**University of Verona**

Department of Computer Science
Master's Degree in Mathematics

Graduation Thesis

# Variational Regularization of Single-Stage Qualitative Photoacoustic Tomography

**Candidate:**
**Andrei Goncharenok**

**Prof. Markus Haltmeier**
**Priv.-Doz. Dr. Lukas Neumann**
**University of Innsbruck**

**Prof. Giacomo Albi**
**University of Verona**

**Academic Year 2024–2025**

# Contents

# 1. Introduction

Quantitative Photoacoustic Tomography (QPAT) is a hybrid imaging technique that combines optical and acoustic wave propagation to infer quantitative tissue parameters from measurements of acoustic signals. By utilizing the high contrast of optics and high resolution of ultrasound, QPAT is an advanced biomedical imaging technique with potential applications in functional imaging, oncology, and vascular studies.

The fundamental principle of QPAT relies on the photoacoustic effect: if biological tissue is exposed to a brief optical pulse, part of the optical energy is absorbed by tissue chromophores. The absorption leads to localized heating, which is followed by thermoelastic expansion and the creation of acoustic waves. These pressure waves propagate through the tissue and are detected by ultrasound transducers placed outside the sample. By analyzing these acoustic signals, QPAT makes it possible to reconstruct images that reflect the optical absorption properties of the tissue.

While classical Photoacoustic Tomography (PAT) focuses on recovering qualitative information about tissue structures, QPAT aims to quantitatively estimate tissue parameters, such as optical absorption and scattering coefficients. This is achieved by solving an ill-posed inverse problem that couples optical light transport and acoustic wave propagation models. A correct solution of this inverse problem enables accurate characterization of biological tissues, which can significantly improve diagnostic and monitoring capabilities.

The inverse problem of QPAT is critically dependent on the accurate modeling of the forward problem that consists of three major components: the modeling of light transport in tissue, computing heat deposition from absorbed photons, and simulation of propagation of the generated acoustic waves. The complexity of these interacting physical processes demands advanced numerical methods.

In this work, we developed a one-step reconstruction method that directly recovers the optical absorption coefficient from measured acoustic signals. This is achieved by formulating a variational regularization problem combining optical and acoustic models. The resulting inverse problem is solved via an iterative gradient-based algorithm, using adjoint techniques to compute the update direction.

We study a numerical method derived from the discrete velocity model, which offers good accuracy and stability. As a result, it allows for precise reconstruction of

tissue parameters from measured photoacoustic signals.

In this paper, we follow the framework of QPAT, incorporating both optical light transport and acoustic wave propagation models, and formulate a variational inverse problem for reconstructing the optical absorption coefficient.

While we outline the theoretical basis for an adjoint formulation of the acoustic wave propagation model, in the numerical implementation we focus solely on the optical part. This choice is motivated by the technical challenges of implementing the acoustic adjoint model, integrating it into the iterative reconstruction scheme, and limited time.

# 2. Modelling of the Optical Part

We consider a bounded, convex domain $\Omega \subset \mathbb{R}^d$, with spatial dimension $d \in \{2, 3\}$, and assume that its boundary $\partial\Omega$ is piecewise smooth. This domain models the region of biological tissue under investigation in quantitative photoacoustic tomography (QPAT).

To describe the propagation of optical radiation within the tissue, we define the photon density function $\Phi : \Omega \times S^{d-1} \to \mathbb{R}$, where $\Phi(x, \theta)$ denotes the density of photons located at spatial position $x \in \Omega$ and traveling in direction $\theta \in S^{d-1}$, with $S^{d-1} \subset \mathbb{R}^d$ denoting the unit sphere of directions.

The propagation of light in biological tissue is modeled by the stationary radiative transfer equation (RTE), which describes the behavior of the photon density function $\Phi$. It satisfies:

$$\theta \cdot \nabla_x \Phi(x, \theta) + (\mu_a(x) + \mu_s(x))\Phi(x, \theta) = \mu_s(x) \int_{S^{d-1}} k(\theta, \theta')\Phi(x, \theta')\, d\theta' + q(x, \theta),$$

for all $(x, \theta) \in \Omega \times S^{d-1}$, where the function $q(x, \theta)$ represents an internal photon source within the tissue.

The interaction of photons with the medium is governed by three spatially dependent optical parameters. The absorption coefficient $\mu_a(x)$ quantifies the likelihood of photon absorption per unit length traveled within the medium. The scattering coefficient $\mu_s(x)$ describes the probability of scattering events occurring per unit length, accounting for the redirection of photons as they propagate through the tissue. Finally, the scattering phase function $k(\theta, \theta')$ determines the angular distribution of scattering; that is, it models the probability density of a photon initially moving in direction $\theta'$ being scattered into direction $\theta$. These parameters together define the behavior of light transport within the tissue.

The phase function $k(\theta, \theta')$ satisfies the normalization condition:

$$\int_{S^{d-1}} k(\theta, \theta')\, d\theta' = 1,$$

ensuring conservation of energy. In this work, we assume that $k$ is known and follows

the Henyey–Greenstein model:

$$k(\theta, \theta') = \frac{1 - g^2}{(1 + g^2 - 2g\,\theta \cdot \theta')^{3/2}}, \quad \text{with } g \in [0, 1),$$

where the anisotropy factor $g$ quantifies the degree of forward scattering in the medium.

To ensure a well-posed problem, we impose boundary conditions by defining the inflow and outflow boundaries as

$$\Gamma^- = \{(x, \theta) \in \partial\Omega \times S^{d-1} \mid \nu(x) \cdot \theta < 0\}, \quad \Gamma^+ = \{(x, \theta) \in \partial\Omega \times S^{d-1} \mid \nu(x) \cdot \theta > 0\},$$

where $\nu(x)$ is the outward-pointing unit normal at $x \in \partial\Omega$. The RTE is supplemented with a zero inflow boundary condition:

$$\Phi(x, \theta) = 0, \quad \text{for } (x, \theta) \in \Gamma^-,$$

where $f(x, \theta)$ represents a prescribed boundary source pattern.

We introduce the following Hilbert spaces to describe the unknowns and data of the problem:

$$\begin{aligned} Q &:= L^2(\Omega \times S^{d-1}) \quad \text{— space of photon sources } q(x, \theta), \\ X &:= L^2(\Omega)^2 \quad \text{— space of optical parameters } (\mu_a, \mu_s), \\ W &:= L^2(\Omega \times S^{d-1}) \quad \text{— space of photon density } \Phi, \\ Y &:= L^2(\partial B_R \times (0, T)) \quad \text{— space of acoustic data on detectors surrounding } \Omega. \end{aligned}$$

Here, $B_R \supset \Omega$ is a ball of radius $R$ centered at the origin and $T > 0$ is the final measurement time. We define the admissible set of tissue parameters as

$$D(T) := \{(\mu_a, \mu_s) \in X \,|\, 0 < \mu_a(x) \leq \bar{\mu}_a, \quad 0 < \mu_s(x) \leq \bar{\mu}_s \text{ for almost all } x \in \Omega\},$$

where $\bar{\mu}_a, \bar{\mu}_s > 0$ are known upper bounds. These conditions reflect the physical constraint that both absorption and scattering coefficients must be strictly positive and bounded in biological tissue.

Assuming a well-posed formulation (see, e.g., Rabanser et al. [4]), the forward photon transport can be modeled as

$$T : Q \times X \to W, \quad (q, \mu_a, \mu_s) \mapsto \Phi,$$

which maps the prescribed internal optical sources and tissue parameters to the resulting photon density distribution $\Phi$ modeled as the solution of the RTE with zero inflow boundary conditions.

The total absorbed optical energy (heating function), serving as the source term

for acoustic wave propagation, is given by:

$$H(x) := \mu_a(x) \int_{S^{d-1}} \Phi(x, \theta)\, d\theta.$$

We introduce the averaging operator $A\colon W \to L^2(\Omega)$, defined as

$$A\Phi(x) := \int_{S^{d-1}} \Phi(x, \theta)\, d\theta,$$

To facilitate the inverse problem analysis and operator-based formulation of QPAT, we define the heating map as

$$\mathcal{H}(q, \mu_a, \mu_s)(x) := \mu_a(x) \cdot A[T(q, \mu_a, \mu_s)](x).$$

In simplified numerical tests aimed at validating the forward model, we set $\mu_a(x) = 1$. This assumption removes the dependence of the heating function on the absorption coefficient, reducing it to $H(x) = A\Phi(x)$. While this does not represent the full QPAT setting, it allows for isolated evaluation of light transport and its numerical discretization.

This operator represents the nonlinear mapping from the optical parameters and source distribution to the spatially resolved absorbed energy profile within the tissue. The function $H(x) = \mathcal{H}(q, \mu_a, \mu_s)(x)$, will serve as the initial condition for the acoustic wave equation in the QPAT framework, which we describe next.

# 3. Modelling of the Acoustic Part

Given the absorbed energy distribution $H(x)$ obtained from the optical model, we now describe how the resulting acoustic wave propagates in tissue and is measured on the boundary.

The local deposition of thermal energy caused by photon absorption gives rise to an acoustic pressure wave. Under the assumption of constant sound speed (normalized to one) and negligible acoustic attenuation, the propagation of the acoustic pressure field $p(x,t)$ in free space is described by the wave equation:

$$\left(\frac{\partial^2}{\partial t^2} - \Delta\right) p(x,t) = 0, \quad \text{for } (x,t) \in \mathbb{R}^d \times (0,\infty),$$

with initial conditions determined by the optical heating:

$$p(x,0) = H(x), \quad \frac{\partial p}{\partial t}(x,0) = 0, \quad x \in \mathbb{R}^d,$$

where $H(x)$ is the absorbed photon energy distribution computed from the optical model, and is assumed to have compact support within a ball $B_R \subset \mathbb{R}^d$ of radius $R$, such that $\Omega \subset B_R$.

The solution $p(x,t)$ of the wave equation is recorded on a subset of the boundary $\partial B_R \times (0,T)$, where measurements are collected. The space of acoustic data is given by

$$Y := L^2(\partial B_R \times (0,T)).$$

We define the acoustic forward operator as:

$$U : L^2(B_R) \to Y, \quad H(x) \mapsto p|_{\partial B_R \times (0,T)},$$

where $p(x,t)$ is the solution of the wave equation with initial data $p(x,0) = H(x)$ and $\partial_t p(x,0) = 0$. The trace of the pressure field $p$ on $\partial B_R \times (0,T)$, denoted by $p|_{\partial B_R \times (0,T)}$, corresponds to the pressure signals measured on the boundary and defines the action of the acoustic forward operator $U$.

We recall the following fundamental result:
Let $H \in C_0(\mathbb{R}^d)$ be supported in $B_R$, and let $p(x,t)$ solve the wave equation with initial

data $p(x,0) = H(x)$, $\partial_t p(x,0) = 0$. Then the following energy identity holds:

$$\int_{\partial B_R} \int_0^\infty p^2(x,t)\, t\, dt\, dS(x) = R \int_{B_R} H^2(x)\, dx.$$

This identity shows that the wave operator $U$ is an isometry up to a constant, preserving energy in the $L^2$-sense. This guarantees stability of the acoustic forward map against initial pressure perturbations. Rigorous proofs of the isometry property in even and odd spatial dimensions can be found in [2] and [3].

# 4. Inverse Problem of Single-Stage QPAT

In quantitative photoacoustic tomography (QPAT), the goal is to reconstruct optical parameters of biological tissue, such as the absorption and scattering coefficients $\mu_a(x)$ and $\mu_s(x)$, from measurements of the acoustic pressure wave generated by optical absorption.

As described earlier, the photon absorption gives rise to an initial pressure distribution $H(x)$, modeled by the heating operator

$$\mathcal{H}(q, \mu_a, \mu_s)(x) := \mu_a(x) \cdot A[T(q, \mu_a, \mu_s)](x),$$

where $T$ solves the stationary radiative transfer equation and $A$ is the angular averaging operator.

This pressure distribution serves as the initial condition for the acoustic wave equation. The measured acoustic data are modeled by an operator

$$M \circ U \colon L^2(B_R) \to Y_{\text{meas}},$$

where $U$ maps the initial pressure to the boundary trace of the acoustic wave, and $M$ is a bounded measurement operator modeling restriction to detector locations or time intervals.

