Joint reconstruction of initial pressure distribution and spatial distribution of acoustic properties of elastic media with application to transcranial photoacoustic tomography

Joemini Poudel\textsuperscript{1} and Mark A Anastasio\textsuperscript{2,}\textsuperscript{*}

\textsuperscript{1} Department of Biomedical Engineering, Washington University in St. Louis, 1 Brookings Dr., St. Louis, MO, 63130, United States of America
\textsuperscript{2} Department of Bioengineering, University of Illinois at Urbana-Champaign, 1406 W Green St, Urbana, IL, 61801, United States of America

E-mail: maa@illinois.edu

Received 17 March 2020, revised 29 October 2020
Accepted for publication 5 November 2020
Published 3 December 2020

Abstract

Photoacoustic computed tomography (PACT) is an emerging computed imaging modality that exploits optical contrast and ultrasonic detection principles to form images of the photoacoustically induced initial pressure distribution within tissue. The PACT reconstruction problem corresponds to a time-domain inverse source problem, where the initial pressure distribution is recovered from the measurements recorded on an aperture outside the support of the source. A major challenge in transcranial PACT brain imaging is to compensate for aberrations in the measured acoustic data that are induced by propagation of the photoacoustic wavefields through the skull. To properly account for these effects, previously proposed image reconstruction methods for transcranial PACT require knowledge of the spatial distribution of the elastic parameters of the skull. However, estimating the spatial distribution of these parameters prior to the PACT experiment remains challenging. To circumvent this issue, in this work a method to jointly reconstruct the initial pressure distribution and a low-dimensional representation of the elastic parameters of the skull is developed and investigated. The joint reconstruction (JR) problem is solved by

\textsuperscript{*}Author to whom any correspondence should be addressed.

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
use of a proximal optimization method that allows constraints and non-smooth regularization terms. The proposed method is evaluated by use of large-scale three-dimensional (3D) computer-simulation studies that mimic transcranial PACT experiments.

Keywords: photoacoustic computed tomography, image reconstruction, transcranial imaging, elastic wave equation, joint image reconstruction

(Some figures may appear in colour only in the online journal)

1. Introduction

Photoacoustic computed tomography (PACT) is an emerging computed imaging modality that exploits optical contrast and ultrasonic detection principles to form images of the absorbed optical energy density within tissue (Kruger et al 1995, Kruger et al 1999, Xu et al 2011). In PACT, the object is irradiated with a short laser pulse. Under the condition of thermal confinement, the absorption of the optical energy within the object results in the generation of pressure waves via the photoacoustic (PA) effect (Xu and Wang 2006a, Oraevsky and Karabutov 2003, Wang and Wu 2012). These pressure waves propagate outside of the object and are subsequently detected by a collection of broadband ultrasound transducers located outside the support of the object. The goal of PACT in its canonical formulation is to reconstruct an image that represents a map of the photoacoustically-induced initial pressure distribution within the object from knowledge of the measured photoacoustically-induced pressure waves (Poudel et al 2019). The initial pressure distribution is proportional to the absorbed optical energy distribution within the object, which can reveal diagnostically useful information based on endogenous hemoglobin contrast or exogenous contrast if molecular probes are utilized (Jones et al 1980, Cheong et al 2003, Xu et al 2011).

The majority of the currently available PACT reconstruction methods are based on idealized imaging models that assume a lossless and acoustically homogeneous fluid medium (Kunyansky 2007, Finch and Rakesh 2004). This assumption, however, is grossly violated in transcranial PACT due to the presence of the skull. As result, images produced by conventional reconstruction methods can contain significant distortions and degraded spatial resolution (Schoonover et al 2012, Xu and Wang 2006b, Jin et al 2008, Xu et al 2011, Yang and Wang 2008, Nie et al 2011, Huang et al 2012).

To mitigate image artifacts caused by acoustic heterogeneities, heuristic methods such as the half-time and partial-time image reconstruction methods have been developed (Poudel et al 2017, Anastasio et al 2005a, Anastasio et al 2005b). This is a heuristic method that has been shown to mitigate distortions in the reconstructed image in certain applications. An attractive feature of such methods is that they do not require knowledge of the spatial distribution of acoustic parameters within the object. Such methods have been demonstrated to be effective for the case of fluid media containing weak acoustic heterogeneities; however, they are generally ineffective in the presence strong acoustic heterogeneities such as the skull (Poudel et al 2019). Moreover, truncation-based reconstruction methods generally require that pressure measurements are acquired on a closed measurement surface that surrounds the object (skull), in order to satisfy certain data-sufficiency requirements and microlocal stability. In 3D transcranial PACT measurement geometries, this cannot generally be achieved.

In addition, numerous image reconstruction methods that can account for (known) acoustic heterogeneities in a fluid medium have been proposed (Huang et al 2013, Jin et al 2008, Poudel et al 2019, Grün et al 2008, Burgholzer et al 2007, Hristova et al 2008, Treeby et al 2008).
However, bones, such as the skull, are more accurately described as elastic media that can support the propagation of both pressure and shear wavefields (Fry and Barger 1978). The methods described above that assume a fluid medium fail to account for shear wavefield propagation within the skull and mode conversion between pressure and shear wavefields at the skull interface (Schoonover and Anastasio 2011, White et al 2006, Fry 1978). The deleterious effects of neglecting these physical effects in transcranial PACT have been studied previously (Schoonover et al 2012, Xu and Wang 2006b, Jin et al 2008, Xu et al 2011, Yang and Wang 2008, Nie et al 2011, Huang et al 2012).

To further advance transcranial PACT, image reconstruction methods based on the 3D elastic wave equation have been developed (Mitsuhashi et al 2017, Poudel et al 2017, Poudel et al 2020, Javaherian and Holman 2018, Poudel et al 2019). It has been demonstrated that the images reconstructed using such methods can contain significantly reduced artifact levels compared to those reconstructed by use of methods that assume a fluid medium (Poudel et al 2017, Mitsuhashi et al 2017, Poudel et al 2020). However, a limitation of these methods is that they require the spatial distribution of the elastic properties of the skull to be estimated prior to the PACT experiment. In practice, such subject-specific information is generally difficult to obtain. Furthermore, as the elastic wave equation-based iterative image reconstruction methods are computationally intensive, it can be difficult to manually tune the parameters of the medium based on the reconstructed initial pressure distribution. A natural approach to mitigating this problem is to reformulate the reconstruction problem as a joint reconstruction (JR) problem in which the initial pressure distribution and elastic medium parameters are concurrently estimated by use of the measured PACT data. This idea is the topic of this article and is developed and explored as described below.

Several theoretical and computational studies of the JR problem for PACT in fluid media have been reported (Stefanov and Uhlmann 2012, Huang et al 2016, Liu and Uhlmann 2015, Kirsch and Scherzer 2012). Theoretical work on the JR problem in fluid media that neglects discrete sampling effects has established that the initial pressure distribution and speed of sound (SOS) distribution can be uniquely determined from the measured PACT data alone only for certain special cases (Hickmann 2010, Hristova et al 2008). However, the uniqueness of the JR problem for more general cases has not been established. Stefanov and Uhlmann considered the joint estimation problem for the linearized wave equation in fluid media and demonstrated that, in general, the SOS and initial pressure distributions could not both be stably recovered (Stefanov and Uhlmann 2013, Stefanov and Uhlmann 2012). A similar conclusion for the full non-linear problem was established via computer-simulation studies (Huang et al 2016). That study also demonstrated that, in certain cases, JR can lead to improved estimates of the initial pressure distribution as compared to conventional methods, even if the acoustic parameters were not estimated accurately (Huang et al 2016).

The idea of stabilizing the JR problem by use of a low-dimensional parameterized representation of the spatial distribution of the acoustic properties of a fluid medium has been proposed (Matthews et al 2018, Zhang et al 2008). It has been demonstrated that the initial pressure distribution and a parameterized version of the SOS distribution can be jointly estimated by minimizing an objective function consisting of a data fidelity term and a regularization term by use of a gradient-based alternating minimization approach (Matthews et al 2018). Acoustic wave propagation was modeled by use of the 2D acoustic wave equation and the adjoint state method was employed to compute the gradients with respect to the initial pressure and acoustic medium parameters. Numerical studies confirmed the robustness of the proposed method when accurate low-dimensional representations of the SOS distribution were employed. A mathematical explanation for the effectiveness of the approach has also been provided (Acosta 2019).
In this work, inspired by the parameterized JR method, we develop and investigate a JR method for 3D transcranial PACT. The formulation of the previously developed parameterized JR problem for PACT with 2D fluid media is extended in several key ways. Firstly, the forward model is replaced with a 3D elastic wave equation-based model that allows for wave propagation in elastic solids. Secondly, motivated by the geometry of a subject’s skull that is visible in adjunct x-ray CT or MRI data, low-dimensional parameterizations of the elastic properties of the skull are introduced. Thirdly, a memory-efficient implementation of the adjoint state method that utilizes binomial checkpointing is utilized for computing gradients of the cost functional with respect to the elastic model parameters. The JR problem is solved by use of a proximal optimization method that allows constraints and non-smooth regularization terms. The computational methods were implemented by use of parallel GPU-based computer code. The proposed JR is evaluated by use of large-scale three-dimensional (3D) computer-simulation studies that mimic transcranial PACT experiments.

