Single-stage reconstruction algorithm for quantitative photoacoustic tomopgraphy

Markus Haltmeier, Lukas Neumann, Simon Rabanser

APPLIEDMATHEMATICS

Single-stage reconstruction algorithm for quantitative photoacoustic tomography

Markus Haltmeier♣, Lukas Neumann◇ and Simon Rabanser♣◇

♣Department of Mathematics, University of Innsbruck
Technikestraße 13, A-6020 Innsbruck, Austria

◇Institute of Basic Sciences in Engineering Science, University of Innsbruck
Technikestraße 21a, A-6020 Innsbruck, Austria

E-mail: {markus.haltmeier,lukas.neumann,simon.rabanser}@uibk.ac.at

Abstract

The development of efficient and accurate image reconstruction algorithms is one of the cornerstones of computed tomography. Existing algorithms for quantitative photoacoustic tomography currently operate in a two-stage procedure: First an inverse source problem for the acoustic wave propagation is solved, whereas in a second step the optical parameters are estimated from the result of the first step. Such an approach has several drawbacks. In this paper we therefore propose the use of single-stage reconstruction algorithms for quantitative photoacoustic tomography, where the optical parameters are directly reconstructed from the observed acoustical data. In that context we formulate the image reconstruction problem of quantitative photoacoustic tomography as a single nonlinear inverse problem by coupling the radiative transfer equation with the acoustic wave equation. The inverse problem is approached by Tikhonov regularization with a convex penalty in combination with the proximal gradient iteration for minimizing the Tikhonov functional. We present numerical results, where the proposed single-stage algorithm shows an improved reconstruction quality at a similar computational cost.

Keywords. Quantitative photoacoustic tomography, stationary radiative transfer equation, wave equation, single-stage algorithm, inverse problem, parameter identification

AMS classification numbers. 44A12, 45Q05, 92C55.

1 Introduction

Photoacoustic tomography (PAT) is a recently developed medical imaging paradigm that combines the high spatial resolution of ultrasound imaging with the high contrast of optical imaging [7, 35, 52, 53, 54]. Suppose a semitransparent sample is illuminated with a short pulse of electromagnetic energy near the visible range. Then parts of the optical energy will be absorbed inside the sample which causes a rapid, non-uniform increase of temperature. The increase of temperature yields a spatially varying thermoelastic expansion which in turn induces an acoustic pressure wave (see Figure 1.1). The induced acoustic pressure wave is measured outside of the object of interest, and mathematical algorithms are used to recover an image of the interior.

Original (and also a lot of recent) work in PAT has been concentrated on the problem of reconstructing the initial pressure distribution, which has been considered as final image (see, for

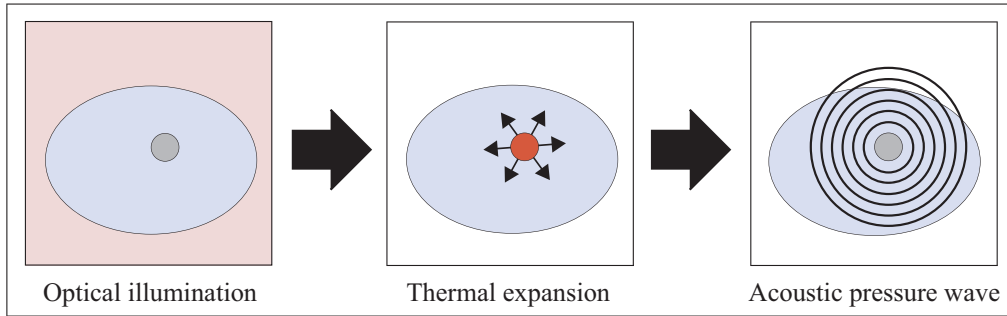| Optical illumination | Thermal expansion | Acoustic pressure wave |

Figure 1.1: BASIC PRINCIPLE OF PAT. A semitransparent sample is illuminated with a short optical pulse. Due to optical absorption and subsequent thermal expansion within the sample an acoustic pressure wave is induced. The acoustic pressure wave is measured outside of the sample and used to reconstruct an image of the interior.

example, [1, 9, 24, 25, 26, 27, 30, 31, 37, 34, 38, 45, 50, 54]). However, the recovered pressure distribution only provides indirect information about the investigated object. This is due to the fact, that the initial pressure distribution is the product of the optical absorption coefficient and the spatially varying optical intensity which again indirectly depends on the tissue parameters. As a consequence, the initial pressure distribution only provides qualitative information about the tissue-relevant parameters. Quantitative photoacoustic tomography (qPAT) addresses exactly this issue and aims at quantitatively estimating the tissue parameters by supplementing the wave inversion with an inverse problem for the light propagation in tissue (see, for example, [2, 5, 6, 10, 15, 16, 14, 19, 36, 39, 41, 44, 46, 47, 51, 56]).

To the best of our knowledge, all existing reconstruction algorithms for qPAT are currently performed via the following two-stage procedure: First, the measured pressure values are used to recover the initial pressure distribution caused by the thermal heating. In a second step, based on an appropriate light propagation model, the spatially varying tissue parameters are estimated from the initial pressure distribution recovered in the first step. However, any algorithm for solving an inverse problem requires prior knowledge abut the parameters to be recovered as well as partial knowledge about the noise. If one approaches qPAT via a two-stage approach, appropriate prior information for the acoustic inverse problem is difficult to model, because the initial pressure depends on parameters not yet recovered. This is particularly relevant for the case that the acoustic data can only be measured on parts of the boundary (limited-angle scenario), in which case the acoustic inverse problems is known to be severely ill-posed. Further, using a two-stage approach, only limited information about the noise for the optical problem is available.

In view of such shortcomings of the two-stage approach, in this paper we propose to recover the optical parameters directly from the measured acoustical data via a single-stage procedure. We work with the stationary radiative transfer equation (RTE) as model for light propagation. Our simulations show improved reconstruction quality of the proposed single-stage algorithm at a computational cost similar to the one of existing single-stage algorithms. Obviously our single-stage strategy can alternatively be combined with the diffusion approximation, which has also frequently been used in qPAT. Here we use the stationary RTE since it is the more realistic model for light propagation in tissue.

1.1 Mathematical modeling of qPAT

Throughout this paper, let $\Omega \subset \mathbb{R}^d$ denote a convex bounded domain with Lipschitz-boundary $\partial\Omega$, where $d \in \{2, 3\}$ denotes the spatial dimension. We model the optical radiation by a function $\Phi \colon \Omega \times \mathbb{S}^{d-1} \to \mathbb{R}$, where $\Phi(x, \theta)$ is the density of photons at location $x \in \Omega$ propagating in direction $\theta \in \mathbb{S}^{d-1}$. The photon density is supposed to satisfy the stationary radiative transfer equation (RTE)

$$\theta \cdot \nabla_x \Phi(x, \theta) + (\sigma(x) + \mu(x)) \Phi(x, \theta)$$
$$= \sigma(x) \int_{\mathbb{S}^{d-1}} k(\theta, \theta') \Phi(x, \theta') \mathrm{d}\theta' + q(x, \theta) \quad \text{for } (x, \theta) \in \Omega \times \mathbb{S}^{d-1}. \quad (1.1)$$

Here $\sigma(x)$ is the scattering coefficient, $\mu(x)$ is the absorption coefficient, and $q(x, \theta)$ is the photon source density. The scattering kernel $k(\theta, \theta')$ describes the redistribution of velocity directions of scattered photons due to interaction with the background. The stationary RTE (1.1) is commonly considered as a very accurate model for light transport in tissue (see, for example, [3, 18, 21, 33]).

In order to obtain a well-posed problem one has to impose appropriate boundary conditions. For that purpose it is convenient to split the boundary $\Gamma := \partial\Omega \times \mathbb{S}^{d-1}$ into inflow and outflow boundaries,

$$\Gamma_- := \left\{ (x, \theta) \in \partial\Omega \times \mathbb{S}^{d-1} \colon \nu(x) \cdot \theta < 0 \right\},$$
$$\Gamma_+ := \left\{ (x, \theta) \in \partial\Omega \times \mathbb{S}^{d-1} \colon \nu(x) \cdot \theta > 0 \right\},$$

with $\nu(x)$ denoting the outward pointing unit normal at $x \in \partial\Omega$. We then augment (1.1) by the inflow boundary conditions

$$\Phi|_{\Gamma_-} = f \quad \text{for some } f \colon \Gamma_- \to \mathbb{R}. \quad (1.2)$$

Under physically reasonable assumptions it can be shown that the stationary RTE (1.1) together with the inflow boundary conditions (1.2) is a well-posed problem. In Section 2.1 we apply a recent result of [20] that guarantees the well-posedness of (1.1), (1.2) even in the presence of voids (parts of the domain under consideration, where μ and σ vanish).

The absorption of photons causes a non-uniform heating of the tissue proportional to the total amount of absorbed photons,

$$h(x) := \mu(x) \int_{\mathbb{S}^{d-1}} \Phi(x, \theta) \mathrm{d}\theta \quad \text{for } x \in \Omega.$$

The heating in turn induces an acoustic pressure wave $p \colon \mathbb{R}^d \times (0, \infty) \to \mathbb{R}$. The initial pressure distribution is given by $p(\cdot, 0) = \Gamma h$, where Γ is the Grüneissen parameter describing the efficiency of conversion of heat into acoustic pressure. For the sake of simplicity we consider the Grüneissen parameter to be constant, known and rescaled to one. We further assume the speed of sound to be constant and also rescaled to one. The photoacoustic pressure then satisfies the following initial value problem for the standard wave equation,

$$\begin{cases} \partial_t^2 p(x, t) - \Delta p(x, t) = 0, & \text{for } (x, t) \in \mathbb{R}^d \times (0, \infty) \\ p(x, 0) = h(x), & \text{for } x \in \mathbb{R}^d \\ \partial_t p(x, 0) = 0, & \text{for } x \in \mathbb{R}^d. \end{cases} \quad (1.3)$$

The goal of qPAT is to reconstruct the parameters μ and σ from measurements of the acoustic pressure p outside Ω.

1.2 The inverse problem of qPAT

In the following we assume that acoustic measurements are available for multiple optical source distributions (illuminations). For that purpose, let (q_i, f_i) for $i = 1, \ldots N$ be given pairs of source patterns and boundary data. We use $\mathbf{T}_i$ to denote the operator that takes the pair (μ, σ) to the solution of the stationary RTE (1.1), (1.2) with q_i and f_i in place of q and f, and denote by

$$\mathbf{H}_i(\mu, \sigma)(x) := \mu(x) \int_{\mathbb{S}^{d-1}} \mathbf{T}_i(\mu, \sigma)(x, \theta) \mathrm{d}\theta \qquad \text{for } x \in \Omega$$

the operator describing the corresponding thermal heating. Further, we write $\mathbf{W}_{\Omega, \Lambda}$ for the operator that maps the initial data h to the solution $\mathbf{W}_{\Omega, \Lambda} h := p|_{\partial \Omega \times (0, \infty)}$ of the wave equation (1.3) restricted to the boundary $\partial \Omega$. Appropriate functional analytic frameworks for $\mathbf{T}_i$, $\mathbf{H}_i$ and $\mathbf{W}_{\Omega, \Lambda}$ will be given in Section 2, where we also study properties of these mapping.

The reconstruction problem of qPAT with multiple illuminations can be written in the form of a nonlinear inverse problem,

$$v_i = (\mathbf{W}_{\Omega, \Lambda} \circ \mathbf{H}_i)(\mu^\star, \sigma^\star) + z_i \quad \text{for } i = 1, \ldots, N. \tag{1.4}$$

Here v_i are the measured noisy data, the operators $\mathbf{W}_{\Omega, \Lambda} \circ \mathbf{H}_i$ model the forward problem of qPAT, z_i are the noise in the data, and $\mu^\star, \sigma^\star$ are the true parameters. The aim is to estimate the parameter pair $(\mu^\star, \sigma^\star)$ from given data v_i, and hence solving the inverse problem (1.4).

1.3 Outline of the paper

In this paper we address the inverse problem (1.4) by Tikhonov regularization,

$$\frac{1}{2} \sum_{i=1}^{N} \| (\mathbf{W}_{\Omega, \Lambda} \circ \mathbf{H}_i)(\mu, \sigma) - v_i \|^2 + \lambda \mathcal{R}(\mu, \sigma) \to \min_{(\mu, \sigma)},$$

where $\mathcal{R}$ is a convex penalty and $\lambda > 0$ is the regularization parameter. We show that Tikhonov regularization applied to single-stage qPAT is well-posed and convergent; see Theorem 3.2. For that purpose we derive regularity results for the heating operators $\mathbf{H}_i$ in Section 2. To establish such properties we use results for the stationary RTE derived recently in [20].

