}_\kappa^\mathcal{C} \sigma_a)(\mathcal{D}_\kappa^\mathcal{C} \mathcal{I}_\mathcal{C}) + \sigma_a^\mathcal{C} \right), \tag{3.14}$$

We now introduce a new state variable $\mathbf{V} \in \mathbb{C}^{N \times 1}$ (called adjoint variable of $\mathbf{U}$) given by

$$-\sum_{d=1}^{N_d} (\mathcal{P}_d \mathbf{U} - z_d^\delta) \mathcal{P}_d (\mathbf{A} - \mathbf{S})^{-1} = \mathbf{V}^T. \tag{3.15}$$

We then say that $\mathbf{V}$ is the solution of the following *adjoint* equation of (3.9)

$$(\mathbf{A} - \mathbf{S})^T \mathbf{V} = -\sum_{d=1}^{N_d} (\mathcal{P}_d \mathbf{U} - z_d^\delta) \mathcal{P}_d^T \tag{3.16}$$

One then arrives at

$$\frac{\partial \mathcal{F}_\alpha}{\partial \sigma_a^\mathcal{C}} = \mathbf{V}^T \frac{\partial (\mathbf{A} - \mathbf{S})}{\partial \sigma_a^\mathcal{C}} \mathbf{U} + \alpha \sum_{\mathcal{C}=1}^{N} \left(\sum_{\kappa=\{x,y,z\}} (\mathcal{D}_\kappa^\mathcal{C} \sigma_a)(\mathcal{D}_\kappa^\mathcal{C} \mathcal{I}_\mathcal{C}) + \sigma_a^\mathcal{C} \right). \tag{3.17}$$

Very similar computation leads to the fact that the derivatives of the objective functional with respect to $\sigma_s^\mathcal{C}$ are given by

$$\frac{\partial \mathcal{F}_\alpha}{\partial \sigma_s^\mathcal{C}} = \mathbf{V}^T \frac{\partial (\mathbf{A} - \mathbf{S})}{\partial \sigma_s^\mathcal{C}} \mathbf{U} + \alpha \sum_{\mathcal{C}=1}^{N} \left(\sum_{\kappa=\{x,y,z\}} (\mathcal{D}_\kappa^\mathcal{C} \sigma_s)(\mathcal{D}_\kappa^\mathcal{C} \mathcal{I}_\mathcal{C}) + \sigma_s^\mathcal{C} \right). \tag{3.18}$$

Remark 1. The above results can be easily extended to the case of multiple sources (say, N_s). The objective functional $\mathcal{F}_a(\sigma_a, \sigma_s)$ becomes the sum of contributions from all sources: $\mathcal{F}_\alpha(\sigma_a, \sigma_s) = \sum_{s=1}^{N_s} \mathcal{F}_\alpha^s(\sigma_a, \sigma_s)$.

4. Numerical implementation. We have implemented the quasi-Newton optimization algorithm to solve the regularized least-square problem (**RLS**) introduced in section 2.2. We have found in practice that this method converged much faster than the nonlinear conjugate gradient method. This is expected from theory [49] and is consistent with practical applications tested in [39]. We employ the BFGS update rule [49] of inverse Hessian matrix for our quasi-Newton method. The usual BFGS method, however, requires the explicit construction of the Hessian matrix, which is unrealistic for large problems. The memory size required to store the Hessian matrix is roughly proportional to the square of the memory used for the unknown parameters. We have thus resorted to a limited-memory version of BFGS method which avoids the explicit construction of the inverse Hessian matrix.

4.1. Numerical optimization. The BFGS algorithm can be viewed as a special case of quasi-Newton method [49]. With $\boldsymbol{\sigma} \in \mathbb{R}^{2N \times 1}$ the vector $(\sigma_a^1, .., \sigma_a^N, \sigma_s^1, .., \sigma_s^N)^T$, the quasi-Newton methods can be characterized by the following iterative process:

$$\boldsymbol{\sigma}_{k+1} = \boldsymbol{\sigma}_k + \alpha_k \mathbf{p}_k, \qquad k \in \mathbb{N}^+ \tag{4.1}$$

where $\mathbf{p}_k$ is a descent direction vector and α_k is the step length. The BFGS algorithm chooses $\mathbf{p}_k$ to be the solution of an approximated solution of Newton-type optimality equation, i.e.,

$$\mathbf{p}_k = H_k \mathbf{g}_k, \tag{4.2}$$

where $\mathbf{g}_k$ is the gradient of the least-square functional, $\mathbf{g}_k = -\nabla_{\boldsymbol{\sigma}} \mathcal{F}_\alpha(\boldsymbol{\sigma}_k)$. H_k is the inverse Hessian matrix at step k. Instead of computing real inverse Hessian matrices, which is very time-consuming, the BFGS algorithm chooses to approximate H_k by the following updating rule

$$H_{k+1} = W_k^T H_k W_k + \rho_k \mathbf{s}_k \mathbf{s}_k^T \tag{4.3}$$

with

$$W_k = \mathbf{I} - \rho_k y_k \mathbf{s}_k^T, \quad \mathbf{s}_k = \boldsymbol{\sigma}_{k+1} - \boldsymbol{\sigma}_k, \quad \mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k, \quad \rho_k = \frac{1}{\mathbf{y}_k^T \mathbf{s}_k}. \tag{4.4}$$

