Acoustic Wave Field Reconstruction From Compressed Measurements With Application in Photoacoustic Tomography

Marta M. Betcke, Ben T. Cox, Nam Huynh, Edward Z. Zhang, Paul C. Beard, and Simon R. Arridge

***Abstract*—We present a method for the recovery of compressively sensed acoustic fields using patterned, instead of point-by-point, detection. From a limited number of such compressed measurements, we propose to reconstruct the field on the sensor plane in each time step independently assuming its sparsity in a Curvelet frame. A modification of the Curvelet frame is proposed to account for the smoothing effects of data acquisition and motivated by a frequency domain model for photoacoustic tomography. An ADMM type algorithm, split augmented Lagrangian shrinkage algorithm, is used to recover the pointwise data in each individual time step from the patterned measurements. For photoacoustic applications, the photoacoustic image of the initial pressure is reconstructed using time reversal in k-Wave Toolbox.**

***Index Terms*—ADMM methods, compressed sensing, curvelet frame, L1 minimization, photoacoustic tomography.**

I. INTRODUCTION

COMPRESSED SENSING (CS) is a new measurement paradigm, which allows for the reconstruction of sparse signals sampled at sub-Nyquist rates. Nowadays, it is common understanding that many digital signals and images admit an adequate representation with far fewer coefficients than their actual length. This phenomena is known as compressibility and it has been a driving force in many image processing applications, most notably the image compression algorithms JPEG and its successor JPEG 2000. CS emerged from the realization that the signals could in fact be acquired directly in their compressed form instead of sampling the signal at Nyquist rate and then compressing it and while doing so discarding most of the laboriously obtained coefficients.

Since the seminal works by Donoho [1] and Càndes, Romberg and Tao [2], [3], there has been an explosion of results in the field and many applications have been suggested starting with the prototype single pixel camera [4].

Here we propose a new application of CS in acoustic wave field sensing. In [5] the authors proved that the acoustic field is almost optimally sparse in Curvelet frame. As the Curvelets essentially describe the wave front sets their propagation is well approximated through geometrical optics (high frequency asymptotic solution to the wave equation). The wave front sets, and hence the Curvelets, propagate along the geometrical rays which are projections on $\mathbb{R}^d$ of phase space solutions of the corresponding Hamilton-Jacobi equation resulting from the high frequency asymptotic. Motivated by this result we investigate the Curvelet representation of the cross-section of the wave field by the planar ultrasound sensor. While the arguments in [5] do not directly apply to this situation, the planar cross-section through the acoustic wave front constitutes a singularity along a smooth curve for which Curvelets have been demonstrated to be a nearly optimal representation [6].

A. *Contribution*

In this paper we focus on the reconstruction problem via data recovery for a novel way of interrogating the high resolution ultrasound sensor using patterns instead of the more conventional sequential point-by-point interrogation. An example of such a system using a Single-Pixel Optical Camera (SPOC) was presented in [7]. The theory of CS predicts that substantially fewer such measurements, of the order $k \log n$, need to be taken in order to capture a signal of length n and sparsity k leading to a substantial reduction of the acquisition time. We propose from such measurements to recover the pressure at the detector at each time step independently using CS recovery algorithms. We discuss the motivation for using a Curvelet tranform as the sparsifying transformation for the acoustic field at the detector and we propose its modification: a low-frequency Curvelet tranform tailored to the frequency range of the acoustic field on the detector. We investigate the appropriate choice of interrogation patterns and recovery algorithm for this problem. The proposed techniques can be used in various applications ranging from ultrasonic field mapping to photoacoustic tomography (PAT). In this work we focus on the latter to illustrate our method.

For PAT applications, the immediate benefit of recovering data independently at each time step is the decoupling of the CS reconstruction from the acoustic inversion, allowing for the

---

Manuscript received October 10, 2016; revised March 1, 2017 and April 29, 2017; accepted April 30, 2017. Date of publication May 18, 2017; date of current version November 6, 2017. This work was supported by the Engineering and Physical Sciences Research Council, U.K. (EP/K009745/1). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Rebecca Willett. *(Corresponding author: Marta M. Betcke.)*

M. M. Betcke and S. R. Arridge are with the Department of Computer Science, University College London, London WC1E 6BT, U.K. (e-mail: m.betcke@ucl.ac.uk; s.arridge@cs.ucl.ac.uk).

B. T. Cox, N. Huynh, E. Z. Zhang, and P. C. Beard are with the Department of Medical Physics, University College London, London WC1E 6BT, U.K. (e-mail: b.cox@ucl.ac.uk; n.huynh@ucl.ac.uk; e.zhang@ucl.ac.uk; paul.beard@ucl.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCI.2017.2706029

recovered time series to be input into any reconstruction algorithm for PAT for which software is readily available.

B. Related Work

For photoacoustic image reconstruction from compressed or subsampled measurements, one and two step approaches have been investigated in the literature. In one step approaches the initial pressure is directly recovered from the compressed measurements. In [8] the authors propose such an approach for a 2D problem with measurements limited in angle and frequency. It was numerically investigated using an analytic approximate inversion formula and a number of sparse representations including Wavelets and Curvelets. An approach for 3D imaging using a translatable circular detection array and Wavelets as a sparse representation was presented in [9]. A generic variational approach for 2D and 3D PAT reconstruction based on an algebraic adjoint was first discussed in [10]. Two variational approaches based on the analytical adjoint were recently proposed, one in a FEM-BEM setting [11] and the other in a $\mathbf{k}$-space setting with an efficient $\mathbf{k}$-Wave implementation [12], [13].

In contrast, in the context of PAT, the present paper is about a two step method. We propose to first recover the photoacoustic data in every time step independently from pattern measurements using sparsity of the data in Curvelet basis, and subsequently to reconstruct it using standard PAT reconstruction methods. Recently, another two stage approach has been suggested in [14], where the authors explore the temporal sparsity of the data by means of a custom made transform in time. From the sparse pressure data subsequently the initial pressure is reconstructed with the universal back-projection formula [15], [16].

C. Outline

The remainder of this paper is organized as follows. In Section II we introduce the forward and inverse problem for PAT and methods for their solution. Section III briefly recapitulates the theory of compressed sensing. An appropriate multiscale representation of the time series PAT data is considered in Section IV, where we derive the frequency model of sensor data and propose a modified version of Curvelet transform tailored to the range of frequencies of the acoustic field. In Section V we discuss specific issues arising when compressively sensing photoacoustic data using patterned interrogation of optical ultrasound detector. We briefly describe the Single-Pixel Optical Camera based PAT scanner. We consider the challenges for Curvelet transforms for approximation of sensor data over the time series. We discuss choice of the interrogation patterns and the algorithm for recovery of the sensor data. In Section VI we present recovery results for the optical sensor data and the final PAT image reconstruction from both simulated and real data.

II. PHOTOACOUSTIC TOMOGRAPHY

Photoacoustic tomography (PAT) is an example of a wider range of hybrid imaging techniques, in which contrast induced by one type of wave is read out by another wave. In this way, both high contrast and high resolution can be simultaneously achieved, which is often difficult with conventional imaging techniques that usually provide either one or the other, but not both. PAT is an emerging biomedical imaging modality with both pre-clinical and clinical applications that can provide complementary information to established imaging techniques [17]–[21]

Many PAT applications require a high resolution, three dimensional image e.g. an image of capillaries of a few tens of microns diameter in a cm sized imaging region. Such highly resolved imaging requires an ultrasound sensor array of tens of thousands of pixels. In one such PAT system [22], the sensor is a Fabry Perot (FP) interferometer interrogated by a laser whose focus is moved to form a raster scan of the desired resolution. For sequential sampling, such as this, the ultimate limit to the data acquisition rate (the rate at which time series are collected) is the propagation time for sound to cross the specimen, e.g. it would take 10 μs for the signal to reach the detector from 15 mm depth, resulting in 100 kHz acquisition rate (not to be confused with the sampling rate, which might be as high as 100 MHz). No sequential scanner is close to approaching this limit, making a sequential acquisition a major practical limitation for high resolution 3D PAT. For *in vivo* applications the required acquisition times at currently achievable rates are typically a few minutes, not only resulting in motion artefacts, but limiting studies to phenomena on such long timescales.

The principle involved in PAT is to send a short (ns) pulse of near-infrared or visible light into tissue, whereupon absorption of the photons e.g. by haemoglobin molecules, generates a small local increase in pressure which propagates to the surface as a broadband, ultrasonic pulse. If the amplitude of this signal is recorded over an array of sensors at the tissue surface, an image reconstruction algorithm can be used to estimate the original 3D pressure increase due to optical absorption; this is the photoacoustic image, p_0.

Mathematically, under an assumption of free space propagation the photoacoustic forward problem is modelled as an initial value problem for the wave equation [23]

$$\frac{1}{c^2(\mathbf{x})}\frac{\partial^2 p(\mathbf{x},t)}{\partial t^2} = \rho_0(\mathbf{x})\nabla\cdot\left(\frac{1}{\rho_0(\mathbf{x})}\nabla\right)p(\mathbf{x},t), \qquad \mathbf{x}\in\mathbb{R}^d,\ t\in(0,T), \tag{1a}$$

$$p(\mathbf{x},0) = p_0(\mathbf{x}), \tag{1b}$$

$$\frac{\partial}{\partial t}p(\mathbf{x},0) = 0, \tag{1c}$$

where $p(\mathbf{x},t)$ denotes the time dependent acoustic pressure in $\mathbb{R}^d\times(0,T), d=2,3, p_0(\mathbf{x})$ its initial value and $c(\mathbf{x})$ and $\rho_0(\mathbf{x})$ are the ambient speed of sound and density, respectively.