For numerically minimizing the Tikhonov functional we apply the proximal gradient algorithm (also named forward backward splitting); see Section 3.4. The proximal gradient algorithm, is an iterative scheme for minimizing functionals that can be written as the sum of a smooth and a convex part [13, 12]. For the classical two-stage approach in qPAT in combination with the diffusion approximation, the proximal gradient algorithm has recently been applied in [58]. Numerical results using the proximal gradient algorithm applied to our single-stage approach are presented in Section 4, where we also include a comparison with the two-stage approach. Of course, our single-stage approach can be combined with classical gradient or Newton-type schemes. The proximal gradient algorithm is our method of choice, since its is very flexible and fast, and can be applied for a large class of smooth or non-smooth penalties.

2 Analysis of the direct problem of qPAT

Before actually studying the inverse problem of qPAT we first make sure that the forward problem is well-posed in suitable spaces and that the data depend continuously on the parameters we

intend to reconstruct. For that purpose we review a recent existence and uniqueness result for the stationary RTE allowing for voids in the domain of interest [20]. The use of a-priori estimates will lead to differentiability results for the operators $\mathbf{T}_i$ and $\mathbf{H}_i$.

2.1 The stationary RTE

The stationary RTE has been studied in various contexts. The most prominent, apart from the transport of radiation in a scattering media, is reactor physics, where the equation is used in the group velocity approximation of the neutron transport problem. An extensive collection of results regarding applications as well as existence and uniqueness of solutions can be found in [18]. The analysis of the RTE becomes considerably more involved if internal voids, i.e. regions where scattering and absorption coefficient become zero, are allowed.

Suppose $1 \leq p \leq \infty$, and denote by $L^p(\Gamma_-, |\nu \cdot \theta|)$ the space of all measurable functions f defined on Γ_- for which

$$\|f\|_{L^p(\Gamma_-, |\nu \cdot \theta|)} := \begin{cases} \sqrt[p]{\int_{\Gamma_-} |\nu(x) \cdot \theta| \, |f(x,\theta)|^p \, \mathrm{d}(x,\theta)} & \text{if } p < \infty \\ \operatorname{ess\,sup}_{(x,\theta) \in \Gamma_-} \{|\nu(x) \cdot \theta| \, |f(x,\theta)|\} & \text{if } p = \infty \end{cases}$$

is finite. We write $W^p(\Omega \times \mathbb{S}^{d-1})$ for the space of all measurable functions defined on $\Omega \times \mathbb{S}^{d-1}$ such that

$$\|\Phi\|^p_{W^p(\Omega \times \mathbb{S}^{d-1})} := \|\Phi\|^p_{L^p(\Omega \times \mathbb{S}^{d-1})} + \|\theta \cdot \nabla_x \Phi\|^p_{L^p(\Omega \times \mathbb{S}^{d-1})} + \|\Phi|_{\Gamma_-}\|^p_{L^p(\Gamma_-, |\nu \cdot \theta|)}$$

is well defined and finite (with the usual modification for $p = \infty$). The subspace of all $\Phi \in W^p(\Omega \times \mathbb{S}^{d-1})$ with $\Phi|_{\Gamma_-} = 0$ will be denoted by $W_0^p(\Omega \times \mathbb{S}^{d-1})$. Further, for a given scattering kernel $k \in L^\infty(\mathbb{S}^{d-1} \times \mathbb{S}^{d-1})$ we write $\mathbf{K} \colon L^p(\Omega \times \mathbb{S}^{d-1}) \to L^p(\Omega \times \mathbb{S}^{d-1})$ for the corresponding scattering operator,

$$(\mathbf{K}\Phi)(x,\theta) = \int_{\mathbb{S}^{d-1}} k(\theta, \theta')\Phi(x,\theta')\mathrm{d}\theta' \qquad \text{for } (x,\theta) \in \Omega \times \mathbb{S}^{d-1}.$$

Throughout this article, the scattering kernel k is supposed to be symmetric and nonnegative, and to satisfy $\int_{\mathbb{S}^{d-1}} k(\theta, \theta') \, \mathrm{d}\theta' = 1$ for all $\theta \in \mathbb{S}^{d-1}$. This reflects the fact that $k(\,\cdot\,, \theta')$ is a probability distribution describing the redistributions of velocity directions due to interaction of the photons with the background. Under these assumption, the scattering operator $\mathbf{K}$ is easily seen to be linear and bounded.

Using the notation just introduced, the stationary RTE (1.1), (1.2) can be written in the compact form

$$\begin{cases} (\theta \cdot \nabla_x + (\mu + \sigma - \sigma\mathbf{K})) \Phi = q & \text{in } \Omega \times \mathbb{S}^{d-1} \\ \Phi|_{\Gamma_-} = f & \text{on } \Gamma_-. \end{cases} \tag{2.1}$$

By definition, a solution of the stationary RTE (1.1), (1.2) in W^p is any function $\Phi \in W^p(\Omega \times \mathbb{S}^{d-1})$ satisfying (2.1). The following theorem which has been derived very recently in [22] states that under physically reasonable assumptions there exists exactly one such solution, that further continuously depends on the source and the boundary data.

Theorem 2.1 (Existence and uniqueness of solutions in W^p). *Let $\overline{\mu}, \overline{\sigma}$ denote positive constants, let μ, σ be measurable functions satisfying $0 \leq \mu \leq \overline{\mu}$ and $0 \leq \sigma \leq \overline{\sigma}$, and let $1 \leq p \leq \infty$. Then, for any source pattern $q \in L^p(\Omega \times \mathbb{S}^{d-1})$ and all boundary data $f \in L^p(\Gamma_-, |\nu \cdot \theta|)$, the stationary*

RTE (1.1), (1.2) *admits a unique solution* $\Phi \in W^p(\Omega \times \mathbb{S}^{d-1})$. *Moreover, there exists a constant* $C_p(\overline{\mu}, \overline{\sigma})$ *only depending on* p, $\overline{\mu}$ *and* $\overline{\sigma}$, *such that the following a-priori estimate holds*

$$\|\Phi\|^p_{W^p(\Omega \times \mathbb{S}^{d-1})} \leq C_p(\overline{\mu}, \overline{\sigma}) \left(\|q\|_{L^p(\Omega \times \mathbb{S}^{d-1})} + \|f\|_{L^p(\Gamma_-, |\nu \cdot \theta|)} \right). \tag{2.2}$$

Proof. See [22]. □

2.2 The parameter-to-solution operator T for the stationary RTE

Throughout this subsection, let $1 \leq p \leq \infty$, and let $q \in L^\infty(\Omega \times \mathbb{S}^{d-1})$ and $f \in L^\infty(\Gamma_-, |\nu \cdot \theta|)$ be given source pattern and boundary data, respectively. Further, for fixed positive numbers $\overline{\mu}, \overline{\sigma} > 0$ we denote

$$\mathcal{D}_p := \{(\mu, \sigma) \in L^p(\Omega) \times L^p(\Omega) : 0 \leq \sigma \leq \overline{\sigma} \text{ and } 0 \leq \mu \leq \overline{\mu}\}. \tag{2.3}$$

Then $\mathcal{D}_p$ is a closed, bounded and convex subset of $L^p(\Omega) \times L^p(\Omega)$, that has empty interior in the case that $p < \infty$.

Definition 2.2 (Parameter-to-solution operator for the stationary RTE)**.** The parameter-to-solution operator for the stationary RTE is defined by

$$\mathbf{T} \colon \mathcal{D}_p \to W^p(\Omega \times \mathbb{S}^{d-1}) \colon (\mu, \sigma) \mapsto \Phi, \tag{2.4}$$

where Φ denotes the unique solution of (1.1), (1.2).

According to Theorem 2.1 the operator $\mathbf{T}$ is well defined. Note further, that $\mathbf{T}$ depends on p, q, f, $\overline{\mu}$ and $\overline{\sigma}$. Since these parameters will be fixed in the following and in order to keep the notation simple we will not indicate the dependence of $\mathbf{T}$ these parameter explicitly.

Now we are in the position to state continuity properties of $\mathbf{T}$ derived in [21]. We include a short proof of these results as its understanding is very useful for the derivation of similar properties of the operator describing the heating that we investigate in the following subsection.

Theorem 2.3 (Lipschitz continuity and weak continuity of $\mathbf{T}$)**.**

(a) *The operator* $\mathbf{T}$ *is Lipschitz-continuous.*

(b) *If* $1 < p < \infty$, *then* $\mathbf{T}$ *is sequentially weakly continuous.*

Proof. (a) Let $(\mu, \sigma), (\hat{\mu}, \hat{\sigma}) \in \mathcal{D}_p$ be two given pairs of absorption and scattering coefficients and denote by $\Phi := \mathbf{T}(\mu, \sigma)$ and $\hat{\Phi} := \mathbf{T}(\hat{\mu}, \hat{\sigma})$ the corresponding solutions of the stationary RTE. Since $q \in L^\infty(\Omega \times \mathbb{S}^{d-1})$ and $f \in L^\infty(\Gamma_-, |\nu \cdot \theta|)$, Theorem 2.1 implies that the difference $\hat{\Phi} - \Phi$ is an element of $W_0^\infty(\Omega \times \mathbb{S}^{d-1})$. Further, this difference is easily seen to satisfy

$$(\theta \cdot \nabla_x + \mu + \sigma - \sigma \mathbf{K})(\hat{\Phi} - \Phi) = (\mu - \hat{\mu})\hat{\Phi} + (\sigma - \hat{\sigma})\hat{\Phi} - (\sigma - \hat{\sigma})\mathbf{K}\hat{\Phi}.$$

Because $\mathbf{K}$ is a bounded linear operator on $L^p(\Omega \times \mathbb{S}^{d-1})$, the right hand side in the above equation is actually contained in $L^p(\Omega \times \mathbb{S}^{d-1})$. Therefore, a further application of Theorem 2.1 yields

$$\|\hat{\Phi} - \Phi\|_{W^p(\Omega \times \mathbb{S}^{d-1})} \leq C_p(\overline{\mu}, \overline{\sigma})\|\hat{\Phi}\|_{L^\infty(\Omega \times \mathbb{S}^{d-1})} \left(\|\mu - \hat{\mu}\|_{L^p(\Omega \times \mathbb{S}^{d-1})} + \|\mathbf{I} - \mathbf{K}\|_p \|\sigma - \hat{\sigma}\|_{L^p(\Omega \times \mathbb{S}^{d-1})} \right),$$

where $\mathbf{I}$ denotes the identity and $\| \cdot \|_p$ the operator norm on $L^p(\Omega \times \mathbb{S}^{d-1})$. Since $\|\hat{\Phi}\|_{L^\infty(\Omega \times \mathbb{S}^{d-1})}$ is bounded independently of $\hat{\Phi}$, this implies the Lipschitz continuity of $\mathbf{T}$.

6

(b) Let $(\mu_n, \sigma_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{D}_p$ that converges weakly to the pair $(\mu, \sigma) \in \mathcal{D}_p$, and denote by $\Phi_n = \mathbf{T}(\mu_n, \sigma_n)$ and $\Phi = \mathbf{T}(\mu, \sigma)$ the corresponding solutions of the stationary RTE. As in (a), one argues that the difference $\Phi_n - \Phi$ is contained in $W_0^\infty(\Omega \times \mathbb{S}^{d-1})$ and satisfies

$$(\theta \cdot \nabla_x + \mu + \sigma - \sigma \mathbf{K})(\Phi_n - \Phi) = (\mu - \mu_n)\Phi_n + (\sigma - \sigma_n)\Phi_n - (\sigma - \sigma_n)\mathbf{K}\Phi_n.$$

Now, from Theorem 2.1 it follows that $\theta \cdot \nabla_x + \mu + \sigma - \sigma \mathbf{K}$ is invertible as an operator from $W_0^p(\Omega \times \mathbb{S}^{d-1})$ to $L^p(\Omega \times \mathbb{S}^{d-1})$. Consequently, the inverse mapping $(\theta \cdot \nabla_x + \mu + \sigma - \sigma \mathbf{K})^{-1}$ is linear and bounded and in particular weakly continuous. It therefore remains to show that $(\mu - \mu_n)\Phi_n + (\sigma - \sigma_n)\Phi_n - (\sigma - \sigma_n)\mathbf{K}\Phi_n$ weakly converges to zero in $L^p(\Omega \times \mathbb{S}^{d-1})$. To see this, denote by $p_* = p/(p-1)$ the dual index and let $\varphi \in L^{p_*}(\Omega \times \mathbb{S}^{d-1})$ be any element in the dual of $L^p(\Omega \times \mathbb{S}^{d-1})$. By Fubini's theorem we have

$$\int_{\Omega \times \mathbb{S}^{d-1}} (\mu(x) - \mu_n(x))\,\Phi_n(x,\theta)\varphi(x,\theta)\,\mathrm{d}(x,\theta) = \int_\Omega (\mu(x) - \mu_n(x))\left(\int_{\mathbb{S}^{d-1}} \Phi_n(x,\theta)\varphi(x,\theta)\,\mathrm{d}\theta\right)\mathrm{d}x.$$

