Master Thesis

Variational Regularization of Single-Stage Qualitative Photoacoustic Tomography

Student: Andrei Goncharenok

Supervisors:

Prof. Marcus Haltmeier

Prof. Lukas Neumann

University of Innsbruck

Prof. Giacomo Albi

University of Verona

Innsbruck & Verona, 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 by coupling 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 address the inverse problem of QPAT by developing and analyzing a numerical method derived from the discrete velocity model. This model improves the accuracy and stability of forward simulations. As a result, it allows for more precise reconstruction of tissue parameters from measured photoacoustic signals.

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 Φ. 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 θ' being scattered into direction θ. 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 the inflow boundary condition

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

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 of the stationary radiative transfer equation (RTE), the forward photon transport operator is defined as

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

which maps the prescribed optical sources and tissue parameters to the resulting photon density distribution.

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

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

For illustrative purposes or in numerical tests, we may set $\mu_a(x) = 1$, in which case the heating function simplifies to $H(x) = A\Phi(x)$.

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 $H(x)$ 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 $H(x)$.

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:

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

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 (μ_a, μ_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. These approaches 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:

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

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

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

In the radiative transfer equation, the angular variable $\boldsymbol{\theta}$ belongs to the unit circle S^1, which implies integration and differentiation over a continuous angular domain. From a functional analysis perspective, this domain is effectively infinite-dimensional 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 system of finite-dimensional equations.

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

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

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

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

The continuous function $\Phi(\mathbf{x}, \boldsymbol{\theta})$ is then approximated by its values along these fixed directions:

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

In this way, instead of a single function of both spatial and angular variables, we work with n scalar functions $\Phi_i(\mathbf{x})$, each describing the photon density traveling in direction $\boldsymbol{\theta}_i$. This discretization enables us to reduce the integral term in the RTE to a finite sum using an angular quadrature formula, thus forming a closed system of differential equations for $\Phi_i(\mathbf{x})$.

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:

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

In this discretized version of the radiative transfer equation, each equation has a contribution from the phase function of scattering. The expression k_{ij} is to be interpreted as the value of the phase function $k(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j)$, which represents the probability that a photon originally traveling in direction $\boldsymbol{\theta}_j$ will be scattered into direction $\boldsymbol{\theta}_i$.

The angular integral is replaced by a finite sum with quadrature weights. For a uniform discretization of the unit circle, these weights are constant and given by $w_j = \frac{2\pi}{n}$, reflecting the equal contribution of each direction to the approximation of the integral.

The source term of the equation is also discretized and given by $q_i(\mathbf{x}) := q(\mathbf{x}, \boldsymbol{\theta}_i)$, which corresponds to the intensity of radiation at point $\mathbf{x}$ in direction $\boldsymbol{\theta}_i$.

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 equations. 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:

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

We also define the scattering matrix $K \in \mathbb{R}^{n \times n}$ as the matrix of phase function

values:

$$K := (k_{ij}) = (k(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j)), \quad i, j = 1, \ldots, n,$$

and the diagonal matrix of quadrature weights:

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

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 differential operators acting along the respective directions. Combining them into a diagonal differential operator, we define:

$$D(\mathbf{x}) := \operatorname{diag}(\boldsymbol{\theta}_1 \cdot \nabla, \ldots, \boldsymbol{\theta}_n \cdot \nabla).$$

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

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

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 $D(\mathbf{x})$ 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 domain Ω 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 $\boldsymbol{\theta}_i \cdot \nabla\Phi(\mathbf{x})$ 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 Δx and Δy. 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 $\boldsymbol{\theta}_i \cdot \nabla \Phi_i(\mathbf{x})$ 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 adirectional 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.

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 stacked vector of angular photon densities $\boldsymbol{\Phi} \in \mathbb{R}^{nN}$, where each block A_i represents transport in the corresponding discrete direction $\boldsymbol{\theta}_i$.

This modular structure allows for efficient implementation using sparse matrix

algebra and is amenable to parallel computation.

As introduced above, each discrete transport direction $\boldsymbol{\theta}_i = (v_x^{(i)}, v_y^{(i)})$ gives rise to a directional transport operator $A_i \in \mathbb{R}^{N\times N}$, constructed using upwind finite differences depending on the sign of the directional components. These operators approximate the directional derivatives $\boldsymbol{\theta}_i \cdot \nabla$ and encode the advection of radiative energy in the corresponding directions.

The previously defined block-diagonal transport operator A_{dir} introduced earlier, acts on the global solution vector $\boldsymbol{\Phi} \in \mathbb{R}^{nN}$, which stacks all directional components $\Phi_i \in \mathbb{R}^N$ corresponding to each transport direction $\boldsymbol{\theta}_i$.

