Photo-acoustic tomographic image reconstruction from reduced data using physically inspired regularization

Rejesh, Nadaparambil Aravindakshan; Kalva, Sandeep Kumar; Pramanik, Manojit;
Arigovindan, Muthuvel

2020

Rejesh, N. A., Kalva, S. K., Pramanik, M., & Arigovindan, M. (2020). Photo-acoustic tomographic image reconstruction from reduced data using physically inspired regularization. *Journal of Instrumentation*, 15(12), P12028-.
[doi:10.1088/1748-0221/15/12/P12028](https://doi.org/10.1088/1748-0221/15/12/P12028)

<https://hdl.handle.net/10356/146416>

<https://doi.org/10.1088/1748-0221/15/12/P12028>

© 2020 IOP Publishing Ltd. All rights reserved. This is an author-created, un-copyedited version of an article accepted for publication in *Journal of Instrumentation*. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The definitive publisher authenticated version is available online at <https://doi.org/10.1088/1748-0221/15/12/P12028>

Downloaded on 26 Apr 2023 12:46:50 SGT

Photo-acoustic tomographic image reconstruction from reduced data using physically inspired regularization

Nadaparambil Aravindakshan Rejesh,^a Sandeep Kumar Kalva,^b Manojit Pramanik,^b
Muthuvel Arigovindan.^{a,1}

^a*Imaging Systems Lab, Electrical Engineering, Indian Institute of Science (IISc), Bangalore 560012, India*

^b*School of Chemical and Biomedical Engineering, Nanyang Technological University, 62 Nanyang Drive, 637459 Singapore, Singapore*

E-mail: mvel@iisc.ac.in

ABSTRACT: We propose a model based image reconstruction method for photoacoustic tomography (PAT) involving a novel form of regularization, and demonstrate its ability to recover good quality images from datasets of significantly reduced size. The regularization is constructed to suit the physical structure of typical PAT images. We construct it by combining second order derivatives and intensity into a non-convex form so as to exploit a structural property of PAT images that we observe: in PAT images, high intensities and high second order derivatives are jointly sparse. The specific form of regularization constructed here is a modification of the form that was proposed for fluorescence image restoration. This regularization is combined with a data fidelity cost and the required image is obtained as the minimizer of this cost. As this regularization is non-convex, efficiency of the minimization method is crucial in obtaining artifact-free reconstructions. We develop a custom minimization method for efficiently handling this non-convex minimization problem. Further, as non-convex minimization requires large number of iterations, and the PAT forward model in the data-fidelity term has to be applied in the iterations, we propose a computational structure for efficient implementation of the forward model with reduced memory requirements. We evaluate the proposed method on both simulated and real measured data sets and compare with a recent reconstruction method that is based on well-known fast iterative shrinkage threshold algorithm (FISTA).

ARXIV EPRINT: [1908.02313](https://arxiv.org/abs/1908.02313)

¹Corresponding author.

Contents

1	Introduction	1
1.1	Forward model in photoacoustic tomography	3
1.2	Discrete forward model	3
1.3	Model based image reconstruction	4
2	The proposed reconstruction method	5
2.1	Proposed Regularization functional	5
2.2	The complete cost functional	5
2.3	Proposed Algorithm	6
2.4	Matrix-free implementation of the forward model	9
3	Reconstruction results	10
3.1	Reconstruction from simulated data	10
3.2	Reconstruction from real measured data	15
4	Conclusions	16

1 Introduction

Photoacoustic tomography (PAT) [1–7] provides high resolution and high contrast images of deep tissues by imaging the spatial distribution of certain substances that can absorb near infra-red optical energy. Upon shining with a laser pulse, the substance under investigation absorbs the optical energy and undergo thermoelastic expansion; as a result, the spatial distribution of the concentration of the substance gets translated into distribution of pressure-rise. This initial pressure rise travels outwards as ultrasound waves which are collected by ultrasound transducers placed at the boundary. From the ultrasound signal measured by the transducers as function of time, a PAT reconstruction method recovers an estimate of the initial pressure-rise by solving the associated inverse problem [8]. Achieving accuracy in image reconstruction is a fundamental challenge in PAT; the reconstructed images suffer from artifacts that depend on the measurement geometry mainly because the reconstruction problem is complex.

There are a class of methods known as analytical inversion methods, which obtain reconstruction (recovery of initial pressure rise) by some transformation on the measured data. Among such methods, filtered back projection (FBP) [9–12] and delay and sum method [13, 14] are fast and memory efficient, but require a large number of measurements for good quality reconstruction. This leads to increased scan time or expensive instrumentation setups. Time reversal methods [15] are the least demanding class of methods in the category of analytical methods, and can be used for any detection geometry [16], acoustic absorption and dispersion [17], and can accommodate the heterogeneities present in the medium [18]. However, in many applications of PAT, due to

the geometrical limitations [19] or by choice to accelerate data acquisition [20], we may have to put restrictions on the spatial and/or temporal sampling of the photoacoustic (PA) signal. In such situations, these methods suffer from blurring and curved-line artifacts.

Model-based reconstruction methods outperform the direct analytic reconstruction methods in the case of limitation in the size of measured data, and yield better quality in reconstruction [21–24]. These methods express the measured pressure samples as a linear transformation on the initial pressure-rise. This transformation is implemented in the form of multiplication of the vector containing image pixels with a very large matrix that represents the PAT forward model (model matrix). This transformation also goes into the data fidelity term, which is minimized along with a regularization functional to achieve reconstruction. The regularization represents a prior belief on the spatial characteristics of image to be recovered [25–28]. The minimization is achieved by means of iterations involving repeated application of the above-mentioned transformation and its adjoint leading to high computational burden and memory-burden. To reduce the memory requirements and computational complexity, several methods have been reported. For example, the memory overhead can be reduced by decreasing the number of measurements [29] or by calculating the matrix-vector products on-the-fly without explicitly storing the model matrix [30]. But since the same operations need to be performed multiple times, the computational time becomes longer. A fast implementation can be achieved by simplifying the forward model with the assumption that the photoacoustic sources lie in a plane [31]. It can also be made faster by decoupling the inversion into smaller subproblems using a discrete wavelet packet decomposition [32]. But the method uses interpolation, which leads to modeling errors. A more accurate model matrix calculation based on the direct discretization of Poisson-type integral showed better reconstruction accuracy in high-noise and low-SNR imaging conditions [33].

We develop a novel model based reconstruction method that yields significantly improved reconstruction from datasets of reduced size. Our contributions are threefold. First, we note that the existing model-based methods in PAT use generic regularization forms developed for general images. There has been no reported application of a regularization that can cater to the structural properties of the photoacoustic images. We observe that PAT images has the following characteristics which is also observed in fluorescence images: in these images high intensities and high second order derivatives are jointly sparse. This property was exploited for fluorescence image reconstruction by constructing a regularization that combines the intensity and the second order derivatives [34]. Here we modify this form such that it is more suitable for the current reconstruction problem. Second, we construct a novel preconditioned gradient method for efficiently minimizing the cost function that combines this regularization and the data-fidelity term. Finally, we develop an efficient method for on-the-fly computation of matrix-vector products derived using a model that uses an exact time propagator to calculate the acoustic field [35, 36]. By taking advantage of the filtering structure present in this formulation, forward and adjoint matrix operations are implemented using filters in a memory efficient way. We evaluate the proposed method and compare it with a recent reconstruction method that is based on well-known fast iterative shrinkage threshold algorithm (FISTA). For this purpose, we use simulated data corresponding to reduced number of transducers and real measured data.

