

Nr. 72 29. July 2020

Preprint-Series: Department of Mathematics - Applied Mathematics

Multi-Scale Factorization of the Wave Equation with Application to Compressed Sensing Photoacoustic Tomography

G. Zangerl, M. Haltmeier



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

1 Multi-Scale Factorization of the Wave Equation with Application to Compressed 2 Sensing Photoacoustic Tomography*

3

4

Gerhard Zangerl[†] and Markus Haltmeier[†]

5Abstract. By performing a large number of spatial measurements, high spatial resolution photoacoustic imaging 6can be achieved without specific prior information. However, the acquisition of spatial measurements 7 is time consuming, costly and technically challenging. By exploiting non-linear prior information, 8 compressed sensing techniques in combination with sophisticate reconstruction algorithms allow a 9 reduction of number of measurements while maintaining a high spatial resolution. For this purpose, 10 in this paper, we propose a multiscale factorization for the wave equation, which separates the data 11 into a low frequency factor and sparse high frequency factors. By extending the acoustic reciprocal 12principle, we transfer sparsity in measurements domain to spatial sparsity of the initial pressure, 13 which allows the use of sparse reconstruction techniques. Numerical results are presented which 14 demonstrate the feasibility of the proposed framework.

Key words. photoacoustic tomography, image reconstruction, limited data, wave equation, cost reduction,
 compressed sensing, multiscale factorization.

17 AMS subject classifications. 45Q05, 65T60, 94A08, 92C55

1. Introduction. Photoacoustic tomography (PAT) is an emerging imaging technique that 18 19 combines the high resolution of ultrasound imaging with the high contrast of optical tomography [45]. As illustrated in Figure 1, in PAT a semitransparent sample is illuminated 20 by short pulses of optical energy which causes the excitation of an acoustic pressure wave 21 $p: \mathbb{R}^3 \times [0,\infty) \to \mathbb{R}$, depending the spatial position $x \in \mathbb{R}^3$ and time t > 0. The initial 22 pressure distribution $f: \mathbb{R}^3 \to \mathbb{R}$ is proportional to the inner light absorption properties of the sample and caries valuable diagnostic information. Detectors located on a measurement 24 25surface S (partially) surrounding the sample measure the acoustic pressure from which the initial pressure distribution is recovered. Throughout we denote by $\mathbf{W} f := p|_{S \times [0,\infty)}$ the 2627restriction of the acoustic pressure to the measurement surface.

Exact reconstruction formulas are available for complete data for certain surfaces S, see, for example, [12, 14, 16, 21, 32, 33, 37]. Efficient reconstruction schemes in PAT accounting for acoustic attenuation or variable sound speed have also been developed [1, 2, 3, 24, 27, 29, 30, 44]. Different detector types such as linear or circular detectors that record integrals of the acoustic pressure have been investigated [5, 46]. In this paper, we consider the constant sound speed case and address the issue of compressed sensing to reduce the amount of data while maintaining high spatial resolution.

1.1. Compressed sensing PAT. Acoustic signals in PAT offer large bandwidth. Therefore, high spatial resolution can be achieved by using a sufficiently large number of measurements [22] as predicted by Shannon's sampling theorem. In practice, however, collecting a high

^{*}Submitted to the editors DATE.

Funding: The work of the authors was supported by Austrian Science Fund (FWF) project P 30747-N32.

[†]Department of Mathematics, University of Innsbruck, Technikestraße 13, 6020 Innsbruck, Austria (gerhard.zangerl@uibk.ac.at, markus.haltmeier@uibk.ac.at).



Figure 1. Basic principles of PAT. Left: A sample object is illuminated by short optical pulses. Middle: Optical energy is absorbed within the sample, causes non-uniform heating and induces a subsequent acoustic pressure wave. Right: Acoustic sensors located outside of the sample capture the pressure signals, which are used to recover an image of the interior. In this paper, we develop a specific compresses sensing algorithm that allows to reduce the number of spatial measurements while keeping high spatial resolution.

number of measurements either needs a high number of parallel data acquisition channels or 38 many sequential measurements. This either results in an increase of the cost and technical 39 complexity of the system, or significantly increases measurement time. Several approaches ac-40 celerating data collection have been proposed. For example, in [38] a phase contrast method 41 42 has been developed where a reconstruction of the initial pressure can be obtained from projections of the acoustic pressure that can be collected in short time. In the present work, 43 we use compressed sensing techniques [6, 11, 17, 18] to reduce the number of measurements 44 while preserving high spatial resolution. Main challenges in compressed sensing are the de-45velopments of sophisticated image formation algorithms. In PAT such CS techniques have 46 been developed [23, 25, 4]. Here, we develop a novel image reconstruction strategy based on 47 a multiscale factorization of the wave equation universally applicable to compressed sensing 48 PAT (CSPAT). 49 Compressed sensing reconstruction techniques are based on the sparsity of the signals to be 50reconstructed. In PAT, a possible approach is to express the initial density in a suitable basis. 5152 Using such a strategy leads to a coupled forward model that might numerically challenging to

be solved. As an alternative, strategies that apply a transform in the time domain to sparsify PAT data have been proposed in [42, 23]. These methods have been demonstrated to very well reconstruct the high frequency content of the initial pressure from a significantly reduced amount of measurements. However, the proposed differential operators used as sparsifying transforms suppress low frequency information of the acoustic signals which results in lowfrequency artefacts in the reconstruction.

1.2. Main Contributions. In this work, we develop the concept multiscale temporal transforms and multiscale factorization for CSPAT. We apply multiscale transforms in the time domain that the separate the data into a low frequency component and several high frequency

components. The main idea of the proposed reconstruction scheme is to use the acoustic 62 reciprocal principle to show that there is a one-to-one correspondence between the transform 63 data in the time domain and spatially transformed initial pressure. This factorization allows 64 the use of sparse recovery techniques for the high frequency factor of the initial pressure while 65 the low frequency part can be inverted with standard methods. To be more specific, for a 66 mother wavelet function $v: \mathbb{R} \to \mathbb{R}$ we set $v_i(t) = 2^j \phi(2^j t)$ for j > 1 and denote by v_0 a 67 function containing missing low frequency content. For example, this might be the associated 68 scaling function of mother wavelet. We then explicitly derive associated functions $u_i \colon \mathbb{R}^2 \to \mathbb{R}$ 69 such that for any initial pressure we have the reciprocal relation

71 (1.1)
$$\mathbf{W}(f \circledast_x u_j) = v_j \circledast_t (\mathbf{W} f) \quad \text{for all } j \in \mathbb{N}.$$

The latter identity is then used in the context of compressed sensing data instead of classical 72

73 data $\mathbf{W} f$. Based on (1.1), we develop a reconstruction strategy for recovering a multiscale decomposition of the initial pressure consisting of several sparse high frequency parts and on 74

smooth version of the initial pressure distribution f. Figure (2) shows a phantom f pressure

75

76 data $\mathbf{W} f$ and the corresponding multiscale factors $v_i \circledast_t (\mathbf{W} f)$ (top) and $f \circledast_x u_i$ (bottom).



Figure 2. Top: Data $\mathbf{W} f$ (left) and convolved pressure data $v_j \circledast_t \mathbf{W} f$ at three different scales. Bottom: Initial pressure f and convolved initial pressure $f \circledast_x u_j$ as the same scales. According to the acoustic reciprocal principle (1.1), the convolved pressure data are corresponding to the convolved initial data for which we have sparsity, that we employ for compresses sensing reconstruction.

The concept of sparsifying temporal transforms for CSPAT has been originally developed for PAT in [42, 23] for two and three spatial dimensions. These earlier approaches use a 78 single scale sparsifying transform. The sparsifying transforms filter out low frequency com-79 ponents which results in low frequency artefacts in the reconstruction. By also considering 80 an additional low frequency component, the proposed multiscale scheme naturally overcomes 81 this drawback of single multiscale transforms. Another solution to the missing low frequency 82

content has been proposed in [25] where we proposed to jointly recover the initial pressure fand a sparsified version of f based on second time derivative. However, the present approach seems more natural from and more accessible to rigorous mathematical analysis.

1.3. Outline. The remainder of the paper is organized as follows. In Section 2 we present required background from PAT and derive a dual form acoustic reciprocal principle. In Section 3 we derive a multiscale factorization for the wave equation. The application to CSPAT is presented Section 4. Numerical illustrations are provided in Section 5. The paper end with a short summary and outlook presented in Section 6.

91 **2. Photoacoustic tomography.** In this section, we provide the required background from 92 PAT and derive a reciprocal version of the acoustic reciprocal principle that is required for 93 our later analysis.

94 **2.1. Wave equation model.** Throughout we consider PAT using constant speed of sound 95 propagation. The acoustic pressure is modeled as a function $p: \mathbb{R}^d \times [0, \infty) \to \mathbb{R}$ that satisfies 96 the following initial value problem for the wave equation

97 (2.1) $p_{tt}(x,t) - \Delta p(x,t) = 0 \qquad \text{for } (x,t) \in \mathbb{R}^d \times (0,\infty) ,$

98 (2.2)
$$p(x,0) = f(x) \quad \text{for } x \in \mathbb{R}^d,$$

$$p_t(x,0) = 0 \qquad \text{for } x \in \mathbb{R}^d.$$

101 Here $f \in C^{\infty}(\mathbb{R}^3)$ is the initial pressure distribution that for simplicity we assume to be 102 smooth. Note that in the practical application the cases d = 2 and d = 3 are relevant 103 [5, 15, 16, 31, 45].

