Leopold-Franzens-Universität Innsbruck

A Novel Compressed Sensing Scheme for Photoacoustic Tomography

M. Sandbichler, F. Krahmer, T.Berer, P.Burgholzer, M. Haltmeier

APPLIEDMATHEMATICS

A Novel Compressed Sensing Scheme for Photoacoustic Tomography

M. Sandbichler[†]   F. Krahmer[‡]   T. Berer[§]   P. Burgholzer[§]

M. Haltmeier[†]

January 2015

Abstract

Speeding up the data acquisition is one of the central aims to advance tomographic imaging. On the one hand, this reduces motion artifacts due to undesired movements, and on the other hand this decreases the examination time for the patient. In this article, we propose a new scheme for speeding up the data collection process in photoacoustic tomography. Our proposal is based on compressed sensing and reduces acquisition time and system costs while maintaining image quality. As measurement data we use random combinations of pressure values that we use to recover a complete set of pressure data prior to the actual image reconstruction. We obtain theoretical recovery guarantees for our compressed sensing scheme and support the theory by reconstruction results on simulated data as well as on experimental data.

Keywords. Photoacoustic imaging, computed tomography, compressed sensing, lossless expanders, wave equation.

AMS classification numbers. 45Q05, 94A08, 92C55.

1   Introduction

Photoacoustic tomography (PAT) is a recently developed non-invasive medical imaging technology whose benefits combine the high contrast of pure optical imaging with the high spatial resolution of pure ultrasound imaging [4, 61, 63]. In order to speed up the measurement process, in this paper we propose a novel compressed sensing approach for PAT that uses random combinations of the induced pressure as measurement data. The proposed strategy yields recovery guarantees and furthermore comes with an efficient numerical implementation allowing high resolution real time imaging. We thereby focus on a variant of PAT using integrating line detectors proposed in [10, 53, 6]. Our strategy, however, can easily be adapted to more classical PAT setups using arrays of point-like detectors.

Our proposal is based on the main components of compressed sensing, namely randomness and sparsity. Compressed sensing is one of the most influential discoveries in applied mathematics and signal processing of the past decade [15, 21]. By combining the benefits of data

---

[†]Department of Mathematics, University of Innsbruck, Technikerstraße 13a, A-6020 Innsbruck, Austria. E-mail: {`michael.sandbichler,markus.haltmeier`}`@uibk.ac.at`

[‡]Institute for Numerical and Applied Mathematics, University of Göttingen, Lotzestraße 16-18, 37083 Göttingen, Germany. E-mail: `f.krahmer@math.uni-goettingen.de`

[§]Christian Doppler Laboratory for Photoacoustic Imaging and Laser Ultrasonics, and Research Center for Non-Destructive Testing (RECENDT), Altenberger Straße 69, 4040 Linz, Austria. E-mail: {`thomas.berer,peter.burholzer`}`@recendt.at`
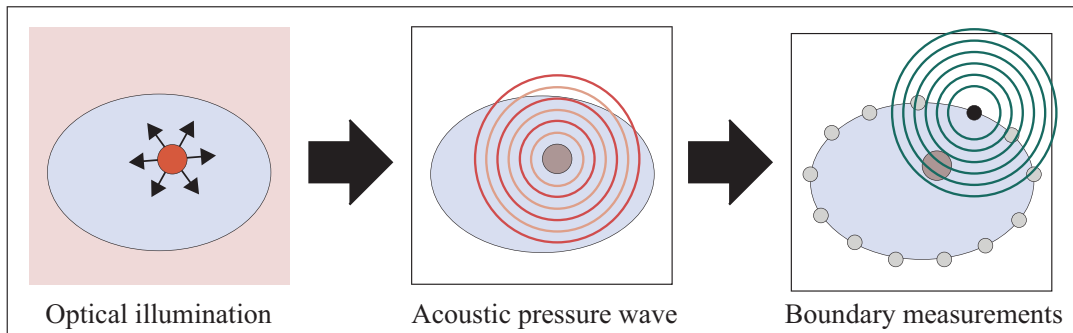
Figure 1.1: BASIC PRINCIPLE OF PAT. A semi-transparent sample is illuminated with a short optical pulse that induces an acoustic pressure wave. The induced pressure is measured outside of the sample and used to recover an image of the interior.

compression and data acquisition it allows to recover a signal from far fewer linear measurements than suggested by Shannon's sampling theorem. It has led to several new proposed sampling strategies in medical imaging, for example for speeding up MRI data acquisition (see [47, 48]) or completing under-sampled CT images [19]. Another prominent application of compressed sensing is the single pixel camera (see [22]) that circumvents the use of several expensive high resolution sensors in digital photography.

1.1 Photoacoustic tomography (PAT)

PAT is based on the generation of acoustic waves by illuminating a semi-transparent sample with short optical pulses (see Figure 1.1). When the sample is illuminated with a short laser pulse, parts of the optical energy become absorbed. Due to thermal expansion a subsequent pressure wave is generated depending on the structure of the sample. The induced pressure waves are recorded outside of the sample and used to recover an image of the interior, see [9, 43, 63, 61].
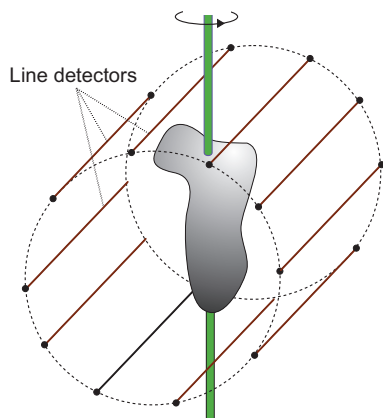

Figure 1.2: PAT WITH INTEGRATING LINE DETECTORS. An array of integrating line detectors measures integrals of the induced acoustic pressure over a certain number of parallel lines. These data are used to recover a linear projection of the object in the first step. By rotating the array of line detectors around a single axis several projection images are obtained and used to recover the actual three dimensional object in a second step.

In this paper, we consider a special variant of photoacoustic tomography that uses integrating line detectors for recording the pressure waves, as proposed in [10]. As illustrated in Figure 1.2, an array of line detectors is arranged around the investigated sample and measures integrals of the pressure wave over a certain number of parallel lines. Assuming constant speed of sound, the pressure integrated along the direction of the line detectors satisfies the

two dimensional wave equation

$$\begin{cases} \partial_t^2 p(x,t) - \Delta p(x,t) = 0\,, & \text{for } (x,t) \in \mathbb{R}^2 \times (0,\infty) \\ \quad\quad\quad\quad p\,(x,0) = f(x)\,, & \text{for } x \in \mathbb{R}^2 \\ \quad\quad\quad\quad \partial_t p\,(x,0) = 0\,, & \text{for } x \in \mathbb{R}^2\,, \end{cases} \tag{1}$$

where the time scaling is chosen in such a way that the speed of sound is normalized to one. The initial datum f in (1) is the two dimensional projection image of the actual, three dimensional initial pressure distribution.

Image reconstruction in PAT with integrating line detectors can be performed via a two-stage approach [9, 52]. In the first step the measured pressure values corresponding to values of the solution of (1) outside the support of f are used to reconstruct a linear projection (the initial data in (1)) of the three dimensional initial pressure distribution. This procedure is repeated by rotating the array of line detectors around a single axis which yields projection images for several angles. In a second step these projection images are used to recover the three dimensional initial pressure distribution by inverting the classical Radon transform. In this work we focus on the first problem of reconstructing the (two dimensional) initial pressure distribution f in (1).

Suppose that the integrating line detectors are arranged on the surface of a circular cylinder of radius $\rho > 0$, and that the object is located inside that cylinder (see Figure 1.2). The data measured by the array of line detectors are then modeled by

$$p_j := p\,(z_j,\,\cdot\,) : [0, 2\rho] \to \mathbb{R}\,, \tag{2}$$

$$z_j := \begin{pmatrix} \rho\cos\left(2\pi(j-1)/N\right) \\ \rho\sin\left(2\pi(j-1)/N\right) \end{pmatrix}\,, \quad \text{for } j = 1, \ldots, N\,, \tag{3}$$

where p_j is the pressure signal corresponding to the j-th line detector. Since the two dimensional initial pressure distribution f is supported in a disc of radius ρ and the speed of sound is constant and normalized to one, no additional information is contained in data $p\,(z_j, t)$ for $t > 2\rho$. This can be seen, for example by exploiting the explicit relations between the two dimensional pressure signals $p\,(z_j,\,\cdot\,)$ and the spherical means (compare with Subsection 3.3) of the initial pressure distribution; see [54].

Of course, in practice also the functions $p_j \colon [0, 2\rho] \to \mathbb{R}$ have to be represented by discrete samples. However, temporal samples can easily be collected at a high sampling rate compared to the spatial sampling, where each sample requires a separate sensor. It is therefore natural to consider the semi-discrete data model (2). Our compressed PAT scheme could easily be adapted to a fully discretized data model.

