

TOPICAL REVIEW**Optical tomography in medical imaging**

S R Arridge

Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK

Received 6 July 1998, in final form 4 November 1998

Abstract. We present a review of methods for the forward and inverse problems in optical tomography. We limit ourselves to the highly scattering case found in applications in medical imaging, and to the problem of absorption and scattering reconstruction. We discuss the derivation of the diffusion approximation and other simplifications of the full transport problem. We develop sensitivity relations in both the continuous and discrete case with special concentration on the use of the finite element method. A classification of algorithms is presented, and some suggestions for open problems to be addressed in future research are made.

 This article features multimedia enhancements available from the abstract page in the online journal; see <http://www.iop.org>.

1. Introduction

Optical tomography has come to mean the use of low-energy visible or near infra-red light to probe highly scattering media, in order to derive qualitative or quantitative images of the optical properties of these media. Of the potential applications, the one that has received a great deal of attention is medical imaging, where optical tomography is hoped to be a low-cost alternative or complement to existing medical imaging technology, with the particular advantage of providing *functional* as opposed to anatomical information. The name is something of a misnomer since the term already exists to describe the use of light to image fast phenomena in non-scattering regions such as gas combustion (see, for example, [1, 2]).

In experimental systems, light is guided by fibre optics to the surface of the subject and detecting fibres are used to measure the transilluminated light. Thus the inverse problem is one of the recovery of coefficients in a domain from data on its boundary, and in many aspects can be considered to be similar to other well-developed fields of research. There are several aspects of the problem however, that make it unusual, and provide a potentially rich set of research topics:

- The data can be acquired either as time-varying intensities giving the system response to an ultra-short input pulse, or as steady-state complex intensity with measureable amplitude and phase. Because there is a high degree of control available over the distribution in space and time (or frequency and phase) of the sources, there is the potential to design optimal acquisition systems in a realizable fashion.
- The forward problem can be interpreted in a variety of ways, either as a particle or wave phenomenon. In the latter framework, the governing equations can actually be set up in any of a continuum of models that range from a transport model at one extreme to a parabolic (or elliptic) partial differential equation at the other extreme.

Table 1. Definition of variables.

Quantity	Meaning	Dimension
\mathbf{r}	position vector	(L, L, L)
\mathbf{m}	measurement position vector constrained to a surface $\partial\Omega$	(L, L, L)
$\hat{\nu}$	outward normal to boundary	
$\hat{\nu}_\perp$	tangent vector on boundary	
$\hat{\mathbf{s}} = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix}$	unit vector in S^{n-1}	
$\phi(\mathbf{r}, \hat{\mathbf{s}}, t)$	number of photons per unit volume at position \mathbf{r} at time t with velocity in direction $\hat{\mathbf{s}}$	L^{-3}
$\mu_s(\mathbf{r})$	scattering cross section at position \mathbf{r}	L^{-1}
$\mu_a(\mathbf{r})$	absorption cross section at position \mathbf{r}	L^{-1}
$\mu_{tr}(\mathbf{r}) = \mu_s(\mathbf{r}) + \mu_a(\mathbf{r})$	transport cross section at position \mathbf{r}	L^{-1}
c	velocity of light	LT^{-1}
$cq(\mathbf{r}, \hat{\mathbf{s}}, t)$	number of photons per unit volume per unit time sourced at position \mathbf{r} at time t with velocity in direction $\hat{\mathbf{s}}$	$L^{-3}T^{-1}$
$\Theta(\hat{\mathbf{a}}, \hat{\mathbf{b}}) = \Theta(\hat{\mathbf{a}} \cdot \hat{\mathbf{b}})$	normalized phase function representing the probability of scattering from direction $\hat{\mathbf{a}}$ to direction $\hat{\mathbf{b}}$	
$\chi = (\mu_a/\kappa)^{\frac{1}{2}}$	effective cross section	L^{-1}
$\hat{\chi}(\omega) = \left(\frac{\mu_a + i\omega/c}{\kappa} \right)^{\frac{1}{2}}$	complex effective cross section	L^{-1}
Special functions:		
$Y_{l,m}(\hat{\mathbf{s}})$	spherical harmonic order l degree m	
$P_l^m(x)$	associated Legendre polynomial order l degree m	
$P_l(x)$	Legendre polynomial order l	
$Q_l(x)$	Legendre polynomial of the second kind, order l	
$K(x)$	modified Bessel function of the second kind	
$k(x)$	modified spherical Bessel function of the second kind	
$B(n+1)$	Gamma function	
Miscellaneous:		
$\binom{n}{r} = \frac{n!}{r!(n-r)!}$	binomial coefficient	

Table 2. Definition of operators.

Operator	Meaning	Mapping
\mathcal{P}_j	projection operator for source j	$(X^{(a)}, X^{(b)}) \rightarrow Y$
\mathcal{P}'_j	Fréchet derivative of \mathcal{P}_j	$(X^{(a)}, X^{(b)}) \rightarrow Y$
\mathcal{P}''_j	Fréchet derivative of \mathcal{P}'_j	$(X^{(a)}, X^{(b)}) \times (X^{(a)}, X^{(b)}) \rightarrow Y$
\mathcal{M}_j	discrete measurement at points $\{\mathbf{m}_j\}$, $j = 1 \dots M_j$ for source j	$\mathbb{R}^D \rightarrow \mathbb{R}^{M_j}$
\mathcal{S}	spatial filter on boundary	$\partial\Omega \rightarrow \mathbb{R}$
\mathcal{T}	temporal filter operator	$[0, T] \rightarrow \mathbb{R}$
\mathcal{E}	DC intensity measurement operator	$[0, T] \rightarrow \mathbb{R}$

- A variety of different coefficients can be considered in the inverse problem, and the degree to which any of these are significant appears to be critical in interpreting the usefulness of the results.

In respect of the last point, we limit ourselves in this paper primarily to the problem

we have called ‘optical absorption and scattering tomography’ (OAST). A related problem is the concept of ‘fluorescence lifetime imaging’ (FLI). Here sources are applied at one optical wavelength (stimulation wavelength λ_1) and the resulting field excites a contrast agent with quantum efficiency f to a metastable state, which produces emission at another wavelength λ_2 as a Poisson process with time-constant τ . The source for the second process is therefore the field from the first process convolved with an exponential decay (in the time domain) or scaled by a Lorentzian-like function (in the frequency domain) [3]. The forward and inverse problems for this case can be formulated in a broadly similar way with an increase in the number of coefficients to be recovered at each point. The formal aspects of this problem have received less attention than OAST, and are left out of the current discussion.

One of the most confusing aspects of optical tomography is as to the appropriate range of coefficients that define the ‘operating point’ of interest. Light in tissue is not only heavily scattered but also attenuated. Therefore the degree of signal realistically obtainable varies by orders of magnitude if the operating point is changed. Furthermore, except in lossless or weakly absorbing applications the dynamic range of the data on the boundary is so large that data acquisition methods have special difficulty in developing an equivalent signal amplitude at all measurement positions. Figure 1 shows a schematic representation of the ranges of interest of the two main applications, breast imaging and brain imaging, together with some of the values used in reported results. Results presented in this paper are given for parameters relevant to brain imaging. Note that the mammography application has a lower degree of both scattering and attenuation which make it in some ways more tractable.

We may define three types of problem:

- (i) Relative or difference imaging: here data is assumed to be available in two sets with some state difference of significance having taken place between their acquisition. The reconstruction sought is a *dynamic* one and may be approximated by a linear representation.
- (ii) Absolute imaging: here it is assumed that only one set of data is available but that enough information is available, for example by calibration, to allow a model to be fitted nonlinearly.
- (iii) Imaging without amplitude information: here it is not assumed that absolute numbers of photons can be determined, only their relative numbers at different spatial, frequency, or temporal windows.

These three types of problem are of increasing difficulty and results need to be set in the context of which problem is being addressed.

In this paper we summarize the theoretical approaches to the inverse problem. The aspects of instrumentation and physical modelling are very complex and have been surveyed recently in other papers [4, 5]. Clinical applications have also been the subject of some recent very good reviews [6]. The reader is encouraged to refer to these papers to appreciate a wider understanding of the technical difficulties and potential applications.

The paper is organized as follows. In section 2 we make some general definitions about the operators and spaces being employed and define notation. In section 3 we discuss in detail the derivation of various approximations to the transport problem. This section is largely bookwork and can be omitted if the reader is already familiar with the assumptions. In section 4 we define the forward problem and in section 5 define the sensitivity relations in both direct and adjoint forms. In section 6 we describe computational methods, in particular the finite element and finite difference schemes. In section 7 we give the equivalent sensitivity relations in the discrete framework. In section 8 we discuss some aspects of uniqueness and illposedness. In section 9 we summarize the reconstruction algorithms that have been reported. In section 10 we present some conclusions and suggestions for future work.

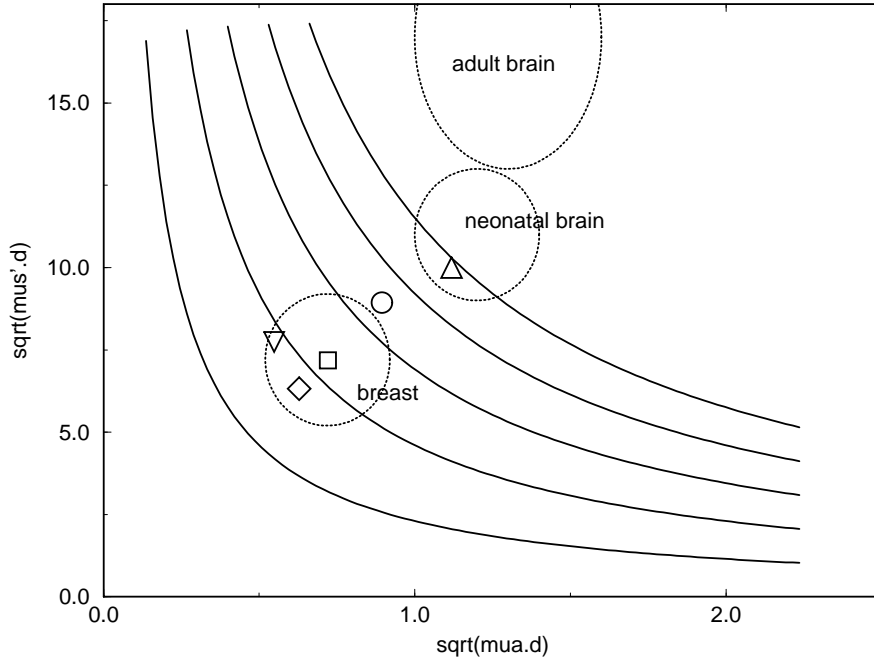


Figure 1. Range of optical properties of tissues of interest and operating point of reported imaging results. Abscissa and ordinate axes are $(\mu_a d)^{\frac{1}{2}}$ and $(\mu_s' d)^{\frac{1}{2}}$ respectively where d is the size of the object being reconstructed and (μ_a, μ_s') are the absorption and (reduced) scattering coefficients. Curved lines are of constant $(\mu_a \mu_s')^{\frac{1}{2}} d$ and represent approximately equal attenuation, at intervals of powers of 10. Dotted circles represent approximate values for tissues given in the literature [7–9]. ∇ [8, 10], \triangle [11–13], \diamond [14], \square [15–18], \circ [19].

2. Formulation of the problem

OAST is an example of an inverse problem over two independent spaces, $(X^{(a)}, X^{(b)})$, illustrated in figure 2. In this framework we will be able to identify both the *transport problem* based on the Boltzmann equation, and the *diffusion problem* based on the simplest non-trivial approximation in spherical harmonics of the Boltzmann equation. In the former, (a, b) are the absorption coefficient μ_a and scattering coefficient μ_s , respectively. In the latter a is the same absorption coefficient but b can be interpreted in a variety of ways of which the diffusion coefficient κ is the most usual.

Let Ω be the domain under consideration, with surface $\partial\Omega$. We will define the *forward problem* as

Given sources $\{q\}$ on $\partial\Omega$ and $(a, b) \in (X^{(a)}(\Omega), X^{(b)}(\Omega))$ find the data $\{y\}$ on $\partial\Omega$,

and the *inverse problem* as

Given sources $\{q\}$ on $\partial\Omega$ and data $\{y\}$ on $\partial\Omega$, find $(a, b) \in (X^{(a)}(\Omega), X^{(b)}(\Omega))$.

Consider S source positions $\mathbf{p}_j \in \partial\Omega$ ($j = 1 \dots S$) and M_j measurement positions $\mathbf{m}_{j,i} \in \partial\Omega$ for the j th source ($i = 1 \dots M_j$), resulting in a total number of measurements $M_{TOT} = \sum_{j=1}^S M_j$, with $M_{UNIQ} \leq M_{TOT}$ distinct measurement positions. The forward problem is nonlinear and is represented by

$$\vec{y} = \vec{\mathcal{P}}[a, b]. \quad (2.1)$$

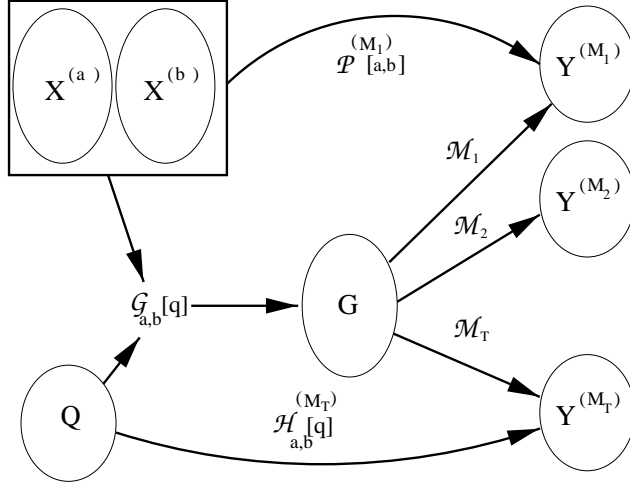


Figure 2. Spaces and operators used in optical tomography (based on figure 3.1 in [87]). $(X^{(a)}, X^{(b)})$ are the solution spaces, Q is the space of sources, G is the space of solutions to the governing equation, Y^{M_d} are the data spaces. \mathcal{G} is the Green function operating on a source, and \mathcal{M}_d are the measurement operators, operating on the solutions to the governing equation to give the data. \mathcal{P} is the forward operator that maps the solution directly to the data. For completeness, the operator \mathcal{H}^{M_d} is defined as the Dirichlet-to-Neumann map for data type M_d .

Under the usual assumptions of a system corrupted by multivariate Gaussian noise, and a maximum-likelihood approach to the solution, we may define the inverse problem as the optimization of an objective function

$$I = \sum_{j=1}^S I_j = \frac{1}{2} \sum_{j=1}^S \sum_{i=1}^{M_j} \left(\frac{y_{j,i} - \mathcal{P}_{j,i}[a, b]}{\sigma_{j,i}} \right)^2 \quad (2.2)$$

which we write in vector form as

$$I = \frac{1}{2} (\vec{y} - \vec{F})^T \mathbf{R}^{-2} (\vec{y} - \vec{F}) = \frac{1}{2} \vec{b}^T \vec{b} \quad (2.3)$$

where $y_{j,i}$ is the data for the i th measurement from source j with standard deviation $\sigma_{j,i}$, $F_{j,i}$ is the modelled data for this source-detector pair, $\mathcal{P}_{j,i}[a, b] = \mathcal{P}_j(\mathbf{m}_{j,i})[a, b]$ is the scalar operator mapping $(X^{(a)}, X^{(b)})$ to $F_{j,i}$ at position $\mathbf{m}_{j,i}$, and $b_{j,i} = \sigma_{j,i}^{-1}(y_{j,i} - F_{j,i})$ is the *residual data* for this measurement. We use subscripted vectors \mathbf{y}_j , \mathbf{F}_j , σ_j to represent all the relevant quantities for a single source j . Thus $\mathcal{P}_j[a, b]$ is the *projection operator* $\mathcal{P}_j : (X^{(a)}, X^{(b)}) \rightarrow Y$ for source j , \mathbf{F}_j is the *projection data* obtained by sampling \mathcal{P}_j at the measurement positions $\{\mathbf{m}_{j,1}, \mathbf{m}_{j,2}, \dots, \mathbf{m}_{j,M_j}\}$, and $\mathbf{b}_j = \mathbf{R}_j^{-1}(\mathbf{y}_j - \mathbf{F}_j)$.

We use \vec{y} , \vec{F} , \vec{b} , as the total vectors, length M_{TOT} , and $\vec{\mathcal{P}}$ as the combined projection operator for all sources.

$$\vec{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_S \end{bmatrix}; \quad \vec{F} = \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \\ \vdots \\ \mathbf{F}_S \end{bmatrix}; \quad \vec{b} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_S \end{bmatrix}; \quad \vec{\mathcal{P}} = \begin{bmatrix} \mathcal{P}_1 \\ \mathcal{P}_2 \\ \vdots \\ \mathcal{P}_S \end{bmatrix}. \quad (2.4)$$

\mathbf{R} is the data-space correlation matrix, which we here take to be

$$\mathbf{R} = \text{diag}(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_S) = \text{diag}(\sigma_{1,1}, \sigma_{1,2}, \dots, \sigma_{j,i}, \dots, \sigma_{S,M_s}). \quad (2.5)$$

We will assume that $(X^{(a)}, X^{(b)}) = (L^\infty(\Omega), L^\infty(\Omega))$, which restricts a and b to be within strictly positive lower and upper bounds. In the transport problem G is taken to be $L^1(\Omega \times S^{n-1})$ in the frequency domain and $L^1(\Omega \times S^{n-1} \times [0, T])$ in the time domain, but in the diffusion problem the question of what space to use has not been resolved and is an important topic for further research. In the closest analogous problem of electrical impedance tomography (EIT), it is usually considered to be a Sobolev space $H^1(\Omega)$ with $Q = H^{\frac{1}{2}}(\Omega)$ and $Y = H^{-\frac{1}{2}}(\Omega)$. Note that we use the term ‘Dirichlet-to-Neumann’ operator for the mapping from Q to Y by analogy with the literature on PDEs [20], although this is certainly a misnomer since it is more typical to use interior sources and Robin boundary conditions for this problem.

For FLI, the corresponding forward problem is

Given sources $\{q\}$ on $\partial\Omega$, $(a(\lambda_1), b(\lambda_1)), (a(\lambda_2), b(\lambda_2)) \in (X^{(a)}(\Omega), X^{(b)}(\Omega))$ and f, τ in Ω
find the data $\{y(\lambda_1)\}, \{y(\lambda_2)\}$ on $\partial\Omega$,

and the inverse problem is in the most general form

Given sources $\{q\}$ on $\partial\Omega$ and data $\{y(\lambda_1)\}, \{y(\lambda_2)\}$ on $\partial\Omega$,
find $(a(\lambda_1), b(\lambda_1), a(\lambda_2), b(\lambda_2)) \in (X^{(a)}(\Omega), X^{(b)}(\Omega))$ and f, τ in Ω .

In fact the general inverse problem is usually simplified by assuming several of the coefficients known, typically $(a(\lambda_1), b(\lambda_1), a(\lambda_2), b(\lambda_2))$ and/or τ .

3. Photon transport models

The Boltzmann transport equation is written as

time domain:

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \hat{\mathbf{s}} \cdot \nabla + \mu_{\text{tr}}(\mathbf{r}) \right) \phi(\mathbf{r}, \hat{\mathbf{s}}, t) = \mu_{\text{s}}(\mathbf{r}) \int_{S^{n-1}} \Theta(\hat{\mathbf{s}} \cdot \hat{\mathbf{s}}') \phi(\mathbf{r}, \hat{\mathbf{s}}', t) d\hat{\mathbf{s}}' + q(\mathbf{r}, \hat{\mathbf{s}}, t) \quad (3.1)$$

frequency domain:

$$\left(\frac{i\omega}{c} + \hat{\mathbf{s}} \cdot \nabla + \mu_{\text{tr}}(\mathbf{r}) \right) \phi(\mathbf{r}, \hat{\mathbf{s}}, \omega) = \mu_{\text{s}}(\mathbf{r}) \int_{S^{n-1}} \Theta(\hat{\mathbf{s}} \cdot \hat{\mathbf{s}}') \phi(\mathbf{r}, \hat{\mathbf{s}}', \omega) d\hat{\mathbf{s}}' + q(\mathbf{r}, \hat{\mathbf{s}}, \omega). \quad (3.2)$$

In the following we will alternate between the time and frequency domain representations as convenient, with the parameter t or ω being sufficient to identify which is being used, and making use of the relation

$$\frac{\partial}{\partial t} \Leftrightarrow i\omega.$$

Two derived quantities that are of interest are

$$\text{photon density:} \quad \Phi(\mathbf{r}, t) = \int_{S^{n-1}} \phi(\mathbf{r}, \hat{\mathbf{s}}, t) d\hat{\mathbf{s}} \quad (3.3)$$

$$\text{photon current:} \quad \mathbf{J}(\mathbf{r}, t) = \int_{S^{n-1}} \hat{\mathbf{s}} \phi(\mathbf{r}, \hat{\mathbf{s}}, t) d\hat{\mathbf{s}} \quad (3.4)$$

which define the *measurables* (see section 4).

3.1. The P_N approximations

The P_N approximations are obtained by spherical harmonic expansion of the quantities in (3.1). Relevant properties of spherical harmonics are summarized in the appendix. We express the

quantities in (3.1) as

$$\phi(\mathbf{r}, \hat{\mathbf{s}}, t) = \sum_l^\infty \sum_{m=-l}^l \left(\frac{2l+1}{4\pi} \right)^{\frac{1}{2}} \psi_{l,m}(\mathbf{r}, t) Y_{l,m}(\hat{\mathbf{s}}) \quad (3.5)$$

$$q(\mathbf{r}, \hat{\mathbf{s}}, t) = \sum_l^\infty \sum_{m=-l}^l \left(\frac{2l+1}{4\pi} \right)^{\frac{1}{2}} q_{l,m}(\mathbf{r}, t) Y_{l,m}(\hat{\mathbf{s}}) \quad (3.6)$$

where the normalization factor $((2l+1)/4\pi)^{1/2}$ is introduced for convenience. The phase function can also be expressed, using (A.3), as:

$$\begin{aligned} \Theta(\hat{\mathbf{s}} \cdot \hat{\mathbf{s}}') &= \sum_l^\infty \left(\frac{2l+1}{4\pi} \right) \Theta_l P_l(\cos \vartheta) \\ &= \sum_l^\infty \sum_{m=-l}^l \Theta_l Y_{l,m}^*(\hat{\mathbf{s}}') Y_{l,m}(\hat{\mathbf{s}}). \end{aligned} \quad (3.7)$$

From (3.3) and (3.4) it follows that

$$\Phi(\mathbf{r}, t) = \psi_{0,0}(\mathbf{r}, t) \quad (3.8)$$

and

$$\mathbf{J}(\mathbf{r}, t) = \begin{pmatrix} \frac{1}{\sqrt{2}}(\psi_{1,-1}(\mathbf{r}, t) - \psi_{1,1}(\mathbf{r}, t)) \\ \frac{1}{i\sqrt{2}}(\psi_{1,-1}(\mathbf{r}, t) + \psi_{1,1}(\mathbf{r}, t)) \\ \psi_{1,0}(\mathbf{r}, t) \end{pmatrix}. \quad (3.9)$$

Equation (3.1) now becomes

$$\begin{aligned} &\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_{\text{tr}}(\mathbf{r}) \right) \sum_l^\infty \sum_{m=-l}^l \left(\frac{2l+1}{4\pi} \right)^{\frac{1}{2}} \psi_{l,m}(\mathbf{r}, t) Y_{l,m}(\hat{\mathbf{s}}) \\ &+ \sum_l^\infty \sum_{m=-l}^l \left(\frac{2l+1}{4\pi} \right)^{\frac{1}{2}} \hat{\mathbf{s}} \cdot \nabla \psi_{l,m}(\mathbf{r}, t) Y_{l,m}(\hat{\mathbf{s}}) \\ &= \mu_s(\mathbf{r}) \int_{S^{n-1}} \left(\sum_l^\infty \sum_{m=-l}^l \Theta_l Y_{l,m}^*(\hat{\mathbf{s}}') Y_{l,m}(\hat{\mathbf{s}}) \right) \\ &\quad \times \left(\sum_l^\infty \sum_{m=-l}^l \left(\frac{2l+1}{4\pi} \right)^{\frac{1}{2}} \psi_{l,m}(\mathbf{r}, t) Y_{l,m}(\hat{\mathbf{s}}') \right) d\hat{\mathbf{s}}' \\ &+ \sum_l^\infty \sum_{m=-l}^l \left(\frac{2l+1}{4\pi} \right)^{\frac{1}{2}} q_{l,m}(\mathbf{r}, t) Y_{l,m}(\hat{\mathbf{s}}). \end{aligned} \quad (3.10)$$

By taking the inner product with $Y_{l,m}^*(\hat{\mathbf{s}})$ the terms in (3.10) decouple with the exception of the transport term $\hat{\mathbf{s}} \cdot \nabla$, which requires the use of the recurrence relations (A.5)–(A.7). After some algebra we obtain the infinite set of coupled equations

$$\begin{aligned} &\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_{\text{tr}}(\mathbf{r}) \right) \psi_{l,m}(\mathbf{r}, t) + \frac{1}{2l+1} \left(\frac{\partial}{\partial z} [(l+1-m)^{\frac{1}{2}} (l+1+m)^{\frac{1}{2}} \psi_{l+1,m}(\mathbf{r}, t) \right. \\ &\quad \left. + (l-m)^{\frac{1}{2}} (l+m)^{\frac{1}{2}} \psi_{l-1,m}(\mathbf{r}, t) \right] \\ &\quad - \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) [(l+m)^{\frac{1}{2}} (l+m-1)^{\frac{1}{2}} \psi_{l-1,m-1}(\mathbf{r}, t) \\ &\quad - (l-m+2)^{\frac{1}{2}} (l-m+1)^{\frac{1}{2}} \psi_{l+1,m-1}(\mathbf{r}, t) \end{aligned}$$

$$\begin{aligned}
& -\frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) [-(l-m)^{\frac{1}{2}} (l-m-1)^{\frac{1}{2}} \psi_{l-1,m+1}(\mathbf{r}, t) \\
& + (l+m+1)^{\frac{1}{2}} (l+m+2)^{\frac{1}{2}} \psi_{l+1,m+1}(\mathbf{r}, t)] \\
& = \mu_s(\mathbf{r}) \Theta_l \psi_{l,m}(\mathbf{r}, t) + q_{l,m}(\mathbf{r}, t).
\end{aligned} \tag{3.11}$$

The P_N approximation is obtained by assuming $\psi_{l,m} = 0, l > N$, giving rise to $(N+1)^2$ coupled first-order PDEs in the general case.

