

Leopold-Franzens-Universität Innsbruck

Nr. 38 09. May 2017

Preprint-Series: Department of Mathematics - Applied Mathematics

Operator learning approach for the limited view problem in photoacoustic tomography

Florian Dreier, Sergiy Pereverzyev Jr., Markus Haltmeier



Technikerstraße 13 - 6020 Innsbruck - Austria Tel.: +43 512 507 53803 Fax: +43 512 507 53898 https://applied-math.uibk.ac.at

# Operator learning approach for the limited view problem in photoacoustic tomography \*

Florian Dreier, Sergiy Pereverzyev Jr. and Markus Haltmeier

May 7, 2017

#### Abstract

In photoacoustic tomography, one is interested to recover the initial pressure distribution inside a tissue from the corresponding measurements of the induced acoustic wave on the boundary of a region enclosing the tissue. In the limited view problem, the wave boundary measurements are given on the part of the boundary, whereas in the full view problem, the measurements are known on the whole boundary. For the full view problem, there exist various fast and accurate reconstruction methods. These methods give severe reconstruction artifacts when they are applied directly to the limited view data. One approach for reducing such artefacts is trying to extend the limited view data to the whole region boundary, and then use existing fast reconstruction methods for the full view data. In this paper, we propose an operator learning approach for constructing an operator that gives an approximate extension of the limited view data. We then consider the behavior of a reconstruction formula on the extended limited view data that is given by our proposed approach. We analyze approximation errors of our approach. This analysis gives recommendations for the choice of the training functions in our operator learning procedure. We also present numerical results that demonstrate the use of the proposed extension of the limited view data in practice. The proposed operator learning approach can also be applied to limited data problems in other tomographic modalities.

**Keywords:** photoacoustic tomography, wave equation, limited view problem, inversion formula, universal back-projection, data extension, operator learning.

AMS subject classifications: 65R32, 35L05, 92C55.

<sup>\*</sup>Department of Mathematics, University of Innsbruck, Technikerstraße 13, A-6020 Innsbruck, Austria, E-mail: florian.dreier@student.uibk.ac.at, sergiy.pereverzyev,markus.haltmeier@uibk.ac.at

# 1 Introduction

Photoacoustic tomography (PAT) is an emerging non-invasive imaging technique. It is based on the photoacoustic effect, and it has a big potential for a successful use in biomedical studies, including preclinical research and clinical practice. Applications include tumor angiogenesis monitoring, blood oxygenation mapping, functional brain imaging, skin melanoma detection [44, 29, 4, 42]

The principle of PAT is the following. When short pulses of non-ionising electromagnetic energy are delivered into a biological (semi-transparent) tissue, then parts of the electromagnetic energy become absorbed. The absorbed energy leads to a nonuniform thermoelastic expansion depending on the tissue structure. This gives rise to an initial acoustic pressure distribution, which further is the source of an ultrasonic wave. These waves are detected by a measurement device on the boundary of the tissue. The mathematical task in PAT is to reconstruct the spatially varying initial pressure distribution using these measurements. The values of the initial pressure distribution inside the tissue allow to make a judgment about the directly unseen structure of the tissue. For example, whether there are some abnormal formations inside the investigated tissue, such as a tumor.

Consider the part of the boundary of a region enclosing the tissue where the wave measurements are available. This part is called observation boundary. If the tissue is fully enclosed by the observation boundary, then one speaks about the full view problem. Otherwise, if some part of the tissue boundary is not accessible, then one has the so-called limited view problem (LVP). The LVP frequently arises in practice, for example in breast imaging (see, e.g., [45, 24]).

The LVP can be approached using iterative reconstruction algorithms (see, e.g., [35, 33, 21, 47, 23, 17, 38]). Although these algorithms can provide accurate reconstruction, they are computationally expensive and time consuming. Approaches for the full view problem, such as time reversal [6, 22], Fourier domain algorithms [14, 27, 46], explicit reconstruction formulas [8, 7, 26, 28, 31], are faster than iterative reconstructions and also accurate. However, when they are directly applied on the limited view data, then one obtains severe reconstruction artifacts.

And so, an idea appears to try to extend the limited view data to the whole boundary, and then use efficient algorithms for the full view data on the extended data to obtain a reconstruction of the initial pressure. Knowing characterizations of the range of the forward operator, which maps the initial pressure distribution to the wave data on the whole boundary of the tissue, may be used for this purpose (see, e.g., [2, 9, 1, 24] and the references therein). This knowledge is expressed with so-called range conditions. In [36, 37], some of these conditions, the so-called moment conditions, were realized for the extension of the limited view data.

The data extension process based on the moment conditions is unstable, and therefore, mostly low frequencies of the limited view wave data can be extended. This instability is connected with the following issue. The observation boundary defines the so-called detection region. It is known (see, e.g., [25, 40, 24]) that if the support of the initial pressure is contained in this detection region, then a stable recovery of the initial pressure



Figure 1: Illustration of the proposed approach for limited view PAT. In the first step, we extend the limited view data to the whole boundary via operator learning. In the second step, we apply a standard direct PAT reconstruction algorithm to the completed data.

from the limited view wave data is possible. However, the data extension process based on the moment conditions does not use information about the support of the initial pressure, and so, it does not employ advantages of the possible stable recovery.

In this paper, we propose a stable method for the extension of the limited view wave data that uses advantages of the mentioned possible stable recovery. Our method is based on the observation that in the case of the stable recovery, there exists a data extension operator that maps the limited view wave data to the unknown wave data on the unobservable part of the boundary. However, this operator is not explicitly known. In our method, we propose to construct an approximate data extension operator using an operator learning approach that is motivated by the methods of the statistical learning theory (see, e.g., [20]). Having an approximately extended limited view wave data, one can employ reconstruction methods for the full view wave data, such as time reversal or methods based on the explicit inversion formulas. As an example, we consider an explicit reconstruction formula for that purpose. We demonstrate that the resulting reconstruction algorithm corrects most limited view reconstruction artifacts, while the computational time remains to be low. The involved steps in the proposed reconstruction approach are illustrated in Figure 1.

The rest of the paper is organized as follows. In section 2, we present a mathematical background for PAT, give the used explicit reconstruction formula, and discuss the LVP. Our operator learning approach to the extension of the limited view wave data is given in section 3. In section 4, we analyze the approximation errors of our approach. We look at the approximation errors for the unknown wave data and for the corresponding reconstructions obtained by explicit reconstruction formulas. We present the numerical results in section 5. Finally, we finish the paper with conclusion and outlook in section 6.

# 2 Mathematics of PAT

Let  $\Omega \subseteq \mathbb{R}^d$  be a bounded domain with a smooth boundary  $\partial\Omega$ , where  $d \geq 2$  denotes the spatial dimension. Further, let  $C_c^{\infty}(\Omega)$  be the set of all smooth functions  $f \colon \mathbb{R}^d \to \mathbb{R}$  that are compactly supported in  $\Omega$ . In PAT, one is interested to recover an unknown function  $f \in C_c^{\infty}(\Omega)$  from the solution of the wave equation given on the boundary of  $\Omega$ . Let us mathematically specify this reconstruction problem.