The averaging lemma (see, for example, [40]) implies that the averaging operator $\mathbf{A} \colon W^{p_*}(\Omega \times \mathbb{S}^{d-1}) \to L^{p_*}(\Omega) \colon \Phi \mapsto \int_{\mathbb{S}^{d-1}} \Phi(\,\cdot\,,\theta)\mathrm{d}\theta$ is compact for $1 < p_* < \infty$. Since $(\Phi_n)_{n \in \mathbb{N}}$ is bounded in $W^\infty(\Omega \times \mathbb{S}^{d-1}) \subset W^{p_*}(\Omega \times \mathbb{S}^{d-1})$, this implies that $\int_{\mathbb{S}^{d-1}} \Phi_n(\,\cdot\,,\theta)\varphi(\,\cdot\,,\theta)\mathrm{d}\theta$ converges to $\int_{\mathbb{S}^{d-1}} \Phi(\,\cdot\,,\theta)\varphi(\,\cdot\,,\theta)\mathrm{d}\theta$ with respect to $\|\cdot\|_{L^{p_*}(\Omega)}$. As $\mu_n \rightharpoonup \mu$ we can conclude that $(\mu - \mu_n)\Phi_n$ converges to zero weakly. In the same manner one shows $(\sigma - \sigma_n)\Phi_n \rightharpoonup 0$. Finally, the equality

$$\int_{\Omega \times \mathbb{S}^{d-1}} (\mu - \mu_n)(x)(\mathbf{K}\Phi_n)(x,\theta)\varphi(x,\theta)\mathrm{d}(x,\theta) = \int_\Omega (\mu - \mu_n)(x)\int_{\mathbb{S}^{d-1}} \Phi_n(x,\theta)(\mathbf{K}\varphi)(x,\theta)\mathrm{d}\theta\mathrm{d}x$$

and the use of similar arguments show that $(\sigma - \sigma_n)\mathbf{K}\Phi_n \rightharpoonup 0$. $\qquad\square$

For the solution of the inverse problem of qPAT we will make use of the derivative of $\mathbf{T}$ that we compute next. For that purpose we call $h \in L^p(\Omega \times \mathbb{S}^{d-1}) \times L^p(\Omega \times \mathbb{S}^{d-1})$ a feasible direction at $(\mu, \sigma) \in \mathcal{D}_p$ if there exists some $\epsilon > 0$ such that $(\mu, \sigma) + \epsilon h \in \mathcal{D}_p$. Due to the convexity of $\mathcal{D}_p$ we have $u + sh \in \mathcal{D}_p$ for all $0 < s \leq \epsilon$. The set of all feasible directions will be denoted by $\mathcal{D}_p(\mu, \sigma)$. One immediately sees that

$$\mathcal{D}_p(\mu, \sigma) = L^p(\Omega \times \mathbb{S}^{d-1}) \times L^p(\Omega \times \mathbb{S}^{d-1}) \quad \text{if } 0 < \mu < \overline{\mu} \text{ and } 0 < \sigma < \overline{\sigma}.$$

For $(\mu, \sigma) \in \mathcal{D}_p$ and any feasible direction $h \in \mathcal{D}_p(\mu, \sigma)$ we denote the one-sided directional derivative of $\mathbf{T}$ at (μ, σ) in direction h by

$$\mathbf{T}'(\mu, \sigma)(h) := \lim_{s \downarrow 0} \frac{\mathbf{T}((\mu, \sigma) + sh) - \mathbf{T}(\mu, \sigma)}{s}, \tag{2.5}$$

provided that the limit on the right hand side of (2.5) exists. If both limits $\mathbf{T}'(\mu, \sigma)(h)$ and $\mathbf{T}'(\mu, \sigma)(-h)$ exist and $h \mapsto \mathbf{T}'(\mu, \sigma)(h)$ is bounded and linear, we say that $\mathbf{T}$ is Gâteaux differentiable at (μ, σ) and call $\mathbf{T}'(\mu, \sigma)$ the Gâteaux derivative of $\mathbf{T}$ at (μ, σ).

Theorem 2.4 (Differentiability of $\mathbf{T}$). *For any $(\mu, \sigma) \in \mathcal{D}_p$, the one-sided directional derivative of $\mathbf{T}$ at (μ, σ) in direction $(h_\mu, h_\sigma) \in \mathcal{D}_p(\mu, \sigma)$ exists. Further, we have $\mathbf{T}'(\mu, \sigma)(h_\mu, h_\sigma) = \Psi$, where Ψ is the unique solution of*

$$\begin{cases} (\theta \cdot \nabla_x + (\mu + \sigma - \sigma \mathbf{K}))\,\Psi = -(h_\mu + h_\sigma - h_\sigma \mathbf{K})\,\mathbf{T}(\mu, \sigma) & \text{in } \Omega \times \mathbb{S}^{d-1} \\ \Psi|_{\Gamma_-} = 0 & \text{on } \Gamma_-. \end{cases} \tag{2.6}$$

If $0 < \mu < \overline{\mu}$ and $0 < \sigma < \overline{\sigma}$, then $\mathbf{T}$ is Gâteaux differentiable at (μ, σ).

Proof. Suppose $(\mu, \sigma) \in \mathcal{D}_p$ and let $h = (h_\mu, h_\sigma) \in \mathcal{D}_p(\mu, \sigma)$ be any feasible direction. For sufficiently small $s > 0$ write $\Phi_s := \mathbf{T}((\mu, \sigma) + sh)$ and $\Phi := \mathbf{T}(\mu, \sigma)$. As in the proof of Theorem 2.3 one shows that $\Psi_s := (\Phi_s - \Phi)/s$ is contained in $W_0^p(\Omega \times \mathbb{S}^{d-1})$ and solves the equation $(\theta \cdot \nabla_x + \mu + \sigma - \sigma\mathbf{K})\Psi_s = -(h_\mu + h_\sigma - h_\sigma\mathbf{K})\Phi_s$. Consequently the difference $\Psi_s - \Psi \in W_0^p(\Omega \times \mathbb{S}^{d-1})$ solves

$$(\theta \cdot \nabla_x + \mu + \sigma - \sigma\mathbf{K})(\Psi_s - \Psi) = -(h_\mu + h_\sigma - h_\sigma\mathbf{K})(\Phi_s - \Phi).$$

Application of the a-priori estimate of Theorem 2.1 shows the inequality $\|\Psi_s - \Psi\|_{W^p(\Omega \times \mathbb{S}^{d-1})} \leq C_p(\overline{\mu}, \overline{\sigma})\|\Phi_s - \Phi\|_{L^\infty(\Omega \times \mathbb{S}^{d-1})}(\|h_\mu\|_{L^p(\Omega \times \mathbb{S}^{d-1})} + \|h_\sigma\|_{L^p(\Omega \times \mathbb{S}^{d-1})})$ Together with the continuity of $\mathbf{T}$ this implies that the one-sided directional derivative $\mathbf{T}'(\mu, \sigma)(h)$ exists and is given by $\lim_{s \to 0} \Psi_s = \Psi$. Finally, if $0 < \mu < \overline{\mu}$ and $0 < \sigma < \overline{\sigma}$, then $h \mapsto \mathbf{T}'(\mu, \sigma)(h)$ is bounded and linear and therefore $\mathbf{T}$ is Gâteaux differentiable at (μ, σ). $\square$

Note that for any parameter pair $(\mu, \sigma) \in \mathcal{D}_p$, the solution of (2.6) depends linearly and continuously on $(h_\mu, h_\sigma) \in L^p(\Omega) \times L^p(\Omega)$. As a consequence, the one-sided directional derivative can be extended to a bounded linear operator

$$\mathbf{T}'(\mu, \sigma) \colon L^p(\Omega) \times L^p(\Omega) \to W^p(\Omega \times \mathbb{S}^{d-1}) \colon (h_\mu, h_\sigma) \mapsto \Psi, \tag{2.7}$$

where Ψ is the unique solution of (2.6). We will refer to this extension as the derivative of $\mathbf{T}$ at (μ, σ).

2.3 The operator H describing the heating

Throughout this subsection, let $q \in L^\infty(\Omega \times \mathbb{S}^{d-1})$ and $f \in L^\infty(\Gamma_-, |\nu \cdot \theta|)$ be given source pattern and boundary data, respectively. As already mentioned in the introduction, photoacoustic signal generation due to the absorption of light is described by the operator

$$\mathbf{H} \colon \mathcal{D}_p \to L^p(\Omega) \colon (\mu, \sigma) \mapsto \mu \int_{\mathbb{S}^{d-1}} \mathbf{T}(\mu, \sigma)(\cdot, \theta)\mathrm{d}\theta.$$

To shorten the notation, in the following we will make use of the averaging operator $\mathbf{A} \colon W^p(\Omega \times \mathbb{S}^{d-1}) \to L^p(\Omega)$ defined by $\mathbf{A}\Phi = \int_{\mathbb{S}^{d-1}} \Phi(\cdot, \theta)\mathrm{d}\theta$. By Hölders inequality the averaging operator is well defined, linear and bounded. Using the averaging operator we can write $\mathbf{H}(\mu, \sigma) = \mu\mathbf{A}\mathbf{T}(\mu, \sigma)$.

Theorem 2.5 (Lipschitz continuity and weak continuity of **H**).

(a) *The operator* **H** *is Lipschitz continuous.*

(b) *For* $1 < p < \infty$, **H** *is sequentially weakly continuous.*

Proof. (a) Suppose that $(\mu, \sigma), (\hat{\mu}, \hat{\sigma}) \in \mathcal{D}_p$ are two pairs of admissible absorption and scattering coefficients. The decomposition $\mathbf{H}(\mu, \sigma) = \mu\mathbf{A}\mathbf{T}(\mu, \sigma)$ and the triangle inequality imply

$$\|\mu\mathbf{A}\mathbf{T}(\mu, \sigma) - \hat{\mu}\mathbf{A}\mathbf{T}(\hat{\mu}, \hat{\sigma})\|_{L^p(\Omega)}$$
$$= \|\mu\mathbf{A}\mathbf{T}(\mu, \sigma) - \hat{\mu}\mathbf{A}\mathbf{T}(\mu, \sigma) + \hat{\mu}\mathbf{A}\mathbf{T}(\mu, \sigma) - \hat{\mu}\mathbf{A}\mathbf{T}(\hat{\mu}, \hat{\sigma})\|_{L^p(\Omega)}$$
$$\leq \|\mathbf{A}\mathbf{T}(\mu, \sigma)\|_{L^\infty(\Omega)}\|\mu - \hat{\mu}\|_{L^p(\Omega)} + \|\hat{\mu}\|_{L^\infty(\Omega)}\|\mathbf{A}\mathbf{T}(\mu, \sigma) - \mathbf{A}\mathbf{T}(\hat{\mu}, \hat{\sigma})\|_{L^p(\Omega)}.$$

According to Theorem 2.3, the operator **T** is Lipschitz continuous. Because **A** is linear and bounded, also the composition **AT** is Lipschitz. Noting that $\|\mathbf{A}\mathbf{T}(\mu, \sigma)\|_{L^\infty(\Omega)}$ and $\|\hat{\mu}\|_{L^\infty(\Omega)}$ are bounded by constants independent of μ, σ and $\hat{\mu}, \hat{\sigma}$, this implies the Lipschitz continuity of **H**.

Let $(\mu_n, \sigma_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{D}_p$ that converges weakly to $(\mu, \sigma) \in \mathcal{D}_p$. Since $\mathbf{T}$ is weakly continuous and $\mathbf{A}$ is linear and bounded, $(\mathbf{AT}(\mu_n, \sigma_n))_{n \in \mathbb{N}}$ converges weakly to $\mathbf{AT}(\mu, \sigma)$. Further, for any function $\varphi \in L^{p_\star}(\Omega)$, the dual space of $L^p(\Omega)$, we have

$$\left| \int_\Omega \left(\mu(x)(\mathbf{AT})(\mu, \sigma)(x) - \mu_n(x)(\mathbf{AT})(\mu_n, \sigma_n)(x) \right) \varphi(x) \, \mathrm{d}x \right|$$
$$\leq \|\mathbf{AT}(\mu, \sigma)\|_{L^\infty(\Omega)} \left| \int_\Omega (\mu(x) - \mu_n(x)) \varphi(x) \, \mathrm{d}x \right|$$
$$+ \|\mu_n\|_{L^\infty(\Omega)} \left| \int_\Omega ((\mathbf{AT})(\mu, \sigma)(x) - (\mathbf{AT})(\mu_n, \sigma_n)(x)) \varphi(x) \mathrm{d}x \right|.$$

The weak convergence of μ_n and $(\mathbf{AT}(\mu_n, \sigma_n))_{n \in \mathbb{N}}$ therefore implies the weak convergence of $\mu_n \mathbf{AT}(\mu_n, \sigma_n)$ to $\mu \mathbf{AT}(\mu, \sigma)$ and shows the weak continuity of $\mathbf{H}$. $\qquad \square$

Note that for the case $1 < p < \infty$, the averaging operator $\mathbf{A}$ is even compact (see [40]) which implies the compactness of the composition $\mathbf{AT}$. As a consequence, for any given μ, the partial mapping $\sigma \mapsto \mu(\mathbf{AT})(\mu, \sigma)$ is compact. It seems unlikely, however, that the full operator $\mathbf{H}$ is compact, too.