104 Continuous PAT data consist of time resolved acoustic pressure restricted in space to a 105 smooth detection surface $S \subseteq \mathbb{R}^3$. The continuous PAT forward operator is given by,

106 (2.4)
$$\mathbf{W} \colon C^{\infty}(\mathbb{R}^3) \to C^{\infty}(S \times (0,\infty)) \colon f \mapsto \mathbf{W} f = p \big|_{S \times (0,\infty)}.$$

107 The corresponding full data inverse problem consists in solving the operator equation $\mathbf{W} f = g$ 108 from possible noisy information. In the last two decades, many methods including exact 109 reconstruction formulas and iterative methods, for different geometries, variable and constant 110 sound speed and detector types have been developed by researchers, see [40] for a recent 111 review. Clearly, in practice only discrete data can be collected. We will present standard 112 discrete sensing as well as compressed sensing strategies in Section 4.

2.2. Acoustic reciprocal principle. Compressed sensing reconstruction techniques are typically based on sparsity of the unknown signals to be reconstructed. To obtain sparsity in PAT we use the acoustic reciprocal principle in combination with sparsifying temporal transforms. The acoustic reciprocal principle states that manipulating photoacoustic data in the temporal domain corresponds to a spatial convolution of the initial pressure with a radial function. An explicit form of acoustic reciprocal principle was first proven in [26] for three spatial dimensions and extended in [20] for arbitrary dimension.

Let $u \in L^1(\mathbb{R}^d)$ be a compactly Proposition 2.1 (Acoustic reciprocal principle of [20]). 120 supported radial function of the form $u = \bar{u} (\|\cdot\|)$ and set 121

122 (2.5)
$$\mathbf{R} \, u \colon \mathbb{R} \to \mathbb{R} \colon t \mapsto \begin{cases} \bar{u}(t) & \text{if } d = 1\\ \omega_{d-2} \int_{|t|}^{\infty} \bar{u}(s) \left(s^2 - t^2\right)^{(d-3)/2} s \mathrm{d}s & \text{if } d > 1, \end{cases}$$

where ω_{d-2} denote the volume of the (d-2)-dimensional unit sphere \mathbb{S}^{d-2} . Then, for every 123124 $f \in C^{\infty}(\mathbb{R}^d)$ we have

$$\begin{array}{l} 125\\ 126 \end{array} \quad (2.6) \qquad \qquad \forall (x,t) \in \mathbb{R}^d \times (0,\infty) \colon \quad \mathbf{W} \left(u \circledast_x f \right) (x,t) = \left((\mathbf{R} \, u) \circledast_t \mathbf{W} f \right) (x,t) \,. \end{array}$$

Here \circledast_x denotes the spatial convolution in \mathbb{R}^d and \circledast_t denotes the one-dimensional con-127volution applied in the second component. 128

129This Lemma serves as the basis for the derived multiscale factorization for the wave equation and the resulting sparse reconstruction strategy. Actually, we use the following dual 130131 version where we prescribe the temporal filter v instead of prescribing the spatial filter u.

Proposition 2.2 (Acoustic reciprocal principle, dual version). Let $v: \mathbb{R} \to \mathbb{R}$ be an even 132function with sufficient decay such that $v \circ \sqrt{|\cdot|} \in C^{\lceil (d-1)/2 \rceil}(\mathbb{R})$ and define 133

134 (2.7)
$$\mathbf{R}^{\sharp} v \colon \mathbb{R}^{d} \to \mathbb{R} \colon x \mapsto \begin{cases} \frac{(-1)^{(d-1)/2}}{\sqrt{\pi^{d-1}}} \left(\left(\frac{1}{2t} \frac{\partial}{\partial t}\right)^{(d-1)/2} v \right) (\|x\|) & \text{for } d \text{ odd} \\ \frac{2(-1)^{(d-2)/2}}{\sqrt{\pi^{d}}} \int_{\|x\|}^{\infty} \frac{\left(\frac{1}{2t} \frac{\partial}{\partial t}\right)^{d/2} v(t)}{\sqrt{t^{2} - \|x\|^{2}}} t \mathrm{d}t & \text{for } d \text{ even} \end{cases}$$

Then, for every $f \in C^{\infty}(\mathbb{R}^d)$, 136

$$\begin{array}{ll} 137\\ 138 \end{array} (2.8) \qquad \forall (x,t) \in \mathbb{R}^d \times (0,\infty) \colon v \circledast_t (\mathbf{W} f) (x,t) = \mathbf{W} \left((\mathbf{R}^{\sharp} v) \circledast_x f \right) (x,t) \,. \end{array}$$

Proof. The proof is given in Appendix A. 139

Note that the assumption $v \circ \sqrt{|\cdot|} \in C^{\lceil (d-1)/2 \rceil}(\mathbb{R})$ is made such that the derivatives in 140 (2.7) are well-defined in the classical sense. 141

3. Multiscale factorizations of the wave equation. Based on the acoustic reciprocal prin-142ciple, in this section we derive convolution factorization for PAT. For that purpose, we fist 143 recall some results for convolutional frames. Then we introduce convolutional frame decompo-144sitions in Subsection 3.2, which are used to derive multiscale factorizations in Subsection 3.3. 145

3.1. Convolutional frames. Let Λ be an at most countable index set and consider a 146 family $(u_{\lambda})_{\lambda \in \Lambda}$ of functions in $L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$. According to the convolution theorem we 147have $f \circledast_x u_{\lambda} = \mathbf{F}_d^{-1}((\mathbf{F}_d f) \cdot (\mathbf{F}_d u_{\lambda}))$ for all $f \in L^2(\mathbb{R}^d)$. Moreover, $f \circledast_x u_{\lambda}$ is well defined almost everywhere and satisfies $f \circledast_x u_{\lambda} \in L^2(\mathbb{R}^d)$. Here and in the following we denote by $\mathbf{F}_d f(\xi) := \int_{\mathbb{R}^d} f(x) e^{-ix \cdot \xi} dx$ for $\xi \in \mathbb{R}^d$ the *d*-dimensional Fourier transform and \mathbf{F}_d^{-1} its 148149150inverse. We write $u^*(x) := u(-x)$ for $u \in L^2(\mathbb{R}^d)$ and note that $\mathbf{F}_d u^* = \overline{[\mathbf{F}_d u]}$, where $\overline{[\cdot]}$ 151denotes complex conjugation. 152

This manuscript is for review purposes only.

5

Definition 3.1 (Convolutional frame). We call a family $\boldsymbol{u} = (u_{\lambda})_{\lambda \in \Lambda} \subseteq (L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d))^{\Lambda}$ 153a convolutional frame in \mathbb{R}^d , if there are constants A, B > 0 such that 154

155 (3.1)
$$f \in L^2(\mathbb{R}^d): \quad A \|f\|_2^2 \le \sum_{\lambda \in \lambda} \|f \circledast_x u_\lambda\|_2^2 \le B \|f\|_2^2.$$

If u is a convolutional frame, we name A, B the frame bounds and call 156

1. $\mathbf{T}_{\boldsymbol{u}} \colon L^2(\mathbb{R}^d) \to \ell^2(\Lambda, L^2(\mathbb{R}^d)) \colon f \mapsto (u_\lambda \circledast_x f)_{\lambda \in \Lambda}$ analysis operator, 157

2. $\mathbf{T}_{\boldsymbol{u}}^* \colon \ell^2(\Lambda, L^2(\mathbb{R}^d)) \to L^2(\mathbb{R}^d) \colon (f_\lambda)_{\lambda \in \Lambda} \mapsto \sum_{\lambda \in \Lambda} u_\lambda^* \circledast_x f_\lambda$ synthesis operator, 3. $\mathbf{T}_{\boldsymbol{u}}^* \mathbf{T}_{\boldsymbol{u}} \colon L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d) \colon f \mapsto \sum_{\lambda \in \Lambda} u_\lambda^* \circledast_x u_\lambda \circledast_x f$ frame operator. 158

159

Finally, we call \boldsymbol{u} tight if $\mathbf{T}_{\boldsymbol{u}}^* \mathbf{T}_{\boldsymbol{u}} = \mathbf{I}$. 160

Note that $\ell^2(\Lambda, L^2(\mathbb{R}^d))$ is a Hilbert space with inner product $\langle \boldsymbol{a}, \boldsymbol{b} \rangle_{\Lambda} := \sum_{\lambda \in \Lambda} \langle a_{\lambda}, b_{\lambda} \rangle$ 161and corresponding norm $\|\cdot\|_{\Lambda}$. Using the analysis operator, we can write the defining iden-tity (3.1) in the form $A \|f\|_2^2 \leq \|\mathbf{T}_u f\|_{\Lambda}^2 \leq B \|f\|_2^2$. Hence the right inequality states that $\mathbf{T}_u: L^2(\mathbb{R}^d) \to \ell^2(\Lambda, L^2(\mathbb{R}^d))$ is well defined an bounded, whereas the left inequality states 162163164that $\mathbf{T}_{\boldsymbol{u}}$ has a bounded Moore-Penrose inverse $\mathbf{T}_{\boldsymbol{u}}^{+} \colon \ell^{2}(\Lambda, L^{2}(\mathbb{R}^{d})) \to L^{2}(\mathbb{R}^{d})$. Further note 165that \mathbf{T}_{u}^{*} is the adjoint of \mathbf{T}_{u} . 166