The photoacoustic inverse problem is to recover this initial pressure $p_0(\mathbf{x}), \mathbf{x}\in\Omega$ compactly supported in the region of interest Ω from a time series measurement $g(\mathbf{x},t) = p(\mathbf{x},t), \mathbf{x}\in\mathcal{S}, t\in(0,T)$ on the surface $\mathcal{S}$ (e.g. boundary of Ω) and it amounts to a solution of the following initial value problem with surface data [24]

$$\frac{1}{c_0^2(\mathbf{x})}\frac{\partial^2 p(\mathbf{x},t)}{\partial t^2} = \rho_0(\mathbf{x})\nabla\cdot\left(\frac{1}{\rho_0(\mathbf{x})}\nabla\right)p(\mathbf{x},t), \quad \mathbf{x}\in\mathbb{R}^d, t\in(0,T), \tag{2a}$$

$$p(\mathbf{x},0) = \mathbf{0}, \tag{2b}$$

$$\frac{\partial}{\partial t}p(\mathbf{x},0) = 0, \tag{2c}$$

$$p(\mathbf{x},t) = g(\mathbf{x},T-t), \quad \mathbf{x}\in\mathcal{S},\ t\in(0,T), \tag{2d}$$

with PAT data fed backwards in time as surface values (also referred to as *time reversal*). The formulation (2) holds exactly in 3D for non-trapping smooth sound speed $c(x)$ if T has been chosen large enough so that $g(\mathbf{x}_{\mathcal{S}},t)=0,\ t\geq T$ and the wave has left the domain Ω. Furthermore, assuming that the measurement surface $\mathcal{S}$ surrounds the region of interest Ω containing the support of the initial pressure p_0, problem (2) has a unique solution. The condition on $\mathcal{S}$ to surround p_0 can be relaxed under additional assumptions on $\mathcal{S}$ and Ω and smoothess of initial pressure $p_0\in H_0^1(\Omega)$; see Theorem 10 in [25] and the citations within. The reconstruction problem is stable (meaning that small errors in the data cause small errors in the reconstruction) for closed surface $\mathcal{S}$ surrounding the region of interest Ω and non-trapping sound speeds, see e.g. [25]. If $\mathcal{S}$ does not fully surround Ω or the sound speed is trapping, some singularities may become invisible in the data leading to the lost of Hölder stability in any Sobolev space [26]. All visible singularities can be reconstructed with Lipschitz stability in appropriate spaces [27].

In a real experiment, we can only acquire a discrete (both in time and space) subset of time series measurements, and frequently it is not possible to acquire measurements on a surface $\mathcal{S}$ surrounding the object. An example of a popular sensor violating this assumption is a planar sensor. In practice, such sensor will have finite size, resulting in invisibility of some interfaces and in turn in artefacts in the reconstructed image.

We should mention that other approaches to image reconstruction exist. An overview of methods for the case when $\mathcal{S}$ is a surface surrounding Ω can be found e.g. in [25]. In this paper we directly solve the time-reversal problem (2) using the pseudospectral method implemented in the $\mathbf{k}$-Wave Toolbox [28], which is an efficient numerical scheme for solving the wave equation in domains with heterogeneous acoustic properties, and is exact in the case of homogeneous media. Furthermore, the methodology proposed here is tailored to high resolution detectors which are planar and so we assume $\mathcal{S}$ to be a finite rectangular section of the xy-plane

$$\mathcal{S} = \{(x,y,z):\ |x|\leq x_d/2,\ |y|\leq y_d/2,\ z=0\}$$

and solve the resulting initial value problem (2) with data on a finite sensor $\mathcal{S}$. The rectangular planar detector shape allows us to use any sparsifying transform derived for natural images.

III. Compressed Sensing

Let Ψ be an isometric sparsifying transform

$$\Psi:\mathbb{R}^n\to\mathbb{R}^N,$$

resulting in possibly an overdetermined representation $N\geq n$ and $\Psi^\dagger=\Psi^{\mathrm{T}}$ denote its left inverse, where the equality is due to isometry. With $f\in\mathbb{R}^N$ we denote the transformation of the original signal $g\in\mathbb{R}^n$,

$$f=\Psi g. \tag{3}$$

In compressed sensing the signal $g\in\mathbb{R}^n$ is projected on a series of sensing vectors $\phi_i\in\mathbb{R}^n, i\in 1,\ldots,m$, where $m\ll n$, yielding a vector $b\in\mathbb{R}^m$ of compressive measurements

$$b=\Phi g+e \quad \text{with} \quad \Phi=[\phi_1,\phi_2,\ldots,\phi_m]^{\mathrm{T}}, \tag{4}$$

where e is the measurement noise, $\|e\|_2\leq\varepsilon$. The following recovery result guarantees that the original signal g, can be robustly recovered from compressed measurements (4) via solution of the minimization problem for f [3]

$$\min_{f\in\mathbb{R}^N}\|f\|_1, \quad \text{s.t.} \quad \|\Phi\Psi^\dagger f-b\|\leq\varepsilon. \tag{5}$$

Theorem RR Robust recovery [29], [30]: Let δ_k be the isometry constant of $\Phi\Psi^\dagger$ defined as a smallest positive number such that

$$(1-\delta_k)\|f\|_2^2\leq\|\Phi\Psi^\dagger f\|_2^2\leq(1+\delta_k)\|f\|_2^2 \tag{6}$$

holds for all k-sparse vectors f i.e. vectors with at most k non-zero elements.

If $\delta_{2k}<\sqrt{2}-1$ (relaxed to $\delta_{2k}<\frac{2}{3+\sqrt{7/4}}$ in [30]) then the error of the solution of (5), f^*, is bounded as follows

$$\|f-f^*\|_2\leq C_1\varepsilon+C_2\frac{\|f-f^k\|_1}{\sqrt{k}}, \tag{7}$$

where f^k denotes best k-term approximation, obtained from f selecting its k largest in magnitude coefficients, and C_1, C_2 are constants dependent only on δ_{2k}.

The robust recovery Theorem RR holds for any vector f, however the error norm $\|f-f^k\|_1$ of the k-term approximation is only small for k-sparse or compressible vector f (i.e. with fast enough decaying magnitude of the coefficients). Furthermore, for k-sparse vectors we have $\|f-f^k\|_1/\sqrt{k}\leq\|f-f^k\|_2$ and hence the the bound can be expressed in L_2 norm

$$\|f-f^*\|_2\leq C_1\varepsilon+C_2\|f-f^k\|_2.$$

IV. Multiscale Representation of Time Series Data

A. Frequency Model of Sensor Data

For constant sound speed and density $c(\mathbf{x})=c_0, \rho(\mathbf{x})=\rho_0$, the wave equation (1) becomes homogeneous and it admits an analytic solution in Fourier domain [31]

$$p(\mathbf{k},t)=\cos(c_0|\mathbf{k}|t)p_0(\mathbf{k}), \tag{8}$$

where $\mathbf{k}$ is the frequency domain wave vector. Equation (8) formally allows us to calculate the PAT data time series

$$\begin{aligned} p(\mathbf{x}_\mathcal{S}, t) &= \mathcal{F}^{-1}(p(\mathbf{k}, t))\big|_{\mathbf{x}=\mathbf{x}_\mathcal{S}} \\ &= \mathcal{F}^{-1}\left(\cos(c_0|\mathbf{k}|t)p_0(\mathbf{k})\right)\big|_{\mathbf{x}=\mathbf{x}_\mathcal{S}} \\ &= 2\int_0^\infty \mathcal{F}_\parallel^{-1}\left(\cos(c_0|\mathbf{k}|t)p_0(\mathbf{k})\right) d\mathbf{k}_\perp, \end{aligned} \quad (9)$$

where we used the tensor decomposition $\mathcal{F} = \mathcal{F}_\parallel \mathcal{F}_\perp$ of the 3D Fourier transform in, and orthogonal to, the sensor plane, and the fact that for $\mathbf{x}_\mathcal{S}$ we have $z = 0$ and hence $e^{i\mathbf{k}_\perp z} = 1$.

Since $p_0(\mathbf{x})$ is a real valued function, $p_0(\mathbf{k})$ and consequently $p(\mathbf{k}, t)$ are symplectic in $\mathbf{k}$ i.e. $p_0(\mathbf{k}) = p_0(-\mathbf{k})^*$. Furthermore, $p(\mathbf{k}, t)$ as given by (8) is even in t. As our domain Ω is positioned in $z \geq 0$ to restore uniqueness we assume $p_0(\mathbf{x})$ to be even w.r.t. $z = 0$ plane. Consequently, the Fourier transform in the z as well as the t direction is equal to the cosine transform and can be calculated integrating from 0 to ∞; thus in what follows the two are used interchangeably.

In experiments however, $p(\mathbf{x}_\mathcal{S}, t)$ is available to us only indirectly, through measurements. To account for limitations of physical equipment, we introduce a degradation operator, $\mathcal{D}$. Note, that here $\mathcal{D}$ is not a measurement operator, but it models the effect of the finite size and temporal response of the measurement system which reads the acoustic time series. In other words, $g(\mathbf{x}_\mathcal{S}, t) = \mathcal{D}p(\mathbf{x}_\mathcal{S}, t)$ describes the filtering of the acoustic pressure $p(\mathbf{x}_\mathcal{S}, t)$ by the physical presence of the sensor before the sensing vector, ϕ_i is applied to it to collect the compressive measurement, b_i.

In this work we assume that $\mathcal{D}$ is a band limiting spatial and temporal filtering operator acting on the time dependent pressure on optical sensor, which for simplicity we describe in the frequency domain

$$\mathcal{D}[p](\mathbf{k}_\mathcal{S}, \omega) = w_t(\omega) w_\parallel(\mathbf{k}_\mathcal{S}) p(\mathbf{k}_\mathcal{S}, \omega), \quad (10)$$

where $w_\parallel(\mathbf{k}_\mathcal{S})$ and $w_t(\omega)$ are some frequency window functions on the sensor and in time.