Theorem 2.6 (Differentiability of $\mathbf{H}$). *For any $(\mu, \sigma) \in \mathcal{D}_p$, the one-sided directional derivative of $\mathbf{H}$ at (μ, σ) in any feasible direction $(h_\mu, h_\sigma) \in \mathcal{D}_p(\mu, \sigma)$ exists. Further, we have*

$$\mathbf{H}'(\mu, \sigma)(h_\mu, h_\sigma) = h_\mu \int_{\mathbb{S}^{d-1}} \mathbf{T}(\mu, \sigma)(\,\cdot\,, \theta) \mathrm{d}\theta + \mu \int_{\mathbb{S}^{d-1}} \mathbf{T}'(\mu, \sigma)(h_\mu, h_\sigma)(\,\cdot\,, \theta) \mathrm{d}\theta, \qquad (2.8)$$

where $\mathbf{T}'(\mu, \sigma)(h_\mu, h_\sigma)$ denotes the one-sides directional derivative of $\mathbf{T}$ at (μ, σ) in direction (h_μ, h_σ) given by (2.6). Finally, if $0 < \mu < \overline{\mu}$ and $0 < \sigma < \overline{\sigma}$, then $\mathbf{H}$ is Gâteaux differentiable at (μ, σ).

Proof. Let $(\mu, \sigma) \in \mathcal{D}_p$ and let $(h_\mu, h_\sigma) \in \mathcal{D}_p(\mu, \sigma)$ be a feasible direction. For sufficiently small $s > 0$, we have

$$\frac{\mathbf{H}((\mu, \sigma) + s(h_\mu, h_\sigma)) - \mathbf{H}(\mu, \sigma)}{s}$$
$$= \frac{(\mu + sh_\mu)(\mathbf{AT})((\mu, \sigma) + s(h_\mu, h_\sigma)) - \mu(\mathbf{AT})(\mu, \sigma)}{s}$$
$$= h_\mu(\mathbf{AT})((\mu, \sigma) + s(h_\mu, h_\sigma)) + \mu \frac{(\mathbf{AT})((\mu, \sigma) + s(h_\mu, h_\sigma)) - (\mathbf{AT})(\mu, \sigma)}{s}.$$

According to Theorem 2.3, the operator $\mathbf{T}$ is continuous and therefore the first term converges to $h_\mu(\mathbf{AT})(\mu, \sigma)$ as $s \to 0$. Because $\mathbf{T}$ is one-sided differentiable, see Theorem 2.4, the second term converges to $\mu\mathbf{AT}'(\mu, \sigma)(h_\mu, h_\sigma)$. Finally, if $0 < \mu < \overline{\mu}$ and $0 < \sigma < \overline{\sigma}$, then $\mathbf{H}'(\mu, \sigma)(h)$ is linear and bounded in the argument h which implies the Gâteaux differentiability of $\mathbf{H}$ at (μ, σ). $\qquad \square$

Recall that for any $(\mu, \sigma) \in \mathcal{D}_p$, the derivative $\mathbf{T}'(\mu, \sigma)$ is bounded and linear. Therefore, the right hand side of (2.8) depends linearly and continuously on $(h_\mu, h_\sigma) \in L^p(\Omega) \times L^p(\Omega)$. As a consequence, the one-sided directional derivative of $\mathbf{H}$ at (μ, σ) can be extended to a bounded linear operator $\mathbf{H}'(\mu, \sigma) \colon L^p(\Omega) \times L^p(\Omega) \to L^p(\Omega)$. We will refer to this extension as the derivative of $\mathbf{H}$ at (μ, σ). The derivative of $\mathbf{H}$ can be written in the form $\mathbf{H}'(\mu, \sigma)(h) = h_\mu \mathbf{A}(\mathbf{T}(\mu, \sigma)) + \mu \mathbf{A}(\mathbf{T}'(\mu, \sigma)(h))$.

9

2.4 The wave operator $\mathbf{W}_{\Omega,\Lambda}$

Let $U \subset \mathbb{R}^d$ be a bounded and convex domain with smooth boundary. We assume that $\bar{\Omega} \subset U$ and write $L_\Omega^2(\mathbb{R}^d)$ for the space of all square integrable functions defined on $\mathbb{R}^d$ that are supported in $\bar{\Omega}$. Likewise we denote by $C_\Omega^\infty(\mathbb{R}^d)$ the space of all infinitely differentiable functions defined on $\mathbb{R}^d$ having support in $\bar{\Omega}$. Further, let $\Lambda \subset \partial U$ be a relatively open subset of ∂U, and denote by $\mathrm{diam}(U)$ the maximal diameter of U and by $\mathrm{dist}(\Omega, \Lambda)$ the minimal distance of Ω from the observation surface Λ.

Definition 2.7 (The wave operator $\mathbf{W}_{\Omega,\Lambda}$). Let $w_{\Omega,\Lambda}\colon (0,\infty) \to \mathbb{R}$ be a smooth, nonnegative, compactly supported function with $w_{\Omega,\Lambda}(t) = 1$ for all $\mathrm{dist}(\Omega, \Lambda) \leq t \leq \mathrm{diam}(U) - \mathrm{dist}(\Omega, \Lambda)$. We then define the wave operator by

$$\mathbf{W}_{\Omega,\Lambda}\colon C_\Omega^\infty(\mathbb{R}^d) \subset L_\Omega^2(\mathbb{R}^d) \to L^2(\Lambda \times (0,\infty)) : h \mapsto w_{\Omega,\Lambda}\, p|_{\Lambda \times (0,\infty)}\,, \qquad (2.9)$$

where p denotes the unique solution of (1.3).

The operator $\mathbf{W}_{\Omega,\Lambda}$ maps the initial data of the wave equation (1.3) to its solution restricted to $\Lambda \subset \partial U$ and models the acoustic part of the forward problem of PAT. The cutoff function $w_{\Omega,\Lambda}$ accounts for the fact, that in the two dimensional case the solution of the wave equation has unbounded support in time but measurements can only be made over a finite time interval.

In the following we use a result from [42] to show that $\mathbf{W}_{\Omega,\Lambda}$ is a bounded linear and densely defined operator, and therefore can be extended to a bounded linear operator on $L_\Omega^2(\mathbb{R}^d)$ in a unique manner.

Theorem 2.8 (Continuity of the wave operator $\mathbf{W}_{\Omega,\Lambda}$). *There exists some constant $c_{\Omega,\Lambda}$ such that $\|\mathbf{W}_{\Omega,\Lambda} h\|_{L^2(\Lambda \times (0,\infty))} \leq c_{\Omega,\Lambda}\|h\|_{L_\Omega^2(\mathbb{R}^d)}$ for all $h \in C_\Omega^\infty(\mathbb{R}^d)$. Consequently, there exists a unique bounded linear extension*

$$\overline{\mathbf{W}}_{\Omega,\Lambda}\colon L_\Omega^2(\mathbb{R}^d) \to L^2(\Lambda \times (0,\infty)) \quad \text{with} \;\; \overline{\mathbf{W}}_{\Omega,\Lambda}|_{C_\Omega^\infty(\mathbb{R}^d)} = \mathbf{W}_{\Omega,\Lambda}\,.$$

With some abuse of notation we again write $\mathbf{W}_{\Omega,\Lambda}$ for $\overline{\mathbf{W}}_{\Omega,\Lambda}$ in the sequel.

Proof. It is sufficient to consider the case when the data are measured on the whole boundary $\Lambda = \partial U$. The well known explicit formulas for the solution of (1.3) in two and three spatial dimensions (see, for example, [23, 32]) imply that for every $(y, t) \in \partial U \times (0,\infty)$ we have

$$(\mathbf{W}_{\Omega,\Lambda} h)(y,t) = \begin{cases} \frac{w_{\Omega,\Lambda}(t)}{2\pi}\, \partial_t \int_0^t \frac{r}{\sqrt{t^2-r^2}} \int_{\mathbb{S}^{d-1}} h(y + r\omega)\, \mathrm{d}\omega \mathrm{d}r & \text{for } d=2 \\ \frac{w_{\Omega,\Lambda}(t)}{4\pi}\, \partial_t \left(t \int_{\mathbb{S}^{d-1}} h(y+t\omega)\, \mathrm{d}\omega \right) & \text{for } d=3\,. \end{cases} \qquad (2.10)$$

We define the spherical mean Radon transform $\mathbf{M}\colon C_\Omega^\infty(\mathbb{R}^d) \to C^\infty(\partial U \times (0,\infty))$ by $\mathbf{M}h(y,t) := 1/\omega_{d-1} \int_{\mathbb{S}^{d-1}} h(y + t\omega)\mathrm{d}\omega$ for $(y,t) \in \partial U \times (0,\infty)$. With the spherical mean Radon transform, the wave operator can be written as $(\mathbf{W}_{\Omega,\Lambda} h)(y,t) = w_{\Omega,\Lambda}(t)\, \partial_t \int_0^t r\mathbf{M}h(y,r)/\sqrt{t^2 - r^2}\mathrm{d}r$ in the case of two spatial dimensions and $(\mathbf{W}_{\Omega,\Lambda} h)(y,t) = w_{\Omega,\Lambda}(t)\, \partial_t(t\mathbf{M}h)(y,t)$ in the three dimensional case.

Next we use a Sobolev estimate derived in [42], which states that for every $\lambda \in \mathbb{R}$ there exists a constant $c_{K,\lambda}$ such that $\|\mathbf{M}h\|_{H^{\lambda+(d-1)/2}(\partial U \times (0,\infty))} \leq c_{K,\lambda}\|h\|_{H^\lambda(U)}$ for any $h \in C_\Omega^\infty(\mathbb{R}^d)$. Application of this identity with $\lambda = 0$ and using the smoothing properties of the Abel transform by degree $1/2$ for the case of two spatial dimensions yields the continuity of $\mathbf{W}_{\Omega,\Lambda}$ with respect to the L^2 topologies. In particular, $\mathbf{W}_{\Omega,\Lambda}$ has a unique bounded linear extension to $L_\Omega^2(\mathbb{R}^d)$. $\qquad \square$

For solving the inverse problem of qPAT we will further require an explicit expression for the adjoint of $\mathbf{W}_{\Omega,\Lambda}$, that we compute next.

Proposition 2.9 (Adjoint of the wave operator). *For $v \in L^2(\Lambda \times (0,\infty)) \cap C^1(\Lambda \times (0,\infty))$, and every $x \in \bar{\Omega}$, we have*

$$\left(\mathbf{W}_{\Omega,\Lambda}^* v\right)(x) = \begin{cases} -\dfrac{1}{2\pi} \displaystyle\int_\Lambda \int_{|x-y|}^\infty \dfrac{\partial_t(w_{\Omega,\Lambda}v)(y,t)}{\sqrt{r^2 - |x-y|^2}} \, \mathrm{d}r \, \mathrm{d}S(y) & \text{if } d = 2 \\ -\dfrac{1}{4\pi} \displaystyle\int_\Lambda \dfrac{\partial_t(w_{\Omega,\Lambda}v)(y,|x-y|)}{|x-y|} \, \mathrm{d}S(y) & \text{if } d = 3 \,. \end{cases} \tag{2.11}$$

Proof. This is a simple application of Fubini's theorem and the explicit expression for $\mathbf{W}_{\Omega,\Lambda}h$ given in (2.10). $\qquad\square$

3 Single-stage approach to qPAT

In this section we solve the inverse problem of qPAT by a single-stage approach. Our setting allows acoustic measurement for multiple sources. Such a strategy has been called multi-source qPAT or multiple illumination qPAT (see [6, 17, 57]). For that purpose, throughout this section $q_i \in L^\infty(\Omega \times \mathbb{S}^{d-1})$ and $f_i \in L^\infty(\Gamma_-, |\nu \cdot \theta|)$ for $i = 1, \ldots N$ denote given source patterns and boundary data, respectively. Recall that $\Omega \subset \mathbb{R}^d$ denotes a bounded convex domain with Lipschitz boundary and Γ_- denotes the inflow boundary consisting of all pairs $(x,\theta) \in \partial\Omega \times \mathbb{S}^{d-1}$ with $\nu(x) \cdot \theta < 0$.
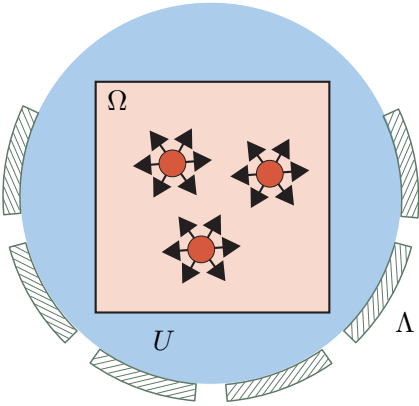

Figure 3.1: SETUP FOR SINGLE-STAGE qPAT. The stationary RTE governs the light propagation in the domain Ω. the absorption of photons induces an initial pressure wave proportional to the heating $\mathbf{H}_i(\mu,\sigma)$. Further, $\bar{\Omega}$ is supposed to be contained in another domain U, and the pressure waves are measured with acoustic detectors located on a open subset $\Lambda \subset \partial U$ of the boundary of U.