Lemma 3.2 (Characterization of convolutional frames). For any family $u = (u_{\lambda})_{\lambda \in \Lambda}$ of 167 functions in $L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, the following statements are equivalent: 168

(i) \boldsymbol{u} is a convolutional frame with frame bounds A, B. 169

(ii) The identity $A \leq \sum_{\lambda \in \Lambda} |\mathbf{F}_d u_\lambda|^2 \leq B$ holds almost everywhere. 170

Proof. The convolution theorem and the isometry property of Fourier transform imply 171 that (3.1) holds if and only if for all $f \in L^2(\mathbb{R}^d)$ we have 172

173
$$A \int_{\mathbb{R}^d} |\mathbf{F}_d f(\xi)|^2 \, \mathrm{d}\xi \le \int_{\mathbb{R}^d} |\mathbf{F}_d f(\xi)|^2 \sum_{\lambda \in \lambda} |\mathbf{F}_d u_\lambda(\xi)|^2 \, \mathrm{d}\xi \le B \int_{\mathbb{R}^d} |\mathbf{F}_d f(\xi)|^2 \, \mathrm{d}\xi.$$

This is in turn equivalent that Item (ii) holds. 174

In particular, \boldsymbol{u} is tight if and only if $\sum_{\lambda \in \Lambda} |\mathbf{F}_d u_\lambda|^2 = 1$ almost everywhere. 175

Lemma 3.3 (and definition of a dual convolutional frame). Let $u = (u_{\lambda})_{\lambda \in \Lambda}$ and w =176 $(w_{\lambda})_{\lambda \in \Lambda}$ be two convolutional frames in \mathbb{R}^{d} . Then the following statements are equivalent: 177

(i) The identity $\sum_{\lambda \in \Lambda} \overline{[\mathbf{F}_d w_\lambda]} \cdot (\mathbf{F}_d u_\lambda) = 1$ holds almost everywhere. 178

(ii) The reproducing formula $\forall f \in L^2(\mathbb{R}^d)$: $\mathbf{T}_{w}^* \mathbf{T}_{u} f = \sum_{\lambda \in \Lambda} w_{\lambda}^* \circledast_x (u_{\lambda} \circledast_x f) = f$ holds. 179If (i), (ii) hold, we call $(w_{\lambda})_{\lambda \in \Lambda}$ a dual convolutional frame of $(u_{\lambda})_{\lambda \in \Lambda}$. 180

Proof. The linearity and continuity of the Fourier transform together with the convolu-181 tion theorem show that the identity in (ii) is equivalent to the identity $(\mathbf{F}_d f) \cdot \sum_{\lambda \in \Lambda} (\mathbf{F}_d w_{\lambda}^*) \cdot$ 182 $(\mathbf{F}_d u_\lambda) = \mathbf{F}_d f$ for all $f \in L^2(\mathbb{R}^d)$. Because $\mathbf{F}_d w_\lambda^* = \overline{[\mathbf{F}_d w_\lambda]}$ this implies the desired equiva-183184lence.

In particular, Item (i) in Lemma 3.3 is satisfied with \boldsymbol{w} taken as the canonical dual 185convolutional frame $\boldsymbol{u}^{+} := (u_{\lambda}^{+})_{\lambda \in \Lambda}$ defined by 186

187 (3.2)
$$\forall \mu \in \Lambda : \quad \mathbf{F}_d \, u_{\mu}^* := \frac{\mathbf{F}_d \, u_{\mu}}{\sum_{\lambda \in \Lambda} |\mathbf{F}_d \, u_{\lambda}|^2} \, .$$

7

- 188 In this case we have $\mathbf{T}_{u^{+}} = \mathbf{T}_{u}^{+}$, where \mathbf{T}_{u}^{+} denotes the Moore-Penrose inverse of the analysis 189 operator \mathbf{T}_{u} .
- 190 **3.2. Convolution factorization.** The following concept is central for this paper.

191 Definition 3.4 (Convolution factorization of the wave equation). Let Λ be an at most count-192 able index set and consider families $\mathbf{u} = (u_{\lambda})_{\lambda \in \Lambda} \in (L^2(\mathbb{R}^d))^{\Lambda}$ and $\mathbf{v} = (v_{\lambda})_{\lambda \in \Lambda} \in (L^2(\mathbb{R}))^{\lambda}$. 193 We call the pair (\mathbf{u}, \mathbf{v}) a convolution factorization for \mathbf{W} if the following hold:

- 194 (CFD1) $(u_{\lambda})_{\lambda \in \Lambda}$ is a convolutional frame of $L^{2}(\mathbb{R}^{d})$,
- 195 (CFD2) $(v_{\lambda})_{\lambda \in \Lambda}$ is a convolutional frame of $L^{2}(\mathbb{R})$,
- 196 (CFD3) The commutation relation holds $\forall f \in L^2(K)$: $\mathbf{W}(u_\lambda \circledast_x f) = v_\lambda \circledast_t (\mathbf{W} f)$.

Given a convolution factorization $(\boldsymbol{u}, \boldsymbol{v})$ for \mathbf{W} and data $g = \mathbf{W} f$ the commutation relation (CFD3) formula shows that is sufficient to solve each equation $\mathbf{W} f_{\lambda} = v_{\lambda} \circledast_t g$. These equations now involve the unknown $f_{\lambda} = u_{\lambda} \circledast_x f$ having specific prior information, that we can exploited for inversion. Moreover, we will later show that the same identity holds for any spatial sampling scheme, enabling it for CSPAT. It is well known that \mathbf{W} is injective when restricted to function having some decay. Therefore, if $(\boldsymbol{u}, \boldsymbol{v})$ is a convolution factorization and $u_{\lambda} \circledast_x f$ has sufficient decay we have the reproducing formula

204 (3.3)
$$f = \sum_{\lambda \in \Lambda} u_{\lambda}^* \circledast_x \mathbf{W}^{-1}(v_{\lambda} \circledast_t (\mathbf{W} f))$$

In fact, the factorization identity (3.3) is the reason why we name a pair (u, v) satisfying (CFD1)-(CFD3) a convolution factorization.

As the main theoretical result, in this paper we construct explicit convolution factorizations for the PAT forward operator. For that purpose, recall

209
$$\mathbf{R}^{\sharp} v \colon \mathbb{R}^{d} \to \mathbb{R} \colon x \mapsto \begin{cases} \frac{(-1)^{(d-1)/2}}{\sqrt{\pi^{d-1}}} \left(\left(\frac{1}{2t} \frac{\partial}{\partial t}\right)^{(d-1)/2} v \right) (\|x\|) & \text{for } d \text{ odd} \\ \frac{2(-1)^{(d-2)/2}}{\sqrt{\pi^{d}}} \int_{\|x\|}^{\infty} \frac{\left(\frac{1}{2t} \frac{\partial}{\partial t}\right)^{d/2} v(t)}{\sqrt{t^{2}-\|x\|^{2}}} t dt & \text{for } d \text{ even} \,, \end{cases}$$

and the dual version of the acoustical reciprocal principle $v \circledast_t (\mathbf{W} f) = \mathbf{W} ((\mathbf{R}^{\sharp} v) \circledast_x f)$ stated in Proposition 2.2.

Theorem 3.5 (Construction of convolution factorization of the wave equation). Let $(v_{\lambda})_{\lambda \in \Lambda} \in L^2(\mathbb{R})^{\Lambda}$ be a convolutional frame consisting of even functions v_{λ} with sufficient decay such that such that $v_{\lambda} \circ \sqrt{|\cdot|} \in C^{\lceil (d-1)/2 \rceil}(\mathbb{R})$ and let $(v_{\lambda}^{*})_{\lambda \in \Lambda}$ be its canonical dual and set $u_{\lambda} = \mathbb{R}^{\sharp} v_{\lambda}$. 1. The pair $((u_{\lambda})_{\lambda \in \Lambda}, (v_{\lambda})_{\lambda \in \Lambda})$ is a convolutional frame decomposition for \mathbb{W} .

216 2. The canonical dual of $(u_{\lambda})_{\lambda \in \Lambda}$ is given by $(\mathbf{R}^{\sharp} v_{\lambda}^{*})_{\lambda \in \Lambda}$.