3.2. P_1 approximation

For the $N = 1$ case we obtain four equations

$$\begin{aligned}
\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_{\text{tr}}(\mathbf{r}) \right) \psi_{0,0}(\mathbf{r}, t) + \frac{\partial}{\partial z} \psi_{1,0}(\mathbf{r}, t) + \frac{1}{\sqrt{2}} \frac{\partial}{\partial x} (\psi_{1,-1}(\mathbf{r}, t) - \psi_{1,1}(\mathbf{r}, t)) \\
+ \frac{1}{i\sqrt{2}} \frac{\partial}{\partial y} (\psi_{1,-1}(\mathbf{r}, t) + \psi_{1,1}(\mathbf{r}, t)) = \mu_s(\mathbf{r}) \Theta_0 \psi_{0,0}(\mathbf{r}, t) + q_{0,0}(\mathbf{r}, t)
\end{aligned} \tag{3.12}$$

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_{\text{tr}}(\mathbf{r}) \right) \psi_{1,0}(\mathbf{r}, t) + \frac{1}{3} \frac{\partial}{\partial z} \psi_{0,0}(\mathbf{r}, t) = \mu_s(\mathbf{r}) \Theta_1 \psi_{1,0}(\mathbf{r}, t) + q_{1,0}(\mathbf{r}, t) \tag{3.13}$$

$$\begin{aligned}
\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_{\text{tr}}(\mathbf{r}) \right) \psi_{1,-1}(\mathbf{r}, t) + \frac{\sqrt{2}}{6} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \psi_{0,0}(\mathbf{r}, t) \\
= \mu_s(\mathbf{r}) \Theta_1 \psi_{1,-1}(\mathbf{r}, t) + q_{1,-1}(\mathbf{r}, t)
\end{aligned} \tag{3.14}$$

$$\begin{aligned}
\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_{\text{tr}}(\mathbf{r}) \right) \psi_{1,1}(\mathbf{r}, t) - \frac{\sqrt{2}}{6} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) \psi_{0,0}(\mathbf{r}, t) \\
= \mu_s(\mathbf{r}) \Theta_1 \psi_{1,1}(\mathbf{r}, t) + q_{1,1}(\mathbf{r}, t).
\end{aligned} \tag{3.15}$$

We make use of the relations (3.8) and (3.9) to arrive at

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_{\text{tr}}(\mathbf{r}) \right) \Phi(\mathbf{r}, t) + \nabla \cdot \mathbf{J}(\mathbf{r}, t) = \mu_s(\mathbf{r}) \Phi(\mathbf{r}, t) + \Theta_0 q_{0,0}(\mathbf{r}, t) \tag{3.16}$$

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_{\text{tr}}(\mathbf{r}) \right) \mathbf{J}(\mathbf{r}, t) + \frac{1}{3} \nabla \Phi(\mathbf{r}, t) = \Theta_1 \mu_s(\mathbf{r}) \mathbf{J}(\mathbf{r}, t) + \mathbf{q}_1 \tag{3.17}$$

where

$$\mathbf{q}_1(\mathbf{r}, t) = \int_{S^{n-1}} \hat{\mathbf{s}} q(\mathbf{r}, \hat{\mathbf{s}}, t) d\hat{\mathbf{s}} = \begin{pmatrix} \frac{1}{\sqrt{2}} (q_{1,-1}(\mathbf{r}, t) - q_{1,1}(\mathbf{r}, t)) \\ \frac{1}{i\sqrt{2}} (q_{1,-1}(\mathbf{r}, t) + q_{1,1}(\mathbf{r}, t)) \\ q_{1,0}(\mathbf{r}, t) \end{pmatrix}. \tag{3.18}$$

If we now make the definitions

$$\text{isotropic source:} \quad q_0 = q_{0,0}, \tag{3.19}$$

$$\text{transport coefficient:} \quad \mu_{\text{tr}} = \mu_a + \mu_s, \tag{3.20}$$

$$\text{reduced scattering coefficient:} \quad \mu'_s = (1 - \Theta_1) \mu_s, \tag{3.21}$$

$$\text{diffusion coefficient:} \quad \kappa = \frac{1}{3(\mu_a + \mu'_s)}, \tag{3.22}$$

and using the fact that $\Theta_0 = 1$ we get the more common form

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_a(\mathbf{r}) \right) \Phi(\mathbf{r}, t) + \nabla \cdot \mathbf{J}(\mathbf{r}, t) = q_0(\mathbf{r}, t) \tag{3.23}$$

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \frac{1}{3\kappa(\mathbf{r})} \right) \mathbf{J}(\mathbf{r}, t) + \frac{1}{3} \nabla \Phi(\mathbf{r}, t) = \mathbf{q}_1. \tag{3.24}$$

3.3. Diffusion approximation

The diffusion approximation results from making the assumptions:

$$\frac{\partial \mathbf{J}}{\partial t} = 0 \quad \mathbf{q}_1 = 0 \quad (3.25)$$

from which (3.24) gives

$$\mathbf{J}(\mathbf{r}, t) = -\kappa(\mathbf{r}) \nabla \Phi(\mathbf{r}, t) \quad (3.26)$$

leading to

$$-\nabla \cdot \kappa(\mathbf{r}) \nabla \Phi(\mathbf{r}, t) + \mu_a \Phi(\mathbf{r}, t) + \frac{1}{c} \frac{\partial \Phi(\mathbf{r}, t)}{\partial t} = q_0(\mathbf{r}, t) \quad (3.27)$$

with the frequency-domain form

$$-\nabla \cdot \kappa(\mathbf{r}) \nabla \Phi(\mathbf{r}, \omega) + \mu_a \Phi(\mathbf{r}, \omega) + \frac{i\omega}{c} \Phi(\mathbf{r}, \omega) = q_0(\mathbf{r}, \omega). \quad (3.28)$$

The assumption that $\dot{\mathbf{J}} = 0$ is clearly erroneous in the time-dependent case, but is usually justified by specifying the condition $\mu_a \ll \mu'_s$. Another possibility is to assume that \mathbf{J} is dominated by an exponentially decaying term with time constant $c\lambda$

$$\mathbf{J}(\mathbf{r}, t) = \mathbf{e}^{-c\lambda t}. \quad (3.29)$$

Then (3.26) becomes

$$(1 - 3\kappa\lambda)\mathbf{J}(\mathbf{r}, t) = -\kappa \nabla \Phi(\mathbf{r}, t) \quad (3.30)$$

leading to

$$\check{\kappa} = \frac{\kappa}{1 - 3\lambda\kappa} = \frac{1}{3(\mu_a + \mu'_s - \lambda)} \quad (3.31)$$

in place of (3.22). Yamada [21] has argued that $\lambda = \mu_a$ which is consistent with the long time limit observation of an exponentially decaying solution and leads to

$$\check{\kappa} = \frac{1}{3\mu'_s}. \quad (3.32)$$

3.4. Telegraph equation

The telegraph equation arises if we attempt to eliminate $\nabla \cdot \mathbf{J}$ from (3.23) and (3.24). Taking the divergence of the second equation we obtain

$$\frac{1}{c} \frac{\partial \nabla \cdot \mathbf{J}}{\partial t} + \nabla \cdot \left(\frac{\mathbf{J}}{3\kappa} \right) + \frac{1}{3} \nabla^2 \Phi = \nabla \cdot \mathbf{q}_1$$

which leads to

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \frac{1}{3\kappa} \right) \nabla \cdot \mathbf{J} - \frac{\mathbf{J} \cdot \nabla \kappa}{3\kappa^2} + \frac{1}{3} \nabla^2 \Phi = \nabla \cdot \mathbf{q}_1. \quad (3.33)$$

Substituting $\nabla \cdot \mathbf{J}$ from (3.23) leads to

$$\left(\frac{3\kappa}{c} \frac{\partial}{\partial t} + 1 \right) \left(q_0 - \mu_a \Phi - \frac{1}{c} \frac{\partial \Phi}{\partial t} \right) - \frac{\mathbf{J} \cdot \nabla \kappa}{\kappa} + \kappa \nabla^2 \Phi = 3\kappa \nabla \cdot \mathbf{q}_1. \quad (3.34)$$

In the case $\kappa = \text{constant}$, the term $\mathbf{J} \cdot \nabla \kappa = 0$ and we have the homogeneous telegraph equation:

$$\frac{3\kappa}{c^2} \ddot{\Phi} + \frac{1}{c} (3\kappa\mu_a + 1) \dot{\Phi} + \mu_a \Phi - \kappa \nabla^2 \Phi = S_1 \quad (3.35)$$

where

$$S_1 = q_0 + \frac{3\kappa}{c}\dot{q}_0 - 3\kappa\nabla \cdot \mathbf{q}_1. \quad (3.36)$$

Alternatively, an inhomogeneous equation can be obtained if we use (3.26):

$$\frac{3\kappa}{c^2}\ddot{\Phi} + \frac{1}{c}(3\kappa\mu_a + 1)\dot{\Phi} + \mu_a\Phi - \nabla \cdot \kappa\nabla\Phi = S_1. \quad (3.37)$$

3.5. Boundary conditions and source terms

3.5.1. Boundary conditions. The boundary condition in the transport equation specifies that no photons travel in an inward direction at the boundary, except for source terms

$$\phi(\mathbf{m}, \hat{\mathbf{s}}, t) = 0 \quad \text{for } \hat{\mathbf{s}} \cdot \hat{\nu} < 0 \quad (3.38)$$

where $\hat{\nu}$ is the outer normal to $\partial\Omega$ at \mathbf{m} . The diffusion equation cannot satisfy this condition exactly. Instead we assume that the *total* inward directed current is zero

$$\int_{\hat{\mathbf{s}} \cdot \hat{\nu} < 0} \hat{\mathbf{s}}\phi(\mathbf{m}, \hat{\mathbf{s}}, t) d\hat{\mathbf{s}} = 0 \quad (3.39)$$

which leads to the Robin condition [22]

$$\Phi(\mathbf{m}) + 2\kappa(\mathbf{m})\frac{\partial\Phi(\mathbf{m})}{\partial\nu} = 0. \quad (3.40)$$

To incorporate diffuse boundary reflection arising from a refractive index mismatch between Ω and the surrounding medium, (3.39) is modified to

$$\int_{\hat{\mathbf{s}} \cdot \hat{\nu} < 0} \hat{\mathbf{s}}\phi(\mathbf{m}, \hat{\mathbf{s}}, t) d\hat{\mathbf{s}} = \int_{\hat{\mathbf{s}} \cdot \hat{\nu} > 0} \hat{\mathbf{s}}R(\hat{\mathbf{s}})\phi(\mathbf{m}, \hat{\mathbf{s}}, t) d^2\hat{\mathbf{s}} \quad (3.41)$$

where $R(\hat{\mathbf{s}})$ is a directionally varying refraction parameter. In the diffusion approximation this becomes [23]

$$\Phi(\mathbf{m}) + 2\kappa(\mathbf{m})\frac{\partial\Phi(\mathbf{m})}{\partial\nu} = R \left[\Phi(\mathbf{m}) - 2\kappa(\mathbf{m})\frac{\partial\Phi(\mathbf{m})}{\partial\nu} \right] \quad (3.42)$$

which we rearrange as

$$\Phi(\mathbf{m}) + 2A\kappa(\mathbf{m})\frac{\partial\Phi(\mathbf{m})}{\partial\nu} = 0 \quad (3.43)$$

where $A = (1 + R)/(1 - R)$. Discussion of the precise value of A to use has been the subject of several studies [24, 25] and is often justified by comparison with experimental or Monte Carlo data [26, 27].

3.5.2. Source conditions. A collimated source incident at $\mathbf{p} \in \partial\Omega$ is commonly represented by a diffuse point source

$$q_0(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_s) \quad (3.44)$$

where \mathbf{r}_s is located at a depth of one scattering length below the surface. This arises as the mean survival depth of an exponentially decaying line source, where photons travelling in a direction $-\hat{\nu}$ travel a distance z with survival probability $P(z) = e^{-\mu_{tr}z}$ (see [22] section 9.3).

Analytic solutions for the exponentially decaying line source, including a spatial distribution over the surface, were developed by Eason *et al* [28], and were included in a finite element model by Paulsen *et al* [29]. Schweiger *et al* [27] investigated several source models in the finite element method, including the specification of the source as a Dirichlet or Neumann boundary condition.

An anisotropic source may also be included by putting, in place of (3.26)

$$\mathbf{J}(\mathbf{r}, t) = 3\kappa(\mathbf{r})\mathbf{q}_1(\mathbf{r}, t) - \kappa(\mathbf{r})\nabla\Phi(\mathbf{r}, t) \quad (3.45)$$

leading to

$$\tilde{q}_0(\mathbf{r}, t) = q_0(\mathbf{r}, t) - 3\nabla \cdot \kappa(\mathbf{r})\mathbf{q}_1(\mathbf{r}, t). \quad (3.46)$$

Although studies reveal a degree of sensitivity to the source model, the majority of authors use the form (3.44). In the inverse problem the niceties of this term are often removed by using either relative intensity data or normalized measurements (see section 4.2).

3.6. Translation to Helmholtz form

Equation (3.28) can be converted to Helmholtz form. We make the change of variables $U = \kappa^{\frac{1}{2}}\Phi$, by which (3.28) becomes

$$-\kappa\nabla^2\Phi(\omega) - 2\kappa^{\frac{1}{2}}\nabla\kappa^{\frac{1}{2}} \cdot \nabla\Phi(\omega) + \mu_a\Phi(\omega) + \frac{i\omega}{c}\Phi(\omega) = q_0(\omega). \quad (3.47)$$

Using

$$\nabla^2U = \kappa^{\frac{1}{2}}\nabla^2\Phi + 2\nabla\Phi \cdot \nabla\kappa^{\frac{1}{2}} + \Phi\nabla^2\kappa^{\frac{1}{2}}$$

leads to

$$-\nabla^2U(\omega) + \hat{\eta}(\omega)U(\omega) = \frac{q_0(\omega)}{\kappa^{\frac{1}{2}}} \quad (3.48)$$

where $\hat{\eta}(\omega) = \eta_0 + i\omega\xi$ with

$$\eta_0 = \left(\frac{\nabla^2\kappa^{\frac{1}{2}}}{\kappa^{\frac{1}{2}}} \right) + \frac{\mu_a}{\kappa} \quad \xi = \frac{1}{c\kappa}. \quad (3.49)$$

The Helmholtz form allows the problem to be related to inverse scattering problems in diffraction tomography, and to make statements about uniqueness, which we return to in section 8.

3.7. P_N approximation to higher order

A frequently asked question is as to why the diffusion approximation is considered adequate. In the time domain it is particularly unsatisfactory that the Green function of the diffusion approximation is causal (in opposition to the fundamental time-reversibility of light propagation) and in violation of relativity (sources of photons give rise to photon densities instantaneously). Some insight into this question is given by considering cases wherein the higher-order solutions are known.

The transport equation in spherical polar coordinates, assuming spherical symmetry is [24, 30]

$$\begin{aligned} & \left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_{tr}(r) \right) \phi(r, \tau_r, t) + \tau_r \frac{\partial \phi(r, \tau_r, t)}{\partial r} + \frac{1 - \tau_r^2}{r} \frac{\partial \phi(r, \tau_r, t)}{\partial \tau_r} \\ & = \mu_s(r) \int_{S^{n-1}} \Theta(\hat{\mathbf{s}} \cdot \hat{\mathbf{s}}') \phi(r, \tau_r', t) d\hat{\mathbf{s}}' + q(r, \tau_r, t) \end{aligned} \quad (3.50)$$

where $\tau_r = \hat{\mathbf{s}} \cdot \hat{\mathbf{r}}$ is the cosine of the angle between the radial position vector and the velocity direction. Then in place of (3.5) we have

$$\phi(r, \tau_r, t) = \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} \psi_l(r, t) P_l(\tau_r) \quad (3.51)$$

which leads, in place of (3.11), to the coupled system

$$(2l+1) \left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_a(r) + (1 - \Theta_l) \mu_s(r) \right) \psi_l(r, t) + (l+1) \left(\frac{\partial}{\partial r} + \frac{n+2}{r} \right) \psi_{l+1}(r, t) \\ + l \left(\frac{\partial}{\partial r} - \frac{n-1}{r} \right) \psi_{l-1}(r, t) = q_l(r, t). \quad (3.52)$$

Converting to the frequency domain, the eigenfunctions of this system lead to the general solution [30]

$$\psi_l(r, \omega) = \sum_j A_j H_l(\lambda_j) \lambda_j k_l(\lambda_j r) \quad (3.53)$$

where $k_l(z) = (\pi/(2z))^{1/2} K_{l+1/2}(z)$ is the modified spherical Bessel function of the second kind. The eigenvalues $\lambda_j(\omega)$ are the roots of the secular equation

$$\begin{vmatrix} \epsilon_0(\omega) & \lambda & 0 & 0 & 0 & \dots & \dots & 0 \\ \lambda & 3\epsilon_1(\omega) & 2\lambda & 0 & 0 & \dots & \dots & 0 \\ 0 & 2\lambda & 5\epsilon_2(\omega) & 3\lambda & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \dots & \vdots \\ 0 & \dots & \dots & \dots & \dots & (N-1)\lambda & (2N-1)\epsilon_{N-1}(\omega) & N\lambda \\ 0 & \dots & \dots & \dots & \dots & 0 & N\lambda & (2N+1)\epsilon_N(\omega) \end{vmatrix} \\ = 0 \quad (3.54)$$

where

$$\epsilon_l(\omega) = \mu_a + (1 - \Theta_l) \mu_s + \frac{i\omega}{c}. \quad (3.55)$$

The eigenfunctions $H_l(\lambda)$ depend on the phase function Θ and satisfy a recurrence relation (see [30], section 17.5.2)

$$(l+1)H_{l+1}(\lambda) + \frac{2l+1}{\lambda} \epsilon_l(\omega) H_l(\lambda) + lH_{l-1}(\lambda) = 0 \quad (3.56)$$

from which

$$H_0(\lambda) = 1 \\ H_1(\lambda) = -\frac{\epsilon_1(\omega)}{\lambda}. \quad (3.57)$$

Note that in the isotropic case the H_l functions are linear combinations of Legendre polynomials of the first and second kind.

In particular, using (3.57), and $k_0(z) = \exp(-z)/z$ leads to

$$\Phi(r, \omega) = \psi_0(r, \omega) = \sum_j A_j \frac{e^{-\lambda_j(\omega)r}}{r} \quad (3.58)$$

$$\mathbf{J}(r, \omega) = \hat{\mathbf{s}} \psi_1(r, \omega) = \sum_j A_j \frac{-\epsilon_1(\omega)}{\lambda_j(\omega)} \left(\frac{2\lambda_j(\omega)}{\pi r} \right)^{1/2} K_{3/2}(\lambda_j(\omega)r). \quad (3.59)$$

Equation (3.58) agrees with [31]. Calculation of $\lambda_j(\omega)$ requires the values of Θ_l ; these are known for both Mie scattering kernel [32], and the Henyey–Greenstein phase function [33]

$$\Theta(\hat{\mathbf{s}} \cdot \hat{\mathbf{s}}') = \Theta(\tau_s) = \frac{1 - \bar{\Theta}^2}{2(1 + \bar{\Theta}^2 - 2\bar{\Theta}\tau_s)^{3/2}}. \quad (3.60)$$

For illustration we follow [31] and use (3.60) for which $\Theta_l = \bar{\Theta}^l$. The distribution of $\lambda_j(\omega)$ are shown for different orders of the P_N approximation, for typical $(\mu_a, \mu_s, \Theta, \omega)$ parameters

in figure 3. The first important characteristic is that successive eigenvalues are much larger than the lowest so that these terms decay extremely quickly with distance from the source. Of more importance is that the lowest eigenvalue in each successive approximation is very similar. Boas [31] argues that P_1 is adequate whenever

$$\left| \frac{108 \epsilon_0 \epsilon_1}{35 \epsilon_2 \epsilon_3} \right| \ll 1 \quad (3.61)$$

which is the case for most applications in optical tomography. Note that the assumption here is that the powers of ω^n for $n > 1$ are ignored. Fishkin *et al* [34] give an argument that these higher powers of frequency, which are the imaginary parts of the roots of the secular equation, are negligible for typical optical parameters.

In the time domain, the Fourier transform of (3.58) for the lowest order eigenvalue gives the Green function of (3.35) in infinite space. This was given by [35] as

$$g_{\text{Tel}}^{(\Phi)}(\mathbf{r}, \mathbf{r}_0, t) = \begin{cases} \frac{\sqrt{3}\kappa_-}{4\pi c\kappa} \frac{e^{-\kappa_+ ct}}{(t^2 - t_r^2)^{\frac{1}{2}}} I_1(\kappa_- c (t^2 - t_r^2)^{\frac{1}{2}}) & t > t_r \\ 0 & t \leq t_r \end{cases} \quad (3.62)$$

where I_1 is a modified Bessel function of the first kind and

$$\kappa_{\pm} = \frac{1 \pm 3\kappa\mu_a}{6\kappa} \quad t_r = \frac{\sqrt{3}(r - r_0)}{c}. \quad (3.63)$$

The observation that the imaginary parts of the higher eigenfunctions decay rapidly is consistent with saying that the higher-order derivatives in the time domain are transient. In figure 4 we plot the discrepancy between the telegraph and diffusion Green functions as a fraction of the peak intensity. The difference is negligible apart from a small time interval around the delayed time $t_d = r/c$ where the first photons arrive.

3.8. The Boltzmann hierarchy

An alternative to the P_N approach was developed by Kaltenbach and Kaschke [35]. Here the moments of ϕ in (3.1) are defined:

$$\phi^{(n)}(\mathbf{r}, t) = \int_{S^{n-1}} \hat{\mathbf{s}}^{\otimes n} \phi(\mathbf{r}, \hat{\mathbf{s}}, t) d\hat{\mathbf{s}} \quad (3.64)$$

where $\hat{\mathbf{s}}^{\otimes n}$ is the n th outer product of unit vector $\hat{\mathbf{s}}$ with itself. Clearly

$$\begin{aligned} \hat{\mathbf{s}}^{\otimes 0} = 1 &\Rightarrow \phi^{(0)} = \Phi \\ \hat{\mathbf{s}}^{\otimes 1} = \hat{\mathbf{s}} &\Rightarrow \phi^{(1)} = \mathbf{J} \end{aligned}$$

with $q^{(n)}$ defined similarly to give $q^{(0)} = q_0$ and $q^{(1)} = \mathbf{q}_1$. The second order tensor $\hat{\mathbf{s}}^{\otimes 2}$ can be written as the sum of a diagonal and a zero-trace tensor

$$\hat{\mathbf{s}}^{\otimes 2} = \hat{\mathbf{s}} \otimes \hat{\mathbf{s}} = \hat{\mathbf{s}}_0^{\otimes 2} + \frac{1}{3}\delta^{\otimes 2} \quad (3.65)$$

where

$$\begin{aligned} \delta^{\otimes 2} &= \sqrt{4\pi} \begin{pmatrix} Y_{0,0} & 0 & 0 \\ 0 & Y_{0,0} & 0 \\ 0 & 0 & Y_{0,0} \end{pmatrix} \\ \hat{\mathbf{s}}_0^{\otimes 2} &= \sqrt{\frac{4\pi}{5}} \begin{pmatrix} \sqrt{\frac{1}{6}}(Y_{2,-2} + Y_{2,2}) - \frac{1}{3}Y_{2,0} & i\sqrt{\frac{1}{6}}(Y_{2,-2} - Y_{2,2}) & \sqrt{\frac{1}{6}}(Y_{2,-1} - Y_{2,1}) \\ -i\sqrt{\frac{1}{6}}(-Y_{2,-2} + Y_{2,2}) & -\sqrt{\frac{1}{6}}(Y_{2,-2} + Y_{2,2}) - \frac{1}{3}Y_{2,0} & i\sqrt{\frac{1}{6}}(Y_{2,-1} + Y_{2,1}) \\ \sqrt{\frac{1}{6}}(Y_{2,-1} - Y_{2,1}) & i\sqrt{\frac{1}{6}}(Y_{2,-1} + Y_{2,1}) & \frac{2}{3}Y_{2,0} \end{pmatrix}. \end{aligned} \quad (3.66)$$

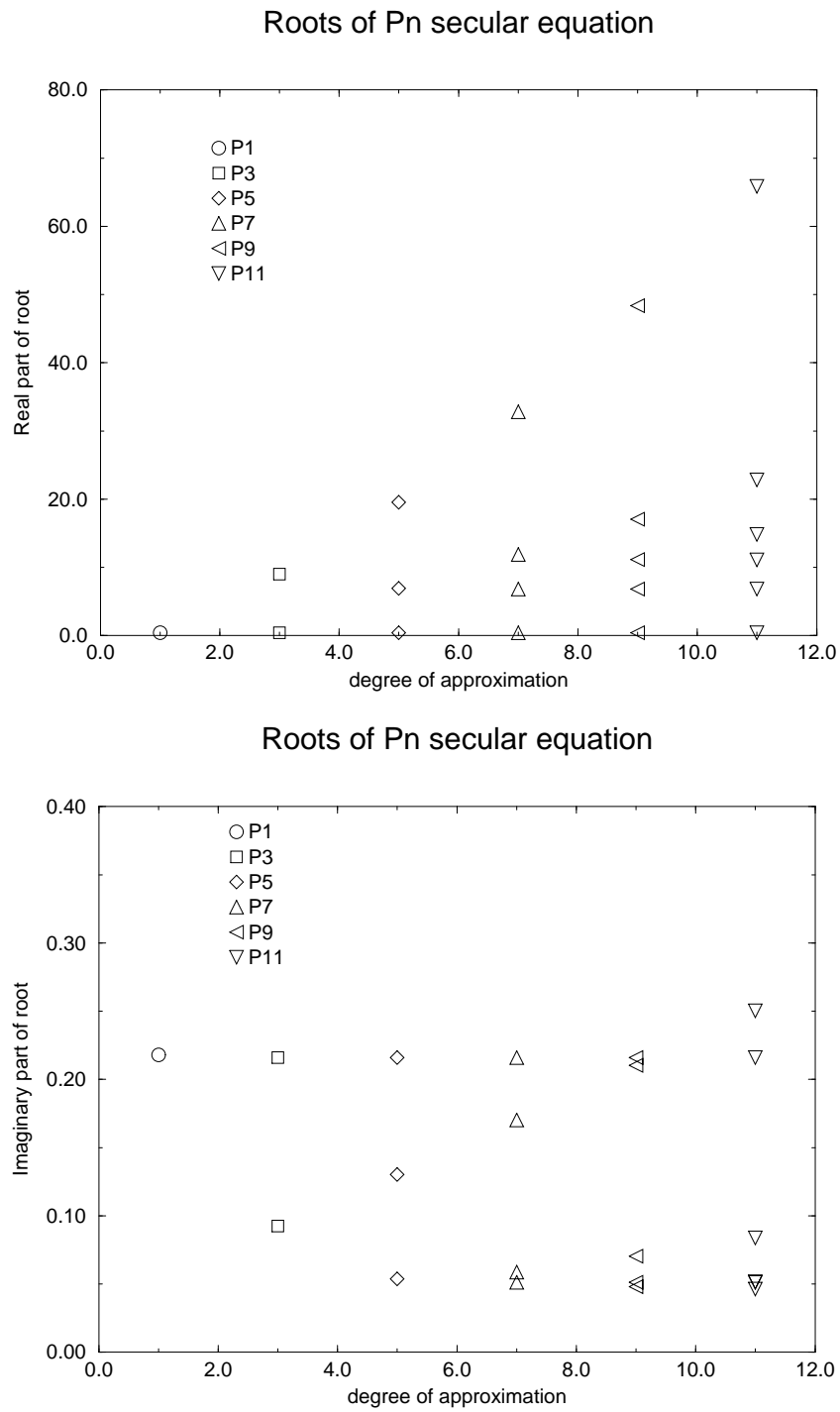


Figure 3. Real and imaginary parts of the roots of the secular equation (3.54) for the case $\mu_a = 0.025 \text{ mm}^{-1}$, $\mu_s = 20 \text{ mm}^{-1}$, $\bar{\Theta} = 0.9$, $\omega = 1 \text{ GHz}$.