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

$$\mu_a = \begin{bmatrix} \mu_a(\mathbf{x}_1) \\ \mu_a(\mathbf{x}_2) \\ \vdots \\ \mu_a(\mathbf{x}_N) \end{bmatrix} \in \mathbb{R}^N, \quad \mu_s = \begin{bmatrix} \mu_s(\mathbf{x}_1) \\ \mu_s(\mathbf{x}_2) \\ \vdots \\ \mu_s(\mathbf{x}_N) \end{bmatrix} \in \mathbb{R}^N.$$

The combined operator describing attenuation of photon density due to local absorption and self-scattering (i.e., without direction change) 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 matrix of values of the scattering phase function $k(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j)$, normalized such that:

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

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

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

Each row of M_{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_{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 densities $\Phi_i(\mathbf{x})$, concatenated into a single vector. The vector $\mathbf{q} \in \mathbb{R}^{nN}$ represents the discretized source term, which may model, for instance, directional injection of light into a localized region.

The complete discretized system is built by assembling the directional transport operators, local absorption and scattering contributions, and inter-directional redistribution. The resulting global system matrix A_{total} reflects the coupling of angular and spatial components of the RTE.

The source term $\mathbf{q} \in \mathbb{R}^{nN}$ can be constructed based on the desired angular emission profile. For example, to simulate a directional Gaussian source, one may populate only the components of q_i corresponding to a selected direction $\boldsymbol{\theta}_i$. For isotropic illumination, the same spatial profile may be replicated across all directions.

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

$$\boldsymbol{\Phi} = \begin{bmatrix}\Phi_1^\top & \Phi_2^\top & \dots & \Phi_n^\top\end{bmatrix}^\top \in \mathbb{R}^{nN},$$

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

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

In the numerical implementation, this integral is approximated by a quadrature rule

over the discrete set of directions:

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

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

Since each directional component $\Phi_i(\mathbf{x})$ 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. It can thus be expressed in vector form as:

$$\mathbf{\Phi}_{\text{total}} := \sum_{i=1}^{n} w_i \Phi_i \in \mathbb{R}^N.$$

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:

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

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 resulting vector $\mathbf{H}$ is used as the initial pressure distribution $p_0(\mathbf{x}) = H(\mathbf{x})$ in the acoustic wave equation. This provides the coupling between the optical transport model and the acoustic response measured in quantitative photoacoustic tomography (QPAT).

This photothermal energy distribution $\mathbf{H}$ defines the initial pressure field in the acoustic wave equation, forming the crucial link between light transport and acoustic signal generation in Quantitative Photoacoustic Tomography (QPAT).

The complete forward process in QPAT can be interpreted as a composition of mappings between functional spaces. First, we define the absorption coefficient $\mu_0 := \mu_a(\mathbf{x}) \in \mathcal{F}_{\text{opt}}$, which belongs to the optical parameter space. The solution of the stationary radiative transfer equation is given by $\Phi_{\text{total}} = \mathcal{A}_{\text{RTE}}(\mu_0)$. The initial pressure distribution is then computed via the photothermal conversion operator as $p_0(\mathbf{x}) = \mu_0(\mathbf{x}) \cdot \Phi_{\text{total}}(\mathbf{x})$. Finally, the measured acoustic data $p(t, \phi)$ is obtained by applying the wave operator $\mathcal{W}$ to p_0, i.e., $p(t, \phi) = \mathcal{W}(p_0)$.

This yields the forward operator composition:

$$\mu_0 \mapsto \mu_0 \cdot \mathcal{A}_{\text{RTE}}(\mu_0) \mapsto p_0 \mapsto \mathcal{W}(p_0) = p(t, \phi).$$

For the inverse problem, we are given measured data $p_{\text{meas}}(t, \phi) \in \mathcal{F}_{t,\phi}$ and seek to recover μ_0, typically via adjoint-based optimization. This involves backpropagating the data through the acoustic adjoint, then optical adjoint, to compute gradients of the loss functional with respect to μ_0.

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 achieve high spatial accuracy, we employ a pseudo-spectral time-domain (PSTD) method based on the Fast Fourier Transform (FFT) for spatial derivatives, combined with a second-order central difference scheme for time integration. The following sections describe the discretization of the domain, the numerical scheme for wave propagation, and the extraction of simulated data at detector locations.

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 Δt 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

$$P_{ij}^n \approx p(x_i, y_j, t_n).$$

To approximate the spatial Laplacian Δp, 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_ϕ 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 Ω.