1.1 Forward model in photoacoustic tomography

The forward problem in PAT accounts for the calculation of the pressure fields in space and time $p(\mathbf{r}, t)$ from a known photoacoustic source $H(\mathbf{r}, t)$, which represents the light energy deposited in the medium per unit volume per unit time. The induced pressure waves $p(\mathbf{r}, t)$ under the condition of thermal and stress confinements obey the following differential equation for an acoustically homogeneous medium [4, 8, 37]

$$\left(\frac{\partial^2}{\partial t^2} - c_0^2 \nabla^2 \right) p(\mathbf{r}, t) = \Gamma \frac{\partial}{\partial t} H(\mathbf{r}, t), \quad (1.1)$$

where Γ is a dimensionless parameter called the Grüneisen coefficient, that describes the conversion efficiency of heat to pressure and c_0 is the speed of sound in the medium. By recognizing that the temporal duration of the laser pulse is shorter than the temporal resolution of the ultrasound detectors in most of the photoacoustic imaging applications, the PA source $H(\mathbf{r}, t)$ may be approximated by $H(\mathbf{r})\delta(t)$, where $H(\mathbf{r})$ is the density of deposited energy. Then the solution to differential equation (1.1) can be written as [8]

$$p(\mathbf{r}, t) = \frac{\Gamma}{4\pi c_0} \frac{\partial}{\partial t} \int_{|\mathbf{r}-\mathbf{r}'|=c_0 t} \frac{H(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}', \quad (1.2)$$

where the initial pressure field $p_0(\mathbf{r}) = p(\mathbf{r}, t = 0)$ can be written as

$$p_0(\mathbf{r}) = \Gamma H(\mathbf{r}). \quad (1.3)$$

The pressure distribution, $p(\mathbf{r}, t)$, can also be expressed as [35, 36]

$$p(\mathbf{r}, t) = \mathcal{F}^{-1} \left\{ \hat{P}_o(\mathbf{k}) \cos(c_0 \|\mathbf{k}\| t) \right\}, \quad (1.4)$$

where $\mathcal{F}^{-1}$ denotes the Fourier inversion, and $\hat{P}_o(\mathbf{k})$ is the Fourier transform of $p_0(\mathbf{r})$ with $\mathbf{k}$ denoting the 2D Fourier frequency.

1.2 Discrete forward model

The discrete representation of the forward problem is the basis of model-based reconstruction algorithms in PAT. Since the measured pressure field is linearly related to the photoacoustic source, the discretization of the forward problem may be written in a matrix form as [22]

$$\mathbf{p}_m = \mathbf{H} \mathbf{p}_0, \quad (1.5)$$

where $\mathbf{p}_m$ is a $LM \times 1$ vector representing the discrete pressure measurements from L transducers each taking M time-samples and $\mathbf{H}$ is the model matrix of size $LM \times N$ with N being the total number of pixels. The initial pressure distribution $p_0(\mathbf{r})$ is represented spatially using a 2D imaging grid having N_x and N_y grid points along x and y directions respectively and is denoted by a $N \times 1$ vector $\mathbf{p}_0$ where $N = N_x \times N_y$. The model matrix can be calculated by discretizing the integral relation in (1.2) in the ideal case of homogeneous lossless medium and point detectors. Several methods to improve the accuracy of the model by incorporating transducer responses, heterogeneity in the

medium and interpolation techniques for accurate discretization has been reported [25, 31, 38]. Including the measurement noise, the modified imaging model can be written as

$$\mathbf{p}_m = \mathbf{H}\mathbf{p}_0 + \boldsymbol{\eta}, \quad (1.6)$$

where $\boldsymbol{\eta}$ represents measurement noise, which is Gaussian.

1.3 Model based image reconstruction

The aim of a reconstruction method in PAT is to recover the initial pressure distribution $\mathbf{p}_0$ from the noisy transducer measurement data $\mathbf{p}_m$. In the limited data case, the PA image reconstruction problem is ill-posed and hence constraints are imposed on the required solution in the form of a regularization. In this case, the image reconstruction problem can be treated as an optimization problem where the solution is obtained by minimizing a cost function. The reconstruction problem can be written as

$$\hat{\mathbf{p}}_0 = \underset{\mathbf{p}_0}{\operatorname{argmin}} J(\mathbf{p}_0) \quad (1.7)$$

where $J(\mathbf{p}_0)$ is the cost function and is given by

$$J(\mathbf{p}_0) = \|\mathbf{p}_m - \mathbf{H}\mathbf{p}_0\|_2^2 + \lambda R(\mathbf{p}_0). \quad (1.8)$$

Here $R(\mathbf{p}_0)$ is the regularization functional and $\|\cdot\|_2$ represents the l_2 norm. The regularization parameter λ controls the amount of regularity in the solution and fidelity to the measured data.

The regularization functional typically should be able to smooth the noise and hence contains derivative terms in their formulation. A quadratic and differentiable regularization functional called Tikhonov regularization has been used in limited data cases [8, 39] and is given by

$$R(\mathbf{p}_0) = \sum_i \|\mathbf{D}_{o,i}\mathbf{p}_0\|_2^2 = \sum_{r=1}^N \sum_i ((\mathbf{D}_{o,i}\mathbf{p}_0)_r)^2 \quad (1.9)$$

where $(\cdot)_r$ denotes the r th component of its vector argument, and $\mathbf{D}_{o,i}$ represents the matrix of i^{th} derivative filter of order o . For example, $\mathbf{D}_{1,i}, i = 1, 2$, are the matrix equivalents of filtering by discrete filters that implement the operators $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$. Further, $\mathbf{D}_{2,i}, i = 1, 2, 3$, are the matrix equivalents of filtering by discrete filters that implement the operators $\frac{\partial^2}{\partial x^2}$, $\frac{\partial^2}{\partial y^2}$ and $\sqrt{2}\frac{\partial}{\partial x\partial y}$. The resulting minimization of the convex quadratic cost function yields a closed form solution given by

$$\mathbf{p}_0 = [\mathbf{H}^T \mathbf{H} + \lambda \sum_i \mathbf{D}_{o,i}^T \mathbf{D}_{o,i}]^{-1} \mathbf{H}^T \mathbf{p}_m. \quad (1.10)$$

As it precludes any large derivative values, Tikhonov regularization tends to smooth edges in the reconstructed image.