#### 2.1 Reconstruction problem

Let  $\mathfrak{U}f: \mathbb{R}^d \times (0, \infty) \to \mathbb{R}$  denote the solution of the following initial value problem for the wave equation:

$$\begin{cases} \left(\partial_t^2 - \Delta_x\right) u(x,t) = 0 & \text{for } (x,t) \in \mathbb{R}^d \times (0,\infty), \\ u(x,0) = f(x) & \text{for } x \in \mathbb{R}^d, \\ \left(\partial_t u\right)(x,0) = 0 & \text{for } x \in \mathbb{R}^d. \end{cases}$$
(1)

Here  $\partial_t$  denotes differentiation with respect to the second variable t, and  $\Delta_x$  is the Laplacian with respect to x. Then the reconstruction problem in PAT consists in recovering the unknown function  $f \in C_c^{\infty}(\Omega)$  from the corresponding wave boundary data

$$u(x,t) = (\mathfrak{U}f)(x,t) \quad \text{for } (x,t) \in \Gamma_1 \times (0,\infty), \qquad (2)$$

where  $\Gamma_1 \subseteq \partial \Omega$ . If  $\Gamma_1 = \partial \Omega$ , then (2) is called full view problem; otherwise, if  $\Gamma_1 \subsetneq \partial \Omega$ , we have the limited view problem (LVP). In this paper, we are particularly interested in the limited view case, which we consider in some detail in subsection 2.3.

Let us denote the unobservable part of the boundary as  $\Gamma_2 := \partial \Omega \setminus \Gamma_1$ . We define also the following restrictions of  $\mathfrak{U}f$ :

$$\mathcal{U}f := \mathfrak{U}f|_{\partial\Omega imes(0,\infty)}, \ \mathcal{U}_1f := \mathfrak{U}f|_{\Gamma_1 imes(0,\infty)}, \ \mathcal{U}_2f := \mathfrak{U}f|_{\Gamma_2 imes(0,\infty)}.$$

Let us note that in practice, the reconstruction problem (2) arises in PAT in spatial dimensions two and three. The three dimensional problem appears when the so-called point-like detectors are used (see, for example, [44, 25, 10]). When one uses linear or circular integrating detectors, then the reconstruction problem (2) is considered in two spatial dimensions (see [5, 13, 34, 48]).

#### 2.2 Explicit inversion formula

The reconstruction problem (2) can be approached by various solution techniques. Among these techniques, the derivation of the explicit inversion formulas of the so-called backprojection type is particularly appealing. A numerical realization of these formulas typically gives reconstruction algorithms that are accurate and robust, and at the same time are faster than iterative approaches.

An inversion formula consists of an explicitly given operator  $\mathcal{G}_d$  that recovers the function f from the data u. Such formulas are currently known only for special domains and only for the full view data, i.e. u must be given for all  $x \in \partial \Omega$ . In this paper, we consider the formula that first has been derived in [43, 26, 5]. In addition to the data u, the formula  $\mathcal{G}_d$  also depends on the boundary  $\partial \Omega$  of the domain  $\Omega \subsetneq \mathbb{R}^d$  and on the reconstruction point  $x_0 \in \Omega$ . The structure of the formula further depends on whether the spatial dimension d is even or odd.

If  $d \ge 2$  is an even integer, then

$$\mathcal{G}_d(\partial\Omega, u, x_0) := \kappa_d \int_{\partial\Omega} \langle \nu_x, x_0 - x \rangle \int_{|x_0 - x|}^{\infty} \frac{\left(\partial_t \mathcal{D}_t^{(d-2)/2} t^{-1} u\right)(x, t)}{\sqrt{t^2 - |x_0 - x|^2}} \, \mathrm{d}t \, \mathrm{d}s(x) \,. \tag{3}$$

Here  $\kappa_d := (-1)^{(d-2)/2} / \pi^{d/2}$  is a constant,  $\nu_x$  denotes the outward pointing unit normal to  $\partial\Omega$ , and  $\mathcal{D}_t := (2t)^{-1}\partial_t$  is the differentiation operator with respect to  $t^2$ . Further,  $\langle \cdot, \cdot \rangle$  and  $|\cdot|$  denote the standard inner product and the corresponding Euclidian norm on  $\mathbb{R}^d$ , respectively.

In the case of odd dimension  $d \geq 3$ , the formula  $\mathcal{G}_d$  is defined as follows:

$$\mathcal{G}_d(\partial\Omega, u, x_0) := \kappa_d \int_{\partial\Omega} \frac{\langle \nu_x, x_0 - x \rangle}{|x_0 - x|} \left( \partial_t \mathcal{D}_t^{(d-3)/2} t^{-1} u \right) (x, |x_0 - x|) \,\mathrm{d}s(x) \,, \tag{4}$$

with constant  $\kappa_d := (-1)^{(d-3)/2} / (2\pi^{(d-1)/2}).$ 

The formula  $\mathcal{G}_d$  has been introduced in [43] for dimension d = 3, and in [5] for dimension d = 2. In [26], it has been studied for the case when  $\Omega$  is a ball in arbitrary dimension. Further, in [30, 15, 16], it has been shown that for any elliptical domain  $\Omega$ , the formula  $\mathcal{G}_d$  exactly recovers any smooth function f with support in  $\Omega$  from data  $u = \mathcal{U}f$ . In [18], it was shown, that the same result also holds for parabolic domains  $\Omega$  with d = 2. The formula  $\mathcal{G}_d$  in arbitrary spatial dimension  $d \geq 2$  on certain quadric hypersurfaces, including the parabolic ones, has been analyzed in [19].

It should be noted that the formula  $\mathcal{G}_d$  can be in fact used for any convex bounded domain  $\Omega$ . Then, however, the formula does not recover the function f exactly, and it introduces an approximation error. The form of this error has been analyzed in [30, 15, 16]. Numerical experiments suggest that this error is rather low for domains that can be well approximated by elliptic domains.

The operator  $\mathcal{U}$  can be defined for functions  $f \in \mathcal{L}^2(\Omega_0)$ , where  $\Omega_0$  is an open set with  $\overline{\Omega_0} \subseteq \Omega$ . Define the image of  $\mathcal{L}^2(\Omega_0)$  under the operator  $\mathcal{U}$  as  $\mathbb{Y} := \mathcal{U}(\mathcal{L}^2(\Omega_0))$ . Then it is known (see, e.g., [25, 40, 24]) that  $\mathbb{Y}$  is a closed subspace of  $\mathcal{L}^2(\partial\Omega \times (0, \infty))$ , and therefore, we will treat  $\mathbb{Y}$  as a Hilbert space with the scalar product of  $\mathcal{L}^2(\partial\Omega \times (0, \infty))$ . Moreover, the operator  $\mathcal{U}: \mathcal{L}^2(\Omega_0) \to \mathbb{Y}$  is bounded, and it has the bounded inverse  $\mathcal{U}^{-1}: \mathbb{Y} \to \mathcal{L}^2(\Omega_0)$ .