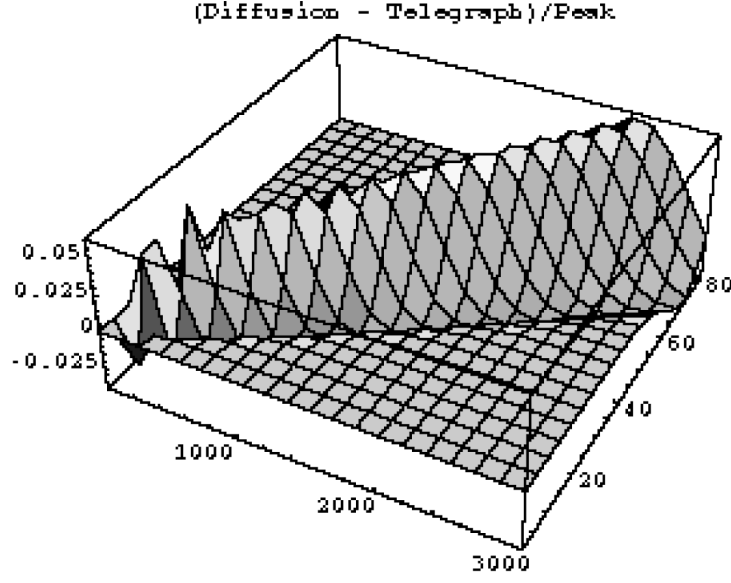


Figure 4. $(g_{\text{Diff}}^{(\Phi)}(r, t) - g_{\text{Tel}}^{(\Phi)}(r, t))/\text{Max}[g_{\text{Diff}}^{(\Phi)}(r, t)]$ plotted against distance and time for the case $\mu_a = 0.025 \text{ mm}^{-1}$, $\mu'_s = 2 \text{ mm}^{-1}$.

We have

$$\text{Tr } \hat{s}^{\otimes 2} = (4\pi)^{\frac{1}{2}} Y_{0,0} = 1 \quad \text{Tr } \phi^{(2)} = \phi^{(0)} \quad (3.67)$$

so we may define another zero-trace tensor

$$\phi_0^{(2)} = \phi^{(2)} - \frac{1}{3} \phi \delta^{\otimes 2}. \quad (3.68)$$

The integral over the phase function in (3.64) couples all spherical harmonics of the same order to the same coefficient, and since $\hat{s}_0^{\otimes 2}$ contains spherical harmonics of order two only, we can separate terms

$$\begin{aligned} \int_{S^{n-1}} \hat{s}^{\otimes 2} \int_{S^{n-1}} \Theta(\hat{s} \cdot \hat{s}') \phi(r, \hat{s}', t) d\hat{s}' d\hat{s} &= \int_{S^{n-1}} \hat{s}_0^{\otimes 2} \int_{S^{n-1}} \Theta(\hat{s} \cdot \hat{s}') \phi(r, \hat{s}', t) d\hat{s}' d\hat{s} \\ &+ \frac{1}{3} \int_{S^{n-1}} \delta^{\otimes 2} \int_{S^{n-1}} \Theta(\hat{s} \cdot \hat{s}') \phi(r, \hat{s}', t) d\hat{s}' d\hat{s} \\ &= \Theta_2 \phi_0^{(2)}(r, t) + \frac{1}{3} \Theta_0 \phi^{(0)}(r, t) \delta^{\otimes 2} \\ &= \Theta_2 \phi^{(2)}(r, t) + \frac{1}{3} (\Theta_0 - \Theta_2) \phi^{(0)}(r, t) \delta^{\otimes 2}. \end{aligned} \quad (3.69)$$

Note that all tensors are symmetric with respect to any permutation of indices and that we have

$$\nabla \cdot a \delta^{\otimes 2} = \frac{\partial}{\partial x_k} a_{i_1 i_2 \dots i_n} \delta_{i_j k} = \frac{\partial}{\partial x_k} a_{i_1 i_2 \dots k \dots i_n} = \nabla a.$$

This leads to the first three equations in the Boltzmann hierarchy in the form given by Kaltenbach and Kaschke [35]

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_a \right) \phi^{(0)} + \nabla \cdot \phi^{(1)} = q^{(0)} \quad (3.70)$$

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_a + (1 - \Theta_1) \mu_s \right) \phi^{(1)} + \frac{1}{3} \nabla \phi^{(0)} + \nabla \cdot \phi_0^{(2)} = q^{(1)} \quad (3.71)$$

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_a + (1 - \Theta_2)\mu_s\right) \phi^{(2)} + \frac{1}{3} (\Theta_2 - \Theta_0) \mu_s \phi^{(0)} \delta^{\otimes 2} + \nabla \cdot \phi^{(3)} = q^{(2)}. \quad (3.72)$$

The Boltzmann hierarchy is different from P_N in that the n th-order term contains terms in derivatives of order $n + 1$. Furthermore, higher-order terms decouple in more complex way due to mixing of lower-order harmonics, with the general term of the form

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_a + (1 - \Theta_n)\mu_s\right) \phi^{(n)} + \sum_{m=1}^{\lfloor n/2 \rfloor} (\Theta_n - \Theta_{n-2m}) \mu_s \underline{\underline{B}}^{(n-2m)}(\phi^{(n-2m)}, \phi^{(n-2m-2)}, \dots) \otimes \delta^{\otimes 2m} + \nabla \cdot \phi^{(n+1)} = q^{(n)} \quad (3.73)$$

where $\underline{\underline{B}}^{(n)}$ is an n -rank tensor. Although an elegant representation this scheme has not been developed beyond the formal results of [35].

3.9. Limitations of the diffusion approximation

Apart from $\mu_a \ll \mu'_s$ there are two main problems in using the diffusion approximation, namely clear regions and jumps. If Ω consists of piecewise continuous regions then the boundary conditions of the transport equation cannot be exactly met at the interfaces [30, 36]. In these circumstances higher-order approximations need to be employed. Hielscher *et al* [37] have used a numerical method for the transport equation to compare the forward solution with a diffusion model in a piecewise continuous model of the brain.

A different, more serious problem is non-scattering regions. Some results for this case are given by Davison [30], section 8.6. Most numerical solutions to the transport equation, apart from Monte Carlo, cannot deal with this case, so special methods need to be developed. Firbank *et al* [38] developed a method for solving the diffusion equation iteratively with radiative propagation in gaps. Comparison with Monte Carlo methods and experimental phantom results show that this is a stable approach [39].

4. The forward problem

4.1. Diffusion model

We combine (3.28) and (3.43) to give

$$-\nabla \cdot \kappa(\mathbf{r}) \nabla \Phi(\mathbf{r}, \omega) + \mu_a(\mathbf{r}) \Phi(\mathbf{r}, \omega) + \frac{i\omega}{c} \Phi(\mathbf{r}, \omega) = q_0(\mathbf{r}, \omega), \quad (4.1a)$$

$$\Phi(\mathbf{m}, \omega) + 2A\kappa(\mathbf{m}) \frac{\partial \Phi(\mathbf{m}, \omega)}{\partial \nu} = 0, \quad (4.1b)$$

with the time-domain form

$$-\nabla \cdot \kappa(\mathbf{r}) \nabla \Phi(\mathbf{r}, t) + \mu_a(\mathbf{r}) \Phi(\mathbf{r}, t) + \frac{1}{c} \frac{\partial \Phi(\mathbf{r}, t)}{\partial t} = q_0(\mathbf{r}, t), \quad (4.2a)$$

$$\Phi(\mathbf{m}, t) + 2A\kappa(\mathbf{m}) \frac{\partial \Phi(\mathbf{m}, t)}{\partial \nu} = 0, \quad (4.2b)$$

Equations (4.1) and (4.2) can be formally solved by applying the appropriate Green operators

$$\hat{\mathcal{G}}(\mathbf{r}, \omega)[q_0] = \int_{\Omega} G^{(\Phi)}(\mathbf{r}, \mathbf{r}', \omega) q_0(\mathbf{r}', \omega) d^n \mathbf{r} \quad (4.3)$$

$$\mathcal{G}(\mathbf{r}, t)[q_0] = \int_{-\infty}^{\infty} \int_{\Omega} g^{(\Phi)}(\mathbf{r}, \mathbf{r}', t - t') q_0(\mathbf{r}', t') d^n \mathbf{r} dt'. \quad (4.4)$$

4.2. Measurement types

The boundary quantity that is measurable (the exitance) can be defined

$$\Gamma(\mathbf{m}) = \mathcal{B}[\Phi] = -\kappa(\mathbf{m}) \frac{\partial \Phi(\mathbf{m})}{\partial \nu} = \hat{\nu} \cdot \mathbf{J}(\mathbf{m}). \quad (4.5)$$

Equation (4.5) combined with the Robin boundary condition (4.1b) or (4.2b) gives the simpler form

$$\Gamma(\mathbf{m}) = \zeta \Phi(\mathbf{m}) \quad (4.6)$$

where $\zeta = 1/2A$. Both $\Gamma(\omega)$ and $\Gamma(t)$ have been discussed as the measurement type appropriate for reconstruction. We may also define

$$\text{DC intensity:} \quad \mathcal{M}_E[\Gamma] = \int_{-\infty}^{\infty} \Gamma(t) dt = \Gamma(\omega = 0) \triangleq E, \quad (4.7)$$

$$\text{AC intensity:} \quad \mathcal{M}_{E(\omega)}[\Gamma] = |\Gamma(\omega)|, \quad (4.8)$$

$$\text{log intensity:} \quad \mathcal{M}_{\log E(\omega)}[\Gamma] = \log|\Gamma|, \quad (4.9)$$

$$\text{modulation depth:} \quad \mathcal{M}_A[\Gamma] = \frac{|\Gamma(\omega)|}{\Gamma(\omega = 0)} = \frac{|\Gamma(\omega)|}{E}, \quad (4.10)$$

$$\text{phase:} \quad \mathcal{M}_\theta[\Gamma] = \arg \Gamma(\omega). \quad (4.11)$$

Note that we formally consider the argument to the operator to be $\Gamma(t)$ since we can obtain $\Gamma(\omega)$ by Fourier transform, e.g.

$$\mathcal{M}_A[\Gamma(\omega)] = \mathcal{M}_A \mathcal{F}[\Gamma(t)].$$

Consider a general spatial-temporal filter operator acting on the measurable Γ

$$\begin{aligned} \mathcal{A}_h[\Gamma] &\triangleq \int_{-\infty}^{\infty} \int_{\partial\Omega} h(\mathbf{r}, t) \Gamma(\mathbf{r}, t) d^{n-1} \hat{\nu}_\perp dt \\ &= \int_{-\infty}^{\infty} \int_{\partial\Omega} h(\mathbf{r}, t) \kappa(\mathbf{r}) \frac{\partial \Phi(\mathbf{r}, t)}{\partial \nu} d^{n-1} \hat{\nu}_\perp dt = \mathcal{A}_h \mathcal{B}[\Phi] = \mathcal{B} \mathcal{A}_h[\Phi]. \end{aligned} \quad (4.12)$$

If $h(\mathbf{r}, t)$ is separable $h(\mathbf{r}, t) = p(\mathbf{r})w(t)$ then we can define a *spatial* operator with general form

$$\mathcal{S}_p[\Gamma] = \int_{\partial\Omega} p(\mathbf{r}) \Gamma(\mathbf{r}, t) d^{n-1} \hat{\nu}_\perp \quad (4.13)$$

and a *temporal* operator \mathcal{T}_w with general form

$$\mathcal{T}_w[\Gamma] = \int_{-\infty}^{\infty} w(t) \Gamma(\mathbf{r}, t) dt \quad (4.14)$$

with $\mathcal{A} = \mathcal{T}\mathcal{S} = \mathcal{S}\mathcal{T}$.

Formally we define the DC measurement operator as the temporal integral with weight unity

$$\mathcal{E} \triangleq \mathcal{T}_1 = \mathcal{M}_E. \quad (4.15)$$

A much discussed data type is then the time-gated intensity, achieved when $w(t)$ is a top-hat function

$$W(T_a, T_b) = \begin{cases} 1 & \text{for } T_a \leq t \leq T_b, \\ 0 & \text{otherwise} \end{cases} \quad (4.16)$$

which leads to two data types, according as whether T_a is set to zero

$$\text{early light: } \mathcal{M}_{E(T)}[\Gamma] = \int_0^T \Gamma(t) dt, \quad (4.17)$$

$$\text{windowed light: } \mathcal{M}_{E(T_a, T_b)}[\Gamma] = \int_{T_a}^{T_b} \Gamma(t) dt. \quad (4.18)$$

4.2.1. *Measurements of integral transform type.* For temporal filters of the form

$$w(t) = At^n e^{-\varsigma t}$$

the measurement can be derived directly (i.e. without explicitly finding $\Gamma(t)$). This filter contains in the general form four special cases

$$\text{Fourier transform: } n = 0, \quad \varsigma \text{ imaginary} \quad (4.19)$$

$$\text{Mellin transform: } n \text{ integer, } \varsigma = 0 \quad (4.20)$$

$$\text{Laplace transform: } n = 0, \quad \varsigma \text{ real} \quad (4.21)$$

$$\text{Mellin-Laplace transform: } n \text{ integer, } \varsigma \text{ real.} \quad (4.22)$$

For types (4.21) and (4.22) the filter is integrable and can be normalized

$$\text{Laplace transform: } w(t) = se^{-st} \quad (4.23)$$

$$\text{Mellin-Laplace transform: } w(t) = \frac{s^{n+1}}{B(n+1)} t^n e^{-st} \quad (4.24)$$

where $B(n+1)$ is a Gamma function. To derive the direct results, we make use of the relation

$$\check{u}^{(n)}(s) \triangleq \int_{-\infty}^{\infty} t^n e^{-st} u(t) dt = i^n \frac{\partial^n}{\partial \omega^n} \int_{-\infty}^{\infty} e^{-i(\omega - is)t} u(t) dt \Big|_{\omega=0}. \quad (4.25)$$

Then consider

$$\left\{ -\nabla \cdot \kappa(\mathbf{r}) \nabla + \left(\mu_a(\mathbf{r}) + \frac{s}{c} + \frac{i\omega}{c} \right) \right\} \Phi(\mathbf{r}, \omega - is) = q_0(\mathbf{r}, \omega) \quad (4.26)$$

which has the solution

$$\Phi(\mathbf{r}, \omega - is) = \hat{\mathcal{G}}(\omega - is)[q_0(\mathbf{r}, \omega - is)]. \quad (4.27)$$

Differentiating both sides of (4.26) gives

$$\left\{ \left[-\nabla \cdot \kappa(\mathbf{r}) \nabla + \left(\mu_a(\mathbf{r}) + \frac{s}{c} + \frac{i\omega}{c} \right) \right] i \frac{\partial}{\partial \omega} - \frac{1}{c} \right\} \Phi(\mathbf{r}, \omega - is) = i \frac{\partial}{\partial \omega} q_0(\mathbf{r}, \omega) \quad (4.28)$$

which is solved by

$$i \frac{\partial \Phi(\mathbf{r}, \omega - is)}{\partial \omega} = \hat{\mathcal{G}}(\omega - is) \left[\frac{1}{c} \Phi(\mathbf{r}, \omega - is) + i \frac{\partial q_0(\mathbf{r}, \omega - is)}{\partial \omega} \right]. \quad (4.29)$$

Iterating this procedure and evaluating at $\omega = 0$ leads to

$$\begin{aligned} \check{\Phi}^{(0)}(s) &= \check{\mathcal{G}}(s)[q_0] \\ \check{\Phi}^{(n)}(s) &= \check{\mathcal{G}}(s) \left[\frac{n}{c} \check{\Phi}^{(n-1)}(s) + q_0^{(n)} \right] \end{aligned} \quad (4.30)$$

where $\check{\mathcal{G}}(s) = \hat{\mathcal{G}}(-is)$ is the Laplace-domain Green operator. In the case where $q_0(t)$ is a δ -function in time, we have

$$\check{\Phi}^{(n)}(s) = \frac{n!}{c^n} \left[\check{\mathcal{G}}(s) \right]^n [q_0]. \quad (4.31)$$

4.2.2. *Self-normalized measurement types.* An important type of measurement operator is one that is *self-normalized*

$$\bar{T} = \frac{T}{\mathcal{E}}. \quad (4.32)$$

Examples are

normalized early light: $\mathcal{M}_{\bar{E}(T)}[\Gamma(t)] = \frac{1}{\mathcal{E}[\Gamma(t)]} \int_0^T \Gamma(t) dt,$ (4.33)

n th temporal moment: $\mathcal{M}_{(t^n)}[\Gamma(t)] = \frac{1}{\mathcal{E}[\Gamma(t)]} \int_{-\infty}^{\infty} t^n \Gamma(t) dt,$ (4.34)

n th central moment: $\mathcal{M}_{c_n}[\Gamma(t)] = \frac{1}{\mathcal{E}[\Gamma(t)]} \int_{-\infty}^{\infty} (t - \langle t \rangle)^n \Gamma(t) dt,$ (4.35)

normalized Laplace transform: $\mathcal{M}_{\bar{L}(s)}[\Gamma(t)] = \frac{1}{\mathcal{E}[\Gamma(t)]} \int_{-\infty}^{\infty} s e^{-st} \Gamma(t) dt,$ (4.36)

normalized Mellin–Laplace transform:

$$\mathcal{M}_{\bar{ML}(n,s)}[\Gamma(t)] = \frac{1}{\mathcal{E}[\Gamma(t)]} \int_{-\infty}^{\infty} \frac{s^{n+1}}{B(n+1)} t^n e^{-st} \Gamma(t) dt. \quad (4.37)$$

4.3. Other data types

The data types given above are all functionals of the exitance (4.5). As we will show in section 5 these functionals are differentiable in the Fréchet sense, and therefore admit of iterative reconstruction algorithms. However, there are other approaches such as the logarithmic slope of the temporal decay of Γ , the peak intensity, and the zero-crossings of the phase of a dipole like source (known as the ‘anti-phase’ measurement).

One possibility that has received some attention is the used of *fitted parameter* data types. The idea is to derive from the data a projected μ_a and/or μ'_s measurement by nonlinear fitting the given data to a known solution such as the infinite space Green function [14]. It is then argued that the data represents a Fredholm integral equation with a kernel that may be derived by simple arguments. For example, for DC intensity data the Green function is

$$G^{(\Phi)}(\mathbf{r}, \mathbf{r}_0) = \frac{\exp(-(\frac{\mu_a}{\kappa})^{\frac{1}{2}} |\mathbf{r} - \mathbf{r}_0|)}{|\mathbf{r} - \mathbf{r}_0|}. \quad (4.38)$$

Therefore the difference of the log E data type between two states (μ_a, κ) and $(\tilde{\mu}_a, \tilde{\kappa})$ is taken to be

$$\log E^\delta(d) \simeq (\chi - \tilde{\chi})d = \int_{\Omega} K(\mathbf{r}', d) \chi^\delta(\mathbf{r}') d^n \mathbf{r}' \quad (4.39)$$

where

$$\chi = \left(\frac{\mu_a}{\kappa}\right)^{\frac{1}{2}} \quad (4.40)$$

and $K(\mathbf{r}', d)$ is a generalized weight kernel. This very simple idea has been used to derive a backprojection-style reconstruction algorithm [40]. More generally it is hoped to derive

$$\begin{aligned} \mu_a^\delta \text{ fitted } d &= \int_{\Omega} K_{\mu_a}(\mathbf{r}', d) \mu_a(\mathbf{r}') d^n \mathbf{r}' \\ \mu'_s \delta \text{ fitted } d &= \int_{\Omega} K_{\mu'_s}(\mathbf{r}', d) \mu'_s(\mathbf{r}') d^n \mathbf{r}'. \end{aligned} \quad (4.41)$$

For example Benaron *et al* fitted μ_a to the logarithmic slope of the temporal decay and used backprojection onto a curvilinear basis [41].

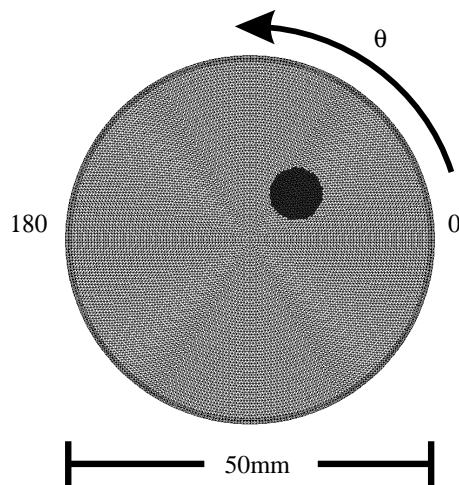


Figure 5. Model for figures 6–9.

4.4. What does the data look like?

Some data types are illustrated in figures 6–9. The case illustrated is a circle radius 25 mm with $\mu_a = 0.025 \text{ mm}^{-1}$, $\mu_s = 2 \text{ mm}^{-1}$, shown in figure 5, containing an off-centre circular region Ω_A , radius 3.125 mm. We show in each case the change in measurement as a $(\vartheta_m, \vartheta_p)$ map over the angular position of the source and detector, separately for the case where Ω_A has an absorption perturbation $\tilde{\mu}_a = 2\mu_a$ or a scatter perturbation $\tilde{\mu}_s = 2\mu_s$. The maps in this form are the equivalent of sinograms in conventional x-ray CT.

Figure 6 shows a set of gated temporal averages of width 200 ps from times 0–2000 ps. Each image is separately normalized and is shown on a log scale. Note that for longer times, the backscattered signal is negligible, and therefore the data at angles close to the source, defined by $|\vartheta_m - \vartheta_p| < 20^\circ$, has been reset to zero. The feature of interest is the increasing width of the perturbing signal with later times. Figure 7 shows the same cases for the μ_s perturbation. Note that the form of the data is very similar suggesting that absorption and scattering objects have a broadly similar effect. Some differences can be seen close to the source and at longer times, where the effect of the scatterer has locally increased the signal intensity, whereas it is always decreased for an absorber. Figures 8 and 9 show the equivalent maps for frequency domain and normalized integral transform domain data for a variety of cases. The features of interest are the extent to which the μ_a and μ_s perturbations give rise to different effects in the data. Note in particular that the real part of the complex intensity has the same sign for both absorber and scatterer, whereas the imaginary parts have opposite sign. This sign change is also seen in the DC intensity and mean time maps.

Choice of appropriate measurement depends on robustness of the experimental measurements with respect to systematic errors such as fluctuations of the source power, detector sensitivity, or fibre coupling losses, the ability of the forward model to generate the corresponding data efficiently, and the relation to stability and nonlinearity of the inverse problem. Some of these aspects are dealt with in later sections, and in other publications [42,43].

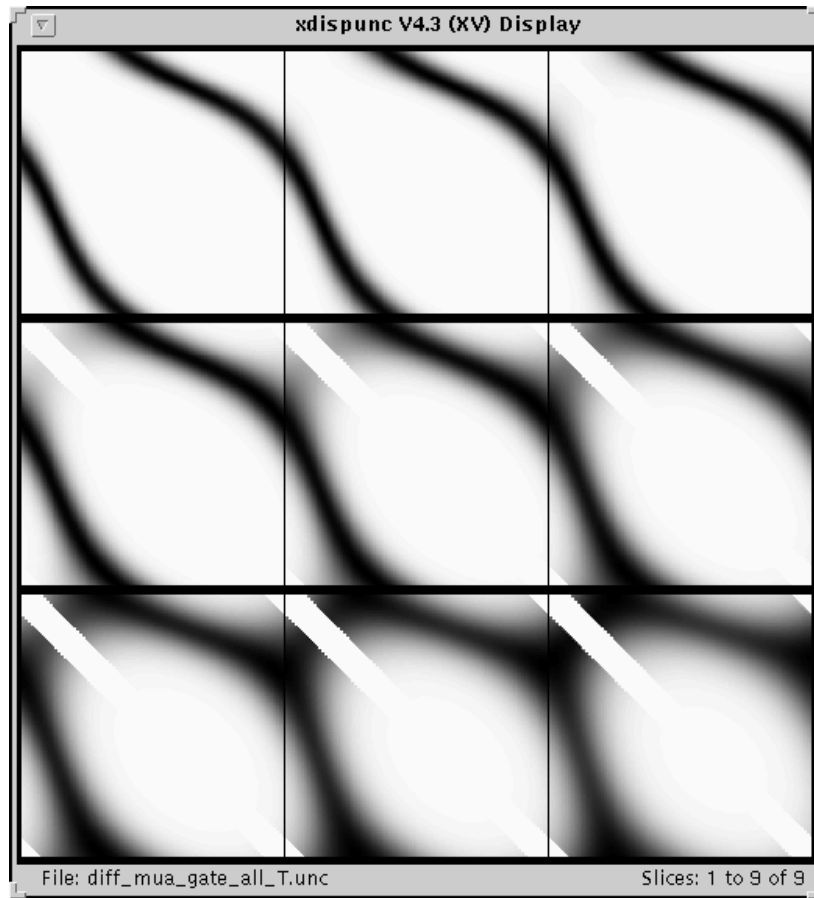



Figure 6. Successive time gates of intervals of 200 ps of the perturbed temporal data $\tilde{\Gamma}(t) - \Gamma(t)$ on a circle radius 25 mm with $\mu_a = 0.025 \text{ mm}^{-1}$, $\mu'_s = 2 \text{ mm}^{-1}$ separately normalized and shown on a log scale. Perturbation is μ_a . Abscissa is the detector position ϑ_m and the ordinate is the source position ϑ_p . For angles $|\vartheta_m - \vartheta_p| < 20^\circ$ the data is set to zero, giving rise to the diagonal stripe in each image.