It should be emphasized that, as of now, only the data-fidelity term requires building a large matrix $\mathbf{H}$, and the regularization does not actually require building the matrices. The matrices, $\mathbf{D}_{o,i}$ s, in the regularization correspond to discrete filters implementing the derivatives $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $\frac{\partial^2}{\partial x^2}$, $\frac{\partial^2}{\partial y^2}$ and $\sqrt{2}\frac{\partial}{\partial x\partial y}$. These filters can be directly applied on the images without building the matrices, which is obviously more efficient. We use the matrix formulation for derivatives also for

achieving notational convenience in describing the minimization method. In the latter part of our development, we will eliminate the need for building the matrix $\mathbf{H}$ as well.

A widely used non-quadratic regularization is the Total Variation (TV) [20, 25, 40, 41] and it is superior to the quadratic functional in its ability to preserve edges and it is robust to noise. The discrete total variation is given by

$$R_{TV}(\mathbf{p}_0) = \sum_{r=1}^N \sqrt{\sum_i (\mathbf{D}_{o,i} \mathbf{p}_0)_r^2} \quad (1.11)$$

Often a differentiable approximation of the total variation is used which can be written as

$$R_{TV}(\mathbf{p}_0) = \sum_{r=1}^N \sqrt{\epsilon + \sum_i (\mathbf{D}_{o,i} \mathbf{p}_0)_r^2} \quad (1.12)$$

where ϵ is a small positive number.

2 The proposed reconstruction method

2.1 Proposed Regularization functional

The photoacoustic images have high contrast due to the differential absorption of light in the near-infrared region by chromophores such as hemoglobin. Due to this, high values of initial pressure, $p_0(\mathbf{r})$, are sparsely distributed. Further, regions having high derivative values are also sparsely distributed. This pattern has also been observed in fluorescence images in the work presented in [34], where the regularization was constructed by adding an intensity term to second order derivatives. The combined point-wise cost goes into a logarithmic function and summed over all pixels. Here we replace the log by a fractional power, and write the proposed regularization as

$$R_{h,1}(\mathbf{p}_0, q) = \sum_{r=1}^N \left(\epsilon + \alpha (\mathbf{p}_0)_r^2 + (1 - \alpha) \sum_i (\mathbf{D}_{o,i} \mathbf{p}_0)_r^2 \right)^q, \quad s.t. \quad 0 < \alpha < 1, \quad (2.1)$$

where the weight $\alpha \in (0, 1)$ controls the relative penalization. The advantage of this modification is that it allows an optimization strategy that can efficiently handle non-convex cost function, which will be demonstrated later. Here, we choose $q < 0.5$ meaning that the resulting cost functional is non-convex. We also consider a variant of the above form, which is given below:

$$R_{h,2}(\mathbf{p}_0, q) = \alpha \sum_{r=1}^N \left(\epsilon + (\mathbf{p}_0)_r^2 \right)^q + (1 - \alpha) \sum_{r=1}^N \left(\epsilon + \sum_i (\mathbf{D}_{o,i} \mathbf{p}_0)_r^2 \right)^q. \quad (2.2)$$

For notational convenience in describing the minimization algorithm, we call q the sparsity index.

2.2 The complete cost functional

The initial pressure distribution in PAT is proportional to the fluence distribution and absorption coefficient in the tissue and hence the recovered PAT images should contain only non-negative

values in it. Therefore, a non-negativity constraint is imposed on the solution of the optimization problem. The modified optimization problem to be solved is given by

$$\hat{\mathbf{p}}_0 = \underset{\mathbf{p}_0 \geq 0}{\operatorname{argmin}} J(\mathbf{p}_0, q) \quad (2.3)$$

where

$$J(\mathbf{p}_0, q) = \|\mathbf{p} - \mathbf{H}\mathbf{p}_0\|_2^2 + \lambda R_{h,l}(\mathbf{p}_0, q) \quad l = 1 \text{ or } 2. \quad (2.4)$$

This constrained optimization problem can be efficiently handled by splitting-based optimization methods. However, such methods have slow convergence and yields reconstructions with artifacts when the regularization is non-convex. Here, we intend to develop a method based on traditional gradient-based iterative schemes, and for this purpose, we introduce an approximate unconstrained formulation. Specifically, we modify cost by adding quadratic penalty term to enforce positivity as given below:

$$J(\mathbf{p}_0, q) = \|\mathbf{p} - \mathbf{H}\mathbf{p}_0\|_2^2 + \lambda R_{h,i}(\mathbf{p}_0, q) + \lambda_p \|\mathcal{P}^-(\mathbf{p}_0)\|_2^2 \quad (2.5)$$

where

$$(\mathcal{P}^-(\mathbf{p}_0))_r = \begin{cases} 0 & \text{if } (\mathbf{p}_0)_r \geq 0 \\ (\mathbf{p}_0)_r & \text{if } (\mathbf{p}_0)_r < 0. \end{cases} \quad (2.6)$$

Note that the problems in the equations (2.3) and (2.5) are equivalent only when λ_p is arbitrarily large. However, setting λ_p to an arbitrarily large value will lead to numerical instability. Fortunately, through a series of reconstruction trials, we found that setting $\lambda_p = 10\lambda$ was sufficient to avoid significant negative values.

2.3 Proposed Algorithm

Here, we first develop the minimization algorithm for reconstruction with $R_{h,1}$ and then describe the modifications necessary to use $R_{h,2}$. We adapt the preconditioned gradient search for minimizing the cost function. To this end, we first need to write the expression of the gradient. For notational convenience, we use $\mathbf{x}$ in the place of $\mathbf{p}_0$. The gradient expression is given by

$$\nabla J(\mathbf{x}, q) = \mathbf{A}^{(\mathbf{x})}\mathbf{x} - \mathbf{H}^T \mathbf{p}_m, \quad (2.7)$$

where

$$\mathbf{A}^{(\mathbf{x})} = \mathbf{H}^T \mathbf{H} + \lambda \alpha \mathbf{W}^{(\mathbf{x})} + \lambda(1 - \alpha) \sum_i \mathbf{D}_{o,i}^T \mathbf{W}^{(\mathbf{x})} \mathbf{D}_{o,i} + \lambda_p \mathbf{N}^{(\mathbf{x})}. \quad (2.8)$$

Here $\mathbf{W}^{(\mathbf{x})}$ and $\mathbf{N}^{(\mathbf{x})}$ are diagonal matrices with j th diagonal element given by

$$\{\mathbf{W}^{(\mathbf{x})}\}_{jj} = q \left(\epsilon + \alpha (\mathbf{x})_j^2 + (1 - \alpha) \sum_i (\mathbf{D}_{o,i} \mathbf{x})_j^2 \right)^{q-1}, \quad (2.9)$$

$$\{\mathbf{N}^{(\mathbf{x})}\}_{jj} = 0.5(1 - \operatorname{sign}((\mathbf{x})_j)). \quad (2.10)$$

The superscript ‘ $(\mathbf{x})$ ’ in the diagonal matrices signify their dependence on $\mathbf{x}$, which makes the operation $\mathbf{A}^{(\mathbf{x})}\mathbf{x}$ a non-linear operation.