We define the full QPAT forward operator as:

$$\mathcal{F}(q, \mu_a, \mu_s) := M \circ U \circ \mathcal{H}(q, \mu_a, \mu_s),$$

which maps optical parameters and sources to measured acoustic data.

The inverse problem of single-stage QPAT consists in recovering the spatially varying optical coefficients $(\mu_a, \mu_s)$ from a known source $q$ and observed acoustic measurements:

$$p_{\text{meas}} = \mathcal{F}(q, \mu_a, \mu_s).$$

This inverse problem is nonlinear and ill-posed due to the composition of non-smooth and compact operators, as well as the limited, noisy, or spatially restricted nature of the measurement data. Regularization procedures, variational formulations,

and adjoint-based methods are typically employed for stabilizing and solving the reconstruction problem. In this work, we adopt a fully discrete approach in which the forward and inverse problems are discretized first, and all computations, including regularization and optimization, are carried out in the discrete setting. These components will be discussed in the subsequent sections.

# 5. Discretization of Optical Part

For the numerical solution of the stationary radiative transfer equation (RTE), which models the propagation of optical radiation in a scattering and absorbing medium, we employ the Discrete Ordinates Method (DOM). The RTE is given by:

$$\theta \cdot \nabla\Phi(x,\theta) + \mu_a(x)\,\Phi(x,\theta) + \mu_s(x)\,\Phi(x,\theta) - \mu_s(x)\int_{S^1} k(\theta,\theta')\,\Phi(x,\theta')\,d\theta' = q(x,\theta)$$

where $x \in \Omega \subset \mathbb{R}^2$, $\theta \in S^1$.

Here, $\Phi(x,\theta)$ is the angular photon density at spatial point $x$ traveling in direction $\theta \in S^1$; $\mu_a(x)$ denotes the absorption coefficient, while $\mu_s(x)$ denotes the scattering coefficient. The scattering phase function $k(\theta,\theta')$ characterizes the probability that a photon originally moving in direction $\theta'$ will be scattered into direction $\theta$. Finally, $q(x,\theta)$ represents the density of an external optical source at point $x$ in direction $\theta$.

As $\theta \in S^1$, the radiative transfer equation includes an integral over all directions, making the problem dependent on a continuous angular variable. From a functional analysis perspective, this results in an effectively infinite-dimensional formulation when dealing with angular-dependent functions.

For numerical approximation, this continuous space is replaced by a finite set of discrete directions, reducing the problem to a finite-dimensional system of equations.

We consider a uniform discretization of $S^1$ into $n$ directions. Let

$$\varphi_i = \frac{2\pi(i-1)}{n}, \quad i = 1,\dots,n,$$

and define the corresponding unit direction vectors as

$$\boldsymbol{\theta}_i := (\cos\varphi_i, \sin\varphi_i) \in S^1.$$

Thus, the set of discrete directions is

$$\Theta := \{\boldsymbol{\theta}_i \in S^1 \mid i = 1,\dots,n\}.$$

The continuous function $\Phi(x,\theta)$ is then approximated by its values along these

fixed directions:

$$\Phi(x,\theta) \approx \Phi_i(x) := \Phi(x,\boldsymbol{\theta_i}), \quad i = 1,\dots,n.$$

... where $\boldsymbol{\theta}_i$ denotes the $i$-th direction in the angular quadrature grid.

In this way, instead of a single function of both spatial and angular variables, we work with $n$ scalar functions $\Phi_i(x)$, each describing the photon density traveling in direction $\boldsymbol{\theta}_i$. This discretization reduces the angular integral to a finite sum, yielding a coupled system of differential equations for $\Phi_i(x)$.

Substituting the angular discretization described above into the original radiative transfer equation and approximating the angular integral using a quadrature rule, we obtain a system of $n$ coupled equations:

$$\boldsymbol{\theta}_i \cdot \nabla\Phi_i(x) + (\mu_a(x) + \mu_s(x))\Phi_i(x) - \mu_s(x)\sum_{j=1}^{n} k_{ij}\Phi_j(x)w_j = q_i(x), \quad i = 1,\dots,n,$$

The values $k_{ij} := k(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j)$ represent the discrete scattering phase function, quantifying the probability that a photon traveling in direction $\boldsymbol{\theta}_j$ is scattered into direction $\boldsymbol{\theta}_i$. These values form the entries of the phase matrix

$$K := (k_{ij}) \in \mathbb{R}^{n\times n},$$

which is used both in the discretized scattering integral and in the construction of the inter-directional scattering operator.

Although the term $\boldsymbol{\theta}_i \cdot \nabla\Phi_i(x)$ appears as a standard scalar product, we do not discretize it directly in that form. Instead, for each discrete direction $\boldsymbol{\theta}_i$, we approximate the directional derivative $\nabla_{\boldsymbol{\theta}_i}\Phi_i(x)$ using a linear combination of upwind-type finite difference operators aligned with the Cartesian grid axes. This approach results in a directionally structured discretization that is consistent with the grid and well-suited for implementation within the Discrete Ordinates Method.

The quadrature weights $w_j = \frac{2\pi}{n}$ are uniform due to the regular angular discretization.

The source vector $\mathbf{q}$ can be constructed based on the desired angular emission profile. For example, a directional Gaussian source may be implemented by populating only the components corresponding to a selected direction $\boldsymbol{\theta}_i$, while isotropic illumination can be modeled by assigning the same spatial profile across all directions.

Thus, after replacing the continuous angular variable with a finite set of directions and approximating the angular integral using a quadrature rule, the original radiative transfer equation is transformed into a system of $n$ differential equations. These equations are linearly coupled through the scattering phase function, allowing the problem to be treated as a standard boundary value problem for a system of equa-

tions. This approach forms the foundation of the Discrete Ordinates Method (DOM).

To represent the above system of equations more compactly and in a form suitable for analytical and numerical analysis, we introduce the vector of unknowns:

$$\mathbf{\Phi}(x) := \begin{bmatrix} \Phi_1(x) \\ \Phi_2(x) \\ \vdots \\ \Phi_n(x) \end{bmatrix}, \quad \mathbf{q}(x) := \begin{bmatrix} q_1(x) \\ q_2(x) \\ \vdots \\ q_n(x) \end{bmatrix}.$$

where each function $\Phi_i(x)$ describes the photon density in direction $\boldsymbol{\theta}_i$ and is discretized over the spatial grid into a vector in $\mathbb{R}^N$, where $N = N_x \cdot N_y$ is the total number of spatial nodes. Consequently, the global vector $\mathbf{\Phi} \in \mathbb{R}^{nN}$ stacks all directional components into a single column vector, which serves as the unknown in the fully discretized system. From this point onward, we omit the argument $x$ and treat $\mathbf{\Phi}$ as the fully discrete solution vector in $\mathbb{R}^{nN}$.

The diagonal matrix of quadrature weights is defined as:

$$W := \operatorname{diag}(w_1, \ldots, w_n).$$

The transport derivatives $\boldsymbol{\theta}_i \cdot \nabla$, appearing in each equation of the system derived via the Discrete Ordinates Method, can be interpreted as directional differential operators along the respective directions. However, in our implementation, we do not work with the scalar product form directly. Instead, we discretize the directional derivatives separately along each direction $\boldsymbol{\theta}_i$, which is more accurately reflected by the notation $\nabla_{\boldsymbol{\theta}_i}$. Note that this denotes the directional derivative with respect to the spatial variable $x$, taken in the direction $\boldsymbol{\theta}_i$, and should not be confused with a derivative with respect to the angular variable $\theta \in S^1$. For theoretical purposes, we define the directional derivative operator as

$$D(x) := \operatorname{diag}(\nabla_{\theta_1}, \ldots, \nabla_{\theta_n}),$$

which provides a compact representation of transport along discrete directions. In the numerical implementation, however, this operator is replaced by the discrete block-diagonal matrix $A_{\text{dir}}$, constructed from upwind finite differences as described below.

Using the introduced notation, the discretized system of equations can be rewritten in vector-matrix form as:

$$D(x)\mathbf{\Phi}(x) + (\mu_a(x) + \mu_s(x))I_n\mathbf{\Phi}(x) - \mu_s(x)KW\mathbf{\Phi}(x) = \mathbf{q}(x),$$

where $I_n$ is the identity matrix of size $n \times n$.

This formulation emphasizes the structure of the problem as a system of linear partial differential equations in the spatial variables with matrix-valued coefficients. The matrix $KW$ represents the scattering operator acting on the vector of angular components, while the operator $D(x)$ encodes directional derivatives along the selected discrete directions. Such a representation is especially useful in numerical schemes based on linear algebra techniques.

To numerically solve the system of equations obtained via the Discrete Ordinates Method (DOM), it is necessary to discretize the spatial domain $\Omega \subset \mathbb{R}^2$.

The Discrete Ordinates Method (DOM) transforms the radiative transfer equation into a coupled system of PDEs by discretizing the angular variable and approximating the integral with quadrature. A comprehensive overview of the DOM and its applications can be found in [6, 1].

The domain $\Omega$ is approximated by a rectangular computational domain $[x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}]$, which is divided into a regular uniform grid with $N_x$ nodes along the $x$-axis and $N_y$ nodes along the $y$-axis. The spatial steps are defined as:

$$\Delta x = \frac{x_{\max} - x_{\min}}{N_x - 1}, \quad \Delta y = \frac{y_{\max} - y_{\min}}{N_y - 1}.$$

The grid nodes are given by:

$$x_m = x_{\min} + (m-1)\Delta x, \quad y_n = y_{\min} + (n-1)\Delta y,$$

where $m = 1, \ldots, N_x$, $n = 1, \ldots, N_y$. The total number of grid nodes is $N = N_x \cdot N_y$.

To approximate the transport term $\nabla_\theta$ appearing in the radiative transfer equation, we employ first-order upwind finite difference schemes that are adapted to each discrete direction $\boldsymbol{\theta}_i \in S^1$. This is essential to ensure numerical stability and monotonicity, particularly in the context of advection-dominated transport.

Let $\Omega \subset \mathbb{R}^2$ be discretized into a uniform Cartesian grid consisting of $N_x \times N_y$ points with step sizes $\Delta x$ and $\Delta y$. The total number of spatial grid points is $N = N_x \cdot N_y$. A grid function $u(x_m, y_n)$ defined at these nodes is represented as a vector in $\mathbb{R}^N$.

To approximate directional derivatives, we define four sparse matrices: $D_x^+$ and $D_x^-$ represent the forward and backward finite difference operators in the $x$-direction, while $D_y^+$ and $D_y^-$ serve the same purpose in the $y$-direction.

These operators are constructed to account for the structured grid topology. For example, $D_x^-$ is a sparse matrix implementing the stencil:

$$\frac{u(x_m, y_n) - u(x_{m-1}, y_n)}{\Delta x},$$

while $D_x^+$ uses:

$$\frac{u(x_{m+1}, y_n) - u(x_m, y_n)}{\Delta x}.$$

Analogous expressions hold for the $y$-direction.

To approximate the transport term $\nabla_{\boldsymbol{\theta}_i} \Phi_i(x)$ numerically for each direction, we define the discrete directional operator $A_i$. Each discrete transport direction $\boldsymbol{\theta}_i = (v_x^{(i)}, v_y^{(i)}) \in \mathbb{R}^2$, normalized to unit length, induces a directional derivative operator $A_i \in \mathbb{R}^{N \times N}$, defined as:

$$A_i := \begin{cases} v_x^{(i)} D_x^- & \text{if } v_x^{(i)} > 0, \\ v_x^{(i)} D_x^+ & \text{if } v_x^{(i)} < 0, \\ 0 & \text{if } v_x^{(i)} = 0, \end{cases} + \begin{cases} v_y^{(i)} D_y^- & \text{if } v_y^{(i)} > 0, \\ v_y^{(i)} D_y^+ & \text{if } v_y^{(i)} < 0, \\ 0 & \text{if } v_y^{(i)} = 0. \end{cases}$$

This formulation ensures that information propagates in the correct upwind direction depending on the sign of each component of $\boldsymbol{\theta}_i$. The derivative is taken against the direction of flow, which is crucial for maintaining numerical stability in advection-type equations [7].

For diagonal directions (e.g., $\boldsymbol{\theta}_i = (\pm 1/\sqrt{2}, \pm 1/\sqrt{2})$), the above construction naturally yields consistent approximations based on combinations of $D_x^\pm$ and $D_y^\pm$.