217 3. For all $f \in L^2(\mathbb{R}^d)$, the factors $f_{\lambda} = u_{\lambda} \circledast_x f$ satisfy

218 (3.4)
$$f = \sum_{\lambda \in \Lambda} u_{\lambda}^{\star} * f_{\lambda}$$

$$210 (3.5) \mathbf{W} f_{\lambda} = v_{\lambda} \circledast_t \mathbf{W} f_{\lambda}$$

Hence any function $f \in L^2(\mathbb{R}^d)$ can be recovered from data **W** f by first solving equation (3.5) for f_{λ} and the evaluating the series (3.4). *Proof.* To show Item 1 we verify (CFD1)-(CFD3) from Definition 3.4. Item (CFD1) is satisfied because $(v_{\lambda})_{\lambda \in \Lambda} \in L^2(\mathbb{R})^{\Lambda}$ is a convolutional frame according to the made assumptions. Item (CFD3) follows from the acoustic reciprocal principle Proposition 2.1. It remains to verify Item (CFD2), namely that the family $(u_{\lambda})_{\lambda \in \Lambda} \in L^2(\mathbb{R})^{\Lambda}$ is a convolutional frame. For that purpose, recall that $\mathbf{R} \mathbf{R}^{\sharp} v_{\lambda} = v_{\lambda}$ where \mathbf{R} denotes the Radon transform of a radial function. According to the Fourier slice theorem we have $\mathbf{F}_d \mathbf{R}^{\sharp} v_{\lambda} = \mathbf{F}_1 v_{\lambda}$. Therefore $\sum_{\lambda \in \Lambda} |\mathbf{F}_d \mathbf{R}^{\sharp} v_{\lambda}|^2 = \sum_{\lambda \in \Lambda} |\mathbf{F}_1 v_{\lambda}|^2$ which implies that $(\mathbf{R}^{\sharp} v_{\lambda})_{\lambda \in \Lambda}$ is a convolutional frame according to Lemma 3.2. Moreover, we have $\mathbf{F}_1 v_{\lambda}^* = \mathbf{F}_1 v_{\lambda} / \sum_{\lambda \in \Lambda} |\mathbf{F}_1 v_{\lambda}|^2$ and therefore

$$\mathbf{F}_{1} \mathbf{R}^{\sharp} v_{\lambda}^{\star} = \mathbf{F}_{1} v_{\lambda}^{\star} = \frac{\mathbf{F}_{1} v_{\lambda}}{\sum_{\lambda \in \Lambda} |\mathbf{F}_{1} v_{\lambda}|^{2}} = \frac{\mathbf{F}_{1} \mathbf{R}^{\sharp} v_{\lambda}}{\sum_{\lambda \in \Lambda} |\mathbf{F}_{d} \mathbf{R}^{\sharp} v_{\lambda}|^{2}},$$

which shows that $(\mathbf{R}^{\sharp} v_{\lambda}^{\dagger})_{\lambda \in \Lambda}$ is the canonical dual of $(\mathbf{R}^{\sharp} v_{\lambda})_{\lambda \in \Lambda}$ which is Item 2. Finally, Item 3 follows Items 1, 2 and the definitions of a CDF and a dual frame.

3.3. Multiscale factorization. As shown in the previous subsection, a convolution factorization splits the image reconstruction into several reconstruction problems, one for each convolved initial data $u_{\lambda} \circledast_x f$. The basic idea is now to take $(u_{\lambda})_{\lambda \in \Lambda}$ as a multiscale system in order to be able bring sparsity into account.

For given $u \in L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ consider the scaled versions

230 (3.6)
$$u_j \colon \mathbb{R}^d \to \mathbb{R} \colon x \mapsto 2^{jd} u(2^j x) \quad \text{for } j \ge 1.$$

According to the scaling property of the Fourier transform we have $\mathbf{F}_{d} u_{j}(\xi) = \mathbf{F}_{d} u(2^{-j}\xi)$. Suppose that $\mathbf{F}_{d} u$ is essentially supported $\{\xi \in \mathbb{R}^{d} \mid b_{0} \leq ||\xi|| < 2b_{0}\}$. Then, as j increases the Fourier transform $\mathbf{F}_{d} u_{j}$ are essentially supported in $\{\xi \in \mathbb{R}^{d} \mid 2^{j-1}b_{0} < ||\xi|| \leq 2^{j}b_{0}\}$. Its union covers all frequencies except the low frequencies that are contained in in the ball $B_{0} = \{\xi \in \mathbb{R}^{d} \mid ||\xi|| < b_{0}\}$. In order to obtain a convolutional frame with reasonable constants we therefore select another function $u_{0} \in L^{2}(\mathbb{R}^{d})$ such that $\mathbf{F}_{d} u_{0}$ covers frequencies in b_{0} .

237 Definition 3.6 (Multiscale convolution decomposition). Let $u_0, u \in L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ and 238 define u_j for $j \ge 1$ by (3.6). We call the family $(u_j)_{j\in\mathbb{N}}$ a multiscale convolution decomposition 239 in $L^2(\mathbb{R}^d)$ if it forms a convolutional frame. For $f \in L^2(\mathbb{R}^d)$, we refer to $f \circledast_x u_0$ as the low 240 frequency factor and to $f * u_j$ for $j \ge 1$ as the high frequency factor at scale j.

According to Lemma 3.2 and the scaling property of the Fourier transform, the family 242 $\boldsymbol{u} = (u_j)_{j \in \mathbb{N}}$ is a multiscale decomposition if and only if there are constants A, B > 0 such 243 that

244 (3.7)
$$A \leq \sum_{j \in \mathbb{N}} |\mathbf{F}_d u_j(\xi)|^2 \leq B \quad \text{for almost every } \xi \in \mathbb{R}^d.$$

Moreover, the canonical dual frame $u^* = (u_j^*)_{j \in \mathbb{N}}$ of u is given by the Fourier representation $\mathbf{F}_d u_j^* := \mathbf{F}_d u_j / \sum_{k \in \mathbb{N}} |\mathbf{F}_d u_k|^2$ for $j \in \mathbb{N}$. In the one-dimensional case we write $(v_j)_{j \in \mathbb{N}}$ for a multiscale decomposition.

Definition 3.7 (Multiscale factorization of the wave equation). We call a pair $(\boldsymbol{u}, \boldsymbol{v})$ a multiscale factorization of \mathbf{W} if it is a convolutional frame decomposition such that $\boldsymbol{u} = (u_j)_{j \in \mathbb{N}}$ and $\boldsymbol{v} = (v_j)_{j \in \mathbb{N}}$ are multiscale decompositions in $L^2(\mathbb{R}^d)$ and $L^2(\mathbb{R})$, respectively.

From Theorem 3.5 we immediately get the following.

Theorem 3.8 (Construction of multiscale factorization of the wave equation). Let $v = (v_j)_{j \in \mathbb{N}}$ be a multiscale decomposition in $L^2(\mathbb{R})$ consisting of even functions with sufficient decay, let v^+ be its canonical dual and define $u = (\mathbf{R}^{\sharp} v_j)_{j \in \mathbb{N}}$. Then the following holds

1. The pair (u, v) is a multiscale factorization for **W**.

256 2. The canonical dual $(u_i^{\dagger})_{i \in \mathbb{N}}$ of **u** has the Fourier representation

257 (3.8)
$$\mathbf{F}_{d} u_{j}^{\star}(\xi) = \frac{1}{\sum_{j \in \mathbb{N}} |\mathbf{F}_{1} v_{j}(\|\xi\|)|^{2}} \mathbf{F}_{1} v_{j}(\|\xi\|)$$

259 3. For all $f \in L^2(\mathbb{R}^d)$ the factors $f_j := u_j \circledast_x f$ satisfy

260 (3.9)
$$f = \sum_{j \in \mathbb{N}} u_j^* \circledast_x f_j$$

$$\frac{261}{262} \qquad (3.10) \qquad \qquad \mathbf{W} f_j = v_j \circledast_t \mathbf{W} f_j$$

Hence any $f \in L^2(\mathbb{R}^d)$ can be recovered from $\mathbf{W} f$ by first solving (3.10) and then evaluating the series (3.9).

265 *Proof.* Follows from Theorem 3.5.

Alternatively, we have the following result that avoids computing the canonical dual.

267 Corollary 3.9 (Multiscale reconstruction for the wave equation). Let $(v_j)_{j\in\mathbb{N}} \in (L^2(\mathbb{R}))^{\mathbb{N}}$ 268 be a multiscale decomposition consisting of even functions v_j with sufficient decay such that 269 $v_j \circ \sqrt{|\cdot|} \in C^{\lceil (d-1)/2 \rceil}(\mathbb{R})$ and set $u_j = \mathbb{R}^{\sharp} v_j$. Then, all $f \in L^2(\mathbb{R}^d)$ the factors $f_j = u_j \circledast_x f$ 270 satisfy

271 (3.11)
$$\Phi \circledast_x f = \sum_{j \in \mathbb{N}} u_j \circledast_x f_j$$

$$\mathbf{W} f_j = v_j \circledast_t \mathbf{W} f \,,$$

274 with $\Phi := \mathbf{F}_d^{-1} \left(\sum_{j \in \mathbb{N}} |\mathbf{F}_d u_j|^2 \right).$

275 *Proof.* Similar to the proof of Theorem 3.5.

From Corollary 3.9 it follows that any function $f \in L^2(\mathbb{R}^d)$ can be recovered from data W f by means of the following consecutive steps

- Solve equation (3.12) for f_j
- Evaluate the series on the right hand side of (3.11)
- Solve the deconvolution problem (3.11) for f.

Because the $(v_{\lambda})_{\lambda \in \Lambda}$ is a convolutional frame, $\mathbf{F}_d \Phi$ is bounded away from zero and therefore the deconvolution problem (3.11) is stably solvable. Because inversion of the wave equation in the full data case is stable, also (3.12) can be stably solved for f_j . In the case of compressed sensing this will not be the case and we will have to include additional prior information for its solution.