As we mentioned above, forming (4.3) takes tremendous computer memory for large problems. To overcome this shortcoming, the limited-memory version of BFGS only stores the vector $\mathbf{y}_k$ and $\mathbf{s}_k$ obtained in the last m ($3 \leq m \leq 7$ usually) iterations [32] and discards the rest. Thus after first m iterations, (4.3) can be expressed as:

$$
\begin{aligned}
H_{k+1} = \ & (V_k^T \cdots V_{k-m}^T) H_{k+1}^0 (V_{k-m} \cdots V_k) \\
& + \ \rho_{k-m}(V_k^T \cdots V_{k-m+1}^T) \mathbf{s}_{k-m} \mathbf{s}_{k-m}^T \times (V_{k-m+1} \cdots V_k) \\
& + \ \rho_{k-m+1}(V_k^T \cdots V_{k-m+2}^T) \mathbf{s}_{k-m+1} \mathbf{s}_{k-m+1}^T \times (V_{k-m+2} \cdots V_k) \\
& \ \vdots \\
& + \ \rho_k \mathbf{s}_k \mathbf{s}_k^T
\end{aligned}
\tag{4.5}
$$

with the sparse initial guess H_{k+1}^0 given by

$$H_{k+1}^0 = \frac{\mathbf{y}_{k+1}^T \mathbf{s}_{k+1}}{\mathbf{y}_{k+1}^T \mathbf{y}_{k+1}} \tag{4.6}$$

We refer interested readers to [10, 32, 49, 60] for more details on the limited-memory BFGS algorithms, and to reference [39] for applications of those algorithms to optical tomographic problems. Convergence of BFGS algorithms has been proved under certain conditions and has been tested on many applications [10, 49].

4.2. Solving algebraic systems. As we have mentioned before, at each step of the minimization process, we have to solve both a discretized transport equation (3.9) and its adjoint problem (3.16) to compute the Fréchet derivatives (3.17) and (3.18) of the objective functional, forming the gradient vector $\mathbf{g}_k$ in (4.2). In fact, almost all of the computational time in the reconstruction process is devoted to the solution of these transport equations. In this work, instead of using the popular source iteration (SI) method, which converges very slowly in diffusive regimes unless it is properly accelerated [1], we choose to solve the forward problems by a preconditioned GMRES(m) algorithm [52, 53], where m denotes the number of iterative steps after which GMRES is restarted. The implementation of the algorithm is based on the template provided in [9]. The preconditioner we employ is the zero fill-in incomplete LU factorization (ILU(0)) [50, 52] that has been proved to be efficient in transport calculations[50]. Details about this factorization can be found in reference [52]. In all of the numerical examples in section 5, we select $m = 7$, and the GMRES algorithm is stopped if the residual is small enough. For example, the stopping criteria $\|\mathbf{G} - (\mathbf{A} - \mathbf{S})\mathbf{U}_k\|_{l^2} \leq 10^{-10}$ is used to solve (3.9).

4.3. Selecting regularization parameter. To choose optimal regularization parameter α in (2.10), we adopt the L-curve method in this study. Although there exists proofs that the L-curve method fails to convergence for certain classes of inverse problems [55], we have observes satisfactory results in our applications. We plot the log of the squared norm of the regularized solution, say, σ_a^α, against the squared norm of the regularized residual, say, r_α, for a range of values of the regularization parameter. The right parameter α is the one at which the L-curve reaches the maximum of its curvature [25, 56]. One can show that the right α maximizes the following curvature function [25, 56]

$$\kappa(\alpha) = -\frac{R(\alpha)S(\alpha)[\alpha R(\alpha) + \alpha^2 S(\alpha)] + [R(\alpha)S(\alpha)]/S'(\alpha)}{[R^2(\alpha) + \alpha^2 S^2(\alpha)]^{3/2}}, \qquad (4.7)$$

where $R(\alpha)$ and $S(\alpha)$ are defined by

$$R(\alpha) := \|r_\alpha\|_{l^2}^2, \qquad S(\alpha) := \|\sigma_a^\alpha\|_{l^2}^2.$$

One notices immediately that the L-curve method requires several reconstructions for any single problem, and thus is very time-consuming. We present in Fig.5.2 (B) and Fig.5.6 (B) the L-curve we have use in example 1 and example 3, respectively, to choose the optimal regularization parameter α.

5. Numerical Examples. We provide in this section three numerical examples that illustrate the performance of our numerical method. In the first example, we reconstruct the spatial distribution of the absorption coefficient while keeping the scattering coefficient fixed. In the second example the spatially varying scattering coefficient is reconstructed, while the absorption coefficient is fixed. Finally we show an example in which both optical properties are reconstructed simultaneously. We focus on 2-dimensional problems here although the discussion in the above sections works in both 2 and 3-dimensional spaces. The setting for our calculations in the following.