Finally, to construct the full system of transport equations in the Discrete Ordinates Method (DOM), we assemble the individual operators $A_i$ into a block-diagonal matrix:

$$A_{\text{dir}} := \text{block\_diag}(A_1, A_2, \ldots, A_n) \in \mathbb{R}^{nN \times nN},$$

which acts on the global solution vector $\boldsymbol{\Phi}$, encoding directional transport before scattering is applied.

This block structure naturally reflects the decoupled transport along discrete directions before scattering is applied, and it enables efficient implementation using sparse linear algebra techniques.

After applying both spatial and angular discretization to the radiative transfer equation, we obtain the final sparse linear system. To construct this system, we now define the required matrix operators.

Let $\mu_a(x), \mu_s(x) : \Omega \to \mathbb{R}_+$ denote the spatial distributions of the absorption and scattering coefficients, defined at the grid nodes $\{x_j\}_{j=1}^N$, where $N = N_x \cdot N_y$. These functions are discretized as vectors:

$$\mu_a = \begin{bmatrix} \mu_a(x_1) \\ \mu_a(x_2) \\ \vdots \\ \mu_a(x_N) \end{bmatrix} \in \mathbb{R}^N, \quad \mu_s = \begin{bmatrix} \mu_s(x_1) \\ \mu_s(x_2) \\ \vdots \\ \mu_s(x_N) \end{bmatrix} \in \mathbb{R}^N.$$

The combined operator describing attenuation of photon density due to absorption and out-scattering (i.e., removal of photons from a given direction) is given by:

$$M_{\text{abs+scat}} := \text{diag}(\mu_a + \mu_s) \otimes I_n \in \mathbb{R}^{nN \times nN},$$

where $\otimes$ denotes the Kronecker product and $I_n$ is the identity matrix of size $n \times n$.

Let $K \in \mathbb{R}^{n \times n}$ denote the phase function matrix (as defined above), assumed to satisfy the normalization condition

$$\sum_{j=1}^{n} k_{ij} w_j = 1 \quad \text{for all } i.$$

The inter-directional scattering operator, which models the redistribution of photons between different directions, is defined as:

$$M_{\text{scat}} := K \otimes \text{diag}(\mu_s) \in \mathbb{R}^{nN \times nN}.$$

It captures the redistribution of photon energy into each direction $\boldsymbol{\theta}_i$ due to scattering from all other directions $\boldsymbol{\theta}_j$, as described by the phase function matrix $K$.

Each row of $M_{\text{scat}}$ describes the energy being scattered into a fixed direction $\boldsymbol{\theta}_i$ from all other directions.

The complete discretized system of equations, incorporating directional transport, absorption, and scattering effects within the Discrete Ordinates Method (DOM), is written in matrix form as

$$A_{\text{total}} \boldsymbol{\Phi} = \mathbf{q},$$

where the system matrix $A_{\text{total}}$ is defined by:

$$A_{\text{total}} := A_{\text{dir}} + M_{\text{abs+scat}} - M_{\text{scat}}.$$

Here, $A_{\text{dir}} \in \mathbb{R}^{nN \times nN}$ is the block-diagonal operator representing directional spatial transport along each discrete direction (see Section 3.2). The vector $\boldsymbol{\Phi} \in \mathbb{R}^{nN}$ contains all directional photon density vectors $\boldsymbol{\Phi}_i \in \mathbb{R}^N$, concatenated into a single column.

The source vector $\mathbf{q} \in \mathbb{R}^{nN}$ is constructed from directional components $\boldsymbol{q}_i \in \mathbb{R}^N$, where each corresponds to a specific angular direction. For example, a directional source can be modeled by setting the source term nonzero only for the component $\boldsymbol{q}_i(x)$ corresponding to a chosen direction $\boldsymbol{\theta}_i$, with a spatial Gaussian profile centered at a specified location. In contrast, isotropic illumination can be modeled by assigning the same spatial profile to all directional components.

The resulting sparse linear system is solved using efficient direct solvers from scientific computing libraries. For example, in Python, we employ the `spsolve` routine from the SciPy package, which provides LU decomposition-based solutions for sparse matrices.

After computing the numerical solution of the system, represented by the vector

$$\mathbf{\Phi} = \begin{bmatrix} \mathbf{\Phi}_1^\top & \mathbf{\Phi}_2^\top & \dots & \mathbf{\Phi}_n^\top \end{bmatrix}^\top \in \mathbb{R}^{nN},$$

the total fluence $\Phi_{\text{total}}(x)$, which corresponds to the angular integral of the photon density over the unit circle, is defined as:

$$\Phi_{\text{total}}(x) := \int_{S^1} \Phi(x, \theta)\, d\theta.$$

In the numerical implementation, this integral is approximated by a quadrature rule over the discrete set of directions:

$$\Phi_{\text{total}}(x) \approx \sum_{i=1}^{n} w_i \Phi_i(x),$$

where $w_i = \frac{2\pi}{n}$ are the weights of the uniform quadrature formula.

Since each directional component $\mathbf{\Phi}_i \in \mathbb{R}^N$ is stored as a discrete vector over all spatial grid nodes, the total fluence vector $\mathbf{\Phi}_{\text{total}} \in \mathbb{R}^N$ is computed by summing the weighted contributions from each direction:

$$\mathbf{\Phi}_{\text{total}} := \sum_{i=1}^{n} w_i \mathbf{\Phi}_i.$$

Note that in earlier expressions such as $H(x) = A\,\Phi(x)$, the quantity $\Phi(x)$ corresponds to the total fluence obtained by integrating $\Phi(x, \theta)$ over all directions. In the discrete setting, this is approximated by the vector $\mathbf{\Phi}_{\text{total}}$.

In practice, since each directional component $\mathbf{\Phi}_i$ is stored as a discrete vector over the spatial grid, the total fluence vector $\mathbf{\Phi}_{\text{total}} \in \mathbb{R}^N$ approximates the continuous fluence function $\Phi_{\text{total}}(x)$ evaluated at the grid nodes. This discretized quantity serves as the basis for further computations, such as photothermal source modeling.

The absorbed optical energy that is converted into heat is proportional to the product of the local fluence and the absorption coefficient. This defines the photothermal source:

$$H(x) := \mu_a(x) \cdot \Phi_{\text{total}}(x).$$

In the discrete setting, the photothermal source is computed as the element-wise

(Hadamard) product of vectors:

$$\mathbf{H} := \mu_a \odot \mathbf{\Phi}_{\text{total}} \in \mathbb{R}^N,$$

where $\odot$ denotes component-wise multiplication.

The photothermal energy distribution $\mathbf{H}$ serves as the initial pressure $p_0(x)$ in the acoustic wave equation, thereby linking the optical model to the acoustic signal in QPAT.

# 6. Discretization of the Acoustic Part

In quantitative photoacoustic tomography (QPAT), the acoustic forward problem models the propagation of pressure waves generated by optical absorption. Once the light fluence $\Phi(x)$ is computed, the initial pressure distribution is given by:

$$H(x) = \mu_a(x)\Phi(x),$$

where $\mu_a(x)$ is the absorption coefficient. This initial pressure acts as a source term for the acoustic wave equation.

We consider the standard second-order wave equation in two spatial dimensions:

$$\begin{cases} \partial_t^2 p(x,t) - c^2 \Delta p(x,t) = 0, & (x,t) \in \Omega \times (0,T), \\ p(x,0) = H(x), & x \in \Omega, \\ \partial_t p(x,0) = 0, & x \in \Omega, \end{cases}$$

where $p(x,t)$ is the acoustic pressure field, $c > 0$ is the (constant) speed of sound in the medium, $\Omega \subset \mathbb{R}^2$ is a rectangular spatial domain, and $T > 0$ is the final time of observation.

The goal of the acoustic forward model is to simulate the time evolution of $p(x,t)$ for a given initial pressure $H(x)$, and to compute the values of $p(x,t)$ at detector locations distributed along the boundary of the domain. These simulated pressure measurements constitute the data used in the inverse problem.

To simulate acoustic wave propagation, we implemented a simplified pseudo-spectral solver in Python, based on FFTs for spatial derivatives and second-order central differences in time. While this solver captures key ideas of PSTD methods, it does not include full features of established toolboxes such as k-Wave.

For realistic forward modeling in quantitative photoacoustic tomography (QPAT), advanced solvers like the `k-Wave` MATLAB toolbox [9] are widely used. This toolbox implements the k-space pseudo-spectral method, supports absorbing boundary conditions, heterogeneous media, and has been widely adopted in the photoacoustic community.

Our custom solver provides a conceptually clear and computationally efficient

model tailored for proof-of-concept reconstructions and gradient-based optimization.

To simulate the acoustic wave propagation numerically, we discretize the domain $\Omega = [-L, L]^2 \subset \mathbb{R}^2$ using a uniform Cartesian grid of size $N \times N$. The grid spacing is defined as

$$\Delta x = \Delta y = \frac{2L}{N},$$

and the spatial grid points are given by

$$x_i = -L + i\Delta x, \quad y_j = -L + j\Delta y, \quad \text{for } i, j = 0, \ldots, N-1.$$

The acoustic pressure field $p(x, t)$ is approximated at discrete spatial nodes and time levels $t_n = n\Delta t$, where $\Delta t$ is the time step and $n = 0, 1, \ldots, N_t$. At each time level, the pressure is represented as a matrix $P^n \in \mathbb{R}^{N\times N}$, where

$$P^n_{ij} \approx p(x_i, y_j, t_n).$$

To approximate the spatial Laplacian $\Delta p$, we employ a pseudo-spectral method based on the two-dimensional discrete Fourier transform (DFT). For a grid function $P \in \mathbb{R}^{N\times N}$, the Laplacian is computed in the frequency domain as:

$$\Delta P \approx \mathcal{F}^{-1}\left[-|\mathbf{k}|^2 \cdot \mathcal{F}[P]\right],$$

where: - $\mathcal{F}[\cdot]$ and $\mathcal{F}^{-1}[\cdot]$ denote the two-dimensional DFT and its inverse, - $|\mathbf{k}|^2 = k_x^2 + k_y^2$ is the squared spatial frequency, with wave number vectors defined by

$$k_x = \frac{\pi}{L} \cdot \text{fftshift}\left(\left[-\frac{N}{2}, -\frac{N}{2}+1, \ldots, \frac{N}{2}-1\right]\right),$$

and similarly for $k_y$.

The multiplication in Fourier space corresponds to applying the Laplacian operator spectrally. This method achieves spectral accuracy in space under the assumption of periodic boundary conditions and is particularly efficient due to the use of fast FFT algorithms.

The temporal domain $[0, T]$ is uniformly discretized using a time step $\Delta t = \frac{T}{N_t - 1}$, resulting in discrete time levels

$$t_n = n\Delta t, \quad \text{for } n = 0, 1, \ldots, N_t - 1.$$

To approximate the second-order time derivative in the wave equation, we employ a second-order accurate central finite difference scheme. The continuous equation

$$\partial_t^2 p(x, t) = c^2 \Delta p(x, t)$$

is discretized as

$$\frac{P_{ij}^{n+1} - 2P_{ij}^{n} + P_{ij}^{n-1}}{\Delta t^2} = c^2 (\Delta P^n)_{ij},$$

where $P_{ij}^n \approx p(x_i, y_j, t_n)$, and $(\Delta P^n)_{ij}$ denotes the spatial Laplacian evaluated at time $t_n$ using the pseudo-spectral method described earlier.

This leads to the following explicit time-stepping scheme:

$$P_{ij}^{n+1} = 2P_{ij}^{n} - P_{ij}^{n-1} + c^2 \Delta t^2 \cdot (\Delta P^n)_{ij}.$$

To initialize this recurrence, we use the given initial conditions:

$$P_{ij}^{0} = H(x_i, y_j), \quad \text{and} \quad P_{ij}^{-1} = P_{ij}^{0} - \frac{\Delta t^2}{2} \cdot c^2 (\Delta P^0)_{ij}.$$

The second expression is derived from a second-order Taylor expansion in time and is consistent with the assumption $\partial_t p(x, 0) = 0$. This initialization ensures correct enforcement of the initial conditions and allows the scheme to advance from $n = 0$ onward.