Remark 3.10 (Examples for multiscale decompositions). Possible multiscale decompositions 286 can be constructed via a dvadic translation invariant wavelet frame, which names a convo-287lutional frame $(\psi_j)_{j\in\mathbb{Z}}$ where $\psi_j := 2^j \psi(2^j(\cdot))$ for a so-called mother wavelet $\psi \colon \mathbb{R} \to \mathbb{R}$. 288If we define $v_j := \psi_j$ for $j \ge 1$ and select v_0 such that $|\mathbf{F}_1 v_0|^2 = \sum_{j \le 0} |\mathbf{F}_1 \psi_j|^2$ we obtain a multiscale decomposition $(u_j)_{j \in \mathbb{N}}$. Alternatively, for the low resolution filter we can take any other function v_0 such that $|\mathbf{F}_1 v_0|^2 + \sum_{j \ge 1} |\mathbf{F}_d \psi_j|^2$ is bounded and away from zero by 289290 291 reasonable constants. Several examples of dyadic translation invariant wavelet frames can be 292extracted from classical wavelet analysis [8, 34]. Among many others, examples for the gen-293erating mother wavelet $\psi = v_0$ are Mexican hat wavelets, Shannon wavelets, spline wavelets 294or Meyer wavelets. For our numerical simulations, we use the Mexican hat wavelet. 295

4. Application to compressed sensing PAT. In this section, we extend the multiscale factorization to the case of compressed sensing data in PAT. Moreover, we derive a corresponding sparse recovery scheme.

4.1. Sampling the wave equation. When realizing a PAT setup, the acoustic data can only be collected for a finite number of sampling points that we denote by $z_{\ell} \in S$ for $\ell \in$ $\{1, \ldots, n\}$. Note that we do not discretize the temporal variable since temporal samples can easily be collected at rate well above the Nyquist sampling rate.

Definition 4.1 (Sampled PAT forward operator). Let **W** be the continuous PAT forward operator defined in (2.4). For sampling points $z_{\ell} \in S$, $\ell \in \{1, ..., n\}$, we set

305 (4.1)
$$\mathbf{S}_n \colon C^{\infty}(S \times (0, \infty)) \to (C^{\infty}(0, \infty))^n \colon g \mapsto (g(z_{\ell}, \cdot))_{\ell=1, \dots, n}$$

 $\underset{336}{366} \quad (4.2) \qquad \qquad \mathbf{W}_n \colon C^{\infty}(\mathbb{R}^d) \to (C^{\infty}(0,\infty))^n \colon f \mapsto \mathbf{S}_n \, \mathbf{W} \, f = (p(z_\ell, \,\cdot\,))_{\ell=1,\dots,n} \, .$

We call \mathbf{S}_n the regular sampling scheme and $\mathbf{W}_n = \mathbf{S}_n \mathbf{W}$ the (regularly) sampled PAT forward operator corresponding to the n-tuple $(z_\ell)_{\ell=1,...,n}$ of spatial sampling points.

The fundamental question of classical Shannon sampling theory in the context of PAT is 310 finding "simple" sets X where the initial pressure belongs to and associated conditions on the 311 sampling points under which the sampled data $\mathbf{W}_n f$ uniquely determine the initial pressure 312distribution $f \in \mathbb{X}$. For equispaced detectors on the boundary of a disc $D \subseteq \mathbb{R}^2$, explicit 313 sampling conditions for PAT have been derived in [22]. Roughly spoken, these results state 314 that any function $f \in C_0^{\infty}(D)$ whose Fourier transform $f(\xi)$ is sufficiently small for $\|\xi\| \ge b_0$, 315where b_0 is the bandwidth, can be stably recovered from sampled PAT data $\mathbf{W}_n f$ provided 316that the sampling condition $n \ge 2R_0b_0$ is satisfied (for a precise statement see [22]). Sampling 317theory for other tomographic inverse problems is discussed, for example, in [43, 28, 13, 10, 36]. 318

4.2. Compressive sampling. In order to reduce the number of detectors while preserving spatial resolution CSPAT has been studied in several works [19, 41, 4, 4, 19]. The basic idea is to use general linear measurements of the form

322 (4.3)
$$y_j = \langle a_j, \mathbf{W}_n(f) \rangle = \sum_{i=1}^n a_{j,i} (\mathbf{W}_n(f))_i \text{ for } j \in \{1, \dots, m\} .$$

Here a_j are measurement vectors with entries $a_{j,i}$ and $\mathbf{A}_{m,n} := (a_{j,i})_{j,i} \in \mathbb{R}^{m \times n}$ is the measurement matrix. The term compressed sensing refers to the fact that the number of measurements

an PREPRINT

MULTI-SCALE FACTORIZATION FOR CSPAT

m is intended to be chosen much smaller than the number of initial sampling points n. Therefore $y = \mathbf{A}_{m,n} \mathbf{W}_n f$ constitutes a highly under determined system of linear equations and can only be solved with additional assumptions on unknown to be reconstructed.

329 For a systematic treatment, we introduce the following notation.

Befinition 4.2 (Generalized sampling in PAT). For sampling points $(z_{\ell})_{\ell=1,...,n} \in S^n$, measurement matrix $\mathbf{A}_{m,n} \in \mathbb{R}^{m \times n}$ and subspace $\mathbb{X} \subseteq C^{\infty}(\mathbb{R}^d)$ we call

- 332 1. $\mathbf{A}_{m,n} \mathbf{S}_n$ a generalized sampling scheme;
- 333 2. $\mathbf{A}_{m,n} \mathbf{W}_n = \mathbf{A}_{m,n} \mathbf{S}_n \mathbf{W}$ CSPAT forward operator;
- 334 3. $\mathbf{A}_{m,n} \mathbf{S}_n$ a complete sampling scheme for \mathbb{X} , if the restriction $\mathbf{A}_{m,n} \mathbf{S}_n \mathbf{W}|_{\mathbb{X}}$ is injective.

The results of [22] essentially show that \mathbf{S}_n gives a complete sampling scheme on the 335space $V = V_{R_0,b_0}$ of all functions being supported in a disc of radius R_0 and having essential 336 bandwidth b_0 for $n \ge 2R_0b_0$ equally distributed sampling points. This implies that for any 337 invertible matrix $\mathbf{A}_{m,n}$, the composition $\mathbf{A}_{m,n} \mathbf{S}_n$ is a complete sampling scheme on V_{R_0,b_0} as 338 well. We are not aware of any results (even for related inverse problems) when the general-339 ized sampling scheme cannot be written in the form \mathbf{BS}_m with \mathbf{S}_m being a regular sampling 340 scheme and $\mathbf{A}_{m,n} \in \mathbb{R}^{n \times n}$ being invertible. The question for which spaces such a general sam-341 pling matrix are complete sampling schemes seems an interesting direction of open research. 342 343 Anyway, in this paper we investigate the case when $\mathbf{A}_{m,n} \mathbf{W}_n$ is not injective on the linear subspace X and develop a nonlinear reconstruction approach based on ℓ^1 -minimization. 344

4.3. Multiscale reconstruction for CSPAT. Let $\mathbf{W}_n = \mathbf{S}_n \mathbf{W}$ be a regularly sampled discrete PAT forward operator with sampling points $(z_\ell)_{\ell=1,...,n} \in S^n$. We suppose that the regular sampling scheme \mathbf{S}_n is complete for a subspace

348 (4.4)
$$\mathbb{X}_n \subseteq \mathcal{B}_{\Omega}(\mathbb{R}^n) := \left\{ f \in L^2(\mathbb{R}^n) \mid f \text{ is } \Omega \text{-bandlimited} \right\}.$$

Here $f \in L^2(\mathbb{R}^d)$ is called Ω -bandlimited if $\operatorname{supp}(\mathbf{F}_d f)(\xi)$ vanishes for $\|\xi\| > \Omega$. Note that we have $\mathcal{B}_{\Omega}(\mathbb{R}^d) \subseteq C^{\infty}(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. Moreover, let $\mathbf{A}_{m,n} \in \mathbb{R}^{m \times n}$ be a measurement matrix with $m < n \operatorname{such} \mathbf{A}_{m,n} \mathbf{S}_n$ is not complete on \mathbb{X}_n . This means that $\mathbf{A}_{m,n} \mathbf{W}_n$ is not injective on \mathbb{X}_n and therefore cannot be uniquely inverted. Our aim is to nevertheless to recover $f \in \mathbb{X}_n$ from data $y = \mathbf{A}_{m,n} \mathbf{W}_n f$ by using suitable prior information.

Below we describe how a multiscale factorization for the wave equation can be used to recover an initial pressure from CSPAT data. The main ingredient for our approach is that the factorizations of Theorems 3.8 and Corollary 3.9 for the full wave equation generalize to the compressed sensing setup. Here we only formulate such an extension of Corollary 3.9, because we use this version for the numerical implementation.