5.1. Domain partition. We consider a computational domain of size 2×2 cm^2, denoted by $\mathcal{D} = [0, 2] \times [0, 2]$. We cover the domain by 80×80 cells of uniform size whose nodes are given by

$$\mathcal{D}_h = \{\mathbf{x}_{i,j} = (x_i, y_j), \ x_i = i\Delta x, \ y_j = j\Delta y, \ i, j = 0, 1, ..., 81\},$$

with $\Delta x = \Delta y = 0.025$. The direction space is discretized into 128 uniformly distributed (over $[0, 2\pi)$) directions with identical quadrature weight:

$$S_{\Delta\theta}^1 = \{\boldsymbol{\theta}_i : \boldsymbol{\theta}_i = (i-1) * \Delta\theta, \ i = 0, 1, ..., 128\}.$$

where $\Delta\theta = 2\pi/128$. The above discretizations yield a total number of 819200 unknowns for one forward problem (also true for the corresponding adjoint problem), which means that the matrices $\mathbf{A}$ (and $\mathbf{S}$) in equation (3.9) have 819200^2 elements. In our computations, we store only the non-zero elements of those matrices by the compressed row storage scheme [9], which already requires very large memory storage.

5.2. Generating synthetic data. In all the numerical reconstructions presented in section 5, the "measurement" are synthetic rather than coming from real experimental data. A straightforward way of generating synthetic data is to use the same discretization for the forward model and for the inversion procedure. This may lead to somewhat simplified inversions of the finite dimensional problem, which is often referred to as *inverse crimes* in the inverse problems community [14, p.133]. To avoid committing these "crimes", the data are generated with a discretization twice as fine (in all variables) as the discretization used in the inversions.

In the following sections, our "exact data" will thus be those obtained from the fine mesh calculations. By noisy data, we mean the "exact data" polluted by additional multiplicative noise. Let z_d be the exact data, the "noisy" data are simulated according to the rule: $z_d^\delta = (1 + \delta * \text{random})z_d$, where "random" is a uniformly distributed random variable in $[-1, 1]$ and $\delta \geq 0$ will vary in our numerical simulations.

It should be noted that the "exact data" seen on the coarse grid used in the inversion actually already contain some "noise" because they were generated by the fine mesh calculations.

5.3. Single parameter reconstructions.

Example 1: Reconstructing σ_a. We first reconstruct a small absorbing disc centered at (1.15 cm, 1.15 cm) of radius 0.2 cm embedded in the computational domain. The optical parameters for the background and the disc are $\sigma_a = 0.1$ cm^{-1}, $\sigma_s = 80$ cm^{-1} and $\sigma_a = 0.2$ cm^{-1}, $\sigma_s = 80$ cm^{-1}, respectively. The anisotropy factor $g = 0.90$. The modulation frequency of the source is $\omega = 600$MHz. Those are typical parameter values as they are encountered in biomedical applications. Fig.5.1 shows the real absorption map and the reconstructed maps for different noise levels. We also provide in Fig.5.2 (A) (solid line) the evolution of the normalized objective function versus the iteration step in the case of reconstructions using noise free synthetic data. Note that although the total number of BFGS iterations can be reduced by stricter line search scheme in the optimization algorithm, the total computational time remains almost constant. We have tested many classical line search algorithms. The results shown here use the one that works best for our application.

The quality of the reconstructions is measured as follows. Denote by $M^o \in \mathbb{R}^{80 \times 80}$ ($M^r \in \mathbb{R}^{80 \times 80}$) an exact (reconstructed) quantity, which can be either the absorption or the scattering map. We then define the relative l^2 error between M^o and M^r by:

$$\mathcal{E}_{l^2} = \frac{\|M^r - M^o\|_{l^2}}{\|M^o\|_{l^2}} := \frac{\sqrt{\sum_{i=1}^{80}\sum_{j=1}^{80}(M_{i,j}^r - M_{i,j}^o)^2}}{\sqrt{\sum_{i=1}^{80}\sum_{j=1}^{80}(M_{i,j}^o)^2}}. \tag{5.1}$$

The quality of the reconstruction in the case of an absorbing disc is given by the parameters shown in the second column of Tab.5.1. In Fig.5.2 (B), we display the
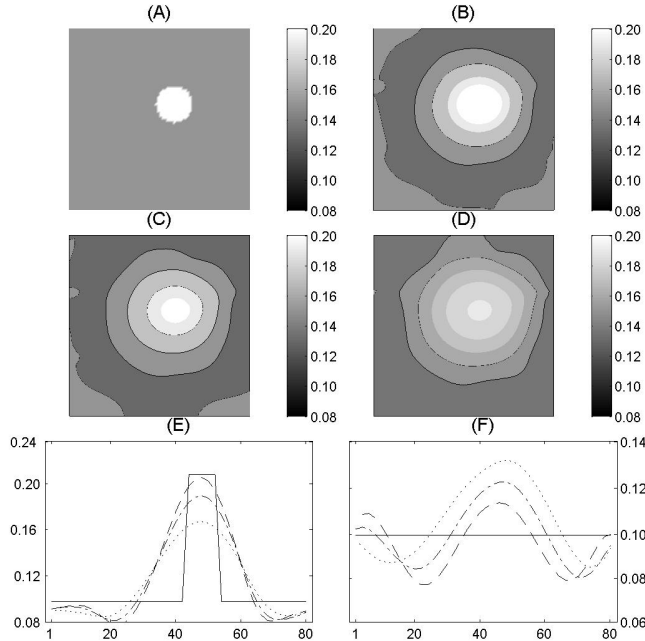
FIG. 5.1. *Maps of the reconstructed absorption coefficient σ_a [cm^{-1}] in Example 1. (A): real absorption map; (B): reconstructed absorption map with exact synthetic data; (C): reconstruction with 10% random noise; (D): reconstruction with 20% random noise; (E): cross sections of map (A) (solid line), (B) (dashed line), (C) (dash-dotted line) and (D) (dotted line) along the bottom left to top right diagonal. (F): same as (E) except that the cross section is along $y = 0.4$.*

L-curve we use to choose the optimal parameter α in the above reconstructions. We mention again that due to the acquisition of our data presented in section 5.2, the noise-free data actually contain "noise" on the coarse grid.