Taking Fourier transform of (9) in both variables we obtain

$$\begin{aligned} p(\mathbf{k}_\mathcal{S}, \omega) &= \mathcal{F}_\mathrm{t}\left(\mathcal{F}_\parallel\left(p(\mathbf{x}_\mathcal{S}, t)\right)\right) \\ &= \mathcal{F}_\mathrm{t}\left(2\int_0^\infty \mathcal{F}_\parallel\left(\mathcal{F}_\parallel^{-1}\left(\cos(c_0|\mathbf{k}|t)p_0(\mathbf{k})\right)\right) d\mathbf{k}_\perp\right) \\ &= \mathcal{F}_\mathrm{t}\left(2\int_0^\infty \frac{\omega/c_0^2 \cos(\omega t)}{\sqrt{(\omega/c_0)^2 - |\mathbf{k}_\mathcal{S}|^2}}\right. \\ &\qquad \left. p_0(\mathbf{k}_\mathcal{S}, \sqrt{(\omega/c_0)^2 - |\mathbf{k}_\mathcal{S}|^2}) d\omega\right) \\ &= \frac{\omega/c_0^2}{\sqrt{(\omega/c_0)^2 - |\mathbf{k}_\mathcal{S}|^2}} p_0(\mathbf{k}_\mathcal{S}, \sqrt{(\omega/c_0)^2 - |\mathbf{k}_\mathcal{S}^2|}), \end{aligned} \quad (11)$$

where changing the integration variable to ω, $\omega/c > |\mathbf{k}_\mathcal{S}|$, in the third line allowed us to interpret the integral as inverse cosine transform in ω. Equation (11) connects the Fourier transform of the pressure time series on the detector with the Fourier transform of initial pressure and is the basis of the reconstruction formula derived in [32], [33]. From (11) it is obvious that application of $\mathcal{D}[p](\mathbf{k}_\mathcal{S}, \omega)$ corresponds to the application of a filter window $w_t(c_0|\mathbf{k}|) w_\parallel(\mathbf{k}_\mathcal{S})$ to $p_0(\mathbf{k})$

$$\begin{aligned} \mathcal{D}[p](\mathbf{k}_\mathcal{S}, \omega) &= \tfrac{w_t(\omega) w_\parallel(\mathbf{k}_\mathcal{S}) \omega/c_0^2}{\sqrt{(\omega/c_0)^2 - |\mathbf{k}_\mathcal{S}|^2}} p_0(\mathbf{k}_\mathcal{S}, \sqrt{(\omega/c_0)^2 - |\mathbf{k}_\mathcal{S}|^2}) \\ &= w_t(c_0|\mathbf{k}|) w_\parallel(\mathbf{k}_\mathcal{S}) \tfrac{|\mathbf{k}|}{|\mathbf{k}_\perp| c_0} p_0(\mathbf{k}) =: \mathcal{D}_\Omega[p_0](\mathbf{k}) \end{aligned}$$

for scalar $\mathbf{k}_\perp > 0$. Thus only a smoothed initial pressure can be recovered from the data. Conversely, when simulating PAT data, the initial pressure can be smoothed with $w_t(c_0|\mathbf{k}|) w_\parallel(\mathbf{k}_\mathcal{S})$ before forward propagation instead of smoothing the sensor data with $\mathcal{D}$, which can be useful to eliminate Gibbs phenomena in $\mathbf{k}$-space methods; see also [34], [35] where a Blackman window was applied to p_0.

B. Curvelets

In this work we suggest using Curvelets to represent each of the measured time series data $g(\mathbf{x}_\mathcal{S}, t)$. For the planar sensor $\mathcal{S}$ this corresponds to a planar cross-section of the wave field $p(\mathbf{x}, t)$ at a given time t (which we observe over a finite section of the plane, $\mathcal{S}$). Due to PAT forward problem being an initial value problem, the corresponding wave field $p(\mathbf{x}, t)$ is essentially smooth away from the p_0-shaped (with corners smoothed) wave front and the same holds for their planar cross-sections. As the time evolves $p(\mathbf{x}, t)$ develops overlapping wave fronts. Those however can be treated as superposition and hence we only need to be able to represent an individual wave front.

In what follows we assume $g(\mathbf{x}_\mathcal{S}, t)$ to be sampled on an $n_1 \times n_2$ grid on the sensor with some steps $h_1 \leq x_d/n_1, h_2 \leq y_d/n_2$, resulting in an image $g^t[i_1, i_2]$, $i_1 = 1, \dots, n_1$ $i_2 = 1, \dots, n_2$, where $(\cdot)$ denotes function evaluation and $[\cdot]$ array indexing.

Curvelets [6], [36] are a multiscale pyramid with many directions and positions at each scale. Curvelets obey parabolic scaling, meaning that at scale 2^{-j} Curvelet has an envelope which aligns along a ridge of length $2^{-j/2}$ and width 2^{-j}, resulting in a location, direction and scale dependent frequency plain tiling. Curvelets have been shown to be nearly optimal representation of functions smooth away from (piecewise) $\mathcal{C}^2$ singularities [37], [38] and hence the error of the k-term Curvelet approximation (corresponding to taking the k largest magnitude coefficients) in continuous setting asympotically with $k \to \infty$ behaves as [6]

$$\|g - g^k\|_2^2 \leq C \cdot (\log k)^3 \cdot k^{-2}. \quad (12)$$

Computation of Curvelets at the finest scale, 2^{-J}, is not straight forward because in order to capture the direction of the wave it is necessary that the Curvelet is sampled more finely than the maximal frequency in the image. As a result the frequency domain support of a Curvelet at the finest scale J with the direction θ_ℓ, $\tilde{U}_{J,\ell}$, does not fit into the fundamental cell $[-n_1/2, n_1/2 - 1] \times [-n_2/2, n_2/2 - 1]$. One solution given in [36] is to wrap the Fourier transform back onto the fundamental cell which effectively periodizes it

$$\begin{aligned} \tilde{U}_{J,\ell}\big[\big(i_1 + \tfrac{n_1}{2}\big) \bmod n_1 - \tfrac{n_1}{2}, \big(i_2 + \tfrac{n_2}{2}\big) \bmod n_2 - \tfrac{n_2}{2}\big] \\ = \tilde{U}_{J,\ell}(2\pi i_1, 2\pi i_2), \end{aligned}$$

where the indices $i_{1,2}$ are chosen in the support of $\tilde{U}_{J,\ell}(\cdot,\cdot)$. The periodization in Fourier domain corresponds to undersampling in space, which in turn causes aliasing. In [36] this effect was shown to account for less than 10% of the squared norm of the coefficients.

After periodization, the Fourier transform of the signal is multiplied by a sampled representation of a $\mathcal{C}^\infty$ partition of unity window i.e. a window which weights the original frequency and its periodic extension such that the sum of their squares is equal to 1. This acts to preserve the norm of the signal throughout periodization thereby guaranteeing that the computed transform is a numerical isometry. The $\mathcal{C}^\infty$ window used in [36] is a tensor product of the following one dimensional $\mathcal{C}^\infty$ windows

$$w[i] = \begin{cases} a(x_i)h_\infty(1-x_i), & i = 1,\dots,m_1,\\ 1, & i = m_1+1,\dots,3m_1+1,\\ a(x_i)h_\infty(x_{i-3m-1}), & i = 3m_1+2,\dots,4m_1+1, \end{cases} \tag{13}$$

where $x_i = (i-1)/(m_1-1),\ i = 1,\dots m_1,\ m_1 = \lfloor n_1/3 \rfloor$, h_∞ is a $\mathcal{C}^\infty$ monotonically decreasing function and

$$a(x) = (h_\infty(1-x)^2 + h_\infty(x)^2)^{-1/2} \tag{14}$$

is a normalizing factor. Expression (13) assumes $n_1 \bmod 3 = 0$, and hence $m_1 = n_1/3$ resulting in a window of length $4m_1 + 1$, while corresponding windows can be derived in the cases $n_1 \bmod 3 = 1, 2$.

C. Low-Frequency Curvelets

The sampled time series data, g^t, is band limited with the high frequencies roll off (see Section IV-A for the discussion of the smoothing effect of the physical measurement process) in contrast to standard images which are "sharp" i.e. high frequencies are not damped. As the Curvelet transform acts on localized regions in frequency space, this raises the question if we should represent those rolling off frequencies in the same way as the undamped frequencies, as the standard Curvelets would do at a cost which scales with size $n = n_1 n_2$ as $n \log n$.

As a compromise, we propose here to use a *low-frequency Curvelet* representation for the high resolution image g^t, i.e. to compute Curvelets corresponding to a subset of the frequencies $[-n_1^{LF}/2, n_1^{LF}/2-1] \times [-n_2^{LF}/2, n_2^{LF}/2-1], n_l^{LF} < n_l, l = 1,2$, excluding the Cartesian annulus of the highest frequencies $\big([-n_1/2, -n_1^{LF}/2-1] \cup [n_1^{LF}/2, n_1/2-1]\big) \times \big([-n_2/2, -n_2^{LF}/2-1] \cup [n_2^{LF}/2, n_2/2-1]\big)$. These highest frequencies in the excluded annulus are however present in the high resolution image g^t, therefore they can be used to fill up the frequency range needed for computation of the low-frequency Curvelets at the finest scale, instead of having to periodize. In this way, we avoid undersampling of the low-frequency Curvelets at the finest scale (note, that by construction the finnest scale of low-frequency Curvelets contains lower frequencies than the finest scale or standard Curvelets). We are going to concentrate on the implementation via wrapping because in this case the original Curvelet transform for the full set of frequencies results in a numerical isometry.

In what follows, we are going to denote the standard Curvelet transform for an $n_1 \times n_2$ image with J scales as $\mathcal{C}_J^{n_1,n_2}$ and the low-frequency Curvelet transform corresponding to frequencies up to $n_1^{LF}/2 \times n_2^{LF}/2$ (we will use such simplified notation to denote both the negative and positive frequencies), $n_l^{LF} < n_l$, $l = 1,2$ for the same image with $\mathcal{C}^{n_1,n_2}_{J,n_1^{LF},n_2^{LF}}$.

Let $\tilde{U}^{LF}_{J,\ell} = [-N_1^{LF}/2, N_1^{LF}/2-1] \times [-N_2^{LF}/2, N_2^{LF}/2-1]$ with $N_l^{LF} = 2\lfloor 2m_l^{LF} \rfloor + 1$, $m_l^{LF} = n_l^{LF}/3$, $l = 1,2$ denote the frequency domain support of the finest scale low-frequency Curvelet. In general we have two cases:

1) $N_l^{LF}/2 \le n_l/2$, $l = 1,2$: the frequency support of the finest scale low-frequency Curvelet $\tilde{U}^{LF}_{J,\ell}$ is contained in the fundamental cell $[-n_1/2, n_1/2-1] \times [-n_2/2, n_2/2-1]$;
2) $N_l^{LF}/2 > n_l/2$, $l = 1,2$: the finest scale low-frequency Curvelet support $\tilde{U}^{LF}_{J,\ell}$ extends to frequencies outside of the fundamental cell.