We assume that acoustic pressure is recorded on a circular array of detectors of radius $R_0$, uniformly distributed around the domain $\Omega \subset \mathbb{R}^2$. The angular positions of the $N_\phi$ detectors are given by

$$\phi_k = \frac{2\pi(k-1)}{N_\phi}, \quad k = 1, 2, \ldots, N_\phi,$$

and the corresponding detector coordinates are

$$x_k = R_0 \cos(\phi_k), \quad y_k = R_0 \sin(\phi_k).$$

The detectors are placed along a circular boundary of radius $R_0 \geq L$, enclosing the computational domain $\Omega$.

These points represent a uniform sampling of the boundary $\partial B_{R_0}$, where the simulated pressure field will be evaluated.

Since the pressure field $p(x, t)$ is computed on a uniform Cartesian grid, the detector coordinates $(x_k, y_k)$ generally do not coincide with grid nodes. Therefore, at each time step $t_n = n\Delta t$, the pressure values at detector locations are obtained via bicubic spline interpolation from the discrete pressure field $P^n \in \mathbb{R}^{N_x \times N_y}$. The interpolated measurement at detector $k$ and time $t_n$ is denoted as:

$$y_k^n := p(x_k, t_n) \approx \text{Interp2D}(P^n, x_k, y_k),$$

where Interp2D$(\cdot)$ denotes a bicubic interpolant constructed from $P^n$.

More specifically, bicubic interpolation builds a smooth approximation over the

grid by fitting a piecewise third-degree polynomial patch over each cell of the 2D grid, using both function values and estimated partial derivatives. In our implementation, this is achieved using a bicubic spline interpolator
(e.g., `scipy.interpolate.RectBivariateSpline`), which precomputes spline coefficients from the discrete grid $P^n$, and then evaluates them at off-grid detector coordinates $(x_k, y_k)$. This approach provides a smooth and accurate interpolation, particularly important for extracting pressure values at finely spaced detector arrays located between grid nodes.

This process is repeated for all $k = 1, \ldots, N_\phi$ and $n = 0, \ldots, N_t - 1$, resulting in a discrete measurement matrix:

$$Y \in \mathbb{R}^{N_\phi \times N_t}, \quad Y[k, n] := p(x_k, t_n).$$

The full matrix $Y$ represents the simulated photoacoustic signal collected by all detectors over time and constitutes the output of the acoustic forward model. This matrix is typically flattened into a vector form for inverse problem formulations:

$$y_{\text{meas}} := \text{vec}(Y) \in \mathbb{R}^{N_\phi \cdot N_t},$$

which is the input data used in variational reconstruction algorithms.

We denote by $p_{\text{meas}}(x, t)$ the continuous pressure signal, and denote by $y_{\text{meas}} \in \mathbb{R}^{N_\phi \cdot N_t}$ the discrete acoustic measurements used in the inverse problem.

The overall method provides a high-order, stable, and efficient approach to solving the acoustic forward problem arising in photoacoustic tomography. It combines spectral accuracy in space with second-order finite differences in time, and facilitates accurate modeling of wave propagation and measurement on circular detector arrays.

After full discretization, the acoustic forward problem can be compactly represented by a (quasi-)linear operator

$$H \mapsto y_{\text{meas}},$$

where $H \in \mathbb{R}^{N_x \times N_y}$ denotes the discretized initial pressure distribution, and $y_{\text{meas}} \in \mathbb{R}^{N_\phi \cdot N_t}$ is the resulting measurement vector collected on the detector ring.

We define the discrete forward operator $M : \mathbb{R}^{N_x \times N_y} \to \mathbb{R}^{N_\phi \cdot N_t}$ such that:

$$y_{\text{meas}} = M(H).$$

The operator $M$ consists of several sequential steps.

First, the spatial Laplacian is approximated spectrally using the Fast Fourier

Transform (FFT), according to the relation

$$\Delta P \approx \mathcal{F}^{-1}\left[-|\mathbf{k}|^2 \cdot \mathcal{F}[P]\right].$$

Second, time stepping is performed using a second-order central difference scheme,

$$P^{n+1} = 2P^n - P^{n-1} + c^2\Delta t^2 \cdot \Delta P^n,$$

with initial conditions given by $P^0 = H$ and

$$P^{-1} = P^0 - \frac{\Delta t^2}{2} \cdot \Delta P^0.$$

Third, at each time step $t_n$, the pressure field $P^n$ is interpolated at the detector locations $(x_k, y_k)$.

Finally, the resulting values are assembled into the measurement matrix $Y \in \mathbb{R}^{N_\phi \times N_t}$, which is then vectorized to form $y_{\text{meas}} = \text{vec}(Y)$.

This forward operator $M$ serves as a building block in the composite model

$$F = M \circ U \circ \mathcal{H},$$

where $\mathcal{H}$ models the optical excitation and fluence, and $M$ incorporates measurement or noise-related transformations applied to the acoustic signal. The full operator $F$ defines the forward model used in the solution of the inverse problem in photoacoustic tomography.

This completes the discretization of the acoustic part. In the next chapter, we consider the inverse problem of reconstructing the initial pressure $H$ from the simulated measurements $y_{\text{meas}}$.

Although the forward operator $F = M \circ M \circ \mathcal{H}$ is naturally decomposed into two components, the inverse problem can be approached in two conceptually distinct ways. One possibility is to decouple the inversion into two stages: first recover the initial pressure distribution $H$ by inverting the acoustic operator $M$, and then solve the optical inverse problem $\mathcal{H}^{-1}$ to retrieve the absorption coefficient $\mu_a$. This two-step strategy allows for modular inversion, which can simplify implementation or analysis in some settings.

However, in our approach, we perform a single-step inversion of the full operator $F = M \circ M \circ \mathcal{H}$, directly optimizing over the optical parameters (e.g., $\mu_a$) with respect to the discrepancy between the simulated and measured data. This strategy avoids intermediate reconstructions of $H$ and leverages automatic differentiation and adjoint-based gradients to compute sensitivities of the loss functional with respect to optical parameters. It also enables a tighter coupling between the acoustic and optical models,

which can improve reconstruction accuracy and stability in practical scenarios.

The complete forward process in QPAT can be understood as a composition of mappings between functional spaces that represent different physical stages of the imaging pipeline. First, we define the absorption coefficient $\mu_a(x) \in \mathcal{F}_{\text{opt}}$, which belongs to the optical parameter space.

Solving the stationary radiative transfer equation (RTE) for a given $\mu_a$ yields the total light fluence $\Phi_{\text{total}}(x)$, defined as the angular integral of the photon density.

This fluence, together with the absorption coefficient, defines the photothermal source via

$$H(x) = \mu_a(x) \cdot \Phi_{\text{total}}(x),$$

which serves as the initial pressure distribution for the acoustic wave equation. The measured acoustic data is then obtained by applying the wave operator $\mathcal{W}$ to the initial pressure:

$$p(t, \phi) = \mathcal{W}(H).$$

The complete forward process can therefore be summarized as the composition:

$$\mu_a \mapsto \mu_a \cdot \Phi_{\text{total}} \mapsto H \mapsto \mathcal{W}(H) \mapsto M[\mathcal{W}(H)] = p(t, \phi),$$

where $M$ denotes the measurement operator applied to the simulated pressure field.

We are given measured data $y_{\text{meas}} \in \mathbb{R}^{N_\phi \cdot N_t}$ and aim to recover the absorption coefficient $\mu_a$, typically via adjoint-based optimization. This involves backpropagating the data through the acoustic and optical adjoint problems to compute the gradient of the loss functional with respect to $\mu_a$.

# 7. Variational Regularization

In inverse problems such as quantitative photoacoustic tomography, the objective is to recover an unknown spatially varying coefficient, specifically the optical absorption coefficient $\mu_a(\mathbf{x})$, from indirect measurements of the resulting acoustic field, denoted by $y_{\text{meas}}(\mathbf{x})$. Let $F := M \circ U \circ \mathcal{H}$ denote the full forward operator that maps the absorption coefficient $\mu_a$ to the predicted acoustic measurements $p_{\text{pred}} = F(\mu_a)$.

A conventional approach is to formulate the reconstruction as an optimization problem that minimizes the discrepancy between the predicted and observed data in the least-squares sense:

$$\min_{\mu_a} \ \frac{1}{2}\|F(\mu_a) - y_{\text{meas}}\|^2_{L^2(\Omega)}.$$

This formulation is typically justified under the assumption of additive Gaussian noise and an accurate forward model, as it corresponds to the maximum likelihood estimator in that setting. It benefits from smoothness and convexity, making it well-suited for numerical optimization. However, its quadratic nature makes it highly sensitive to outliers and modeling errors, and it does not address ill-posedness or instability intrinsic to many inverse problems.

To address these challenges, we adopt a more general variational framework by replacing the squared error term with a robust data fidelity measure $D(\cdot, \cdot)$, and by including a regularization term $R(\mu_a)$ that encodes prior knowledge or enforces desirable properties in the solution. The resulting variational problem takes the form:

$$\min_{\mu_a} \ D(F(\mu_a), y_{\text{meas}}) + \lambda R(\mu_a),$$

where $\lambda \geq 0$ is a regularization parameter that balances the trade-off between fidelity to the data and regularity of the solution.

To define the specific form of the data fidelity term $D$, we employ the Huber loss function [5], which offers a principled compromise between least-squares sensitivity and robustness to outliers by interpolating between the quadratic and linear regimes.

In the context of inverse problems such as photoacoustic tomography, the choice of loss function plays a crucial role in balancing sensitivity to measurement errors and robustness to outliers. While the classical squared error loss (least-squares) is optimal for Gaussian noise, it can be overly sensitive to large deviations in the data. Conversely,

the absolute error (L1 norm) is more robust to outliers but introduces non-smoothness that complicates optimization.

Let $r(\mathbf{x}) := p_{\text{pred}}(\mathbf{x}) - y_{\text{meas}}(\mathbf{x})$ denote the residual between the predicted acoustic pressure $p_{\text{pred}}(\mathbf{x})$, computed from the forward model, and the measured acoustic data $y_{\text{meas}}(\mathbf{x})$. The Huber loss is defined pointwise as:

$$\ell_\delta(r) = \begin{cases} \frac{1}{2}r^2, & |r| \le \delta, \\ \delta\left(|r| - \frac{1}{2}\delta\right), & |r| > \delta, \end{cases}$$

where $\delta > 0$ is a tunable threshold parameter that governs the transition between the quadratic and linear regimes. This continuous transition allows the Huber loss to maintain sensitivity near the solution and suppress the influence of outliers.

More specifically, when $|r| \le \delta$, the loss behaves like a least-squares penalty, yielding a smooth and strongly convex region near the minimizer. In contrast, for $|r| > \delta$, it grows linearly, which reduces the influence of outliers and enhances robustness.

Its derivative, which is required for computing the gradient of the loss functional with respect to the model parameters, is given by:

$$\ell'_\delta(r) = \begin{cases} r, & |r| \le \delta, \\ \delta \cdot \text{sign}(r), & |r| > \delta. \end{cases}$$

The Huber loss is continuously differentiable due to its smooth transition between quadratic and linear regions, making it particularly suitable for gradient-based optimization methods.

To define the predicted pressure, we solve the forward radiative transfer equation (RTE) to obtain the directional photon densities $\Phi_i(\mathbf{x})$, from which we compute the total fluence:

$$\Phi_{\text{total}}(\mathbf{x}) = \sum_{i=1}^{n} w_i \Phi_i(\mathbf{x}),$$

using uniform quadrature weights $w_i$ as defined above.

The absorbed optical energy, which serves as the initial condition for the acoustic simulation, is given by:

$$H(\mathbf{x}) = \mu_a(\mathbf{x}) \cdot \Phi_{\text{total}}(\mathbf{x}).$$

The full predicted acoustic data is then computed as:

$$p_{\text{pred}} = M \circ U \circ \mathcal{H}(\mu_a) = F(\mu_a).$$

The objective functional is then given by:

$$\mathcal{J}[\mu_a] := \sum_{\mathbf{x}\in\Omega_h} \ell_\delta \left( p_{\text{pred}}(\mathbf{x}) - y_{\text{meas}}(\mathbf{x}) \right),$$

where $\ell_\delta$ denotes the Huber loss function, as defined previously, and $\Omega_h$ is the computational grid.

Minimizing $\mathcal{J}[\mu_a]$ leads to an estimate of the absorption coefficient that best explains the measured acoustic response. Thanks to the Huber loss, this formulation offers a robust and smooth optimization landscape, which ensures practical numerical stability even in the presence of measurement noise or modeling inaccuracies.