1.2   Compressed sensing PAT

When using data of the form (2), high resolution imaging requires the number N of detector locations to be sufficiently large. As the fabrication of an array of parallel line detectors is demanding, most experiments using integrating line detectors have been carried out using a single line detector, scanned on circular paths using scanning stages [51, 30]. Recently, systems using arrays of 64 parallel line detectors have been demonstrated [29, 2]. The most costly building blocks in such devices are the analog to digital converters (ADC). For completely parallel readout, a separate ADC is required for every detector channel. In order to reduce costs, in these practical implementations two to four line detector channels are multiplexed to one ADC. For collecting the complete pressure data, the measurements have to be performed two (respectively four) times, because only 32 (respectively 16) of the 64 line measurements can be read out in parallel. This, again, leads to an increased overall measurement time. For

example, using an excitation laser with a repetition rate of 10 Hz and two times multiplexing, a measurement time of 0.2 s for a projection image has been reported in [29]. Without multiplexing, this measurement time would reduce to the half.

In order to speed up the scanning process and to reduce system costs, in this paper we propose a novel compressed sensing approach that allows to perform a smaller number of random measurements with a reduced number of ADCs, while retaining high spatial resolution. For that purpose, instead of collecting individually sampled data $p_j(t)$ as in (2), we use random combinations

$$y_i(t) = \sum_{j \in J_i} p_j(t) \quad \text{for } i \in \{1, \dots, m\} \text{ and } t \in [0, 2\rho], \tag{4}$$

where $m \ll N$ is the number of compressed sensing measurements and $J_i \subset \{1, \dots, N\}$ corresponds to the random set of detector locations contributing to the i-th measurement. In the reconstruction process the random linear combinations $y_i(t)$ are used to recover the full set of pressure data $p_1(t), \dots, p_N(t)$ using compressed sensing techniques. The initial pressure distribution f in (1) is subsequently recovered from the completed pressure data by applying standard PAT reconstruction algorithms such as time reversal [11, 34, 60] or filtered backprojection [9, 23, 24, 32, 44, 62].

A naive approach for recovering the pressure data from the random measurements $y_i(t)$ would be to solve (4) for $p_j(t)$ separately for each $t \in [0, 2\rho]$. Since $m \ll N$, this is a severely underdetermined system of linear equations and its solution requires appropriate prior knowledge of the unknown parameter vector. Compressed sensing suggests to use the sparsity of the parameter vector in a suitable basis for that purpose. However, recovery guarantees for zero/one measurements of the type (4) are basis-dependent and require the parameter to be sparse in the standard basis rather than sparsity in a different basis such as orthonormal wavelets (see Subsection 2.2). However, for pressure signals (2) of practical relevance such sparsity assumption in the original basis does not hold.

In this work we therefore propose a different approach for solving (4) by exploiting special properties of the data in PAT. For that purpose we apply a transformation that acts in the temporal variable only, and makes the transformed pressure values sufficiently sparse in the spatial (angular) component. In Subsection 3.3 we present an example of such a transform. The application of a sparsifying transform to (4) yields linear equations with unknowns being sparse in the angular variable. It therefore allows to apply sparse recovery results for the zero/one measurements under consideration.

1.3 Relations to previous work

A different compressed sensing approach for PAT has been considered in [55, 31]. In these articles, standard point samples (such as (2)) have been used as measurement data and no recovery guarantees have been derived. Further, in [55, 31] the phantom is directly reconstructed from the incomplete data, whereas we first complete the data using sparse recovery techniques. Our approach is more related to a compressed sensing approach for PAT using a planar detector array that has been proposed in [35] and also uses random zero/one combinations of pressure values and recovers the complete pressure prior to the actual image reconstruction. However, in [35] the sparsifying transform is applied in spatial domain where recovery guarantees are not available as noted above. We finally notice that our proposal of using a sparsifying temporal transform can easily be extended to planar detector arrays in two or three spatial dimensions; compare Section 5.

1.4 Outline

The rest of this paper is organized as follows. In Section 2 we review basic results from compressed sensing that we require for our proposal. We therefore focus on recovery guarantees for zero/one matrices modeled by lossless expanders; see Subsection 2.2. In Section 3 we present the mathematical framework of the proposed PAT compressed sensing scheme. The sparsity in the spatial variable, required for ℓ^1-minimization, is obtained by applying a transformation acting in the temporal variable. An example of such a transformation is given in Subsection 3.3. In Section 4 we present numerical results supporting our theoretical investigations. The paper concludes with a short discussion in Section 5.

2 Background from compressed sensing

In this section we shortly review basic concepts and results of compressed sensing (sometimes also termed compressive sampling). Our main focus will be on recovery results for lossless expanders, which are the basis of our PAT compressed sensing scheme.

2.1 Compressed sensing

Suppose one wants to sense a high dimensional data vector $\mathbf{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$, such as a digital image. The classical sampling approach is to measure each component of x_i individually. Hence, in order to collect the whole data vector one has to perform N separate measurements, which may be too costly. On the other hand it is well known and the basis of data compression algorithms, that many datasets are compressible in a suitable basis. That is, a limited amount of information is sufficient to capture the high dimensional vector $\mathbf{x}$.

Compressed sensing incorporates this compressibility observation into the sensing mechanism [15, 17, 21]. Instead of measuring each coefficient of the data vector individually, one collects linear measurements

$$\mathbf{A}\mathbf{x} = \mathbf{y}\,, \tag{5}$$

where $\mathbf{A} \in \mathbb{R}^{m \times N}$ is the measurement matrix with $m \ll N$, and $\mathbf{y} = (y_1, \ldots, y_m) \in \mathbb{R}^m$ is the measurement vector. Any component of the data vector can be interpreted as a scalar linear measurement performed on the unknown $\mathbf{x}$, and the assumption $m \ll N$ means that far fewer measurements than parameters are available. As $m \ll N$, the system (5) is highly underdetermined and cannot be uniquely solved (at least without additional information) by standard linear algebra.

Compressed sensing overcomes this obstacle by utilizing randomness and sparsity. Recall that the vector $\mathbf{x} = (x_1, \ldots, x_N)$ is called s-sparse if the support

$$\mathrm{supp}\,(\mathbf{x}) := \{j \in \{1, \ldots, N\} : x_j \neq 0\} \tag{6}$$

contains at most s elements. Results from compressed sensing state that for suitable $\mathbf{A}$, any s-sparse $\mathbf{x} \in \mathbb{R}^n$ can be found via the optimization problem

$$\begin{aligned} \underset{\mathbf{z} \in \mathbb{R}^N}{\text{minimize}} \quad & \|\mathbf{z}\|_1 = \sum_{j=1}^N |z_j| \\ \text{such that} \quad & \mathbf{A}\mathbf{z} = \mathbf{y}\,. \end{aligned} \tag{7}$$

By relaxing the equality constraint $\mathbf{A}\mathbf{z} = \mathbf{y}$, the optimization problem (7) can be adapted to data which are only approximately sparse and noisy [16].

A sufficient condition to guarantee recovery is the so called *restricted isometry property* (RIP), requiring that for any s-sparse vector x, we have

$$(1 - \delta)\|\mathbf{x}\|_2 \leq \|\mathbf{A}\mathbf{x}\|_2 \leq (1 + \delta)\|\mathbf{x}\|_2 \quad \text{for some small } \delta \in (0, 1).$$

The smallest constant δ satisfying this inequality is called the s-restricted isometry constant of A and denoted by δ_s. Under certain conditions on δ_s, recovery guarantees for sparse and approximately sparse data can be obtained, see for example [14, 12].

While the restricted isometry itself is deterministic, to date all constructions that yield near-optimal embedding dimensions m are based on random matrices. Sub-gaussian random matrices satisfy the RIP with high probability for an order-optimal embedding dimension $m = \mathcal{O}(s \log(N/s))$, see e.g. [1]. Partial random Fourier matrices (motivated by MRI measurements) and subsampled random convolutions (motivated by remote sensing) have been shown to allow for order-optimal embedding dimensions up to logarithmic factors, see [58, 56] and [57, 40], respectively.

The sparsity is often not present in the standard basis of $\mathbb{R}^N$, but in a special *sparsifying* basis, such as wavelets. For matrices with subgaussian rows this does not cause a problem, as the rotation invariance of subgaussian vectors ensures that after incorporating an orthogonal transform, the resulting random matrix construction still yields RIP matrices with high probability. As a consequence, the sparsifying basis need not be known for designing the measurement matrix $\mathbf{A}$, which is often referred to as universality of such measurements [1].