The proposed preconditioned gradient search proceeds as follows: given a current estimate of the minimum, $\mathbf{x}^{(k)}$, we update the estimate as $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \beta_k \mathbb{P}(\nabla J(\mathbf{x}^{(k)}, q))$, where $\mathbb{P}(\cdot)$ is

the preconditioning, and β_k is the step-size. To describe how the preconditioning is done, let $\mathbf{g}^{(k)} = \nabla J(\mathbf{x}^{(k)}, q)$ and $\hat{\mathbf{g}}^{(k)} = \mathbb{P}(\nabla J(\mathbf{x}^{(k)}, q))$. Then $\hat{\mathbf{g}}^{(k)}$ is determined by solving the linear system of equations $\mathbf{A}^{(\mathbf{x}^{(k)})} \hat{\mathbf{g}}^{(k)} = \mathbf{g}^{(k)}$, where

$$\mathbf{A}^{(\mathbf{x}^{(k)})} = \mathbf{H}^T \mathbf{H} + \lambda \alpha \mathbf{W}^{(\mathbf{x}^{(k)})} + \lambda(1 - \alpha) \sum_i \mathbf{D}_{o,i}^T \mathbf{W}^{(\mathbf{x}^{(k)})} \mathbf{D}_{o,i} + \lambda_p \mathbf{N}^{(\mathbf{x}^{(k)})}. \quad (2.11)$$

The idea behind this preconditioning is to mimic damped Newton approach as done in [42]. Here the matrix $\mathbf{A}^{(\mathbf{x}^{(k)})}$, which is the matrix involved in the gradient expression of equations (2.7), (2.8), works as an approximation for Hessian of $J(\cdot, q)$. To solve the equation $\mathbf{A}^{(\mathbf{x}^{(k)})} \hat{\mathbf{g}}^{(k)} = \mathbf{g}^{(k)}$, we use the well-known method of conjugate gradients (CG) and denote the operation by $\hat{\mathbf{g}}^{(k)} = CG(\mathbf{A}^{(\mathbf{x}^{(k)})}, \mathbf{g}^{(k)}, \epsilon_{cg})$ where ϵ_{cg} denotes the termination tolerance for CG iteration. Next, the step size β_k is chosen such that $\frac{J(\mathbf{x}^{(k+1)}, q)}{J(\mathbf{x}^{(k)}, q)} < \epsilon_s$ by means of a back-tracking procedure, where ϵ_s is an another user-defined tolerance for back-tracking. Specifically, starting with $\beta_k = 1$, the required β_k is determined by series of checks on the condition $\frac{J(\mathbf{x}^{(k+1)}, q)}{J(\mathbf{x}^{(k)}, q)} < \epsilon_s$ with iterated multiplication of β_k with a factor $\rho \in (0, 1)$. Finally, the iterative update on $\mathbf{x}^{(k)}$ is terminated upon the attainment of the condition $\frac{\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_2}{\|\mathbf{x}^{(k)}\|_2} < \epsilon_o$ where ϵ_o is a yet another user-defined tolerance. This iterative method is summarized in the panel **Algorithm 1**, where the first input $\mathbf{x}^{(0)}$ is an appropriate initialization. **Algorithm 1** calls the back-tracking line search method described above. This line search is summarized in the panel **Algorithm 2**.

Algorithm 1 Regularized Reconstruction Algorithm: $RR(\mathbf{x}^{(0)}, \mathbf{p}_m, \alpha, \lambda, \lambda_p, \epsilon_s, \epsilon_{cg}, \epsilon_o, \rho, q)$

Initialization : $k \leftarrow 0, r_k \leftarrow 1 + \epsilon_o$

while $r_k \geq \epsilon_o$ **do**

$\mathbf{g}^{(k)} \leftarrow \nabla J(\mathbf{x}^{(k)}, q)$ $\hat{\mathbf{g}}^{(k)} \leftarrow CG(\mathbf{A}^{(\mathbf{x}^{(k)})}, \mathbf{g}^{(k)}, \epsilon_{cg})$ $\beta_k \leftarrow LS(\mathbf{x}^{(k)}, \hat{\mathbf{g}}^{(k)}, \rho, \epsilon_s)$ $\mathbf{x}^{(k+1)} \leftarrow \mathbf{x}^{(k)} - \beta_k \hat{\mathbf{g}}^{(k)}$ $r_{k+1} \leftarrow \frac{\ \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\ _2}{\ \mathbf{x}^{(k)}\ _2}$ $k \leftarrow k + 1$	<p>▷ Use Eqs. (2.7), (2.8), (2.9), and (2.10)</p> <p>▷ Use Eq. (2.8). CG: Conjugate gradient method</p> <p>▷ Line search (Algorithm 2)</p>
--	---

end

return $\mathbf{y} = \mathbf{x}^{(k)}$

Algorithm 2 Line Search Algorithm: $\beta_k = LS(\mathbf{x}^{(k)}, \hat{\mathbf{g}}^{(k)}, \rho, \epsilon_s)$

Initialization : $\hat{\beta} \leftarrow 1, D \leftarrow J(\mathbf{x}^{(k)}), N \leftarrow J(\mathbf{x}^{(k)} - \hat{\beta} \hat{\mathbf{g}}^{(k)})$ ▷ Use Eq. (2.5)

while $\frac{N}{D} \geq \epsilon_s$ **do**

$\hat{\beta} \leftarrow \rho \hat{\beta}$ $N \leftarrow J(\mathbf{x}^{(k)} - \hat{\beta} \hat{\mathbf{g}}^{(k)})$	<p>▷ Use Eq. (2.5)</p>
--	------------------------

end

return $\beta_k = \hat{\beta}$

Algorithm 3 Graduated Non-Convexity Algorithm: GNC($\mathbf{p}_m, \alpha, \lambda, \lambda_p, \epsilon_s, \epsilon_{cg}, \epsilon_o, \rho, q, n_s$)

Initialization : $\bar{\mathbf{A}} \leftarrow \mathbf{H}^T \mathbf{H} + \lambda \alpha \mathbf{I} + \lambda(1 - \alpha) \sum_i \mathbf{D}_{o,i}^T \mathbf{D}_{o,i}$
 $\mathbf{y}^{(-1)} \leftarrow CG(\bar{\mathbf{A}}, \mathbf{H}^T \mathbf{p}_m, \epsilon_{cg})$
for $m = 0, 1, \dots, n_s$ **do**
 $q_m \leftarrow 0.5 - m \frac{0.5 - q}{n_s}$
 $\mathbf{y}^{(m)} = RR(\mathbf{y}^{(m-1)}, \mathbf{p}_m, \alpha, \lambda, \lambda_p, \epsilon_s, \epsilon_{cg}, \epsilon_o, \rho, q_m)$