Proposition 4.3 (Multiscale CSPAT decomposition). Let (u, v) be a multiscale factorization of W. Then for all $f \in L^2(\mathbb{R}^d)$ the factors $f_j = u_j \circledast_x f$ satisfy

361 (4.5)
$$\Phi \circledast_x f = \sum_{j \in \mathbb{N}} (\mathbf{R}^{\sharp} v_j) * f_j$$

364 with $\Phi := \mathbf{F}_d^{-1} \left(\sum_{j \in \mathbb{N}} |\mathbf{F}_d u_j|^2 \right).$

³⁶⁵ *Proof.* This follows from Corollary 3.9 by simply noting that $\mathbf{A}_{m,n} \mathbf{S}_n$ acts in the spatial ³⁶⁶ variable and therefore commutes with the temporal convolution.

4.4. Proposed multiscale reconstruction. Consider the situation as described in Subsection 4.3. Based on the factorization of Corollary 4.3 we propose the following reconstruction scheme form CSPAT data:

Algorithm 4.4 (Recover initial pressure f from CSPAT data $\mathbf{A}_{m,n} \mathbf{W}_n f$).

- 371 (S1) Solve equation (4.6) for f_j as described below.
- (S2) Evaluate the series on the right hand side of (4.5).
- 373 (S3) Solve the deconvolution problem (4.5) for f.

Because the $(v_j)_{j \in \mathbb{N}}$ is a convolutional frame, $\mathbf{F}_d \Phi$ is bounded away from zero and therefore the deconvolution problem (4.5) in step (S3) is stably solvable. However, because $\mathbf{A}_{m,n} \mathbf{W}_n$ is non-injective, solving (4.6) in step (S1) requires prior information. The proposed procedure is described in following in Remark 4.5.

378 Remark 4.5 (Solution of step (S1)). Appropriate prior information is different for the low 379 frequency factor f_0 and the high frequency factors f_j for $j \ge 1$. We therefore recover the low 380 frequency factor and the high frequency factors differently.

• LOW FREQUENCY FACTOR: Assuming that v_0 has essential bandwidth in b_0 , then low frequency factor $f_0 = u_0 \circledast_x f$ also has essential bandwidth b_0 . This means that the Fourier transform $\mathbf{F}_d f_0(\xi)$ is sufficiently small for $\|\xi\| \ge b_0$. Therefore, classical sampling theory in the context of PAT [22] shows that m regular samples are sufficient to recover $f_0 = u_0 \circledast_x f$ as solution of

386 (4.7)
$$\min_{h \in \mathbb{X}_n} \|h\|_2 \quad \text{such that } \mathbf{A}_{m,n} \mathbf{W}_n h = y_0.$$

We use (4.7) when $\mathbf{A}_{m,n} \in \mathbb{R}^{m \times n}$ is a either a subsampling matrix or random sensing matrix.

• HIGH FREQUENCY FACTORS: Assuming that the Fourier transform $\mathbf{F}_1 v(\omega)$ is neg-389 ligible in a suitable sense when $|\omega|$ is outside the interval $[b_0, 2b_0]$, then the Fourier 390transforms are of high factors $\mathbf{F}_d u(\xi)$ are negligible outside the annulus $D_i := \{\xi \in$ 391 $\mathbb{R}^d \mid [2^{j-1}b_0 \leq \|\xi\| \leq 2^k b_0\}$. However if we perform compressive sampling, $\mathbf{A}_{m,n} \mathbf{W}_n$ 392 will not be injective on the space of all functions that are essentially supported in D_i 393 and therefore (4.5) cannot be inverted uniquely without additional prior information. 394As we observe from Figure 2 the high frequency factors $f_0 = u_0 \otimes_x f$ are sparse in the 395 spatial domain. Therefore in this paper we propose to use ℓ^1 -minimizing solutions 396

397 (4.8)
$$\min_{h \in \mathbb{X}_n} \|h\|_1 \quad \text{such that } \mathbf{A}_{m,n} \mathbf{W}_n h = y_j.$$

Here $||h||_1 := \sum_{i \in \mathbb{Z}^2} |h(i\pi/\Omega)|$ is the ℓ^1 -norm of h at the discrete samples $x_i = \pi/\Omega$ advices by Shannon's sampling theorem.

We can derive uniqueness of (4.8) from compressed sensing conditions. Here we make use of the [18] dealing with the reconstruction of individual elements, which in our context read as follows:

• There exists $\eta \in \mathbb{R}^m$ with $(\mathbf{A}_{m,n} \mathbf{W}_n)^* \eta \in \mathrm{Sign}(u_j \circledast_x f)$

404 • $\forall i \in I_n \setminus \operatorname{supp}(u_{0,j} \circledast_x f) \colon |\langle e_n, (\mathbf{A}_{m,n} \mathbf{W}_n)^* \eta \rangle| < 1.$

• The restricted mapping $(\mathbf{A}_{m,n} \mathbf{W}_n)_{\mathrm{supp}(x^{\dagger})}$ is injective on \mathbb{X}_j .

A detailed error analysis for CSPAT with Algorithm 4.4 using (4.7) for the low resolution factor and (4.8) for the high resolution factors is an interesting line of future research and beyond the scope of this paper.

5. Numerical experiments. In this section, we present details on the implementation of the sparse reconstruction scheme from compressed sensing measurements presented in Section 4. In our numerical experiments, we consider the two dimensional case when the initial pressure is supported in the unit disc in \mathbb{R}^2 of radius 0.9 and measurements are taken the unit sphere \mathbb{S}^1 . This situation appears in PAT with integrating line detectors [5, 14].



Figure 3. Left: Initial pressure on a square grid of side length 2. The point-like detectors are equidistantly distributed on the black circle. Right: Full data.

5.1. Numerical implementation. For all presented numerical implementations we replace f by its discrete values $(f(x_i))_i$ on a Cartesian grid of side length 2 at nodes $x_i = i 2/N_x$ for $i \in \{-N_x/2, \dots, N_x/2 - 1\}^2$ with $N_x = 100$. We assume that f is sampled at the Nyquist rate 417 such that the maximal Bandwidth is given by $\Omega = N_x(\pi/2)$. The space X_n in (4.4) is taken 418 as the space of all $f \in \mathcal{B}_{\Omega}$ whose samples vanish outside the disc of radius 0.9. We implement \mathbf{W}_n and $\mathbf{A}_{m,n}$ with n = 300 and m = 75 as described below. Note that the fully sampled PAT forward operator \mathbf{W}_n satisfies the classical sampling conditions of [22]. The measurements matrix $\mathbf{A}_{m,n}$ corresponds to a subsampling factor of 4.

• SAMPLED PAT FORWARD OPERATOR: The discretization of \mathbf{W}_n is based on the 422 Fourier representation $\mathbf{F}_d p(\xi, t) = \cos(||\xi|| t) \mathbf{F}_d f(\xi)$ for the solution of the wave equa-423tion (2.1)-(2.3). For the numerically computation we replace the Fourier transform 424 by the discrete Fourier transform on the square grid of side length 4 with spatial 425nodes $x_i = i 2/N_x$ for $i \in \{-N_x, \ldots, N_x - 1\}^2$ and frequency nodes $\xi_k = k \Omega/N_x$ for 426 $k \in \{-N_x, \ldots, N_x - 1\}^2$. Here, the bandwidth Ω and the spatial sampling step size 427 $2/N_x$ are satisfy the Nyquist condition $2/N_x = \pi/\Omega$ and the larger numerical domain 428 is $[-2,2] \times [-2,2]$ is chosen to avoid boundary effects. We then define discrete fully 429sampled PAT forward operator \mathbf{W}_n by nearest neighbor interpolation at the detec-430tor locations. The adjoint of \mathbf{W}_n^* is numerically computed via the backprojection 431algorithm described in [5]. 432

• MULTISCALE FILTERS: The high frequency filters v_j for ≥ 1 are taken as Mexican hat wavelets

435
$$v_j(t) := 2^3 2^j (1 - (2^j 8t)^2) \exp\left(-\frac{(2^j 8t)^2}{2}\right)$$

and the corresponding low resolution temporal filter v_0 is taken as a Gaussian $v_0(t) :=$ 436 $2^3 \exp(-(8t)^2/2)$. The width of the filter v_0 is taken such that $v_0 \circledast_x f$ is can be 437 recovered from n = 75 samples according to classical sampling theory. For the high 438frequency components $v_j \circledast_x f$ with $j \ge 1$ this is not the case, and therefore we use 439sparsity as described in Section (4). The spatial filters $u_j = \mathbf{R}^{\sharp} v_j$ are computed 440 analytically by evaluating (2.7) for d = 2 and $v = v_j$. For $j \ge 2$, the essential support 441 v_i lies outside the considered frequency regime [-b, b] and therefore for the numerical 442 simulations we use the three filters (v_0, v_1, v_2) . All temporal and spatial convolutions 443 are replaced by discrete convolution via the discrete Fourier transform. 444

• MEASUREMENT MATRIX: For the measurement matrix $\mathbf{A}_{m,n} \in \mathbb{R}^{m \times n}$ we consider two choices. First, we take $\mathbf{A}_{m,n}$ as uniform subsampling matrix which has entries $a_{j,i} = 1$ if j = 4(i-1) + 1 and $a_{j,i} = 0$ otherwise. Second we take $\mathbf{A}_{m,n}$ as Gaussian random matrix where each entry $a_{j,i}$ is the realization of an independent Gaussian random variables with zero mean.

The initial pressure used for the numerical simulations, the corresponding fully sampled data as well as the subsampled data and the Gaussian measurement data are shown in Figure 3. The filtered data for the subsampling scheme are shown in the left column of Figure 4 and the filtered data using Gaussian measurements in Figure 5.