The remainder of the manuscript is organized as follows. In section 2, background information on the imaging model and existing image reconstruction methods is provided. In section 3, the proposed parameterized JR method for 3D transcranial PACT is introduced. A description of the computer-simulation studies and the results of these studies are presented in sections 4 and 5, respectively. Finally, the article concludes with a summary in section 6.

2. Background

2.1. Imaging physics

Let the photoacoustically-induced stress tensor at location \( \mathbf{r} \in \mathbb{R}^3 \) and time \( t \geq 0 \) be defined as

\[
\sigma(\mathbf{r}, t) \equiv \begin{bmatrix}
\sigma_{11}(\mathbf{r}, t) & \sigma_{12}(\mathbf{r}, t) & \sigma_{13}(\mathbf{r}, t) \\
\sigma_{21}(\mathbf{r}, t) & \sigma_{22}(\mathbf{r}, t) & \sigma_{23}(\mathbf{r}, t) \\
\sigma_{31}(\mathbf{r}, t) & \sigma_{32}(\mathbf{r}, t) & \sigma_{33}(\mathbf{r}, t)
\end{bmatrix},
\]

where \( \sigma_{ij}(\mathbf{r}, t) \) represents the stress in the \( i \)th direction acting on a plane perpendicular to the \( j \)th direction. Additionally, let \( p_0(\mathbf{r}) \) denote the photoacoustically-induced initial pressure distribution within the object, referred to as the object function, and let \( \dot{\mathbf{u}}(\mathbf{r}, t) \equiv (\ddot{u}^1(\mathbf{r}, t), \ddot{u}^2(\mathbf{r}, t), \ddot{u}^3(\mathbf{r}, t)) \) represent the vector-valued acoustic particle velocity. For conciseness, the velocity \( \partial_t \mathbf{u} \) and the acceleration \( \partial_t \dot{\mathbf{u}} \) are represented by \( \dot{\mathbf{u}} \) and \( \dot{\dot{\mathbf{u}}} \) respectively. Here, \( \rho(\mathbf{r}) \), \( \lambda(\mathbf{r}) \) and \( \mu(\mathbf{r}) \) represent the spatial distribution of the medium’s density and the Lamé parameters, respectively. The compressional and shear wave propagation speeds are given by \( c_p(\mathbf{r}) = \sqrt{\frac{\lambda(\mathbf{r}) + 2\mu(\mathbf{r})}{\rho(\mathbf{r})}} \) and \( c_s(\mathbf{r}) = \sqrt{\frac{\mu(\mathbf{r})}{\rho(\mathbf{r})}} \). The object function \( p_0(\mathbf{r}) \) and all functions that describe the elastic isotropic medium are assumed to be bounded with compact supports.

The acoustic absorption of the skull is described by a diffusive absorption model (Moczo et al. 2007) in this study. The diffusive absorption model assumes that the wavefield absorption is independent of temporal frequency; this model does not accurately describe the power law based frequency dependent absorption characteristics of soft tissue and bone (Szabo 1994, Treeby and Cox 2014). This model, however, can be sufficiently accurate in cases where the bandwidth of the photoacoustic signals is limited.

In a heterogeneous isotropic elastic medium with a diffusive acoustic absorption distribution \( \alpha(\mathbf{r}) \), the propagation of \( \dot{\mathbf{u}}(\mathbf{r}, t) \) and \( \sigma(\mathbf{r}, t) \) can be modeled by the following two coupled...
the discrete forward operator described below will be employed. Additional details can be
wavefield is discretized temporally and spatially at specific transducer locations. In this work,
found in a previous study that utilized the same operator (Mitsuhashi et al. 1998, Alterman and Jr 1968):

\[ \partial_t \mathbf{u}(\mathbf{r}, t) + \alpha(\mathbf{r}) \mathbf{u}(\mathbf{r}, t) = \frac{1}{\rho(\mathbf{r})} \nabla \cdot \mathbf{\sigma}(\mathbf{r}, t) \]  
\[ \partial_t \mathbf{\sigma}(\mathbf{r}, t) = \lambda(\mathbf{r}) \mathbf{tr}(\nabla \mathbf{u} \mathbf{u}(\mathbf{r}, t)) \mathbf{I} + \mu(\mathbf{r})(\nabla \mathbf{u} \mathbf{u}(\mathbf{r}, t) + \nabla \mathbf{u} \mathbf{u}(\mathbf{r}, t)^T), \]  

subject to the initial conditions

\[ \mathbf{\sigma}_0(\mathbf{r}) \equiv \mathbf{\sigma}|_{t=0} = -\frac{1}{3} p_0(\mathbf{r}) \mathbf{I}, \quad \mathbf{u}|_{t=0} = 0. \]  

Here, \( \mathbf{tr}(\cdot) \) is the operator that calculates the trace of a matrix, \( \mathbf{I} \in \mathbb{R}^{3 \times 3} \) is the identity matrix, and \( \nabla \cdot \) is the divergence operator. In equation (2c), \( p_0(\mathbf{r}) \) is assumed to be compactly supported in a fluid medium where the shear modulus \( \mu(\mathbf{r}) = 0 \). Hence, the formulation of the initial value condition assumes that there are no optical absorbers inside the elastic solid.

2.2. Forward model in continuous and discrete forms

Let \( p(\mathbf{r}) \equiv \mathbf{tr}(\mathbf{\sigma}) \in L^2(\Omega \times [0, T]) \). The forward model can be described by a continuous-to-continuous (C–C) mapping as:

\[ p|_{\partial \Omega \times (0, T]} = \mathcal{M} \mathcal{H} p_0, \]  

where \( p|_{\partial \Omega \times [0, T]} \in L^2(\partial \Omega \times [0, T]) \) denotes the measured pressure wavefield data, \( \partial \Omega \subset \mathbb{R}^3 \) denotes a measurement aperture that encloses the domain \( \Omega \) and \( p_0 \in L^2(\Omega) \) is the sought-after photoacoustically-induced initial pressure distribution. The linear operator \( \mathcal{H} : L^2(\Omega) \mapsto L^2(\Omega \times [0, T]) \) maps \( p_0 \) to the pressure wavefield \( p \) via the initial value problem in equation (2), and the operator \( \mathcal{M} \) is the restriction of \( \mathcal{H} \) to \( \partial \Omega \times [0, T] \). In practice, the detected pressure wavefield is discretized temporally and spatially at specific transducer locations. In this work, the discrete forward operator described below will be employed. Additional details can be found in a previous study that utilized the same operator (Mitsuhashi et al. 2017).

Let \( \mathbf{p} \in \mathbb{R}^{N_L} \) denote the discretized pressure signal, where \( N_t \) represents the number of transducers and \( L \) represents the total number of discrete temporal samples. In the forward model described in equation (3), the effects of the acousto-electric impulse response as well as the spatial impulse response of the transducer have been ignored. However, the transducer responses can be incorporated into the imaging model readily (Huang et al. 2013).

To obtain a discrete-to-discrete (D–D) imaging model for use with an optimization-based reconstruction method, finite-dimensional approximate representations of the object function \( p_0(\mathbf{r}) \) and the material parameters \( \rho(\mathbf{r}), \alpha(\mathbf{r}), \lambda(\mathbf{r}), \) and \( \mu(\mathbf{r}) \) need to be introduced, whose forms are prescribed by the numerical method employed to solve the wave equation. In this study, the numerical method employed to solve the wave equation in equation (2) was a 10th-order staggered grid finite difference time domain (FDTD) scheme. Note that for the staggered finite-difference (FD) scheme, the material properties, stress, and particle velocity functions are sampled at different points of a staggered FDTD cell. Let \( \rho, \alpha, \lambda, \mu \), and \( p_0 \in \mathbb{R}^N \) be the finite dimensional representations of \( \rho(\mathbf{r}), \alpha(\mathbf{r}), \lambda(\mathbf{r}), \mu(\mathbf{r}) \) and \( p_0(\mathbf{r}) \), respectively, where \( N \) is the total number of grid points on the 3D grid.