▶ Algorithm 1

 $m \leftarrow m + 1$
end
return $\mathbf{y}^* = \mathbf{y}^{(n_s)}$

Since we choose q to be less than 0.5, the regularization functional is non-convex. Hence the cost function can have several local minima, and the tolerances used in the inner iterations (ϵ_{cg} and ϵ_s) might influence the actual minimum attained by the overall algorithm. Such minimum is likely to contain artifacts. To alleviate this problem, we adopt the well-known graduated non-convexity (GNC) approach [43]. Specifically, we introduce an outer loop in reconstruction method, where a series of n_s regularized reconstructions are performed with sparsity indices $\{q_m := 0.5 - m(0.5 - q)/n_s, m = 0, 1, \dots, n_s\}$. For each sparsity index q_m , the reconstruction is denoted by $\mathbf{y}^{(m)}$, and it is obtained by calling regularized reconstruction (Algorithm 1) with the initialization set to $\mathbf{y}^{(m-1)}$, which is the reconstruction corresponding to the previous sparsity index q_{m-1} . Note that, for $m = 0$, $q_0 = 0.5$, and for this value of q_0 , the cost is convex and hence the parameters ϵ_{cg} and ϵ_s do not affect the reconstruction obtained ($\mathbf{y}^{(0)}$). For subsequent calls for reconstruction with iteration index $m > 0$, since the sparsity index q_m is typically close to the index of previous reconstruction, q_{m-1} , the reconstruction $\mathbf{y}^{(m)}$ is likely to be close to $\mathbf{y}^{(m-1)}$. As a result, the sequence $\{\mathbf{y}^{(m)}, m = 0, 1, \dots, n_s\}$ is mostly determined by the sequence $\{q_m, m = 0, 1, \dots, n_s\}$. This will reduce the dependence of the final reconstruction on the parameters ϵ_{cg} and ϵ_s and will lead a good quality final reconstruction. For $m = 0$, we need an efficient initialization for calling the regularized reconstruction. We propose to use the result of minimizing $J(\cdot, q)$ with $q = 1$. With $q = 1$, $J(\cdot, q)$ is quadratic and its minimum is determined by the following linear system of equations:

$$[\mathbf{H}^T \mathbf{H} + \lambda \alpha \mathbf{I} + \lambda(1 - \alpha) \sum_i \mathbf{D}_{o,i}^T \mathbf{D}_{o,i}] \mathbf{y}^{(-1)} = \mathbf{H}^T \mathbf{p}_m. \quad (2.12)$$

We solve this again by calling the conjugate gradient method. The overall algorithm is given in Algorithm 3. Note that the possibility of using GNC method is the main advantage of replacing the logarithm used in the original formulation [34] by the fractional power q .

Now, we will specify the modifications required if $R_{h,2}$ is used. It is only required to change the operator $\mathbf{A}^{(x)}$ present in the equation (2.8). The operator corresponding to the regularization $R_{h,2}$ can be expressed as

$$\mathbf{A}^{(x)} = \mathbf{H}^T \mathbf{H} + \lambda \alpha \bar{\mathbf{W}}^{(x)} + \lambda(1 - \alpha) \sum_i \mathbf{D}_{o,i}^T \hat{\mathbf{W}}^{(x)} \mathbf{D}_{o,i} + \lambda_p \mathbf{N}^{(x)}. \quad (2.13)$$

where $\bar{\mathbf{W}}^{(x)}$ and $\hat{\mathbf{W}}^{(x)}$ are diagonal matrices with its elements given by

$$\{\bar{\mathbf{W}}^{(x)}\}_{jj} = q \left(\epsilon + (\mathbf{x}_j^2) \right)^{q-1}, \quad (2.14)$$

$$\{\hat{\mathbf{W}}^{(\mathbf{x})}\}_{jj} = q \left(\epsilon + \sum_i (\mathbf{D}_{o,i} \mathbf{x})_j^2 \right)^{q-1}. \quad (2.15)$$

2.4 Matrix-free implementation of the forward model

Finding the gradient $\mathbf{g}^{(k)}$ of the cost function and the search direction $\hat{\mathbf{g}}^{(k)}$ using CG involves the repeated application of $\mathbf{H}^T \mathbf{H}$; this is computationally very expensive and needs large amount of memory. For example, generating $\mathbf{H}$ for image size 512×512 corresponding to 128 transducers with each transducer taking 1024 samples results in 10^5 elements in $\mathbf{H}$, and thus requires 256 GB RAM [31]. In this work, we derive a formula for matrix free implementation of the proposed algorithm by doing on-the-fly computation of multiplication with $\mathbf{H}^T \mathbf{H}$ without explicitly constructing $\mathbf{H}$.

We first note that $\mathbf{H}$ is of the form $\mathbf{H} = [\mathbf{H}_1^T \mathbf{H}_2^T \cdots \mathbf{H}_M^T]^T$, where $\mathbf{H}_i$ represents the operation of obtaining samples from L transducers at i th time instant. Let $\mathbf{v}$ be some vector that undergo multiplication by $\mathbf{H}$. Let $v(\mathbf{r})$ be the image obtained by putting elements of $\mathbf{v}$ into image form. Then the operator that is the equivalent of multiplying $\mathbf{v}$ with $\mathbf{H}_i$ can be written by using the forward model of the equation (1.4) as

$$\mathbf{H}_i \mathbf{v} = \mathcal{P} \mathcal{F}^{-1} \{ \mathcal{F}(\mathcal{A} \mathbf{v}) \cos(c_0 |\mathbf{k}| t_i) \}, \quad (2.16)$$

where $\mathcal{P}$ is the operation that represents retrieval of samples from transducer locations $\{\mathbf{r}_s, s = 1, \dots, L\}$ into vector form, and $\mathcal{A}$ represents the operation of assembling the $N \times 1$ vector into an $N_x \times N_y$ image. Then, multiplication with $\mathbf{H}_i^T$ can be expressed as

$$\mathbf{H}_i^T \mathbf{u} = \mathcal{A}^a \mathcal{F}^{-1} \{ \cos(c_0 |\mathbf{k}| t_i) \mathcal{F} \{ \mathcal{P}^a \mathbf{u} \} \}, \quad (2.17)$$

where $\mathcal{A}^a$ and $\mathcal{P}^a$ are the adjoints of $\mathcal{A}$ and $\mathcal{P}$ respectively. $\mathcal{A}^a$ represents scanning a $N_x \times N_y$ image into a vector of size $N \times 1$ ($N = N_x N_y$). Further, assuming that the transducer locations $\{\mathbf{r}_s, s = 1, \dots, L\}$ are in a subset of image grid points, $\mathcal{P}^a$ becomes the operation of embedding an $L \times 1$ vector into an zero-image of size $N_x \times N_y$. Then, the operator-equivalent of $\mathbf{H}_i^T \mathbf{H}_i$ can be expressed as

$$\mathbf{H}_i^T \mathbf{H}_i \mathbf{v} = \mathcal{A}^a \mathcal{F}^{-1} \{ \cos(c_0 |\mathbf{k}| t_i) \mathcal{F} \{ \mathcal{P}^a \mathcal{P} \mathcal{F}^{-1} \{ \mathcal{F}(\mathcal{A} \mathbf{v}) \cos(c_0 |\mathbf{k}| t_i) \} \} \} \} \quad (2.18)$$

It can be shown that $\mathcal{P}^a \mathcal{P}$ becomes equivalent to multiplication by a binary images of 1s and 0s, with 1s corresponding to the transducer locations. Let $S(\mathbf{r})$ be this binary image.