To indicate the dependence of the solution of the stationary RTE on the pair (q_i, f_i) we write $\mathbf{T}_i \colon \mathcal{D}_2 \to L^2(\Omega \times \mathbb{S}^{d-1})$ for the solution operator of the stationary RTE (1.1), (1.2) with sources (q_i, f_i). Here $\mathcal{D}_2$ is the set of all admissible pairs (μ, σ) defined in (2.3). Further, we use

$$\mathbf{H}_i \colon \mathcal{D}_2 \to L^2(\Omega) \colon (\mu, \sigma) \mapsto \mu \int_{\mathbb{S}^{d-1}} \mathbf{T}_i(\mu, \sigma)(\,\cdot\,, \theta) \mathrm{d}\theta$$

to denote the corresponding operator describing the heating. For the coupling with the acoustic problem, it will be convenient to consider $\mathbf{H}_i(\mu, \sigma) \in L^2(\Omega)$ as an element of $L^2_\Omega(\mathbb{R}^d)$, by extending it to a function defined on $\mathbb{R}^d$ that is equal to zero on $\mathbb{R}^d \setminus \Omega$.

Further, recall the Definition 2.7 of the wave operator $\mathbf{W}_{\Omega,\Lambda}$ modeling the acoustic problem, that maps the initial pressure h in the wave equation (1.3) to its solution restricted to $\Lambda \subset \partial U$.

11

Here U is a convex domain with smooth boundary that contains the support of f. To apply the results of Section 2, in the following we assume that $\bar{\Omega} \subset U$. Then, according to the Section 2, the operators $\mathbf{H}_i$ are Lipschitz continuous, weakly continuous and one-sided directional differentiable, and $\mathbf{W}_{\Omega,\Lambda}$ is linear and bounded. A practical geometry of these domains is illustrated in Figure 3.1.

3.1 Formulation as operator equation

In order to apply standard techniques for the solution of inverse problems we write the reconstruction problem of (multiple-source) qPAT as a single operator equation. For that propose we denote by

$$\mathbf{F}\colon \mathcal{D}_2 \to \left(L^2(\Lambda \times (0,\infty))\right)^N$$
$$(\mu,\sigma) \mapsto (\mathbf{W}_{\Omega,\Lambda} \circ \mathbf{H}_1(\mu,\sigma), \ldots, \mathbf{W}_{\Omega,\Lambda} \circ \mathbf{H}_N(\mu,\sigma))$$

the operator describing the entire forward problem of qPAT. Further we denote by $\|v\|_N^2 := \sum_{i=1}^N \|v_i\|_{L^2(\Lambda \times (0,\infty))}^2$ the squared standard norm on $L^2(\Lambda \times (0,\infty))^N$.

Theorem 3.1 (Properties of the forward operator of qPAT)**.**

(a) *The operator* $\mathbf{F}$ *sequentially weakly continuous.*

(b) *The operator* $\mathbf{F}$ *is Lipschitz continuous.*

(c) *For any* $(\mu,\sigma) \in \mathcal{D}_2$, *the one-sided directional derivative in any feasible direction* $h \in \mathcal{D}_2(\mu,\sigma)$ *exists. Further,*

$$\mathbf{F}'(\mu,\sigma)(h) = \left(\mathbf{W}_{\Omega,\Lambda} \circ \mathbf{H}_1'(\mu,\sigma)(h), \ldots, \mathbf{W}_{\Omega,\Lambda} \circ \mathbf{H}_N'(\mu,\sigma)(h)\right) \tag{3.1}$$

where $\mathbf{H}_i'(\mu,\sigma)(h)$ *is given by* (2.8) *with* $\mathbf{T}$ *replaced by* $\mathbf{T}_i$.

(d) *If* $0 < \mu \le \overline{\mu}$ *and* $0 < \sigma < \overline{\sigma}$, *then* $\mathbf{F}$ *is Gâteaux differentiable at* (μ,σ).

Proof. All claims follow from the corresponding properties of the operators $\mathbf{H}_i$ (see Theorems 2.5 and 2.6) and the boundedness of $\mathbf{W}_{\Omega,\Lambda}$ discussed in Theorem 2.8. $\square$

The inverse problem of qPAT with multiple illuminations consists in solving the nonlinear equation

$$v = \mathbf{F}(\mu^\star, \sigma^\star) + z, \tag{3.2}$$

where $(\mu^\star, \sigma^\star)$ is the unknown, $v = (v_1, \ldots, v_N)$ are the given noisy data, $\mathbf{F}\colon \mathcal{D}_2 \to L^2(\Lambda \times (0,\infty))^N$ is the forward operator, and z is the noise in the data. Our single-stage approach for qPAT consists in estimating the parameter pair $(\mu^\star, \sigma^\star)$ directly from (3.2). In contrast, existing two-stage approaches for qPAT first construct estimates h_i for the heating functions $\mathbf{H}_i(\mu^\star, \sigma^\star)$ from data v_i by numerically inverting $\mathbf{W}_{\Omega,\Lambda}$, and subsequently solve $(h_1, \ldots h_N) = (\mathbf{H}_1(\mu,\sigma), \ldots, \mathbf{H}_N(\mu,\sigma))$ for (μ,σ).

There are at least two common methods for tackling an inverse problem of the form (3.2): Tikhonov type regularization methods on the one and iterative regularization methods on the other hand. In the following we apply Tikhonov regularization to the inverse problem of qPAT.

12

3.2 Tikhonov regularization for single-stage qPAT

We address the inverse problem (3.2) by Tikhonov regularization with general convex penalty. For that purpose, let $\mathcal{R}\colon L^2(\Omega) \times L^2(\Omega) \to \mathbb{R} \cup \{\infty\}$ be a convex, and lower semicontinuous functional with domain $\mathcal{D}(\mathcal{R}) := \{(\mu, \sigma) \in L^2(\Omega) \times L^2(\Omega) : \mathcal{R}(\mu, \sigma) < \infty\}$. We assume that $\mathcal{R}$ is chosen such that $\mathcal{D}_2 \cap \mathcal{D}(\mathcal{R})$ is non-empty.

Tikhonov regularization with penalty $\mathcal{R}$ consists in computing a minimizer of the generalized Tikhonov functional

$$
\begin{aligned}
\mathcal{T}_{v,\lambda}\colon & L^2(\Omega) \times L^2(\Omega) \to \mathbb{R} \cup \{\infty\} \\
& (\mu, \sigma) \mapsto \begin{cases} \frac{1}{2} \|\mathbf{F}(\mu, \sigma) - v\|_N^2 + \lambda \mathcal{R}(\mu, \sigma) & \text{if } (\mu, \sigma) \in \mathcal{D}_2 \cap \mathcal{D}(\mathcal{R}) \\ \infty & \text{otherwise} . \end{cases}
\end{aligned}
\tag{3.3}
$$

Here $\lambda > 0$ is the so called regularization parameter which has to be chosen accordingly, to balance between stability with respect to noise and accuracy in the case of exact data. The data-fidelity term $\frac{1}{2}\|\mathbf{F}(\mu, \sigma) - v\|_N^2$ guarantees that any minimizer of (3.3) predicts the given data sufficiently well. The regularization term $\lambda \mathcal{R}(\mu, \sigma)$ on the other hand avoids over-fitting of the data and makes the reconstruction process well-defined and stable.

Note that Tikhonov regularization with penalty $\mathcal{R}$ is designed to stably approximate a solution of the constrained optimization problem

$$
\mathcal{R}(\mu, \sigma) \to \min_{(\mu,\sigma) \in \mathcal{D}_2 \cap \mathcal{D}(\mathcal{R})} \qquad \text{such that } \mathbf{F}(\mu, \sigma) = v^\star .
\tag{3.4}
$$

Here $v^\star \in \operatorname{ran}(\mathbf{F})$ is an element in the range of $\mathbf{F}$ and is referred to as exact data. Any solution of (3.4) is called $\mathcal{R}$-minimizing solution of the equation $\mathbf{F}(\mu, \sigma) = v^\star$. Under the given assumptions there exists at least one $\mathcal{R}$-minimizing solution, which however is not necessarily unique, see [48, 49].

The properties of the operator $\mathbf{F}$ derived above and the use of general results from regularization theory yield the following result.

Theorem 3.2 (Well-posedness and convergence of Tikhonov regularization)**.**

(a) *For data $v \in L^2(\Lambda \times (0, \infty))^N$ and every $\lambda > 0$, the Tikhonov functional $\mathcal{T}_{v,\lambda}$ has at least one minimizer.*

(b) *Let $\lambda > 0$, $v \in L^2(\Lambda \times (0, \infty))^N$, and let $(v_n)_{n \in \mathbb{N}}$ be a sequence in $L^2(\Lambda \times (0, \infty))^N$ with $\|v - v_n\|_N \to 0$. Then every sequence of minimizers $(\mu_n, \sigma_n) \in \arg\min \mathcal{T}_{v_n, \lambda}$ has a weakly convergent subsequence. Further, the limit u of every weakly convergent subsequence $(\mu_{\tau(n)}, \sigma_{\tau(n)})_{n \in \mathbb{N}}$ is a minimizer (μ, σ) of $\mathcal{T}_{\lambda,v}$ and satisfies $(\mu_{\tau(n)}, \sigma_{\tau(n)}) \to \mathcal{R}(\mu, \sigma)$ for $n \to \infty$.*

(c) *Let $v^\star \in \operatorname{ran}(\mathbf{F})$, let $(\delta_n)_{n \in \mathbb{N}} \subset (0, \infty)$ be a sequence converging to zero, and let $(v_n)_{n \in \mathbb{N}} \subset V$ be a sequence of data with $\|v^\star - v_n\|_N \leq \delta_n$. Suppose further that $(\lambda_n)_{n \in \mathbb{N}} \subset (0, \infty)$ satisfies $\lambda_n \to 0$ and $\delta_n^2 / \lambda_n \to 0$ as $n \to \infty$. Then the following hold:*

- *Every sequence $(\mu_n, \sigma_n) \in \arg\min \mathcal{T}_{v_k, \lambda_k}$ has a weakly converging subsequence.*
- *The limit of every weakly convergent subsequence $(\mu_{\tau(n)}, \sigma_{\tau(n)})_{n \in \mathbb{N}}$ of $(\mu_n, \sigma_n)_{n \in \mathbb{N}}$ is an $\mathcal{R}$-minimizing solution $(\mu^\star, \sigma^\star)$ of $\mathbf{F}(\mu, \sigma) = v^\star$ and satisfies $\mathcal{R}(\mu_{\tau(n)}, \sigma_{\tau(n)}) \to \mathcal{R}(\mu^\star, \sigma^\star)$.*
- *If the $\mathcal{R}$-minimizing solution of $\mathbf{F}(\mu, \sigma) = v^\star$ is unique, then $(\mu_n, \sigma_n) \rightharpoonup (\mu^\star, \sigma^\star)$.*

Proof. Since **F** is sequentially weakly continuous (see Theorem 3.1) and $\mathcal{D}_2$ is closed and convex, this follows from general results of Tikhonov regularization with convex penalties, see for example, [48, Thm. 3.3, Thm. 3.4, Thm. 3.5]. $\qquad\square$

3.3 Gradient of the data-fidelity term

For numerically minimizing the Tikhonov functional we require the gradient of the data-fidelity term

$$\mathcal{F}(\mu,\sigma) := \frac{1}{2}\left\|\mathbf{F}(\mu,\sigma) - v\right\|_N^2 = \frac{1}{2}\sum_{i=1}^N \left\|\mathbf{W}_{\Omega,\Lambda}\mu\mathbf{A}\mathbf{T}_i(\mu,\sigma) - v_i\right\|_{L^2(\Lambda\times(0,\infty))}^2. \tag{3.5}$$

Recall that $v_i \in L^2(\Omega)$ are the given data, $\mathbf{T}_i$ is the solution operator for the stationary RTE with source patterns q_i and boundary data f_i, $\mathbf{A}$ is the averaging operator, and $\mathbf{W}_{\Omega,\Lambda}$ is the solution operator for the wave equation.