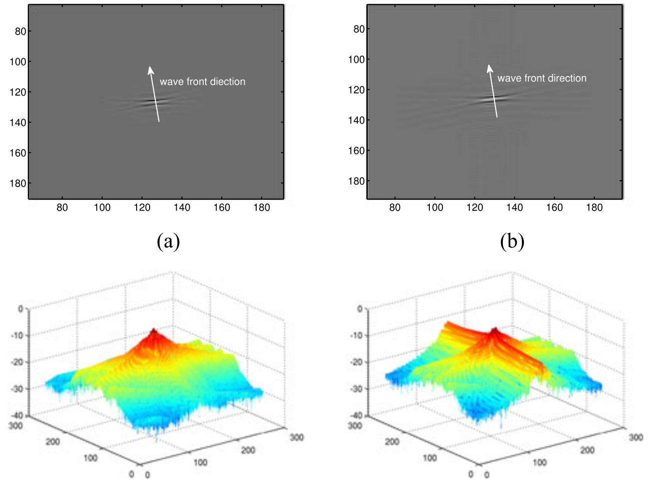
In the first case, to compute the standard Curvelets corresponding to image with maximal frequency $n_1^{LF}/2 \times n_2^{LF}/2$, $\mathcal{C}_J^{n_1^{LF},n_2^{LF}}$, one would compute the Fourier transform of g^t, restrict it to frequencies up to $n_1^{LF}/2 \times n_2^{LF}/2$ and subsequently periodize it while weighting with the C^∞ window (13). Instead, to compute the low-frequency Curvelets, $\mathcal{C}^{n_1,n_2}_{J,n_1^{LF},n_2^{LF}}$, we apply to the full Fourier transform of g^t (with frequencies up to $n_1/2 \times n_2/2$) a rectangular low-pass window with a cut-off frequency $N_l^{LF}/2$, $l = 1,2$ being the highest frequency in the support $\tilde{U}^{LF}_{J,\ell}$. Then we proceed as for standard Curvelets $\mathcal{C}_J^{n_1^{LF},n_2^{LF}}$ to compute the finest scale coefficients.

In the second case, to compute the low-frequency Curvelets, $\mathcal{C}^{n_1,n_2}_{J,n_1^{LF},n_2^{LF}}$, we still need to use periodization to fill up the support $\tilde{U}^{LF}_{J,\ell}$. However, we only periodically extend the Fourier transform of g^t beyond the range of available frequencies, here $n_1/2 \times n_2/2$, rather then $n_1^{LF}/2 \times n_2^{LF}/2$ as would be done using standard Curvelets, $\mathcal{C}_J^{n_1^{LF},n_2^{LF}}$. This results in a C^∞ window with shorter (and steeper) flanks as we only need to fill in the frequencies in the range $[-N_1^{LF}/2, -n_1/2-1] \cup [n_1/2, N_1^{LF}/2-1] \times [-N_2^{LF}/2, -n_2/2-1] \cup [n_2/2, N_2^{LF}/2-1]$. The corresponding window is a tensor product of the following one dimensional $\mathcal{C}^\infty$ windows

$$w^{LF}[i] = \begin{cases} a(x_i)h_\infty(1-x_i), & i = 1,\dots,N_1^{LF}-n_1,\\ 1, & i = N_1^{LF}-n_1+1,\dots,n_1,\\ a(x_i)h_\infty(x_{i-n_1}), & i = n_1+1,\dots,N^{LF}, \end{cases} \tag{15}$$

where $x_i = (i-1)/(N_1^{LF}-n_1-1),\ i = 1,\dots,N_1^{LF}-n_1$, and a and h_∞ are as before.

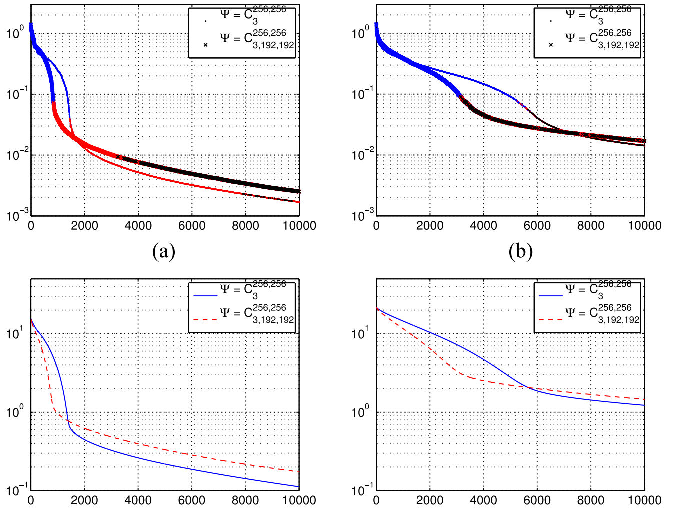
Subsequently, all the lower scales of low-frequency Curvelets can be computed exactly as for the standard Curvelets $\mathcal{C}_J^{n_1^{LF},n_2^{LF}}$. The construction of low-frequency Curvelets above results in a numerical isometry on the restriction to frequency range up to $\min\{N_1^{LF}/2, n_1/2\} \times \min\{N_2^{LF}/2, n_2/2\}$. Thus low-frequency Curvelet transform is a different transformation to the original Curvelet transform which however for a choice



(a) (b)

(c) (d)

Fig. 1. Zoomed in image of (a) standard Curvelet $\mathcal{C}_3^{256,256}$, (b) low-frequency Curvelet $\mathcal{C}_{3,192,192}^{256,256}$, at the finest level 3. (c), (d) Surface plot of log amplitude of (a), (b).



(a) (b)

(c) (d)

Fig. 2. Clock phantom. The decay of log amplitude of Curvelet coefficients of $g(\mathbf{x}_{\mathcal{S}}, t_i)$ at time steps (a) t_{100}, (b) t_{230}; the colors correspond to coefficients at different scales. (c), (d) The corresponding compression error in log scale.

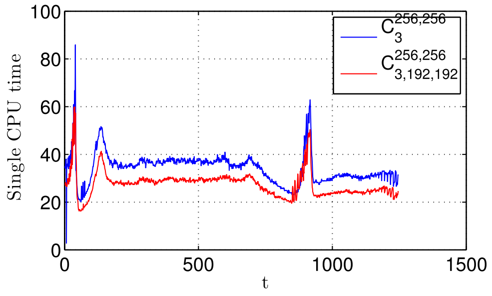


Fig. 3. Single CPU time for data reconstruction plotted in each time step.

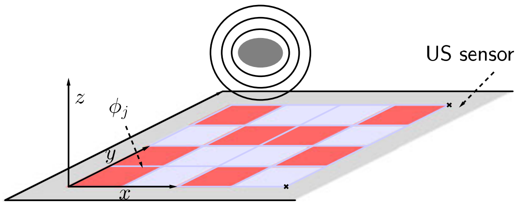


Fig. 4. Single-pixel optical camera.

of $n_l^{LF}/2$ such that $N_l^{LF}/2 \geq n_l/2$, $l = 1, 2$, is also an isometry on $[-n_1/2, n_1/2-1] \times [-n_2/2, n_2/2-1]$ as the standard Curvelets $\mathcal{C}_J^{n_1,n_2}$. Fig. 1(a) and (b) shows the finest scale Curvelet and low-frequency Curvelet, respectively. The plots of the logarithm of the amplitude in Fig. 1(c) and (d) illustrate that the decay of the low-frequency Curvelets along the wavefront direction is slower than that of the standard Curvelets, which is due to the frequencies higher then $n_l^{LF}/2$, $l = 1, 2$ used in computation of Curvelets at the finest scale.

Recapitulating, there are two major benefits of such low-frequency Curvelet transform. First, is the super linear reduction in computation cost of the transform, which is particularly beneficial for solution of the CS recovery problem (5) which involves repeated application of $\Phi\Psi^{\mathrm{T}}$ and its adjoint $\Psi\Phi^{\mathrm{T}}$. Second, is the ability to effectively represent realistic PAT data, in which due to measurement process the high frequencies are damped as in the model derived in Section IV-A. This results in amplitudes of low-frequency Curvelet coefficients being higher and exhibiting a quicker decay than those of the standard Curvelet coefficients (see Fig. 2 and the accompanying discussion in Section VI-A), and consequently in higher robustness to noise and imperfect compressibility.

V. COMPRESSED SENSING OF OPTICAL ULTRASOUND DETECTOR

A. *Single-Pixel Optical Camera*

In a series of publications [7], [39], [40] we introduced a single-pixel optical camera (SPOC) for ultrasonic and photoacoustic imaging. With the SPOC, instead of recording the pressure on the detector point-by-point, the entire active area of the optical ultrasound sensor is illuminated, and using a digital micro-mirror device (DMD) a pattern ϕ_j is applied to the wide field light reflected from the sensor resulting in a compressed measurement

$$b_j = \phi_j^{\mathrm{T}} g(\mathbf{x}_{\mathcal{S}}, t_i), \quad t_i \in (0, T). \tag{16}$$

Fig. 4 shows a sketch of the operational principle of SPOC while for technical details of the system we refer to [7].

B. *Sparse Representation of the Sensor Data*

In PAT the entire time series for one point or pattern is acquired with one excitation, $\phi_j^{\mathrm{T}} g(\mathbf{x}, t_i)$, $\mathbf{x} \in \mathcal{S}, t_i \in (0, T)$. This has the consequence that we acquire the same number of compressed/point measurements of the wave field at each time step t_i. Furthermore, due to the rate at which the DMD can change patterns, at least at present, we are limited to use of only one pattern throughout the acoustic propagation.

However, as the wave propagates, the complexity of the sensor data varies with time. After a wave front reaches the detector, its cross-section with the detector plane expands and more wave fronts, corresponding to features farther away, reach the detector. As a result, in general the complexity of the wave field at the sensor over time first increases and at some point it starts to decrease again corresponding to the tail of the wave field. If the complexity is reflected by the sparsity, the second term in the error bound in the robust recovery Theorem RR varies throughout the time series and with it the actual recovery error.

Furthermore, due the sound intensity, which is $\propto p(\mathbf{x}, t)^2$, obeying the inverse square law, $p(\mathbf{x}, t)$ decays as inverse distance of $\mathbf{x}$ to the source. This means that as with increasing t, $p(\mathbf{x}_{\mathcal{S}}, t)$ encodes information about p_0 further away from the sensor, the recorded pressure amplitudes decrease linearly for the same initial pressure amplitude value, resulting in a lower point wise signal to noise ratio for longer propagation times and ultimately larger recovery error ($\|f\|/\varepsilon$ decreases in Theorem RR).

Thus the error of the PAT data recovery is inhomogeneous across the time series as a consequence of the measurement setup.