For a given \( \rho, \alpha, \lambda, \mu \), and \( p_0 \), the FDTD method for solving the elastic wave equation can be described in operator form as

\[ \mathbf{p} = \mathcal{M} \mathcal{H}(\mathbf{c}) p_0. \]
where $c \in \mathbb{R}^{4N \times 1}$ is a vector that represents the acoustic properties of the elastic medium and is given by $c \equiv [\rho, \alpha, \lambda, \mu]^T$. $\mathbb{H}(c) \in \mathbb{R}^{NL \times N}$ is the discrete approximation of the wave operator $\mathcal{H}$ for a fixed $c$ that solves the initial value problem defined in equation (2), and $M \in \mathbb{R}^{NL \times NL}$ is a sampling matrix that maps pressure data sampled on the computational grid onto the transducer locations. The transducers were assumed to be point-like and thus when the receiver and grid point locations do not coincide, an interpolation method is required. The elements of $M$ were chosen such that trilinear interpolation is performed. The goal of image reconstruction in a discrete setting is to determine an estimate of $p_0$ given the measured data $p$ and the forward model in equation (4).

2.3. Reconstruction of initial pressure distribution when the elastic medium is known

For a fixed $c$, an estimate of $p_0$ can be obtained as (Poudel et al 2020, Poudel et al 2019)

$$\hat{p}_0 = \arg\min_{p_0 \geq 0} F(p_0, c) + \gamma R(p_0),$$

(5)

where

$$F(p_0, c) = \frac{1}{2} \| p - M\mathbb{H}(c)p_0 \|_2^2,$$

(6)

with $p$ denoting the recorded pressure data. The non-negativity constraint $p_0 \geq 0$ is imposed because negative initial pressure distribution values are nonphysical. The regularization term $R(p_0)$ incorporates a priori information about the sought-after initial pressure distribution with $\gamma$ denoting the regularization parameter that controls the relative weight of the data fidelity and regularization terms. The regularization function in the studies below is chosen to be the total-variation (TV) seminorm (Chambolle et al 2010)

$$R(p_0) = \| p_0 \|_{TV}$$

$$= \sum_{n=1}^{N} \left\{ (\| p_0 \|_{n} - [p_0]_{n,1-})^2 + (\| p_0 \|_{n} - [p_0]_{n,2-})^2 + (\| p_0 \|_{n} - [p_0]_{n,3-})^2 \right\}^{\frac{1}{2}}.$$

(7)

Here, $[p_0]_{n}$ denotes the $n$th grid node, and $[p_0]_{n,1-}$, $[p_0]_{n,2-}$, and $[p_0]_{n,3-}$ are the neighboring nodes before the $n$th node along the first, second and third dimension, respectively. The fast iterative shrinkage/thresholding algorithm (FISTA) with backtracking linesearch and adaptive restart was employed to solve the optimization problem in equation (5) (O’donoghue and Candes 2015, Huang et al 2013, Beck and Teboulle 2009b, 2009a).

3. Parameterized joint reconstruction (JR)

A JR method for transcranial PACT that circumvents the need to know the elastic properties of the skull is described below.

3.1. Parameterized representation of the elastic properties of the skull

To constrain and stabilize the JR problem, a low-dimensional parameterized representation of the spatially variant elastic properties of the skull is considered. Consider the vector $c \equiv [\rho, \alpha, \lambda, \mu]^T \in \mathbb{R}^{4N \times 1}$ that represents the concatenation of the four material properties defined
on a 3D grid as described above. The vector \( c_p \equiv [\rho_p, \alpha_p, \lambda_p, \mu_p]^T \in \mathbb{R}^{4Q \times 1} \) represents the parameterized representation of the elastic properties of the medium, where \( Q < N \). Here, the subscript \( p \) is used to represent the parameterized representation of the elastic properties of the medium. The original and parameterized elastic properties of the medium can be related as

\[
\begin{align*}
\mathbf{c} &= \Phi \mathbf{c}_p, \\
\Phi \in \mathbb{R}^{4N \times 4Q} &\text{ is a real-valued binary matrix defined as}
\end{align*}
\]

\[
[\Phi]_{i,j} = \begin{cases} 
1 & \text{if } i \in I_j \\
0 & \text{otherwise}
\end{cases},
\]

where \( I_j \) denote the set of indices of pixels that correspond to the \( j \)th parameter value, and \( [\Phi]_{i,j} \) denotes the \((i, j)\) element of the matrix \( \Phi \).

### 3.2. Formulation of JR problem

By use of the parameterization in equation (8), the JR problem can be formulated as

\[
\left( \hat{p}_0, \hat{c}_p \right) = \arg\min_{p_0 \geq 0, c_p \in S_p} F(p_0, \Phi c_p) + \gamma \mathcal{R}(p_0),
\]

The quantities \( \hat{p}_0, \hat{c}_p, \) and \( \hat{c} \equiv \Phi \hat{c}_p \) denote estimates of \( p_0, c_p, \) and \( c \), respectively. The estimate \( \hat{c}_p \) in equation (10) is constrained to lie in the closed, convex set

\[
S_p = \{ c_p : [c_p]_j \leq c^\text{max}_j \text{ and } [c_p]_j \geq c^\text{min}_j \},
\]

where \( c^\text{max}_j \) and \( c^\text{min}_j \) represent the maximum and minimum value associated with the \( j \)th parameter. The computation of the gradient \( \nabla_{c_p} F(p_0, \Phi c_p) \) can be accomplished by use of the adjoint state method described in appendix A. To mitigate the significant computational and memory burden involved in this computation, a binomial checkpointing method can be employed (Symes 2007, Yang et al 2016, Griewank and Walther 2000). Given \( \nabla_{c_p} F(p_0, \Phi c_p) \), the gradient of the data fidelity term with respect to \( c_p \) can be expressed as

\[
\nabla_{c_p} F(p_0, \Phi c_p) = \Phi^T \nabla_c F(p_0, \Phi c_p).
\]

For the choice of \( \Phi \) defined in equation (9), one obtains

\[
[\nabla_{c_p} F(p_0, \Phi c_p)]_j = \sum_{i \in I_j} [\Phi^T \nabla_c F(p_0, \Phi c_p)]_i.
\]

According to equation (13), the magnitude of the \( j \)th component of \( \nabla_{c_p} F(p_0, \Phi c_p) \) will depend on the cardinality of the indicator set \( I_j \), which represents the number of pixels in \( c \) whose values are specified by the \( j \)th component of \( c_p \). This can render the gradient poorly scaled, resulting in a slow convergence. To mitigate this, a diagonal, positive definite scaling matrix is introduced and defined as

\[
[B]_j = \frac{1}{|I_j|}.
\]
Algorithm 1. Parameterized joint reconstruction method.

Input: \( c_p^{(0)}, p_0^{(0)}, g, \Phi, \epsilon \).
Output: \( \hat{c}_p, \hat{p}_0 \).
1: \( k = 0 \), where \( k \) is the iteration number.
2: while \( \epsilon < \epsilon_F \) and \( k > 0 \) do
3: \( k = k + 1 \), where \( k \) is the iteration number.
4: \( p_0^{(k+1)} \leftarrow \text{prox}_{\gamma R}(p_0^{(k)} - \hat{c}_p^{(k)} G_p^{(k)}) \).
5: \( c_p^{(k+1)} \leftarrow \text{prox}_{\gamma R}(c_p^{(k)} - \hat{c}_p^{(k)} G_p^{(k)}) \).
6: end while

This matrix can be employed to generate a new search direction \( B \nabla c_p F(p_0, \Phi c_p) \) that represents the average gradient over the pixels associated with each component of \( c_p \). This new search direction is always a descent direction because \( B \) is positive definite by construction. Other choices for \( B \) could result in even faster convergence; however, the choice given by equation (14) requires little additional computational cost and ensures that the magnitude of each gradient component does not depend strongly on the number of pixels corresponding to each parameter.