In the following, we will work with functions  $f \in \mathcal{L}^2(\Omega_0)$ , and we will assume that the domain  $\Omega$  is such that the formula  $\mathcal{G}_d$  gives exact recovery of the function f from its wave data  $u = \mathcal{U}f$ , i.e. it holds that

$$f = \mathcal{G}_d \,\mathcal{U}f.\tag{5}$$

As we already mentioned, this is, for example, the case for circular and elliptical domains. In such a situation, it can be shown that  $\mathcal{G}_d$  is a continuous extension of  $\mathcal{U}^{-1}$  to  $\mathcal{L}^2(\partial\Omega \times (0,\infty))$ .

#### 2.3 Limited view problem

In practice, the wave data u is frequently given on a subset  $\Gamma_1$  of the boundary  $\partial \Omega$  (Figure 2). This subset  $\Gamma_1$ , called observation boundary, defines the so-called detection region



Figure 2: Setting of LVP.

 $\mathbb{D}(\Gamma_1)$  (see, e.g., [33, 25]). If  $\operatorname{supp}(f) \subsetneq \mathbb{D}(\Gamma_1)$ , then the function f in (2) can be stably recovered from data on  $\Gamma_1$ . The detection region  $\mathbb{D}(\Gamma_1)$  contains points x such that any line going through x intersects  $\Gamma_1$ . For example, if  $\Gamma_1$  is a spherical or elliptical cap, then  $\mathbb{D}(\Gamma_1) = \operatorname{conv}(\Gamma_1)$ .

Let us mathematically specify the stable recovery of f. Let  $\Omega_1$  be an open set with  $\overline{\Omega_1} \subseteq \mathbb{D}(\Gamma_1)$ . Define the image of  $\mathcal{L}^2(\Omega_1)$  under the operator  $\mathcal{U}_1$  as  $\mathbb{Y}_1 := \mathcal{U}_1(\mathcal{L}^2(\Omega_1))$ . Then  $\mathbb{Y}_1$  is a closed subspace of  $\mathcal{L}^2(\Gamma_1 \times (0, \infty))$ , and we will consider it as a Hilbert space with the scalar product of  $\mathcal{L}^2(\Gamma_1 \times (0, \infty))$ . The operator  $\mathcal{U}_1 : \mathcal{L}^2(\Omega_1) \to \mathbb{Y}_1$  is bounded, and it has bounded inverse  $\mathcal{U}_1^{-1} : \mathbb{Y}_1 \to \mathcal{L}^2(\Omega_1)$ , which implies stable recovery of f from data in (2).

Denote  $\mathbb{Y}_2 := \mathcal{L}^2(\Gamma_2 \times (0, \infty))$ . From the boundness of the operator  $\mathcal{U} : \mathcal{L}^2(\Omega_0) \to \mathbb{Y}$ , we can derive the boundness of the operator  $\mathcal{U}_2 : \mathcal{L}^2(\Omega_1) \to \mathbb{Y}_2$ . We will use this for the data extension operator below.

In spite of the stable recovery of f in (2), the use of formula  $\mathcal{G}_d$  on the data u given on  $\Gamma_1 \subsetneq \partial \Omega$  leads to serious artifacts in the reconstruction; see, e.g., [18], where the numerical results of the application of  $\mathcal{G}_2$  on finite parabolas are presented. The reconstruction artefacts in the case of the limited view data are also discussed in [45, 11, 41, 3, 12, 32].

At the same time, the use of formula  $\mathcal{G}_d$  for reconstructing function f can be attractive from various points of view. For example, as we already pointed out, the reconstruction using a numerical realization of  $\mathcal{G}_d$  is very fast. Another point may be connected with the nature of the software development. Namely, having already a tested and trusted computer code of the numerical realization of formula  $\mathcal{G}_d$ , it could be easier to develop its extensions for the LVP, than to develop new computer software based on the other approaches to (2).

An extension of the limited view data u from the observable part of the boundary  $\Gamma_1 \subsetneq \partial \Omega$  to the whole boundary  $\partial \Omega$  may give a possibility to improve the reconstruction quality of the formula  $\mathcal{G}_d$ . In this paper, we propose to realize this extension using the operator learning approach, which we consider in the next section.

# 3 Data extension using operator learning approach

The extension of the limited view data to the whole boundary can be in principle done by the extension operator that we define in the next subsection. This operator is however not explicitly known, and we propose an operator learning approach to construct its approximation in subsection 3.2. In subsection 3.3, we discuss computational aspects of the proposed learned approximation of the extension operator.

#### 3.1 Extension operator

The operator  $\mathcal{A}: \mathbb{Y}_1 \to \mathbb{Y}_2$  that maps functions  $\mathcal{U}_1 f$  to functions  $\mathcal{U}_2 f$  for  $f \in \mathcal{L}^2(\Omega_1)$ realizes the extension of the limited view data  $u_1 = \mathcal{U}_1 f$  to the unobservable part of the boundary  $\Gamma_2$ . This operator  $\mathcal{A}$  can be written as  $\mathcal{A} = \mathcal{U}_2 \circ \mathcal{U}_1^{-1}$ . Because of this representation and the assumptions on  $\Gamma_1$  and  $\Omega_1$ , the operator  $\mathcal{A}$  is a linear continuous operator.

With the introduced extension operator  $\mathcal{A}$ , one could extend the limited view data  $u_1$  to the whole boundary  $\partial\Omega$ , and then use the formula  $\mathcal{G}_d$  on this extended data. In this way, the disadvantages of the use of the formula  $\mathcal{G}_d$  on the limited view data can be eliminated. However, the form of the operator  $\mathcal{A}$  is not explicitly known.

#### 3.2 Learned extension operator

In this paper, we propose to construct an operator  $\hat{\mathcal{A}}_n$  that approximates the operator  $\mathcal{A}$ . The parameter n is defined below. The approximate operator  $\hat{\mathcal{A}}_n$  must satisfy the following two requirements. The first requirement concerns the approximation quality:  $\hat{\mathcal{A}}_n u_1$  must be close to  $\mathcal{A}u_1$ . The second requirement is related to the computational effort of the numerical evaluation of  $\hat{\mathcal{A}}_n u_1$ . This evaluation must be fast such that the evaluation of the formula  $\mathcal{G}_d$  on the extended limited view data with the help of  $\hat{\mathcal{A}}_n$  remains to be computationally efficient.