C. Sensing Patterns

In our experiments we used scrambled Hadamard patterns

$$H_j^s = P_r H_j P_c, \tag{17}$$

where H_j is the $2^j \times 2^j$ Hadamard matrix and $P_c, P_r \in \{0,1\}^{2^j \times 2^j}$ denote permutation matrices for columns and rows, respectively. For compressed sensing we only select first $m \ll n = 2^j$ rows of H_j^s. The application of H_j^s to a vector $H_j^s v = P_r H_j P_c v$ amounts to application of the permutation matrix P_c to v, performing Hadamard transform on the permuted vector $P_c v$, and subsequently permuting the rows. Thus the scrambled Hadamard transform can be computed at essentially the same cost as the fast Hadamard transform, while scrambled Hadamard matrices have recovery properties similar to those of random Bernoulli matrices see e.g. [41].

In practice using the DMD, it was only possible to apply binary patterns. In order to make use of properties of Hadamard transform such as othogonality and self inversion, the experimental Hadamard matrix $H_j^{(0,1)}$ needs to be transformed into the Hadamard matrix H_j using the simple relation

$$H_j = \left(2H_j^{(0,1)} - \mathbf{1}\mathbf{1}^{\mathrm{T}}\right)/\sqrt{2^j}, \tag{18}$$

and correspondingly the measured data $w = H_j^{(0,1)} g$ into

$$H_j g = \frac{1}{\sqrt{2^j}}\left(2H_j^{(0,1)} - \mathbf{1}\mathbf{1}^{\mathrm{T}}\right) g = \frac{1}{\sqrt{2^j}}\left(2w - \mathbf{1}w(1)\right). \tag{19}$$

Here, we used that $w(1) = \mathbf{1}^{\mathrm{T}} g$ corresponds to the measurement acquired with 'all-1s' pattern, which is the first row of $H^{(0,1)}$. It is immediately clear that the same transformation (18), (19) holds for scrambled Hadamard matrices (17) (with the index 1 in $w(1)$ replaced by the 'all-1s' row number after row permutation).

Algorithm 1: Split Augmented Lagrangian Shrinkage Algorithm (SALSA), [42].

```
1: Choose $\mu > 0$, $v_0$ and $d_0$ and tolerance $\epsilon$
2: $i := 0$
3: repeat
4:     $f_{i+1} = \arg\min_f \|\Phi\Psi^{\mathrm{T}} f - b\|_2^2 + \mu\|f - v_i - d_i\|_2^2$
5:     $v_{i+1} = \arg\min_v \tau\|v\|_1 + \mu/2\|f_{i+1} - v - d_i\|_2^2$
6:     $d_{i+1} = d_i - (f_{i+1} - v_{i+1})$
7:     $i = i + 1$
8: until $|\zeta(f_{i+1}) - \zeta(f_i)|/\zeta(f_i) < \epsilon$
```

The light reflected from the DMD is integrated by a photodiode. In order to best utilize the dynamic range of the photo diode, it is necessary to keep the optical power incident on the photodiode in the same range for each pattern. All but the 'all-1s' pattern are composed of an equal number of 0s and 1s. Therefore, the 'all-1s' pattern was replaced with a vector which first half entries are 0 and the second half 1. As the negative of this vector (0 becomes 1 and vice versa) is exactly the $(2^j/2+1)$th row of $H_j^{(0,1)}$, the data corresponding to the 'all-1s' pattern can be constructed by adding the data from the modified first row pattern and $(2^j/2+1)$th row pattern. Again, this construction is not affected by scrambling.

D. Recovery of the Sensor Data

The sampled PAT data at each time step, $g(\mathbf{x}_{\mathcal{S}}, t)$, can be recovered by solving the optimization problem (5). The sensing matrix Φ in (5) is set to be the first m rows of the scrambled Hadamard matrix $H^s_{\log_2(n)}$, $n = n_1 n_2$,

$$\Phi = S H^s_{\log_2(n)} = S P_r H_{\log_2(n)} P_c, \tag{20}$$

where $S \in \{0,1\}^{m \times n}$ is a binary subsamplig matrix such that $S^{\mathrm{T}} S$ is an $n \times n$ diagonal matrix with ones at positions corresponding to the chosen and zeros to the omitted rows, respectively, while $SS^{\mathrm{T}} = I^{m \times m}$ is an $m \times m$ identity matrix. Consequently, we have $\Phi\Phi^{\mathrm{T}} = I^{m \times m}$.

The vector of measurements, b, is computed from experimental measurements using (19). The sparsifying transform Ψ is chosen as an orthonormal low-frequency or standard Curvelet transform on the $n_1 \times n_2$ optical sensor $\mathcal{S}$ and f^t is the (sparse) vector of sought for coefficients, $f^t = \Psi g^t$.

To take advantage of the described structure of the problem we solve (5) using the Split Augmented Lagrangian Shrinkage Algorithm (SALSA) proposed in [42]. SALSA is an ADMM scheme which solves the unconstrained problem

$$\min_f \zeta(f) := \frac{1}{2}\|\Phi\Psi^{\mathrm{T}} f - b\|_2^2 + \tau\|f\|_1. \tag{21}$$

A version used in this work is summarized in Algorithm 1.

The quadratic minimization problem in line 4 leads to the linear system

$$f_{i+1} = (\Psi\Phi^{\mathrm{T}}\Phi\Psi^{\mathrm{T}} + \mu I)^{-1}(\Psi\Phi^{\mathrm{T}} b + \mu(v_i + d_i)), \tag{22}$$

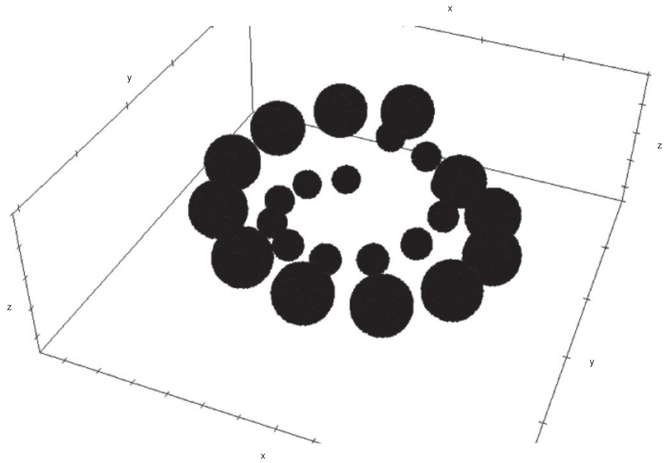


Fig. 5. Clock phantom.

where using Sherman Morrison Woodbury formula the inverse can be expressed in terms of application of $\Phi\Psi^{\mathrm{T}}$ and its adjoint

$$(\Psi\Phi^{\mathrm{T}}\Phi\Psi^{\mathrm{T}} + \mu I)^{-1} = \frac{1}{\mu}\left(I - \frac{1}{\mu+1}\Psi\Phi^{\mathrm{T}}\Phi\Psi^{\mathrm{T}}\right). \tag{23}$$

The proximal operator in line 5 amounts to point wise soft thresholding

$$\mathcal{T}_{\tau/\mu}(x) = \mathrm{sign}(x)(|x| - \tau/\mu)_+. \tag{24}$$

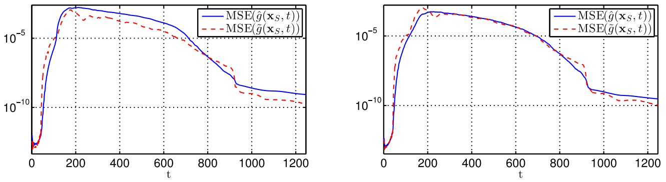
VI. PAT RECONSTRUCTION WITH THE RECOVERED SENSOR DATA

We present reconstruction results for three photoacoustic image reconstruction problems, reconstruction from: simulated pattern data, synthesized pattern data from a point-by-point scanner data, and data acquired with the SPOC (Section V-A). After the PAT data has been recovered using the proposed method for acoustic field reconstruction, the PAT images are reconstructed using time reversal via first order method in **k**-Wave Toolbox. All the computations were carried out on a desktop workstation with 12 Intel Xeon X5690 cores @3.47 GHz with 12MB cache using Matlab 2013a.

A. Simulated Data: Clock Phantom

In our first example we simulate the PAT data for initial pressure distribution p_0 depicted in Fig. 5. The "clock phantom" is a collection of balls which give rise to spherical wave forms expanding uniformly in all directions. This is naturally a difficult case for a directional basis like Curvelets which represents each direction independently resulting in potentially not the most economical representation.

We consider a volume of $256 \times 256 \times 92$ voxels of size $0.05\,\mathrm{mm}^3$ and the sensor of matching resolution, 256×256, placed at $z = 0$. We assume homogeneous ambient speed of sound and density of 1500 m/s and 1000 kg/m^3, respectively. The pressure at the sensor is sampled every $10\,\mathrm{ns}$. Both, the forward and inverse PAT problems are solved with first order method from **k**-Wave Toolbox. Before the PAT time series is simulated, p_0 is smoothed with the 3D Blackman window.



(a) (b)

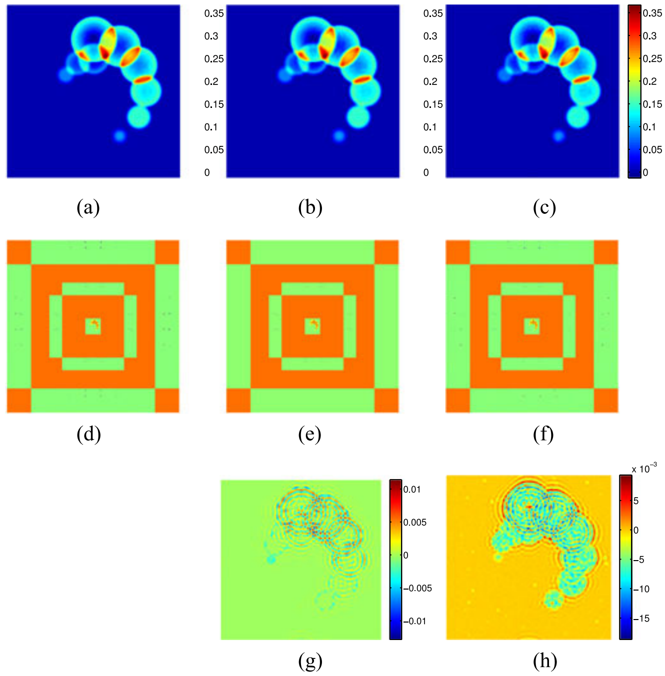
Fig. 6. MSE of the compressed PAT data $\hat{g}(\mathbf{x}_{\mathcal{S}}, t)$ versus the reconstructed PAT data $\tilde{g}(\mathbf{x}_{\mathcal{S}}, t)$ over time, for (a) standard Curvelet transform, $\mathcal{C}_3^{256,256}$; (b) low-frequency Curvelet transform, $\mathcal{C}_{3,192,192}^{256,256}$.