Many structured random measurement systems including the partial random Fourier and the subsampled random convolution scenarios mentioned above, however, do not exhibit universality. For example, one can easily see that subsampled Fourier measurements cannot suffice if the signal is sparse in the Fourier basis. While it has been shown that this problem can be overcome by randomizing the column signs [41], such an alteration often cannot be implemented in the sensing setup. Another way to address this issue is by requiring incoherence between the measurement basis and the sparsity basis [13]. That is, one needs that inner products between vectors of the two bases are uniformly small. If not all, but most of these inner products are small, one can still recover, provided that one adjusts the sampling distribution accordingly; this scenario includes the case of Fourier measurements and Haar wavelet sparse signals [42].

Incoherence is also the key to recovery guarantees for gradient sparse signals. Namely, many natural images are observed to have an approximately sparse discrete gradient. As a consequence, it has been argued using a commutation argument that one can recover the signal from uniformly subsampled Fourier measurements via minimizing the ℓ^1 norm of the discrete gradient, the so-called total variation (TV) [15]. TV minimization had already proven to be a favorable method in image processing, see, for example [59, 18]. A problem with this approach is that the compressed sensing recovery guarantees then imply good recovery of the gradient, not of the signal itself. Small errors in the gradient, however, can correspond to substantial errors in the signal, which is why this approach can only work if no noise is present. A refined analysis that allows for noisy measurements requires the incoherence of the measurement basis to the Haar wavelet basis [50]. Again, TV minimization is considered for recovery. By adjusting the sampling distribution, these results have also been shown to extend to Fourier measurements and other systems with only most measurement vectors incoherent to the Haar basis [42].

For the measurement matrices considered in this work, namely zero/one matrices based on expander graphs, recovery guarantees build on an ℓ^1-version of the restricted isometry property, namely one requires that

$$(1 - \delta)\|\mathbf{x}\|_1 \leq \|\mathbf{A}\mathbf{x}\|_1 \leq (1 + \delta)\|\mathbf{x}\|_1$$

for all sufficiently sparse $\mathbf{x}$ and some constant $\delta > 0$; see [7] and Subsection 2.2 below. As the

ℓ^1-norm is not rotation invariant, basis transformations typically destroy this property. That is, not even incoherence based recovery guarantees are available; recovery results only hold in the standard basis. Thus an important aspect of our work will be to ensure (approximate) sparsity in the standard basis. This will be achieved by applying a particular transformation in the time variable.

2.2 Recovery results for lossless expanders

Recall that we seek recovery guarantees for a measurement setup, where each detector is switched on exactly d out of m times. That is, one obtains a binary measurement matrix $\mathbf{A} \in \{0,1\}^{m \times N}$ with exactly d ones in each column. It therefore can be interpreted as the adjacency matrix of a left d-regular bipartite graph. Under certain additional conditions, such a bipartite graph is a lossless expander (see Definition 2.1) which, as we will see, guarantees stable recovery of sparse vectors. Expander graphs have been used since the 1970s in theoretical computer science, originating in switching theory from modeling networks connecting many users, cf. [39] for further applications. They have also been useful in measure theory, where it was possible to solve the Banach-Ruziewicz problem using tools from the construction of expander graphs, see [45] for a detailed examination of this connection. For a survey on expander graphs and their applications, see for example [33, 46].

Compressed sensing with expander graphs has been considered in [7, 8, 36, 37, 64], where also several efficient algorithms for the solution of compressed sensing problems using expander graphs have been proposed. A short review of sparse recovery algorithms using expander like matrices is given in [26]. In this subsection we recall main compressed sensing results using lossless expanders as presented in the recent monograph [25], where the proofs of all mentioned theorems can be found in Section 13.
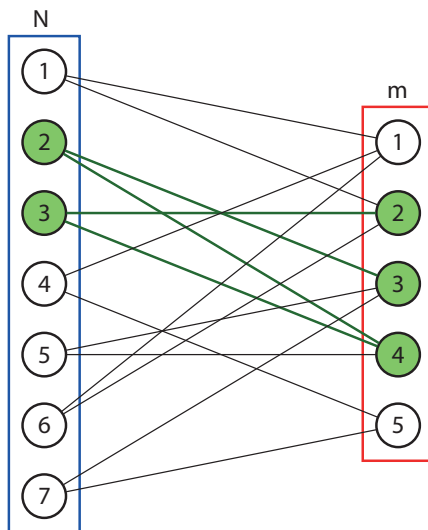

Figure 2.1: Example of a left d-regular bipartite graph with $d = 2$, $N = 7$ left vertices, and $m = 5$ right vertices. Here $d = 2$ because exactly 2 edges emerge at each left vertex. For $J = \{2, 3\}$ the set of right vertices connected to J is given by $R(J) = \{2, 3, 4\}$.

Recall that a bipartite graph consists of a triple (L, R, E), where L is the set of left vertices, R the set of right vertices, and $E \subset L \times R$ is the set of edges. Any element $(j, i) \in E$ represents an edge with left vertex $j \in L$ and right vertex $i \in R$.

A bipartite graph (L, R, E) is called d-regular for some $d \geq 0$, if for every given left vertex

$j \in L$, the number of edges $(j, i) \in E$ emerging from j is exactly equal to d. Finally, for any subset $J \subset L$ of left vertices, let

$$R(J) := \{i \in R \colon \text{ there exists some } j \in J \text{ with } (j, i) \in E\}$$

denote the set of right vertices connected to J. Obviously, for any left d-regular bipartite graph and any $J \subset L$, we have $|R(J)| \leq d|J|$.

Definition 2.1 (Lossless expander)**.** Let (L, R, E) be a left d-regular bipartite graph, $s \in \mathbb{N}$, and $\theta \in [0, 1]$. Then (L, R, E) is called an (s, d, θ)-lossless expander, if

$$|R(J)| \geq (1 - \theta)d|J| \quad \text{for all } J \subset L \text{ with } |J| \leq s. \tag{8}$$

For any left d-regular bipartite graph, the smallest number $\theta \geq 0$ satisfying (8) is called the s-th restricted expansion constant and denoted by θ_s.

The following theorem states that a randomly chosen d-regular bipartite graph will be a lossless expander with high probability.

Theorem 2.2 (Regular bipartite graphs are expanders with high probability)**.** *For every $0 < \epsilon < \frac{1}{2}$, every $\theta \in (0, 1)$ and every $s \in \mathbb{N}$, the proportion of (s, d, θ)-lossless expanders among the set of all left d-regular bipartite graphs having N left vertices and m right vertices exceeds $1 - \epsilon$, provided that*

$$d = \left\lceil \frac{1}{\theta} \ln\left(\frac{eN}{\epsilon s}\right) \right\rceil \quad and \quad m \geq c_\theta s \ln\left(\frac{eN}{\epsilon s}\right). \tag{9}$$

Here c_θ is a constant only depending on θ, e is Euler's constant, $\ln(\,\cdot\,)$ denotes the natural logarithm and $\lceil x \rceil$ denotes the smallest integer larger or equal to x.

According to Theorem 2.2, any randomly chosen left d-regular bipartite graph is a lossless expander with high probability, provided that (9) is satisfied. The following theorem states that the adjacency matrix $\mathbf{A} \in \{0, 1\}^{m \times N}$,

$$\mathbf{A}_{ij} = 1 \colon \iff (j, i) \in E, \tag{10}$$

of any lossless expander with left vertices $L = \{1, \ldots, N\}$ and right vertices $R = \{1, \ldots, m\}$ yields stable recovery of any sufficiently sparse parameter vector. The result was first established in [7], we present the version found in [25].

Theorem 2.3 (Recovery guarantee for lossless expanders)**.** *Let $\mathbf{A} \in \{0, 1\}^{m \times N}$ be the adjacency matrix of a left d-regular bipartite graph having $\theta_{2s} < 1/6$. Further, let $\mathbf{x} \in \mathbb{C}^N$, $\eta > 0$ and $\mathbf{e} \in \mathbb{C}^m$ satisfy $\|\mathbf{e}\|_1 \leq \eta$, set $\mathbf{b} := \mathbf{A}\mathbf{x} + \mathbf{e}$, and denote by $\mathbf{x}_\star$ a solution of*

$$\begin{aligned} \underset{\mathbf{z} \in \mathbb{C}^N}{\text{minimize}} \quad & \|\mathbf{z}\|_1 \\ \text{such that} \quad & \|\mathbf{A}\mathbf{z} - \mathbf{b}\|_1 \leq \eta. \end{aligned} \tag{11}$$

Then

$$\|\mathbf{x} - \mathbf{x}_\star\|_1 \leq \frac{2(1 - 2\theta_{2s})}{(1 - 6\theta_{2s})}\sigma_s(\mathbf{x})_1 + \frac{4}{(1 - 6\theta_{2s})d}\eta.$$

Here the quantity $\sigma_s(\mathbf{x})_1 := \inf\{\|\mathbf{x} - \mathbf{z}\|_1 \colon \mathbf{z} \text{ is } s\text{-sparse}\}$ measures by how much the vector $\mathbf{x} \in \mathbb{C}^N$ fails to be s-sparse.