Let $(\mu,\sigma) \in \mathcal{D}_2$ be some admissible pair of parameters and let $(h_\mu, h_\sigma) \mapsto \mathcal{F}'(\mu,\sigma)(h_\mu, h_\sigma)$ denote the one-sided directional derivative of $\mathcal{F}$ at (μ,σ). We define the gradient $\nabla\mathcal{F}(\mu,\sigma)$ of $\mathcal{F}$ at (μ,σ) to be any element in $L^2(\Omega\times\mathbb{S}^{d-1})\times L^2(\Omega\times\mathbb{S}^{d-1})$ satisfying

$$\langle\nabla\mathcal{F}(\mu,\sigma),(h_\mu,h_\sigma)\rangle_{L^2(\Omega\times\mathbb{S}^{d-1})^2} = \mathcal{F}'(\mu,\sigma)(h_\mu,h_\sigma) \quad \text{for } (h_\mu,h_\sigma) \in \mathcal{D}_2(\mu,\sigma). \tag{3.6}$$

From Theorem 3.1 and the chain rule, it follows that $\mathcal{F}'(\mu,\sigma)(h_\mu,h_\sigma)$ exists for any feasible direction $(h_\mu,h_\sigma) \in \mathcal{D}_2(\mu,\sigma)$. Further, in the case that μ and σ are strictly positive, we have $\mathcal{D}_2(\mu,\sigma) = L^2(\Omega\times\mathbb{S}^{d-1})\times L^2(\Omega\times\mathbb{S}^{d-1})$, which implies that $\nabla\mathcal{F}(\mu,\sigma)$ is uniquely defined by (3.6).

In order to compute the gradient we require a more explicit expression for the one-sided directional derivative, that we shall derive next.

Proposition 3.3 (One-sided directional derivative of the data-fidelity term). *Let $(\mu,\sigma) \in \mathcal{D}_2$ be an admissible pair of parameters and let $(h_\mu, h_\sigma) \in \mathcal{D}_2(\mu,\sigma)$ be a feasible direction. Then we have*

$$\mathcal{F}'(\mu,\sigma)(h_\mu,h_\sigma) = \sum_{i=1}^N \left\langle\mathbf{A}\Phi_i\mathbf{W}_{\Omega,\Lambda}^*\left[\mathbf{W}_{\Omega,\Lambda}\mu\mathbf{A}\mathbf{T}_i(\mu,\sigma) - v_i\right] - \mathbf{A}(\Phi_i\Phi_i^*), h_\mu\right\rangle_{L^2(\Omega)}$$

$$+ \sum_{i=1}^N \left\langle\mathbf{A}\left(-\Phi_i\Phi_i^* + (\mathbf{K}\Phi_i)\Phi_i^*\right), h_\sigma\right\rangle_{L^2(\Omega)}, \tag{3.7}$$

where $\Phi_i := \mathbf{T}_i(\mu,\sigma)$, and Φ_i^* is the unique solution of the adjoint equation

$$(-\theta\cdot\nabla_x + (\mu+\sigma-\sigma\mathbf{K}))\,\Phi_i^* = \mathbf{A}^*\mu\mathbf{W}_{\Omega,\Lambda}^*\left[\mathbf{W}_{\Omega,\Lambda}\mu\mathbf{A}\Phi_i - v_i\right] \quad \text{in } \Omega\times\mathbb{S}^{d-1} \tag{3.8}$$

satisfying the zero outflow boundary condition $\Phi^*|_{\Gamma_+} = 0$.

Proof. Obviously it is sufficient to consider the case $N = 1$, where we write v, $\mathbf{T}$, Φ and Φ^* in place of v_i, $\mathbf{T}_i$, Φ_i and Φ_i^*. By (3.5) we have

$$\mathcal{F}'(\mu,\sigma)(h_\mu,h_\sigma)$$
$$= \left\langle\mathbf{W}_{\Omega,\Lambda}\mu\mathbf{A}\mathbf{T}(\mu,\sigma) - v, \mathbf{W}_{\Omega,\Lambda}h_\mu\mathbf{A}\mathbf{T}(\mu,\sigma) + \mathbf{W}_{\Omega,\Lambda}\mu\mathbf{A}\mathbf{T}'(\mu,\sigma)(h_\mu,h_\sigma)\right\rangle_{L^2(\Lambda\times(0,\infty))}$$
$$= \left\langle\mathbf{W}_{\Omega,\Lambda}\mu\mathbf{A}\Phi - v, \mathbf{W}_{\Omega,\Lambda}h_\mu\mathbf{A}\Phi\right\rangle_{L^2(\Lambda\times(0,\infty))}$$

$$+ \left\langle \mathbf{W}_{\Omega,\Lambda} \mu \mathbf{A} \Phi - v, \mathbf{W}_{\Omega,\Lambda} \mu \mathbf{A} \mathbf{T}'(\mu,\sigma)(h_\mu, h_\sigma) \right\rangle_{L^2(\Lambda \times (0,\infty))}$$

$$= \left\langle \mathbf{A} \Phi \mathbf{W}_{\Omega,\Lambda}^* \left[\mathbf{W}_{\Omega,\Lambda} \mu \mathbf{A} \Phi - v \right], h_\mu \right\rangle_{L^2(\Omega)}$$

$$+ \left\langle \mathbf{A}^* \mu \mathbf{W}_{\Omega,\Lambda}^* \left[\mathbf{W}_{\Omega,\Lambda} \mu \mathbf{A} \Phi - v \right], \mathbf{T}'(\mu,\sigma)(h_\mu, h_\sigma) \right\rangle_{L^2(\Omega \times \mathbb{S}^{d-1})}. \tag{3.9}$$

Recall that Φ^* is the solution of (3.8) with source term $q = \mathbf{A}^* \mu \mathbf{W}_{\Omega,\Lambda}^* \left[\mathbf{W}_{\Omega,\Lambda} \mu \mathbf{A} \Phi - v \right]$ and zero outflow boundary conditions $\Phi^*|_{\Gamma_+} = 0$. Further, according to Theorem 2.4, the one-sided directional derivative of $\mathbf{T}$ is given by $\mathbf{T}'(\mu,\sigma)(h_\mu, h_\sigma) = \Psi$, where Ψ is the unique solution of $(\theta \cdot \nabla_x + (\mu + \sigma - \sigma \mathbf{K})) \Psi = -(h_\mu + h_\sigma - h_\sigma \mathbf{K}) \Phi$ with inflow boundary conditions $\Psi|_{\Gamma_-} = 0$. The zero outflow and zero inflow boundary conditions of Φ^* and Ψ, respectively, and one integration by parts, show $\langle -\theta \cdot \nabla_x \Phi^*, \Psi \rangle_{L^2(\Omega \times \mathbb{S}^{d-1})} = \langle \Phi^*, \theta \cdot \nabla_x \Psi \rangle_{L^2(\Omega \times \mathbb{S}^{d-1})}$. Further, notice that the averaging operator $\mathbf{A}$ has adjoint $\mathbf{A}^*: L^2(\Omega) \to L^2(\Omega \times \mathbb{S}^{d-1})$, $(\mathbf{A}^* g)(x,v) = g(x)$, and that the scattering operator $\mathbf{K}$ is self-adjoint.

Using these considerations, the second term in (3.9) can be written as

$$\left\langle \mathbf{A}^* \mu \mathbf{W}_{\Omega,\Lambda}^* \left[\mathbf{W}_{\Omega,\Lambda} \mu \mathbf{A} \Phi - v \right], \mathbf{T}'(\mu,\sigma)(h_\mu, h_\sigma) \right\rangle_{L^2(\Omega \times \mathbb{S}^{d-1})}$$
$$= \left\langle \left(-\theta \cdot \nabla_x + (\mu + \sigma - \sigma \mathbf{K}) \right) \Phi^*, \Psi \right\rangle_{L^2(\Omega \times \mathbb{S}^{d-1})}$$
$$= \left\langle \Phi^*, \left(\theta \cdot \nabla_x + (\mu + \sigma - \sigma \mathbf{K}) \right) \Psi \right\rangle_{L^2(\Omega \times \mathbb{S}^{d-1})}$$
$$= \left\langle \Phi^*, -\left(h_\mu + h_\sigma - h_\sigma \mathbf{K} \right) \Phi \right\rangle_{L^2(\Omega \times \mathbb{S}^{d-1})}$$
$$= -\left\langle \Phi \Phi^*, h_\mu + h_\sigma \right\rangle_{L^2(\Omega \times \mathbb{S}^{d-1})} + \left\langle (\mathbf{K}\Phi)\Phi^*, h_\sigma \right\rangle_{L^2(\Omega \times \mathbb{S}^{d-1})}$$
$$= -\left\langle \Phi \Phi^*, \mathbf{A}^* (h_\mu + h_\sigma) \right\rangle_{L^2(\Omega \times \mathbb{S}^{d-1})} + \left\langle (\mathbf{K}\Phi)\Phi^*, \mathbf{A}^* h_\sigma \right\rangle_{L^2(\Omega \times \mathbb{S}^{d-1})}$$
$$= \left\langle -\mathbf{A}(\Phi \Phi^*), h_\mu \right\rangle_{L^2(\Omega)} + \left\langle -\mathbf{A}(\Phi \Phi^*) + \mathbf{A}((\mathbf{K}\Phi)\Phi^*), h_\sigma \right\rangle_{L^2(\Omega)}.$$

Together with (3.9) this yields the desired identity (3.7). $\qquad \square$

Let $(\mu,\sigma) \in \mathcal{D}_2$ be an admissible pair of absorption and scattering coefficient. If μ, σ are both strictly positive, then one concludes from Proposition 3.3 that the gradient of $\mathcal{F}$ at (μ,σ) is uniquely defined and given by $\nabla \mathcal{F}(\mu,\sigma) = (\nabla_\mu \mathcal{F}(\mu,\sigma), \nabla_\sigma \mathcal{F}(\mu,\sigma))$ with

$$\nabla_\mu \mathcal{F}(\mu,\sigma) = \sum_{i=1}^N \left(\mathbf{A} \Phi_i \mathbf{W}_{\Omega,\Lambda}^* \left[\mathbf{W}_{\Omega,\Lambda} \mu \mathbf{A} \Phi_i - v_i \right] - \mathbf{A}(\Phi_i \Phi_i^*) \right) \tag{3.10}$$

$$\nabla_\sigma \mathcal{F}(\mu,\sigma) = \sum_{i=1}^N \mathbf{A} \left(-\Phi_i \Phi_i^* + (\mathbf{K}\Phi_i)\Phi_i^* \right). \tag{3.11}$$

Here $\Phi_i := \mathbf{T}_i(\mu,\sigma)$, and Φ_i^* is the solution of the adjoint equation (3.8) with zero outflow boundary condition. In the case that μ, σ are not both strictly positive, the gradient is not uniquely defined by (3.6). However, Proposition 3.3 implies that the vector $\nabla \mathcal{F}(\mu,\sigma)$ defined by (3.10), (3.11) still satisfies (3.3). We therefore take (3.10), (3.11) as gradient of $\mathcal{F}$ at any $(\mu,\sigma) \in \mathcal{D}_2$.

3.4 Proximal gradient algorithm for single-stage qPAT

In order to minimize the Tikhonov functional we apply the proximal gradient (or forward backward splitting) algorithm, which is an iterative algorithm for minimizing functionals that can be written as a sum $\mathcal{F} + \mathcal{G}$, where $\mathcal{F}$ is smooth and $\mathcal{G}$ is convex [13, 12]. The proximal gradient algorithm

computes a sequence of iterates by alternating application of explicit gradient steps for the first functional $\mathcal{F}$ and implicit proximal steps for the second functional $\mathcal{G}$.

To apply the proximal gradient algorithm for minimizing the Tikhonov functional (3.3) we take $\mathcal{F}(\mu, \sigma) = \frac{1}{2}\|\mathbf{F}(\mu, \sigma) - v\|_N^2$ for the first and $\mathcal{G}(\mu, \sigma) = \lambda\mathcal{R}(\mu, \sigma)$ for the second functional. The proximal gradient algorithm then generates a sequence (μ_n, σ_n) of iterates defined by

$$(\mu_{n+1}, \sigma_{n+1}) := \text{prox}_{s_n\lambda\mathcal{R}}\left((\mu_n, \sigma_n) - s_n\nabla\mathcal{F}(\mu_n, \sigma_n)\right) \quad \text{for } n \in \mathbb{N}. \tag{3.12}$$

Here $(\mu_0, \sigma_0) \in \mathcal{D}_2 \cap \mathcal{D}(\mathcal{R})$ is an initial guess, $s_n > 0$ is the step size in the n-th iteration, $\nabla\mathcal{F}(\mu, \sigma) = (\nabla_\mu\mathcal{F}(\mu, \sigma), \nabla_\sigma\mathcal{F}(\mu, \sigma))$ is the gradient of $\mathcal{F}$ given by (3.10), (3.11), and

$$\text{prox}_{s_n\lambda\mathcal{R}}(\hat{\sigma}, \hat{\mu}) := \underset{(\sigma, \mu)\in\mathcal{D}_2\cap\mathcal{D}(\mathcal{R})}{\arg\min} \frac{1}{2}\|(\mu, \sigma) - (\hat{\sigma}, \hat{\mu})\|^2 + s_n\lambda\mathcal{R}(\mu, \sigma) \tag{3.13}$$

is the proximity operator corresponding to the functional $s_n\lambda\mathcal{R}(\mu, \sigma)$.