These points represent a uniform sampling of the boundary ∂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.

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.

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 $\mathcal{U} : \mathbb{R}^{N_x \times N_y} \to \mathbb{R}^{N_\phi \cdot N_t}$ such that:

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

The operator $\mathcal{U}$ 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 $\mathcal{U}$ serves as a building block in the composite model

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

where $\mathcal{H}$ models the optical excitation and fluence, and $\mathcal{M}$ 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_{meas}.

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 $y_{\text{target}}(\mathbf{x})$. Let $\mathcal{M}$ represent the forward operator that maps μ_a to predicted pressure data $p_{\text{pred}} = \mathcal{M}(\mu_a)$.

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:

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

This formulation assumes that the noise is Gaussian and the forward model is accurate. 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:

$$\min_{\mu_a} \ D(\mathcal{M}(\mu_a), y_{\text{target}}) + \lambda R(\mu_a),$$

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 choose the Huber loss, which offers a principled compromise between least-squares sensitivity and robustness to outliers.

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.

To address this trade-off, we employ the Huber loss function [1], which interpolates between the squared loss and the absolute loss. Let $r(\mathbf{x}) := p_{\text{pred}}(\mathbf{x}) - y_{\text{target}}(\mathbf{x})$ denote the residual between the predicted acoustic pressure $p_{\text{pred}}(\mathbf{x})$, computed from the forward model, and the measured acoustic data $y_{\text{target}}(\mathbf{x})$. 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:

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

The piecewise nature of the Huber loss ensures that it is continuously differentiable, which makes it particularly well-suited for use in gradient-based optimization algorithms.

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 is converted to pressure via the photoacoustic effect, is modeled as:

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

The objective functional is then given by:

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

where ℓ_δ denotes the Huber loss function, as defined previously, and Ω_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 minimize the objective functional $\mathcal{J}[\mu_a]$, we require its gradient with respect to the absorption coefficient field $\mu_a(\mathbf{x})$. Due to the implicit dependence of the fluence $\Phi(\mathbf{x}, \boldsymbol{\theta})$ on μ_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. **Forward Problem:** Given a current estimate of the absorption coefficient vector $\mu_a \in \mathbb{R}^N$, solve the discretized forward 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 density fields:

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

The total fluence is computed as:

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

with w_i as previously defined. The predicted pressure is:

$$p_{\text{pred}} = \mu_a \odot \Phi_{\text{total}},$$

with $\odot$ denoting element-wise multiplication.

2. **Residual and Huber Gradient:** Given the predicted pressure field p_{pred} and

measured data y_{target}, define the residual:

$$r = p_{\text{pred}} - y_{\text{target}}.$$

Apply the Huber loss derivative pointwise to compute:

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

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.

3. **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 Φ_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:

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

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

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

   where $\lambda \in \mathbb{R}^{nN}$ is the adjoint variable, 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 λ_i corresponds to the adjoint field in direction i, and the full vector λ contains stacked components for all directions.

5. **Gradient Evaluation:** The gradient of the objective functional with respect to the absorption vector $\mu_a \in \mathbb{R}^N$ is then computed as:

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

where the first term accounts for the implicit dependence of the fluence on μ_a (through the adjoint system), and the second term captures the explicit dependence of the predicted pressure on μ_a.

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 (e.g., a smoothed version of the measured data).

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 μ_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 ϵ, that is,

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

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 $p_{\text{pred}}^{(k)} = \mu_a^{(k)} \cdot \Phi_{\text{total}}$;
   (c) Evaluate Huber gradient of residual and form adjoint RHS;
   (d) Solve adjoint system to obtain λ;
   (e) Compute gradient $\nabla \mathcal{J}[\mu_a^{(k)}]$;
   (f) Update: $\mu_a^{(k+1)} = \mu_a^{(k)} - \alpha \cdot \nabla \mathcal{J}$;

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

   (g) Compute loss $\mathcal{J}^{(k+1)}$;
   (h) 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

This section presents numerical experiments on reconstructing the absorption coefficient map μ_a based on simulated photoacoustic data. The simulation is conducted in a two-dimensional setting using the discrete ordinates method (DOM) for light transport and the acoustic wave equation for pressure propagation.

For the numerical experiments, a two-dimensional square domain is considered, where the propagation of light and acoustic waves is simulated within the framework of quantitative photoacoustic tomography (QPAT). Light transport is modeled by the radiative transfer equation, discretized using the discrete ordinates method (DOM) with eight angular directions. The fluence distribution is computed by accounting for both absorption and scattering, where the scattering phase function follows the Henyey–Greenstein model with an anisotropy factor $g = 0.85$. Scattering is assumed to be isotropic and homogeneous, with a fixed scattering coefficient μ_s.