Combining Theorems 2.2 and 2.3, we can conclude that the adjacency matrix $\mathbf{A}$ of a randomly chosen left d-regular bipartite graph will, with high probability, recover any sufficiently sparse vector $\mathbf{x} \in \mathbb{C}^N$ by basis pursuit reconstruction (11).

3 Mathematical framework of the proposed PAT compressed sensing scheme

In this section we describe our proposed compressed sensing strategy. As mentioned in the introduction, we focus on PAT with integrating line detectors, which is governed by the two dimensional wave equation (1). In the following we first describe the compressed sensing measurement setup in Subsection 3.1 and describe the sparse recovery strategy in Subsection 3.2. As the used pressure data are not sparse in the original domain we introduce a temporal transform that makes the data sparse in the spatial domain. In Subsection 3.3 we present an example of such a sparsifying temporal transform. In Subsection 3.4 we finally summarize the whole PAT compressed sensing scheme.
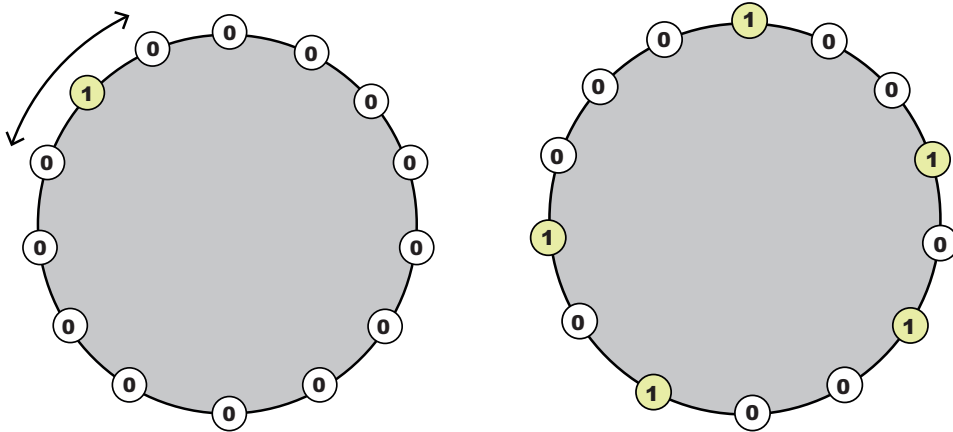

Figure 3.1: LEFT: Classical PAT sampling, where a single detector is moved around the object to collect individual pressure signal p_j. RIGHT: Compressed sensing approach where each measurement consists of a random zero/one combination of individual pressure values.

3.1 Compressed sensing PAT

We define the unknown full sample pressure $p_j \colon [0, 2\rho] \to \mathbb{R}$, for $j \in \{1, \dots, N\}$ by (2), (3), where p is the solution of the two dimensional wave equation (1). We suppose that the (two dimensional) initial pressure distribution f is smooth and compactly supported in $B_R(0)$ which implies that any p_j is smooth and vanishes in a neighborhood of zero. Furthermore we assume that the data (2) are sampled finely enough to allow for f to be reconstructed from p_j by standard PAT reconstruction algorithms such as time reversal [11, 34, 60] or filtered backprojection [9, 23, 24, 32, 44, 62].

Instead of measuring each pressure signal p_j separately, we take m compressed sensing measurements. For each measurement, we select sensor locations z_j with $j \in J_i$ at random and take the sum of the corresponding pressure values p_j. Thus the i-th measurement is given by

$$y_i(t) := \sum_{j \in J_i} p_j(t) \quad \text{for } i \in \{1, \dots, m\} \,, \tag{12}$$

where $J_i \subset \{1, \dots, N\}$ corresponds to the set of all detector locations selected for the i-th measurement, and $m \ll N$ is the number of compressed sensing measurements.

In practice, the compressed sensing measurements could be realized by summation of several detector channels, using a configurable matrix switch and a summing amplifier. Even more simply, a summing amplifier summing over all N channels could be used, while individual

9

detector channels z_j are turned `on` or `off`, using solid state relays. Thereby, only one ADC is required for one compressed sensing measurement. Performing m compressed sensing measurements in parallel is facilitated by using m ADCs in parallel, and the according number of matrix switches, relays, and summing amplifiers.

In the following we write

$$\mathbf{p}\colon (0,\infty) \to \mathbb{R}^N \colon t \mapsto \mathbf{p}(t) = (p_1(t),\dots,p_N(t))^\mathsf{T}\,, \tag{13}$$

$$\mathbf{y}\colon (0,\infty) \to \mathbb{R}^m \colon t \mapsto \mathbf{y}(t) = (y_1(t),\dots,y_m(t))^\mathsf{T}\,, \tag{14}$$

for the vector of unknown complete pressure signals and the vector of the corresponding measurement data, respectively. Further, we denote by

$$\mathbf{A} \in \{0,1\}^{m\times N} \text{ with entries } \mathbf{A}_{ij} := \begin{cases} 1, & \text{for } j \in J_i \\ 0, & \text{for } j \notin J_i\,, \end{cases}$$

the matrix whose entries in the i-th row correspond the sensor locations selected for the i-th measurement. In order to apply the exact recovery guarantees from Subsection 2.2, we require that each row of $\mathbf{A}$ contains exactly d ones, where $d \in \mathbb{N}$ is some fixed number. Practically, this means that each detector location contributes to exactly d of the m measurements, which also guarantees that the measurements are well calibrated and there is no bias towards some of the detector locations.

Recovering the complete pressure data (2) from the compressed sensing measurements (4) can be written as an uncoupled system of under-determined linear equations,

$$\mathbf{A}\mathbf{p}(t) = \mathbf{y}(t) \quad \text{for any } t \in [0,2\rho]\,, \tag{15}$$

where $\mathbf{A}\mathbf{p}(t) := \mathbf{A}\,(\mathbf{p}(t))$ for any t. From (15), we would like to recover the complete set of pressure values $\mathbf{p}(t)$ for all times $t \in [0,2\rho]$. Compressed sensing results predict that under certain assumptions on the matrix $\mathbf{A}$ any s-sparse vector $\mathbf{p}(t)$ can be recovered from $\mathbf{A}\mathbf{p}(t)$ by means of sparse recovery algorithms like ℓ^1-minimization.

Similar to many other applications (cf. Section 2 above), however, we cannot expect sparsity in the original domain. Instead, one has $\mathbf{p}(t) = \Psi\,\mathbf{x}(t)$, where Ψ is an appropriate orthogonal transform and $\mathbf{x}(t) \in \mathbb{R}^N$ is a sparse coefficient vector. This yields a sparse recovery problem for $\mathbf{x}(t)$ involving the matrix $\mathbf{A}\,\Psi$, which does not inherit the the recovery guarantees of Subsection 2.2. Hence we have to find a means to establish sparsity without considering different bases. Our approach will consist of applying a transformation in the time domain that sparsifies the pressure in the spatial domain. A further advantage of working in the original domain is that the structure of $\mathbf{A}$ allows the use of specific efficient algorithms like sparse matching pursuit [8] or certain sublinear-time algorithms like [37, 64].

3.2 Reconstruction strategy

We denote by $\mathcal{G}([0,2\rho])$ the set of all infinitely differentiable functions $g\colon [0,2\rho] \to \mathbb{R}$ that vanish in a neighborhood of zero. To obtain the required sparsity, we will work with a sparsifying transformation

$$\mathbf{T}\colon \mathcal{G}([0,2\rho]) \to \mathcal{G}([0,2\rho])\,, \tag{16}$$

that is, $\mathbf{T}\mathbf{p}(t) \in \mathbb{R}^N$ can be sufficiently well approximated by a sparse vector for any $t \in [0,2\rho]$ and certain classes of practically relevant data $\mathbf{p}(t)$. Here we use the convention that $\mathbf{T}$ applied to a vector valued function $\mathbf{g} = (g_1,\dots,g_k)$ is understood to be applied in each component separately, that is,

$$\mathbf{T}\mathbf{g}(t) := ((\mathbf{T}g_1)(t),\dots,(\mathbf{T}g_k)(t))\,, \quad \text{for } t \in [0,2\rho]\,. \tag{17}$$

We further require that $\mathbf{T}$ is an injective mapping, such that any $g \in \mathcal{G}([0, 2\rho])$ can be uniquely recovered from the transformed data $\mathbf{T}g$. See Subsection 3.3 for the design of such a sparsifying transformation.