# 8. Gradient Computation

To enable gradient-based minimization of the objective functional $\mathcal{J}[\mu_a]$, we need to compute its gradient with respect to the absorption coefficient field $\mu_a(\mathbf{x})$. Due to the implicit dependence of the fluence $\Phi(\mathbf{x}, \theta)$ on $\mu_a$ through the forward model, a direct computation of the derivative is inefficient and often impractical. Instead, we employ the adjoint-state method, which enables efficient evaluation of the gradient by solving an additional linear system (the adjoint system).

The derivation proceeds as follows:

1. **Optical Forward Problem:** Given a current estimate of the absorption coefficient vector $\mu_a \in \mathbb{R}^N$, solve the discretized transport model:

$$A_{\text{total}}[\mu_a] \cdot \mathbf{\Phi} = \mathbf{q},$$

where $\mathbf{\Phi} \in \mathbb{R}^{nN}$ is the stacked vector of directional photon densities:

$$\mathbf{\Phi} = \begin{bmatrix} \Phi_1 \\ \Phi_2 \\ \vdots \\ \Phi_n \end{bmatrix}, \quad \Phi_i \in \mathbb{R}^N.$$

The total fluence is computed as:

$$\Phi_{\text{total}} = \sum_{i=1}^{n} w_i \Phi_i,$$

with $w_i$ as quadrature weights.

2. **Acoustic Forward Problem:** Using the absorbed energy $H = \mu_a \odot \Phi_{\text{total}}$, solve the acoustic wave equation to simulate pressure propagation. The governing equation is:

$$\partial_t^2 p(x,t) - c^2 \Delta p(x,t) = 0, \quad p(x,0) = H(x), \quad \partial_t p(x,0) = 0,$$

where $c$ is the speed of sound. This equation is discretized and solved using a

numerical wave solver (see Section 6) to compute the pressure field $p(x,t)$ in the domain.

Here, $\odot$ denotes element-wise (Hadamard) multiplication between vectors of size $N$.

The predicted pressure signal at the detector positions is then obtained by evaluating $p(x,t)$ at the measurement points and times:

$$p_{\text{pred}} = \mathcal{W}(H),$$

where $\mathcal{W}$ represents the numerical operator implementing wave propagation and extraction of the signal on the detection surface.

This two-stage forward computation—first solving the optical transport equation, then simulating acoustic wave propagation—reflects the physical process of QPAT. It is essential for correctly formulating the gradient of the loss functional via the adjoint-state method, which requires differentiating through both the optical and acoustic models.

3. **Residual and Huber Gradient:** Given the predicted pressure field $p_{\text{pred}}$ and the measured data $y_{\text{meas}}$, we define the residual as

$$r = p_{\text{pred}} - y_{\text{meas}},$$

where $r(x,t)$ is defined on the measurement domain, i.e., over the detector space $Y_{\text{meas}}$, which corresponds to spatial locations $x \in \partial B_R$ (where detectors are placed) and times $t \in [0,T]$, with $T$ denoting the final measurement time.

To obtain a spatially defined residual vector suitable for use in the gradient computation, we apply a temporal reduction operator $\mathcal{T}$, which aggregates the time-resolved residual into a scalar value per spatial location. Specifically, let

$$\mathcal{T}[r](x_j)$$

denote a temporal aggregation (e.g., time integration, averaging, or pointwise evaluation) of the signal $r(x_j,t)$ over the measurement interval.

Apply the Huber loss derivative pointwise to compute:

$$g_\delta(x_j) = \begin{cases} \mathcal{T}[r](x_j), & \text{if } |\mathcal{T}[r](x_j)| \le \delta, \\ \delta \cdot \operatorname{sign}(\mathcal{T}[r](x_j)), & \text{if } |\mathcal{T}[r](x_j)| > \delta. \end{cases}$$

where $x_j \in \Omega$, $j = 1, \ldots, N$, denote the spatial grid nodes.

This yields the vector $g_\delta \in \mathbb{R}^N$, representing the pointwise gradient of the Huber loss. It will be used in constructing the adjoint right-hand side, and is defined at each spatial grid node $x_j$.

4. **Construction of the Adjoint Right-Hand Side:** The right-hand side of the adjoint equation reflects the derivative of the data misfit with respect to the forward solution $\mathbf{\Phi}$. Since the predicted pressure depends linearly on the fluence $\Phi_i$, and the residual is measured pointwise, the sensitivity with respect to each direction is identical up to a scaling.

   Thus, the adjoint right-hand side is constructed by replicating the same vector $\mu_a \odot g_\delta$ across all $n$ angular directions:

$$\mathrm{RHS}_\lambda = \frac{1}{n} \begin{bmatrix} \mu_a \odot g_\delta \\ \mu_a \odot g_\delta \\ \vdots \\ \mu_a \odot g_\delta \end{bmatrix} \in \mathbb{R}^{nN}.$$

5. **Adjoint Solve:** Solve the transpose of the forward system:

$$A_{\text{total}}^\top \lambda = \mathrm{RHS}_\lambda,$$

   Here, $\lambda \in \mathbb{R}^{nN}$ is the adjoint variable associated with the forward photon density vector $\mathbf{\Phi}$, and encodes the sensitivity of the loss functional to changes in each directional component of the fluence.

   It is decomposed into directional components:

$$\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}, \quad \lambda_i \in \mathbb{R}^N.$$

   Each vector $\lambda_i$ corresponds to the adjoint field in direction $i$, and the full vector $\lambda$ contains stacked components for all directions.

6. **Gradient Evaluation:** To minimize the objective functional $\mathcal{J}[\mu_a]$, we need to compute its gradient with respect to the absorption coefficient $\mu_a$. Given that the predicted pressure field depends on $\mu_a$ through a composition of the optical and acoustic models, direct computation of the full Jacobian is computationally infeasible.

Instead, we employ the *adjoint-state method*, a widely used technique in PDE-constrained optimization. This method allows for efficient gradient computation by solving two PDEs: one forward problem to compute the predicted pressure $p_{\text{pred}}$, and one adjoint (backward) problem to propagate the data misfit backward through the model.

The adjoint-state method avoids explicitly forming the Jacobian of the forward operator and instead uses the solution of the adjoint system to project the residual sensitivity onto the parameter space of $\mu_a$. This leads to an efficient and scalable way to evaluate $\nabla \mathcal{J}[\mu_a]$, even in high-dimensional settings.

Before computing the gradient, a full forward simulation must be carried out. First, the optical forward problem is solved for all angular directions $\boldsymbol{\theta}_i$, yielding the directional fluence fields $\Phi_i$ and the total fluence $\Phi_{\text{total}} = \sum_i w_i \Phi_i$.

Then, the absorbed energy distribution is computed as $H = \mu_a \odot \Phi_{\text{total}}$. After that, the acoustic wave equation is solved with $H$ as the initial pressure, resulting in the predicted pressure signal $p_{\text{pred}} = \mathcal{W}(H)$, where $\mathcal{W}$ denotes the numerical wave propagation operator.

Thus, the full forward modeling pipeline can be expressed as the composition:

$$\mu_a \mapsto \Phi_{\text{total}}[\mu_a] \mapsto H = \mu_a \odot \Phi_{\text{total}} \mapsto p_{\text{pred}} = \mathcal{W}(H).$$

Recall that this composition includes both an explicit and implicit dependence on $\mu_a$, via $H = \mu_a \odot \Phi_{\text{total}}$ and $\Phi_{\text{total}} = \Phi_{\text{total}}[\mu_a]$, respectively.

To obtain the gradient $\nabla_{\mu_a} \mathcal{J}$, we differentiate the loss with respect to $\mu_a$, accounting for both the explicit dependence of $p_{\text{pred}}$ on $\mu_a$ via the term $H = \mu_a \odot \Phi_{\text{total}}$, and the implicit dependence of $\Phi_{\text{total}}$ on $\mu_a$, via the optical forward model.

As defined earlier, we use the Huber-weighted residual $g_\delta := \nabla \rho_\delta(r)$ to represent the sensitivity of the data fidelity term, where $r = p_{\text{pred}} - y_{\text{meas}}$ is the residual between predicted and measured acoustic signals.

Here, $\rho_\delta$ denotes the Huber loss function $\ell_\delta(r)$, as introduced in Chapter 7.

To backpropagate this error through the acoustic model, we solve the adjoint wave equation with source term $g_\delta$, yielding an adjoint pressure field $\lambda_H$. This field is then projected back through the optical system.

To formalize backpropagation via the acoustic model, we use an explicit expression for the adjoint operator of the acoustic forward problem $\mathcal{W}^*$. This operator maps the time-resolved residual signal measured at the detectors back to the domain $\Omega$, where it serves as a source term for the optical adjoint equation.

Let $r(x,t) \in L^2(\partial B_R \times (0,\infty))$ be the temporal residual signal recorded at the measurement surface $\partial B_R$. Then the adjoint field $\lambda_H = \mathcal{W}^* r$ is given by the following expression derived in [4]:

$$\mathcal{W}^* r(x) = \begin{cases} \dfrac{1}{2\pi} \displaystyle\int_{\partial B_R} \int_{|x-y|}^{\infty} \dfrac{\partial_t r(y,t)}{\sqrt{t^2 - |x-y|^2}} \, dt \, dS(y), & \text{for } d = 2, \\ \dfrac{1}{4\pi} \displaystyle\int_{\partial B_R} \dfrac{\partial_t r(y,|x-y|)}{|x-y|} \, dS(y), & \text{for } d = 3. \end{cases}$$

In our case, when the detectors are placed in a circle along the measurement boundary $\partial B_R$, the expression for the two-dimensional case ($d = 2$) is used. Here $r(y,t)$ is the pressure signal at the point $y \in \partial B_R$ at the time $t$.

This expression represents a continuous backprojection of the time derivative of the pressure residual onto each point $x \in \Omega$, accounting for the geometry of wave propagation. This representation follows from the theory of time-reversible wave propagation [8], and plays a key role in the method based on solving the adjoint problem.

The operator $\mathcal{W}^*$ effectively implements an integral backprojection over circles (in 2D) or spheres (in 3D) centered at the reconstruction points, which allows for efficient transfer of information from measurements back to the spatial domain.

Numerically, this operator is implemented as a filtered backprojection over circular wavefronts — in our implementation, this corresponds to the function `wavstarprop`. The resulting field $\lambda_H = \mathcal{W}^* r$ serves as the adjoint variable in the acoustic problem and is used as a source term in the optical adjoint equation.

The adjoint field $\lambda_H$ is then used to construct the right-hand side of the optical adjoint problem, which is solved independently for each illumination direction $\boldsymbol{\theta}_i$. This yields the directional adjoint fields $\lambda_i$, which represent the sensitivity of the loss functional to perturbations in the fluence $\Phi_i$.

The final expression for the gradient becomes:

$$\frac{\partial \mathcal{J}}{\partial \mu_a} = \sum_{i=1}^{n} \Phi_i \odot \lambda_i + \Phi_{\text{total}} \odot g_\delta,$$

This expression reflects the application of the chain rule to the composite forward model. The first term accounts for the sensitivity of the fluence to $\mu_a$ via the radiative transfer equation (RTE), and is obtained by solving the adjoint optical problem. The second term captures the explicit dependence of the absorbed energy on $\mu_a$, arising from the product $H = \mu_a \odot \Phi_{\text{total}}$.

This formulation ensures that both explicit and implicit dependencies of the

model on $\mu_a$ are captured, and the gradient is consistent with the chain rule through the composite forward model.

This expression is evaluated pointwise at each spatial node and yields the full gradient vector needed for gradient-based optimization.

This adjoint formulation is highly efficient and scalable with respect to the number of unknowns. It avoids differentiating through the full forward solver, making it suitable for high-dimensional inverse problems.

# 9. Iterative Scheme

To solve the inverse problem of estimating the spatially varying absorption coefficient $\mu_a(\mathbf{x})$, we adopt a first-order iterative optimization strategy based on gradient descent. This method is particularly suited for high-dimensional problems where second-order methods become computationally infeasible.

We begin by specifying an initial guess $\mu_a^{(0)} \in \mathbb{R}^N$ for the absorption coefficient, where $N$ denotes the number of spatial grid points. This initial estimate may be uniform (e.g., a constant value over the domain) or informed by prior knowledge.