The reconstructions are classical examples of what we may expect in the field. Because of the severe ill-posedness of the inverse problem, it is difficult to reconstruct localized objects unless additional information is included in the reconstruction (which we do not want to do here). Note that the center and the integral of the absorption inclusion are more or less well reconstructed, even if the localized absorption map is somewhat smeared over a relatively large domain (whose size increases as the noise level, whence the regularization parameter α, increases). This is consistent with results obtained from asymptotic theories; see for instance [6].

Example 2: Reconstructing σ_s. We now replace the highly absorbing disc in the previous case by a highly scattering disc at the same position and with the same size. The optical parameters for the background and the disc are $\sigma_a = 0.1$ cm^{-1}, $\sigma_s = 70$ cm^{-1} and $\sigma_a = 0.1$ cm^{-1}, $\sigma_s = 80$ cm^{-1}, respectively. Again, the anisotropy factor $g = 0.9$, the modulation frequency $\omega = 600$MHz. Fig.5.3 shows the exact scattering map and the reconstructed maps for different noise levels. Error estimates are presented the third column of Tab.5.1.

The quality of reconstructions is very similar in the above two examples and is satisfactory. In spite of the fact that stronger regularizations have to be imposed as the
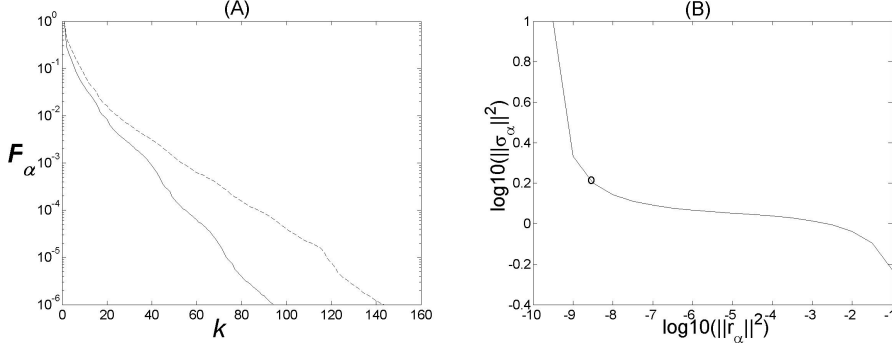
FIG. 5.2. *(A). Evolution of normalized objective functional $\mathcal{F}_\alpha$ with respect to the number of iteration steps k for example 1. Solid line: reconstruction of an absorbing disc; dashed line: reconstruction of a scattering disc. (B). L-curve used to choose optimal regularization parameter α for reconstruction with noise-free data in the reconstruction of an absorbing disc. The circle ($\circ$) denotes the place where α is chosen.*
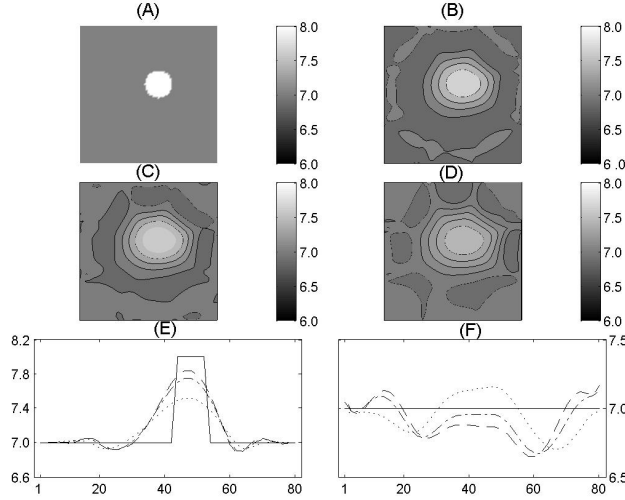

FIG. 5.3. *Maps of reconstructed reduced scattering coefficients $\sigma_s' = (1 - g)\sigma_s$ $[cm^{-1}]$ for example 1. (A): real scattering map. (B): reconstructed scattering map with noise free synthetic data. (C): reconstruction with 10% random noise. (D): reconstruction with 20% random noise. (E): cross sections of map (A) (solid line), (B) (dashed line), (C) (dash-dotted line) and (D) (dotted line) along the bottom left to top right diagonal. (F): same as (E) except that the cross section is along $y = 0.4$.*

noise level increases, the localization and the estimate for the optical parameters in the presence of moderate noise indeed allow us to obtain reasonably accurate information towards diagnostic purposes.