 An animated GIF of this figure is available from the article's abstract page in the online journal; see <http://www.iop.org>.

5. Sensitivity relations—continuous case

5.1. Reciprocity theorem for the Boltzmann equation

A standard reciprocity theorem for the Boltzmann equation is given in Case and Zweifel [24], section 2.7:

$$G^{(\phi)}(\mathbf{r}, \hat{\mathbf{s}}; \mathbf{r}_0, \hat{\mathbf{s}}_0) = G^{(\phi)}(\mathbf{r}_0, -\hat{\mathbf{s}}_0; \mathbf{r}, -\hat{\mathbf{s}}) \quad (5.1)$$

which states that the angular density at \mathbf{r} in direction $\hat{\mathbf{s}}$ due to a source at \mathbf{r}_0 in direction $\hat{\mathbf{s}}_0$ is the same as the angular density at \mathbf{r}_0 in direction $-\hat{\mathbf{s}}_0$ due to a source at \mathbf{r} in direction $-\hat{\mathbf{s}}$. By integrating (5.1) over $\hat{\mathbf{s}}_0$ we obtain

$$G^{(\phi)}(\mathbf{r}, \hat{\mathbf{s}}; q_0(\mathbf{r}_0)) = G^{(\phi)}(\mathbf{r}_0; \mathbf{r}, -\hat{\mathbf{s}}) \quad (5.2)$$

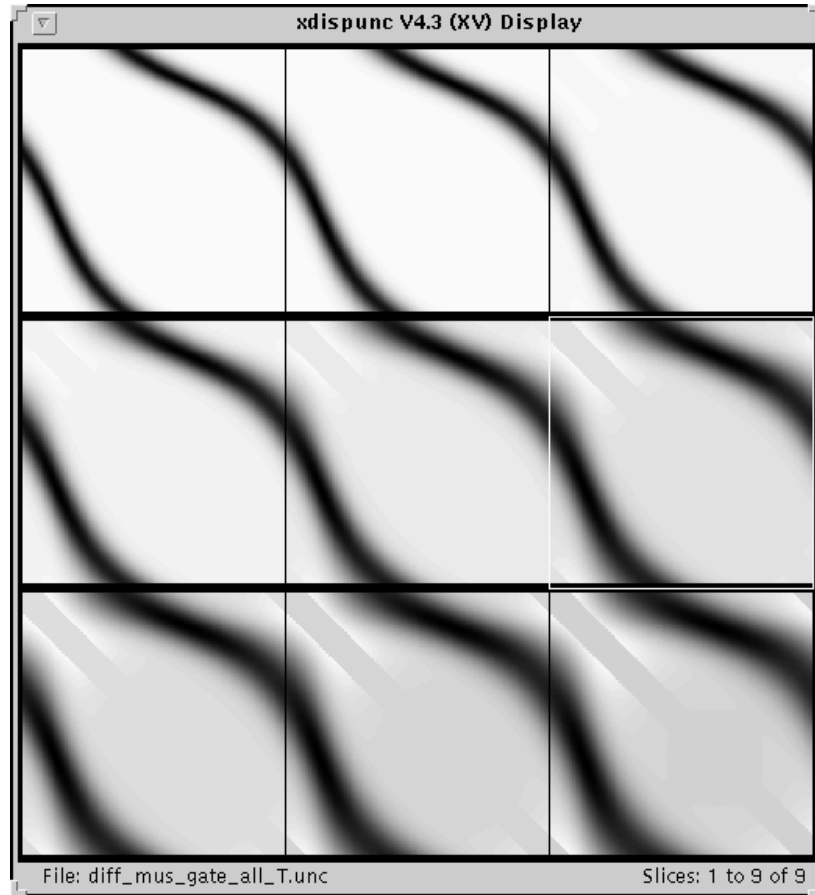


Figure 7. As figure 6, for the μ'_s perturbation.

[M] An animated GIF of this figure is available from the article's abstract page in the online journal; see <http://www.iop.org>.

which states that the angular density at r in direction \hat{s} due to a point isotropic source at r_0 is the same as the photon density at r_0 due to a source at r in direction $-\hat{s}$.

These basic relations give rise to the adjoint formulation of the sensitivity relations for the Boltzmann equation, which we return to in section 5.6. First we will derive these relations directly for the diffusion equation.

5.2. Diffusion model

We will develop the sensitivity relations in the frequency domain and obtain the other domains by transformation. We define the operators

$$\hat{\mathcal{D}}_{a,b}[u] = \left\{ -\nabla \cdot b \nabla + \left(a + \frac{i\omega}{c} \right) \right\} u \quad (5.3)$$

$$\mathcal{D}_{a,b}^\delta[u] = \{ -\nabla \cdot b \nabla + a \} u. \quad (5.4)$$

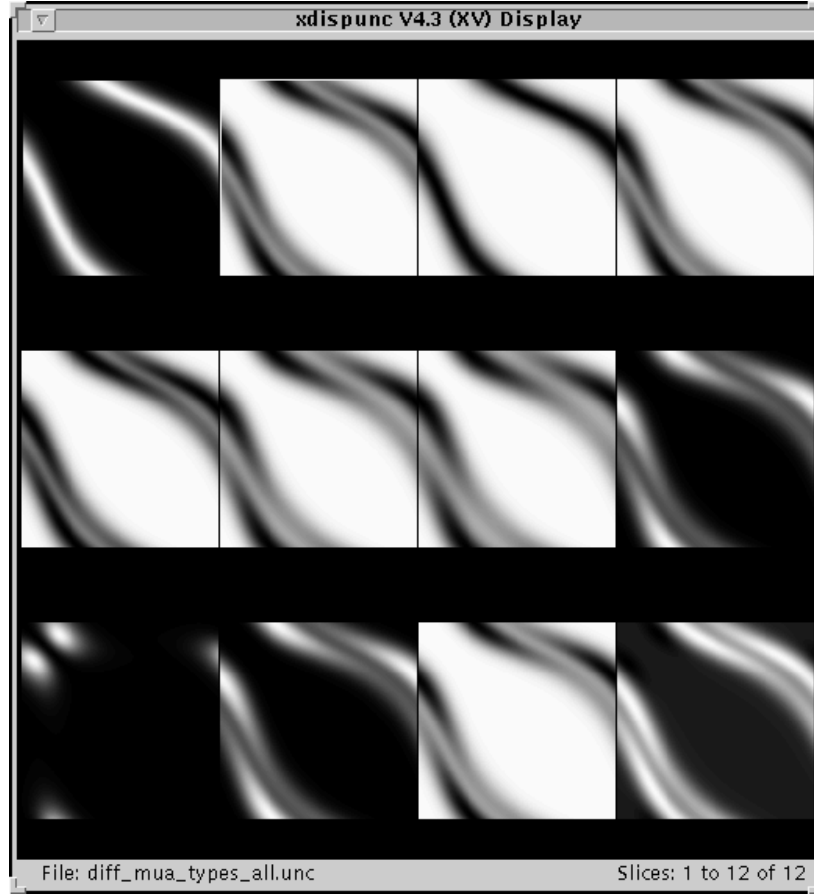


Figure 8. Comparison of data, μ_a perturbation: real and imaginary parts of $\log \Gamma$, $\log E$, $\langle t \rangle$, $\langle t^2 \rangle$, c_2 , c_3 , $L(s = 0.001)$, $L(s = 0.01)$, $ML(s = 0.002, n = 1)$, $ML(s = 0.002, n = 5)$, $ML(s = 0.0004, n = 1)$.

5.2.1. *Direct form.* Let Φ be a solution to (4.1), for the parameters (μ_a, κ) , and let $\tilde{\Phi} = \Phi + \Phi^\delta$ be the solution for the same source, but a different set of parameters $(\tilde{\mu}_a, \tilde{\kappa})$ where

$$\tilde{\mu}_a = \mu_a + \alpha \quad \tilde{\kappa} = \kappa + \beta. \quad (5.5)$$

By expansion and elimination of second-order terms we have

$$\hat{D}_{\mu_a, \kappa}[\Phi^\delta(\mathbf{r}, \omega)] = -\mathcal{D}_{\alpha, \beta}^\delta[\Phi(\mathbf{r}, \omega)] \quad (5.6)$$

which we can express as

$$-\nabla \cdot \kappa(\mathbf{r}) \nabla \Phi^\delta(\mathbf{r}, \omega) + \left(\mu_a(\mathbf{r}) + \frac{i\omega}{c} \right) \Phi^\delta(\mathbf{r}, \omega) = \nabla \cdot \beta(\mathbf{r}) \nabla \Phi(\mathbf{r}, \omega) - \alpha(\mathbf{r}) \Phi(\mathbf{r}, \omega) \quad (5.7a)$$

$$\Phi^\delta(\mathbf{m}, \omega) + 2A\kappa(\mathbf{m}) \frac{\partial \Phi^\delta(\mathbf{m}, \omega)}{\partial \nu} = 0 \quad (5.7b)$$

where the Robin boundary condition is the same as (4.1b). Equation (5.7) can be formally solved by applying the Green operator $\hat{\mathcal{G}}$ which leads to

$$\Phi^\delta(\mathbf{r}, \omega) = -\hat{\mathcal{G}}_{\alpha, \beta}^\delta[\Phi]$$

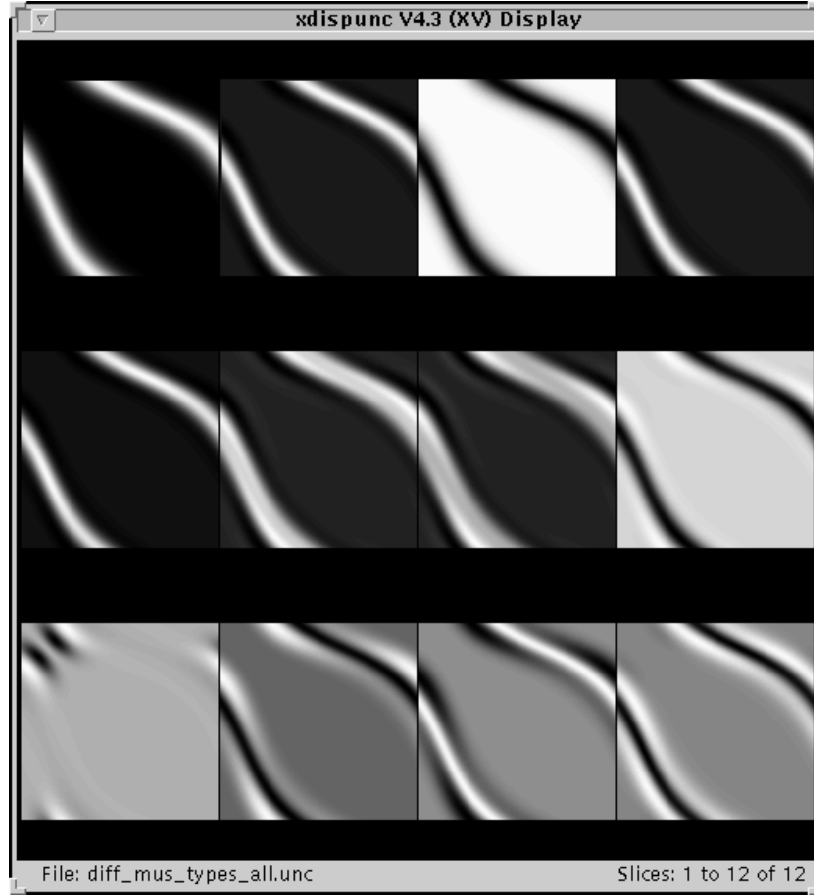


Figure 9. As figure 8, for the μ'_s perturbation.

$$\begin{aligned}
 &= -\hat{\mathcal{G}}[{-\nabla \cdot \beta(\mathbf{r})\nabla + \alpha(\mathbf{r})}\Phi(\mathbf{r}, \omega)] \\
 &= \int_{\Omega} [G^{(\Phi)}(\mathbf{r}, \mathbf{r}', \omega) \{\nabla_{\mathbf{r}'} \cdot \beta(\mathbf{r}')\nabla_{\mathbf{r}'} - \alpha(\mathbf{r}')\}\Phi(\mathbf{r}', \omega)] d^n \mathbf{r}'. \quad (5.8)
 \end{aligned}$$

Applying the divergence theorem and insisting that $\beta|_{\partial\Omega} = 0$ yields the Born approximation

$$\Phi^\delta(\mathbf{r}, \omega) = - \int_{\Omega} [\beta(\mathbf{r}')\nabla_{\mathbf{r}'} G^{(\Phi)}(\mathbf{r}, \mathbf{r}', \omega) \cdot \nabla_{\mathbf{r}'} \Phi(\mathbf{r}', \omega) + \alpha(\mathbf{r}')G^{(\Phi)}(\mathbf{r}, \mathbf{r}', \omega)\Phi(\mathbf{r}', \omega)] d^n \mathbf{r}'. \quad (5.9)$$

Of particular interest is the change in measurement

$$\Gamma_{\mu_a, \kappa}^\delta \left(\begin{array}{c} \alpha \\ \beta \end{array} \right) = \Gamma_{\mu_a + \alpha, \kappa + \beta} - \Gamma_{\mu_a, \kappa} = -(\kappa + \beta) \frac{\partial(\Phi + \Phi^\delta)}{\partial \nu} + \kappa \frac{\partial \Phi}{\partial \nu}. \quad (5.10)$$

If we again apply the condition that $\beta|_{\partial\Omega} = 0$ then

$$\Gamma_{\mu_a, \kappa}^\delta \left(\begin{array}{c} \alpha \\ \beta \end{array} \right) = -\kappa \frac{\partial \Phi^\delta}{\partial \nu}. \quad (5.11)$$

Applying the measurement operator \mathcal{B} to (5.9) gives

$$\Gamma_{\mu_a, \kappa}^{\delta} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} (\mathbf{m}, \omega) = - \int_{\Omega} [\beta(\mathbf{r}') \nabla_{\mathbf{r}'} G^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \omega) \cdot \nabla_{\mathbf{r}'} \Phi(\mathbf{r}', \omega) + \alpha(\mathbf{r}') G^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \omega) \Phi(\mathbf{r}', \omega)] d^n \mathbf{r}'. \quad (5.12)$$

The equivalent forms in the time domain are obtained by Fourier transform:

$$\Gamma_{\mu_a, \kappa}^{\delta} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} (\mathbf{m}, t) = - \int_{-\infty}^{\infty} \int_{\Omega} [\beta(\mathbf{r}') \nabla_{\mathbf{r}'} g^{(\Gamma)}(\mathbf{m}, \mathbf{r}', t') \cdot \nabla_{\mathbf{r}'} \Phi(\mathbf{r}', t - t') + \alpha(\mathbf{r}') g^{(\Gamma)}(\mathbf{m}, \mathbf{r}', t') \Phi(\mathbf{r}', t - t')] d^n \mathbf{r}' dt. \quad (5.13)$$

5.2.2. *Adjoint form.* Let Φ be the solution to (4.1) for a source q and let Ψ be a solution to the *adjoint problem*

$$-\nabla \cdot \kappa(\mathbf{r}) \nabla \Psi(\mathbf{r}, \omega) + \mu_a(\mathbf{r}) \Psi(\mathbf{r}, \omega) - \frac{i\omega}{c} \Psi(\mathbf{r}, \omega) = 0, \quad (5.14a)$$

$$\Psi(\mathbf{m}, \omega) + 2A\kappa(\mathbf{m}) \frac{\partial \Psi(\mathbf{m}, \omega)}{\partial \nu} = q^+(\mathbf{m}, \omega), \quad (5.14b)$$

where q^+ is a ‘Robin source’ on $\partial\Omega$ (not an internal source). We make use of the relation, for any functions u, w

$$\nabla \cdot (\kappa u \nabla w) = u \nabla \cdot (\kappa \nabla w) + \kappa \nabla u \cdot \nabla w \quad (5.15)$$

whence, by the divergence theorem

$$\int_{\Omega} -u \nabla \cdot \kappa \nabla w = \int_{\partial\Omega} -u \kappa \frac{\partial w}{\partial \nu} + \int_{\Omega} \kappa \nabla u \cdot \nabla w. \quad (5.16)$$

Consider the integral of the left-hand side in (5.7a):

$$\begin{aligned} \int_{\Omega} \Psi^* \left[-\nabla \cdot \kappa \nabla \Phi^{\delta} + \left(\mu_a + \frac{i\omega}{c} \right) \Phi^{\delta} \right] d^n \mathbf{r} \\ = \left(\int_{\Omega} \Phi^{\delta*} \left[-\nabla \cdot \kappa \nabla \Psi + \left(\mu_a - \frac{i\omega}{c} \right) \Psi \right] d^n \mathbf{r} \right)^* \\ + \int_{\partial\Omega} \left[\Phi^{\delta*} \kappa \frac{\partial \Psi^*}{\partial \nu} - \Psi^* \kappa \frac{\partial \Phi^{\delta}}{\partial \nu} \right] d^{n-1} \hat{\nu}_{\perp}. \end{aligned} \quad (5.17)$$

The first term on the right is zero by (5.14a). The second term is simplified by applying the Robin boundary condition (5.7b) for Φ^{δ} :

$$\begin{aligned} \int_{\partial\Omega} \left[\Phi^{\delta*} \kappa \frac{\partial \Psi^*}{\partial \nu} - \Psi^* \kappa \frac{\partial \Phi^{\delta}}{\partial \nu} \right] d^{n-1} \hat{\nu}_{\perp} &= \int_{\partial\Omega} \kappa \left[\left(-2A\kappa \frac{\partial \Phi^{\delta}}{\partial \nu} \right) \frac{\partial \Psi^*}{\partial \nu} - \Psi^* \frac{\partial \Phi^{\delta}}{\partial \nu} \right] d^{n-1} \hat{\nu}_{\perp} \\ &= \int_{\partial\Omega} -\kappa \frac{\partial \Phi^{\delta}}{\partial \nu} \left(2A\kappa \frac{\partial \Psi^*}{\partial \nu} + \Psi^* \right) d^{n-1} \hat{\nu}_{\perp} \\ &= \int_{\partial\Omega} q^{+*} \Gamma^{\delta} d^{n-1} \hat{\nu}_{\perp} \end{aligned} \quad (5.18)$$

where we used the boundary condition (5.14b) for Ψ .

Applying the divergence theorem the right-hand side of (5.7a) gives the adjoint form of (5.12):

$$\int_{\partial\Omega} q^{+*} \Gamma_{\mu_a, \kappa}^{\delta} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} d^{n-1} \hat{\nu}_{\perp} = - \int_{\Omega} \beta \nabla \Psi^* \cdot \nabla \Phi + \alpha \Psi^* \Phi d^n \mathbf{r}. \quad (5.19)$$

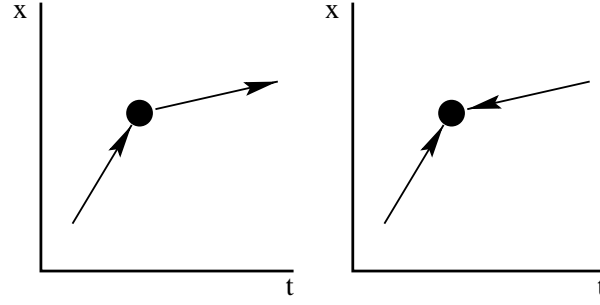


Figure 10. Feynman diagram for the direct (left) and adjoint (right) form.

If we now specify that q^+ is a δ -function at \mathbf{m}_0 , then the surface integral on the left in (5.19) ‘picks out’ the measurement at that point only, and Ψ is an adjoint Green function $G_R^{(\Phi)}$ with the property:

$$G_R^{(\Phi)*}(\mathbf{r}', \mathbf{m}_0, \omega) = G^{(\Gamma)}(\mathbf{m}_0, \mathbf{r}', \omega) \quad (5.20)$$

where the subscript R indicates that the δ -function source is Robin rather than the more usual Dirichlet or Neumann source. Equation (5.20) is the diffusion model equivalent of the reciprocity relations of the Boltzmann equation (5.2).

In the time domain we obtain by Fourier transform:

$$\begin{aligned} & \int_{-\infty}^{\infty} \int_{\partial\Omega} q^+(-t') \Gamma_{\mu_a, \kappa}^{\delta} \left(\begin{array}{c} \alpha \\ \beta \end{array} \right) (t-t') d^{n-1} \hat{\nu}_{\perp} dt' \\ &= \int_{-\infty}^{\infty} \int_{\Omega} [\alpha \Psi(-t') \Phi(t-t') + \beta \nabla \Psi(-t') \cdot \nabla \Phi(t-t')] d^n \mathbf{r} dt' \end{aligned} \quad (5.21)$$

where $\Psi(t)$ is a solution of the backward time-diffusion equation

$$-\nabla \cdot \kappa(\mathbf{r}) \nabla \Psi(\mathbf{r}, t) + \mu_a(\mathbf{r}) \Psi(\mathbf{r}, t) - \frac{1}{c} \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = 0, \quad (5.22a)$$

$$\Psi(\mathbf{m}, t) + 2A\kappa(\mathbf{m}) \frac{\partial \Psi(\mathbf{m}, t)}{\partial \nu} = q^+(\mathbf{r}, t), \quad (5.22b)$$

Specifying $q^+ = \delta(\mathbf{m}_0) \delta(t_0)$ gives rise to the adjoint Green function $g_R^{(\Phi)}$ with the property:

$$g_R^{(\Phi)}(\mathbf{r}', \mathbf{m}_0, t_0 - t) = g^{(\Gamma)}(\mathbf{m}_0, \mathbf{r}', t - t_0). \quad (5.23)$$

The Feynman diagram for (5.13) becomes the interaction of a forward and backward propagator as shown in figure 10.

5.3. Functional forms

If the source term is a δ -function at \mathbf{r}_0 , then we arrive at a functional form

$$\begin{aligned} \Phi^{\delta}(\mathbf{r}, \mathbf{r}_0, \omega) = & - \int_{\Omega} [\beta(\mathbf{r}') \nabla_{\mathbf{r}'} G^{(\Phi)}(\mathbf{r}, \mathbf{r}', \omega) \cdot \nabla_{\mathbf{r}'} G^{(\Phi)}(\mathbf{r}', \mathbf{r}_0, \omega) \\ & + \alpha(\mathbf{r}') G^{(\Phi)}(\mathbf{r}, \mathbf{r}', \omega) G^{(\Phi)}(\mathbf{r}', \mathbf{r}_0, \omega)] d^n \mathbf{r}'. \end{aligned} \quad (5.24)$$

The functional form of (5.12) is

$$\begin{aligned} \Gamma_{\mu_a, \kappa}^{\delta} \left(\begin{array}{c} \alpha \\ \beta \end{array} \right) (\mathbf{m}, \mathbf{r}_0, \omega) = & - \int_{\Omega} [\beta(\mathbf{r}') \nabla_{\mathbf{r}'} G^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \omega) \cdot \nabla_{\mathbf{r}'} G^{(\Phi)}(\mathbf{r}', \mathbf{r}_0, \omega) \\ & + \alpha(\mathbf{r}') G^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \omega) G^{(\Phi)}(\mathbf{r}', \mathbf{r}_0, \omega)] d^n \mathbf{r}' \end{aligned} \quad (5.25)$$

which in the time domain is

$$\begin{aligned} \Gamma_{\mu_a, \kappa}^{\delta} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} (\mathbf{m}, \mathbf{r}_0, t) = & - \int dt' \int_{\Omega} [\beta(\mathbf{r}') \nabla_{\mathbf{r}'} g^{(\Gamma)}(\mathbf{m}, \mathbf{r}', t') \cdot \nabla_{\mathbf{r}'} g^{(\Phi)}(\mathbf{r}', \mathbf{r}_0, t - t') \\ & + \alpha(\mathbf{r}') g^{(\Gamma)}(\mathbf{m}, \mathbf{r}', t') g^{(\Phi)}(\mathbf{r}', \mathbf{r}_0, t - t')] d^n \mathbf{r}'. \end{aligned} \quad (5.26)$$

Since (5.12) is linear in both α and β we can define the Fréchet derivatives

$$\begin{aligned} \frac{\partial \Gamma(\mathbf{m}, \omega)}{\partial \beta(\mathbf{r}')} &= -\nabla_{\mathbf{r}'} G^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \omega) \cdot \nabla_{\mathbf{r}'} \Phi(\mathbf{r}', \omega) \\ \frac{\partial \Gamma(\mathbf{m}, \omega)}{\partial \alpha(\mathbf{r}')} &= -G^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \omega) \Phi(\mathbf{r}', \omega) \end{aligned} \quad (5.27)$$

with the time-domain form defined similarly from (5.13). If we assume a δ -function source we obtain the functional forms:

$$\begin{aligned} \rho_{\kappa}^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \mathbf{r}_0, \omega) &= -\nabla_{\mathbf{r}'} G^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \omega) \cdot \nabla_{\mathbf{r}'} G^{(\Phi)}(\mathbf{r}', \mathbf{r}_0, \omega) \\ \rho_{\mu_a}^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \mathbf{r}_0, \omega) &= -G^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \omega) G^{(\Phi)}(\mathbf{r}', \mathbf{r}_0, \omega). \end{aligned} \quad (5.28)$$

5.4. Other measurement types

The fundamental forms for Φ^{δ} act as generating forms for the sensitivity relations for other measurement types. We can distinguish:

$$\text{functional transformations:} \quad \mathcal{M}_f[\Phi] = f(\Phi) \Rightarrow \mathcal{M}_f^{\delta}[\Phi] = f'(\Phi) \Phi^{\delta}, \quad (5.29)$$

$$\text{integral transformations:} \quad \mathcal{T}_w[\Phi] = \int_{-\infty}^{\infty} w(t) \Phi(t) dt \Rightarrow \mathcal{T}_w^{\delta}[\Phi] = \mathcal{T}_w[\Phi^{\delta}], \quad (5.30)$$

normalized integral transformations:

$$\bar{\mathcal{T}}_w[\Phi] = \frac{\int_{-\infty}^{\infty} w(t) \Phi(t) dt}{\int_{-\infty}^{\infty} \Phi(t) dt} \Rightarrow \bar{\mathcal{T}}_w^{\delta}[\Phi] = \frac{1}{\mathcal{E}[\Phi]} (\mathcal{T}_w[\Phi^{\delta}] - \mathcal{E}[\Phi^{\delta}] \bar{\mathcal{T}}_w[\Phi]). \quad (5.31)$$

Some important forms arise directly:

Equation (5.29) with $f = \log$ gives the Rytov approximation

$$\begin{aligned} \log \Gamma^{\delta}(\mathbf{m}, \omega) = & \frac{-1}{\Gamma(\mathbf{m}, \omega)} \int_{\Omega} [\beta(\mathbf{r}') \nabla_{\mathbf{r}'} G^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \omega) \cdot \nabla_{\mathbf{r}'} \Phi(\mathbf{r}', \omega) \\ & + \alpha(\mathbf{r}') G^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \omega) \Phi(\mathbf{r}', \omega)] d^n \mathbf{r}' \end{aligned} \quad (5.32)$$

with the time-domain form

$$\begin{aligned} \log \Gamma^{\delta}(\mathbf{m}, t) = & \frac{-1}{\Gamma(\mathbf{m}, t)} \int_{-\infty}^{\infty} \int_{\Omega} [\beta(\mathbf{r}') \nabla_{\mathbf{r}'} g^{(\Gamma)}(\mathbf{m}, \mathbf{r}', t') \cdot \nabla_{\mathbf{r}'} \Phi(\mathbf{r}', t - t') \\ & + \alpha(\mathbf{r}') g^{(\Gamma)}(\mathbf{m}, \mathbf{r}', t') \Phi(\mathbf{r}', t - t')] d^n \mathbf{r}' dt'. \end{aligned} \quad (5.33)$$

Note that (5.33) is not the Fourier transform of (5.32).