$$\mathbf{H}_i^T \mathbf{H}_i \mathbf{v} = \mathcal{A}^a \mathcal{F}^{-1} \{ \cos(c_0 |\mathbf{k}| t_i) \mathcal{F} \{ S(\mathbf{r}) \mathcal{F}^{-1} \{ \mathcal{F}(\mathcal{A} \mathbf{v}) \cos(c_0 |\mathbf{k}| t_i) \} \} \} \} \quad (2.19)$$

This results in the following form for $\mathbf{y} = \mathbf{H}^T \mathbf{H} \mathbf{v}$

$$\mathbf{y} = \mathbf{H}^T \mathbf{H} \mathbf{v} = \sum_{i=1}^M \mathbf{H}_i^T \mathbf{H}_i \mathbf{v} \quad (2.20)$$

$$= \mathcal{A}^a \sum_{i=1}^M \mathcal{F}^{-1} \{ \cos(c_0 |\mathbf{k}| t_i) \mathcal{F} \{ S(\mathbf{r}) \mathcal{F}^{-1} \{ \mathcal{F}(\mathcal{A} \mathbf{v}) \cos(c_0 |\mathbf{k}| t_i) \} \} \} \} \quad (2.21)$$

In actual implementation, we do not build vectors. We used vector matrix notation in the description given above only for notational convenience. Every vector involved in the algorithm is

kept in image form. Specifically, the input to the above operator will be image form and output will also be in image form. If $v(\mathbf{r})$ is the $N \times N$ image equivalent of the $N^2 \times 1$ vector $\mathbf{v}$, and $y(\mathbf{r})$ is the image equivalent of the vector $\mathbf{y}$ the above operator can be represented by

$$y(\mathbf{r}) = \sum_{i=1}^M \mathcal{F}^{-1} \left\{ \cos(c_0 |\mathbf{k}| t_i) \mathcal{F} \left\{ S(\mathbf{r}) \mathcal{F}^{-1} \left\{ \mathcal{F}(v(\mathbf{r})) \cos(c_0 |\mathbf{k}| t_i) \right\} \right\} \right\} \quad (2.22)$$

3 Reconstruction results

3.1 Reconstruction from simulated data

We used three numerical phantoms that are commonly used to evaluate PAT reconstruction methods viz. blood vessel, Derenzo, and PAT, which are given in Figure 1. All were normalized to the size 128×128 with corresponding physical size of $12.8 \text{ mm} \times 12.8 \text{ mm}$ as model for generating the synthetic data. The synthetic data was generated as per the geometry given in the Figure 2, where

Figure 1. Numerical phantoms used for evaluation of the proposed method

the dotted circle, whose radius is 12 mm, denotes the trajectory of possible locations for transducers. The required image (image to be reconstructed) was defined to be on a larger grid of size 512×512 with the equivalent physical size of $51.2 \text{ mm} \times 51.2 \text{ mm}$. This was done to accommodate for the boundary effect caused by Fourier based convolutions represented in Equation (2.22). The forward data was generated using Equation (1.4) and added with Gaussian noise to form simulated data having SNR levels of 20 dB, 30 dB and 40 dB. Values for t in the model of Eq. (1.4) was chosen as $\{\delta_t i, i = 1, \dots, M\}$ with $1/\delta_t = 100 \text{ MHz}$ and $M = 1600$. The number of transducers to mimic the limited data scenario in our experiments was taken as 16, 32, 64, and 128. Note that most reported methods in the literature typically use not less than 128 transducers. The speed of the sound in the medium was assumed to be $1.5 \text{ mm}/\mu\text{s}$ and we considered the medium to be homogeneous with no dispersion or absorption of sound. We choose second-order derivatives for the regularization, i.e., we set $o = 2$ in the equation (2.1) and (2.2). The parameters λ was determined using the model itself, as done in most methods that focus on the design of regularization. Next, for positivity, we set $\lambda_p = 10\lambda$. Further, we found that $q = 0.25$ with $n_s = 10$ was adequate for all cases. All the tolerances ($\epsilon, \epsilon_s, \epsilon_{cg}$, and ϵ_o) were set to 10^{-6} . Further, for all test cases, setting $\alpha = 0.5$ gave good quality reconstructions. We choose a recent FISTA-based method [25] for comparison. For the blood vessel phantom, we conducted experiments with number of transducers set to 64, and 128. For the remaining phantoms, we did experiments with 16 and 32 as the number of transducers.

Figure 2. a) Schematic diagram of PA data acquisition geometry with ultrasound transducers (shown by dots) around the imaging region of size 12.8 mm $\times$ 12.8 mm. The computational imaging grid size is 51.2 mm $\times$ 51.2 mm.

For a given q , each reconstruction takes about 38 mins for the proposed method in Dell Precision T7820, whereas the FISTA-based methods takes about 30 mins.

In the first experiment, we generated measurements corresponding to 128 transducers with SNR of 20 dB. Figure 3 compares the reconstructed results obtained from this measurement set. It is clear that the proposed method with both forms of regularization yields better reconstruction over FISTA-based method, although the result with second form of regularization gives slightly inferior result. This is reflected by the improved SSIM score (0.7622 and 0.7548 against 0.7052). The proposed method was also able to reconstruct the peak amplitude of the initial pressure distribution (1 Pa) more accurately than FISTA algorithm as evident from the Figures 3.A, 3.B and 3.C, 3.D. Figure 4 shows the scanline based intensity profile of reconstructions in Figure 3. The intensity profiles show that our method closely follows the ground truth compared to the FISTA based method which is also confirmed by superior SSIM scores. Next, we consider two more noise levels viz. 30 dB and 40 dB. For this phantom, since the first form of proposed regularization outperforms the second, we present the reconstruction results only for the first form for these additional cases of input SNR. The SSIM scores corresponding to 30 dB, and 40 dB are 0.7636, and 0.7643 respectively, and the corresponding SSIM scores yielded by the FISTA-based method are 0.7370, and 0.7427. Note that, the proposed method yielded SSIM score of 0.7622 for 20 dB noise level, whereas the FISTA-based produced 0.7052. This indicates that the proposed method retrieves most of the quality that can be obtained from nearly noise-free data from the noisy data. On the other hand, the improvement in reconstruction yielded by FISTA-based method with respect to increase in the input SNR, is more gradual, as indicated by the SSIM scores 0.7052, 0.7370, and 0.7427 corresponding to input SNRs 20 dB, 30 dB, and 40 dB. Further, it is also clear that the difference in performance between the proposed and FISTA-based methods is higher for lower input SNR. Next, we repeated the above trials with 64 transducers instead of 128. The SSIM scores of the reconstructions obtained by proposed method with first form of regularization are 0.7588, 0.7623, and 0.7626 with corresponding input SNRs 20 dB, 30 dB, and 40 dB. On the other hand, the

reconstructions obtained by the FISTA-based method have SSIM scores of 0.6408, 0.7236, and 0.7437 for the same set of input SNRs.

Figure 3. Comparison of reconstructions obtained from simulated data with 128 transducers and 20 dB measurement noise. (A): Reference phantom model (the maximum initial Pressure rise is assumed to be 1Pa); (B): Reconstruction obtained by proposed method with form I regularization (SSIM score: 0.7622); (C) Reconstruction from FISTA-based method with λ chosen for best SSIM with SSIM score of 0.7052; (D): Reconstruction obtained by proposed method with form II regularization (SSIM score: 0.7548)

Figure 4. Scan line based intensity profiles of reconstructed images from Fig. 3.