5.4. Frequency-domain versus steady-state. The very reason for introducing frequency-domain reconstructions is that they allow for a better separation between the scattering and absorption properties of the inclusions [45]. Diffusion-based theories show that both coefficients cannot be reconstructed simultaneously without

Cases	absorbing disc			scattering disc		
Noise level	0%	10%	20%	0%	10%	20%
$\alpha \times 10^8$	1.0	2.8	4.6	1.3	2.9	5.7
$\mathcal{E}_{l^2}$	0.064	0.069	0.072	0.065	0.074	0.080

<div align="center">TABLE 5.1</div>

Optimal regularization parameters α and errors in reconstructions for different cases in example 1 and example 2, respectively.

additional geometrical hypotheses [5, 31]. We now show on a numerical example that frequency-domain data indeed substantially improve the reconstruction of both coefficients.
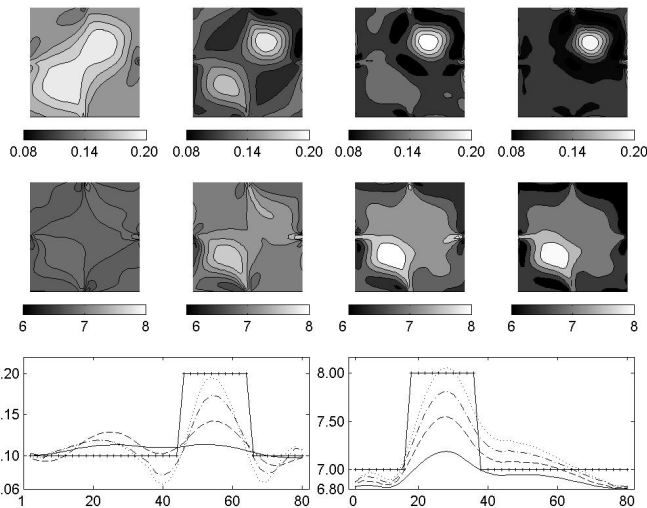
FIG. 5.4. *Top row: Maps of reconstructed absorption coefficients σ_a [cm^{-1}] at BFGS iteration $k = 40$, 80, 120 and 156(final), respectively, for the frequency domain reconstruction in example 3. Middle row: same as the top row but for the reduced reconstructed scattering coefficients $\sigma'_s = (1-g)\sigma_s$ [cm^{-1}]. Bottom row: cross section of real (solid line with +) and reconstructed absorption maps (left), reduced scattering maps (right) along the diagonal at iterations $k = 40$ (solid line), 80 (dashed line), 120 (dash-dotted line) and 156 (dotted line). The reconstructions are done with noise-free synthetic data.*

Example 3: Reconstructing σ_a and σ_s simultaneously. We reconstruct here two small discs of radius 0.2 cm and centered at (1.35 cm, 1.35 cm) and (0.65 cm, 0.65 cm), respectively. The first disc is highly absorbing and the second one is highly scattering. Optical properties for the two discs are $\sigma_a = 0.2$ cm^{-1}, $\sigma_s = 70$ cm^{-1} and $\sigma_a = 0.1$ cm^{-1}, $\sigma_s = 80$ cm^{-1}, respectively. The background parameters are $\sigma_a = 0.1$ cm^{-1} and $\sigma_s = 70$ cm^{-1}. As before, $g = 0.9$, $\omega = 600$MHz.

We compare the reconstructions based on frequency-domain ERT with those based on steady-state ERT. The latter is obtained by setting the frequency $\omega = 0$ in our formulation and keeping everything else the same. We present in Fig.5.4 and Fig.5.5 the reconstructions obtained by the frequency-domain method and the steady-state method, respectively. We also list the parameters which measure the quality of

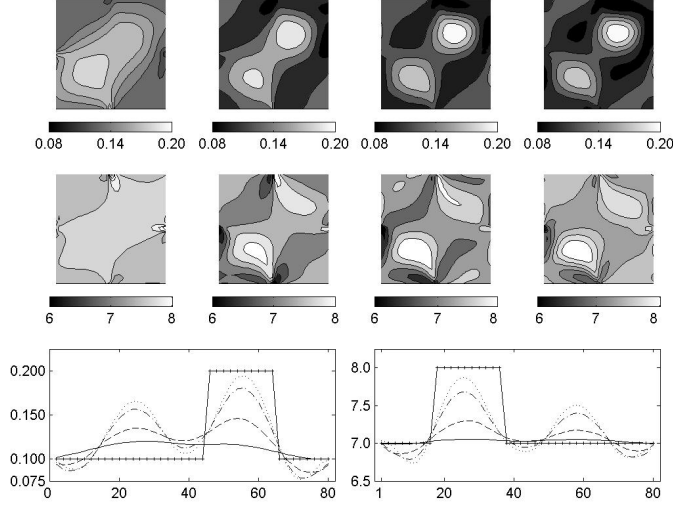the reconstructions at different iteration steps in Tab.5.2.

FIG. 5.5. *Top row: Maps of reconstructed absorption coefficients σ_a [cm^{-1}] at BFGS iteration $k = 40, 80, 120$ and 354(final), respectively, for the steady state reconstruction in example 3. Middle row: same as the top row but for the reduced reconstructed scattering coefficients $\sigma'_s = (1 - g)\sigma_s$ [cm^{-1}]. Bottom row: cross section of real (solid line with +) and reconstructed absorption maps (left), reduced scattering maps (right) along the diagonal at iterations $k = 40$ (solid line), 80 (dashed line), 120 (dash-dotted line) and 354 (dotted line). The reconstructions are done with noise-free synthetic data.*
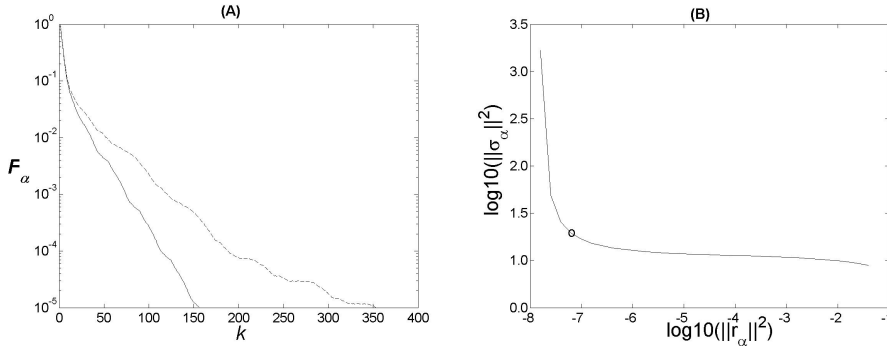