Equation (5.30) with $w = t^n \exp(-st)$ gives

$$\begin{aligned} \check{\Gamma}^{(n)\delta}(\mathbf{m}, s) = & - \int_{\Omega} \left[\beta(\mathbf{r}') \sum_{n'=1}^n \binom{n}{n'} \right. \\ & \times \left\{ \nabla_{\mathbf{r}'} \left(\frac{\partial^{n-n'}}{\partial \omega^{n-n'}} G^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \omega - is) \right) \cdot \nabla_{\mathbf{r}'} \left(\frac{\partial^{n'}}{\partial \omega^{n'}} \Phi(\mathbf{r}', \omega - is) \right) \right\} \\ & + \alpha(\mathbf{r}') \sum_{n'=1}^n \binom{n}{n'} \\ & \times \left. \left[\left(\frac{\partial^{n-n'}}{\partial \omega^{n-n'}} G^{(\Gamma)}(\mathbf{m}, \mathbf{r}', \omega - is) \right) \left(\frac{\partial^{n'}}{\partial \omega^{n'}} \Phi(\mathbf{r}', \omega - is) \right) \right] \right] \Big|_{\omega=0} d^n \mathbf{r}' \end{aligned} \quad (5.34)$$

with the normalized form

$$\overline{\Gamma^{(n)\delta}}(\mathbf{m}, s) = \frac{-1}{E(\mathbf{m})} [\check{\Gamma}^{(n)\delta}(\mathbf{m}, s) - \overline{\check{\Gamma}^{(n)}}(\mathbf{m}, s) E^\delta(\mathbf{m})] \quad (5.35)$$

where $E(\mathbf{m}) = \Gamma(\mathbf{m}, 0)$ is the DC measurement, and $E^\delta(\mathbf{m}) = \Gamma^\delta(\mathbf{m}, 0)$ is the respective Fréchet derivative.

Application of this procedure to the case where Φ is a Green function leads to a variety of functional forms for the *photon measurement density functions* (PMDFs) in extension of (5.28). For special cases such as infinite space, half spaces and slabs, the Green function can be given explicitly which allows closed forms of the PMDFs to be tabulated [44].

5.5. Other parameter bases

We have developed the PMDFs in the (μ_a, κ) form. It is simple to obtain these in other forms using the following relations:

$$\frac{d}{d\mu_a} = \frac{\partial}{\partial \mu_a} + \frac{\partial \kappa}{\partial \mu_a} \frac{\partial}{\partial \kappa} = \frac{\partial}{\partial \mu_a} - 3\kappa^2 \frac{\partial}{\partial \kappa} \quad (5.36)$$

$$\frac{d}{d\mu'_s} = \frac{\partial \kappa}{\partial \mu'_s} \frac{\partial}{\partial \kappa} = -3\kappa^2 \frac{\partial}{\partial \kappa} \quad (5.37)$$

$$\frac{d}{d\mu_s} = \frac{\partial \mu'_s}{\partial \mu_s} \frac{\partial}{\partial \mu'_s} = -3\kappa^2 (1 - \Theta_1) \frac{\partial}{\partial \kappa}. \quad (5.38)$$

This leads to

$$\begin{aligned} (\mu_a, \mu'_s): \quad \rho_{\mu_a} &= \Psi^* \Phi - 3\kappa^2 \nabla \Psi^* \cdot \nabla \Phi \\ \rho_{\mu'_s} &= -3\kappa^2 \nabla \Psi^* \cdot \nabla \Phi \end{aligned} \quad (5.39)$$

$$\begin{aligned} (\mu_a, \mu_s): \quad \rho_{\mu_a} &= \Psi^* \Phi - 3\kappa^2 \nabla \Psi^* \cdot \nabla \Phi \\ \rho_{\mu_s} &= -3(1 - \Theta_1) \kappa^2 \nabla \Psi^* \cdot \nabla \Phi. \end{aligned} \quad (5.40)$$

5.6. Sensitivity relations for the Boltzmann equation

In place of (4.5) we have

$$\Gamma(\mathbf{m}) = \mathcal{B}[\phi] = \int_{\hat{\nu} \cdot \hat{s} > 0} w(\hat{s}) \hat{s} \phi(\mathbf{m}, \hat{s}) d\hat{s} \quad (5.41)$$

where $w(\hat{s})$ is a weighting term for the angular intensity. We define the operators

$$\begin{aligned} \hat{\mathcal{L}}_{a,b}[\phi(\mathbf{r}, \hat{s}, \omega)] &= \left\{ \hat{s} \cdot \nabla + \left(a(\mathbf{r}) + b(\mathbf{r}) + \frac{i\omega}{c} \right) \right\} \phi(\mathbf{r}, \hat{s}, \omega) \\ &\quad - b(\mathbf{r}) \int_{S^{n-1}} \Theta(\hat{s} \cdot \hat{s}') \phi(\mathbf{r}, \hat{s}', \omega) d\hat{s}' \end{aligned} \quad (5.42)$$

$$\mathcal{L}_{a,b}^\delta[\phi(\mathbf{r}, \hat{s}, \omega)] = \{a(\mathbf{r}) + b(\mathbf{r})\} \phi(\mathbf{r}, \hat{s}, \omega) - b(\mathbf{r}) \int_{S^{n-1}} \Theta(\hat{s} \cdot \hat{s}') \phi(\mathbf{r}, \hat{s}', \omega) d\hat{s}'. \quad (5.43)$$

5.6.1. Direct form. Let ϕ be a solution to (3.2), for the parameters (μ_a, μ_s) , and let $\tilde{\phi} = \phi + \phi^\delta$ be the solution for the same source, but a different set of parameters $(\tilde{\mu}_a, \tilde{\mu}_s)$ where

$$\tilde{\mu}_a = \mu_a + \alpha \quad \tilde{\mu}_s = \mu_s + \gamma. \quad (5.44)$$

The equivalent of (5.6) is

$$\hat{\mathcal{L}}_{\mu_a, \mu_s}[\phi^\delta(\mathbf{r}, \hat{s}, \omega)] = -\mathcal{L}_{\alpha, \gamma}^\delta[\phi(\mathbf{r}, \hat{s}, \omega)]. \quad (5.45)$$

5.6.2. *Adjoint form.* The equivalent to (5.19) is given by (Dorn [45])

$$\int_{\partial\Omega} \psi^* \Gamma_{\mu_a, \mu_s}^\delta \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} d^{n-1} \hat{\nu}_\perp = \int_\Omega \int_{S^{n-1}} \left[-\alpha \psi^* \phi + \gamma \psi^* \left(\int_{S^{n-1}} \Theta(\hat{s} \cdot \hat{s}') \phi d\hat{s}' - \phi \right) \right] d\hat{s} d^n r \quad (5.46)$$

where ψ is a solution to the adjoint transport problem

$$\left(-\frac{i\omega}{c} - \hat{s} \cdot \nabla + \mu_{tr}(\mathbf{r}) \right) \psi(\mathbf{r}, \hat{s}, \omega) - \mu_s(\mathbf{r}) \int_{S^{n-1}} \Theta(\hat{s} \cdot \hat{s}') \psi(\mathbf{r}, \hat{s}', \omega) d\hat{s}' = 0, \quad (5.47a)$$

$$\psi(\mathbf{r}, \hat{s}, \omega) = q^+(\mathbf{r}, \omega) \quad \text{on} \quad \partial\Omega \times \{\hat{\nu} \cdot \hat{s} > 0\}. \quad (5.47b)$$

5.7. Discussion

Sensitivity relations in various forms have been developed many times, in the majority of treatments only for the absorption term. The form (5.26) was presented by Arridge *et al* [11] and the general framework of PMDFs by Arridge [44, 46]. The absorption term in (5.26) was derived by Schotland *et al* [47] by taking the first term in the Feynman path integral expression for the Hamiltonian operator. Feng *et al* [48] derived the absorption term in (5.25) by taking the limiting value of the exact perturbed intensity of a spherical inhomogeneity as the radius of the inhomogeneity reduced to zero. Sevick *et al* [49] derived the absorption term in (5.25) using a Monte Carlo argument. Equation (5.40) was derived from the Boltzmann equation directly by Chang *et al* [50] using an adjoint transport source.

The absorption term in (5.33) is what Schotland *et al* [47] termed the ‘photon hitting density’. Arridge *et al* [51] compared reconstruction from intensity and log intensity measurements and found the latter provided considerable improvement. This was later confirmed by O’Leary *et al* [10] who pointed out that the logarithmic intensity is equivalent to the Rytov approximation whereas absolute intensity is the Born approximation.

6. Computational methods

6.1. The finite element method (FEM)

The domain Ω is divided into P elements, joined at D vertex nodes. The solution Φ is approximated by the piecewise function $\Phi^h(\mathbf{r}, t) = \sum_i^D \Phi_i(t) u_i(\mathbf{r}) \in \mathcal{U}^h$, where \mathcal{U}^h is a finite-dimensional subspace spanned by basis functions $u_i(\mathbf{r})$ ($i = 1 \dots D$) chosen to have limited support. The problem of solving for Φ^h becomes one of sparse matrix inversion. The advantage of the FEM approach is its versatility which makes it applicable to complex geometries and highly inhomogeneous parameter distributions.

As developed in [27, 52], the diffusion equation in the FEM framework is expressed, in the frequency domain as:

$$(\mathbf{K}(\kappa) + \mathbf{C}(\mu) + \zeta \mathbf{A} + i\omega \mathbf{B}) \Phi = \mathbf{q}_0 \quad (6.1)$$

where $\zeta = 1/2A$, and in the time domain as:

$$(\mathbf{K}(\kappa) + \mathbf{C}(\mu) + \zeta \mathbf{A}) \Phi(t) + \mathbf{B} \frac{\partial \Phi(t)}{\partial t} = \mathbf{q}_0(t) \quad (6.2)$$

where the *system matrices* $\mathbf{K}, \mathbf{C}, \mathbf{A}$ and \mathbf{B} have entries given by:

$$K_{ij} = \int_\Omega \kappa(\mathbf{r}) \nabla u_i(\mathbf{r}) \cdot \nabla u_j(\mathbf{r}) d^n r \quad (6.3)$$

$$C_{ij} = \int_{\Omega} \mu_a(\mathbf{r}) u_i(\mathbf{r}) u_j(\mathbf{r}) d^n \mathbf{r} \quad (6.4)$$

$$B_{ij} = \frac{1}{c} \int_{\Omega} u_i(\mathbf{r}) u_j(\mathbf{r}) d^n \mathbf{r} \quad (6.5)$$

$$A_{ij} = \int_{\partial\Omega} u_i(\mathbf{r}) u_j(\mathbf{r}) d^{n-1} \hat{\nu}_{\perp}. \quad (6.6)$$

To relate the FEM approach to the forward model, we define the discrete–continuous operator

$$\mathcal{P}_j = \mathcal{M}[\Phi_j] \quad (6.7)$$

where Φ_j is the solution to (6.1) for the j th source, and $\mathcal{M} : \mathbf{G} \rightarrow \mathbf{Y}$ is a *measurement operator*. In discrete mode, we define also $\mathcal{M}_j : \mathbb{R}^D \rightarrow \mathbb{R}^{M_j}$. The equivalent of (6.7) is the discrete–discrete operator

$$\mathbf{F}_j = \mathcal{M}_j[\Phi_j]. \quad (6.8)$$

Equation (6.1) is formally solved by matrix inversion:

$$\Phi = (\mathbf{K}(\kappa) + \mathbf{C}(\mu) + \zeta \mathbf{A} + i\omega \mathbf{B})^{-1} \mathbf{q}_0 = \hat{\mathbf{G}}(\omega) \mathbf{q}_0 \quad (6.9)$$

where $\hat{\mathbf{G}}(\omega)$ is the discrete representation of the Green operator $\hat{\mathcal{G}}$. In the time domain, (6.2) is solved using a finite-differencing scheme in time [52] which effectively increases the time for the forward problem solution by $T_{MAX} = \frac{1}{\Delta t} (T_{final} - T_{initial})$. Here $T_{MAX} > T_{j,i}$ is dictated by stability and accuracy considerations, so that even if only a small number of time samples are used in the problem specification, the computation time of the problem is still increased by an order of magnitude. Typical values of T_{MAX} are 100–1000. Note that in the finite difference scheme outlined below, operator-splitting methods can be employed, so that the computation time becomes reasonable.

Since the matrices \mathbf{K} , \mathbf{C} , \mathbf{A} , \mathbf{B} are sparse, the solution is obtained by standard methods such as Cholesky reduction and substitution [52], or biconjugate gradients in the complex case [53]. For very large problems, such as occur in 3D, a direct sparse solver such as conjugate gradients can be employed [54].

6.1.1. Measurement types. Measurement types are derived by applying the measurement operators to the left-hand side of (6.9). The simplest measurement operator to consider is the *boundary operator* \mathcal{B} which implements (4.5).

$$\Gamma = \mathbf{D} \hat{\mathbf{G}} \mathbf{q}_0. \quad (6.10)$$

If the measurement i is considered to be a δ -function at \mathbf{m}_i , then let $\tau(\mathbf{m}_i)$ be the single boundary element that contains this point. The form of \mathbf{D} is a $M_j \times D$ matrix with entries given by

$$D_{ij} = \begin{cases} \zeta u_j(\mathbf{m}_i) & \text{if node } j \in \tau(\mathbf{m}_i), \\ 0 & \text{otherwise} \end{cases} \quad (6.11)$$

where we make use of (4.6). To implement measurements of integral transform type we follow the method of section 4.2.1. The iterative scheme (4.30) becomes [55]

$$\begin{aligned} \check{\Phi}^{(0)}(s) &= \check{\mathbf{G}}(s)[\mathbf{q}_0] \\ \check{\Phi}^{(n)}(s) &= \check{\mathbf{G}}(s)[n\mathbf{B}\check{\Phi}^{(n-1)}(s) + \mathbf{q}_0^{(n)}]. \end{aligned} \quad (6.12)$$

6.2. The finite difference method

The finite difference method (FDM) can be considered as a special case of the FEM on a regular grid. Generally speaking it is much harder to adapt FDM to complex boundary shapes and conditions, but this is offset by the considerably easier solution methods available when the system matrices have high degrees of symmetry. Not surprisingly therefore, many authors have used FDM to investigate the inverse problem for simulated data, but some experimental results have also been obtained, for the difference imaging problem [15].

The FDM replaces the continuous operators by a differencing representation, so that (4.1a) becomes (in three dimensions)

$$-\left(\frac{\partial}{\partial x}\kappa\frac{\partial\Phi}{\partial x}+\frac{\partial}{\partial y}\kappa\frac{\partial\Phi}{\partial y}+\frac{\partial}{\partial z}\kappa\frac{\partial\Phi}{\partial z}+\mu_a\Phi\right)\Big|_{l,m,p}+\frac{i\omega}{c}\Phi_{l,m,p}=q_{l,m,p}. \quad (6.13)$$

We evaluate each operator on a regular grid with spacing h in each dimension. To give a stable scheme, we evaluate first derivatives at midpoint positions, then form the second derivatives by differencing from these positions, with the scalar term interpolated

$$\begin{aligned} \frac{\partial}{\partial x}\kappa\frac{\partial\Phi}{\partial x}\Big|_{l,m,p} &= \frac{1}{h}\left[\kappa_{l+1/2,m,p}\frac{(\Phi_{l+1,m,p}-\Phi_{l,m,p})}{h}-\kappa_{l-1/2,m,p}\frac{(\Phi_{l,m,p}-\Phi_{l-1,m,p})}{h}\right] \\ \frac{\partial}{\partial y}\kappa\frac{\partial\Phi}{\partial y}\Big|_{l,m,p} &= \frac{1}{h}\left[\kappa_{l,m+1/2,p}\frac{(\Phi_{l,m+1,p}-\Phi_{l,m,p})}{h}-\kappa_{l,m-1/2,p}\frac{(\Phi_{l,m,p}-\Phi_{l,m-1,p})}{h}\right] \\ \frac{\partial}{\partial z}\kappa\frac{\partial\Phi}{\partial z}\Big|_{l,m,p} &= \frac{1}{h}\left[\kappa_{l,m,p+1/2}\frac{(\Phi_{l,m,p+1}-\Phi_{l,m,p})}{h}-\kappa_{l,m,p-1/2}\frac{(\Phi_{l,m,p}-\Phi_{l,m,p-1})}{h}\right]. \end{aligned}$$

The resultant discrete system can be solved by a variety of schemes of which the multigrid method is optimal [56].

In the time domain the time derivative in (4.2a) can be discretized using a forward, backward, or Crank–Nicholson scheme. A particular advantage is gained by using the alternating direction implicit (ADI) scheme, wherein Φ is updated in three intervals per time step

$$\begin{aligned} &\left(-\Delta t\frac{\partial}{\partial x}\kappa\frac{\partial\Phi^{(n+1/3)}}{\partial x}+\frac{\Delta t}{3}\mu_a\Phi^{(n+1/3)}+\frac{3}{c}\Phi^{(n+1/3)}\right)\Big|_{l,m,p} \\ &= \Delta tq_{l,m,p}^{(n)}+\left(\Delta t\frac{\partial}{\partial y}\kappa\frac{\partial\Phi^{(n)}}{\partial y}+\Delta t\frac{\partial}{\partial z}\kappa\frac{\partial\Phi^{(n)}}{\partial z}-\frac{2\Delta t}{3}\mu_a\Phi^{(n)}+\frac{3}{c}\Phi^{(n)}\right)\Big|_{l,m,p} \end{aligned} \quad (6.14)$$

$$\begin{aligned} &\left(-\Delta t\frac{\partial}{\partial y}\kappa\frac{\partial\Phi^{(n+2/3)}}{\partial y}+\frac{\Delta t}{3}\mu_a\Phi^{(n+2/3)}+\frac{3}{c}\Phi^{(n+2/3)}\right)\Big|_{l,m,p} \\ &= \Delta tq_{l,m,p}^{(n+1/3)}+\left(\Delta t\frac{\partial}{\partial x}\kappa\frac{\partial\Phi^{(n+1/3)}}{\partial x}+\Delta t\frac{\partial}{\partial z}\kappa\frac{\partial\Phi^{(n+1/3)}}{\partial z}\right. \\ &\quad \left.-\frac{2\Delta t}{3}\mu_a\Phi^{(n+1/3)}+\frac{3}{c}\Phi^{(n+1/3)}\right)\Big|_{l,m,p} \end{aligned} \quad (6.15)$$

$$\begin{aligned} &\left(-\Delta t\frac{\partial}{\partial z}\kappa\frac{\partial\Phi^{(n+1)}}{\partial z}+\frac{\Delta t}{3}\mu_a\Phi^{(n+1)}+\frac{3}{c}\Phi^{(n+1)}\right)\Big|_{l,m,p} \\ &= \Delta tq_{l,m,p}^{(n+2/3)}+\left(\Delta t\frac{\partial}{\partial x}\kappa\frac{\partial\Phi^{(n+2/3)}}{\partial x}+\Delta t\frac{\partial}{\partial y}\kappa\frac{\partial\Phi^{(n+2/3)}}{\partial y}\right. \\ &\quad \left.-\frac{2\Delta t}{3}\mu_a\Phi^{(n+2/3)}+\frac{3}{c}\Phi^{(n+2/3)}\right)\Big|_{l,m,p}. \end{aligned} \quad (6.16)$$

Not only is this scheme unconditionally stable and second-order accurate in time, but with appropriate solution re-indexing, each of the time steps is solved by a tridiagonal matrix inversion that is linear in the number of gridpoints.

6.3. Boltzmann equation models

Monte Carlo methods have a long pedigree, especially in transport theory [57]. The method proceeds by defining regions of Ω that have particular values of μ_a , μ_s and possibly $\Theta(\hat{s}, \hat{s}', \mathbf{r})$. Individual photons are run through the model, undergoing scattering and absorption events according to the local values of parameters, until they either have negligible contribution, or escape the surface $\partial\Omega$, thus contributing to a measurement [58]. Such methods offer great flexibility in modelling arbitrarily complex geometries and parameter distributions, but they are prohibitively costly in computational time—typically photon paths are several hundred interactions in length, and many millions of photons need to be followed to obtain sufficient statistics. Care has to be taken with the statistical properties of the algorithm used in order to avoid bias [5, 59, 60]. It has even been suggested to use Monte Carlo as the basis of an approach to the inverse problem [61] but results are very limited and this method has been largely superseded by FEM or FDM techniques.

Computational schemes for the Boltzmann equation have a similarly long pedigree in transport theory, generally classified into spherical harmonic or discrete ordinate schemes. Most schemes are of finite-difference type [45], but a finite element method for the Boltzmann equation is described by de Oliveira [62]. Debate as to the relative merits of these schemes is outside the scope of this paper, but attention is turning to them in optical tomography [37, 45].

6.4. Other methods

Other methods such as random walk theory (RWT) [63, 64] and the Markov random field method of Grünbaum and Patch [65–67] have been far less widely applied to the inverse problem and are left out of the discussion in this paper. Some further discussion of these methods can be found in another review paper [5].

7. Sensitivity relations—discrete case

We will develop the discret sensitivity relations in the FEM context, although the equivalent formulations in the FDM context are readily apparent. We assume that $\mu_a(\mathbf{r})$ and $\kappa(\mathbf{r})$ are expressed in a basis:

$$\kappa(\mathbf{r}) = \sum_{k=1}^{N_\kappa} \kappa_k v_k^{(\kappa)}(\mathbf{r}) \quad \mu_a(\mathbf{r}) = \sum_{k=1}^{N_\mu} \mu_{a_k} v_k^{(\mu)}(\mathbf{r}). \quad (7.1)$$

The equivalent of (6.7) is

$$\mathcal{P}'_j = \mathcal{M}'[\Phi_j] \quad (7.2)$$

and the equivalent of (6.8) is

$$\mathbf{F}'_j = \mathcal{M}'_j[\Phi_j]. \quad (7.3)$$

7.1. Finite element method

Using the basis expansion of $(\kappa(\mathbf{r}), \mu(\mathbf{r}))$ given in (7.1) we can further express \mathbf{K}, \mathbf{C} as:

$$\begin{aligned}\mathbf{K}(\kappa) &= \sum_k^{N_\kappa} \kappa_k \mathbf{V}_k(\kappa) \\ \mathbf{C}(\mu) &= \sum_k^{N_\mu} \mu_{a_k} \mathbf{V}_k(\mu)\end{aligned}\quad (7.4)$$

where $\mathbf{V}_k(\kappa), \mathbf{V}_k(\mu)$ represent *basis system matrices* whose entries are given by

$$\mathbf{V}_{k,ij}(\kappa) = \int_{\Omega} v_k^{(\kappa)}(\mathbf{r}) \nabla u_i(\mathbf{r}) \cdot \nabla u_j(\mathbf{r}) \, d^n \mathbf{r} \quad (7.5)$$

$$\mathbf{V}_{k,ij}(\mu) = \int_{\Omega} v_k^{(\mu)}(\mathbf{r}) u_i(\mathbf{r}) u_j(\mathbf{r}) \, d^n \mathbf{r}. \quad (7.6)$$

Note that the matrices $\mathbf{V}_k(\kappa), \mathbf{V}_k(\mu)$ will be sparse if the chosen basis functions have compact support.

7.1.1. Direct form. Following section 5.2.1 we differentiate (6.1) with respect to each basis coefficient defined in (7.4)

$$(\mathbf{K}(\kappa) + \mathbf{C}(\mu) + \zeta \mathbf{A} + i\omega \mathbf{B}) \frac{\partial \Phi}{\partial x_k} = \frac{\partial \mathbf{q}}{\partial x_k} - \mathbf{V}_k(x) \Phi \quad (7.7)$$

where x is either μ_a or κ .

Henceforward we will make the assumption that the source term is constant, so that the first term on the right is zero. Making use of (6.9) we obtain the analogy of (5.27)

$$\frac{\partial \Gamma}{\partial x_k} = \mathbf{D} \frac{\partial \Phi}{\partial x_k} = -\mathbf{D} \hat{\mathbf{G}} \mathbf{V}_k(x) \Phi. \quad (7.8)$$

7.1.2. Adjoint form. The discrete form of the reciprocity relation (5.20) is expressed, using the fact that $\hat{\mathbf{G}}$ is symmetric

$$\mathbf{D}^T(\mathbf{m}_i) \hat{\mathbf{G}} \mathbf{q}(\mathbf{r}_j) = \mathbf{q}^T(\mathbf{r}_j) \hat{\mathbf{G}} \mathbf{D}(\mathbf{m}_i) \quad (7.9)$$

where $\mathbf{D}^T(\mathbf{m}_i)$ is the single row of \mathbf{D} corresponding to measurement i . We can now state [46]

$$\begin{aligned}\mathbf{D}^T(\mathbf{m}_i) \frac{\partial \Phi_j}{\partial x_k} &= -\mathbf{D}^T(\mathbf{m}_i) \hat{\mathbf{G}} \mathbf{V}_k(x) \Phi_j \\ &= -\Phi_j^T \mathbf{V}_k(x) \hat{\mathbf{G}} \mathbf{D}(\mathbf{m}_i) \\ &= -\Phi_j^T \mathbf{V}_k(x) \Psi_i^* \\ &= -\Psi_i^{*T} \mathbf{V}_k(x) \Phi_j \triangleq -\langle \Psi_i, \mathbf{V}_k(x) \Phi_j \rangle\end{aligned}\quad (7.10)$$

where Ψ_i is the solution to the adjoint problem

$$(\mathbf{K} + \mathbf{C} + \zeta \mathbf{A} - i\omega \mathbf{B}) \Psi_i = \mathbf{q}_i^+ \quad (7.11)$$

where \mathbf{q}_i^+ is an *adjoint source*, initializing the boundary nodes of the element containing \mathbf{m}_i .