At each iteration $k$, the absorption estimate is updated according to:

$$\mu_a^{(k+1)} = \mu_a^{(k)} - \alpha \cdot \nabla \mathcal{J}[\mu_a^{(k)}],$$

where $\alpha > 0$ is the step size or learning rate, which controls the magnitude of the update and $\nabla \mathcal{J}[\mu_a^{(k)}] \in \mathbb{R}^N$ is the gradient of the objective functional $\mathcal{J}$ with respect to $\mu_a$, evaluated at iteration $k$, and computed via the adjoint formulation.

In principle, optional post-processing techniques—such as block averaging over $b \times b$ patches or total variation (TV) denoising—can help suppress numerical artifacts and improve smoothness, we omit these steps in the current implementation and rely solely on the robustness of the Huber loss.

The optimization process continues until a stopping condition is met. Two criteria are considered in this work. First, the algorithm terminates if the number of iterations exceeds a predefined maximum threshold $k_{\max}$. Second, convergence is assumed if the relative change in the objective function between consecutive iterations falls below a user-specified tolerance $\epsilon$, that is, Second, convergence is assumed if the relative change in the objective function between consecutive iterations falls below a user-specified tolerance $\epsilon$, that is,

$$\frac{|\mathcal{J}[\mu_a^{(k+1)}] - \mathcal{J}[\mu_a^{(k)}]|}{\mathcal{J}[\mu_a^{(k)}]} < \epsilon.$$

To assess convergence and reconstruction quality, the loss $\mathcal{J}[\mu_a^{(k)}]$ is recorded at each iteration. This allows for plotting and analyzing the descent curve, which should ideally decrease monotonically.

The full iterative scheme can be summarized as follows:

**Algorithm: Gradient Descent for Absorption Reconstruction**

1. Initialize $\mu_a^{(0)} \leftarrow$ constant or prior;
2. For $k = 0, 1, 2, \ldots$:
   (a) Compute $\mathbf{\Phi}[\mu_a^{(k)}]$ via forward solve;
   (b) Compute absorbed energy $H^{(k)} = \mu_a^{(k)} \odot \Phi_{\text{total}}$;
   (c) Solve acoustic model: $p_{\text{pred}}^{(k)} = \mathcal{W}(H^{(k)})$;
   (d) Evaluate Huber gradient of residual and form adjoint RHS;
   (e) Solve adjoint system to obtain $\lambda$;
   (f) Compute gradient $\nabla \mathcal{J}[\mu_a^{(k)}] = \sum_i \Phi_i \odot \lambda_i + \Phi_{\text{total}} \odot g_\delta$;
   (g) Update: $\mu_a^{(k+1)} = \mu_a^{(k)} - \alpha \cdot \nabla \mathcal{J}$;

Optional: Apply projection or denoising to $\mu_a^{(k+1)}$;

   (h) Compute loss $\mathcal{J}^{(k+1)}$;
   (i) If convergence or max iteration reached: stop.

This algorithm forms the core optimization routine for solving the inverse problem in quantitative photoacoustic tomography.

# 10. Numerical Results

In this study, a numerical model of photothermal imaging was implemented, in which light transport and acoustic propagation are modeled using the Discrete Ordinates Method (DOM) and the wave equation, respectively.

This section presents numerical experiments on the reconstruction of the absorption coefficient map $\mu_a$ based on simulated photoacoustic data. A modified Shepp–Logan phantom is used to define the $\mu_a$ distribution, scaled and embedded into the simulation domain. Two cases are considered to analyze the influence of both the optical properties of the medium—specifically, absorption contrast—and the illumination geometry, i.e., the source location, on the reconstruction results. The light source is modeled as a uniform distribution along a horizontal line at height $y = 0.85$, with constant intensity.

In the numerical experiments, a two-dimensional square domain is considered, in which the propagation of light and acoustic waves is simulated within the framework of quantitative photoacoustic tomography (QPAT). Light transport is described by the radiative transfer equation, discretized using the Discrete Ordinates Method (DOM) with eight angular directions. The fluence distribution is computed taking into account both absorption and scattering, with the scattering coefficient fixed at $\mu_s = 50$, and the absorption distribution specified by a spatially inhomogeneous function $\mu_a(x, y)$.

The assembly of the transport and scattering operators is performed using sparse matrices, and the forward problem is solved via LU decomposition using the `spsolve` function from the `SciPy` library. The propagation of acoustic waves is modeled using the spectral-temporal method. In this case, the absorption coefficient values at the boundary of the region are assumed to be known and fixed to zero; in the reconstruction, they are treated as constraints.

The solution of the radiative transfer equation is obtained using a matrix approach: the discretized system is formed using directional transfer operators along eight angular directions, as well as absorption and scattering operators. The angular redistribution is modeled using the Henyey–Greenstein phase function with anisotropy coefficient $g = 0.85$. The final fluence distribution $\Phi(x, y)$ is calculated as the sum of contributions from all directions:

$$\Phi(x, y) = \sum_{i=1}^{n} w_i \Phi_i(x, y),$$

where $n = 8$ is the number of discrete directions, and $w_i$ are the corresponding weighting coefficients by solid angle.

The absorbed optical energy, which serves as the source of acoustic pressure, is defined as:

$$H(x, y) = \mu_a(x, y) \cdot \Phi(x, y).$$

The subsequent propagation of acoustic waves excited by the distribution $H(x, y)$ is simulated numerically using the wave equation in the time domain. To avoid boundary reflections and ensure accurate wave propagation, the computational domain is extended by zero-padding the initial pressure distribution beyond the physical phantom. The resulting acoustic measurements are then used to reconstruct the absorption coefficient map $\mu_a(x, y)$ by solving a variational optimization problem.

A total of $N_\phi = 100$ detectors are uniformly distributed on a circle of radius $R_0 = 1.0$. For each detector, a time-dependent pressure signal is recorded, forming a two-dimensional measurement matrix $p(t, \phi)$, where $t \in [0, T_0]$, $\phi \in [0, 2\pi)$. The number of time steps is set to $N_t = 300$, and the final simulation time is $T_0 = 2.2$.

As a result, a synthetic dataset is generated that mimics measurements typically obtained in quantitative photoacoustic tomography (QPAT) experiments.

The goal of the inverse problem is to reconstruct the spatial distribution of the absorption coefficient $\mu_a(x, y)$ from simulated acoustic pressure measurements on a ring detector. To do this, an optimization problem is solved in which the functional reflects the discrepancy between the simulated and observed acoustic signals.

The Huber loss function introduced in Chapter 7 is used to compare the predicted pressure $p_{\text{pred}}(\mu_a)$ with the synthetic measurements $y_{\text{meas}}$. This loss is robust to noise and outliers. In all experiments, the threshold parameter is set to $\delta = 100$.

The gradient of the loss function is calculated using the adjoint method implemented through the `wavstarprop` function, which performs a temporal inversion of the residual signal (backprojection).

The reconstruction is performed using the gradient descent method with fixed hyperparameters: initial absorption value $\mu_a = 0.1$, number of iterations $N_{\text{iter}} = 10$, gradient descent step $\alpha = 0.1$.

Although total variational deviation (TV) regularization and block-averaged projections (Q1) were implemented in the experiment, they were intentionally disabled to evaluate the reconstruction quality without additional smoothing or structural priors.

In the first numerical experiment, the effect of spatial contrast in the absorption coefficient $\mu_a$ is investigated under fixed illumination geometry. The distribution $\mu_a(x, y)$ is based on a modified Shepp–Logan phantom, scaled to provide two absorption levels: the background value is set to 0.05, while the bright inclusions reach 0.15.

This allows us to simulate a realistic inhomogeneous medium with pronounced optical contrasts.

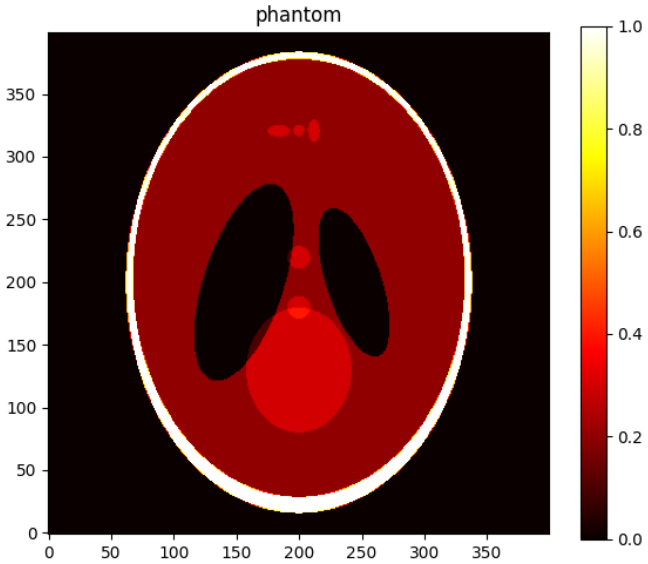


Figure 10.1: True absorption coefficient distribution $\mu_a(x, y)$, based on the modified Shepp–Logan phantom.

The illumination geometry and optical parameters of the medium correspond to those described above: the source is distributed along a horizontal line at height $y = 0.85$, with a scattering coefficient $\mu_s = 50$ and an anisotropy coefficient $g = 0.85$.

At the first stage, the fluence $\Phi(x, y)$ is calculated using the Discrete Ordinates Method (DOM), which provides eight directional components $\Phi_i(x, y)$.

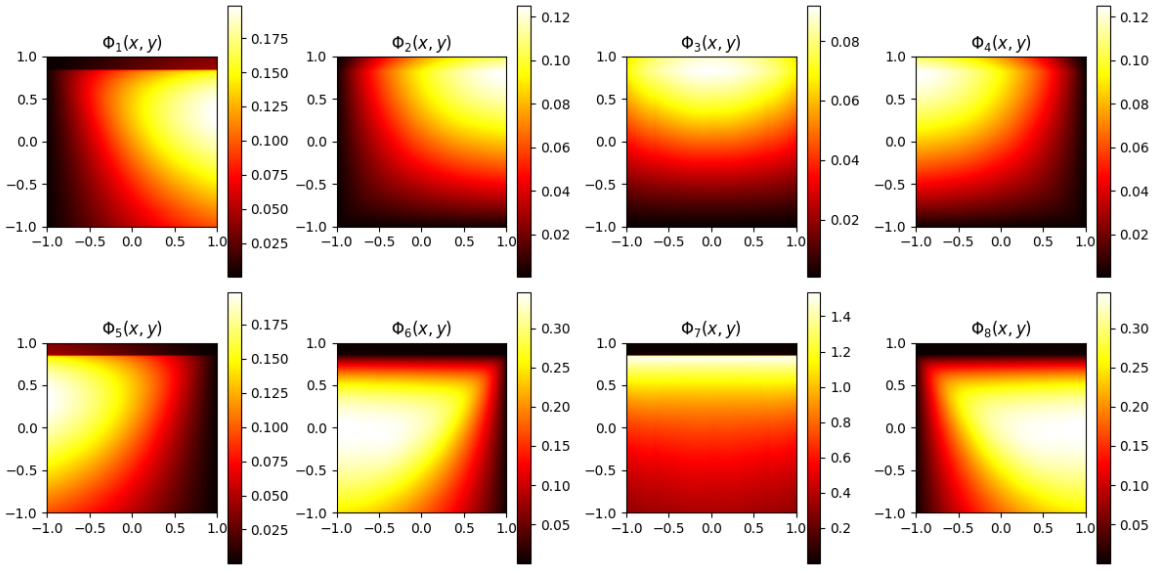


Figure 10.2: Directional components of fluence $\Phi_i(x, y)$, $i = 1, \ldots, 8$, obtained by the DOM method.

The total fluence field $\Phi(x, y)$ is then obtained by summing all directional contributions:

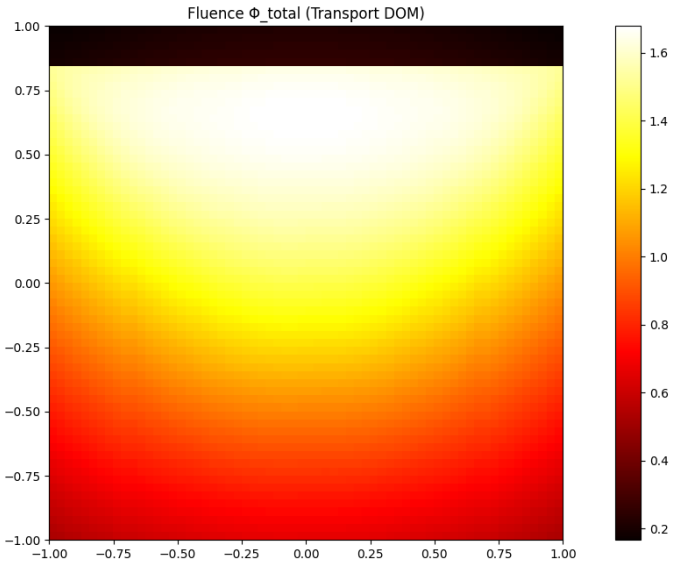


Figure 10.3: Total fluence distribution $\Phi(x, y)$.