3.3. Reconstruction algorithm

Equation (10) can be solved by use of the weighted proximal gradient descent method summarized in algorithm 1. This algorithm contains two major parts, one in which \( p_0 \) is updated via a proximal gradient descent step and one in which \( c_p \) is updated via a weighted gradient descent step. This is possible since the constraint and non-smooth regularization function \( R \) only involve \( p_0 \). For brevity, the following notation is employed to describe the algorithm:

\[
\begin{align*}
F^{(k)} &\equiv F(p_0^{(k)}, \Phi c_p^{(k)}), \\
C^{(k)} &\equiv F(p_0^{(k)}, \Phi c_p^{(k)}) + \gamma R(p_0^{(k)}), \\
C(p_0, c_p) &\equiv F(p_0, \Phi c_p) + \gamma R(p_0), \\
G_p^{(k)} &\equiv \nabla p_0 F(p_0^{(k)}, \Phi c_p^{(k)}), \\
G_p^{(k)} &\equiv B \nabla c_p F(p_0^{(k+1)}, \Phi c_p^{(k)}) .
\end{align*}
\]  

where the superscript \( k \) refers to the \( k \)th iteration of the algorithm. In addition, the notation \( \text{prox}_{iR} (\cdot) \) denotes the proximal operator of \( iR \) and is defined as

\[
\text{prox}_{iR}(v) \equiv \arg \min_x \left( R(x) + \frac{1}{2i} \| x - v \|_2^2 \right) .
\]
A two-parameter line search method, described in appendix C, can be employed to choose separate scalar step sizes for updating the initial pressure and sound speed distributions.

The action of the projection operator \( \mathcal{P}_{c_p} \) in algorithm 1 on an element of the vector \( c_p \) is defined as

\[
\mathcal{P}_{c_p} [c_p]_i = \begin{cases} 
& c_i^{\text{max}} \quad \text{if } [c_p]_i > c_i^{\text{max}}, \\
& c_i^{\text{min}} \quad \text{if } [c_p]_i < c_i^{\text{min}}, \\
& [c_p]_i \quad \text{otherwise}
\end{cases}
\]

(17)

4. Description of computer-simulation studies

Three-dimensional computer-simulation studies were performed to validate the proposed JR algorithm.

4.1. Methods

4.1.1. Parameterized representation of the skull model. Developing a low-dimensional acoustic representation of the skull’s elastic properties that permits accurate JR is a challenging task. In this study, two different low-dimensional representations of the skull were employed to validate the proposed JR algorithm.

The first low-dimensional skull model employs a three layer model, where the skull is made up of three distinct layers: the inner table, the outer table and the diploë layer (Fry and Barger 1978, Estrada et al 2018, Chang et al 2016). It is assumed that both the inner and outer table are made up of cortical bone that have distinct acoustic and elastic properties compared to the diploë layer and the background fluid medium (Estrada et al 2018, Fry and Barger 1978, Chang et al 2016). Furthermore, the diploë layer is comprised of a layer of trabecular spongy bone that is described by another distinct set of acoustic parameter values (Estrada et al 2018, Fry and Barger 1978, Chang et al 2016). Hence, in this form of parameterization, it is assumed that \( Q = 3 \) and \( c_p \in \mathbb{R}^{12 \times 1} \).

The mapping function \( \Phi \) in equation (9) employed for this low dimensional skull model was generated based on 3D x-ray CT images of a human skull. The CT images from an intact human skull were employed to infer the thickness and contour of the inner table, outer table and the diploë layer. To extract the contour and location of the skull layers from CT images, a segmentation algorithm was employed. The segmentation algorithm generated a mask specifying the location of the inner table, outer table and the diploë layer within the 3D volume. A 2D slice of the CT image acquired from the human skull and the corresponding mask generated by use of the segmentation algorithm are shown in figures 1(a) and (b), respectively. Using the 3D segmented CT skull, the mapping function \( \Phi \) for \( Q = 3 \) was established.

The second low-dimensional skull model also assumes that the skull can be characterized by a three layer model. However, in this model, it is assumed that the cortical bone comprising the inner table and outer table can be further divided into distinct sub-types, the frontal cortical bone, the left parietal cortical bone and the right parietal cortical bone, as illustrated by figure 2(a). In figure 2(a), the maximum amplitude projection of a 3D CT image of the skull along the z-direction is shown. From figure 2(a), one can observe the presence of three bones types fusing along the skull sutures. For illustration purposes, the three cortical bones are segmented and are shown in figure 2(b). The cortical bones are now described as the frontal cortical bone, left parietal cortical bone and right parietal cortical bone. Each of these bone types are assigned distinct acoustic and elastic parameter values. Similar to the \( Q = 3 \) model, it is also assumed that the diploë layer is comprised of a layer of trabecular spongy bone that is
described by another distinct set of acoustic parameter values. Hence, for this low-dimensional representation of the skull, it is assumed that $Q = 5$ and $c_p \in \mathbb{R}^{20 \times 1}$.

Similar to the $Q = 3$ skull model, the mapping function $\Phi$ employed in the $Q = 5$ was generated based on 3D x-ray CT images of a human skull. The CT images from an intact human skull were employed to infer the thickness and contour of the frontal cortical bone, left parietal cortical bone, right parietal cortical bone and the diploë layer. In order to extract the contour and location of the skull layers from CT images, a segmentation algorithm was employed. The segmentation algorithm generated a mask specifying the location of the frontal cortical bone, left parietal cortical bone, right parietal cortical bone and the diploë layer within the 3D volume. A 2D slice of the CT image acquired from the human skull and the corresponding mask generated by use of the segmentation algorithm are shown in figures 3(a) and (b), respectively.
Figure 3. (a) A 2D slice through the CT image obtained from a skull. (b) The $Q = 5$ skull model mask generated by use of the segmentation algorithm. The diploë layer is represented by the yellow color, the frontal cortical bone is represented by the blue color, the left parietal cortical bone is represented by the red color, and the right parietal cortical bone is represented by the green color.

Figure 4. The maximum intensity projections of the numerical phantom that represents the 3D initial pressure distribution along (a) $x$-axis, (b) $y$-axis and (c) $z$-axis, respectively.

4.1.2. Initial pressure distribution phantom. A numerical blood vessel structure extracted from magnetic resonance angiography (MRA) data was employed as a numerical phantom that represented $p_0$. The blood vessel phantom, shown in figure 4, was placed 6 mm below the inner surface of the skull.

4.1.3. Imaging system. The 3D hemispherical measurement system shown in figure 5(a) was assumed in the computer-simulation studies. The measurement system consisted of a total 38 400 ideal, point-like transducers that were distributed evenly in a spiral configuration over a hemispherical surface of radius 115.2 mm. The position of the measurement system with respect to the skull is shown in figure 5(b).

4.1.4. Simulation of pressure data. In order to avoid inverse crime (Colton 2013), the measured pressure data were simulated with different temporal and spatial sampling rates than those employed for image reconstruction. When generating the measured data, the simulation
grid consisted of a 3D Cartesian grid with 896 $\times$ 896 $\times$ 448 pixels and a pixel size of 0.3 mm. The pressure data were recorded for 7000 time steps at a temporal sampling rate of 100 MHz. Additive Gaussian white noise with zero mean and a standard deviation of 10 percent of the maximum pressure amplitude was added to the measured data. A MPI-based FDTD numerical wave solver (Mitsuhashi et al 2017) was employed to simulate the measured pressure data.

Two sets of pressure data were generated for the simulation studies. The first set of pressure data were generated using the initial pressure distribution shown in figure 4. Moreover, the spatial distribution of the acoustic properties of the skull employed for the generation of first set of pressure data correspond to the $Q = 3$ skull model. In this skull model, the diploë layer, cortical bone layers and the background fluid media were assigned distinct acoustic property values. The second set of pressure data were generated using the same initial pressure distribution. However, the pressure data was generated by use of the $Q = 5$ skull model. In this skull model, the diploë layer, the left parietal cortical bone, the right parietal cortical bone, the frontal cortical bone and the background fluid media were assigned distinct acoustic property values. The acoustic property values assigned to each of the parameters to generate the two sets of forward pressure data are summarized in table 1.

4.1.5. Image reconstruction. During image reconstruction, the computational grid corresponded to a 3D Cartesian grid with 448 $\times$ 448 $\times$ 224 pixels and a pixel size of 0.6 mm. The pressure wavefield was simulated for 3500 time steps with a temporal sampling rate of 50 MHz. During the image reconstruction process, the pressure wavefield was simulated by use of a MPI-based FDTD numerical wave solver (Mitsuhashi et al 2017), implemented using NVIDIA’s CUDA framework (CUDA C Programming Guide 2015). The image reconstruction algorithms were implemented by use of four NVIDIA V100 GPU cards. The gradient of the cost functional with respect to $p_0$ was computed using the action of the matched adjoint operator, while the gradient of the cost functional with respect to $c$ was computed using the adjoint state method, described in appendix A.