FIG. 5.6. *(A). Evolution of normalized objective functional $\mathcal{F}_\alpha$ with respect to the number of iteration steps k for example 3. Solid line: frequency domain reconstruction of both coefficient simultaneously; dashed line: steady state reconstruction of both coefficients simultaneously. (B.) L-curve used to choose optimal regularization parameter α for reconstruction with noise-free data in the frequency domain simultaneous reconstruction of absorbing and scattering coefficients. The circle (○) denotes the place where α is chosen. Note that here $\|\boldsymbol{\sigma}_\alpha\| := \|(\sigma_a + i\sigma_s)_\alpha\|$.*

We first observe from the results on Fig.5.6 (A) and in Tab.5.2 that the frequency domain reconstruction converges faster than the steady state reconstruction. This has been confirmed in many other geometrical settings we have tested: the speed

of convergence of the steady-state reconstruction presented here is one of the most favorable we have obtained, whereas the speed of convergence of the frequency-domain reconstructions was very often similar to what we have presented here. This issue will be further explored on three-dimensional situations of practical interest in the future.

As far as quality of the reconstruction is concerned, we observe an significant improvement in the frequency domain reconstructions versus the steady-state reconstructions. In both simulations, the stopping time is the same: $\frac{\mathcal{F}_\alpha(\sigma_a^k, \sigma_s^k)}{\mathcal{F}_\alpha(\sigma_a^0, \sigma_s^0)} \leq 10^{-5}$. Although the L^2 errors may not enjoy a dramatic improvement (see Tab.5.2), they are still significantly reduced. More importantly, the last rows in Fig.5.4 and Fig.5.5 show dramatic reductions (at least by a factor 2) of the cross-talk between the absorption and scattering reconstructions: the spurious bumps (left of the left picture on the bottom row and right of the right picture on the bottom row in Figs. 5.4 and 5.5) are clearly much stronger in the steady-state calculations than in the frequency-domain calculations. This is the major advantage of the frequency-domain calculations. As predicted by theory, we have observed that an increases in ω led to reduced cross-talks. How much this effect depends on the choice of the frequency (as well as possible combinations of different frequencies) and the geometrical setting will be further explored in future works.

# Iteration	$k = 40$		$k = 80$		$k = 120$		Final	
	σ_a	σ_s	σ_a	σ_s	σ_a	σ_s	σ_a	σ_s
$\mathcal{E}_{l^2}(f)$	0.121	0.144	0.092	0.112	0.080	0.092	0.063	0.076
$\mathcal{E}_{l^2}(s)$	0.181	0.224	0.127	0.132	0.113	0.112	0.094	0.106

TABLE 5.2
Error estimates for the reconstructions of example 3 for several iteration steps in the optimization process. Here, f refers to frequency-domain calculations and s to steady state calculations.

6. Concluding remarks. We have formulated an inverse problem in optical tomography as a regularized least square problem based on the frequency-domain equation of radiative transfer to model light propagation in biological tissues. In the inversion procedure, the forward model is discretized by using a finite volume method and a discrete ordinates method. We solve the regularized least square problem by using a limited-memory Quasi-Newton method with BFGS type updating rule for the Hessian matrix, and have incorporated positivity constraints and L^∞ bounds on the optical parameters. Numerical reconstructions based on synthetic data provide results that are in agreement with the expected reconstructions. Notably, the crosstalk between two optical parameters is significantly reduced in frequency-domain reconstructions.

The method presented here also overcomes several of the shortcomings of diffusion-equation based optical tomography [4, 15], which provides a very useful tool in many problems but fails to adequately model strongly absorbing regions (e.g. large blood-filled spaces such as brain hematoma), low-scattering void-like inclusions (e.g. spaces filled with cerebrospinal fluid, amniotic fluid, or synovial fluid) and optically relatively thin media such as fingers and small animals. One of the main domains of application of the proposed method will be functional imaging of rheumatoid arthritis (RA) of human finger joints, where the diffusion equation typically fails to generate accurate forward predictions. Extensive numerical reconstructions in practically relevant geometries for this application and based on possibly multi-frequency experimental data will be performed in the near future.

REFERENCES

[1] M. L. ADAMS AND E. W. LARSEN, *Fast iterative methods for discrete-ordinates particle transport calculations*, Prog. Nucl. Energy, 40 (2002), pp. 3–150.

[2] V. AGOSHKOV, *Boundary Value Problems for the Transport Equations*, Birkhauser, Boston, 1998.