Our construction of the approximate operator  $\hat{\mathcal{A}}_n$  is motivated by the statistical learning approach (see, e.g., [20]). For  $i = 1, \ldots, n$ , consider training functions  $f_i: \Omega_1 \to \mathbb{R}$ . For each training function  $f_i$ , we can determine the corresponding wave data  $u_{1,i} := \mathcal{U}_1 f_i$ ,  $u_{2,i} := \mathcal{U}_2 f_i$ . By the definition of the extension operator  $\mathcal{A}$  we have that  $u_{2,i} = \mathcal{A} u_{1,i}$ . In the context of statistical learning, the set  $\mathcal{Z} := \{ (u_{1,i}, \mathcal{A} u_{1,i}), i = 1, \ldots, n \}$  is called a training set. Define for future reference the set  $\mathbf{U}_{1,n} := \{ u_{1,i}, i = 1, \ldots, n \}$ .

So, how to construct (or, using the terminology of the statistical learning, how to learn) an approximation  $\hat{\mathcal{A}}_n u_1$  of  $\mathcal{A}u_1$  using the training set  $\mathcal{Z}$ ? We propose to use a projection operator for this purpose. Define a linear subspace

$$V_{n} := \left\{ \sum_{j=1}^{n} c_{j} u_{1,j}, \ c_{j} \in \mathbb{R} \right\}, \ V_{0} := \{ \ 0 \ \} \subseteq \mathbb{Y}_{1}, \tag{6}$$

and let  $\mathcal{P}_n: \mathcal{L}^2(\Gamma_1 \times (0, \infty)) \to V_n$  be the orthogonal projection on  $V_n$  in  $\mathcal{L}^2(\Gamma_1 \times (0, \infty))$ .

Then we define the learned approximation  $\hat{\mathcal{A}}_n u_1$  as follows:

$$\hat{\mathcal{A}}_n u_1 := \mathcal{A} \mathcal{P}_n u_1. \tag{7}$$

Note that  $V_n \subseteq \mathbb{Y}_1$ , and therefore, the operator composition  $\mathcal{AP}_n$  is well-defined, and  $\hat{\mathcal{A}}_n : \mathcal{L}^2(\Gamma_1 \times (0,\infty)) \to \mathbb{Y}_2$ . Further, note that for all  $u_1 \in \mathcal{L}^2(\Gamma_1 \times (0,\infty))$ ,  $\hat{\mathcal{A}}_0 u_1 = 0 \in \mathbb{Y}_2$ .

#### 3.3 Computation of learned approximation

How to compute the learned approximation  $\mathcal{A}_n u_1$  using the training set  $\mathcal{Z}$  for  $n \geq 1$ ? First of all, observe that since  $\mathcal{P}_n u_1 \in V_n$ , the projection  $\mathcal{P}_n u_1$  has the following representation:

$$\mathcal{P}_n u_1 = \sum_{j=1}^n c_j u_{1,j},\tag{8}$$

where the coefficients  $c_j \in \mathbb{R}$  can be determined from the conditions  $\langle \mathcal{P}_n u_1 - u_1, u_{1,i} \rangle = 0$ for  $i = 1, \ldots, n$ . These conditions can be written in the form of the system of linear equations for the coefficients  $c_j$ 

$$\sum_{j=1}^{n} c_j \langle u_{1,i}, u_{1,j} \rangle = \langle u_1, u_{1,i} \rangle, \ i = 1, \dots, n.$$
(9)

Denote the matrix corresponding to the above linear system as  $\mathbf{P}_n$ , i.e. the elements of  $\mathbf{P}_n$  are  $(\mathbf{P}_n)_{ij} = \langle u_{1,i}, u_{1,j} \rangle$ . Further, denote the vector of unknowns as  $\mathbf{c}_n$ , and the right-hand side as  $\mathbf{u}_n$ , i.e.  $(\mathbf{c}_n)_i = c_i$  and  $(\mathbf{u}_n)_i = \langle u_1, u_{1,i} \rangle$ .

The matrix  $\mathbf{P}_n$  is the Gram matrix of the functions in  $\mathbf{U}_{1,n}$ , and it is invertible if the set  $\mathbf{U}_{1,n}$  is linearly independent. Since the operator  $\mathcal{U}_1$  is invertible, the set  $\mathbf{U}_{1,n}$  is linearly independent if the set  $\{f_i, i = 1, ..., n\}$  is linearly independent, and for the following, we assume that this is the case.

Note that the matrix  $\mathbf{P}_n$  does not depend on the limited view wave data  $u_1$  that we want to extend. Therefore, the inverse matrix  $\mathbf{P}_n^{-1}$  can be precomputed once the set of the learning inputs  $\mathbf{U}_{1,n}$  is given. This will make the determination of the coefficients  $c_j$  very fast.

Finally, with the coefficients  $c_j$  in (8), i.e.  $\mathbf{c}_n = \mathbf{P}_n^{-1}\mathbf{u}_n$ , the approximation  $\hat{\mathcal{A}}_n u_1$  is calculated as follows:

$$\hat{\mathcal{A}}_n u_1 = \mathcal{A} \mathcal{P}_n u_1 = \mathcal{A} \left( \sum_{j=1}^n c_j u_{1,j} \right) = \sum_{j=1}^n c_j u_{2,j} = \sum_{j=1}^n c_j \mathcal{U}_2 f_j.$$

### 4 Approximate reconstructions and their error analysis

For obtaining an approximate reconstruction of f using the limited view data  $u_1 = \mathcal{U}_1 f$ and the formula  $\mathcal{G}_d$ , we can now proceed as follows. First, we extend the limited view data  $u_1$  to the whole boundary  $\partial \Omega$  using the learned extension operator  $\hat{\mathcal{A}}_n$  in this way:

$$\hat{u}_n(x,t) = \begin{cases} u_1(x,t) & \text{if } x \in \Gamma_1, \\ \left(\hat{\mathcal{A}}_n u_1\right)(x,t) & \text{if } x \in \Gamma_2. \end{cases}$$

And then we apply the formula  $\mathcal{G}_d$  to this extended wave data  $\hat{u}_n$  in order to obtain an approximate reconstruction  $\hat{f}_n$ :

$$\hat{f}_n = \mathcal{G}_d \hat{u}_n. \tag{10}$$

Note that  $\hat{u}_0$  is obtained by extending the limited view data  $u_1$  to the whole boundary  $\partial \Omega$  with zero values on  $\Gamma_2$ . As we already discussed, the corresponding approximate reconstruction  $\hat{f}_0$  contains significant errors, and we wish to have better reconstructions of f using  $u_1$ . In particular, one may desire that the reconstruction  $\hat{f}_n$  improves as n increases.

In the following theorem, we estimate the  $\mathcal{L}^2$ -error of the approximation of  $\mathcal{A}u_1$  by  $\hat{\mathcal{A}}_n u_1$ and of the approximation of f by  $\hat{f}_n$ . From the derived estimates in this theorem, we will see that the above wish can be realized if the training functions  $f_i$ ,  $i = 1, \ldots, n$ , are chosen appropriately.