4.1.6. Reconstruction of $p_0$ when the acoustic properties of the skull are known. The initial pressure distribution was reconstructed for a fixed $c \equiv \Phi c_p$ for both the $Q = 3$ and $Q = 5$ cases specified in table 2. The values for $c_p$ specified in table 2 were also used as initial guess for

---

Figure 5. (a) The map of transducer locations employed for the simulation studies. (b) The position of the skull relative to the transducer locations.
Table 1. The values assigned to each parameter in the low-dimensional skull model that were used to generate the pressure data. The values are based on experimental studies that measured the acoustical properties of the skull (Chang et al 2016, Motherway et al 2009, Fry and Barger 1978, Auperrin et al 2014, McElhaney et al 1970). The density $\rho$ is given in units of $\frac{kg}{m^3}$, the absorption $\alpha$ is given in units of $\frac{1}{\mu s}$ and the longitudinal speed $c_l$, shear speed $c_s$ are given in units of $\frac{mm}{\mu s}$.

|                | Q = 3                      | Q = 5                      |
|----------------|----------------------------|----------------------------|
| Diploë layer   | $\rho = 1000$, $\alpha = 0.75$,  | $\rho = 1000$, $\alpha = 0.75$,  |
|                | $c_l = 2.0785$, $c_s = 1.4$  | $c_l = 2.0785$, $c_s = 1.4$  |
| Cortical bone  | $\rho = 1850$, $\alpha = 0.5$,  | $\rho = 1850$, $\alpha = 0.5$,  |
|                | $c_l = 2.7607$, $c_s = 1.5$  | $c_l = 2.7607$, $c_s = 1.5$  |
| Frontal cortical bone | —                      | $\rho = 1850$, $\alpha = 0.5$,  |
|                |                           | $c_l = 2.0322$, $c_s = 1.4$  |
| Left parietal cortical bone | —                      | $\rho = 1850$, $\alpha = 0.75$,  |
|                |                           | $c_l = 2.7607$, $c_s = 1.5$  |
| Right parietal cortical bone | —                      | $\rho = 1850$, $\alpha = 0.75$,  |
|                |                           | $c_l = 2.5468$, $c_s = 1.45$ |
| Background fluid media | $\rho = 1000$, $\alpha = 0.0$,  | $\rho = 1000$, $\alpha = 0.0$,  |
|                | $c_l = 1.5$, $c_s = 0.0$  | $c_l = 1.5$, $c_s = 0.0$  |

Table 2. The values assigned to each parameter in the low-dimensional skull model that were used as initial guesses to the JR algorithm. The density $\rho$ is given in units of $\frac{kg}{m^3}$, the absorption $\alpha$ is given in units of $\frac{1}{\mu s}$ and the longitudinal speed $c_l$, shear speed $c_s$ are given in units of $\frac{mm}{\mu s}$.

|                | Q = 3                      | Q = 5                      |
|----------------|----------------------------|----------------------------|
| Diploë layer   | $\rho = 1000$, $\alpha = 0.75$,  | $\rho = 1000$, $\alpha = 0.75$,  |
|                | $c_l = 1.6$, $c_s = 1.3$  | $c_l = 1.6$, $c_s = 1.3$  |
| Cortical bone  | $\rho = 1850$, $\alpha = 0.5$,  | $\rho = 1850$, $\alpha = 0.5$,  |
|                | $c_l = 2.0$, $c_s = 1.35$ | $c_l = 2.0$, $c_s = 1.35$ |
| Frontal cortical bone | —                      | $\rho = 1850$, $\alpha = 0.5$,  |
|                |                           | $c_l = 1.6$, $c_s = 1.3$  |
| Left parietal cortical bone | —                      | $\rho = 1850$, $\alpha = 0.5$,  |
|                |                           | $c_l = 2.0$, $c_s = 1.35$ |
| Right parietal cortical bone | —                      | $\rho = 1850$, $\alpha = 0.75$,  |
|                |                           | $c_l = 2.0$, $c_s = 1.35$ |
| Background fluid media | $\rho = 1000$, $\alpha = 0.0$,  | $\rho = 1000$, $\alpha = 0.0$,  |
|                | $c_l = 1.5$, $c_s = 0.0$  | $c_l = 1.5$, $c_s = 0.0$  |

the JR algorithm. The forward pressure data were generated by use of both the $Q = 3$ and $Q = 5$ skull models. The initial pressure distribution was then reconstructed by solving the optimization problem described in equation (5) for both the $Q = 3$ and $Q = 5$ skull models. The TV-regularization parameter ($\gamma$) was tuned so as to minimize the root mean squared error (RMSE) between the reconstructed initial pressure distribution and the ground truth initial pressure distribution. The value of the regularization parameter should, ideally, be determined so as to maximize a task-based image quality metric. Appendix B contains images corresponding to different values of the regularization parameter and a relevant discussion about parameter tuning. The initial guess for initial pressure distribution was the vector of all zeros. The stopping criterion employed for the algorithm was when the relative change in cost function between
successive iterations was less than $10^{-6}$. The average time per iteration for the image reconstruction algorithm was approximately 8 min, and the results shown in section 5.1 converged within 350 iterations.

4.1.7 Joint reconstruction. For all the computer-simulation studies, the density and the absorption coefficient were assumed to be known. Hence, the JR problem was restricted to the problem of estimating the spatial distribution of the longitudinal sound speed of the medium ($c_l$), the spatial distribution of the shear sound speed of the medium ($c_s$) along with $p_0$. It has been shown that the spatial distribution of the density and the absorption coefficient of the skull can be estimated using raw 3D x-ray CT images expressed in Hounsfield units (Aubry et al 2003). The methodology outlined by Aubrey et al (Aubry et al 2003) was employed to compute and assign the density values and the longitudinal wave absorption coefficient values to the low-dimensional acoustic representation of the skull.

The mapping function $\Phi$ employed for the $Q = 3$ and $Q = 5$ skull models were generated using the segmented 3D x-ray CT images of the skull. The initial guess for $p_0$ was the vector of all zeros. The initial guess for the parameterized sound speed distribution depended on the number of parameters employed and is summarized in table 2. The two versions of $c_l$ employed for the reconstruction of the initial pressure distribution in section 4.1.6 were employed as initial guesses to the JR algorithm. The stopping criterion employed for the JR algorithm was the same as the one employed in reconstruction of $p_0$ for known acoustic properties of the skull. Similar to section 4.1.6, the TV-regularization parameter ($\gamma$) in the JR algorithm was tuned so as to minimize the RMSE between the reconstructed initial pressure distribution and the ground truth initial pressure distribution. The average time per iteration for the joint image reconstruction algorithm was approximately 24 min, and the results shown in section 5.2 converged within 300 iterations.

5. Results

5.1. Reconstruction of $p_0$ when the acoustic properties of the skull are known

The maximum intensity projection (MIP) images of the estimates of $p_0$ reconstructed by solving equation (5) when the $Q = 3$ and $Q = 5$ skull models were employed are shown in figures 6 and 7. Here, inverse crime with respect to the dimension of the skull model occurred when solving the reconstruction problem. From figures 6 and 7, it can be observed that the images reconstructed by solving the optimization problem defined in equation (5) contains artifacts. The artifacts are present because the acoustic property values assigned to the skull when reconstructing the initial pressure distributions were inaccurate. This demonstrates that image quality degrades considerably when the acoustic property values assigned to the parameters in the low-dimensional skull model are inaccurate. The artifacts present in figures 6 and 7 motivate the need to develop JR algorithms whereby the acoustic property values associated with the low-dimensional representation of the skull can be refined concurrently with $p_0$.

5.2. Joint reconstruction

The 3D MIP images of the estimates of $p_0$ reconstructed by use of the JR method from the pressure data generated corresponding to the $Q = 3$ and $Q = 5$ skull models are shown in figures 8 and 9, respectively. Here, inverse crime with respect to the dimension of the skull model occurred. For the pressure data generated using $Q = 5$ skull model, JR of $p_0$ and the acoustic properties of a $Q = 3$ skull model was also performed. This study, i.e. non-inverse
Figure 6. The maximum intensity projections of the reconstructed initial pressure distribution along the (a) $x$-axis, (b) $y$-axis and (c) $z$-axis, respectively. The initial pressure distribution was reconstructed using a conventional (non-JR) iterative image reconstruction algorithm for a fixed $Q = 3$ skull model.