[3] R. ARONSON, R. L. BARBOUR, J. LUBOWSKY, AND H. GRABER, *Application of transport theory to infra-red medical imaging*, in Modern Mathematical methods in Transport Theory, W. Greenberg and J. Polewczak, eds., Birkhäuser, 1991.

[4] S. R. ARRIDGE, *Optical tomography in medical imaging*, Inverse Problems, 15 (1999), pp. R41–R93.

[5] S. R. ARRIDGE AND W. R. B. LIONHEART, *Nonuniqueness in diffusion-based optical tomography*, Opt. Lett., 23 (1998), pp. 882–884.

[6] G. BAL, *Optical tomography for small volume absorbing inclusions*, Inverse Problems, 19 (2003), pp. 371–386.

[7] ———, *Transport through diffusive and non-diffusive regions, embedded objects, and clear layers*, SIAM J. Appl. Math., 62 (2002), pp. 1677–1697.

[8] G. BAL AND K. REN, *Generalized diffusion model in optical tomography with clear layers*, J. Opt. Soc. Am. A, 20 (2003), pp. 2355–2364.

[9] R. BARRETT, M. BERRY, T. F. CHAN, J. DEMMEL, J. DONATO, J. DONGARRA, V. EIJKHOUT, R. POZO, C. ROMINE, AND H. V. DER VORST, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, SIAM, Philadelphia, 2nd ed., 1994.

[10] R. H. BYRD, P. LU, J. NOCEDAL, AND C. ZHU, *A limited memory algorithm for bound constrained optimization*, SIAM J. Sci. Comput., 16 (1995), pp. 1190–1208.

[11] M. T. CHAHINE, *Inverse problems in radiative transfer: Determination of atmospheric parameters*, J. Atmos. Sci., 27 (1970), pp. 960–967.

[12] B. CHANCE, R. R. ALFANO, B. J. TROMBERG, AND A. KATZIR, eds., *Optical Tomography and Spectroscopy of Tissue*, vol. V, Bellingham, WA, 2003, SPIE.

[13] M. CHOULLI AND P. STEFANOV, *Reconstruction of the coefficients of the stationary transport equation from boundary measurements*, Inverse Problems, 12 (1996), pp. L19–L23.

[14] D. COLTON AND R. KRESS, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer-Verlag, New York, 1998.

[15] R. DAUTRAY AND J.-L. LIONS, *Mathematical Analysis and Numerical Methods for Science and Technology, Vol 6*, Springer-Verlag, Berlin, 1993.

[16] H. DEHGHANI, D. T. DELPY, AND S. R. ARRIDGE, *Photon migration in non-scattering tissue and the effects on image reconstruction*, Phys. Med. Biol., 44 (1999), pp. 2879–2906.

[17] O. DORN, *A transport-backtransport method for optical tomography*, Inverse Problems, 14 (1998), pp. 1107–1130.

[18] ———, *Scattering and absorption transport sensitivity functions for optical tomography*, Optics Express, 7 (2000), pp. 492–506.

[19] R. A. ELLIOTT, T. DURACZ, N. J. MCCORMICK, AND D. R. EMMONS, *Experimental test of a time-dependent inverse radiative transfer algorithm for estimating scattering parameters*, J. Opt. Soc. Am. A, 5 (1988), pp. 366–373.

[20] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, 1996.

[21] R. EYMARD, T. GALLOUET, AND R. HERBIN, *Finite volume methods*, in Handbook of Numerical Analysis VII, P. Ciarlet and J. L. Lions, eds., North-Holland, Armsterdam, 2000.

[22] M. FIRBANK, S. R. ARRIDGE, M. SCHWEIGER, AND D. T. DELPY, *An investigation of light transport through scattering bodies with non-scattering regions*, Phys. Med. Biol., 41 (1996), pp. 767–783.

[23] L. FUKSHANSKY, N. FUKSHANSKY-KAZARINOVA, AND A. M. V. REMISOWSKY, *Estimation of optical properties in a living tissue by solving the inverse problem of the multiflux radiative transfer*, Appl. Opt., 30 (1991), pp. 3145–3153.

[24] F. GOLSE, S. JIN AND C. D. LEVERMORE, *The convergence of numerical transfer schemes in diffusive regimes I: The discrete-ordinate method*, SIAM J. Numerical Analysis, 36 (1999),

pp. 1333–1369.

[25] P. C. Hansen and D. P. O'Leary, *The use of l-curve in the regularization of discrete ill-posed problems*, SIAM Journal on Scientific Computing, 14 (1993), pp. 1487–1503.

[26] L. G. Henyey and J. L. Greenstein, *Diffuse radiation in the galaxy*, Astrophys. J., 90 (1941), pp. 70–83.

[27] A. H. Hielscher, R. E. Alcouffe, and R. L. Barbour, *Comparison of finite-difference transport and diffusion calculations for photon migration in homogeneous and heterogeneous tissue*, Phys. Med. Biol., 43 (1998), pp. 1285–1302.

[28] A. H. Hielscher, A. Y. Bluestone, G. S. Abdoulaev, A. D. Klose, J. Lasker, M. Stewart, U. Netz, and J. Beuthan, *Near-infrared diffuse optical tomography*, Disease Markers, 18 (2002), pp. 313–337.

[29] A. H. Hielscher, A. Klose, A. Scheel, B. Moa-Anderson, M. Backhaus, U. Netz, and J. Beuthan, *Sagittal laser optical tomography for imaging of rheumatoid finger joints*, Physics in Medicine and Biology, 49 (2004), pp. 1147–1163.