Since any temporal transformation interchanges with $\mathbf{A}$, application of the sparsifying temporal transformation $\mathbf{T}$ to the original system (15) yields

$$\mathbf{A}\left(\mathbf{T}\mathbf{p}(t)\right) = \mathbf{T}\mathbf{y}(t) \quad \text{for any } t \in [0, 2\rho]. \tag{18}$$

According to the choice of the temporal transform, the transformed pressure $\mathbf{T}\mathbf{p}(t)$ can be well approximated by a vector that is sparse in the spatial component. We therefore solve, for any $t \in [0, T]$, the following ℓ^1-minimization problem

$$\begin{aligned} \underset{\mathbf{q} \in \mathbb{R}^N}{\text{minimize}} \quad & \|\mathbf{q}\|_1 \\ \text{such that} \quad & \|\mathbf{A}\mathbf{q} - \mathbf{T}\mathbf{y}(t)\|_1 \leq \eta, \end{aligned} \tag{19}$$

for some error threshold $\eta > 0$. As follows from Theorem 2.3, the solution $\mathbf{q}_\star(t)$ of the ℓ^1-minimization problem (19) provides an approximation to $\mathbf{T}\mathbf{p}(t)$ and consequently we have $\mathbf{T}^{-1}\mathbf{q}_\star(t) \simeq \mathbf{p}(t)$ for all $t \in [0, 2\rho]$. Note that the use of ℓ^1-norm $\|\cdot\|_1$ for the constraint in (19) is required for the stable recovery guarantees for our particular choice of the matrix $\mathbf{A}$ containing zero/one entries using data that is noisy and only approximately sparse.

As a further benefit, compressed sensing measurements may show an increased signal to noise ratio. For that purpose consider Gaussian noise in the pressure data, where instead of the exact pressure data $\mathbf{p}(t)$, we have noisy data $\tilde{\mathbf{p}}(t) = \mathbf{p}(t) + \boldsymbol{\eta}$. For the sake of simplicity assume that the entries $\eta_j \sim \mathcal{N}(0, \sigma^2)$ of $\boldsymbol{\eta}$ are independent and identically distributed. The corresponding noisy (and rescaled) measurements are then given by

$$\frac{1}{|J_i|} \sum_{j \in J_i} (p_j(t) + \eta_j) = \frac{1}{|J_i|} \sum_{j \in J_i} p_j(t) + \frac{1}{|J_j|} \sum_{j \in J_i} \eta_j.$$

The variance in the compressed sensing measurements is therefore $\sigma^2/|J_i|$ compared to σ^2 in the individual data $\tilde{p}_j(t)$. Assuming some coherent averaging in the signal part this yields an increased signal to noise ratio reflecting the inherent averaging of compressed sensing measurements.

3.3 Example of a sparsifying temporal transform

As we have seen above, in order to obtain recovery guarantees for our proposed compressed scheme, we require a temporal transformation that sparsifies the pressure signals in the angular component. In this section, we construct an example of such a sparsifying transform.

Since the solution of the two dimensional wave equation can be reduced to the spherical means, we will construct such a sparsifying transform for the spherical means

$$\mathbf{M}f(z, r) := \frac{1}{2\pi} \int_{\mathbb{S}^1} f(z + r\omega) \mathrm{d}\sigma(\omega) \quad \text{for } (z, r) \in \partial B_R(0) \times (0, \infty).$$

In fact, the solution of the two dimensional wave equation (1) can be expressed in terms of the spherical means via $p(z, t) = \partial_t \int_0^t r \, \mathbf{M}f(z, r)/\sqrt{t^2 - r^2} \mathrm{d}r$, see [38]. By using standard tools for solving Abel type equations, the last expression can be explicitly inverted, resulting in $(\mathbf{M}f)(z, r) = 2/\pi \int_0^r p(z, t)/\sqrt{r^2 - t^2} \mathrm{d}t$ (see [9, 27]). Hence any sparsifying transformation for the spherical means $\mathbf{M}f$ also yields a sparsifying transformation for the solution of the wave equation (1), and vice versa.

We found empirically, that $\partial_r r \, \mathbf{H}_r \, \partial_r \, \mathbf{M} \, f$ is sparse in the spatial direction for any function f that is the superposition of few indicator functions of regular domains. Here $\partial_r g$ denotes the derivative in the radial variable,

$$\mathbf{H}_r \, g(z, r) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{g(z, s)}{r - s} \mathrm{d}s \,, \quad \text{for } (z, r) \in \partial B_R(0) \times (0, \infty) \,,$$

the Hilbert transform of the function $g \colon \partial B_R(0) \times \mathbb{R} \to \mathbb{R}$ in the second component, and r the multiplication operator that maps the function $(z, r) \mapsto g(z, r)$ to the function $(z, r) \mapsto r g(z, r)$. Further, the spherical means $\mathbf{M} \, f \colon \partial B_R(0) \times \mathbb{R} \to \mathbb{R}$ are extended to an odd function in the second variable. For a simple radially symmetric phantom the sparsity of $\partial_r r \, \mathbf{H}_r \, \partial_r \, \mathbf{M} \, f$ is illustrated in Figure 3.2. Thus we can choose $\mathbf{T} = \partial_r r \, \mathbf{H}_r \, \partial_r$ as a sparsifying temporal transform for the spherical means.
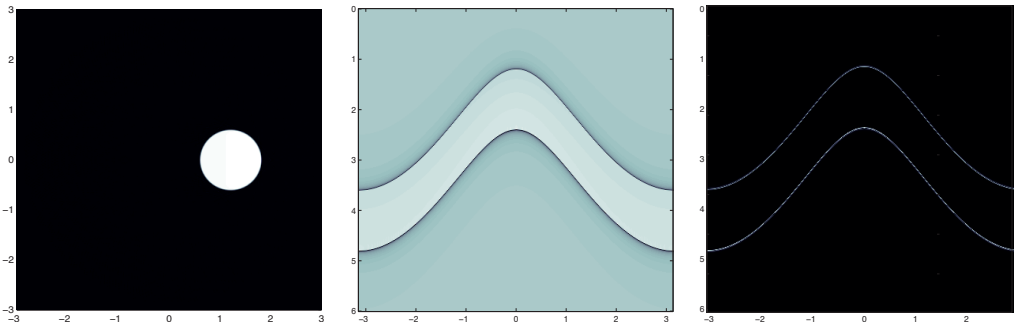

Figure 3.2: SPARSITY INDUCED BY $\partial_r(r \, \mathbf{H}_r \, \partial_r)$. The left image shows the simple disc-like phantom f (characteristic function of a disc), the middle image shows the filtered spherical means $r \, \mathbf{H}_r \, \partial_r \, \mathbf{M} \, f$ and the right image shows the sparse data $\partial_r(r \, \mathbf{H}_r \, \partial_r \, \mathbf{M} \, f)$.

The function $r \, \mathbf{H}_r \, \partial_r \, \mathbf{M} \, f$ also appears in the following formula for recovering a function from its spherical means derived in [23].

Theorem 3.1 (Exact reconstruction formula for spherical means)**.** *Suppose $f \in C^\infty(\mathbb{R}^2)$ is supported in the closure of $B_R(0)$. Then, for any $x \in B_R(0)$, we have*

$$f(x) = \frac{1}{2\pi R} \int_{\partial B_R(0)} (r \, \mathbf{H}_r \, \partial_r \, \mathbf{M} \, f)\,(z, |x - z|) \, \mathrm{d}s(z) \,. \tag{20}$$

Proof. See [23, Corollary 1.2]. $\qquad\square$

In the practical implementation the spherical means are given only for a discrete number of centers $z_j \in \partial B_R(0)$ yielding semi-discrete data similar to (2), (3). Formula (20) can easily be adapted to discrete or semi-discrete data yielding a filtered backprojection type reconstruction algorithm; compare [23, Section 4]. So if we can find the filtered spherical means

$$(r \, \mathbf{H}_r \, \partial_r \, \mathbf{M} \, f)\,(z_j, \cdot) \quad \text{for all detector locations } z_j \in \partial B_R(0) \tag{21}$$

we can obtain the desired reconstruction of f by applying the backprojection operator (the outer integration in (20)) to $r \, \mathbf{H}_r \, \partial_r \, \mathbf{M} \, f$.

3.4 Summary of the reconstruction procedure

In this section, we combine and summarize the compressed sensing scheme for photoacoustic tomography as described in the previous sections. Our proposed compressed sensing and sparse recovery strategy takes the following form.

(CS1) Create a matrix $\mathbf{A} \in \{0,1\}^{m \times N}$ as the adjacency matrix of a randomly selected left d-regular bipartite graph. That is, $\mathbf{A}$ is a random matrix consisting of zeros and ones only, with exactly d ones in each column.

(CS2) Perform m measurements, whereby in the i-th measurement pressure signals corresponding to the nonzero entries in i-th row of $\mathbf{A}$ are summed up, see Equations (4) and (15). This results in measurement data $\mathbf{A}\mathbf{p}(t) = \mathbf{y}(t) \in \mathbb{R}^m$ for any $t \in [0, 2\rho]$.