Figure 7. The maximum intensity projections of the reconstructed initial pressure distribution along the (a) $x$-axis, (b) $y$-axis and (c) $z$-axis, respectively. The initial pressure distribution was reconstructed using a conventional (non-JR) iterative image reconstruction algorithm for a fixed $Q = 5$ skull model.

crime with respect to the dimension of the skull model, was performed to help elucidate if performing JR with an inaccurate lower dimensional parameterized model had significant effect on the image quality of the reconstructed estimates of $p_0$. In figure 10, MIP images of the reconstructed $p_0$ estimated with a JR algorithm employing a $Q = 3$ skull model and utilizing pressure data generated using a $Q = 5$ skull model, is shown. Also, shown for comparison in figures 8(d)–(f) and 9(d)–(f) are the MIP images of $p_0$ reconstructed using the half-time backprojection (BP) method (Anastasio et al 2005). The uniform SOS value employed in the method was chosen by considering a range of values and selecting the value that gave us the best reconstructed image quality, as subjectively judged via visual inspection. Heuristic methods such as the half-time BP methods have been demonstrated to be effective for the case of fluid media containing weak acoustic heterogeneities; however, they are expected to be generally ineffective in the presence strong acoustic heterogeneities such as the skull. This is evident
Figure 8. The maximum intensity projections of the reconstructed 3D initial pressure distribution. The first row of images correspond to the initial pressure distribution reconstructed using a parameterized JR algorithm with a $Q = 3$ skull model. The pressure data used to produce the images were generated using a $Q = 3$ skull model. The second row of images correspond to initial pressure distribution reconstructed using a half-time reconstruction method. Each column of images correspond to MIP projections along $x$-axis, $y$-axis, and $z$-axis, respectively.

when comparing the quality of the images reconstructed using the parameterized JR algorithm and the half-time BP method in figures 8 and 9. The uniform SOS values of $1.602 \mu s$ and $1.610 \mu s$ were used to reconstruct the images shown in figure 8 and 9, respectively.

Comparing the reconstructed images shown in figure 6 with figure 8, and the images shown in figure 7 with figure 9, one can observe that the use of the proposed JR algorithm qualitatively improves the image quality of the reconstructed initial pressure distribution. It can also be observed from figure 10 that allowing the JR algorithm to effectively tune the acoustic properties of the medium, even when the parameterization is crude, can improve the image quality associated with the reconstructed initial pressure distribution. In addition to qualitative improvements, the proposed JR algorithm also quantitatively improves the accuracy of the estimated initial pressure distribution as demonstrated by the line profiles shown in figures 11 and 12.

Table 3 summarizes the values of the acoustic properties of the skull estimated using the JR algorithm for the two skull models. In addition, table 3 also summarizes the values of the
Figure 9. The maximum intensity projections of the reconstructed 3D initial pressure distribution. The first row of images correspond to the initial pressure distribution reconstructed using a parameterized JR algorithm with a $Q = 5$ skull model. The pressure data used to produce the images were generated using a $Q = 5$ skull model. The second row of images correspond to initial pressure distribution reconstructed using a half-time reconstruction method. Each column of images correspond to MIP projections along x-axis, y-axis, and z-axis, respectively.

Figure 10. The maximum intensity projections of the reconstructed initial pressure distribution along the (a) x-axis, (b) y-axis and (c) z-axis, respectively. The initial pressure distribution reconstructed using a parameterized JR algorithm with a $Q = 3$ skull model. The pressure data used to produce the images were generated using a $Q = 5$ skull model.
acoustic properties of the $Q = 3$ skull model estimated when applying the JR algorithm on the data generated using a $Q = 5$ skull model. Comparing the results shown in table 3 with the original values shown in table 1, one observes that the JR algorithm can accurately estimate the acoustic and elastic property values associated with the low-dimensional skull model. When non-inverse crime with respect to the skull model was performed, i.e. when the JR algorithm
Table 3. The values reconstructed for each parameter in the low-dimensional skull model using the proposed JR algorithm. The density $\rho$ is given in units of $\text{kg/m}^3$, the absorption $\alpha$ is given in units of $\mu\text{s}$ and the longitudinal speed $c_l$, shear speed $c_s$ are given in units of $\mu\text{s}$.

|                              | $Q = 3$                              | $Q = 5$                              | $Q = 3$ reconstruction from $Q = 5$ pressure data |
|------------------------------|--------------------------------------|--------------------------------------|--------------------------------------------------|
| Diploë layer                | $\rho = 1000, \alpha = 0.75, c_l = 2.0842, c_s = 1.3976$ | $\rho = 1000, \alpha = 0.75, c_l = 2.0940, c_s = 1.3933$ | $\rho = 1000, \alpha = 0.75, c_l = 2.1430, c_s = 1.3946$ |
| Cortical bone               | $\rho = 1850, \alpha = 0.5, c_l = 2.7491, c_s = 1.4990$ | $\rho = 1000, \alpha = 0.75, c_l = 2.6087, c_s = 1.4454$ | $\rho = 1000, \alpha = 0.75, c_l = 2.5316, c_s = 1.4484$ |
| Frontal cortical bone       | $\rho = 1850, \alpha = 0.5, c_l = 2.0191, c_s = 1.3974$ | $\rho = 1850, \alpha = 0.5, c_l = 2.7324, c_s = 1.4987$ | $\rho = 1850, \alpha = 0.75, c_l = 2.5316, c_s = 1.4484$ |
| Left parietal cortical bone | $\rho = 1850, \alpha = 0.5, c_l = 2.0191, c_s = 1.3974$ | $\rho = 1850, \alpha = 0.5, c_l = 2.7324, c_s = 1.4987$ | $\rho = 1850, \alpha = 0.75, c_l = 2.5316, c_s = 1.4484$ |
| Right parietal cortical bone| $\rho = 1850, \alpha = 0.5, c_l = 2.0191, c_s = 1.3974$ | $\rho = 1850, \alpha = 0.5, c_l = 2.7324, c_s = 1.4987$ | $\rho = 1850, \alpha = 0.75, c_l = 2.5316, c_s = 1.4484$ |
| Background fluid media      | $\rho = 1000, \alpha = 0.0, c_l = 1.5, c_s = 0.0$ | $\rho = 1000, \alpha = 0.0, c_l = 1.5, c_s = 0.0$ | $\rho = 1000, \alpha = 0.0, c_l = 1.5, c_s = 0.0$ |
with a $Q = 3$ skull model was applied on the pressure data generated utilizing the $Q = 5$ skull model, the true value of the acoustic properties of the skull were not recovered due to the choice of the parameterization. However, in this case, the JR algorithm estimated the effective/average value of the acoustic properties of the medium for certain regions.

6. Conclusion

In this work, a JR method for transcranial PACT was proposed in which the initial pressure distribution and the spatial distribution of the acoustic properties of the skull are jointly estimated from PACT measurements alone. This approach overcomes the instability associated with the JR problem from PACT measurements by utilizing a low dimensional representation of the spatial distribution of the acoustic properties of the skull. The developed reconstruction methodology was validated by use of computer-simulation studies that mimicked transcranial PACT experiments.

The computer-simulation studies compared and contrasted the performance of the proposed JR algorithm with a conventional optimization-based image reconstruction algorithm that assumed a fixed acoustic skull model. It was demonstrated that the conventional optimization-based image reconstruction algorithm yielded images with significant artifacts when the assumed skull model was inaccurate. The proposed JR algorithm, on the other hand, was able to refine the spatial distribution of the acoustic properties of the skull concurrently with the initial pressure distribution to yield accurate, high quality images of the initial pressure distribution. In addition, the proposed JR algorithm was also able to accurately estimate the acoustic property values associated with the low-dimensional skull model.

In addition, parameterized JR may offer advantages compared with manually tuning a parameterized representation of the low-dimensional skull model, particularly as the number of parameters in the skull model is increased. The number of parameters employed in the skull model is a design parameter that needs to be carefully considered. When the assumed parameterization is too simple to describe the true variations in acoustic properties within the skull, model error will lead to errors in the reconstructed images. When the assumed parameterization is very complex, the inverse problem may be poorly conditioned and may have many local minima or saddle points. Future studies can focus on considering a range of different parameterizations with different numbers of parameters to carefully examine the trade-off between model error and the conditioning of the inverse problem.

There remain several additional topics for further investigation. For instance, the parameterized low-dimensional representations of the skull developed in this work were developed based on 3D x-ray CT images of the skull. Hence, adjunct image data that provided information about the contour, location of the skull was required. There remains a need to investigate other sources of adjunct image that can help provide the contour, location of the skull without employing ionizing radiation. In addition, future investigations should also include validating the proposed JR algorithm using experimental data. Validating the proposed JR algorithm using experimental data will provide insight into the complexity of skull model required for accurate reconstruction of initial pressure distribution in transcranial PACT. Furthermore, future work can also focus on developing more efficient optimization methods for solving the JR problem, such as accelerated first-order or second-order methods, that would reduce the computational cost.
Acknowledgments

This work was supported in part by National Institutes of Health Grants R01 NS102213, T32EB01485505 and National Science Foundation Grant DMS1614305. The authors also thank Professor Lihong V Wang for his years of collaboration that inspired this work.