We attempt to recover the PAT data from 18% of noiseless compressed measurements obtained with binary scrambled Hadamard patterns. We solve the recovery problem in each time step using SALSA with the data dependent regularization parameter $\tau = 0.01\max(|\Psi\Phi^{\mathrm{T}}b^t|)$ and $\mu = 5\max(|\Psi\Phi^{\mathrm{T}}b^t|)/\|b^t\|_2$. We stop the algorithm if the relative change in the objective function $\zeta(f)$ (21) drops below $\epsilon = 5 \cdot 10^{-4}$ or after 100 iterations.

We start by examining the utility of the Curvelet transform as a sparsifying transform for PAT data recovery problem. In particular we compare the standard Curvelet transform $\mathcal{C}_3^{256,256}$ with the low-frequency Curvelet transform $\mathcal{C}_{3,192,192}^{256,256}$, the lowest resolution transform which is still an isometry on the full frequency range (up to $256/2 \times 256/2$).

First, we consider the approximation properties of the standard Curvelet and low-frequency Curvelet transforms. Fig. 2(a) and (b) shows the decay of the amplitudes of Curvelet coefficients for standard Curvelets $\mathcal{C}_3^{256,256}$ and low-frequency Curvelets $\mathcal{C}_{3,192,192}^{256,256}$ of the photoacoustic field at time steps $t_i = 100, 230$ corresponding to the fields depicted in Figs. 7(a) and 9(a). For the largest amplitude coefficients the low-frequency Curvelets consistently exhibit a quicker decay and correspondingly lower approximation error shown in Fig. 2(c) and (d). While in the early time steps $t_i = 100$ the decay rate difference is most pronounced for later time steps the difference gets smaller but is distributed over more coefficients. In all cases eventually the approximation error of standard Curvelet transform falls below that of the low-frequency Curvelet transform but this is only after all the significant coefficients have been captured.

Next, we compare the compression with 3% of coefficients to recovery from 18% of measurements which corresponds to 6 times the assumed sparsity, an empirically chosen factor for binary scrambled Hadamard patterns of size 256^2. Fig. 6 shows the mean square error (MSE) of the compressed versus recovered PAT data for both transforms. While the MSEs of the reconstructed data are almost identical, for the data compressed with the low-frequency Curvelets, after some initial time steps the MSE becomes lower than for the standard Curvelets, and most importantly it matches closer the MSE of the reconstructed data. This demonstrates that the low-frequency Curvelets are an adequate (while cheaper) representation of the PAT data. The series of Figs. 7–10 shows the PAT data $g(\mathbf{x}_{\mathcal{S}}, t)$, its compression $\hat{g}(\mathbf{x}_{\mathcal{S}}, t)$ and reconstruction $\tilde{g}(\mathbf{x}_{\mathcal{S}}, t)$ at different time

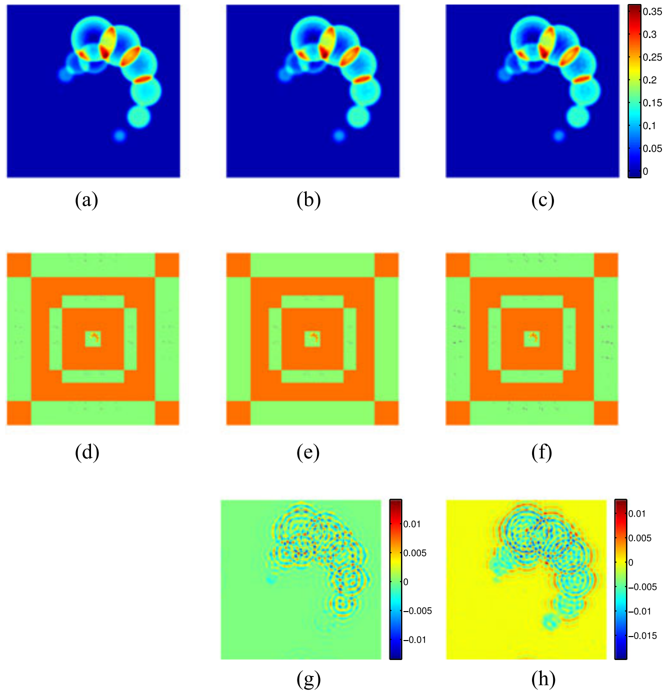


(a) (b) (c)

(d) (e) (f)

(g) (h)

Fig. 7. Clock phantom. PAT data at time step t_{100}, (a) simulated full data $g(\mathbf{x}_\mathcal{S}, t)$, (b) compressed $\hat{g}(\mathbf{x}_\mathcal{S}, t)$ and (c) reconstructed $\tilde{g}(\mathbf{x}_\mathcal{S}, t)$ data with standard Curvelet transform, $\mathcal{C}_3^{256,256}$. The colorbar applies to the entire row. The corresponding Curvelet coefficients (d-f) and (g) compression error $\hat{g}(\mathbf{x}_\mathcal{S}, t) - g(\mathbf{x}_\mathcal{S}, t)$, (h) recovery error $\tilde{g}(\mathbf{x}_\mathcal{S}, t) - g(\mathbf{x}_\mathcal{S}, t)$.

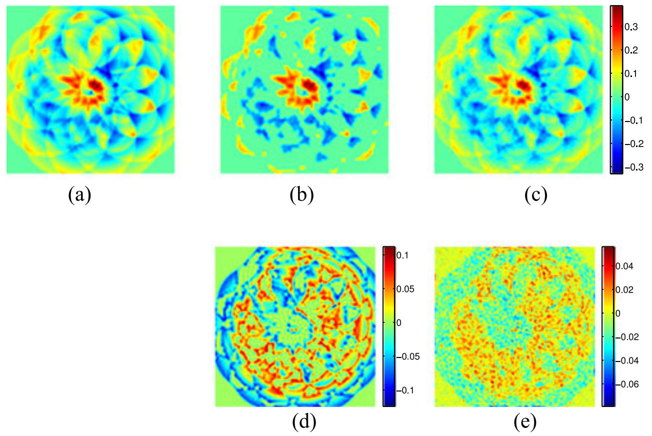


(a) (b) (c)

(d) (e) (f)

(g) (h)

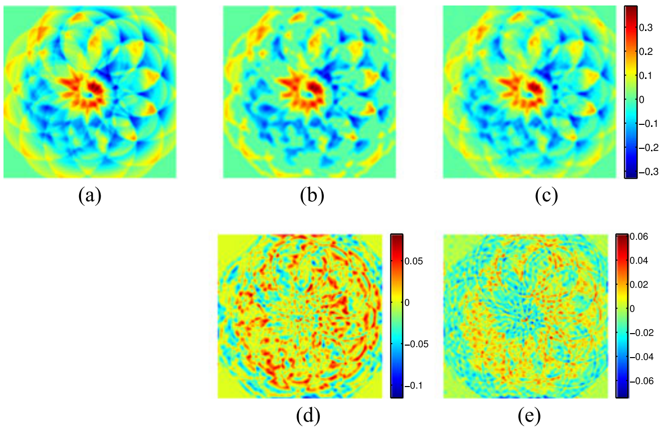
Fig. 8. Clock phantom. PAT data at time step t_{100}, (a) simulated full data $g(\mathbf{x}_\mathcal{S}, t)$, (b) compressed $\hat{g}(\mathbf{x}_\mathcal{S}, t)$ and (c) reconstructed $\tilde{g}(\mathbf{x}_\mathcal{S}, t)$ data with low-frequency Curvelet transform, $\mathcal{C}_{3,192,192}^{256,256}$. The colorbar applies to the entire row. The corresponding Curvelet coefficients (d-f) and (g) compression error $\hat{g}(\mathbf{x}_\mathcal{S}, t) - g(\mathbf{x}_\mathcal{S}, t)$, (h) recovery error $\tilde{g}(\mathbf{x}_\mathcal{S}, t) - g(\mathbf{x}_\mathcal{S}, t)$.



(a) (b) (c)

(d) (e)

Fig. 9. Clock phantom. PAT data at time step t_{230}, (a) simulated full data $g(\mathbf{x}_\mathcal{S}, t)$, (b) compressed $\hat{g}(\mathbf{x}_\mathcal{S}, t)$ and (c) reconstructed $\tilde{g}(\mathbf{x}_\mathcal{S}, t)$ data with standard Curvelet transform, $\mathcal{C}_3^{256,256}$. The colorbar applies to the entire row. (d) Compression error $\hat{g}(\mathbf{x}_\mathcal{S}, t) - g(\mathbf{x}_\mathcal{S}, t)$, (e) recovery error $\tilde{g}(\mathbf{x}_\mathcal{S}, t) - g(\mathbf{x}_\mathcal{S}, t)$.



(a) (b) (c)

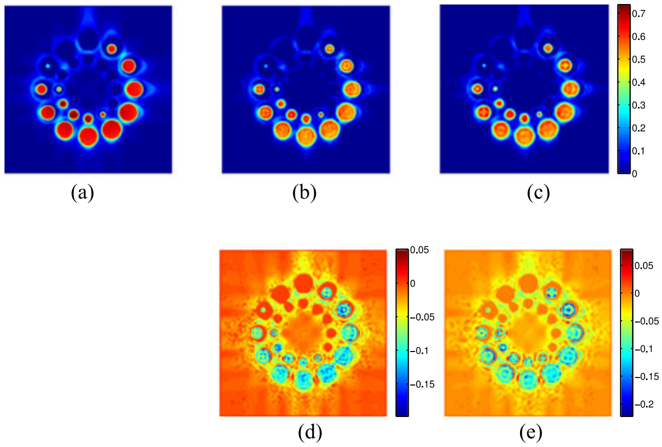
(d) (e)

Fig. 10. Clock phantom. PAT data at time step t_{230}, (a) simulated full data $g(\mathbf{x}_\mathcal{S}, t)$, (b) compressed $\hat{g}(\mathbf{x}_\mathcal{S}, t)$ and (c) reconstructed $\tilde{g}(\mathbf{x}_\mathcal{S}, t)$ data with low-frequency Curvelet transform, $\mathcal{C}_{3,192,192}^{256,256}$. The colorbar applies to the entire row. (d) Compression error $\hat{g}(\mathbf{x}_\mathcal{S}, t) - g(\mathbf{x}_\mathcal{S}, t)$, (e) recovery error $\tilde{g}(\mathbf{x}_\mathcal{S}, t) - g(\mathbf{x}_\mathcal{S}, t)$.

steps. Consistently, we observe that the higher scale coefficients are eliminated by compression while they partially reappear in the reconstruction hinting that the factor 6 maybe somewhat pessimistic. This is also evident in the higher frequency appearance of the error of the reconstruction in comparison to the compression.