(CS3) Choose a transform $\mathbf{T}$ acting in the temporal direction, which sparsifies the pressure data $\mathbf{p}$ along the spatial direction; compare Equation (17).

(CS4) For any $t \in [0, 2\rho]$ and some given threshold η, perform ℓ^1-minimization (19) resulting in a sparse vector $\mathbf{q}_\star(t)$ satisfying $\|\mathbf{A}\mathbf{q}_\star(t) - \mathbf{T}\mathbf{y}(t)\|_1 \leq \eta$.

(CS5) Use $\mathbf{p}_\star(t) = \mathbf{T}^{-1}\mathbf{q}_\star(t)$ as the input for a standard PAT inversion algorithm for complete data, such as time reversal or filtered backprojection.

As we have seen in Subsection 2.2 the procedure (CS1)–(CS5) yields a close approximation to the original function f if the transformed data $\mathbf{T}\mathbf{p}(t)$ are sufficiently sparse in the spatial direction. The required sparsity level is hereby given by the expander-properties of the matrix $\mathbf{A}$. Note that for exact data and exactly sparse data, we can use the error threshold $\eta = 0$. In the more realistic scenario of noisy data and $\mathbf{T}\mathbf{p}(t)$ being only approximately sparse, we solve the optimization problem (19) to yield a near optimal solution with error level bounded by the noise level.

4 Numerical results

To support the theoretical examinations in the previous sections, in this section we present some simulations using the proposed compressed sensing method. We first present reconstruction results using simulated data and then show reconstruction results using experimental data.

4.1 Results for simulated data

As in Subsection 3.3, we work with the equivalent notion of the spherical means instead of directly working with the solution of the wave equation (1). In this case the compressed sensing measurements provide data

$$y_i(r) = \sum_{j \in J_i} m_j(r) \quad \text{for } i \in \{1, \ldots, m\} \text{ and } t \in [0, 2\rho], \tag{22}$$

where $m_j = (\mathbf{M}f)(z_j, \cdot)$ denote the spherical means collected at the j-the detector location z_j. We further denote by $\mathbf{m}(t) = (m_1(t), \ldots, m_N(t))^\mathsf{T}$ the vectors of unknown complete spherical means and by $\mathbf{y}(t) = (y_1(t), \ldots, y_m(t))^\mathsf{T}$ the vector of compressed sensing measurement data. Finally, we denote by $\mathbf{A} \in \{0,1\}^{m \times N}$ the compressed sensing matrix such that (22) can be rewritten in the form $\mathbf{A}\mathbf{m} = \mathbf{y}$.

As proposed in Subsection 3.3 we use $\mathbf{T} = \partial_r(r\,\mathbf{H}_r\,\partial_r)$ as a sparsifying transform for the spherical means. An approximation to $\partial_r r\,\mathbf{H}_r\,\partial_r\,\mathbf{M}f$ can be obtained from compressed sensing measurements in combination via ℓ^1-minimization. For the recovery of the original function from the completed measurements, we use one of the inversion formulas of [23] presented given in Theorem 3.1. Recall that this inversion formulas can be implemented by applying the circular back-projection to the filtered spherical means $r\,\mathbf{H}_r\,\partial_r\,\mathbf{M}f$.

In order to obtain an approximation to the data (21) from the sparse intermediate reconstruction $\partial_r(r \mathbf{H}_r \partial_r \mathbf{M} f)(\,\cdot\,, r)$, one has to perform one numerical integration along the second dimension after the ℓ^1-minimization process. We found that this numerical integration introduces artifacts in the reconstruction of $r \mathbf{H}_r \partial_r \mathbf{M} f$, required for application of the inversion formula (20), see the middle image in Figure 4.1. These artifacts also yield some undesired blurring in the final reconstruction of f; see the middle image in Figure 4.2.
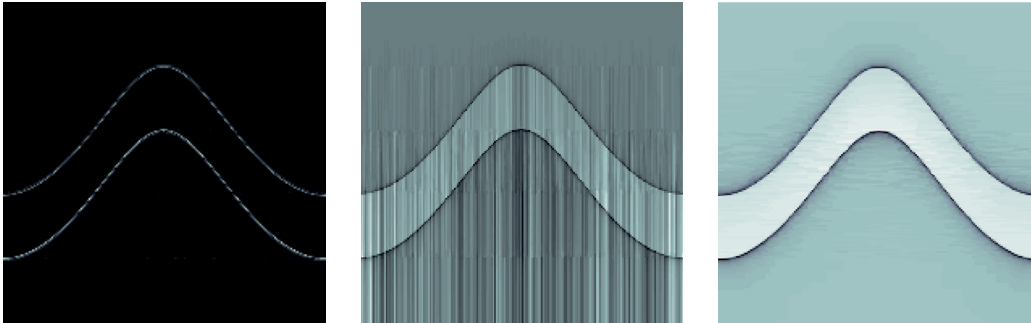
Figure 4.1: RECONSTRUCTION OF FILTERED SPHERICAL MEANS FOR $N = 200$ AND $m = 100$. Left: Reconstruction of $\partial_r r \mathbf{H}_r \partial_r \mathbf{M} f$ using ℓ^1-minimization. Center: Result of integrating the ℓ^1-reconstruction in the radial direction. Right: Result of directly recovering $r \mathbf{H}_r \partial_r \mathbf{M} f$ by TV-minimization.

In order to overcome the artifact introduced by the numerical integration, instead of applying an additional radial derivative to $r \mathbf{H}_r \partial_r$ to obtain sparsity in the spatial direction, we will apply one-dimensional total variation minimization (TV) for directly recovering $(r \mathbf{H}_r \partial_r \mathbf{M} f)(\,\cdot\,, r)$. Thereby we avoid performing numerical integration on the reconstructed sparse data. Furthermore, this yields much better results in terms of image quality of the final image f.

In our interpretation, this performance discrepancy is comparable to the difference between uniform and variable density samples for subsampled Fourier measurements. While [15] proves recovery of the discrete gradient, this does not carry over to the signal in a stable way – a refined analysis was required [50, 42]. Similarly, we expect that a refined analysis to be provided in subsequent work can help explain the quality gap between ℓ^1 and TV minimization that we observe in our scenario.

In order to approximately recover $r \mathbf{H}_r \partial_r \mathbf{M} f$ from the compressed sensing measurements, we perform, for any $r \in [0, 2\rho]$, one-dimensional discrete TV-minimization

$$\|\mathbf{A}\mathbf{q} - (r \mathbf{H}_r \partial_r \mathbf{y})(r)\|_2^2 + \lambda \|\mathbf{q}\|_{\mathrm{TV}} \to \min_{\mathbf{q} \in \mathbb{R}^N} . \tag{23}$$

Here $\|\mathbf{q}\|_{\mathrm{TV}} = 2\pi/N \sum_{j=1}^N |q_{j+1} - q_j|$ denotes the discrete total variation using the periodic extension $q_{N+1} := q_1$. The one-dimensional total variation minimization problem (23) can be efficiently solved using the fast iterative shrinkage thresholding algorithm (FISTA) of Beck and Teboulle [5]. The required proximal mapping for the total variation can be computed in $\mathcal{O}(N)$ operation counts by the tautstring algorithm (see [49, 20, 28]). The approximate solution of (23) therefore only requires $\mathcal{O}(N m N_{\mathrm{iter}})$ floating point operations, with N_{iter} denoting the number of iterations in the FISTA. Assuming the radial variable to be discretized using $\mathcal{O}(N)$ samples, the whole data completion procedure by (23) only requires $\mathcal{O}(N^2 m N_{\mathrm{iter}})$ operation counts. Since we found that fewer than 100 iterations in the FISTA are often sufficient for accurate results, the numerical effort of data completion is only a few times higher than that of standard reconstruction algorithms in PAT.