Appendix A. Application of the adjoint state method to the elastic wave equation

Here, the application of the adjoint state method for computing the gradients of the data fidelity term given in equation (6) with respect to the spatial distribution of the acoustic parameters of the elastic medium is described (Fabien-Ouellet et al 2016, Tromp et al 2005).

For simplicity, the presentation focuses on the first-order velocity–stress formulation of the elastic wave equation for transcranial PACT given in equation (2). The approaches for the second-order elastic wave equation follow the same basic method and have been detailed previously (Tromp et al 2005, Plessix 2006, Norton 1999).

To start, consider the constraints correspond to the elastic wave equation given by equation (2),

\[ h_0(p, \dot{u}, \sigma, C) = \rho \partial_t \dot{u}(r, t) + \rho c \dot{u}(r, t) - \nabla \cdot \sigma(r, t) \]  

(A.1a)

\[ h_1(p, \dot{u}, \sigma, C) = \partial_t \sigma - C : \nabla \dot{u} \]  

(A.1b)

\[ h_2(p, \dot{u}, \sigma, C) = \sigma(r, 0) - \sigma_0(r) \]  

(A.1c)

\[ h_3(p, \dot{u}, \sigma, C) = \dot{u}(r, 0) \]  

(A.1d)

where \( \sigma \) and \( \dot{u} \) are the state variables and \( C \) is a 4D tensor that describes the elastic moduli of the media which are the sought-after model parameters. A method for computing the values of the stress tensor and particle velocity given the elastic moduli and initial stress distributions, i.e. a method for solving the elastic wave equation, was previously discussed in section 2.2.

Consider the following functional to be minimized:

\[ f(p) = \frac{1}{2} \int_0^T \int_{\partial \Omega} \mathrm{d}\mathbf{r}_0 \left( p(r_0, t) - M \sigma(r, t) \right)^2, \]

(A.2)

where \( p(r, t) \) is the pressure recorded over all transducer and \( M \) is the restriction of the diagonal components of the stress tensor over the whole domain to the pressure at the location of the transducers. Hence, the operator \( M \) computes the trace of the stress tensor over the whole grid and maps it to the pressure at the transducer locations. Note that this function depends only on the stress tensor and not directly on the particle velocity. The corresponding augmented functional is given by

\[ \mathcal{L} = \frac{1}{2} \int_0^T \int_{\partial \Omega} \mathrm{d}\mathbf{r}_0 \left( p(r_0, t) - M \sigma(r, t) \right)^2 \]

\[ - \int_0^T \mathrm{d}t \langle \gamma_1, \rho \partial_t \dot{u}(r, t) + \rho c \dot{u}(r, t) - \nabla \cdot \sigma(r, t) \rangle_U \]

\[ - \int_0^T \mathrm{d}t \langle \gamma_2, \partial_t \sigma(r, t) - C : \nabla \dot{u} \rangle_V - \langle \phi^0, \dot{u}(0) \rangle_U - \langle \phi^1, \sigma(0, t) - \sigma_0 \rangle_V, \]

(A.3)
where
\[
\langle x, y \rangle_U = \sum_{i=1}^{n} \int_U x_i(r) y_i(r) \quad \text{and} \quad \langle a, b \rangle_V = \sum_{i=1}^{m} \sum_{j=1}^{m} \int_V a_{ij}(r) b_{ij}(r)
\]
are the inner products for vector-valued and tensor-valued quantities, respectively. Here, \(n\) is the number of components in the vector-valued quantity, and \(x_i\) represents the \(i\)th component of \(x\). Furthermore, let \(m\) be the rank of the tensor and \(a_{ij}\) represents the \((i, j)\) element of the tensor \(a\).

To compute the derivatives needed to solve the adjoint state equations, it is useful to rearrange the constraint terms. This is done by use of integration-by-parts and by the fact that the adjoint of the gradient operator is the negative divergence operator. With this, the first and the second term in the Lagrangian corresponding to equation (A.1a) can be rewritten as
\[
\int_0^T dt \langle \gamma_1, \rho \partial_t \dot{u} \rangle_U = \langle \rho \gamma_1, \dot{u} \rangle_U \bigg|_{T=0} - \int_0^T dt \langle \partial_t \gamma_1, \rho \dot{u} \rangle_U. \tag{A.4a}
\]
\[
\int_0^T dt \langle \gamma_2, \rho \dot{u} \rangle_U = \int_0^T dt \langle \gamma_2, \rho \dot{u} \rangle_U. \tag{A.4b}
\]
The third term in the Lagrangian corresponding to equation (A.1a) can be rewritten as
\[
\int_0^T dt \langle \gamma_1, \nabla \cdot \sigma \rangle_U = \int_0^T dt \int_U dr \gamma_1 \cdot (\nabla \cdot \sigma).
\]
Using integration by parts
\[
\Rightarrow \int_U dr \gamma_1 \cdot (\nabla \cdot \sigma) = \int_{\partial U} dr [\hat{n} \times \gamma_1] : \sigma - \int_U \nabla \gamma_1 : \sigma
\]
\[
\Rightarrow \int_0^T dt \langle \gamma_1, \nabla \cdot \sigma \rangle_U = \int_0^T dt \int_{\partial U} dr [\hat{n} \times \gamma_1] : \sigma - \int_0^T dt \langle \nabla \gamma_1, \sigma \rangle_U. \tag{A.4c}
\]
Similarly, the first term in the Lagrangian corresponding to equation (A.1b) can be rewritten as
\[
\int_0^T dt \langle \gamma_2, \partial_t \sigma \rangle_V = \langle \gamma_2, \sigma \rangle_V \bigg|_{T=0} - \int_0^T dt \langle \partial_t \gamma_2, \sigma \rangle_V. \tag{A.4d}
\]
The second term in the Lagrangian corresponding to equation (A.1b) can be rewritten as
\[
\int_0^T dt \langle \gamma_2, C : \nabla \dot{u} \rangle_V = \int_0^T dt \int_V dr \gamma_2 : (C : \nabla \dot{u}).
\]
Using the symmetric nature of elastic moduli \(C\),
\[
\Rightarrow \int_V dr \gamma_2 : (C : \nabla \dot{u}) = \int_V dr \nabla \dot{u} : C \gamma_2 dx,
\]
and integration by parts yields
\[
\Rightarrow \int_V dr \nabla \dot{u} : C \gamma_2 = \int_{\partial V} dr [\hat{n} \times \dot{u}] : [C : \gamma_2] - \int_V dr \dot{u} \cdot [\nabla \cdot (C : \gamma_2)].
Hence,
\[
\int_0^T \! \! \frac{d}{dt} \langle \gamma_2, C : \nabla \hat{u} \rangle_V = \int_0^T \! \! \frac{d}{dt} \int_{\partial \Omega} \! \! d\mathbf{r} \ [\mathbf{n} \times \hat{u}] : [C : \gamma_2] - \int_0^T \! \! \frac{d}{dt} \langle \hat{u}, \nabla \cdot (C : \gamma_2) \rangle_V.
\]

(A.4e)

Since, the wave equation in equation (2) is solved overall space,
\[
[\mathbf{n} \times \hat{u}] : [C : \gamma_2] = 0 \quad \text{on } \partial V
\]
\[\text{(A.4f)}\]
\[
[\mathbf{n} \times \gamma_1] : \sigma = 0 \quad \text{on } \partial U.
\]

(A.4g)

Combining all subequations in equation (A.4), the overall Lagrangian can be rewritten as
\[
L = \frac{1}{2} \int_0^T \frac{d}{dt} \int_{\partial \Omega} \! \! d\mathbf{r}_0 \left( M \sigma(r, t) - p(r_0, t) \right)^2 - \langle \rho \gamma_1, \hat{u} \rangle |_T^0 - \langle \phi^0, \hat{u}(0) \rangle |_V + \int_0^T \frac{d}{dt} \langle \hat{u}, \nabla \cdot (C : \gamma_2) \rangle_V.
\]

(A.5)

Now the derivatives of the Lagrangian with respect to the state variables can be computed as
\[
\frac{\partial L}{\partial \hat{u}} = \int_0^T \frac{d}{dt} \left( \rho \partial_t \gamma_1 - \rho \alpha \gamma_1 + \nabla \cdot (C : \gamma_2) \right) d\mathbf{r} - \langle \rho \gamma_1, \hat{u} \rangle |_T^0 - \phi^0 |_{t=0}
\]
\[
\frac{\partial L}{\partial \sigma} = \int_0^T \frac{d}{dt} \int_{\partial \Omega} \! \! d\mathbf{r}_0 \quad M^*(M \sigma(r, t) - p(r_0, t))
\]
\[
+ \int_0^T \frac{d}{dt} \langle \partial_t \gamma_2 + \nabla \gamma_1 - \gamma_2 |_{T} - \phi^1 |_{t=0} \rangle |_V.
\]