The fluence field is used to compute the absorbed optical energy $H(x, y)$, which is the source of the acoustic wave equation:

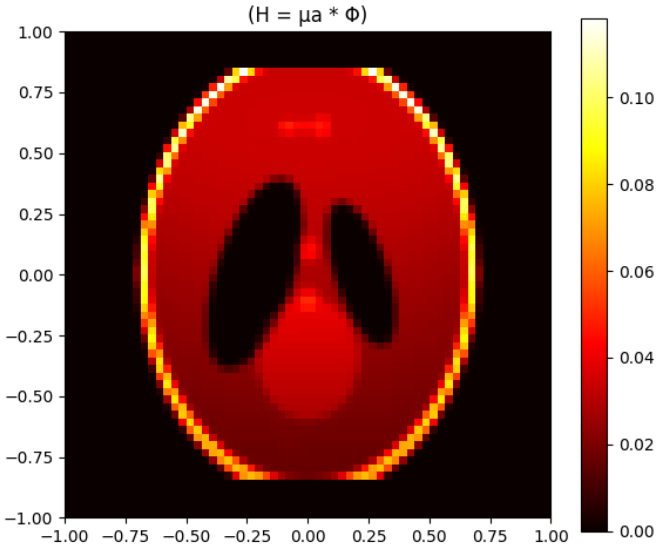


Figure 10.4: Distribution of absorbed energy $H(x, y)$, computed from the map $\mu_a(x, y)$ and the fluence $\Phi(x, y)$.

To illustrate the forward problem more clearly, Figure 10.12 shows the simulated photoacoustic data used for reconstruction, the true absorbed energy distribution $H(x, y)$, and the resulting wavefield at the final time step. These components together

define the input and intermediate stages of the quantitative photoacoustic tomography process.

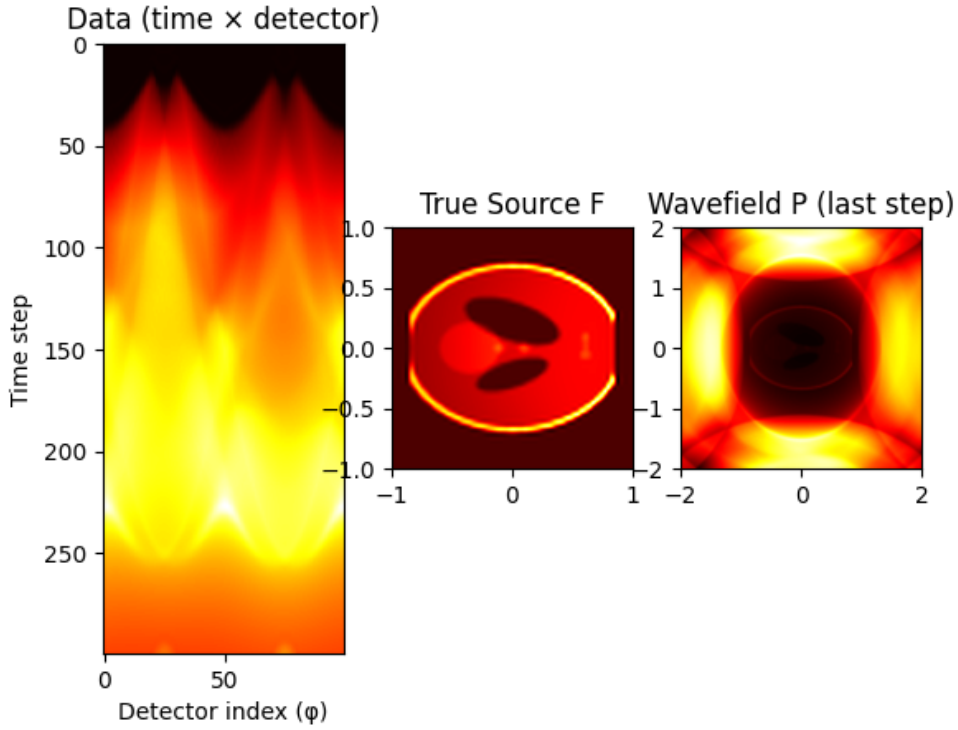


Figure 10.5: Left: synthetic data $y_{\text{meas}}(t, \phi)$; center: true absorbed energy $H(x, y)$; right: final step of the simulated wavefield $p(x, y, t)$.

The simulated acoustic wavefields are used to generate synthetic pressure data $p(t, \phi)$, recorded by a ring array of 100 detectors. Then, the absorption coefficient map $\mu_a(x, y)$ is numerically reconstructed by minimizing the Huber loss function. Reconstruction in this case is performed using the gradient descent method with the previously described parameters, without using additional regularization.

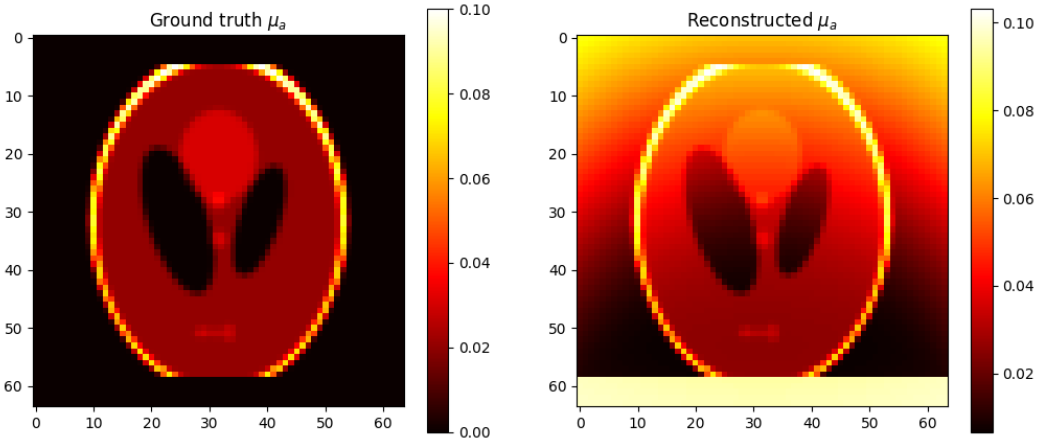


Figure 10.6: Comparison between the true and reconstructed absorption maps $\mu_a(x, y)$: left — ground truth, right — reconstruction.

The accuracy of the reconstruction is evaluated using the absolute error:

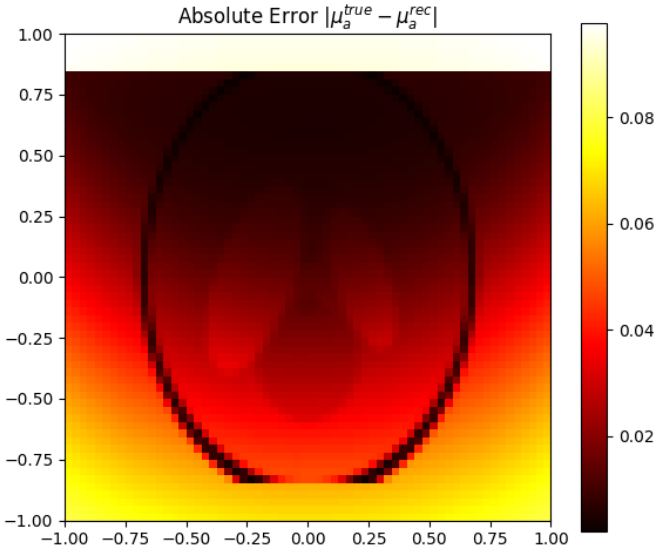


Figure 10.7: Absolute error map: $|\mu_a^{\text{true}}(x, y) - \mu_a^{\text{rec}}(x, y)|$.

The convergence of the optimization is illustrated by the evolution of the Huber loss function over the iterations:

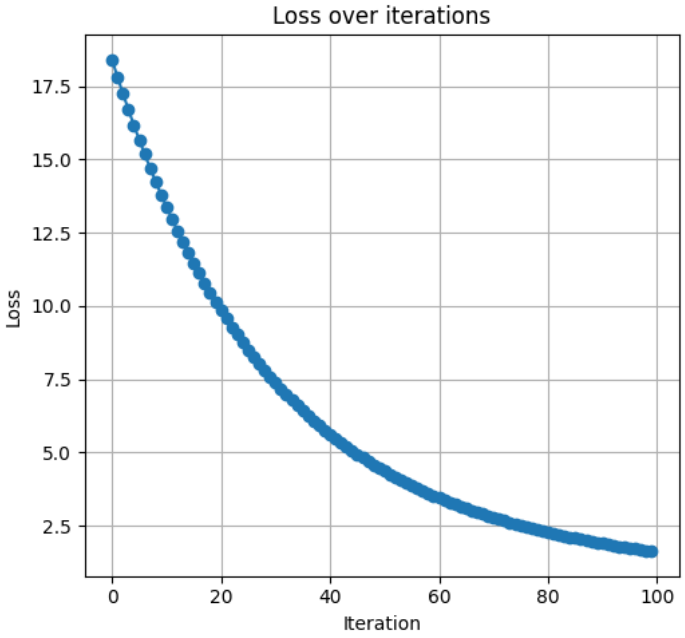


Figure 10.8: Huber loss function over iterations during the reconstruction process.

In the second numerical experiment, we study the influence of the illumination geometry on the accuracy of the absorption coefficient map reconstruction $\mu_a(x, y)$, assuming a fixed optical inhomogeneity of the medium. The distribution $\mu_a(x, y)$ remains unchanged from the previous experiment (see Figure 10.1) and is specified by a modified Shepp–Logan phantom. The main difference between the scenarios is the location and direction of the light source: in one case, it is uniformly distributed along a horizontal line at a height of $y = 0.85$; in the other, along the lower boundary of the region at $y = -0.85$, with light propagating upward into the domain.

The simulation and reconstruction results for the case of illumination from below ($y = -0.85$) are presented below.

The remaining simulation parameters are the same as in the previous case: the Discrete Ordinates Method (DOM) with eight directions is used, the scattering coefficient is set to $\mu_s = 50$, the Henyey–Greenstein phase function with anisotropy $g = 0.85$ is applied, and the acoustic measurements are simulated based on the wave equation. The reconstruction of $\mu_a(x, y)$ is performed by minimizing the Huber loss function using the same optimization parameters as before.

This approach allows us to analyze how changing the illumination geometry affects the fluence distribution, the shape of the acoustic signal, and the quality of reconstruction under conditions of the same internal structure of the medium.

At the first stage, the directional components of the fluence $\Phi_i(x, y)$ are calculated using the DOM method, corresponding to eight discrete light propagation directions:

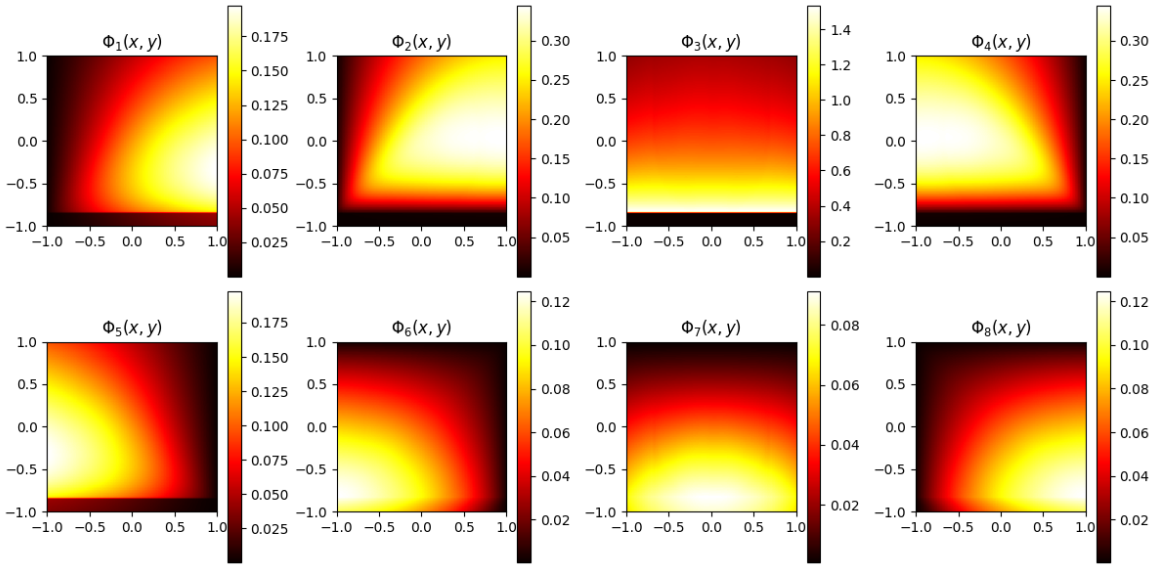


Figure 10.9: Directional components of the fluence $\Phi_i(x, y)$ for Case 2.