We can now state the analogy to (5.28)

$$\begin{aligned}\rho_\kappa(i, k, j) &= -\Psi_i^{*T} \mathbf{V}_k(\kappa) \Phi_j = -\langle \Psi_i, \mathbf{V}_k(\kappa) \Phi_j \rangle \\ \rho_\mu(i, k, j) &= -\Psi_i^{*T} \mathbf{V}_k(\mu) \Phi_j = -\langle \Psi_i, \mathbf{V}_k(\mu) \Phi_j \rangle.\end{aligned}\quad (7.12)$$

7.1.3. *Time domain.* Proceeding in the same manner as (7.7), we can write:

$$\left(\mathbf{K}(\kappa) + \mathbf{C}(\mu) + \zeta \mathbf{A} + \mathbf{B} \frac{\partial}{\partial t} \right) \frac{\partial \Phi}{\partial x_k} = \frac{\partial \mathbf{q}}{\partial x_k} - \mathbf{V}_k(x) \Phi(t). \quad (7.13)$$

The adjoint form leads to

$$\int_{T_{\text{initial}}}^{T_{\text{final}}} \Phi_i^{\text{T}}(-t') \mathbf{V}_k(x) \Phi_j(t-t') dt' \quad (7.14)$$

where $\Psi_i(t)$ is the solution to the adjoint problem:

$$\left(\mathbf{K}(\kappa) + \mathbf{C}(\mu) + \zeta \mathbf{A} - \mathbf{B} \frac{\partial}{\partial t} \right) \Psi_i(t) = \mathbf{q}_i^+(t) \quad (7.15)$$

which propagates backward in time from the initial condition $\Psi_i(t) = 0$ for $t > T_{\text{final}}$.

7.1.4. *Other measurement types.* As in section 5.4 we can derive the sensitivity functions $\rho_x^{(\mathcal{M})}(i, k, j)$ for other types from the form (7.12). The particular forms of interest are:

Logarithmic intensity (Rytov approximation)

$$\rho_x^{(\log)}(i, k, j) = \frac{1}{\Gamma_{j,i}(\omega)} \rho_x(i, k, j) \quad (7.16)$$

Integral transform ((5.30) with $w = t^n \exp(-st)$)

$$\rho_x^{(\mathcal{T}(n,s))}(i, k, j) = - \sum_{n'=1}^n \binom{n}{n'} \frac{\partial^{n'}}{\partial \omega^{n'}} \Psi_i^{*\text{T}}(\omega - is) \mathbf{V}_k(x) \frac{\partial^{n-n'}}{\partial \omega^{n-n'}} \Phi_j(\omega - is) \Big|_{\omega=0}. \quad (7.17)$$

Normalized integral transform ((5.31) with $w = t^n \exp(-st)$)

$$\rho_x^{(\bar{\mathcal{T}}(n,s))}(i, k, j) = \frac{1}{E_{j,i}} [\rho_x^{(\mathcal{T}(n,s))}(i, k, j) - \check{\Gamma}^{(n)}(s) \rho_x^{(\mathcal{E})}(i, k, j)]. \quad (7.18)$$

8. Illposedness and nonlinearity

8.1. Uniqueness

Uniqueness of the inverse problem for the Boltzmann equation is an open question. Choulli and Stefanov give a uniqueness result for (μ_a, μ_s, Θ) [68], and some special cases have been addressed by Anikonov [69, 70]. By contrast, (4.1) and (4.2) fit into the general framework of Isakov [20].

It is convenient to return to the Helmholtz version of the problem (3.48). If it were the case that the source q_0 in (3.48) was of Neumann type, as discussed in section 3.5.2, and we had complete knowledge of all pairs of Dirichlet and Neumann data on the boundary, then the result of Sylvester and Uhlmann for the real case [71] suggests that $\hat{\eta}$ is uniquely determined [72]. The situation is complicated if we use an interior source, and is a possible topic of research. However, the mechanism of an interior source is in a sense artificial, since in practice almost equivalent results are obtained for the different source models [27]. A second complication is that the diffusion approximation will not be valid very close to the source, or at short times. Therefore, technically, we cannot have complete data that satisfies (4.1). However, Arridge and Lionheart [73] recently produced a valuable result that lends insight into the difficulty of separating absorption and scattering parameters in a simultaneous reconstruction.

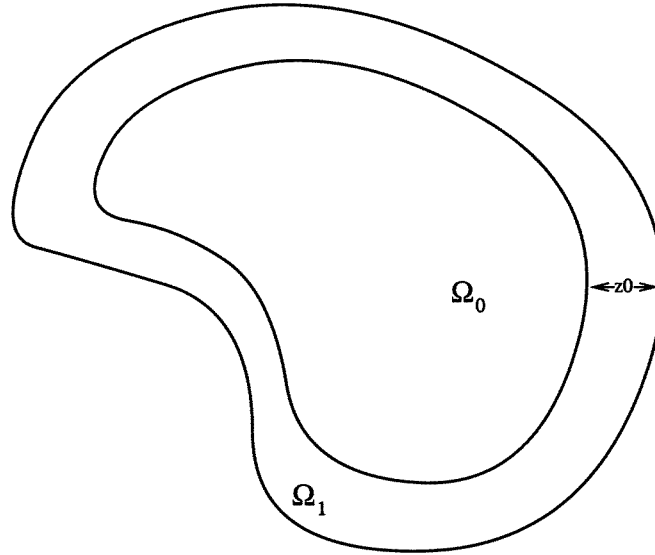


Figure 11. Definition of Ω_0 and Ω_1 (based on figure 1 in [73]). Ω_0 is a region with every point at least a distance z_0 from the domain boundary $\partial\Omega$.

8.2. Non-uniqueness of the DC problem

For consistency with the normal source definition as give in (3.44), we make the assumption that $\Omega = \Omega_1 \cup \Omega_0$ with $\partial\Omega_0$ strictly contained in the interior of Ω as shown in figure 11, with the closest distance between the boundaries being z_0 .

Suppose that we have two ordered sets of functions (κ, μ_a) and $(\tilde{\kappa}, \tilde{\mu}_a)$ with the equivalent canonical parameters $\hat{\eta}$ and $\tilde{\hat{\eta}}$. We can state two conditions:

- Condition 1: $\tilde{\hat{\eta}} = \hat{\eta}$ everywhere in Ω
- Condition 2: $\tilde{\kappa} = \kappa$ everywhere in Ω_1 .

If these hold, then it follows immediately that both sets will give rise to the same solution U and thus to the same measured data Γ as defined by (4.5).

Consider the DC case ($\omega = 0$). Suppose we add a function β to κ and α to μ_a :

$$\tilde{\kappa} = \kappa + \beta \quad \tilde{\mu}_a = \mu_a + \alpha. \tag{8.1}$$

Condition 1 is now just $\tilde{\eta}_0 = \eta_0$. If condition 2 also holds, then the identical data will be measured if

$$\left(\frac{\nabla^2(\kappa + \beta)^{1/2}}{(\kappa + \beta)^{1/2}} \right) + \frac{\mu_a + \alpha}{(\kappa + \beta)} = \left(\frac{\nabla^2\kappa^{1/2}}{\kappa^{1/2}} \right) + \frac{\mu_a}{\kappa}. \tag{8.2}$$

Thus, for any β satisfying $\beta = 0$ in Ω_1 , we can find α from:

$$\alpha = (\kappa + \beta) \left(\left(\frac{\nabla^2\kappa^{1/2}}{\kappa^{1/2}} \right) + \frac{\mu_a}{\kappa} - \left(\frac{\nabla^2(\kappa + \beta)^{1/2}}{(\kappa + \beta)^{1/2}} \right) \right) - \mu_a. \tag{8.3}$$

Although in reality the bounds on κ, μ_a rule out some possibilities, the above result proves that there exists an infinite set of functions that give rise to the same data. Only if $\omega \neq 0$ and $\hat{\eta}$ can be found uniquely can we state that κ and μ_a can be determined uniquely from

$$\kappa = \frac{\omega}{c\Im[\hat{\eta}]} \quad \mu_a = \kappa \left(\Re[\hat{\eta}] - \left(\frac{\nabla^2\kappa^{1/2}}{\kappa^{1/2}} \right) \right). \tag{8.4}$$

8.3. Non-uniqueness for other measurement types

It would be attractive if the analysis of section 8.2 could be extended to the other measurement types given in section 4.2, but formal results for other cases have not been discovered to date. However, we can gain some insight into their form by explicitly plotting the error surface (2.2) for a simple case. To illustrate this idea we consider the simple problem of a homogeneous circle with parameters $\bar{\mu}_a$ and $\bar{\mu}'_s$ and a single embedded perturbation region. We thus consider

$$I(\mu_a, \mu'_s) = \sum_{i,j} \left(\frac{\mathcal{P}_{j,i}[\bar{\mu}_a, \bar{\mu}'_s] - \mathcal{P}_{j,i}[\mu_a, \mu'_s]}{\sigma_{j,i}} \right)^2. \quad (8.5)$$

The mesh for which the homogeneous reference data $F_{j,i}(\bar{\mu}_a, \bar{\mu}'_s)$ and perturbation data $F_{j,i}(\mu_a, \mu'_s)$ were generated was as figure 5. Data were calculated at 16 equally spaced source positions and 16 detector positions, each placed between two sources and the error maps $I(\bar{\mu}_a, \bar{\mu}'_s)$ were sampled at 41 $\bar{\mu}_a$ values in the range from 0.0125–0.0375 mm⁻¹ (vertical axis) and at 41 $\bar{\mu}'_s$ values in the range from 1–3 mm⁻¹ (horizontal axis).

Figure 12 shows error maps for the measurement types $\log \Gamma$ (real and imaginary), $\log E$, $\langle t \rangle$, c_3 , $L(s = 0.01)$. Characteristic for all types is an elongated ‘valley’, while the orientation of the valley differs between measurement types.

The shape of the error functions suggests that the simultaneous reconstruction of μ_a and μ'_s will be ambiguous for any single measurement type. The reconstruction algorithm will descend into the valley at some point, but due to the small gradient along the valley the convergence towards the true solution will be slow. In order to overcome this problem we have to use multiple measurement types for the reconstruction using a data vector composed of several measurement types, as described in [74]. We require a combination of measurement types for which the sums of the error norms exhibit a clearly defined minimum, to avoid the μ_a/μ'_s ambiguity of the single-measure case. We summarize our position with the following conjecture.

Conjecture 1. *In 2D, reconstruction of two functions of two variables $\mu_a(x, y)$, $\mu_s(x, y)$ requires two maps of two variables $\Lambda_1(\vartheta_p, \vartheta_m)$, $\Lambda_2(\vartheta_p, \vartheta_m)$.*

In 3D, reconstruction of two functions of three variables $\mu_a(x, y, z)$, $\mu_s(x, y, z)$ requires two maps of four variables $\Lambda_1(\vartheta_p, \varphi_p, \vartheta_m, \varphi_m)$, $\Lambda_2(\vartheta_p, \varphi_p, \vartheta_m, \varphi_m)$.

8.4. Singular value analysis of the linearized problem

The uniqueness results discussed above are ‘local’ in the sense of making statements about the ability to separate the recovered properties at the same point in space. Of course we should also address the question of image resolution which we expect to be worse in the interior of the object rather than at the boundary because of the strong smoothing property of the second differential operator in the forward problem. Again, formal results are lacking, but we can illustrate the problem by computing the singular values of the linearized problem (i.e. the Jacobian).

Figure 13 shows singular values for the data types E , $\log E$, $\langle t \rangle$, c_3 and $L(s = 0.01)$, computed for the model problem of figure 5, discretized into a 16×16 pixel basis with 16 equally spaced sources and 16 detectors. The behaviour of the singular values is quite similar, although the fall-off of the absolute intensity type is much worse than for its logarithm, consistent with saying that the Born approximation is much more illposed than the Rytov approximation. Significantly, there is no sharp transition in the singular value spectrum. Therefore, we suppose that recovery of fine resolution image detail is difficult with respect to noise in the data in a predictable manner—we do not suppose that any particular change in

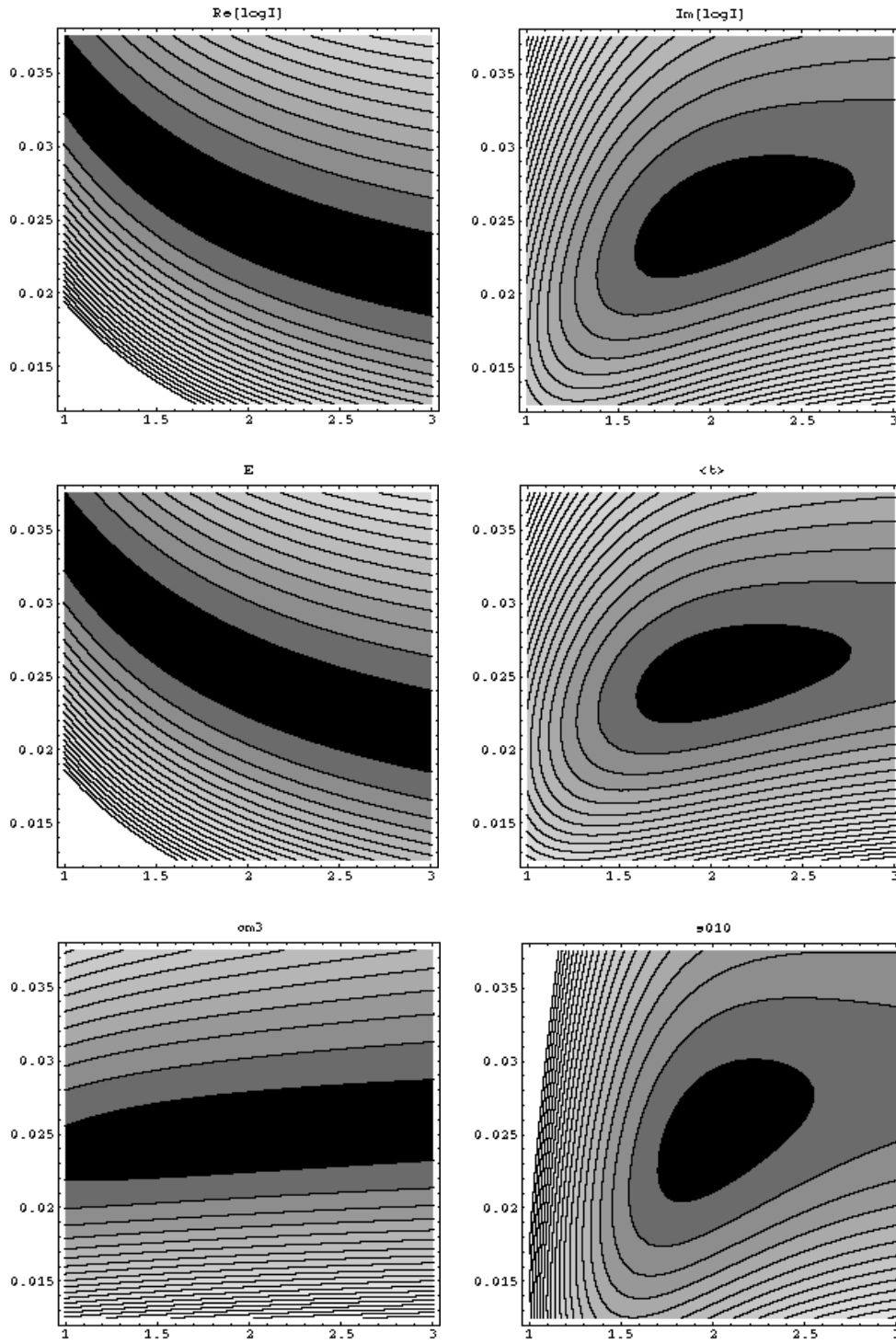


Figure 12. Measurement error norms as a function of global $\bar{\mu}_a$ and $\bar{\mu}'_s$ for different measurement types. Abscissae: $\bar{\mu}'_s$ (mm⁻¹), ordinates: $\bar{\mu}_a$ (mm⁻¹). Top to bottom, left to right, data types are: real and imaginary parts of $\log \Gamma$, $\log E$, $\langle t \rangle$, c_3 , $L(s = 0.01)$.

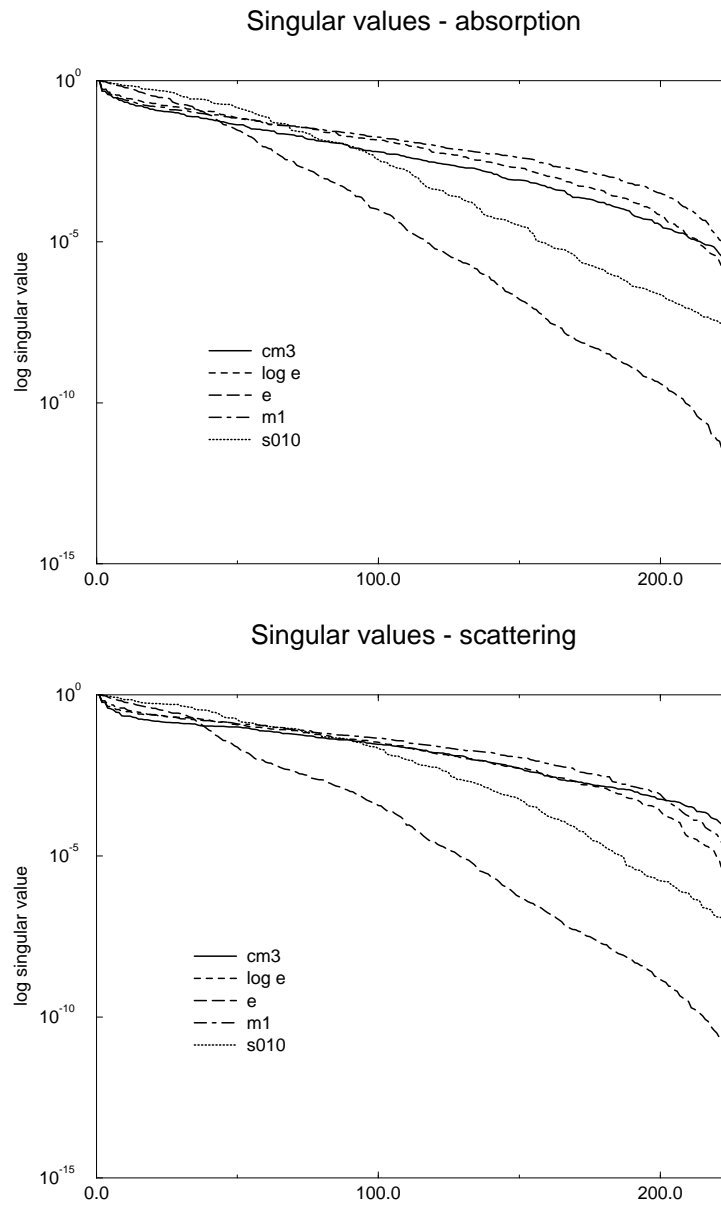


Figure 13. Singular values of the linearized problem for μ_a (top) and μ_s (bottom). All values are normalized to $\lambda_1 = 1$ for each data type. Data types are E , $\log E$, (t) , c_3 and $L(s = 0.01)$.

sampling strategy will have a dramatic difference on image quality. Clearly these results are anecdotal and much more study could be made on this important aspect of the problem.

9. Reconstruction methods

Many medical imaging modalities are governed by a radiation propagation model represented by the Boltzmann equation. For example, if μ_s and q are zero, and the system is assumed to be

steady-state, then (3.1) in the steady state becomes the differential form of the Radon transform for x-ray CT. Similarly if μ_a and μ_s are zero and q is non-zero and isotropic, we obtain the Radon transform for SPECT. If both μ_a and q are non-zero, then the problem becomes the attenuated Radon transform, with the exponential Radon transform representing the case where μ_a is non-zero and constant.

Not surprisingly, many attempts have been made to translate the optical tomography problem into one that has been solved in another field. In particular we find:

- (i) CT-like methods: here it is argued that a Radon transform like formulation can be applied. In this respect both closed form solutions (in analogy to the inverse Radon transform) and backprojection schemes (in analogy to backprojection convolution) methods have been proposed.
- (ii) SPECT-like methods: here it is argued that a linear matrix formulation is appropriate in analogy to the transfer function of SPECT systems.
- (iii) EIT-like methods: here it is argued that the problem is a nonlinear optimization and iterative nonlinear solutions are required.

All the above represent the inversion of an integral equation problem. For completeness we note a separate approach by Klivanov *et al* [75] in which the approach is to derive a differential form for the inverse problem.

9.1. Analytic methods

Sotland [76] obtained an inversion formula for the absorption term for the CW case. The infinite space Green function is given in integral form as

$$G^{(\Phi)}(\mathbf{r}, \mathbf{r}', \omega) = \frac{1}{(2\pi)^3} \int \frac{e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}}{\kappa k^2 + \mu_a + i\omega/c} d^3k. \quad (9.1)$$

Using (5.25) and (4.6) we have

$$\Gamma_{\mu_a, \kappa}^{\delta} \left(\begin{array}{c} \alpha \\ \beta \end{array} \right) (\mathbf{m}, \mathbf{p}, \omega) = -\frac{\zeta}{(2\pi)^6} \int d^3k_m \int d^3k_p \int d^3r' \left[(\alpha(\mathbf{r}') + \mathbf{k}_m \cdot \mathbf{k}_p \beta(\mathbf{r}')) \times \frac{e^{i\mathbf{k}_m \cdot (\mathbf{m}-\mathbf{r}') + i\mathbf{k}_p \cdot (\mathbf{p}-\mathbf{r}')}}{\kappa^2(k_m^2 + \hat{\chi}^2(\omega))(k_p^2 + \hat{\chi}^2(\omega))} \right] \quad (9.2)$$

where

$$\hat{\chi}^2(\omega) = \frac{\mu_a + i\omega/c}{\kappa}.$$

We consider the four-dimensional integral of \mathbf{m}, \mathbf{p} over $\partial\Omega$. If \mathbf{k} is separable then

$$\mathbf{k} \cdot \mathbf{r} = \mathbf{k}_{\parallel} \cdot \mathbf{q} + k_{\perp} v.$$

Sotland's result considered $\partial\Omega$ as the half-space $z = 0$ with $v = z$, which leads to

$$\Lambda_{\mu_a, \kappa}^{\delta} \left(\begin{array}{c} \alpha \\ \beta \end{array} \right) (\mathbf{k}_{m\parallel}, \mathbf{k}_{p\parallel}, \omega) = -\frac{\zeta}{(2\pi)^2} \int dk_{mz} \int dk_{pz} \times \int d^3r' \left[(\alpha(\mathbf{r}') + (\mathbf{k}_{m\parallel} \cdot \mathbf{k}_{p\parallel} + k_{mz}k_{pz})\beta(\mathbf{r}')) \times \frac{e^{i(\mathbf{k}_{m\parallel} + \mathbf{k}_{p\parallel}) \cdot \mathbf{q}' - i(k_{mz} + k_{pz})z'}}{\kappa^2(k_{mz}^2 + k_{m\parallel}^2 + \hat{\chi}^2(\omega))(k_{pz}^2 + k_{p\parallel}^2 + \hat{\chi}^2(\omega))} \right]. \quad (9.3)$$

Note that this approach formally requires that \mathbf{p} is a boundary source rather than an internal one. Completing the two k_z integrals in the upper half-plane contributes one simple pole each with the result

$$\Lambda_{\mu_a, \kappa}^{\delta} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} (\mathbf{k}_{m\parallel}, \mathbf{k}_{p\parallel}, \omega) = -\frac{\zeta}{4\kappa^2} \int d^3\mathbf{r}' \left\{ (\alpha(\mathbf{r}') + s(\mathbf{k}_m, \mathbf{k}_p, \omega)\beta(\mathbf{r}')) \right. \\ \left. \times \frac{e^{i(\mathbf{k}_{m\parallel} + \mathbf{k}_{p\parallel}) \cdot \mathbf{d}' + [(k_{m\parallel}^2 + \hat{\chi}^2(\omega))^{\frac{1}{2}} + (k_{p\parallel}^2 + \hat{\chi}^2(\omega))^{\frac{1}{2}}]z'}}{(k_{m\parallel}^2 + \hat{\chi}^2(\omega))^{\frac{1}{2}} (k_{p\parallel}^2 + \hat{\chi}^2(\omega))^{\frac{1}{2}}} \right\} \\ \text{where } s(\mathbf{k}_m, \mathbf{k}_p, \omega) = \mathbf{k}_{m\parallel} \cdot \mathbf{k}_{p\parallel} + (k_{m\parallel}^2 + \hat{\chi}^2(\omega))^{\frac{1}{2}} (k_{p\parallel}^2 + \hat{\chi}^2(\omega))^{\frac{1}{2}}. \quad (9.4)$$

By change of variables

$$\mathbf{a} = \mathbf{k}_{m\parallel} + \mathbf{k}_{p\parallel} \\ b_m = 2 (k_{m\parallel}^2 + \hat{\chi}^2(\omega))^{\frac{1}{2}} \\ b_p = 2 (k_{p\parallel}^2 + \hat{\chi}^2(\omega))^{\frac{1}{2}}$$

we get

$$\kappa^2 b_m b_p \Lambda_{\mu_a, \kappa}^{\delta} (\alpha) (\mathbf{a}, b_m, b_p, \omega) = -\zeta \int d^3\mathbf{r}' e^{i\mathbf{a} \cdot \mathbf{d}' + \frac{1}{2}[b_m + b_p]z'} \alpha(\mathbf{r}') \quad (9.5)$$

$$\kappa^2 \frac{b_m b_p}{\tilde{s}(\mathbf{a}, b_m, b_p)} \Lambda_{\mu_a, \kappa}^{\delta} (\beta) (\mathbf{a}, b_m, b_p, \omega) = -\zeta \int d^3\mathbf{r}' e^{i\mathbf{a} \cdot \mathbf{d}' + \frac{1}{2}[b_m + b_p]z'} \beta(\mathbf{r}') \quad (9.6)$$

where $\tilde{s}(\mathbf{a}, b_m, b_p, \omega)$ is the function $s(\mathbf{k}_m, \mathbf{k}_p, \omega)$ expressed under the change of variables. Equations (9.5) and (9.6) may be formally inverted via a Fourier–Laplace transform wherein the four spatial dimensions are restricted to a three-dimensional manifold $b = \frac{1}{2}(b_m + b_p)$. At least in principle the recovery of a three-dimensional function from four-dimensional data is stable, but the effect of finite sampling and band-limiting in the discrete case has not been investigated. Furthermore the method is subject to the restriction that second-order effects are negligible.