The PAT image reconstruction from the recovered PAT data for both transforms is depicted in Fig. 11, which are visually very similar with MSE of $4.4243 \cdot 10^{-4}$ for the standard Curvelets and $4.6885 \cdot 10^{-4}$ for the low-frequency Curvelets, where the reconstruction from full data has been used as the ground truth in MSE calculations.

The reconstruction of the entire time series data with standard Curvelets takes 710.6 min on a single CPU as opposed to 571 min when low frequency Curvelets are used. Fig. 3 shows the respective single CPU reconstruction times over all



(a) (b) (c)

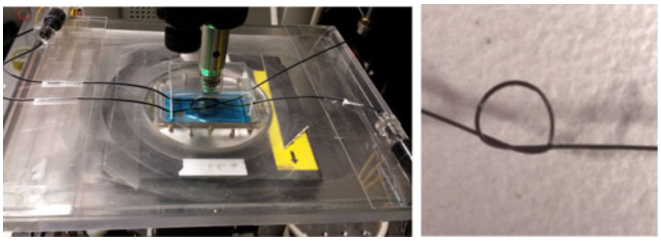
(d) (e)

Fig. 11. Clock phantom. Central slice through reconstructed PAT image $\tilde{p}_0$ from (a) full data $g(\mathbf{x}_S, t)$, (b) data reconstructed with standard Curvelet transform $\mathcal{C}_3^{256,256}$, (c) data reconstructed with low-frequency Curvelet transform $\mathcal{C}_{3,192,192}^{256,256}$. The colorbar applies to the entire row. The corresponding error (d), (e).

time steps. Note that each iteration of SALSA has the same complexity (but dependent on the choice of Curvelets), which for the standard Curvelets on average takes 0.8607 s while for the low frequency Curvelets only 0.582 s of a single CPU time. The qualitative behaviour is the same in both cases. The first phase up to the first peak corresponds to a very small pressure which is an artefact of k-space propagation method – we see the tail of the band limited interpolant. The phase up to the next peak corresponds to the pressure on the FP sensor gradually becoming more “complex” due to multiple arrivals, until the complexity of the pressure is higher than it can be fully captured by the k-term approximation resulting from the choice of the regularization parameter manifesting as a plateau. The following decrease corresponds to the wavefront leaving the detector, which then peaks again for the tail when the amplitude values are so small that the data fitting is ineffective. We would like to point out that the iteration numbers are dependent of the choice of the stoping criteria (here the relative change of the objective function). This criterium leads us to take to many iterations when the pressure magnitude is very small. Thus the final number of iterations could be further reduced using stopping criteria carefully tailored to the problem. As all the time steps can be reconstructed independently, the problem admits trivial parallelization and the single CPU time can be effectively divided by the number of available CPUs, in our case 12 resulting in approximately 59.2 min and 47.6 min for standard and low frequency Curvelets, respectively.

B. Synthesized Pattern Data: Knotted Tubes Filled With Ink

Next, we present reconstruction from compressed measurements synthesized from the point-by-point FP sensor measurements. The purpose of such data is to demonstrate the reconstruction with realistic noise but good signal to noise ratio, which at present is a limiting factor for SPOC.



(a) (b)

Fig. 12. (a) Knotted ink tubes phantom on the FP sensor. (b) Photo of the artificial hair phantom.

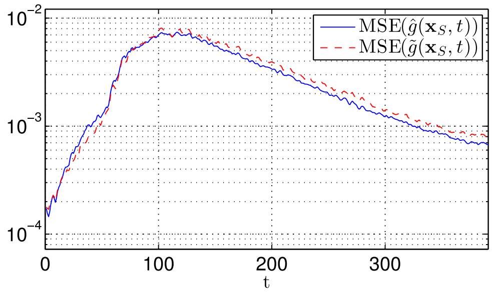


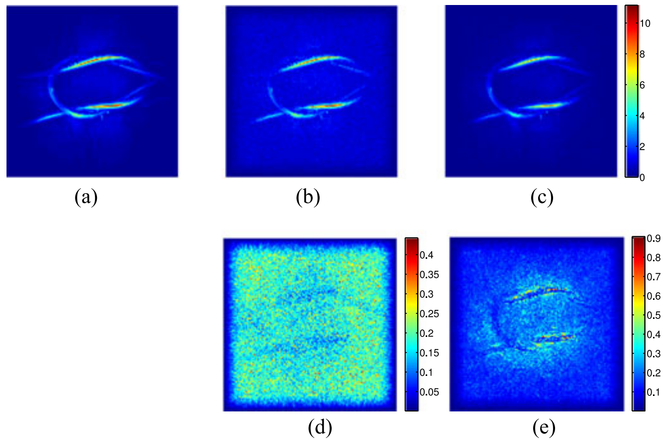
Fig. 13. Knotted tubes. MSE of the compressed PAT data $\hat{g}(\mathbf{x}_S, t)$ versus the reconstructed PAT data $\tilde{g}(\mathbf{x}_S, t)$ over time for low-frequency Curvelet transform $\mathcal{C}_{3,96,96}^{128,128}$.

Two polythene tubes were filled with 10% and 100% ink and tied into a knot, see Fig. 12(a). The tubes were immersed in a 1% intralipid solution. The wavelength of the excitation laser was 1064 nm delivering energy of approximately 20 mJ. A full scan data consists of 128 × 128 locations corresponding to spatial resolution of 150 μm × 150 μm, sampled at 625 time points corresponding to time resolution of 12 ns. The excitation laser rate was 20 Hz, which resulted in an acquisition time of 14 min for the *fully sampled* point data set. There was no averaging of the data.

For the chosen subsampling ratio of 25%, Fig. 13 shows good agreement between the MSEs of the compressed and reconstructed PAT data using low-frequency Curvelets. As a gold standard we take the reconstruction from a full point-by-point data set shown in Fig. 14(a). The linear reconstruction from 25% of patterns obtained by setting the missing Hadamard pattern measurements to 0 is shown in Fig. 14(b) while the nonlinear reconstruction with SALSA (with parameters $\tau = 0.01 \max(|\Psi\Phi^{\mathrm{T}} b^t|)$, $\mu = 5 \max(|\Psi\Phi^{\mathrm{T}} b^t|)/\|b^t\|_2$ and stopping tolerance $\epsilon = 5 \cdot 10^{-4}$ or after 100 iterations) in Fig. 14(c). The nonlinear reconstruction effectively restores the contrast lost in the linear 0-padded reconstruction. The MSE with respect to the full data reconstruction is $6.7582 \cdot 10^{-3}$ for the linear 0-padded reconstruction and $2.0824 \cdot 10^{-3}$ for the nonlinear reconstruction. The single CPU time for the nonlinear reconstruction of the entire time series data was 94 min with a total of 28591 SALSA iterations (average time per iteration 0.1973 s). Our computation was carried out on 12 cores simultaneously resulting in an actual real computing time of 470 s.

C. Experimental Pattern Data: Hair Knot

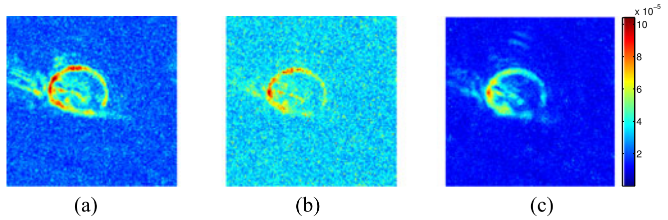
Our last example is a synthetic hair knot phantom of diameter ~150 μm, immersed in 1% intralipid solution and positioned approximately 2 mm above the sensor and 3 mm deep below



(a) (b) (c)

(d) (e)

Fig. 14. Knotted tubes. Maximum intensity projection of reconstructed PAT image $\tilde{p}_0$ from (a) full data $g(\mathbf{x}_{\mathcal{S}}, t)$, (b) linearly reconstructed data $\tilde{g}^{lin}(\mathbf{x}_{\mathcal{S}}, t)$ (here plotted scaled by a factor 3.77), (c) nonlinearly reconstructed data $\tilde{g}(\mathbf{x}_{\mathcal{S}}, t)$ with low-frequency Curvelet transform $\mathcal{C}^{256,256}_{3,192,192}$ (the colorbar applies to the entire row), and (d), (e) the corresponding error.



(a) (b) (c)

Fig. 15. Hair knot. Maximum intensity projection of reconstructed PAT image $\tilde{p}_0$ from (a) full data $g(\mathbf{x}_{\mathcal{S}}, t)$, (b) linearly reconstructed data $\tilde{g}^{lin}(\mathbf{x}_{\mathcal{S}}, t)$ (here plotted scaled by a factor 4.12), (c) nonlinearly reconstructed data $\tilde{g}(\mathbf{x}_{\mathcal{S}}, t)$ with low-frequency Curvelet transform $\mathcal{C}^{128,128}_{3,96,96}$. The colorbar applies to the entire row.

the intralipid surface. The photo of the phantom is shown in Fig. 12(b). The area of DMD corresponding to 640×640 micromirrors grouped in 5×5 was used to form 128×128 pixels. Due to the angle of the optical path, each such pixel corresponds to an area of 62.12 μm × 68 μm on the FP sensor. The time series was sampled every 20 ns, resulting in 81 time steps. In the actual experiment, we collected the entire set of 128^2 patterns. The repetition rate of the excitation laser was 20 Hz. The measurement was averaged over four excitations resulting in a total acquisition time of 54 min for the complete 4-times averaged pattern set. If only 18% of the data was collected this would correspond to acquisition time of 9 min and 43 s. Please note that this is not indicative of the actual frame rates achievable by the system, as a slow excitation laser was used in this proof of concept experiment. The speed of sound used for time reversal was 1490 m/s.