Remark 3.4 (Lipschitz continuity of $\nabla\mathcal{F}$). *Note that the gradient $\nabla\mathcal{F}$ of the data-fidelity term is easily be shown to be Lipschitz continuous. This either can be deduced from the explicit expressions (3.10), (3.11) or by using the expression $\nabla\mathcal{F}(\mu, \sigma) = \mathbf{F}'(\mu, \sigma)^*(\mathbf{F}(\mu, \sigma) - v)$. In any case, the Lipschitz continuity of $(\mu, \sigma) \mapsto \nabla\mathcal{F}(\mu, \sigma)$ is shown similarly to proofs of the Lipschitz-continuity of $\mathbf{T}$ and $\mathbf{H}$, given in Section 2.2 and Section 2.3, respectively.*

Convergence of the proximal gradient algorithm (3.12) is well known for the case that $\mathcal{F}$ is convex with β-Lipschitz continuous gradient and that the chosen step sizes satisfy $s_n \in [\epsilon, 2/\beta - \epsilon]$ for some constant $\epsilon > 0$, see [13, 12]. These results are also valid for infinite dimensional Hilbert spaces. Because our forward operator $\mathbf{F}$ is nonlinear, the data-fidelity term $\mathcal{F}$ is non-convex and these results are not directly applicable to qPAT. Recently, the converge analysis of the proximal gradient analysis has been extended to the case of non-convex functionals; see [4, 8, 11].

4 Numerical implementation

Our numerical simulations are carried out in $d = 2$ spatial dimensions. The stationary RTE is solved on a square domain $\Omega = [0, 1]^2$. For the scattering kernel we choose the two dimensional version of the Henyey-Greenstein kernel,

$$k(\theta, \theta') := \frac{1}{2\pi}\frac{1 - g^2}{1 + g^2 - 2g\cos(\theta \cdot \theta')} \quad \text{for } \theta, \theta' \in \mathbb{S}^1,$$

where $g \in (0, 1)$ is the anisotropy factor. Before we present results of our numerical simulations we first outline how we numerically solve the stationary RTE in two spatial dimensions that is required for evaluating both, the forward operator $\mathbf{F}$ and the gradient $\nabla\mathcal{F}$ of the data-fidelity term.

4.1 Numerical solution of the RTE

For the numerical solution of the stationary RTE (1.1), (1.2) we employ a finite element method. For that purpose one calculates the weak form of equation (1.1), (1.2) by integrating the equation against a test function $w: \Omega \times \mathbb{S}^1 \to \mathbb{R}$. Integrating by parts in the transport term yields

$$\int_\Omega \int_{\mathbb{S}^1} (-\theta \cdot \nabla_x w + \mu w + \sigma w - \sigma\mathbf{K}w)\Phi\,\mathrm{d}\theta\,\mathrm{d}x + \int_{\partial\Omega\times\mathbb{S}^1}\Phi w(\theta \cdot \nu)\,\mathrm{d}\sigma = \int_\Omega\int_{\mathbb{S}^1} qw\,\mathrm{d}\theta\,\mathrm{d}x. \tag{4.1}$$

16

Here we dropped all dependencies on the variables to shorten notation and $\mathrm{d}\sigma$ denotes the usual surface measure on $\partial\Omega \times \mathbb{S}^1$.

The numerical scheme replaces the exact solution by a linear combination in the finite element space

$$\Phi^{(h)}(x,\theta) = \sum_{i=1}^{N_h} c_i^{(h)} \psi_i^{(h)}(x,\theta)\,, \tag{4.2}$$

where any $\psi_i^{(h)}(x,\theta)$ is the product of a basis function in space and a basis function in velocity and the sum ranges over all possible combinations. The spatial domain is triangulated uniformly with mesh size h as illustrated in Figure 4.1. The velocity direction on the circle is divided into 16 equal subintervals. We use P_1 Lagrangian elements, i.e. piecewise affine functions, in the two dimensional spatial domain as well as for the angle.
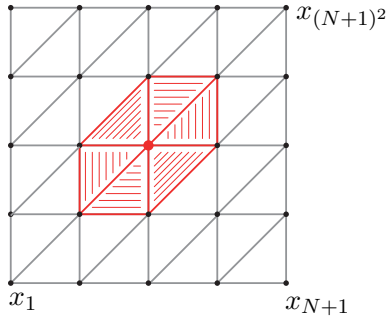

Figure 4.1: SPATIAL FINITE ELEMENT DISCRETIZATION. The square domain $\Omega = [0,1]^2$ is divided into $2N^2$ triangles. To any of the $(N+1)^2$ grid points $x_1,\ldots,x_{(N+1)^2}$ a piecewise affine basis function is associated, that takes the value one at one grid point and the value zero on all other grid points, and is affine on every triangle.

To increase stability in low scattering areas we add some artificial diffusion in the transport direction. This is called the streamline diffusion method, see for example [33] and the references therein. In the streamline diffusion method the solution Φ is approximated in the usual way by (4.2). However, the test functions take the form

$$w = \sum_{j=1}^{N_h} w_j (\psi_j(x,\theta) + \delta(x)\,\theta \cdot \nabla_x \psi_j(x,\theta))\,, \tag{4.3}$$

where the additional term introduces some artificial diffusion. In our experiments, the stabilization parameter is taken as $\delta(x) = 3h/100$ for $\sigma(x) + \mu(x) < 1$ and zero otherwise. Note that the streamline diffusion method provides a fully consistent stabilization of the original problem.

Making the ansatz (4.2) for the numerical solution and using test functions of the form (4.3), equation (4.1) yields a system of linear equations $M^{(h)} c^{(h)} = b^{(h)}$ for the coefficient vector of the numerical solution. The entries of $M^{(h)}$ and $b^{(h)}$ can be calculated by setting $\Phi = \psi_i$ and $w = \psi_j + \delta\theta \cdot \nabla_x \psi_j$. For simplicity we only consider the case $q = 0$. Then, similar to (4.1) we obtain

$$\int_\Omega \int_{\mathbb{S}^1} (\delta\theta \cdot \nabla_x \psi_i - \psi_i)\,\theta \cdot \nabla_x \psi_j \mathrm{d}\theta \mathrm{d}x + \int_{\Gamma_+} |\theta \cdot \nu|\,\psi_i \psi_j \mathrm{d}\sigma$$

$$\int_\Omega \int_{\mathbb{S}^1} (\mu + \sigma - \sigma\mathbf{K})\,(\psi_j + \delta\theta \cdot \nabla_x \psi_j)\,\psi_i \mathrm{d}\theta \mathrm{d}x = \int_{\Gamma_-} |\theta \cdot \nu|\,\psi_i \psi_j \mathrm{d}\sigma\,. \tag{4.4}$$

The entries of the matrix $M^{(h)}$ now can be calculated by evaluating the integrals on the left hand side of (4.4). The right hand side of (4.4) together with the prescribed boundary values on Γ_- yields the entries of $b^{(h)}$.

17

4.2 Numerical results

For the following numerical results, the stationary RTE is solved by the finite element method outlined in Subsection 4.1. For that purpose the domain $\Omega = [0,1]^2$ is discretized by a mesh containing 7442 triangular elements (compare Figure 4.1). The angular domain is divided into 16 subintervals of equal length. The anisotropy factor is taken as $g = 0.6$ and the scattering coefficient is taken as $\sigma = 3$. We use a single source distribution representing a planar illumination along the lower edge $[0,1] \times \{0\}$.

The solution of the two-dimensional wave equation (1.3) is computed by numerically evaluating the solution formula (2.10), where the detection curve $\Lambda = \{3/2(\cos\varphi, \sin\varphi) : \varphi \in (-\pi, 0)\}$ is a half-circle on the boundary of $B_{3/2}(0)$. The adjoint $\mathbf{W}^*_{\Omega,\Lambda} h$ is evaluated by numerically implementing (2.11). This can be done efficiently by a filtered backprojection algorithm as described in [9, 25]. The geometry of Ω and Λ is similar to the one illustrated in Figure 3.1.
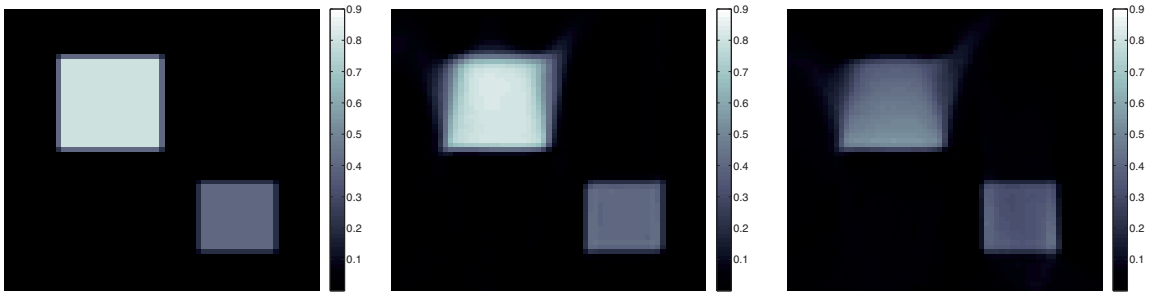

Figure 4.2: RECONSTRUCTION RESULTS FOR SIMULATED DATA. True absorption coefficient (left), reconstructed absorption coefficient using our single-stage approach (middle), and reconstructed absorption coefficient using the usual two-stage approach (right).

For our initial experiments we assume the scattering coefficient σ to be known. In such a situation, the proximal gradient algorithm outlined in the Subsection 3.4 reads

$$\mu_{n+1} := \mathrm{prox}_{s_n \lambda \mathcal{R}} \left(\mu_n - s_n \nabla_\mu \mathcal{F}(\mu_n, \sigma)\right) \quad \text{for } n \in \mathbb{N}$$

where $s_n > 0$ is the step size, $\nabla_\mu \mathcal{F}(\mu, \sigma)$ is the gradient of $\mathcal{F}$ in the first component, given by (3.10), and $\mathrm{prox}_{s_n \lambda \mathcal{R}}(\cdot)$ is the proximity operator similar as in (3.13). In the presented numerical examples the regularization term is taken as a quadratic functional $\mathcal{R}(\mu) = \frac{1}{2}\|\partial_x \mu\|^2_{L^2(\Omega)} + \frac{1}{2}\|\partial_y \mu\|^2_{L^2(\Omega)}$. In order to speed up the iterative scheme we compute the proximity operator only approximately by projecting the unconstrained minimizer $\arg\min_\mu \frac{1}{2}\|\mu - \hat{\mu}\|^2 + s_n \lambda \mathcal{R}(\mu)$ on $\mathcal{D}_2$. Therefore the main numerical cost in the proximal step the solution of a linear equation, which is relatively cheap compared to the evaluation of the gradient $\nabla_\mu \mathcal{F}$.

In Figure 4.2 we present results of our numerical simulations. The left image shows the true absorption coefficient and the central image shows the numerical reconstruction with the proposed single-stage approach (using 40 iterations of the proximal gradient algorithm). For comparison purpose, the right image in Figure 4.2 shows reconstruction results using the classical two-stage approach. For that purpose we apply Tikhonov regularization and the proximal gradient algorithm (again using 40 iterations) to the inverse problem $h = \mathbf{H}_i(\mu) + z_h$. The approximate heating h is computed numerically by applying the two dimensional universal backprojection formula [9, 38, 29] the wave data $v = \mathbf{W}_{\Omega,\Lambda} \circ \mathbf{H}_i(\mu) + z$. All computations have been performed in MATLAB on a

MacBook Pro with 2.3 GHz Intel Core i7 processor. The total computation times have been 26 minutes for the two-stage approach and 38 minutes for the single-stage approach.

One notices that in both reconstructions some boundaries in the upper half are blurred. Such artifacts are expected and arise from the ill-posedness of the acoustical problem when using limited-angle data; see [28, 43, 55]. However these artifacts are less severe for the single-stage algorithm than for the classical two-stage algorithm. Further, in this example, the single-stage algorithm also yields a better quantitative estimation of the values of the absorption coefficient.
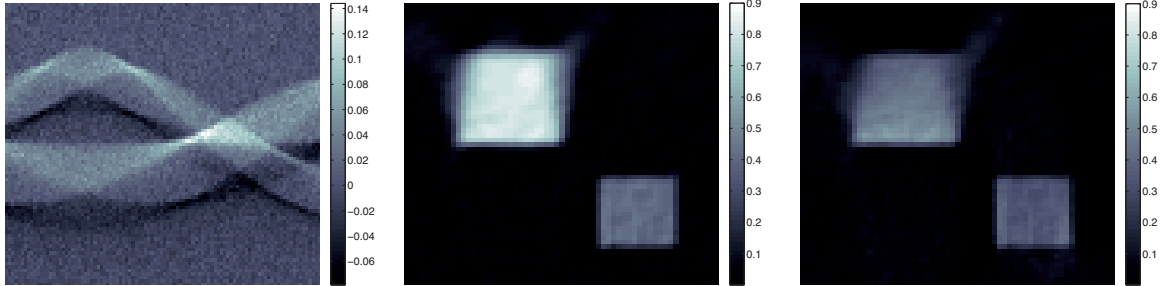

Figure 4.3: RECONSTRUCTION RESULTS FOR NOISY DATA. Pressure data with 5% noise (left), reconstructed absorption coefficient using our single-stage approach (middle) the classical two-stage approach (right).