Another analytic approach using a Fourier transform over the data space was reported by Li *et al* [77]. Here a degree of assumption as to the depth of inhomogeneities was required.

9.2. Backprojection methods

Backprojection schemes were already introduced in section 4.3 where the idea is to formulate a Fredholm integral equation such as (4.39). The backprojection scheme simply applies some operator to the left-hand side of (4.39). The suggestion of Feng *et al* [48] is to form modal lines of the PMDF functions, but a variety of other forms are possible, the main criterion being to span the space of the solution adequately.

The use of backprojection schemes can be made more rigorous using the scheme of Barber [78] where for any backprojection matrix \mathbf{X} the filter $\mathbf{W} = (\mathbf{J}\mathbf{X})^{-1}$ solves a generalized inverse of the underdetermined linear problem [79]. For the particular case $\mathbf{X} = \mathbf{J}^T$ the generalized inverse is the Moore–Penrose one.

9.3. Linear methods

If we have an estimate that is close to the ideal solution, then its projection is close to $\vec{\mathbf{y}}$. We can expand equation (2.1) in a Taylor series

$$\vec{\mathbf{y}} = \vec{\mathbf{y}}_0 + \vec{\mathcal{P}}'[x_0](x - x_0) + (x - x_0)^T \vec{\mathcal{P}}''[x_0](x - x_0) + \dots \quad (9.7)$$

Neglecting terms after the first, linear term constitutes the perturbation approach and the problem reduces to inversion of the matrix representation of the Jacobian

$$\vec{b} = Jh \tag{9.8}$$

where

$$h = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

is the linear update vector. The majority of reported results use this approach, which can be compared to a similar technique developed in EIT [80]. Without exception they require, either explicitly or implicitly, a difference experiment that measures \vec{b} as the difference between two states. Resultant images are usually presented qualitatively. Recently Boas [81] has given an argument for an intrinsic limitation to the quantitative information recoverable in a linearized problem, based on an exact solution of the forward problem for a sphere in an infinite medium.

The difference between different approaches to the linearized problem depends on how the matrix J is constructed. We may distinguish

- (i) Semi-analytic methods: here the functional form of the PMDFs is used assuming an infinite space or infinite half-space geometry wherein the closed form expressions can be written down directly [10].
- (ii) Monte Carlo methods: here a Monte Carlo simulation is run to assemble weight functions which are assumed to be reusable for a variety of cases [82].
- (iii) Numerical PDE methods: here a FDM or FEM solution to the forward problem is used to generate the PMDFs such as (7.12) described in section 7.

We define the Jacobian

$$J = R^{-1} \vec{F}' \tag{9.9}$$

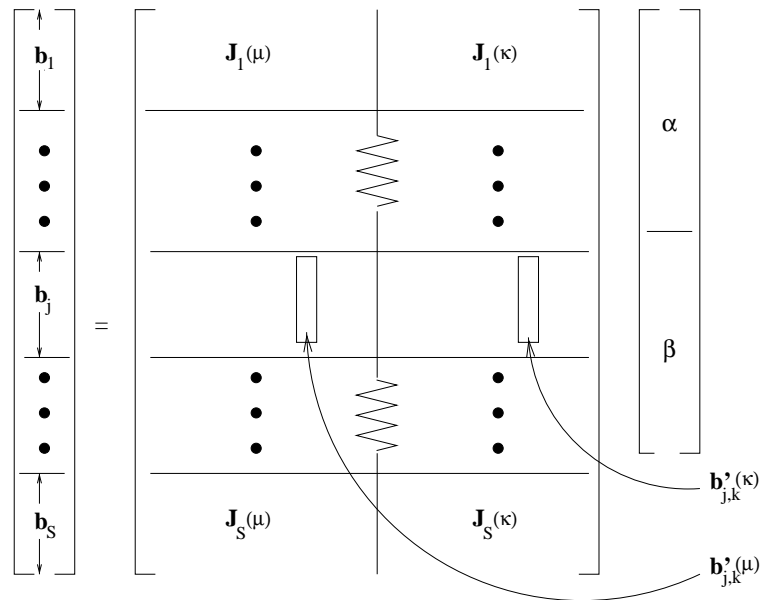


Figure 14. Terms in the Jacobian calculated by the direct method (based on figure 6.1 in [87]).

of the forward problem as a $M_{TOT} \times N_{TOT}$ matrix where $M_{TOT} = \sum M_j$ and $N_{TOT} = N_\mu + N_\kappa$ with the following structure

$$\begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{b}_S \end{pmatrix} = \begin{pmatrix} \mathbf{J}_1(\mu) & \mathbf{J}_1(\kappa) \\ \mathbf{J}_2(\mu) & \mathbf{J}_2(\kappa) \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \mathbf{J}_S(\mu) & \mathbf{J}_S(\kappa) \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (9.10)$$

where we define \mathbf{J}_j as the $M_j \times N_{TOT}$ matrix that is the Jacobian for the sub-objective function I_j and $\mathbf{J}_j(\mu), \mathbf{J}_j(\kappa)$ as the sub-matrices corresponding to the basis coefficients with respect to the two variables μ, κ . Note that we absorbed the covariance matrix \mathbf{R} into both the definition of the residual $\vec{\mathbf{b}}$ and the Jacobian in (9.9).

Solving (7.7) and using (6.10) yields two vectors of length M_j , given by:

$$\begin{aligned} \mathbf{b}'_{j,k}(\mu_a) &= \mathbf{R}_j^{-1} \mathbf{D}_j \frac{\partial \Phi_j}{\partial \mu_k} \\ \mathbf{b}'_{j,k}(\kappa) &= \mathbf{R}_j^{-1} \mathbf{D}_j \frac{\partial \Phi_j}{\partial \kappa_k} \end{aligned} \quad (9.11)$$

which are located into the Jacobian as in figure 14. We can thus state the direct method, algorithm 1, to construct the Jacobian.

Algorithm 1. *Construction of the Jacobian using the direct method.*

```

for all sources  $j = 1 \dots S$ 
  Solve  $(\mathbf{K} + \mathbf{C} + \zeta \mathbf{A} + i\omega \mathbf{B}) \Phi_j = \mathbf{q}_j$ 
  for all basis functions  $v_k^{(x)}$ 
    Solve  $(\mathbf{K} + \mathbf{C} + \zeta \mathbf{A} + i\omega \mathbf{B}) \frac{\partial \Phi_j}{\partial x_k} = -\mathbf{V}_k(x) \Phi_j$ 
    set  $k$ th column of  $\mathbf{J}_j(x)$  to:  $\mathbf{b}'_{j,k}(x) = \mathbf{R}_j^{-1} \mathbf{D}_j \frac{\partial \Phi_j}{\partial x_k}$ 
  end for
end for

```

From (7.12), the adjoint method to build the Jacobian can be stated as in algorithm 2, figure 15. We introduce the notation $\mathbf{j}_{j,i}$ for the i th row of matrix \mathbf{J}_j from which

$$\mathbf{j}_{j,i} = -\sigma_{j,i}^{-1} \boldsymbol{\rho}(i, j) \quad (9.12)$$

where $\boldsymbol{\rho}(i, j)$ is the vector with components $\{\rho_x(i, k, j)\}$, $k = 1, \dots, N$; $x = \mu_a, \kappa$. For convenience we also establish a single-indexing scheme $m = \sum_{j'}^{j-1} M_{j'} + i$ for a row of \mathbf{J} whereby $\boldsymbol{\rho}(m) = \boldsymbol{\rho}(i, j)$ and $\mathbf{j}_m = \mathbf{j}_{j,i}$.

Algorithm 2. *Construction of the Jacobian using the adjoint method.*

```

for all sources  $j = 1 \dots S$ 
  Solve  $(\mathbf{K} + \mathbf{C} + \zeta \mathbf{A} + i\omega \mathbf{B}) \Phi_j = \mathbf{q}_j$ 
end for
for all measurements  $m = 1 \dots M_{UNIQ}$ 
  Solve  $(\mathbf{K} + \mathbf{C} + \zeta \mathbf{A} - i\omega \mathbf{B}) \Psi_m = \mathbf{q}_m^+$ 
end for
for all sources  $j = 1 \dots S$ 
  for all measurements  $i$  from this source and all basis functions  $v_k^{(x)}$ 
    let  $m(i)$  be the index of this measurement in the set of unique measurements
    set  $(i, k)$ th element of  $\mathbf{J}_j(x)$  to:  $J_{j,ik}(x) = \sigma_{j,i}^{-1} \langle \Psi_{m(i)}, \mathbf{V}_k(x) \Phi_j \rangle = -\sigma_{j,i}^{-1} \rho_x(i, k, j)$ 
  end for
end for

```

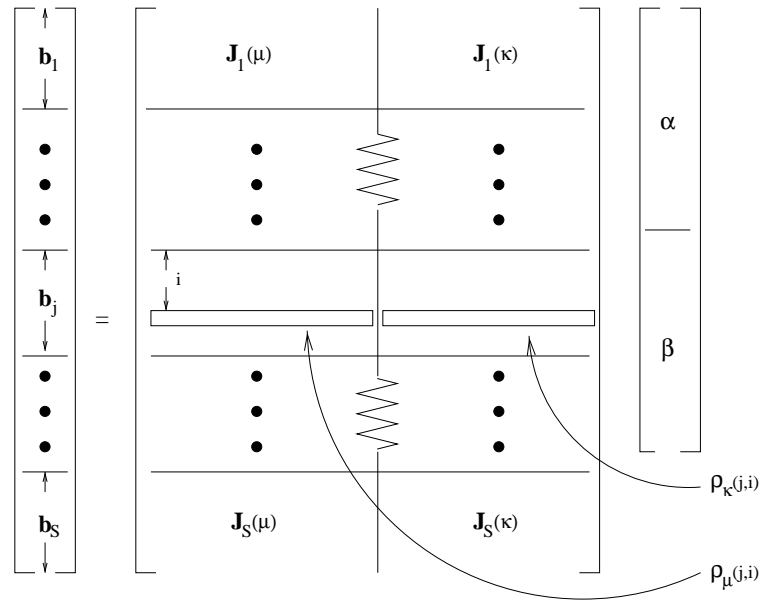


Figure 15. Terms in the Jacobian calculated by the adjoint method (based on figure 6.2 in [87]).

A third method can be used to generate the Jacobian, namely to explicitly perturb the model in a region and subtract the two projected data sets. This can have an advantage if the direct or adjoint methods cannot be expressed directly. It has been reported for the fully time-domain problem [83] and in fluorescence imaging [3].

9.3.1. *Other measurement types.* Applying the discrete measurement operator \mathcal{M}_j yields, in place of (9.11)

$$\mathbf{b}_{j,k}^{\mathcal{M}}(x) = (\mathbf{R}_j^{\mathcal{M}})^{-1} \mathcal{M}_j \left[\mathbf{D}_j \frac{\partial \Phi_j}{\partial x_k} \right] \quad (9.13)$$

with the adjoint form

$$(\sigma_{j,i}^{\mathcal{M}})^{-1} \rho_x^{(\mathcal{M})}(i, k, j) = - (\sigma_{j,i}^{\mathcal{M}})^{-1} \mathcal{M}[\langle \Psi_{m(i)}, \mathbf{V}_k(x) \Phi_j \rangle].$$

9.3.2. *Linear inversion methods.* The linear problem is illposed and large, and so a variety of regularizing inversion methods have been suggested including truncated SVD, conjugate gradients and the Kaczmarz method, commonly known as ART (algebraic reconstruction technique). The latter is the most popular method since it allows efficient storage representation of the huge matrix \mathbf{J} , as well as being easily adaptable to projection onto convex sets (POCS), for example to impose range constraints. Although the Jacobian is dense in the strict sense, the sensitivity functions $\rho(i, j)$ are well-localized so that many matrix elements can be considered to be zero with reasonable accuracy. This leads to the use of sparse matrix representations and solvers. Nevertheless, the matrix inversion is still costly, and subject to numerical error, especially in 3D.

9.3.3. *Symmetry exploitation.* An important simplification can be made if the acquisition scheme possesses cylindrical symmetry. The approach was developed by Metherall [84], and

has been applied to optical tomography [85]. We require also cylindrical symmetry in the solution, both in its sampling, and in the coefficients. Thus the method usually depends on a homogeneous starting point, although in fact a radially varying starting solution could be employed.

If the symmetry is p -fold, then the linear problem can be written

$$\begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_p \end{pmatrix} = \begin{pmatrix} A_1 & A_2 & A_3 & \dots & A_p \\ A_p & A_1 & A_2 & \dots & A_{p-1} \\ A_{p-1} & A_p & A_1 & \dots & A_{p-2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ A_2 & A_3 & A_4 & \dots & A_1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \\ \vdots \\ h_p \end{pmatrix}. \quad (9.14)$$

This matrix can be block-diagonalized by a Fourier transform in the angular variable only

$$\vec{Z}\vec{b} = (ZA Z^*)Zh = \text{diag}(\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \dots, \mathcal{S}_p)\hat{\mathbf{h}} \quad (9.15)$$

then the inversion of $ZA Z^*$ is reduced to the inversion of p block-matrices. An inverse Fourier transform is applied to recover the solution.

9.4. Nonlinear methods

To solve the optimization problem (2.2) we consider the gradient of the objective function

$$\frac{\partial I}{\partial x_k} = \sum_{j=1}^S \sum_{i=1}^{M_j} \left(\frac{y_{j,i} - \mathcal{P}_{j,i}[\mu_a, \kappa]}{\sigma_{j,i}^2} \right) \left(-\frac{\partial \mathcal{P}_{j,i}[\mu_a, \kappa]}{\partial x_k} \right) \quad (9.16)$$

where x is either μ_a or κ , and k indexes the basis functions defined in (7.1). We can further state

$$\frac{\partial I}{\partial x_k} = -\vec{\mathcal{P}}^T \mathbf{R}^{-2}(\vec{\mathbf{y}} - \vec{\mathbf{F}}) \quad (9.17)$$

$$= -\vec{\mathcal{P}}^T \mathbf{R}^{-1} \vec{\mathbf{b}} \triangleq \mathbf{z} \quad (9.18)$$

where \mathbf{z} is a $(N_\mu + N_\kappa)$ vector representing the gradient of the objective function in the chosen basis, and $\mathcal{P}'_j : (\mathbf{X}^{(\mu)}, \mathbf{X}^{(\kappa)}) \rightarrow \mathbf{Y}$ is the Fréchet derivative of \mathcal{P}_j .

Two classes of optimization scheme will be discussed:

- (i) Newton methods. These methods seek a zero of I' by an iterative method, using a Taylor expansion around the current estimate $\mathbf{x}^{(n)}$

$$\begin{aligned} I'(\mathbf{x}^{(n+1)}) &= I'(\mathbf{x}^{(n)}) + I''(\mathbf{x}^{(n)})\mathbf{h}^{(n)} = 0 \\ \Rightarrow \mathbf{h}^{(n)} &= -(I''(\mathbf{x}^{(n)}))^{-1} I'(\mathbf{x}^{(n)}). \end{aligned} \quad (9.19)$$

These methods imply an explicit linearization of the inverse problem and become computationally intractable when the number of dimensions of the inverse search space is large.

- (ii) Gradient methods. These methods require only knowledge of the gradient \mathbf{z} , and are generally preferable for an optimization problem over a large number of dimensions.

9.4.1. Newton methods. The prototypical algorithm for Newton methods is the Levenberg–Marquardt algorithm. In optical tomography this was discussed in some detail in [29, 86].

From (9.19) we require to solve a linear step:

$$\mathbf{h} = -(I''(\mathbf{x}))^{-1} I'(\mathbf{x}). \quad (9.20)$$

Using (9.17) we have

$$\begin{aligned} I'(x) &= -\vec{\mathcal{P}}'^T \mathbf{R}^{-1} \vec{\mathbf{b}} = -\mathbf{J}^T \vec{\mathbf{b}} \\ I''(x) &= \vec{\mathcal{P}}'^T \mathbf{R}^{-2} \vec{\mathcal{P}}' - \mathbf{R}^{-1} \vec{\mathcal{P}}'' \vec{\mathbf{b}} \\ &= \mathbf{J}^T \mathbf{J} - \mathbf{H} \vec{\mathbf{b}}. \end{aligned} \quad (9.21)$$

Therefore we can represent (9.20) as

$$\mathbf{h} = (\mathbf{J}^T \mathbf{J} - \mathbf{H} \vec{\mathbf{b}})^{-1} \mathbf{J}^T \vec{\mathbf{b}}. \quad (9.22)$$

In the Levenberg–Marquardt method the Hessian \mathbf{H} is ignored and a control term is introduced. In the underdetermined case, the number of rows of \mathbf{J} is less than the number of columns, i.e. $M_{TOT} < N_{TOT}$ so that the $N_{TOT} \times N_{TOT}$ matrix $\mathbf{J}^T \mathbf{J}$ is certainly rank-deficient, whereas in the overdetermined case $M_{TOT} > N_{TOT}$ so that $\mathbf{J}^T \mathbf{J}$ may be invertible. In practice even the overdetermined case is usually ill-conditioned so that both cases require stabilization via the Levenberg–Marquardt method:

$$\mathbf{h} = (\mathbf{J}^T \mathbf{J} + \lambda \mathbf{I})^{-1} \mathbf{J}^T \vec{\mathbf{b}}. \quad (9.23)$$

9.4.2. Implicit linear methods. The principal difficulty with the Levenberg–Marquardt algorithm is the explicit formation and inversion of the matrix \mathbf{J} . Iterative linear techniques proceed by repeatedly solving the linear problem using one of the methods of section 9.3.2. Since the linear problem is itself solved iteratively, the optimization can be thought of as occurring in two loops: a linear step which computes an update based on the current state \mathbf{x} , followed by a recomputation of the linearization. A trade-off occurs between the effort expended in the inner iteration to compute the linear step, and the effort expended in computing the linearization. Usually the cost of recomputing \mathbf{J} is thought to dominate, but the use of adjoint methods mean that this cost can be cut to a fraction.

The advantage of ART over other algorithms for the linearized problem, such as truncated SVD [11], arises from the use of the adjoint method for the Jacobian, described in algorithm 2. By precomputing the forward vectors $\{\Phi_j\}$, $j = 1 \dots S$, and the adjoint vectors $\{\Psi_m\}$, $m = 1 \dots M_{UNIQ}$, the vectors $\mathbf{j}_{m'}$ (in ART), or matrices $\mathbf{J}_{j'}$ (in Block-ART) can be generated efficiently on the fly.

We illustrate this method in the Block-ART variation in algorithm 3. For other variations see [87].

Algorithm 3. *Nonlinear Block-ART algorithm.*

```

for  $L$  sweeps
  for  $j = 1 \dots S$ 
     $j'$  = next index from permutation set  $\pi(1, 2, \dots, S)$ 
    Solve  $(\mathbf{K} + \mathbf{C} + \zeta \mathbf{A} + i\omega \mathbf{B}) \Phi_{j'} = \mathbf{q}_{j'}$ 
     $\tilde{\mathbf{b}}_{j'}^M = \mathbf{W}_{j'} (\mathbf{R}_{j'}^M)^{-1} (\mathbf{y}_{j'}^M - \mathbf{F}_{j'}^M)$ 
     $\mathbf{a}_{j'} = \sum_{i=1}^{M_{j'}} \tilde{\mathbf{b}}_{j',i}^M \mathbf{q}_i^+$ 
    Solve  $(\mathbf{K} + \mathbf{C} + \zeta \mathbf{A} - i\omega \mathbf{B}) \Psi_{j'} = \mathbf{a}_{j'}$ 
    for all basis functions  $v_k^{(x)}$ 
      Set  $h_k = \mathcal{M}[\langle \Psi_{j'}, v_k^{(x)} \Phi_{j'} \rangle]$ 
    end for
     $\mathbf{x} = \mathbf{x} + \lambda \mathbf{h}$ 
  end for
end for

```

We note the following:

- (i) Row-ordering: a well known improvement to ART uses non-sequential ordering of rows (see for example [88]).
- (ii) Parallel projection: in the standard Block-ART algorithm to solve a linear problem [89], $W_{j'}$ is a diagonal matrix whose entries are the norms of the rows of $J_{j'}$

$$W_{j',ii} = \|\mathbf{j}_{m''}\|^{-2} \quad m'' = \sum_{j''}^{j'-1} M_{j''} + i. \quad (9.24)$$

- (iii) Minimum norm projection: if instead we use

$$W_{j'} = (J_{j'} J_{j'}^T + \lambda I)^{-1} \quad (9.25)$$

then we are implementing the (regularized) underdetermined version of the Moore–Penrose generalized inverse [90].

- (iv) Block steepest descent: if we use $W = I$, the identity matrix, then each nonlinear step is a steepest descent direction in the I_j defined in (2.2). This method has been used for ultrasound tomography [91].

A variety of flavours of ART are possible, according to how much effort is put into the construction of $W_{j'}$, that range from a steepest-descent step to a (underdetermined) Newton step. The cost increases if $W_{j'}$ is recomputed, but the convergence may be faster. A sensible choice might be to compute $W_{j'}$ from (9.25) once, and store these for each subsequent iteration since they are small.

Finally, note that the use of all data from one source, premultiplied by a matrix and projected into the solution via an adjoint operator is similar to a filtered backprojection algorithm, which is the method of choice for x-ray CT, and can be contrasted with the backprojection convolution approach of section 9.2.

9.4.3. Gradient methods. Consider the continuous case. Let Ψ be a solution to the adjoint problem

$$-\nabla \cdot \kappa(\mathbf{r}) \nabla \Psi(\mathbf{r}, \omega) + \mu_a(\mathbf{r}) \Psi(\mathbf{r}, \omega) - \frac{i\omega}{c} \Psi(\mathbf{r}, \omega) = 0, \quad (9.26a)$$

$$\Psi(\mathbf{m}, \omega) + 2A\kappa(\mathbf{m}) \frac{\partial \Psi(\mathbf{m}, \omega)}{\partial \nu} = b(\mathbf{m}, \omega), \quad (9.26b)$$

then by the argument leading to (5.19)

$$\int_{\partial\Omega} b^* \Gamma_{\mu_a, \kappa}^\delta \begin{pmatrix} \alpha \\ \beta \end{pmatrix} d^{n-1} \hat{\nu}_\perp = - \int_{\Omega} \beta \nabla \Psi^* \cdot \nabla \Phi + \alpha \Psi^* \Phi d^n \mathbf{r}. \quad (9.27)$$

Thus the right-hand side of (9.27) is an update direction that corrects for the discrepancy in the data for one source. In the discrete case, the left-hand side of (9.27) is simply a matrix transposition and so the right-hand side can be seen to be the gradient \mathbf{z} of the objective function. This can be in principle be calculated by explicitly calculating the Jacobian using either of algorithms 1 or 2 and forming

$$\mathbf{z} = - \mathbf{J}^T \vec{\mathbf{b}}. \quad (9.28)$$

However, an alternative is to generate it directly using the adjoint scheme. The use of an adjoint scheme to calculate \mathbf{z} has been discussed previously for the time-domain diffusion model using a finite difference scheme [92], the time-domain transport model using a finite

difference scheme [45, 93], and the steady-state DC case using a finite element scheme [94]. From (7.12) we may write

$$z_k(x) = - \sum_{j=1}^S \sum_{i=1}^{M_j} \frac{b_{j,i}}{\sigma_{j,i}} \langle \Psi_{m(i)}, \mathbf{V}_k(x) \Phi_j \rangle \quad (9.29)$$

where Ψ_m is the solution to

$$(\mathbf{K} + \mathbf{C} + \zeta \mathbf{A} - i\omega \mathbf{B}) \Psi_m = \mathbf{q}_m^+ \quad (9.30)$$

so we form a vector

$$\mathbf{a}_j = \sum_{i=1}^{M_j} \frac{b_{j,i}}{\sigma_{j,i}} \mathbf{q}_{m(i)}^+ \quad (9.31)$$

and solve

$$(\mathbf{K} + \mathbf{C} + \zeta \mathbf{A} - i\omega \mathbf{B}) \Psi_j = \mathbf{a}_j \quad (9.32)$$

leading to

$$z_k(x) = - \sum_{j=1}^S \langle \Psi_j, \mathbf{V}_k(x) \Phi_j \rangle. \quad (9.33)$$

The normalized measurement types have to be handled in a more complex way. From (7.18) we have

$$\mathbf{J}_{j_i}^{\bar{T}} = \frac{1}{F_{j,i}^{\mathcal{E}}} (\mathbf{J}_{j_i}^T - F_{j,i}^{\bar{T}} \mathbf{J}_{j_i}^{\mathcal{E}}) \quad (9.34)$$

which leads to

$$\mathbf{J}^{\bar{T}T} \bar{\mathbf{b}}^{\bar{T}} = \sum_{j=1}^S \sum_i^{M_j} \mathbf{J}_{j_i}^{\bar{T}} b_{j,i}^{\bar{T}} \quad (9.35)$$

$$= \sum_{j=1}^S \sum_i^{M_j} \frac{1}{F_{j,i}^{\mathcal{E}}} (\mathbf{J}_{j_i}^T - F_{j,i}^{\bar{T}} \mathbf{J}_{j_i}^{\mathcal{E}}) b_{j,i}^{\bar{T}} \quad (9.36)$$

$$= \sum_{j=1}^S \sum_i^{M_j} \frac{b_{j,i}^{\bar{T}}}{\sigma_{j,i}^{\bar{T}} F_{j,i}^{\mathcal{E}}} (\mathcal{T}[\langle \Psi_{m(i)}, \mathbf{V}_k(x) \Phi_j \rangle] - F_{j,i}^{\bar{T}} \langle \Psi_{m(i)}, \mathbf{V}_k(x) \Phi_j \rangle). \quad (9.37)$$

We therefore need to construct two adjoint sources

$$\mathbf{a}_j^{\bar{T}}(0) = \sum_{i=1}^{M_j} \frac{b_{j,i}^{\bar{T}} F_{j,i}^{\bar{T}}}{\sigma_{j,i}^{\bar{T}} F_{j,i}^{\mathcal{E}}} \mathbf{q}_{m(i)}^+ \quad (9.38)$$

$$\mathbf{a}_j^{\bar{T}}(1) = \sum_{i=1}^{M_j} \frac{b_{j,i}^{\bar{T}}}{\sigma_{j,i}^{\bar{T}} F_{j,i}^{\mathcal{E}}} \mathbf{q}_{m(i)}^+. \quad (9.39)$$

From here we proceed as before, finding two solutions $\Psi_j^{\bar{T}}(0)$, $\Psi_j^{\bar{T}}(1)$ and computing the gradient as

$$z_k(x)^{\bar{T}} = - \sum_{j=1}^S \mathcal{T}[\langle \Psi_j^{\bar{T}}(1), \mathbf{V}_k(x) \Phi_j \rangle] - \langle \Psi_j^{\bar{T}}(0), \mathbf{V}_k(x) \Phi_j \rangle. \quad (9.40)$$

This method is applicable for measurement types (4.34) and (4.36). The n th central moment given by (4.35) can be constructed as a linear combination of the first n temporal moments and so the gradient requires n adjoint sources, constructed in a straightforward manner.