**Theorem 1.** Let a set of linearly independent training functions  $\{f_i, i = 1, ..., n\} \subseteq \mathcal{L}^2(\Omega_1)$  be given, and  $W_n := \{\sum_{i=1}^n c_i f_i, c_i \in \mathbb{R}\}, W_0 := \{0\} \subseteq \mathcal{L}^2(\Omega_1)$ . Define the training limited view wave data  $u_{1,i} := \mathcal{U}_1 f_i$ , the corresponding linear subspace  $V_n$  in (6), and the learned extension operator  $\hat{\mathcal{A}}_n$  in (7). Consider a function  $f \in \mathcal{L}^2(\Omega_1)$ , its limited view wave data  $u_1 := \mathcal{U}_1 f$ , and its approximation  $\hat{f}_n$  defined in (10). Then the following  $\mathcal{L}^2$ -error estimate for the unobservable data holds:

$$\left\| \mathcal{A}u_1 - \hat{\mathcal{A}}_n u_1 \right\| \le \left\| \mathcal{A} \right\| \cdot \left\| \mathcal{U}_1 \right\| \cdot \min_{g \in W_n} \left\| f - g \right\|.$$
(11)

If additionally, the domain  $\Omega$  is such that (5) holds, then we have the following  $\mathcal{L}^2$ -error estimate for the reconstruction:

$$\left\| f - \hat{f}_n \right\| \le \left\| \mathcal{G}_d \right\| \cdot \left\| \mathcal{A} \right\| \cdot \left\| \mathcal{U}_1 \right\| \cdot \min_{g \in W_n} \left\| f - g \right\|.$$
(12)

*Proof.* We first prove (11). From the definition of the operator  $\mathcal{A}_n$ , we have that

$$\left\| \mathcal{A}u_1 - \hat{\mathcal{A}}_n u_1 \right\| = \left\| \mathcal{A}\mathcal{U}_1 f - \mathcal{A}\mathcal{P}_n \mathcal{U}_1 f \right\| \le \left\| \mathcal{A} \right\| \cdot \left\| \mathcal{U}_1 f - \mathcal{P}_n \mathcal{U}_1 f \right\|.$$
(13)

From the properties of the projection operators, we also have that

$$\left\|\mathcal{U}_{1}f - \mathcal{P}_{n}\mathcal{U}_{1}f\right\| = \min_{h \in V_{n}}\left\|\mathcal{U}_{1}f - h\right\|.$$
(14)

For an element  $h \in V_n$ , there are unique constants  $c_i \in \mathbb{R}$ , i = 1, ..., n such that

$$h = \sum_{i=1}^{n} c_i \, u_{1,i} = \sum_{i=1}^{n} c_i \, \mathcal{U}_1 f_i = \mathcal{U}_1 \left( \sum_{i=1}^{n} c_i \, f_i \right),$$

and therefore, there exists an element  $g \in W_n$  such that  $h = \mathcal{U}_1 g$ . Using this fact, we can estimate

$$\min_{h \in V_n} \|\mathcal{U}_1 f - h\| = \min_{g \in W_n} \|\mathcal{U}_1 f - \mathcal{U}_1 g\| \le \|\mathcal{U}_1\| \cdot \min_{g \in W_n} \|f - g\|.$$
(15)

Then combining (13),(14),(15), we obtain estimate (11) for the  $\mathcal{L}^2$ -error  $\left\|\mathcal{A}u_1 - \hat{\mathcal{A}}_n u_1\right\|$ . Now, consider (12). Using (5) and (10), we have

$$\left\| f - \hat{f}_n \right\| = \left\| \mathcal{G}_d \mathcal{U}f - \mathcal{G}_d \hat{u}_n \right\| \le \left\| \mathcal{G}_d \right\| \cdot \left\| \mathcal{U}f - \hat{u}_n \right\|.$$
(16)

Since  $(\mathcal{U}f)(x,t) = \hat{u}_n(x,t) = u_1(x,t)$  for  $x \in \Gamma_1$ , then

$$\left\|\mathcal{U}f - \hat{u}_n\right\| = \left\|\mathcal{U}_2 f - \hat{\mathcal{A}}_n u_1\right\| = \left\|\mathcal{A}u_1 - \hat{\mathcal{A}}_n u_1\right\|.$$
(17)

Thus, the error estimate (12) is obtained from (16), (17), and the error estimate (11).  $\Box$ 

**Remark 1.** Let  $\mathcal{Q}_n \colon \mathcal{L}^2(\Omega_1) \to W_n$  be the orthogonal projection on  $W_n$  in the space  $\mathcal{L}^2(\Omega_1)$ . Then, since we have that  $\min_{g \in W_n} ||f - g|| = ||f - \mathcal{Q}_n f||$ , we can write  $||f - \mathcal{Q}_n f||$  instead of  $\min_{g \in W_n} ||f - g||$  in (11) and (12).

As we see from Theorem 1, the estimates of the  $\mathcal{L}^2$ -errors given by our learning procedure depend on the minimal distance from the unknown function f to the linear subspace  $W_n$ defined by the training functions  $f_i$ . This gives us an indication for the choice of the training functions. Namely, one should choose the training functions  $f_i$  such that the unknown function f can be well approximated by their linear combination.

Estimates (11), (12) also allow us to state the condition for the exact approximation given by our learning procedure and for the convergence of the learned approximation when the number of the training functions n goes to infinity. We present these conditions in the following two corollaries.

**Corollary 1.** If  $f \in W_n$ , then the learned approximation  $\hat{\mathcal{A}}_n u_1$  and the reconstruction  $\hat{f}_n$  are exact, i.e.

$$\left\|\mathcal{A}u_1 - \hat{\mathcal{A}}_n u_1\right\| = \left\|f - \hat{f}_n\right\| = 0.$$

**Corollary 2.** If  $\overline{\bigcup_{n\geq 1} W_n} = \mathcal{L}^2(\Omega_1)$ , then the learned approximation  $\hat{\mathcal{A}}_n u_1$  and the reconstruction  $\hat{f}_n$  converge correspondingly to  $\mathcal{A}u_1$  and f as  $n \to \infty$ , i.e.

$$\lim_{n \to \infty} \left\| \mathcal{A}u_1 - \hat{\mathcal{A}}_n u_1 \right\| = \lim_{n \to \infty} \left\| f - \hat{f}_n \right\| = 0.$$

Let us now compare the errors of the approximations  $\hat{f}_n$  with  $n \ge 1$  and  $\hat{f}_0$ . The  $\mathcal{L}^2$ -error estimates (11),(12) for n = 0 become:

$$\|\mathcal{A}u_1 - 0\| \le \|\mathcal{A}\| \cdot \|\mathcal{U}_1\| \cdot \|f\|,$$
 (18)

$$\left\| f - \hat{f}_0 \right\| \le \left\| \mathcal{G}_d \right\| \cdot \left\| \mathcal{A} \right\| \cdot \left\| \mathcal{U}_1 \right\| \cdot \left\| f \right\|.$$
(19)

Comparing the error estimates (11),(12) for the learned approximations with  $n \ge 1$  and the error estimates (18),(19) for the approximations using zero extension of the limited view wave data, one sees that these error estimates differ regarding the following factors:

$$\mathcal{E}_{n}(f) := \min_{g \in W_{n}} \|f - g\|, \quad \mathcal{E}_{0}(f) := \|f\|,$$
(20)

correspondingly for learned approximations with  $n \ge 1$  and approximation using zero extension.