Fig. 15(a) shows the PAT image reconstruction obtained from the full set of scrambled binary Hadamard 128^2 patterns (the full point data was computed inverting scrambled Hadamard transform). The $\sim$4.12 amplified linear reconstruction from 0-padded 18% measurements is shown in Fig. 15(b) and the nonlinear reconstruction obtained with SALSA ($\tau = 0.05 \max(\Psi\Phi^{\mathrm{T}} b^t), \mu = 0.75 \max(\Psi\Phi^{\mathrm{T}} b^t)/\|b^t\|_2$ and the stopping tolerance $\epsilon = 5 \cdot 10^{-3}$ or maximum 100 iterations) is illustrated in Fig. 15(c).

The single CPU time for the entire time series data reconstruction was 372.5 s with total 1566 SALSA iterations (average time per iteration 0.2379 s). The computation was carried out on 12 cores in parallel, resulting in actual real computing time of 31 s.

VII. CONCLUSIONS AND DISCUSSION

We presented a method for acoustic field reconstruction from a limited number of patterned measurements from an optical ultrasound detector. When compressively sensing photoacoustic signals, our method recovered the pressure field on the sensor (the PAT data) at each time step independently using the sparsity of the data in a low-frequency Curvelet frame, which is a modification of the standard Curvelets tailored to account for the smoothing of the wave front during optical acquisition. The major advantage of such a scheme is that the series of problems to solve are standard (2D) CS recovery problems and as they are independent they can be solved in parallel. Furthermore, decoupling the CS and acoustic reconstruction affords more flexibility in PAT modeling, for instance including absorption or nonlinear effects, and allows for the use of highly optimized readily available software for photoacoustic image reconstruction.

One of the major challenges for compressed sensing of photoacoustic signals is that the same number of interrogation patterns is applied in each time step. As the sparsity of the wave field on the detector varies, in the proposed scheme this inevitably results in different quality reconstruction in different time steps. In particular, we first lose the "sharpness" of the wave front, because this is reflected in the coefficients at the highest scale which magnitude is generally smaller. This effect is partially counterweight by the proposed low-frequency Curvelet representation which is tailored to frequency range of photoacoustic data and hence boosts those coefficients. Furthermore, in our experiments we observed that the data recovery errors were partially alleviated during the acoustic inversion. As the initial pressure p_0 is mapped to the entire time series, the entire time series carries the information about the wave front and the acoustic inversion acts to average out data errors. In [12], [13] this problem was tackled utilizing the sparsity directly in the photoacoustic image p_0, rather than in the data. There we solve one large (3D) CS recovery problem where the CS sensing operator is a composition of the acoustic propagation and pattern measurements. Such forward operator is significantly more expensive to apply and its incoherence properties have to be analyzed. Finally, nonlinear effects such as acoustic absorption are out of scope of the standard linear CS framework.

In future work, we intend to extend the here proposed acoustic field reconstruction method to efficiently handle dynamic problems.

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] E. J. Candés, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[3] E. J. Candés, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[4] M. F. Duarte *et al.*, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, Mar. 2008.

[5] E. J. Candes and L. Demanet, "The curvelet representation of wave propagators is optimally sparse," *Commun. Pure Appl. Math.*, vol. 58, no. 11, pp. 1472–1528, 2005.

[6] E. J. Candés and D. L. Donoho, "New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities," *Commun. Pure Appl. Math*, vol. 57, no. 2, pp. 219–266, Feb. 2004.

[7] N. Huynh, E. Zhang, M. Betcke, S. Arridge, P. Beard, and B. Cox, "Single-pixel optical camera for video rate ultrasonic imaging," *Optica*, vol. 3, no. 1, pp. 26–29, Jan. 2016.

[8] J. Provost and F. Lesage, "The application of compressed sensing for photo-acoustic tomography," *IEEE Trans. Med. Imag.*, vol. 28, no. 4, pp. 585–594, Apr. 2009.

[9] Z. Guo, C. Li, L. Song, and L. V. Wang, "Compressed sensing in photoacoustic tomography in vivo," *J. Biomed. Opt.*, vol. 15, no. 2, 2010, Art. no. 021311.

[10] C. Huang, K. Wang, L. Nie, L. V. Wang, and M. A. Anastasio, "Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 1097–1110, Jun. 2013.

[11] Z. Belhachmi, T. Glatz, and O. Scherzer, "A direct method for photoacoustic tomography with inhomogeneous sound speed," *Inverse Probl.*, vol. 32, no. 4, Apr. 2016, Art. no. 045005.

[12] S. R. Arridge, M. M. Betcke, B. T. Cox, F. Lucka, and B. E. Treeby, "On the adjoint operator in photoacoustic tomography," *Inverse Probl.*, vol. 32, no. 11, 2016, Art. no. 115012. [Online]. Available: http://stacks.iop.org/0266–5611/32/i=11/a=115012

[13] S. Arridge *et al.*, "Accelerated high-resolution photoacoustic tomography via compressed sensing," *Phys. Med. Biol.*, vol. 61, no. 24, 2016, Art. no. 8908. [Online]. Available: http://stacks.iop.org/0031–9155/61/i=24/a=8908

[14] M. Haltmeier, T. Berer, S. Moon, and P. Burgholzer, "Compressed sensing and sparsity in photoacoustic tomography," *J. Opt.*, vol. 18, no. 11, 2016, Art. no. 114004.

[15] M. Xu and L. V. Wang, "Universal back-projection algorithm for photoacoustic computed tomography," *Phys. Rev. E*, vol. 71, Jan 2005, Art. no. 016706. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevE.71.016706

[16] M. Haltmeier, "Universal inversion formulas for recovering a function from spherical means," *SIAM J. Math. Anal.*, vol. 46, no. 1, pp. 214–232, 2014.

[17] L. V. Wang, *Photoacoustic Imaging and Spectroscopy*. Boca Raton, FL, USA: CRC Press, Mar. 2009, vol. 144.

[18] P. Beard, "Biomedical photoacoustic imaging," *Interface Focus*, vol. 1, no. 4, pp. 602–631, Aug. 2011.

[19] K. S. Valluru, K. E. Wilson, and J. K. Willmann, "Photoacoustic imaging in oncology: Translational preclinical and early clinical experience," *Radiology*, vol. 280, no. 2, pp. 332–349, 2016.

[20] Y. Zhou, J. Yao, and L. V. Wang, "Tutorial on photoacoustic tomography," *J. Biomed. Opt.*, vol. 21, no. 6, 2016, Art. no. 061007.

[21] J. Xia and L. V. Wang, "Small-animal whole-body photoacoustic tomography: A review," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1380–1389, May 2014.

[22] E. Zhang, J. Laufer, and P. Beard, "Backward-mode multiwavelength photoacoustic scanner using a planar Fabry–Perot polymer film ultrasound sensor for high-resolution three-dimensional imaging of biological tissues," *Appl. Opt.*, vol. 47, no. 4, pp. 561–577, Feb. 2008.

[23] B. E. Treeby, J. Jaros, A. P. Rendell, and B. T. Cox, "Modeling nonlinear ultrasound propagation in heterogeneous media with power law absorption using a k-space pseudospectral method." *J. Acoust. Soc. Amer.*, vol. 131, no. 6, pp. 4324–36, 2012.

[24] D. Finch, R. Sarah, and K. Patch, "Determining a function from its mean values over a family of spheres," *SIAM J. Math. Anal.*, vol. 35, no. 5, pp. 1213–1240, Jan. 2004.

[25] P. Kuchment and L. Kunyansky, *Mathematics of Photoacoustic and Thermoacoustic Tomography*. New York, NY, USA: Springer, 2011, pp. 817–865.

[26] L. V. Nguyen, "On singularities and instability of reconstruction in thermoacoustic tomography," *Tomography and Inverse Transport Theory (ser. Contemporary Mathematics)*. Providence, RI, USA: Amer. Math. Soc., 2011, vol. 559.

[27] P. Stefanov and G. Uhlmann, "Thermoacoustic tomography with variable sound speed," *Inverse Probl.*, vol. 25, no. 7, 2009, Art. no. 075011.

[28] B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *J. Biomed. Opt.*, vol. 15, no. 2, 2010, Art. no. 021314.

[29] E. J. Candés, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, pp. 589–592, May 2008.

[30] M. Fornasier and H. Rauhut, "Compressive sensing," in *Handbook of Mathematical Methods in Imaging*. New York, NY, USA: Springer, 2011, pp. 187–228.

[31] B. T. Cox and P. C. Beard, "Fast calculation of pulsed photoacoustic fields in fluids using k-space methods," *J. Acoust. Soc. Amer.*, vol. 117, no. 6, pp. 3616–3627, 2005.

[32] K. P. Köstli, M. Frenz, H. Bebie, and H. P. Weber, "Temporal backward projection of optoacoustic pressure transients using Fourier transform methods," *Phys. Med. Biol.*, vol. 46, no. 7, pp. 1863–1872, Jul. 2001.

[33] Y. Xu, D. Feng, and L. V. Wang, "Exact frequency-domain reconstruction for thermoacoustic tomography. I. Planar geometry," *IEEE Trans. Med. Imag.*, vol. 21, no. 7, pp. 823–828, Jul. 2002.

[34] M. Tabei, T. D. Mast, and R. C. Waag, "A k-space method for coupled first-order acoustic propagation equations," *J. Acoust. Soc. Amer.*, vol. 111, no. 1, pp. 53–63, Jan. 2002.

[35] B. E. Treeby and B. T. Cox, "A k-space Green's function solution for acoustic initial value problems in homogeneous media with power law absorption," *J. Acoust. Soc. Amer.*, 2011.

[36] E. Candés, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Model. Simul.*, vol. 5, no. 3, pp. 861–899, 2006.

[37] D. L. Donoho, "Sparse components of images and optimal atomic decompositions," *Constructive Approx.*, vol. 17, no. 3, pp. 353–382, 2001.

[38] B. S. Kashin, "Approximation properties of complete orthonormal systems," *Trudy Mat. Inst. Steklov.*, vol. 172, pp. 187–191, 353, 1985.

[39] 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," *Proc. SPIE*, vol. 8943, Mar. 2014, Art. no. 894327.

[40] N. Huynh, E. Zhang, M. Betcke, S. Arridge, P. Beard, and B. Cox, "A real-time ultrasonic field mapping system using a Fabry–Perot single pixel camera for 3D photoacoustic imaging," *Proc. SPIE*, vol. 9323, 2015, Art. no. 93 231O.

[41] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressed Sensing*. New York, NY, USA: Springer, 2010.

[42] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, Sep. 2010.

Authors' photographs and biographies not available at the time of publication.