In the next experiment, we generated measurements using derenzo phantom image corresponding to 32 transducers with SNR of 20 dB. Figure 5 compares the reconstructed results obtained from

Figure 5. Comparison of reconstructions obtained from simulated data with 32 transducers and 20 dB measurement noise. (A.1): Derenzo phantom model (the maximum initial Pressure rise is assumed to be 1Pa); (B.1): Reconstruction obtained by proposed method with form I regularization (SSIM score: 0.9896); (C.1) Reconstruction from FISTA-based method with λ chosen for best SSIM with SSIM score of 0.9148; (D.1): Reconstruction obtained by proposed method with form II regularization (SSIM score: 0.9947). A.2, B.2, C.2, and D.2 are cropped regions from A.1, B.1, C.1, and D.1.

this measurement set. Here too, the proposed method with both forms of regularization yields better reconstruction over FISTA-based method as seen from the Figures 5.A.1, 5.B.1, 5.C.1, and 5.D.1. Figures 5.A.2, 5.B.2, 5.C.2, and 5.D.2 display a cropped region from the images displayed in 5.A.1, 5.B.1, 5.C.1, and 5.D.1 for a closer view. The improvement seen in the displayed images is also reflected by the improved SSIM score (0.9896 and 0.9947 against 0.9148). Next, we considered two other noise levels, viz., 30 dB and 40 dB for the same phantom with the same number of transducers (32). For these cases, we present the reconstruction result only for the second form since the second form of proposed regularization outperforms the first for the current phantom. The SSIM scores of the reconstructions from data sets with input SNRs 30 dB, and 40 dB are 0.9994, and 0.9997 respectively, which are better than the corresponding SSIM scores of FISTA-based method which are 0.9820 and 0.9892 respectively. This once again proves that the proposed method retrieves most of the quality that can be obtained from nearly noise-free data from the noisy data. The performance between the proposed and FISTA is higher for lower input SNR in this experiment also. Next, we repeated the same set of trials with 16 transducers instead of 32. The proposed method outperformed the FISTA method for all SNRs levels in this case also as shown in the Figure 6. The quality of the reconstruction for FISTA method was heavily degraded for all SNRs (20dB, 30 dB and 40 dB) and the reconstructed images contained artifacts as seen from the Figure 6.D, Figure 6.E and Figure 6.F respectively. However, the quality of reconstruction was not much affected for the proposed method although number of transducers used is very less. The improved performance of

Figure 6. Reconstructed images from input datasets corresponding to 16 transducers for various noise levels. (A), (B), and (C): reconstructions obtained from the proposed method with second form of regularization corresponding to SNRs 20 dB, 30 dB, and 40 dB with SSIM scores 0.9830, 0.9974, and 0.9998 respectively. (D), (E), and (F): reconstructions obtained from the FISTA method corresponding to SNRs 20 dB, 30 dB, and 40 dB with SSIM scores 0.6828, 0.7255, and 0.7401.

the proposed method in comparison with the FISTA method is also reflected in the large difference of SSIM scores for all three SNR levels: the proposed method gives 0.9830, 0.9974, and 0.9998 for the inputs SNRs 20dB, 30 dB and 40 dB respectively, whereas the FISTA-based method gives 0.6828, 0.7255, and 0.7401. This results clearly demonstrates the robustness of our algorithm in the limited data scenario. The above observations was further verified by repeating the same pattern of experiments on the PAT phantom. Reconstruction results for this phantom along with the results of all other trials reported in this section are displayed **Table 1**. For visual comparison of performance evaluated using PAT phantom, we display the images corresponding to trials with 16 transducers in **Figure 7**, which clearly confirms the superiority of the proposed method.

Blood Vessel Phantom					Derenzo Phantom				PAT Phantom			
SNR (dB)	64 trans.		128 trans.		16 trans.		32 trans.		16 trans.		32 trans.	
	FBM	Ours	FBM	Ours	FBM	Ours	FBM	Ours	FBM	Ours	FBM	Ours
20	.640	.758	.705	.762	.682	.983	.914	.994	.691	.907	.805	.957
30	.723	.762	.737	.763	.725	.997	.982	.999	.774	.976	.976	.999
40	.743	.762	.742	.764	.740	.999	.989	.999	.780	.997	.985	.999

Table 1. SSIM scores of reconstruction with varying input noise levels and different number of transducers for all phantoms. FBM: FISTA-based method of Huang et al. [25]

Figure 7. Reconstructed images from input data sets corresponding to 16 transducers for various noise levels. (A), (B), and (C): reconstructions obtained from the proposed method with second form of regularization corresponding to SNRs 20 dB, 30 dB, and 40 dB with SSIM scores 0.9070, 0.9760, and 0.9979 respectively. (D), (E), and (F): reconstructions obtained from the FISTA method corresponding to SNRs 20 dB, 30 dB, and 40 dB with SSIM scores 0.6917, 0.7749, and 0.7808.

3.2 Reconstruction from real measured data

We use a triangular shaped physical phantom constructed using horse hair to generate real measured data. The details of the experimental set up can be found in Fig.2 of Ref. [44]. A Q-switched Nd:YAG laser operating at 532 nm delivered laser pulses with 5 ns width at 10 Hz repetition rate onto the sample. One uncoated Plano-concave lens (LC1715, Thorlabs) and four right-angle uncoated prisms (PS911, Thorlabs) were utilized to provide a light fluence of 9 mJ/cm² (< 20 mJ/cm² : ANSI safety limit). The hair phantom having the side-length and diameter of 10 and 0.15 mm, respectively, was attached to the pipette tips adhered on acrylic slab. For recording the PA data, a 2.25 MHz flat ultrasound transducer (Olympus-NDT, V306-SU) of 13 mm diameter active area and 70% nominal bandwidth was rotated continuously for 360 deg around the sample. A pulse amplifier (Olympus-NDT, 5072PR) first amplified and filtered the acquired PA signals and then a data acquisition (DAQ) card (GaGe, compuscope 4227) recorded the signals using a sampling frequency of 25 MHz. A sync signal from laser was used for synchronization of data acquisition with laser illumination.

The reconstructed PA imaging region has a size of 12.8 mm by 12.8 mm containing 128 by 128 pixels and data from 400 transducer positions is used to do the reconstruction. Since the actual values of initial pressure rise are unknown here, we have used following figure-of-merit as used in the reference [45] to compare different methods.

$$FOM = 20 \times \log_{10} \left(\frac{S}{n} \right) \quad (3.1)$$

Figure 8. Reconstructed images from horse hair phantom data using 400 transducers. (A): reconstruction from proposed method with first form of regularization (FOM 47.42 dB). (B): reconstruction from proposed method with second form of regularization (FOM 47.02 dB); (C): reconstruction from FISTA-based method (FOM 44.21 dB).

where S is peak intensity value of the reconstructed image and n is the standard deviation of the intensity.