The factors (20) can be seen as indicators for the expected approximation quality of the considered algorithms. For a fixed non-zero function f, the factor  $\mathcal{E}_0(f)$  is a fixed non-zero value, while the factor  $\mathcal{E}_n(f)$  can be zero, or can be made arbitrary small, see Corollaries 1,2. Therefore, we may expect that the approximation quality of the learned approximations is better than of the approximations using zero extension of the data. This expectation will be confirmed by the numerical results in the next section.

In fact, one can show (see Remark 2 below) that the factor  $\mathcal{E}_n(f)$  is always less or equal than the factor  $\mathcal{E}_0(f)$ , and the strict inequality  $\mathcal{E}_n(f) < \mathcal{E}_0(f)$  holds under rather mild condition on the function f and the training functions  $f_i$ . Generally, this condition can be expected to hold in practice. This additionally demonstrates the effectiveness of our approach based on the learned extension of the data in comparison to the approach using zero extension.

**Remark 2.** Using properties of the projection operators in Hilbert spaces, one can show that the sequence  $\mathcal{E}_n(f)$  is nonincreasing, i.e.

$$\mathcal{E}_n(f) \le \mathcal{E}_m(f) \quad for \quad n > m \ge 0.$$
 (21)

If additionally

$$\langle f, f_i \rangle \neq 0 \quad \text{for some} \quad i \in \{ m+1, \dots, n \},$$

$$(22)$$

then inequality (21) is strict, i.e.

$$\mathcal{E}_n(f) < \mathcal{E}_m(f) \quad for \quad n > m \ge 0.$$
(23)

Condition (22) is also necessary for (23), i.e. if (23) holds, then we have (22).

### 5 Numerical results

In this section, we present results of the numerical realization of the proposed operator learning approach.

We consider the spatial dimension d = 2, and we take the elliptical domain

$$\Omega = \left\{ \left( x_1, x_2 \right) \in \mathbb{R}^2 \mid \left( x_1/a_1 \right)^2 + \left( x_2/a_2 \right)^2 < 1 \right\},\$$



Figure 3: Left: the function f that we use in our numerical experiments and the chosen observation boundary  $\Gamma_1$ . Right: the corresponding numerical full view wave boundary data  $\mathcal{U}f$ . The region between two white vertical lines corresponds to the unknown part of the data on the unobservable part of the boundary  $\Gamma_2$ .

with  $a_1 = 2, a_2 = 1$ . We use the following parametrization of the boundary

$$\partial \Omega = \{ (a_1 \cos \theta, a_2 \sin \theta) \mid \theta \in [-\pi, \pi) \},\$$

and we assume that the unobservable part of the boundary is (see Figure 3(left))

$$\Gamma_2 = \{ (a_1 \cos \theta, a_2 \sin \theta) \mid \theta \in [0.97, 2.17) \}$$

We work with the function f presented in Figure 3(left). Its numerical full view wave boundary data  $u = \mathcal{U}f$  is given in Figure 3(right), and we use the corresponding limited view wave boundary data  $u_1 = \mathcal{U}_1 f$ . The observation boundary  $\Gamma_1$  is discretized such that the distance between two consecutive points is in the interval [0.0099, 0.0101]. We take the time step size as 0.01.

We further assume that we know a rectangular region

$$K = \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid -1.25 \le x_1 < 0.5, \ -0.7 \le x_2 < 0.1752 \right\}$$

containing supp(f) (Figure 4(top and bottom)). We use this region K for defining training functions  $f_i$ . Namely, we consider partitions of the region K into squares  $K_i$ ,  $i \in \{1, \ldots, n\}$ . The square  $K_i$  contains points  $(x_1, x_2) \in \mathbb{R}^2$  such that

$$-1.25 + (\lceil i/n_h \rceil - 1) w/n_w \le x_1 < -1.25 + \lceil i/n_h \rceil w/n_w, -0.7 + (i \mod n_h - 1) h/n_h \le x_2 < -0.7 + (i \mod n_h) h/n_h,$$

where w = 1.75 (width of K), h = 0.8752 (height of K),  $n_w = \sqrt{2n}$ ,  $n_h = n_w/2$  (see Figure 4(middle)). Then we define the training function  $f_i$  as the indicator function of the square  $K_i$ . We take the number of the training functions in the form  $n = n_1 \times n_2$ , where  $n_1$  and  $n_2$  are the numbers of the partitioning intervals along the coordinate  $x_1$  and  $x_2$  correspondingly. We present the numerical results for  $n = 4 \times 2$ ,  $8 \times 4$ ,  $16 \times 8$ ,  $32 \times 16$ .



Figure 4: Top: the rectangular region K containing  $\operatorname{supp}(f)$ . Middle: the example of the partition of K into  $8 \times 4$  squares. The training functions  $f_i$  are numbered starting from the bottom-left square from bottom to top and from left to right. Bottom: the position of  $\operatorname{supp}(f)$  in K with the partition of K into  $8 \times 4$  squares.

The extended limited view data  $\hat{u}_n$  using the learned extension operator  $\hat{\mathcal{A}}_n$  for the considered values of n are presented in Figure 5. We observe that as n increases, the extended data  $\hat{u}_n$  approaches the full view data u in Figure 3(right). Note that the chosen training functions  $f_i$  satisfy the condition of Corollary 2. Therefore, the approach of  $\hat{u}_n$  to the full view data u is in agreement with our theoretical analysis.

The reconstructions  $\hat{f}_n$  using the extended data  $\hat{u}_n$  are presented in Figure 6(2nd and 3rd rows). For comparison purpose, we also present the reconstruction  $\hat{f}$  using the full view wave boundary data u, and the reconstruction  $\hat{f}_0$  using the zero extended data  $\hat{u}_0$  (Figure 6(1st row)). We evaluate the reconstructions at the points from the discrete set

$$\Omega_h := \left\{ \left( -2.2 + n_1 h, -2.2 + n_2 h \right) \in \mathbb{R}^2 \mid n_1, n_2 \in \{ 0, 1, \dots, 300 \} \right\} \cap \Omega,$$

with h = 11/750. We also consider the discrete  $\mathcal{L}^2$ -error of a reconstruction  $\hat{f}_*$  defined as follows:

$$E_2\left(\hat{f}_*\right) := \left(\sum_{x \in \Omega_h} \left| f(x) - \hat{f}_*(x) \right|^2 \cdot h^2 \right)^{1/2}.$$



Figure 5: The extended limited view data  $\hat{u}_n$  using the learned extension operator  $\hat{\mathcal{A}}_n$  for  $n = 4 \times 2, 8 \times 4, 16 \times 8, 32 \times 16$  (from left to right and from top to bottom). The gray scaling is as in Figure 3 (right).