Finally, in order to investigate the stability of the derived algorithms with respect to noise, we applied the single-stage and the two-stage algorithm after adding Gaussian white noise to the data with standard deviation equal to 5% of the maximal absolute value of the pressure data. The reconstruction results for noisy data are shown in Figure 4.3. As can be seen both algorithms are quite stable with respect to data perturbations. However, again, the single-stage approach yields better results and less artifacts than the two-stage algorithm.

5   Conclusion

In this paper we proposed a single-stage approach for quantitative PAT. For that purpose we derive algorithms that directly recover the optical parameters from the measured acoustical data. This is in contrast to the usual two-stage approach, where the absorbed energy distribution is estimated in a first step, and the optical parameters are reconstructed from the estimated energy distribution in a second step. Our single-stage algorithm is based on generalized Tikhonov regularization and minimization of the Tikhonov functional by the proximal gradient algorithm. In order to show that Tikhonov regularization is well-posed and convergent we analyzed the stationary radiative transfer equation (1.1), (1.2) in a functional analytic framework. For that purpose we used recent results of [20] that guarantees the well-posedness even in the case of voids.

We presented results of our initial numerical studies using a simple limited angle scenario, where the scattering coefficient is assumed to be known. In this situation our single-stage algorithms has led to less artifacts than the two-stage procedure. More detailed numerical studies will be presented in future work. In that context, we will also investigate the use of multiple illuminations and multiple wavelength, which allows to also reconstruct the in general unknown scattering coefficient Grüneissen parameter. We further plan to investigate the use of more general regularization functionals such as the total variation in combination with the two-stage approach.

Acknowledgements

References

[1] M. Agranovsky, P. Kuchment, and L. Kunyansky. On reconstruction formulas and algorithms for the thermoacoustic tomography. In L. V. Wang, editor, *Photoacoustic imaging and spectroscopy*, chapter 8, pages 89–101. CRC Press, 2009.

[2] H. Ammari, E. Bossy, V. Jugnon, and H. Kang. Reconstruction of the optical absorption coefficient of a small absorber from the absorbed energy density. *SIAM J. Appl. Math.*, 71(3):676–693, 2011.

[3] S. R. Arridge. Optical tomography in medical imaging. *Inverse Probl.*, 15(2):R41–R93, 1999.

[4] H. Attouch, J. Bolte, and B. F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. *Math. Program.*, 137(1-2):91–129, 2013.

[5] G. Bal, A. Jollivet, and V. Jugnon. Inverse transport theory of photoacoustics. *Inverse Probl.*, 26:025011, 2010.

[6] G. Bal and K. Ren. Multi-source quantitative photoacoustic tomography in a diffusive regime. *Inverse Probl.*, 27(7):075003, 20, 2011.

[7] P. Beard. Biomedical photoacoustic imaging. *Interface focus*, 1(4):602–631, 2011.

[8] J. Bolte, S. Sabach, and M. Teboulle. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Math. Program.*, 146(1-2, Ser. A):459–494, 2014.

[9] P. Burgholzer, J. Bauer-Marschallinger, H. Grün, M. Haltmeier, and G. Paltauf. Temporal back-projection algorithms for photoacoustic tomography with integrating line detectors. *Inverse Probl.*, 23(6):S65–S80, 2007.

[10] J. Chen and Y. Yang. Quantitative photo-acoustic tomography with partial data. *Inverse Probl.*, 28(11):115014 (15pp), 2012.

[11] E. Chouzenoux, J.-C. Pesquet, and A. Repetti. Variable metric forward-backward algorithm for minimizing the sum of a differentiable function and a convex function. *J. Optim. Theory Appl.*, 162(1):107–132, 2014.

[12] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering*, pages 185–212. Springer, 2011.

[13] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Model. Sim.*, 4(4):1168–1200, 2005.

[14] B. Cox, J. G. Laufer, S. R. Arridge, and Paul C. Beard. Quantitative spectroscopic photoacoustic imaging: a review. *J. Biomed. Opt.*, 17(6):0612021, 2012.

[15] B. T. Cox, S. A. Arridge, and P. C. Beard. Gradient-based quantitative photoacoustic image reconstruction for molecular imaging. In *Proc. SPIE 6437*, page 64371T, 2007.

[16] B. T. Cox, S. R. Arridge, P. Köstli, and P. C. Beard. Two-dimensional quantitative photoacoustic image reconstruction of absorption distributions in scattering media by use of a simple iterative method. *Appl. Opt.*, 45(8):1866–1875, 2006.

[17] B. T. Cox, T. Tarvainen, and S. R. Arridge. Multiple illumination quantitative photoacoustic tomography using transport and diffusion models. In G. Bal, D. Finch, P. Kuchment, J. Schotland, P. Stefanov, and G. Uhlmann, editors, *Tomography and Inverse Transport Theory*, volume 559 of *Contemporary Mathematics*, pages 1–12. AMS, 2011.

[18] R. Dautray and J. Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 6.* Springer-Verlag, Berlin, 1993.

[19] A. De Cezaro and T. F. De Cezaro. Regularization approaches for quantitative photoacoustic tomography using the radiative transfer equation. http://arXiv:1307.3201, 2013.

[20] H. Egger and M. Schlottbom. An L^p theory for stationary radiative transfer. *Appl. Anal.*, 93(6):1283–1296, 2014.

[21] H. Egger and M. Schlottbom. Numerical methods for parameter identification in stationary radiative transfer. *Comput. Optim. Appl.*, pages 1–17, 2014.

[22] H. Egger and M. Schlottbom. Stationary radiative transfer with vanishing absorption. *Math. Models Methods Appl. Sci.*, 24(5):973–990, 2014.

[23] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. Amer. Math. Soc., Providence, RI, 1998.

[24] F. Filbir, S. Kunis, and R. Seyfried. Effective discretization of direct reconstruction schemes for photoacoustic imaging in spherical geometries. *SIAM J. Numer. Anal.*, 52(6):2722–2742, 2014.

[25] D. Finch, M. Haltmeier, and Rakesh. Inversion of spherical means and the wave equation in even dimensions. *SIAM J. Appl. Math.*, 68(2):392–412, 2007.

[26] D. Finch, S. K. Patch, and Rakesh. Determining a function from its mean values over a family of spheres. *SIAM J. Math. Anal.*, 35(5):1213–1240, 2004.

[27] D. Finch and Rakesh. The spherical mean value operator with centers on a sphere. *Inverse Probl.*, 23(6):37–49, 2007.

[28] J. Frikel and E. T. Quinto. Artifacts in incomplete data tomography - with applications to photoacoustic tomography and sonar. arXiv:1407.3453 [math.AP], 2014.

[29] M. Haltmeier. Inversion of circular means and the wave equation on convex planar domains. *Comput. Math. Appl.*, 65(7):1025–1036, 2013.

[30] M. Haltmeier. Universal inversion formulas for recovering a function from spherical means. *SIAM J. Math. Anal.*, 46(1):214–232, 2014.

[31] M. Haltmeier, T. Schuster, and O. Scherzer. Filtered backprojection for thermoacoustic computed tomography in spherical geometry. *Math. Method. Appl. Sci.*, 28(16):1919–1937, 2005.

[32] F. John. *Partial Differential Equations*, volume 1 of *Applied Mathematical Sciences*. Springer Verlag, New York, fourth edition, 1982.

[33] G. Kanschat. Solution of radiative transfer problems with finite elements. In *Numerical methods in multidimensional radiative transfer*, pages 49–98. Springer, Berlin, 2009.

[34] R. Kowar. On time reversal in photoacoustic tomography for tissue similar to water. *SIAM J. Imaging Sci.*, 7(1):509–527, 2014.

[35] R. A. Kruger, K. K. Kopecky, A. M. Aisen, Reinecke D. R., G. A. Kruger, and W. L. Kiser. Thermoacoustic ct with radio waves: A medical imaging paradigm. *Radiology*, 200(1):275–278, 1999.

[36] R. A Kruger, P. Lui, Y. R. Fang, and R. C. Appledorn. Photoacoustic ultrasound (PAUS) – reconstruction tomography. *Med. Phys.*, 22(10):1605–1609, 1995.

[37] P. Kuchment and L. A. Kunyansky. Mathematics of thermoacoustic and photoacoustic tomography. *Eur. J. Appl. Math.*, 19:191–224, 2008.

[38] L. A. Kunyansky. Explicit inversion formulae for the spherical mean Radon transform. *Inverse Probl.*, 23(1):373–383, 2007.

[39] A. V. Mamonov and K. Ren. Quantitative photoacoustic imaging in radiative transport regime. *Comm. Math. Sci.*, 12(2):201–234, 2014.

[40] M. Mokthar-Kharroubi. *Mathematical topics in neutron transport theory*. World Scientific, 1997.

[41] W. Naetar and O. Scherzer. Quantitative photoacoustic tomography with piecewise constant material parameters. *SIAM J. Imaging Sci.*, 7(3):1755–1774, 2014.

[42] V. P. Palamodov. Remarks on the general Funk–Radon transform and thermoacoustic tomography. *Inverse Probl. Imaging*, 4(4):693–702, 2010.

[43] G. Paltauf, R. Nuster, M. Haltmeier, and P. Burgholzer. Experimental evaluation of reconstruction algorithms for limited view photoacoustic tomography with line detectors. *Inverse Probl.*, 23(6):S81–S94, 2007.

[44] K. Ren, H. Gao, and H. Zhao. A hybrid reconstruction method for quantitative PAT. *SIAM J. Imaging Sci.*, 6(1):32–55, 2013.

[45] A. Rosenthal, V. Ntziachristos, and D. Razansky. Acoustic inversion in optoacoustic tomography: A review. *Curr. Med. Imaging Rev.*, 9(4):318–336, 2013.

[46] A. Rosenthal, D. Razansky, and V. Ntziachristos. Fast semi-analytical model-based acoustic inversion for quantitative optoacoustic tomography. *IEEE Trans. Med. Imag.*, 29(6):1275–1285, 2010.

[47] T. Saratoon, T. Tarvainen, B. T. Cox, and S. R. Arridge. A gradient-based method for quantitative photoacoustic tomography using the radiative transfer equation. *Inverse Probl.*, 29(7):075006, 2013.

[48] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational methods in imaging*, volume 167 of *Applied Mathematical Sciences*. Springer, New York, 2009.

[49] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski. *Regularization methods in Banach spaces*, volume 10 of *Radon Series on Computational and Applied Mathematics*. Walter de Gruyter, Berlin, 2012.

[50] P. Stefanov and G. Uhlmann. Thermoacoustic tomography with variable sound speed. *Inverse Probl.*, 25(7):075011, 16, 2009.

[51] T. Tarvainen, B. T. Cox, J. P. Kaipio, and S. A. Arridge. Reconstructing absorption and scattering distributions in quantitative photoacoustic tomography. *Inverse Probl.*, 28(8):084009 (17pp), 2012.

[52] K. Wang and M. Anastasio. Photoacoustic and thermoacoustic tomography: Image formation principles. In *Handbook of Mathematical Methods in Imaging*, chapter 18, pages 781–815. Springer, 2011.

[53] L. V. Wang. Multiscale photoacoustic microscopy and computed tomography. *Nat. Photonics*, 3(9):503–509, 2009.

[54] M. Xu and L. V. Wang. Photoacoustic imaging in biomedicine. *Rev. Sci. Instrum.*, 77(4):041101 (22pp), 2006.

[55] Y. Xu, L. V. Wang, G. Ambartsoumian, and P. Kuchment. Reconstructions in limited-view thermoacoustic tomography. *Med. Phys.*, 31(4):724–733, 2004.

[56] L. Yao, Y. Sun, and J. Huabei. Transport-based quantitative photoacoustic tomography: simulations and experiments. *Phys. Med. Biol.*, 55(7):1917–1934, 2010.

[57] R. J. Zemp. Quantitative photoacoustic tomography with multiple optical sources. *Appl. Opt.*, 49(18):3566–3572, 2010.

[58] X. Zhang, W. Zhou, X. Zhang, and H. Gao. Forward–backward splitting method for quantitative photoacoustic tomography. *Inverse Probl.*, 30(12):125012, 2014.