The light source is modeled as a uniform distribution along a horizontal line positioned near the upper boundary of the domain. A modified Shepp–Logan phantom, scaled and embedded into the domain, is used to define the absorption coefficient map μ_a.

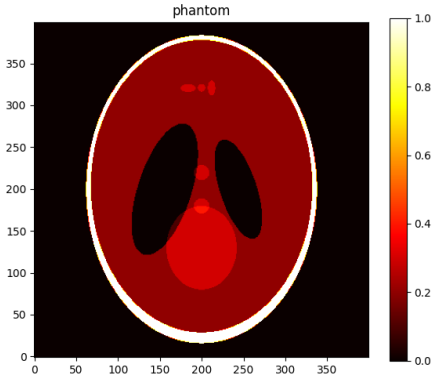

Figure 10.1: Modified Shepp–Logan phantom used as the ground truth absorption coefficient μ_a.

The absorption values at the boundary are assumed to be known and fixed at zero, and they are treated as constraints during the reconstruction process.

The assembly of the transport and scattering operators is performed using sparse matrices, and the forward problem is solved via LU factorization using the `spsolve` function from the SciPy library. Acoustic wave propagation is simulated using a spectral-temporal method, with pressure data recorded on a circular detector array surrounding the domain. These acoustic measurements are then used to reconstruct the absorption coefficient map via a variational optimization approach.

In the first stage of the simulation, the radiative transfer equation is solved using the discrete ordinates method (DOM) with a given spatial distribution of the absorption coefficient $\mu_a(x, y)$, a fixed scattering coefficient $\mu_s = 50$, and scattering anisotropy $g = 0.85$. The light source is modeled as a uniform illumination along a horizontal line at $y = 0.85$, implemented as a stripe with constant intensity.

The solution to the radiative transfer equation is obtained using a matrix-based approach: the discretized system is constructed using directional transport operators along eight angular directions, along with absorption and scattering operators. Angular redistribution is modeled using the Henyey–Greenstein phase function. The resulting fluence distribution $\Phi(x, y)$ is computed by summing contributions from all directions:

$$\Phi(x, y) = \sum_{i=1}^{N_{\text{dirs}}} \Phi_i(x, y) \cdot \Delta\Omega_i,$$

where $N_{\text{dirs}} = 8$ is the number of discrete directions, and $\Delta\Omega_i$ are the corresponding solid angle weights.

Next, the absorbed optical energy, which serves as the initial condition for the acoustic pressure, is computed as:

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

The propagation of acoustic waves generated by $H(x, y)$ is then simulated numerically using the time-domain wave equation. A larger computational domain is used to accommodate both the phantom and the sensor ring.

A total of $N_\phi = 100$ detectors are placed uniformly 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, mimicking measurement data typically obtained in quantitative photoacoustic tomography (QPAT) experiments.

The objective of the inverse problem is to recover the spatial distribution of the absorption coefficient $\mu_a(x, y)$ from simulated measurements of the acoustic pressure

on a detector ring. This is achieved by solving an optimization problem, where the cost function measures the discrepancy between the simulated and observed acoustic signals.

As a loss function, we employ the Huber loss, which is robust to noise and outliers. It is defined as:

$$L(a) = \sum_{i,j} \text{Huber}\left(p_{ij}^{\text{sim}}(a) - p_{ij}^{\text{obs}}\right),$$

where p_{ij}^{sim} denotes the simulated pressure signal at time t_i and detector ϕ_j, and p_{ij}^{obs} denotes the synthetic observed data computed from the true absorption distribution.

The Huber loss is defined as follows:

$$\text{Huber}(r) = \begin{cases} \frac{1}{2}r^2, & \text{if } |r| \leq \delta, \\ \delta(|r| - \frac{1}{2}\delta), & \text{if } |r| > \delta, \end{cases}$$

where $\delta = 100$ is the threshold parameter that controls the transition between the quadratic and linear regimes.

The gradient of the loss function is computed using the adjoint method (back-projection), implemented via the function `wavstarprop`, which performs time reversal of the residual signal.

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

Although TV (total variation) regularization and Q1 block-averaging projections were implemented in the framework, they were deliberately disabled in this experiment to assess the raw reconstruction performance without any additional smoothing or structural prior.

Bibliography

[1] Peter J. Huber. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1):73–101, 1964.