Figures 4.1 and 4.2 show results of simulation studies for a simple phantom, where the initial pressure distribution f is the characteristic function of a disc. For the compressed sensing
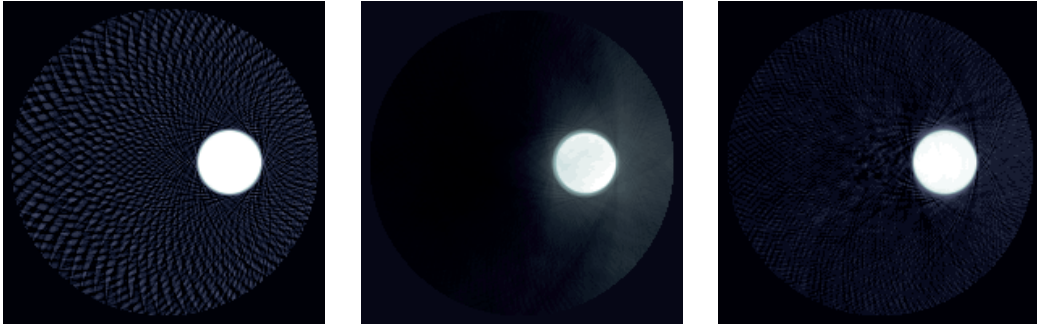
Figure 4.2: RECONSTRUCTION RESULTS FOR DISC-LIKE PHANTOM USING $N = 200$ AND $m = 100$. Left: Reconstruction from 100 standard measurements. Center: Compressed sensing reconstruction using ℓ^1-minimization. Right: Compressed sensing reconstruction using TV-minimization. The reconstruction from standard measurements contains under-sampling artifacts which are not present in the compressed sensing reconstructions. Further, the use of TV-minimization yields much less blurred results than the use of ℓ^1-minimization.

reconstruction, we used $m = 100$ random measurements instead of $N = 200$ standard point measurements. As one can see from the right image in Figure 4.2, the completed data $r\,\mathbf{H}_r\,\partial_r\,\mathbf{M}\,f$, for this simple phantom, is recovered almost perfectly from the compressed sensing measurements by means of TV-minimization. The reconstruction results in Figure 4.2 show that the combination of our compressed sensing approach with TV-minimization yields much better results than the use of ℓ^1-minimization. Fur comparison purpose, the left image in Figure 4.2 shows the reconstruction from 100 standard measurements. Observe that the use of $m = 100$ random measurements yields better result than the use of the 100 standard measurements, where artifacts due to spatial under-sampling are clearly visible.

4.2   Results for real measurement data

Experimental data have been acquired by scanning a single integrating line detector on a circular path around a phantom. A bristle with a diameter of $120\,\mu$m was formed to a knot and illuminated from two sides with pulses from a frequency doubled Nd:YAG laser with a wavelength of $532\,$nm. The radiant exposure for each side was below $7.5\,\mathrm{mJ/cm^2}$. Generated photoacoustic signals have been detected by a graded index polymer optical fiber being part of a Mach-Zehnder interferometer, as described in [30]. Ultrasonic signals have been demodulated using a self-made balanced photodetector, the high-frequency output of which was sampled with a 12-bit data acquisition device. A detailed explanation of the used photo-detector and electronics can be found in [3]. The polymer fiber detector has been scanned around the phantom on a circular path with a radius of $6\,$mm and photoacoustic signals have been acquired on 121 positions. The scanning curve was not closed, but had an opening angle of $\pi/2\,$rad. Hence photoacoustic signals have been acquired between $\pi/8\,$rad and $15\pi/8\,$rad.

Using these experimental data, compressed sensing data have been generated, where each detector location was used $d = 10$ times and $m = 60$ measurements are made in total. The reconstruction of the complete measurement data has been obtained by one-dimensional discrete TV-minimization (23) as suggested in the Section 4.1. The measured and the recovered complete pressure data are shown in the top row in Figure 4.3. The bottom row in Figure 4.3 shows the reconstruction results from 121 standard measurements (bottom left) and the reconstruction from 60 compressed sensing measurements (bottom center). Observe that there is only a small difference between the reconstruction results. This clearly demonstrates

the potential of our compressed sensing scheme (CS1)–(CS5) for decreasing the number of measurements while keeping image quality. For comparison purpose we also display the reconstruction using 60 standard measurement (bottom right). Compared to the compressed sensing reconstruction using the same number of measurements the use of standard measurements shows significantly more artifacts which are due to spatial under-sampling.
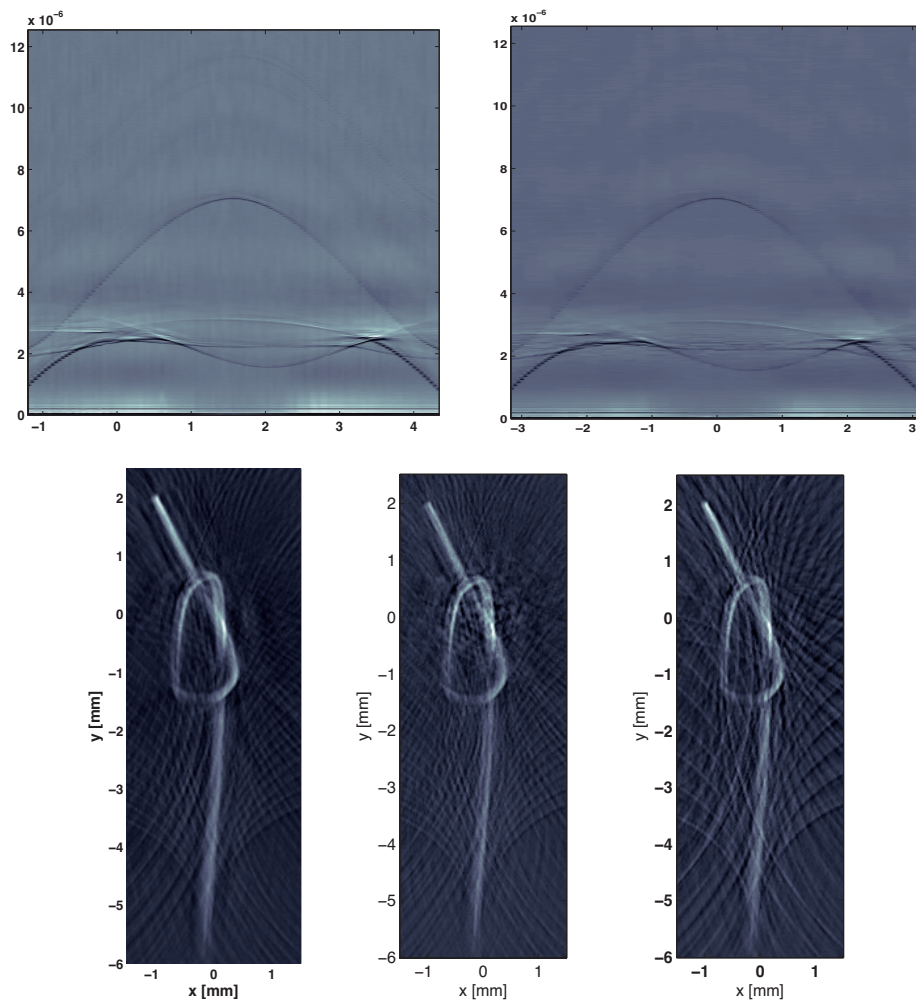

Figure 4.3: RECONSTRUCTION RESULTS FOR PAT MEASUREMENTS. Top left: Measured pressure data for $N = 121$ detector positions. Top right: Compressed sensing reconstruction of the full pressure data from $m = 60$ compressed measurements. Bottom left: Reconstruction from the full measurement data using $N = 121$ detector positions. Bottom center: Reconstruction from $m = 60$ compressed sensing measurements. Bottom right: Reconstruction from half of the measurement data using 60 detector positions.

5 Discussion and outlook

In this paper, we proposed a novel approach to compressive sampling for photoacoustic tomography using integrating line detectors providing recovery guarantees for suitable datasets. Instead of measuring pressure data p_j at any of the N individual line detectors, our approach uses m random combinations of p_j with $m \ll N$ as measurement data. The reconstruc-

tion strategy consists of first recovering the pressure values p_j for $j \in \{1, \ldots, N\}$ and then applying a standard PAT reconstruction for obtaining the final photoacoustic image. For recovering the individual pressure data, we propose to apply a sparsifying transformation that acts in the temporal variable and makes the data sparse in the angular component. After applying such a transform the complete pressure data is recovered by solving a set of one dimensional ℓ^1-minimization problems. This decomposition also makes our reconstruction algorithm numerically efficient.

Although we focused on PAT using integrating line detectors, we emphasize that a similar framework can be developed for standard PAT based on the three dimensional wave equation and a two dimensional array of point-like detectors. In such a situation, finding a sparsifying transform is even simpler. Recalling that N-shaped profile of the thermoacoustic pressure signal induced by the characteristic function of a ball suggests to use $\mathbf{p} \mapsto \partial_t^2 \mathbf{p}$ as a sparsifying transform.

Note that the recovery guarantees in this paper crucially depend on the choice of an appropriate sparsifying transform. In Subsection 3.3 we proposed a candidate for such a transformation that works well in our numerical examples. A more theoretical study of such sparsifying transforms and the resulting recovery guarantees for simple (piecewise constant) phantoms is postponed to further research. In this context, we will also investigate the use of different sparsifying temporal transforms, such as the 1D wavelet transform in the temporal direction. Further research includes using a fixed number of detectors in each measurement process. This requires novel results for right k-regular expander graphs and compressive sampling.