Figure 8 compares reconstructed images for horse hair phantom data. For the proposed method, since we do not have the ground truth for evaluation, we present reconstruction results obtained by using both form of regularization. Figure 8.A shows the image reconstructed using the first form of regularization term and Figure 8.B shows the result obtained by using second form regularization. Figure 8.C shows the corresponding result obtained from FISTA based method. The proposed methods were able to give sharp images while giving a 3dB improvement in the FOM values when compared with the FISTA based result as given in figure.

4 Conclusions

A novel model based method that can reconstruct high quality PAT images from significantly reduced measurements was developed. The joint sparse nature of intensity and derivatives in PAT images was exploited to construct the regularization, and the reconstruction was formulated as the minimizer of this regularization along with data fidelity cost and a cost that quantifies the deviation from positivity. A novel computational method was developed to carry out the minimization. Further, a new computational formula was derived to implement the forward model and was used in the data fidelity term to reduce the memory requirement. The algorithm was compared against FISTA based method of Huang et al. [25] for various levels of reduction in the measured data (16, 32, 64 and 128 transducers) with various levels measurement noise (20, 30 and 40 dB). We

considered both real and simulated data sets for our experiments and the proposed method yielded superior reconstruction quality in all cases.

References

- [1] Wang L V and Yao J 2016 *Nature methods* **13** 627
- [2] Pramanik M, Ku G, Li C and Wang L V 2008 *Medical physics* **35** 2218–2223
- [3] Beard P 2011 *Interface focus* **1** 602–631
- [4] Zhou Y, Yao J and Wang L V 2016 *Journal of biomedical optics* **21** 061007
- [5] Upputuri P K and Pramanik M 2016 *Journal of Biomedical Optics* **22** 041006
- [6] Li L, Zhu L, Ma C, Lin L, Yao J, Wang L, Maslov K, Zhang R, Chen W, Shi J *et al.* 2017 *Nature biomedical engineering* **1** 0071
- [7] Han Y, Ding L, Ben X L D, Razansky D, Prakash J and Ntziachristos V 2017 *Opt. Lett.* **42** 979–982
- [8] Rosenthal A, Ntziachristos V and Razansky D 2013 *Current medical imaging reviews* **9** 318–336
- [9] Finch D and Patch S K 2004 *SIAM journal on mathematical analysis* **35** 1213–1240
- [10] Kunyansky L A 2007 *Inverse problems* **23** 373
- [11] Xu M and Wang L V 2002 *IEEE transactions on medical imaging* **21** 814–822
- [12] Xu M and Wang L V 2005 *Physical Review E* **71** 016706
- [13] Kalva S K and Pramanik M 2016 *Journal of Biomedical Optics* **21** 086011
- [14] Pramanik M 2014 *JOSA A* **31** 621–627
- [15] Xu Y and Wang L V 2004 *Physical review letters* **92** 033902
- [16] Burgholzer P, Matt G J, Haltmeier M and Paltauf G 2007 *Physical Review E* **75** 046706
- [17] Treeby B E, Zhang E Z and Cox B T 2010 *Inverse Problems* **26** 115003
- [18] Hristova Y, Kuchment P and Nguyen L 2008 *Inverse Problems* **24** 055006
- [19] Xu Y, Wang L V, Ambartsoumian G and Kuchment P 2004 *Medical physics* **31** 724–733
- [20] Arridge S, Beard P, Betcke M, Cox B, Huynh N, Lucka F, Ogunlade O and Zhang E 2016 *Physics in Medicine & Biology* **61** 8908
- [21] Buehler A, Rosenthal A, Jetzfellner T, Dima A, Razansky D and Ntziachristos V 2011 *Medical physics* **38** 1694–1704
- [22] Paltauf G, Viator J, Prah S and Jacques S 2002 *The Journal of the Acoustical Society of America* **112** 1536–1544
- [23] Schwab J, Pereverzyev Jr S and Haltmeier M 2018 *SIAM Journal on Numerical Analysis* **56** 160–184
- [24] Prakash J, Mandal S, Razansky D and Ntziachristos V 2019 *IEEE Transactions on Biomedical Engineering* 1–1 ISSN 0018-9294
- [25] Huang C, Wang K, Nie L, Wang L V and Anastasio M A 2013 *IEEE Transactions on Medical Imaging* **32** 1097–1110
- [26] Arridge S R, Betcke M M, Cox B T, Lucka F and Treeby B E 2016 *Inverse Problems* **32** 115012

- [27] Boink Y E, Lagerwerf M J, Steenbergen W, van Gils S A, Manohar S and Brune C 2018 *Physics in Medicine & Biology* **63** 045018
- [28] Saratoon T, Tarvainen T, Cox B and Arridge S 2013 *Inverse Problems* **29** 075006
- [29] Han Y, Tzoumas S, Nunes A, Ntziachristos V and Rosenthal A 2015 *Medical physics* **42** 5444–5452
- [30] Wang K, Huang C, Kao Y J, Chou C Y, Oraevsky A A and Anastasio M A 2013 *Medical physics* **40**
- [31] Rosenthal A, Razansky D and Ntziachristos V 2010 *IEEE transactions on medical imaging* **29** 1275–1285
- [32] Rosenthal A, Jetzfellner T, Razansky D and Ntziachristos V 2012 *IEEE transactions on medical imaging* **31** 1346–1357
- [33] Liu H, Wang K, Peng D, Li H, Zhu Y, Zhang S, Liu M and Tian J 2016 *IEEE transactions on medical imaging* **35** 2546–2557
- [34] Arigovindan M, Fung J C, Elnatan D, Mennella V, Chan Y H M, Pollard M, Branlund E, Sedat J W and Agard D A 2013 *Proceedings of the National Academy of Sciences* 201315675
- [35] Cox B T, Kara S, Arridge S R and Beard P C 2007 *The Journal of the Acoustical Society of America* **121** 3453–3464
- [36] Cox B and Beard P 2005 *The Journal of the Acoustical Society of America* **117** 3616–3627
- [37] Xu M and Wang L V 2006 *Review of scientific instruments* **77** 041101
- [38] Wang K, Ermilov S A, Su R, Brecht H P, Oraevsky A A and Anastasio M A 2011 *IEEE transactions on medical imaging* **30** 203–214
- [39] Wang K, Su R, Oraevsky A A and Anastasio M A 2012 *Physics in Medicine & Biology* **57** 5399
- [40] Tao M, Yang J and He B 2009 *TR0918, Department of Mathematics, Nanjing University*
- [41] Wang Y, Yang J, Yin W and Zhang Y 2008 *SIAM Journal on Imaging Sciences* **1** 248–272
- [42] Skariah D G and Arigovindan M 2017 *IEEE Transactions on Image Processing* **26** 4471–4482
- [43] Nikolova M, Ng M K and Tam C 2010 *IEEE Transactions on Image Processing* **19** 3073–3088
- [44] Awasthi N, Kalva S K, Pramanik M and Yalavarthy P K 2018 *Journal of Biomedical Optics* **23** 091413
- [45] Li L, Zhu L, Shen Y and Wang L V 2017 *Journal of biomedical optics* **22** 076017