Let us discuss the reconstructions in Figure 6. First of all, as expected, one observes strong artifacts in the reconstruction  $\hat{f}_0$ , especially outside of  $\operatorname{supp}(f)$ . These artifacts are considerably corrected in the reconstruction  $\hat{f}_{4\times 2}$ , and as the number of the training functions *n* increases, the artifacts become weaker such that the reconstruction  $\hat{f}_{32\times 16}$  is very similar to the reconstruction  $\hat{f}$ . This observation is also reflected in  $E_2$ -errors that are presented in Figure 7. Note that  $\hat{f}$  differs from *f* due to the discretization error of the numerical realization of the formula  $\mathcal{G}_2$ . Thus, as in the case of the data  $\hat{u}_n$ , the approach of  $\hat{f}_n$  to *f* is in agreement with Corollary 2.

Finally, in Table 1, we present the calculation times for the parts involved in the proposed reconstruction approach. Our numerical results are performed with MATLAB version R2015b on the PC lenovo e31 with four processors Intel(R) Xeon(R) CPU 3.20GHz. We see that the most time consuming part is the calculation of the matrix  $\mathbf{P}_n^{-1}$ , which is used for solving the system of linear equations (9). But for a given set of the training functions, this matrix has to be calculated only once and prior to the actual image reconstruction process.

The calculation of the learned data extension  $\hat{\mathcal{A}}_n u_1$  is fast. In particular, for the biggest considered number  $n = 32 \times 16$  of the training functions, the calculation time for  $\hat{\mathcal{A}}_n u_1$ is near the calculation time for the formula  $\mathcal{G}_2$ . Thus, our proposed operator learning approach fulfills the requirements that we stated at the beginning of Section 3.2. Namely,



Figure 6: From left to right and from top to bottom: the reconstructions  $\hat{f}$ ,  $\hat{f}_0$ , and  $\hat{f}_n$ , for  $n = 4 \times 2$ ,  $8 \times 4$ ,  $16 \times 8$ ,  $32 \times 16$ .

the closeness of the approximation  $\hat{\mathcal{A}}_n u_1$  to  $\mathcal{A} u_1$ , and the fast evaluation of  $\hat{\mathcal{A}}_n u_1$  are realized.

# 6 Conclusion and outlook

In this paper, we demonstrated that an approximate extension of limited view data in PAT can be realized using an operator learning approach. Our error analysis gives recommendations for the choice of the training functions in the proposed operator learning. Our numerical results show that the learned extension of the limited view data with a high approximation quality and a low computational cost is possible. This makes the proposed learned data extension attractive for the algorithms that are designed for the full view data. As an example, we demonstrated a satisfactory performance of a reconstruction formula with the proposed learned data extension.



Figure 7:  $E_2$ -errors of the considered reconstructions  $\hat{f}$ ,  $\hat{f}_0$ , and  $\hat{f}_n$ , for  $n = 4 \times 2$ ,  $8 \times 4$ ,  $16 \times 8$ ,  $32 \times 16$ .

Table 1: Calculation times in seconds for the parts involved in the proposed reconstruction approach.

| n              | $\mathbf{P}_n^{-1}$ | $\hat{\mathcal{A}}_n u_1$ | $\mathcal{G}_2$ |
|----------------|---------------------|---------------------------|-----------------|
| $4 \times 2$   | 1179.73             | 0.53                      | 4.20            |
| $8 \times 4$   | 4707.31             | 0.68                      | 3.55            |
| $16 \times 8$  | 19036.23            | 1.41                      | 3.90            |
| $32 \times 16$ | 75874.87            | 6.07                      | 4.33            |

It could be interesting to look at the behavior of the proposed learned data extension without knowledge of a rectangular region K containing  $\operatorname{supp}(f)$ . In this case, one could consider partitions of the whole detection region  $\Omega_1$ . Also other training functions, such as generalized Kaiser-Bessel functions (see, e.g., [39]), can be tried.

It is appealing to consider a comparison of the reconstruction quality and computation time of the proposed reconstruction approach and iterative reconstruction algorithms. Realization of the proposed learned extension of the limited view data to three spatial dimensions is also interesting to do.

Finally, it seems to be worth to examine applications of the presented operator learning approach to the limited data problems in other tomographic modalities, such as sparse angle computed tomography, region of interest fan beam computed tomography, positron emission tomography, magnetic resonance imaging.

# Acknowledgements

Authors gratefully acknowledge the support of the Tyrolean Science Fund (TWF). Sergiy Pereverzyev Jr. gratefully acknowledges the support of the Austrian Science Fund (FWF): project P 29514-N32. He also would like to thank Alessandro Verri, Vera Kurkova, Linh Nguyen, Jürgen Frikel, Xin Guo, Ding-Xuan Zhou, and members of Ding-Xuan Zhou's group at the City University of Hong Kong for discussions concerning this work.

# References

- M. Agranovsky, D. Finch, and P. Kuchment. Range conditions for a spherical mean transform. *Inverse Probl. Imaging*, 3(3):373–383, 2009.
- [2] G. Ambartsoumian and P. Kuchment. A range description for the planar circular Radon transform. SIAM J. Math. Anal., 38(2):681–692, 2006.
- [3] L. L. Barannyk, J. Frikel, and L. V. Nguyen. On artifacts in limited data spherical Radon transform: curved observation surface. *Inverse Probl.*, 32(1):015012, 2015.
- [4] P. Beard. Biomedical photoacoustic imaging. Interf. Focus, 1(4):602–631, 2011.
- [5] 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.
- [6] P. Burgholzer, G. J. Matt, M. Haltmeier, and G. Paltauf. Exact and approximate imaging methods for photoacoustic tomography using an arbitrary detection surface. *Phys. Rev. E*, 75(4):046706, 2007.
- [7] 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.
- [8] D. Finch, S. Patch, and Rakesh. Determining a function from its mean values over a family of spheres. SIAM J. Math. Anal., 35(5):1213–1240, 2004.
- [9] D. Finch and Rakesh. The range of the spherical mean value operator for functions supported in a ball. *Inverse Probl.*, 22(3):923–938, 2006.
- [10] D. Finch and Rakesh. Recovering a function from its spherical mean values in two and three dimensions. In L. V. Wang, editor, *Photoacoustic imaging and spectroscopy*, chapter 7, pages 77–88. CRC Press, 2009.
- [11] J. Frikel and E. T. Quinto. Characterization and reduction of artifacts in limited angle tomography. *Inverse Probl.*, 29(12):21, 2013.
- [12] J. Frikel and E. T. Quinto. Artifacts in incomplete data tomography with applications to photoacoustic tomography and sonar. SIAM J. Appl. Math., 75(2):703–725, 2015.
- [13] H. Grün, T. Berer, P. Burgholzer, R. Nuster, and G. Paltauf. Three-dimensional photoacoustic imaging using fiber-based line detectors. J. Biomed. Optics, 15(2):021306, 2010.
- [14] M. Haltmeier. Frequency domain reconstruction for photo- and thermoacoustic tomography with line detectors. Math. Mod. Meth. Appl. Sci., 19(2):283–306, 2009.