Whenever the gradient is available, gradient based algorithms such as nonlinear conjugate gradients can be used to produce fast nonlinear optimization schemes. A comparison of this method with Block-ART and steepest-descent was reported in [73].

9.5. Regularization and prior knowledge

We developed our discussion in terms of a maximum likelihood approach, but as in any illposed inverse problem, the instability of the solution with respect to small variations in the data suggests the need for regularization. Regularization consists of the modification of the objective function to a form such as

$$I + \tau \Upsilon(x) \quad (9.41)$$

where $\Upsilon(x)$ is a function of the solution only, and represents *a priori* information, and τ is a *hyper-parameter* whose value is often hard to determine. Under these circumstances the Levenberg–Marquardt method (9.23) is modified to

$$\mathbf{h} = (\mathbf{J}^T \mathbf{J} + \tau \Upsilon'' + \lambda \mathbf{I})^{-1} (\mathbf{J}^T \vec{\mathbf{b}} - \tau \Upsilon') \quad (9.42)$$

and the gradient in conjugate-gradient methods is modified to

$$\mathbf{z} + \tau \Upsilon'. \quad (9.43)$$

These modifications are independent of whichever data type(s) are being used, and so add constant computational effort to the inverse problem. Regularization using quadratic constraints on the solution derivatives was discussed in [95], and the use of an $L1$ -norm (total variation) was discussed in [96]. Linear reconstruction schemes are limited to quadratic regularization terms, which may be a limiting factor in their use.

9.6. Restriction to real solutions

It should be noted that in the complex case the operators such as (5.12) are $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}$, so the resulting PMDF functions and Jacobians are complex, as are the updates specified by any of the reconstruction schemes in the previous sections. All authors make the natural choice of restricting the update to its real part, which is physically correct. If instead the complex Helmholtz form of the problem (3.48) is solved then the commensurate operator is $\mathbb{C} \rightarrow \mathbb{C}$ and no such restriction is required. This may lead to a better reconstruction scheme since some component of the data is being ignored in the restriction case.

10. Conclusions

We have discussed optical tomography as the problem of recovery of absorption and scattering coefficients in the interior of a domain given boundary data. By far the most prevalent approach is the diffusion model which casts the problem into the form of recovery of coefficients of an elliptic partial differential equation. Although much of the theory of such problems is well known there are some open questions which can still be addressed:

- The ability to localize changes in one or other coefficient appears to be achievable although there are vastly varying claims as to image resolution and reconstruction times. On the other hand the ability to differentiate between two coefficients appears to be very difficult.
- Since the diffusion approximation is not completely accurate, it needs to be investigated to what degree uniqueness and stability of the inverse problem could be improved using a higher order approximation to the Boltzmann equation. From a practical point of view the fact that the diffusion approximation is nearly accurate suggests that the improvement will be minor, but formal results are missing.
- As a corollary to the above point there is scope to investigate the properties of the inverse problem when data is incomplete, i.e. if we restrict ourselves to regions of $\partial\Omega$ where the diffusion approximation is valid to within some tolerance limit.

- The nonlinearity and illposedness of the inverse problem is transformed by the use of different measurement operators. A theoretical understanding as to the optimal measurement types to use is lacking.

Optical tomography is a complex and fast-moving area. Our intention here was to present an overview of the methods developed for the inverse problem. It is hoped that this will stimulate further interest from theoreticians in developing deeper understanding of the several difficult problems that are being addressed in this subject.

Further information

The software suite developed at UCL by the author and his colleagues for solving the forward and inverse problems in optical tomography using finite elements is available from the webpage: <http://www.medphys.ucl.ac.uk/toast/index.htm>.

Acknowledgments

The author has benefited from long-standing collaboration with physicists and clinicians at University College Hospital. In particular Professor David Delpy, Dr Mark Cope and Dr Jem Hebden are amongst the pioneers in the field. I am especially indebted to Dr Martin Schweiger for being the principal architect of the FEM and other computational methods that have been hugely influential. In preparing this article the author benefited from access to excellent PhD theses by Dr David Boas and Dr Oliver Dorn amongst others. The article benefited from helpful comments given by Dr Oliver Dorn, Mr Ville Kohlemainen, Thomas Dierkes and others. In particular the author acknowledges a discussion with Thomas Dierkes concerning the restriction to real space mentioned in section 9.6. I am very grateful for the considerable long-standing encouragement and support from Professor Frank Natterer, and to many discussions over the years with Bill Lionheart, Roy Pike, Mike Patterson, Alberto Grünbaum, Ken Hanson, and Harry Barrett. Funding has been generously received from the Wellcome Trust, Action Research, the CEC and the EPSRC.

Appendix. Properties of spherical harmonics

The spherical harmonics are given by:

$$Y_{l,m}(\hat{\mathbf{s}}) = \left(\left(\frac{2l+1}{4\pi} \right) \frac{(l-|m|)!}{(l+|m|)!} \right)^{1/2} (-1)^{\frac{1}{2}(m+|m|)} P_l^{|m|}(\cos \vartheta) e^{im\varphi} \quad (\text{A.1})$$

where $P_l^m(x)$ is the associated Legendre polynomial (see for example [97]). We make use of special properties of these functions:

Orthogonality:

$$\int_{S^{n-1}} Y_{l,m}^*(\hat{\mathbf{s}}) Y_{l',m'}(\hat{\mathbf{s}}) d\hat{\mathbf{s}} = \delta_{ll',mm'}. \quad (\text{A.2})$$

Addition theorem:

$$P_l(\hat{\mathbf{a}} \cdot \hat{\mathbf{b}}) = \frac{4\pi}{2l+1} \sum_{m=-l}^l Y_{l,m}^*(\hat{\mathbf{a}}) Y_{l,m}(\hat{\mathbf{b}}). \quad (\text{A.3})$$

Unit vector expansion:

$$\hat{\mathbf{s}} = \begin{pmatrix} s_x \\ s_y \\ s_z \end{pmatrix} = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix} = \sqrt{\frac{4\pi}{3}} \begin{pmatrix} \frac{1}{\sqrt{2}} (Y_{1,-1}(\hat{\mathbf{s}}) - Y_{1,1}(\hat{\mathbf{s}})) \\ \frac{i}{\sqrt{2}} (Y_{1,-1}(\hat{\mathbf{s}}) + Y_{1,1}(\hat{\mathbf{s}})) \\ Y_{1,0}(\hat{\mathbf{s}}) \end{pmatrix}. \quad (\text{A.4})$$

Recurrence relations:

$$\cos \vartheta Y_{l,m} = \left(\frac{(l+m)(l-m)}{(2l+1)(2l-1)} \right)^{\frac{1}{2}} Y_{l-1,m} + \left(\frac{(l+m+1)(l-m+1)}{(2l+1)(2l+3)} \right)^{\frac{1}{2}} Y_{l+1,m} \quad (\text{A.5})$$

$$\sin \vartheta e^{i\varphi} Y_{l,m} = \left(\frac{(l-m)(l-m-1)}{(2l+1)(2l-1)} \right)^{\frac{1}{2}} Y_{l-1,m+1} - \left(\frac{(l+m+1)(l+m+2)}{(2l+1)(2l+3)} \right)^{\frac{1}{2}} Y_{l+1,m+1} \quad (\text{A.6})$$

$$\sin \vartheta e^{-i\varphi} Y_{l,m} = - \left(\frac{(l+m)(l+m-1)}{(2l+1)(2l-1)} \right)^{\frac{1}{2}} Y_{l-1,m-1} + \left(\frac{(l-m+1)(l-m+2)}{(2l+1)(2l+3)} \right)^{\frac{1}{2}} Y_{l+1,m-1}. \quad (\text{A.7})$$

References

- [1] Hesselink L 1989 Optical tomography *Handbook of Flow Visualisation* ed W-J Yang (New York: Hemisphere) ch 20
- [2] Piersen R E, Chen E Y, Bishop K P and McMackin L 1996 Modeling and measurement of optical turbulence by tomographic imaging of a heated air flow *Proc. SPIE* **2827** 130–41
- [3] Paithankar D Y, Chen A U, Pogue B W, Patterson M S and Sevick-Muraca E M 1997 Imaging of fluorescent yield and lifetime from multiply scattered light reemitted from random media *Appl. Opt.* **36** 2260–72
- [4] Hebden J C, Arridge S R and Delpy D T 1997 Optical imaging in medicine: I. Experimental techniques *Phys. Med. Biol.* **42** 825–40
- [5] Arridge S R and Hebden J C 1997 Optical imaging in medicine: II. Modelling and reconstruction *Phys. Med. Biol.* **42** 841–53
- [6] Chance B, Delpy D T, Cooper C E and Reynolds E O R (ed) 1997 Near-infrared spectroscopy and imaging of living systems *Phil. Trans. R. Soc. B* **352**
- [7] van der Zee P 1993 Measurement and modelling of the optical properties of human tissue in the near infrared *PhD Thesis* University of London
- [8] O'Leary M A 1996 Imaging with diffuse photon density waves *PhD Thesis* University of Pennsylvania
- [9] Tromberg B J, Coquoz O, Fishkin J B, Pham T, Anderson E R, Butler J, Cahn M, Gross J D, Venugopalan V and Pham D Non-invasive measurements of breast tissue optical properties using frequency domain migration *Phil. Trans. R. Soc. B* **352** 661–8
- [10] O'Leary M A, Boas D A, Chance B and Yodh A G 1995 Experimental images of heterogeneous turbid media by frequency-domain diffusing-photon tomography *Opt. Lett.* **20** 426–8
- [11] Arridge S R, van der Zee P, Delpy D T and Cope M 1991 Reconstruction methods for infra-red absorption imaging *Time-Resolved Spectroscopy and Imaging of Tissues (Proc. SPIE 1431)* ed B Chance and A Katzir pp 204–15
- [12] Schweiger M 1994 Application of the finite element method in infrared image reconstruction of scattering media *PhD Thesis* University of London
- [13] Arridge S R and Schweiger M 1995 Sensitivity to prior knowledge in optical tomographic reconstruction *Optical Tomography, Photon Migration, and Spectroscopy of Tissue and Model Media: Theory, Human Studies, and Instrumentation (Proc. SPIE 2389)* ed B Chance and R R Alfano, pp 378–88
- [14] Walker S A, Fantini S and Gratton E 1997 Image reconstruction by backprojection from frequency-domain optical measurements in highly scattering media *Appl. Opt.* **36** 170–9
- [15] Pogue B W, Patterson M S, Jiang H and Paulsen K D 1995 Initial assessment of a simple system for frequency domain diffuse optical tomography *Phys. Med. Biol.* **40** 1709–29
- [16] Jiang H, Paulsen K D, Osterberg U L, Pogue B W and Patterson M S 1995 Optical image reconstruction using frequency-domain data: Simulations and experiments *J. Opt. Soc. Am. A* **13** 253–66
- [17] Jiang H, Paulsen K D, Osterberg U L and Patterson M S 1997 Frequency domain optical image reconstruction in turbid media: An experimental study of single target detectability *Appl. Opt.* **36** 52–63
- [18] Jiang H, Paulsen K D, Österberg U L and Patterson M S 1998 Frequency-domain near-infrared photo diffusion imaging: Initial evaluation in multitarget tissuelike phantoms *Med. Phys.* **25** 183–93
- [19] Ueda Y, Ohta K, Oda M, Miwa M, Yamashita Y and Tsuchiya Y 1995 Optical imaging reconstruction using the average value as the reference *Optical Tomography and Spectroscopy of Tissue: Theory, Instrumentation, Model, and Human Studies II (Proc. SPIE 2979)* ed B Chance and R R Alfano, pp 795–806
- [20] Isakov V 1998 *Inverse Problems in Partial Differential Equations* (New York: Springer)

- [21] Yamada Y 1995 Diffusion coefficient in the photon diffusion equation *Optical Tomography, Photon Migration, and Spectroscopy of Tissue and Model Media: Theory, Human Studies, and Instrumentation (Proc. SPIE 2389)* ed B Chance and R R Alfano, pp 87–97
- [22] Ishimaru A 1978 *Wave Propagation and Scattering in Random Media* vol 1 (New York: Academic)
- [23] Groenhuis R A J, Ferwerda H A and Ten Bosch J J 1983 Scattering and absorption of turbid materials determined from reflection measurements (parts 1 and 2) *Appl. Opt.* **22** 2456–67
- [24] Case M C and Zweifel P F 1967 *Linear Transport Theory* (New York: Addison-Wesley)
- [25] Aronson R 1993 Extrapolation distance for diffusion of light *Photon Migration and Imaging in Random Media and Tissues (Proc. SPIE 1888)* ed B Chance and R R Alfano, pp 297–305
- [26] Moulton J D 1990 Diffusion modelling of picosecond laser pulse propagation in turbid media *MEng Thesis* McMaster University, Hamilton, Ontario
- [27] Schweiger M, Arridge S R, Hiraoka M and Delpy D T 1995 The finite element model for the propagation of light in scattering media: Boundary and source conditions *Med. Phys.* **22** 1779–92
- [28] Eason G, Veitch A R, Nisbet R M and Turnbull F W 1978 The theory of back-scattering of light by blood *J. Phys. D: Appl. Phys.* **11** 1463–79
- [29] Paulsen K D and Jiang H 1995 Spatially-varying optical property reconstruction using a finite element diffusion equation approximation *Med. Phys.* **22** 691–701
- [30] Davison B 1957 *Neutron Transport Theory* (Oxford: Oxford University Press)
- [31] Boas D A 1996 Diffuse photon probes of structural and dynamical properties of turbid media: theory and biomedical applications *PhD Thesis* University of Pennsylvania
- [32] Arridge S R 1989 A note on the spherical harmonic expansion of the Mie scattering kernel *J. Mod. Opt.* **36** 685–92
- [33] Henyey L G and Greenstein J L 1941 Diffuse radiation in the galaxy *AstroPhys. J.* **93** 70–83
- [34] Fishkin J B, Gratton E, van de Ven M J and Mantulin W W 1991 Diffusion of intensity modulated near infrared light in turbid media *Time-Resolved Spectroscopy and Imaging of Tissues (Proc. SPIE 1431)* ed B Chance and A Katzir, pp122–35
- [35] Kaltenbach J P and Kaschke M 1993 Frequency- and time-domain modelling of light transport in random media *Medical Optical Tomography: Functional Imaging and Monitoring* ed G Muller *et al* (Bellingham, WA: SPIE) pp 65–86
- [36] Anikonov D S 1998 Tomography through the transport equation *Computational Radiology and Imaging: Therapy and Diagnosis (IMA Volumes in Mathematics and its Applications 110)* ed C Borgers and F Natterer (Berlin: Springer) in press
- [37] Hielscher A H, Alcouffe R E and Barbour R L 1995 Transport and diffusion calculations on MRI-generated data *Optical Tomography, Photon Migration, and Spectroscopy of Tissue and Model Media: Theory, Human Studies, and Instrumentation (Proc. SPIE 2389)* ed B Chance and R R Alfano, pp 500–8
- [38] Firbank M, Arridge S R, Schweiger M and Delpy D T 1996 An investigation of light transport through scattering bodies with non-scattering regions *Phys. Med. Biol.* **41** 767–83
- [39] Okada E, Firbank M, Schweiger M, Arridge S R, Cope M and Delpy D T 1997 Theoretical and experimental investigation of near-infrared light propagation in a model of the adult head *Appl. Opt.* **36** 21–31
- [40] Colak S B, Papaioannou D G, 't Hooft G W, van der Mark M B, Schomberg H, Paasschens J C J, Melissen J B M and van Asten N A A J 1997 Tomographic image reconstruction from optical projections in light-diffusing media *Appl. Opt.* **36** 180–213
- [41] Benaron D A, Ho D C, Spilman S D, Van Houten J P and Stevenson D K 1994 Non-recursive linear algorithms for optical imaging in diffusive media *Oxygen Transport to Tissue (Adv. Exp. Med. Biol. XVI)* (New York: Plenum) pp 609–17
- [42] Schweiger M and Arridge S R 1997 Optimal data types in optical tomography *Information Processing in Medical Imaging (IPMI'97 Proceedings) (Lecture Notes in Computer Science 1230)* (Berlin: Springer) pp 71–84
- [43] Grosenick D, Wabnitz H and Rinneberg H 1997 Time-resolved imaging of solid phantoms for optical mammography *Appl. Opt.* **36** 221–31
- [44] Arridge S R 1995 Photon measurement density functions. Part 1: Analytical forms *Appl. Opt.* **34** 7395–409
- [45] Dorn O 1997 Das inverse transportproblem in der lasertomographie *PhD Thesis* University of Münster
- [46] Arridge S R and Schweiger M 1995 Photon measurement density functions. Part 2: Finite element calculations *Appl. Opt.* **34** 8026–37
- [47] Schotland J C, Haselgrove J C and Leigh J S 1993 Photon hitting density *Appl. Opt.* **32** 448–53
- [48] Feng S, Zeng Z-A and Chance B 1995 Photon migration in the presence of a single defect: a perturbation analysis *Appl. Opt.* **34** 3826–37
- [49] Sevick E M, Burch C L, Frisoli J K and Lakowicz J R 1994 Localization of absorbers in scattering media by

- use of frequency domain measurements of time-dependent photon migration *Appl. Opt.* **33** 3562–71
- [50] Chang J, Aronson R, Graber H L and Barbour R L 1995 Imaging diffusive media using time-independent and time-harmonic sources: Dependence of image quality on imaging algorithms, target volume, weight matrix, and view angles *Optical Tomography, Photon Migration, and Spectroscopy of Tissue and Model Media: Theory, Human Studies, and Instrumentation* (Proc. SPIE 2389) ed B Chance and R R Alfano, pp 448–64
- [51] Arridge S R, Schweiger M and Delpy D T 1992 Iterative reconstruction of near-infrared absorption images. ed M A Fiddy *Inverse Problems in Scattering and Imaging* (Proc. SPIE 1767) pp 372–83
- [52] Arridge S R, Schweiger M, Hiraoka M and Delpy D T 1993 A finite element approach for modeling photon transport in tissue *Med. Phys.* **20** 299–309
- [53] Schweiger M and Arridge S R 1997 The finite element model for the propagation of light in scattering media: Frequency domain case *Med. Phys.* **24** 895–902
- [54] Schweiger M and Arridge S R 1998 Comparison of 2D and 3D reconstruction algorithms in optical tomography *Appl. Opt.* **37** 7419–28
- [55] Arridge S R and Schweiger M 1995 Direct calculation of the moments of the distribution of photon time of flight in tissue with a finite-element method *Appl. Opt.* **34** 2683–7
- [56] Hackbusch W 1980 *Multigrid Methods and Applications* (Berlin: Springer)
- [57] Duderstadt J J and Hamilton L J 1976 *Nuclear Reactor Analysis* (New York: Wiley)
- [58] Wilson B C and Adam G 1983 A Monte Carlo model for the absorption and flux distribution of light in tissue *Med. Phys.* **10** 824–30
- [59] Hiraoka M, Arridge S R and Delpy D T 1994 A new Monte Carlo algorithm for the description of light transport in tissue *Proc. 7th Int. Workshop on Multiple Scattering Lidar Experiments (MUSCLE7) (21st–23rd July, Chiba University, Japan)* pp 74–7
- [60] Arridge S R, Hiraoka M and Schweiger M 1995 Statistical basis for the determination of optical pathlength in tissue *Phys. Med. Biol.* **40** 1539–58
- [61] Barbour R L, Graber H L, Lubowsky J and Aronson R 1990 Model for 3-D optical imaging of tissue *10th Annual IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS)* vol II, ed J Ormsby, pp 1395–9
- [62] de Oliveira C R E 1986 An arbitrary geometry finite element method for multigroup neutron transport with anisotropic scattering *Prog. Nucl. Energy* **18** 227–36
- [63] Bonner R F, Nossal R, Havlin S and Weiss G H 1987 Model for photon migration in turbid biological media. *J. Opt. Soc. Am. A* **4** 423–32
- [64] Gandjbakhche A H and Weiss G H 1995 Random walk and diffusion-like models of photon migration in turbid media *Progress in Optics* vol 34, ed E Wolf (Amsterdam) pp 333–401
- [65] Patch S K 1994 Recursive recovery of Markov transition probabilities from boundary value data *PhD Thesis* University of Berkeley, San Francisco
- [66] Grünbaum F A 1992 Diffuse tomography: the isotropic case *Inverse Problems* **8** 409–19
- [67] Grünbaum F A and Zubelli J P 1992 Diffuse tomography: Computational aspects of the isotropic case *Inverse Problems* **8** 421–33
- [68] Choulli M and Stefanov P 1996 Inverse scattering and inverse boundary value problems for the linear Boltzmann equation *Commun. Part. Diff. Eq.* **21** 763–85
- [69] Anikonov D S 1984 Uniqueness of simultaneous determination of two coefficients of the transport equation *Sov. Math. Dokl.* **30** 149–51
- [70] Anikonov D S 1985 Uniqueness of the determination of a coefficient of the transport equation for a special type of source *Sov. Math. Dokl.* **32** 511–15
- [71] Sylvester J and Uhlmann G 1987 A global uniqueness theorem for an inverse boundary value problem *Ann. Math.* **125** 153–69
- [72] Uhlman G Private communication
- [73] Arridge S R and Lionheart W R B 1998 Non-uniqueness in diffusion-based optical tomography *Opt. Lett.* **23** 882–4
- [74] Arridge S R and Schweiger M 1993 The use of multiple data types in time-resolved optical absorption and scattering tomography (TOAST) ed J N Wilson and D C Wilson, *Mathematical Methods in Medical Imaging* (Proc. SPIE 2035) vol II, pp 218–29
- [75] Klibanov M V, Lucas T R and Frank R M 1997 A fast and accurate imaging algorithm in optical diffusion tomography *Inverse Problems* **13** 1341–61
- [76] Schotland J C 1997 Continuous-wave diffusion imaging *J. Opt. Soc. Am. A* **14** 275–9
- [77] Li X D, Pattanayak N, Chance B and Yodh A G 1997 K-space approach to biomedical imaging with diffusive photon density waves *Proc. 1997 Conf. on Lasers and Electro-Optics* vol 11 (Lasers and Electro-Optics Society) p 39
- [78] Barber D C and Seagar A D 1987 Fast reconstruction of resistive images *Clin. Phys. P. A* **8** 47–54

- [79] Santosa F and Vogelius M 1990 A backprojection algorithm for electrical impedance imaging *SIAM J. Appl. Math.* **50** 216–43
- [80] Cheney M, Isaacson D, Newell J C, Simske S and Goble J 1990 NOSER: An algorithm for solving the inverse conductivity problem *Int. J. Imag. Syst. Technol.* **2** 66–75
- [81] Boas D A 1997 A fundamental limitation of linearized algorithms for diffuse optical tomography *Opt. Express* **1** 404–13
- [82] Graber H L, Chang J, Lubowsky J, Aronson R and Barbour R L 1993 Near infrared absorption imaging of dense scattering media by steady-state diffusion tomography *Photon Migration and Imaging in Random Media and Tissues (Proc. SPIE 1888)* ed B Chance and R R Alfano, pp 372–86
- [83] Model R, Orlt M, Walzel M and Hünlich R 1997 Reconstruction algorithm for near-infrared imaging in turbid media by means of time-domain data *Appl. Opt.* **14** 313–24
- [84] Metherall P, Barber D C, Smallwood R H and Brown B H 1996 Three-dimensional electrical impedance tomography. *Nature* **380** 509–12
- [85] Hampel U and Freyer R 1998 Fast image reconstruction for optical tomography in media with radially symmetric boundaries *Med. Phys.* **25** 92–101
- [86] Schweiger M, Arridge S R and Delpy D T 1993 Application of the finite-element method for the forward and inverse models in optical tomography *J. Math. Imag. Vision* **3** 263–83
- [87] Arridge S R and Schweiger M 1998 A general framework for iterative reconstruction algorithms in optical tomography, using a finite element method *Computational Radiology and Imaging: Therapy and Diagnosis (IMA Volumes in Mathematics and its Applications 110)* ed C Borgers and F Natterer (Berlin: Springer) in press
- [88] Mueller K, Yagel R and Cornhill F 1997 The weighted-distance scheme: A globally optimizing projection ordering method for ART *IEEE Med. Imag.* **16** 223–30
- [89] Censor Y and Zenios S A 1997 *Parallel Optimization* (New York: Oxford University Press)
- [90] Eggermont P P B, Herman G T and Lent A 1981 Iterative algorithms for large partitioned systems, with applications to image reconstruction *Linear Algebra Appl.* **40** 37–67
- [91] Natterer F and Wübbeling F 1995 A propagation-backpropagation method for ultrasound tomography *Inverse Problems* **11** 1225–32
- [92] Saquib S S, Hanson K M and Cunningham G S 1997 Model-based image reconstruction from time-resolved diffusion data *Medical Imaging: Image Processing (Proc. SPIE 3034)* ed K M Hanson, pp 369–380
- [93] Dorn O 1998 A transport-backtransport method for optical tomography *Inverse Problems* **14** 1107–30
- [94] Roy R 1997 Image reconstruction from light measurements on biological tissue *PhD Thesis* University of Hertfordshire
- [95] Arridge S R, Schweiger M, Hiraoka M and Delpy D T 1993 Performance of an iterative reconstruction algorithm for near infrared absorption and scatter imaging *Photon Migration and Imaging in Random Media and Tissues (Proc. SPIE 1888)* ed B Chance and R R Alfano, pp 360–71
- [96] Paulsen K D and Jiang H 1996 Enhanced frequency-domain optical image reconstruction in tissues through total variation minimization *Appl. Opt.* **35** 3447–58
- [97] Wang Z X and Guo D R 1989 *Special Functions* (Singapore: World Scientific)
- [98] Chance B and Katzir A (ed) 1991 *Time-Resolved Spectroscopy and Imaging of Tissues (Proc. SPIE 1431)*
- [99] Chance B and Alfano R R (ed) 1995 *Optical Tomography, Photon Migration, and Spectroscopy of Tissue and Model Media: Theory, Human Studies, and Instrumentation (Proc. SPIE 2389)*
- [100] Chance B and Alfano R R (ed) 1993 *Photon Migration and Imaging in Random Media and Tissues (Proc. SPIE 1888)*
- [101] Borgers C and Natterer F (ed) 1998 *Computational Radiology and Imaging: Therapy and Diagnosis (IMA Volumes in Mathematics and its Applications 110)* (Berlin: Springer)