The total fluence field $\Phi(x, y)$ is then obtained:

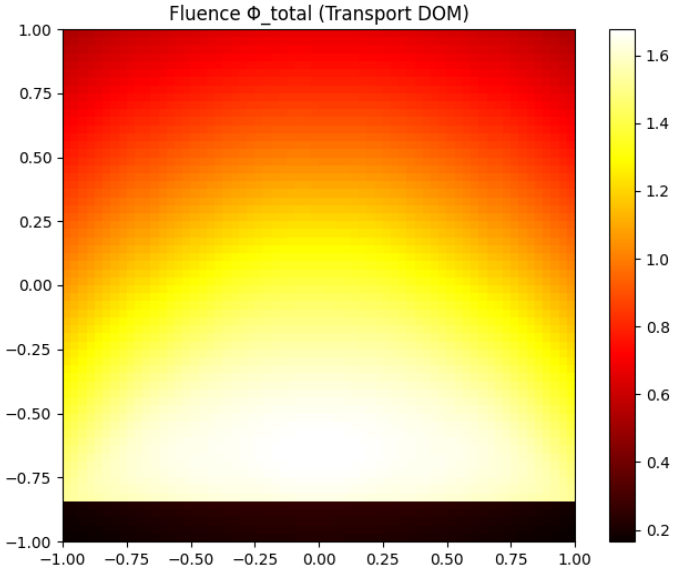


Figure 10.10: Total fluence field $\Phi(x, y)$ corresponding to illumination from below (Case 2).

The absorbed optical energy is calculated by the formula $H(x, y) = \mu_a(x, y) \cdot \Phi(x, y)$, and is visualized in the following figure:

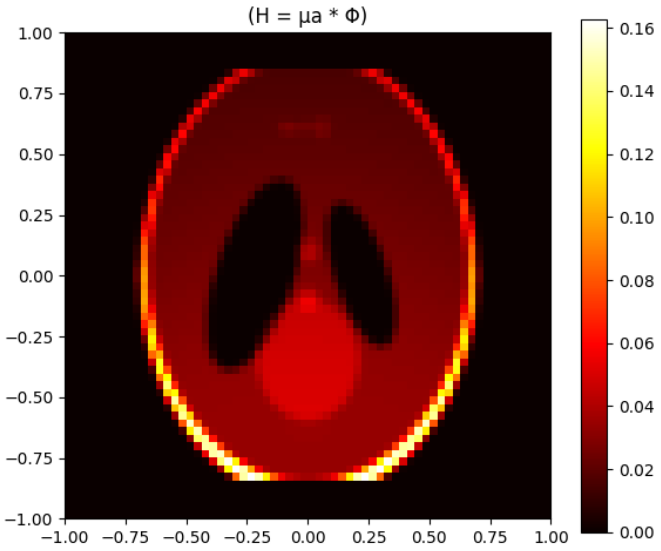


Figure 10.11: Distribution of absorbed energy $H(x, y)$ under lower illumination.

The resulting acoustic wavefield at the final simulation time step is illustrated below:

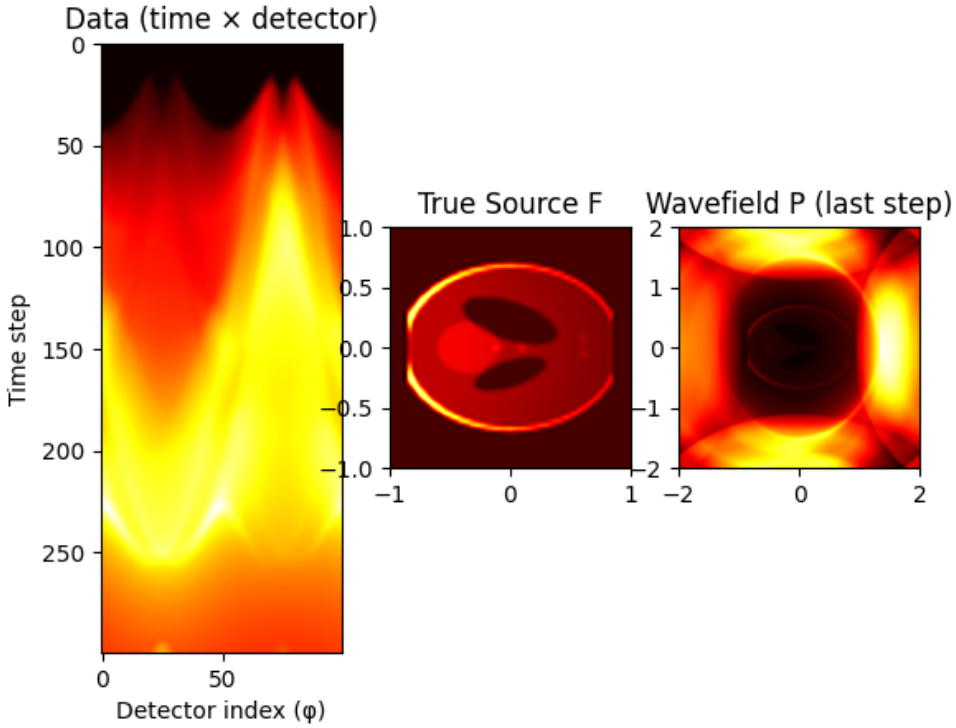


Figure 10.12: Left: synthetic data $y_{\text{meas}}(t, \phi)$; center: true absorbed energy $H(x, y)$; right: final step of the simulated wavefield $p(x, y, t)$.

The reconstruction results based on synthetic acoustic data:

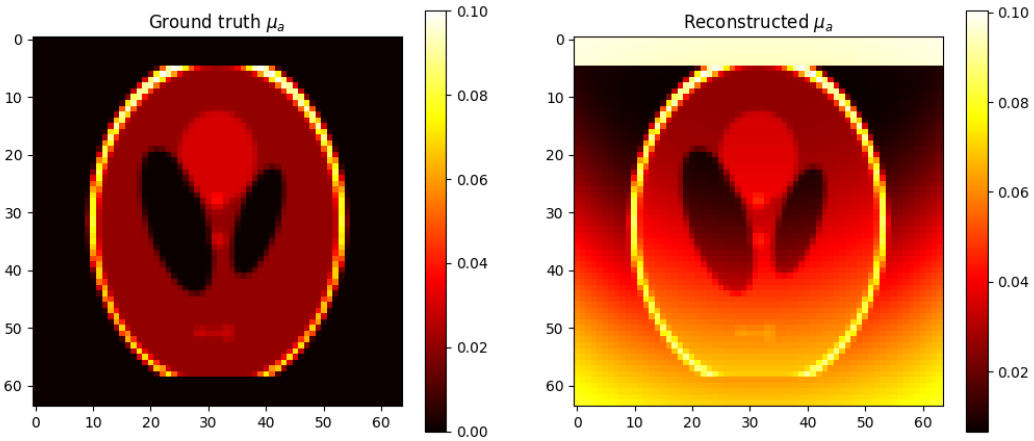


Figure 10.13: Comparison of the true and reconstructed distributions $\mu_a(x, y)$ for Case 2.

The reconstruction accuracy is assessed using the absolute error map:

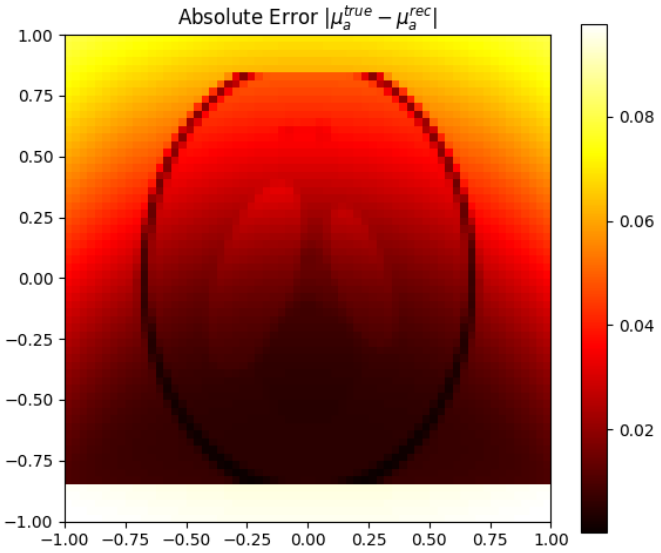


Figure 10.14: Absolute error map $|\mu_a^{true}(x, y) - \mu_a^{rec}(x, y)|$ for Case 2.

The convergence of the optimization process is illustrated by the evolution of the Huber loss function over iterations:

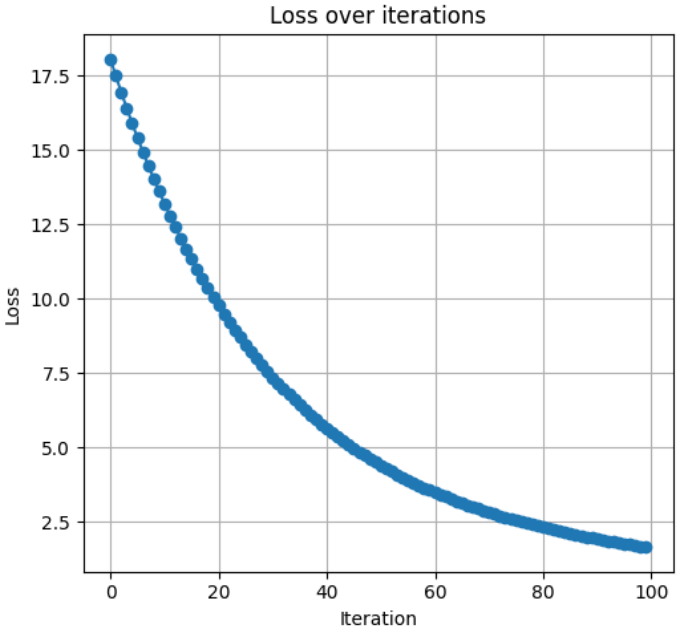


Figure 10.15: Huber loss function over iterations for bottom illumination.

# Bibliography


[1] Pedro J. Coelho. Discrete ordinates and finite volume methods. *Thermopedia*, 2011.

[2] David Finch, Markus Haltmeier, and Rakesh. Inversion of spherical means and the wave equation in even dimensions. *SIAM Journal on Applied Mathematics*, 68(2):392–412, 2007.

[3] David Finch and Sarah K Patch. Determining a function from its mean values over a family of spheres. *SIAM journal on mathematical analysis*, 35(5):1213–1240, 2004.

[4] Markus Haltmeier, Lukas Neumann, and Simon Rabanser. Single-stage reconstruction algorithm for quantitative photoacoustic tomography. *arXiv preprint arXiv:1501.04603*, 2015.

[5] Peter J. Huber. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 1964.

[6] K. D. Lathrop and B. G. Carlson. *Transport Theory: The Method of Discrete Ordinates*. Gordon and Breach Science Publishers, 1968.

[7] Randall J. LeVeque. *Numerical Methods for Conservation Laws*. Springer, 2 edition, 1992.

[8] Plamen Stefanov and Gunther Uhlmann. Thermoacoustic tomography with variable sound speed. *Inverse Problems*, 2009.

[9] Bradley E Treeby and Benjamin T Cox. k-wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave fields. *Journal of biomedical optics*, 15(2):021314–021314, 2010.