- [15] M. Haltmeier. Inversion of circular means and the wave equation on convex planar domains. *Comput. Math. Appl.*, 65(7):1025–1036, 2013.
- [16] M. Haltmeier. Universal inversion formulas for recovering a function from spherical means. SIAM J. Math. Anal., 41(1):214–232, 2014.
- [17] M. Haltmeier and L. V. Nguyen. Analysis of iterative methods in photoacoustic tomography with variable sound speed. *SIAM J. Imaging Sci.*, 2017, to appear. arXiv:1611.07563.
- [18] M. Haltmeier and S. Pereverzyev Jr. Recovering a function from circular means or wave data on the boundary of parabolic domains. SIAM J. Imaging Sci., 8(1):592–610, 2015.
- [19] M. Haltmeier and S. Pereverzyev Jr. The universal back-projection formula for spherical means and the wave equation on certain quadric hypersurfaces. J. Math. Anal. Appl., 429(1):366–382, 2015.
- [20] T. Hastie, R. Tibshirani, and J. Friedman. The Elements of Statistical Learning. Springer Series in Statistics, 2009.
- [21] G. T. Herman. Fundamentals of computerized tomography: image reconstruction from projections. Springer, 2009.
- [22] Y. Hristova, P. Kuchment, and L. Nguyen. Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media. *In*verse Probl., 24(5):055006 (25pp), 2008.
- [23] C. Huang, K. Wang, L. Nie, L. V. Wang, and M. A. Anastasio. Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media. *IEEE Trans. Med. Imaging*, 32(6):1097–1110, 2013.
- [24] P. Kuchment and L. Kunyansky. Mathematics of photoacoustic and thermoacoustic tomography. In *Handbook of Mathematical Methods in Imaging*, pages 817–865. Springer, 2011.
- [25] P. Kuchment and L. A. Kunyansky. Mathematics of thermoacoustic and photoacoustic tomography. *Eur. J. Appl. Math.*, 19:191–224, 2008.
- [26] L. A. Kunyansky. Explicit inversion formulae for the spherical mean Radon transform. *Inverse Probl.*, 23(1):373–383, 2007.
- [27] L. A. Kunyansky. A series solution and a fast algorithm for the inversion of the spherical mean Radon transform. *Inverse Probl.*, 23(6):S11–S20, 2007.
- [28] L. A. Kunyansky. Reconstruction of a function from its spherical (circular) means with the centers lying on the surface of certain polygons and polyhedra. *Inverse Probl.*, 27(2):025012, 2011.
- [29] C. Li and L. V. Wang. Photoacoustic tomography and sensing in biomedicine. *Phys. Med. Biol.*, 54(19):R59, 2009.

- [30] F. Natterer. Photo-acoustic inversion in convex domains. Inverse Probl. Imaging, 6(2):315–320, 2012.
- [31] L. V. Nguyen. On a reconstruction formula for spherical Radon transform: A microlocal analytic point of view. Anal. Math. Phys., 4(3):199–220, 2014.
- [32] L. V. Nguyen. On artifacts in limited data spherical Radon transform: flat observation surfaces. SIAM J. Math. Anal., 47(4):2984–3004, 2015.
- [33] 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.
- [34] G. Paltauf, R. Nuster, M. Haltmeier, and P. Burgholzer. Photoacoustic tomography using a Mach-Zehnder interferometer as an acoustic line detector. App. Opt., 46(16):3352–3358, 2007.
- [35] G. Paltauf, J. A. Viator, S. A. Prahl, and S. L. Jacques. Iterative reconstruction algorithm for optoacoustic imaging. J. Acoust. Soc. Am., 112(4):1536–1544, 2002.
- [36] S. K. Patch. Thermoacoustic tomography consistency conditions and the partial scan problem. *Phys. Med. Biol.*, 49:2305–2315, 2004.
- [37] S. K. Patch. Photoacoustic and thermoacoustic tomography: Consistency conditions and the partial scan problem. In *Photoacoustic Imaging and Spectroscopy*, pages 103– 116. CRC Press, 2009.
- [38] A. Rosenthal, V. Ntziachristos, and D. Razansky. Acoustic inversion in optoacoustic tomography: A review. Curr. Med. Imaging Rev., 9(4):318–336, 2013.
- [39] J. Schwab, S. Pereverzyev Jr., and M. Haltmeier. A Galerkin least squares approach for photoacoustic tomography. Technical report, University of Innsbruck, Department of Mathematics, Applied Mathematics Group, 2017. Preprint Nr. 30, arXiv:1612.08094 [math.NA].
- [40] P. Stefanov and G. Uhlmann. Thermoacoustic tomography with variable sound speed. *Inverse Probl.*, 25(7):075011, 2009.
- [41] P. Stefanov and G. Uhlmann. Is a curved flight path in SAR better than a straight one? SIAM J. Appl. Math., 73(4):1596–1612, 2013.
- [42] J. Xia, J. Yao, and L. V. Wang. Photoacoustic tomography: principles and advances. Prog. Electromagnetics Res., 147:1–22, 2014.
- [43] M. Xu and L. V. Wang. Universal back-projection algorithm for photoacoustic computed tomography. *Phys. Rev. E*, 71(1):0167061–0167067, 2005.
- [44] M. Xu and L. V. Wang. Photoacoustic imaging in biomedicine. Rev. Sci. Instruments, 77(4):041101 (22pp), 2006.
- [45] Y. Xu, L. V. Wang, G. Ambartsoumian, and P. Kuchment. Reconstructions in limitedview thermoacoustic tomography. *Med. Phys.*, 31(4):724–733, 2004.

- [46] Y. Xu, M. Xu, and L. V. Wang. Exact frequency-domain reconstruction for thermoacoustic tomography–II: Cylindrical geometry. *IEEE Trans. Med. Imag.*, 21:829–833, 2002.
- [47] L. Yao and H. Jiang. Photoacoustic image reconstruction from few-detector and limited-angle data. *Biomed. Opt. Express*, 2(9):2649–2654, 2011.
- [48] G. Zangerl, O. Scherzer, and M. Haltmeier. Exact series reconstruction in photoacoustic tomography with circular integrating detectors. *Commun. Math. Sci.*, 7(3):665– 678, 2009.