Acknowledgements

References

[1] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constr. Approx.*, 28(3):253–263, 2008.

[2] J. Bauer-Marschallinger, K. Felbermayer, K.-D. Bouchal, I. A. Veres, H. Grün, P. Burgholzer, and T. Berer. Photoacoustic projection imaging using a 64-channel fiber optic detector array. In *Proc. SPIE*, volume 9323, 2015.

[3] J. Bauer-Marschallinger, K. Felbermayer, A. Hochreiner, H. Grün, G. Paltauf, P. Burgholzer, and T. Berer. Low-cost parallelization of optical fiber based detectors for photoacoustic imaging. In *Proc. SPIE*, volume 8581, pages 85812M–85812M–8, 2013.

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

[5] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, 2(1):183–202, 2009.

[6] T. Berer, I. A. Veres, H. Grün, J. Bauer-Marschallinger, K. Felbermayer, and P. Burgholzer. Characterization of broadband fiber optic line detectors for photoacoustic tomography. *J. Biophotonics*, 5(7):518–528, 2012.

[7] R. Berinde, A. C. Gilbert, P. Indyk, H. Karloff, and M. J. Strauss. Combining geometry and combinatorics: A unified approach to sparse signal recovery. In *46th Annual Allerton Conference on Communication, Control, and Computing, 2008*, pages 798–805, 2008.

[8] R. Berinde, P. Indyk, and M. Ruzic. Practical near-optimal sparse recovery in the l_1 norm. In *46th Annual Allerton Conference on Communication, Control, and Computing*, pages 198–205, 2008.

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

[10] P. Burgholzer, C. Hofer, G. Paltauf, M. Haltmeier, and O. Scherzer. Thermoacoustic tomography with integrating area and line detectors. *IEEE Trans. Ultrason., Ferroeletr., Freq. Control*, 52(9):1577–1583, 2005.

[11] 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.

[12] T. Cai and A. Zhang. Sparse representation of a polytope and recovery of sparse signals and low-rank matrices. *IEEE Trans. Inf. Theory*, 60:122 – 132, 2014.

[13] E. Candes and J. Romberg. Sparsity and incoherence in compressive sampling. *Inverse Probl.*, 23(3):969, 2007.

[14] E. J. Candes. The restricted isometry property and its implications for compressed sensing. *C. R. Math. Acad. Sci. Paris*, 346(9):589–592, 2008.

[15] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory*, 52(2):489–509, 2006.

[16] E. J. Candès, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.*, 59(8):1207–1223, 2006.

[17] E. J. Candès and T. Tao. Near-optimal signal recovery from random projections: universal encoding strategies? *IEEE Trans. Inf. Theory*, 52(12), 2006.

[18] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numer. Math.*, 76(2):167–188, 1997.

[19] G.-H. Chen, J. Tang, and S. Leng. Prior image constrained compressed sensing (piccs): a method to accurately reconstruct dynamic ct images from highly undersampled projection data sets. *Med. Phys.*, 35(2):660–663, 2008.

[20] P. L. Davies and A. Kovac. Local extremes, runs, strings and multiresolution. *Ann. Stat.*, pages 1–48, 2001.

[21] D. L. Donoho. Compressed sensing. *IEEE Trans. Inf. Theory*, 52(4):1289–1306, 2006.

[22] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Process. Mag.*, 25(2):83–91, 2008.

[23] 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.

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

[25] S. Foucart and H. Rauhut. *A mathematical introduction to compressive sensing*. Springer, 2013.

[26] A. Gilbert and P. Indyk. Sparse recovery using sparse matrices. *Proc. IEEE*, 98(6):937–947, 2010.

[27] R. Gorenflo and S. Vessella. *Abel integral equations*, volume 1461 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1991. Analysis and applications.

[28] M. Grasmair and A. Obereder. Generalizations of the taut string method. *Numer. Funct. Anal. Optim.*, 29(3-4):346–361, 2008.

[29] S. Gratt, R. Nuster, G. Wurzinger, M. Bugl, and G. Paltauf. 64-line-sensor array: fast imaging system for photoacoustic tomography. *Proc. SPIE*, 8943:894365–894365–6, 2014.

[30] 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):021306–021306–8, 2010.

[31] Z. Guo, C. Li, L. Song, and L. V. Wang. Compressed sensing in photoacoustic tomography in vivo. *J. Biomed. Opt.*, 15(2):021311–021311, 2010.

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

[33] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 43(4):439–561, 2006.

[34] Y. Hristova, P. Kuchment, and L. Nguyen. Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media. *Inverse Probl.*, 24(5):055006 (25pp), 2008.

[35] N. Huynh, E. Zhang, M. Betcke, S. Arridge, P. Beard, and B. Cox. Patterned interrogation scheme for compressed sensing photoacoustic imaging using a fabry perot planar sensor. In *Proc. SPIE*, pages 894327–894327, 2014.

[36] P. Indyk and M. Ruzic. Near-optimal sparse recovery in the l_1 norm. In *49th Annual IEEE Symposium on Foundations of Computer Science*, pages 199–207, 2008.

[37] S. Jafarpour, W. Xu, B. Hassibi, and R. Calderbank. Efficient and robust compressed sensing using optimized expander graphs. *IEEE Trans. Inf. Theory*, 55(9):4299–4308, 2009.

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

[39] M. Klawe. Limitations on explicit constructions of expanding graphs. *SIAM J. Comput.*, 13:156–166, 1984.

[40] F. Krahmer, S. Mendelson, and H. Rauhut. Suprema of chaos processes and the restricted isometry property. *Comm. Pure Appl. Math.*, 67(11):1877–1904, 2014.

[41] F. Krahmer and R. Ward. New and improved Johnson-Lindenstrauss embeddings via the restricted isometry property. *SIAM J. Math. Anal.*, 43(3):1269–1281, 2011.

[42] F. Krahmer and R. Ward. Stable and robust sampling strategies for compressive imaging. *IEEE Trans. Image Proc.*, 23(2):612–622, 2014.

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

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

[45] A. Lubotzky. *Discrete groups, expanding graphs and invariant measures*, volume 125. Springer, 1994.

[46] A. Lubotzky. Expander graphs in pure and applied mathematics. *Bull. Amer. Math. Soc.*, 49(1):113–162, 2012.

[47] M. Lustig, D. Donoho, and J. M. Pauly. Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic resonance in medicine*, 58(6):1182–1195, 2007.

[48] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly. Compressed sensing mri. *IEEE Sig. Proc. Mag.*, 25(2):72–82, 2008.

[49] E. Mammen, S. van de Geer, et al. Locally adaptive regression splines. *Ann. Stat.*, 25(1):387–413, 1997.

[50] D. Needell and R. Ward. Stable image reconstruction using total variation minimization. *SIAM J. Imaging Sci.*, 6(2):1035–1058, 2013.

[51] R. Nuster, M. Holotta, C. Kremser, H. Grossauer, P. Burgholzer, and G. Paltauf. Photoacoustic microtomography using optical interferometric detection. *J. Biomed. Optics*, 15(2):021307–021307–6, 2010.

[52] 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.

[53] G. Paltauf, R. Nuster, M. Haltmeier, and P. Burgholzer. Photoacoustic tomography using a Mach-Zehnder interferometer as an acoustic line detector. *Appl. Opt.*, 46(16):3352–3358, 2007.

[54] G. Paltauf, R. Nuster, M. Haltmeier, and P. Burgholzer. Photoacoustic tomography with integrating area and line detectors. In *Photoacoustic imaging and spectroscopy*, chapter 20, pages 251–263. CRC Press, 2009.

[55] J. Provost and F. Lesage. The application of compressed sensing for photo-acoustic tomography. *IEEE Trans. Med. Imag.*, 28(4):585–594, 2009.

[56] H. Rauhut. Random sampling of sparse trigonometric polynomials. *Appl. Comp. Harm. Anal.*, 22(1):16–42, 2007.

[57] H. Rauhut, J. Romberg, and J. Tropp. Restricted isometries for partial random circulant matrices. *Appl. Comput. Harmon. Anal.*, 32(2):242–254, 2012.

[58] M. Rudelson and R. Vershynin. On sparse reconstruction from fourier and gaussian measurements. *Communications on Pure and Applied Mathematics*, 61(8):1025–1045, 2008.

[59] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1–4):259–268, 1992.

[60] B. E. Treeby and B. T. Cox. k-wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave-fields. *J. Biomed. Opt.*, 15:021314, 2010.

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

[62] M. Xu and L. V. Wang. Universal back-projection algorithm for photoacoustic computed tomography. *Phys. Rev. E*, 71(1):0167061–0167067, 2005.

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

[64] W. Xu and B. Hassibi. Efficient compressive sensing with deterministic guarantees using expander graphs. In *IEEE Information Theory Workshop*, pages 414–419, 2007.