[30] H. Jiang, K. D. Paulsen, U. L. Österberg, B. W. Pogue and M. S. Patterson, *Simultaneous reconstruction of optical absorption and scattering maps turbid media from near-infrared frequency-domain data*, Opt. Lett. 20 (1995), pp. 2128–2130.

[31] V. Isakov, *Inverse Problems for Partial Differential Equations*, Springer-Verlag, New York, 1998.

[32] C. T. Kelley, *Iterative Methods for Optimization*, Frontiers in Applied Mathematics, Society of Industrial and Applied Mathematics, Philadelphia, 1999.

[33] A. Kienle, F. K. Forster, and R. Hibst, *Influence of the phase function on determination of the optical properties of biological tissue by spatially resolved reflectance*, Optics Lett., 26 (2001), pp. 1571–1573.

[34] A. D. Kim and J. B. Keller, *Light propagation in biological tissue*, J. Opt. Soc. Am. A 20 (2003), pp. 92–98.

[35] A. D. Kim and M. Moscoso, *Chebyshev spectral methods for radiative transfer*, SIAM J. Sci. Comput. 23 (2002), pp. 2075–2095.

[36] A. D. Klose, U. Netz, J. Beuthan, and A. H. Hielscher, *Optical tomography using the time-independent equation of radiative transfer. part 1: Forward model*, J. Quant. Spectrosc. Radiat. Transfer, 72 (2002), pp. 691–713.

[37] A. D. Klose and A. H. Hielscher, *Iterative reconstruction scheme for optical tomography based on the equation of radiative transfer*, Med. Phys., 26 (1999), pp. 1698–1707.

[38] ———, *Optical tomography using the time-independent equation of radiative transfer. part 2: Inverse model*, J. Quant. Spectrosc. Radiat. Transfer, 72 (2002), pp. 715–202.

[39] ———, *Quasi-Newton methods in optical tomographic image reconstruction*, Inverse Problems, 19 (2003), pp. 387–409.

[40] E. W. Larsen, *Solution of the inverse problem in multigroup transport theory*, J. Math. Phys., 22 (1981), pp. 158–160.

[41] ———, *Solution of three dimensional inverse transport problems*, Transport Theory and Statistical Physics, 17 (1988), pp. 147–167.

[42] E. E. Lewis and W. F. Miller, *Computational Methods of Neutron Transport*, American Nuclear Society, La Grange Park, IL, 1993.

[43] E. H. Lieb and M. Loss, *Analysis*, American Mathematical Society, Providence, RI, Second ed., 2001.

[44] V. Markel and J. C. Schotland, *Effects of sampling and limited data in optical tomography*, App. Phys. Lett. 81 (2002), pp. 1180–1182.

[45] T. O. McBride, B. W. Pogue, U. L. Österberg, and K. D. Paulsen, *Separation of absorption and scattering heterogeneities in NIR tomographic imaging of tissue*, in OSA Technical Digest, Washington DC, 2000, Optical Society of America. Biomedical Topical Meetings.

[46] N. J. McCormick, *Recent developments in inverse scattering transport methods*, Transport Theory and Statistical Physics, 13 (1984), pp. 15–28.

[47] ———, *Inverse radiative transfer problems: A review*, Nuclear Science and Engineering, 112 (1992), pp. 185–198.

[48] M. Mokhtar-Kharroubi, *Mathematical Topics in Neutron Transport Theory: New Aspects*, Advances in Mathematics for Applied Sciences, World Scientific, Singapore, 1997.

[49] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer-Verlag, New York, 1999.

[50] B. W. Patton and J. P. Holloway, *Application of preconditioned GMRES to the numerical solution of the neutron transport equation*, Annals of Nuclear Energy, 29 (2002), pp. 109–136.

[51] K. Ren, G. Abdoulaev, G. Bal, and A. H. Hielscher, *Algorithm for solving the equation of radiative transfer in the frequency domain*, Optics Lett., 29 (2004), pp. 578–580.

[52] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, 2nd ed., 2003.

[53] Y. SAAD AND M. H. SCHULTZ, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.

[54] J. C. SCHOTLAND, *Continuous-wave diffusion imaging*, J. Opt. Soc. Am. A 14 (1997), pp. 275-279.

[55] C. R. VOGEL, *Non-convergence of the l-curve regularization parameter selection method*, Inverse Problems, 12 (1996), pp. 535–547.

[56] ———, *Computational Methods for Inverse Problems*, Frontiers in Applied Mathematics, SIAM, Philadelphia, 2002.

[57] A. P. WANG AND S. UENO, *An inverse problem in a three-dimensional radiative transfer*, Astrophys. Space Sci., 155 (1989), pp. 105–111.

[58] A. J. WELCH AND M. J. C. VAN-GEMERT, *Optical-thermal Response of Laser Irradiated Tissue*, Plenum Press, New York, 1995.

[59] A. G. YODH AND B. CHANCE, *Spectroscopy and imaging with diffusing light*, Physics Today, 48 (1995), pp. 34–40.

[60] C. ZHU, R. H. BYRD, P. LU, AND J. NOCEDAL, *L-BFGS-B–FORTRAN subroutines for large-scale bound constrained optimization*, ACM Trans. Math. Software, 23 (1997), pp. 550–560.