5.2. Reconstruction results. Following the strategy proposed in Section 4 (see Algorithm 4.4 and Remark 4.5), we recover the initial phantom via the following three steps:

• First recover the factors $u_j \otimes_x f$ from data $y_j = v_j \otimes_x (\mathbf{A}_{m,n} \mathbf{W}_n f)$. For that purpose, we use the Landweber iteration for recovering the low frequency factor $u_0 \otimes_x f$ and the iterative soft thresholding algorithm

459
$$h_k = \operatorname{soft}_{s\lambda} \left(h_k + s \, \mathbf{W}_n^* \, \mathbf{A}_{m,n}^* \left(\mathbf{T}_j f y - \mathbf{A}_{m,n} \, \mathbf{W}_n \, f_k \right) \right)$$



Figure 4. Left: The convolved data $y_j = v_j \otimes_t (\mathbf{A}_{m,n} \mathbf{W}_n f)$ for j = 0, 1, 2 where $\mathbf{A}_{m,n}$ is the subsampling matrix with subsampling factor 4. Right. Corresponding reconstructions of convolved initial pressure $u_j \circledast_x f$ using the Landweber method (for j = 0) and iterative soft thresholding (for j = 1, 2).

460	for recovering the sparse high frequency factors $u_1 \circledast_x f$ and $u_2 \circledast_x f$. Here $\operatorname{soft}_{s\lambda} f =$
461	$\mathrm{sign}(f)\max\{ f -s\lambda,0\}$ is the soft thresholding operation, s the step size and λ the
462	regularization parameter.

463464 Second we evaluate f_{conv} := ∑_{j=0}² u_j ⊛_x f_j.
As a final reconstruction step we recover an approximation to f by deconvolution

G. ZANGERL, M. HALTMEIER



Figure 5. Left: The convolved data $y_j = v_j \circledast_t (\mathbf{A}_{m,n} \mathbf{W}_n f)$ for j = 0, 1, 2 where $\mathbf{A}_{m,n}$ is a Gaussian random matrix with subsampling factor 4. Right. Corresponding reconstructions of convolved initial pressure $u_j \circledast_x f$ using the Landweber method (for j = 0) and iterative soft thresholding (for j = 1, 2).

465 f_{conv} with kernel $\Phi = \mathbf{F}_d^{-1} \sum_{j=0}^2 |\mathbf{F}_d u_j|^2$. In this work we again use the iterative soft 466 thresholding algorithm for performing the deconvolution.

The reconstructions of the convolved phantoms for the subsampling measurements are shown in the right column in Figure 4 and the for the Gaussian measurements in the right column in Figure 5. The right column in Figure 6 shows the resulting reconstructions from

This manuscript is for review purposes only.





Figure 6. Top: Reconstruction from sparse measurements $y = \mathbf{A}_{m,n} \mathbf{W}_n f$ with subsampling factor 4 using standard ℓ^1 -minimization (left) and the proposed algorithm (right). Bottom: Same for Gaussian measurements.

subsampled measurements and Gaussian measurements. The left column shows the reconstructions using standard ℓ^1 -minimization without out multiscale sparsifying transforms. The relative ℓ^2 -reconstruction errors are 0.17 (sparse sampling) and 0.19 (Gaussian measurements) for the proposed method and 0.22 (both cases) for standard ℓ^1 -minimization with iterative soft thresholding. Notably, the high-resolution pattern is reconstructed significantly better for the proposed multiscale approach that for standard ℓ^1 -minimization.

476 **6.** Conclusion. In this paper, we derived a multiscale factorization of the wave equation. We applied the multiscale factorization to CSPAT, where reconstructions are obtained from 477 only a few compressed sensing measurements that consist of linear combinations of the signals 478 recorded by individual point-like detectors. We present a novel multiscale reconstruction 479approach that uses the acoustic reciprocal principle to achieve a multiscale decomposition of 480 the desired initial pressure by applying a family of operators that act on acoustic data in the 481 time domain. In this way sparsity of the desired initial pressure distribution can be introduced 482 for the high scales of the operators. Our numerical results demonstrate that the proposed 483 484 method improves the reconstructions in the case of compressed sensing measurements.

In future work, we will improve and analyze the reconstruction algorithm associated to the multiscale factorization. In particular, we analyze the theoretical conditions for unique recovery. Other interesting lines of future research is the extension of the proposed method to PAT with variable sound speed as well as other tomographic image reconstruction modalities. During finalization of the manuscript we found that in the context of Radon inversion with filtered backprojection, related multiscale factorizations have been proposed in [39, 7].

491 The combination of such results with compressed sensing and more advanced reconstruction

492 techniques seems an interesting line of research.

493 Appendix A. Proof of Proposition 2.2.

According to Proposition 2.1 it is sufficient to show that $\mathbf{R}^{\sharp} v$ is a solution of the equation $v = \mathbf{R} u$. Recalling the definition of \mathbf{R} in (2.5) this amounts in showing that $\mathbf{R}^{\sharp} v = \bar{u} \circ \| \cdot \|$ satisfies the integral equation $v(t) = \omega_{d-2} \int_{|t|}^{\infty} s\bar{u}(s)(s^2 - t^2)^{(d-3)/2} ds$ for $t \in \mathbb{R}$. We note that $(\theta, t) \mapsto \mathbf{R} u(t)$ is the Radon transform of the radially symmetric function u. Therefore, there is exactly one radial function satisfying the above integral equation. An explicit expression for the solution has been given in [9]. Using elementary computation, a formula has been derived in [35, p. 23]. By slight modification we obtain the following results.

Lemma A.1. The solution $u = \bar{u} \circ || \cdot ||$ of the equation $v = \mathbf{R} u$ is given by

502 (A.1)
$$\forall r > 0: \quad \bar{u}(r) := \frac{2(-1)^{d-1}}{\pi^{(d-1)/2}\Gamma((d-1)/2)} \mathbf{D}_r^{d-1} \int_r^\infty (t^2 - r^2)^{(d-3)/2} v(t) t \, \mathrm{d}t.$$

503 **Proof.** In [35, p. 23] the identity $\mathbf{D}_r^{d-1} \int_r^{\infty} (t^2 - r^2)^{(d-3)/2} v(t) t dt = 2^{-1} (-1)^{d-1} \omega_{d-2} c(d) (d-504) 2! \bar{u}(r)$ has been derived with $c(d) := 2^{1-d} \int_{-1}^{1} (1-s^2)^{(d-3)/2} ds$. Together with the identities $\omega_{d-2} = 2\pi^{(d-1)/2} / \Gamma((d-1)/2)$ and $c(d) = 2^{1-d} \pi^{1/2} \Gamma((d-1)/2) / \Gamma(d/2)$ as well as 506 $\Gamma(d/2) \Gamma((d-1)/2) = 2^{2-d} \pi^{1/2} (d-2)!$ imply the explicit solution formula (A.1).

507 In remains to bring the right hand side of Equation (A.1) in the desired form. We do this 508 separately for the even and odd dimensional case.

509 • If d is odd we have

510
$$\mathbf{D}_{r}^{(d-1)/2} \mathbf{D}_{r} \mathbf{D}_{r}^{(d-3)/2} \int_{r}^{\infty} (t^{2} - r^{2})^{(d-3)/2} v(t) t \, \mathrm{d}t$$

511
$$= (-1)^{(d-3)/2} \left((d-3)/2 \right)! \mathbf{D}_r^{(d-1)/2} \mathbf{D}_r \int_r^\infty v(t) t \, \mathrm{d}t$$

512
513 =
$$(-1)^{(d-1)/2} \frac{((d-3)/2)!}{2} \mathbf{D}_r^{(d-1)/2} v(r).$$

514 Together with Lemma A.1 this gives (2.7) for d odd.

515 •	If	d	is	even	we	first	compute
-------	----	---	----	------	----	-------	---------

516
$$\mathbf{D}_r \int_r^\infty \frac{v(t)}{\sqrt{t^2 - r^2}} t \mathrm{d}t$$

$$= \mathbf{D}_r \int_r^\infty \left(\partial_t \sqrt{t^2 - r^2}\right) \phi(t) dt$$
$$= -\mathbf{D}_r \int^\infty \sqrt{t^2 - r^2} \left(\partial_t \phi(t)\right) dt$$

519
$$= \int_{r}^{\infty} \frac{1}{2} \frac{1}{\sqrt{t^2 - r^2}} (\partial_s \phi(t)) dt$$

520 (A.2)
$$= \int_{r}^{\infty} \frac{\mathbf{D}_{t} \, \phi(t)}{\sqrt{t^{2} - r^{2}}} t \mathrm{d}t.$$

522 Therefore

523
$$\mathbf{D}_r^{d/2} \; \mathbf{D}_r^{(d-2)/2} \int_r^\infty (t^2 - r^2)^{(d-3)/2} v(t) \, t \, \mathrm{d}t$$

524 =
$$(-1)^{(d-2)/2} \Gamma((d-1)/2) \mathbf{D}_r^{d/2} \int_r^\infty (t^2 - r^2)^{-1/2} v(t) t \, \mathrm{d}t$$

525
526
$$= (-1)^{(d-2)/2} \Gamma((d-1)/2) \int_r^\infty \frac{\left(\frac{1}{2t}\frac{\partial}{\partial t}\right)^{d/2} v(t)}{\sqrt{t^2 - r^2}} t \, \mathrm{d}t \,,$$

527 where the last equality follows after (d/2)-times applying equality (A.2). This gives 528 (2.7) for d even.

529

517

518

REFERENCES

- [1] S. ACOSTA AND C. MONTALTO, Multiwave imaging in an enclosure with variable wave speed, Inverse
 Probl., 31 (2015), p. 065009.
- [2] M. AGRANOVSKY AND P. KUCHMENT, Uniqueness of reconstruction and an inversion procedure for ther moacoustic and photoacoustic tomography with variable sound speed, Inverse Probl., 23 (2007), p. 2089.
- [3] H. AMMARI, E. BRETIN, J. GARNIER, AND A. WAHAB, Time reversal in attenuating acoustic media, Contemporary Mathematics, 548 (2011), pp. 151–163.
- [4] S. ARRIDGE, P. BEARD, ET AL., Accelerated high-resolution photoacoustic tomography via compressed
 sensing, Phys. Med. Biol., 61 (2016), pp. 8908–8940.
- [5] P. BURGHOLZER, J. BAUER-MARSCHALLINGER, ET AL., Temporal back-projection algorithms for photoacoustic tomography with integrating line detectors, Inverse Probl., 23 (2007), p. S65.
- 540 [6] E. J. CANDÈS, J. K. ROMBERG, AND T. TAO, Stable signal recovery from incomplete and inaccurate
 541 measurements, Comm. Pur. Appl. Math., 59 (2006), pp. 1207–1223.
- [7] M. COSTIN, D. LAZARO-PONTHUS, ET AL., A 2D multiresolution image reconstruction method in x-ray
 computed tomography, J. X-ray Sci. Technol., 19 (2011), pp. 229–247.
- 544 [8] I. DAUBECHIES, Ten lectures on wavelets, vol. 61, Siam, 1992.
- 545 [9] S. DEANS, Gegenbauer transforms via the Radontransform, SIAM J. Math. Anal., 10 (1979), pp. 577–585.
- 546 [10] L. DESBAT, Efficient sampling on coarse grids in tomography, Inverse Probl., 9 (1993), p. 251.
- 547 [11] D. L. DONOHO, Compressed sensing, IEEE Trans. Inf. Theory, 52 (2006), pp. 1289–1306.
- 548 [12] F. DREIER AND M. HALTMEIER, Explicit inversion formulas for the two-dimensional wave equation from 549 neumann traces, SIAM J. Imaging Sci., 13 (2020), pp. 589–608.
- 550 [13] A. FARIDANI, Fan-beam tomography and sampling theory, in The Radon transform, inverse problems, and 551 tomography, vol. 63, AMS, 2006, pp. 43–66.

- [14] D. FINCH, The spherical mean value operator with centers on a sphere, Inverse Probl., 23 (2007), p. S37. 552553[15] D. FINCH, M. HALTMEIER, AND RAKESH, Inversion of spherical means and the wave equation in even 554dimensions, SIAM J. Appl. Math., 68 (2007), pp. 392-412. [16] D. FINCH, S. K. PATCH, AND RAKESH, Determining a function from its mean values over a family of 555556spheres, SIAM J. Math. Anal., 35 (2004), pp. 1213–1240. [17] S. FOUCART AND H. RAUHUT, A mathematical introduction to compressive sensing, Bull. Am. Math, 54 557558 (2017), pp. 151–165. 559[18] M. GRASMAIR, O. SCHERZER, AND M. HALTMEIER, Necessary and sufficient conditions for linear con-560vergence of ℓ^1 -regularization, Comm. Pure Appl. Math., 64 (2011), pp. 161–182. 561[19] Z. GUO, C. LI, L. SONG, AND L. V. WANG, Compressed sensing in photoacoustic tomography in vivo, J. 562 Biomed. Opt., 15 (2010), p. 021311. 563[20] M. HALTMEIER, A mollification approach for inverting the spherical mean Radon transform, SIAM J. 564Appl. Math., 71 (2011), pp. 1637–1652. 565[21] M. HALTMEIER, Universal inversion formulas for recovering a function from spherical means, SIAM J. 566Math. Anal., 46 (2014), pp. 214–232. 567[22] M. HALTMEIER, Sampling conditions for the circular Radon transform, IEEE Trans. Image Process., 25 568(2016), pp. 2910–2919. 569[23] M. HALTMEIER, T. BERER, S. MOON, AND P. BURGHOLZER, Compressed sensing and sparsity in pho-570toacoustic tomography, J. Opt., 18 (2016), p. 114004. [24] M. HALTMEIER AND . NGUYEN, Analysis of iterative methods in photoacoustic tomography with variable 571572sound speed, SIAM J. Imaging Sci., 10 (2017), pp. 751-781. 573[25] M. HALTMEIER, M. SANDBICHLER, ET AL., A sparsification and reconstruction strategy for compressed 574sensing photoacoustic tomography, J. Acoust. Soc. Am., 143 (2018), pp. 3838-3848. 575[26] M. HALTMEIER AND G. ZANGERL, Spatial resolution in photoacoustic tomography: effects of detector size 576and detector bandwidth, Inverse Probl., 26 (2010), p. 125002. 577[27] C. HUANG, K. WANG, L. NIE, AND M. A. WANG, L. V.AND ANASTASIO, Full-wave iterative image 578reconstruction in photoacoustic tomography with acoustically inhomogeneous media, IEEE Trans. Med. 579Imag., 32 (2013), pp. 1097–1110. 580[28] A. KATSEVICH, A local approach to resolution analysis of image reconstruction in tomography, SIAM J. 581Appl. Math., 77 (2017), pp. 1706–1732. 582[29] R. KOWAR, On time reversal in photoacoustic tomography for tissue similar to water, SIAM J. Imaging 583Sci., 7 (2014), pp. 509–527. 584[30] R. KOWAR AND O. SCHERZER, Photoacoustic imaging taking into account attenuation, in Mathematics 585and Algorithms in Tomography, vol. 18, Springer, 2012, pp. 54–56. 586[31] P. KUCHMENT AND L. KUNYANSKY, Mathematics of thermoacoustic tomography, Eur. J. Appl. Math., 19 587(2008), pp. 191–224. 588[32] L. KUNYANSKY, Reconstruction of a function from its spherical (circular) means with the centers lying 589on the surface of certain polygons and polyhedra, Inverse Probl., 27 (2011), p. 025012. 590[33] L. KUNYANSKY, Inversion of the spherical means transform in corner-like domains by reduction to the 591 classical Radon transform, Inverse Probl., 31 (2015), p. 095001. 592 [34] S. MALLAT, A wavelet tour of signal processing: The sparse way, Elsevier/Academic Press, Amsterdam, 593third ed., 2009. [35] F. NATTERER, Computerized tomography, in The Mathematics of Computerized Tomography, Springer, 5945951986, pp. 1-8. 596[36] F. NATTERER, Sampling and resolution in CT, in Computerized tomography (Novosibirsk, 1993), VSP, 597 Utrecht, 1995, pp. 343–354. 598 [37] L. V. NGUYEN, A family of inversion formulas for thermoacoustic tomography, Inverse Probl., 3 (2009), 599pp. 649–675. 600 [38] R. NUSTER, G. ZANGERL, M. HALTMEIER, AND G. PALTAUF, Full field detection in photoacoustic tomog-
- 601 raphy, Opt. Express, 18 (2010), pp. 6288–6299.
 602 [39] F. PEYRIN, M. ZAIM, AND R. GOUTTE, Multiscale reconstruction of tomographic images, in Proc. IEEE
 603 Int. Symp. on Time-Frequency and Time-Scale Analysis, 1992, pp. 219–222.
- 604 [40] J. POUDEL, Y. LOU, AND M. A. ANASTASIO, A survey of computational frameworks for solving the 605 acoustic inverse problem in three-dimensional photoacoustic computed tomography, Phys. Med. Biol.,

606 64 (2019), p. 14TR01.
607 [41] J. PROVOST AND F. LESAGE, The application of compressed sensing for photo-acoustic tomography, IEEE
608 Trans. Med. Imag., 28 (2009), pp. 585–594.

Trans. Med. Imag., 28 (2009), pp. 585–594.
[42] M. SANDBICHLER, F. KRAHMER, ET AL., A novel compressed sensing scheme for photoacoustic tomography, SIAM J. Appl. Math., 75 (2015), pp. 2475–2494.

611 [43] P. STEFANOV, Semiclassical sampling and discretization of certain linear inverse problems, 612 arXiv:1811.01240, (2018).

613 [44] P. STEFANOV AND G. UHLMANN, *Thermoacoustic tomography with variable sound speed*, Inverse Probl., 614 25 (2009), p. 075011.

615 [45] M. XU AND L. V. WANG, Photoacoustic imaging in biomedicine, Rev. Sci. Instrum., 77 (2006), p. 041101.

616 [46] G. ZANGERL, O. SCHERZER, AND M. HALTMEIER, Exact series reconstruction in photoacoustic tomogra-617 phy with circular integrating detectors, Commun. Math. Sci., 7 (2009), pp. 665–678.