To obtain an expression for the adjoint state equation, the computed $\frac{\partial L}{\partial \hat{u}}$ and $\frac{\partial L}{\partial \sigma}$ are set to zero and the terms are grouped by their time dependence. This yields
\[
\rho \partial_t \gamma_1 - \rho \alpha \gamma_1 + \nabla \cdot (C : \gamma_2) = 0
\]
\[\text{(A.6a)}\]
\[
-\partial_t \gamma_2 - \nabla \gamma_1 = \int_{\partial \Omega} \! \! d\mathbf{r}_0 \quad M^*(M \sigma(r, t) - p(r_0, t))
\]
\[\text{(A.6b)}\]
\[
\gamma_2(r, T) = 0
\]
\[\text{(A.6c)}\]
\[
\rho \gamma_1(r, T) = 0
\]
\[\text{(A.6d)}\]
\[
\gamma_2(r, 0) = \phi^1
\]
\[\text{(A.6e)}\]
\[
\rho \gamma_1(r, 0) = \phi^0.
\]

(A.6f)

The set of differential equations in equation (A.6) bear a close resemblance to the first-order acoustic wave equation in equation (2). This suggests that a solution to these differential equations could possibly be computed by a similar approach. However, one difficulty is that...
equation (A.6) involves final conditions on $\phi^0$ and $\phi^1$ rather than the initial conditions specified in equation (2). A change of variables can be employed to change the final conditions to initial conditions:

$$\eta_i(r, t) = \gamma_i(r, T - t).$$  \hspace{1cm} (A.7)

This gives a set of differential equations for the adjoint state variables that can be solved by the same method as employed to solve equation (2):

$$\partial_t \eta_1 + \alpha \eta_1 = \frac{1}{\rho} \left( \nabla \cdot (C : \eta_2) \right)$$  \hspace{1cm} (A.8a)

$$\partial_t \eta_2 - \nabla \eta_1 = \int_{\partial V} d\mathbf{r}_0 \mathcal{M}^* \left( \mathcal{M} \sigma(r, t) - p(r_0, t) \right)$$  \hspace{1cm} (A.8b)

$$\eta_2(r, 0) = 0$$  \hspace{1cm} (A.8c)

$$\eta_1(r, 0) = 0.$$  \hspace{1cm} (A.8d)

As such, the computational cost of calculating the adjoint state variables is the same as solving the acoustic wave equation for the forward problem.

The last step is to compute the gradients with respect to the model parameters by use of the previously calculated state and adjoint state variables. This can be done as

$$\frac{\partial L}{\partial \sigma_0} = \phi^1 = \eta_2(r, T)$$  \hspace{1cm} (A.9a)

$$\frac{\partial L}{\partial \rho} = \int_0^T dt \partial_t \bar{u}(r, t) \cdot \eta_1(r, T - t)$$  \hspace{1cm} (A.9b)

$$\frac{\partial L}{\partial \alpha} = -\int_0^T dt \rho(r) \eta_1(r, T - t).$$  \hspace{1cm} (A.9c)

In an isotropic media,

$$C_{ij,kl}(\mathbf{r}) = \left( \kappa(\mathbf{r}) - \frac{2}{3} \mu(\mathbf{r}) \right) \delta_{ij} \delta_{kl} + \mu(\mathbf{r}) (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}),$$

where $\kappa(\mathbf{r})$ is the bulk modulus and $\mu(\mathbf{r})$ is the shear modulus. The bulk modulus is related to the Lamé parameters by the following relation

$$\kappa(\mathbf{r}) = \lambda(\mathbf{r}) - \frac{2}{3} \mu(\mathbf{r}).$$

The only term in the Lagrangian dependent on the elastic moduli of the medium is the term $\int_0^T dt \langle \eta, C : \nabla u \rangle_V$. Also,

$$\int_0^T dt \langle \eta_2, C : \nabla u \rangle_V = \int_0^T \int V \rangle_d \langle \nabla u, C : \eta_2 \rangle_V,$$

24
due to the symmetric property of the stress tensor. We can write
\[
\langle \nabla \hat{u}, C : \eta \rangle_v = \int \left( \eta_1^I (r, T - t) \left( \kappa + \frac{4}{3} \mu \right) \partial_t \hat{u}^I (r, t) + \left( \kappa - \frac{2}{3} \mu \right) \partial_r \hat{u}^I (r, t) \right)
+ \eta_2^I (r, T - t) \left( \kappa + \frac{4}{3} \mu \right) \partial_r \hat{u}^I (r, t) + \left( \kappa - \frac{2}{3} \mu \right) \partial_t \hat{u}^I (r, t)
+ \eta_3^I (r, T - t) \left( \kappa + \frac{4}{3} \mu \right) \partial_t \hat{u}^I (r, t) + \left( \kappa - \frac{2}{3} \mu \right) \partial_r \hat{u}^I (r, t)
+ \eta_4^I (r, T - t) \left( \kappa + \frac{4}{3} \mu \right) \partial_r \hat{u}^I (r, t) + \left( \kappa - \frac{2}{3} \mu \right) \partial_t \hat{u}^I (r, t)
+ \eta_5^I (r, T - t) \left[ \mu \left( \partial_t \hat{u}^I (r, t) + \partial_r \hat{u}^I (r, t) \right) \right] + \eta_6^I (r, T - t) \left[ \mu \left( \partial_t \hat{u}^I (r, t) + \partial_r \hat{u}^I (r, t) \right) \right]
+ \eta_7^I (r, T - t) \left[ \mu \left( \partial_t \hat{u}^I (r, t) + \partial_r \hat{u}^I (r, t) \right) \right]
+ \eta_8^I (r, T - t) \left[ \mu \left( \partial_t \hat{u}^I (r, t) + \partial_r \hat{u}^I (r, t) \right) \right] \right) \, dr.
\]
Hence, the gradient with respect to the same parameters \( \kappa, \mu \) are given by,
\[
\frac{\partial L}{\partial \kappa} = \int_0^T \int \, dr \, \text{tr} \left( \nabla \hat{u} (r, t) \right) \text{tr} \left( \eta_2 (r, T - t) \right),
\]
where \( \text{tr}(\cdot) \) computes the trace of a matrix and
\[
\frac{\partial L}{\partial \mu} = \int_0^T \int \, dr \left( \frac{4}{3} \eta_2^1 (r, T - t) - \frac{2}{3} \eta_2^2 (r, T - t) - \frac{2}{3} \eta_2^3 (r, T - t) \right) \partial_t \hat{u}^1 (r, t)
+ \left( \frac{4}{3} \eta_2^2 (r, T - t) - \frac{2}{3} \eta_2^3 (r, T - t) - \frac{2}{3} \eta_2^1 (r, T - t) \right) \partial_r \hat{u}^2 (r, t)
+ \left( \frac{4}{3} \eta_2^3 (r, T - t) - \frac{2}{3} \eta_2^1 (r, T - t) - \frac{2}{3} \eta_2^2 (r, T - t) \right) \partial_t \hat{u}^3 (r, t)
+ \eta_2^1 (r, T - t) \left( \partial_t \hat{u}^1 (r, t) + \partial_r \hat{u}^1 (r, t) \right) + \eta_2^2 (r, T - t) \left( \partial_t \hat{u}^2 (r, t) + \partial_r \hat{u}^2 (r, t) \right)
+ \eta_2^3 (r, T - t) \left( \partial_t \hat{u}^3 (r, t) + \partial_r \hat{u}^3 (r, t) \right) + \eta_2^4 (r, T - t) \left( \partial_t \hat{u}^4 (r, t) + \partial_r \hat{u}^4 (r, t) \right)
+ \eta_2^5 (r, T - t) \left( \partial_t \hat{u}^5 (r, t) + \partial_r \hat{u}^5 (r, t) \right) + \eta_2^6 (r, T - t) \left( \partial_t \hat{u}^6 (r, t) + \partial_r \hat{u}^6 (r, t) \right)
\]
Instead of using \( \kappa \) and \( \mu \) as the independent parameters but rather the compressional and shear speed value as independent parameters, we arrive at the following gradient values
\[
\frac{\partial L}{\partial c_1} = 2 \rho(r) c_1(r) \frac{\partial L}{\partial \kappa},
\]
\[
\frac{\partial L}{\partial c_4} = 2 \rho(r) c_4(r) \left[ \frac{\partial L}{\partial \kappa} - \frac{4}{3} \frac{\partial L}{\partial \mu} \right],
\]
where \( c_1(r), c_4(r) \) are the compressional and shear speed distributions respectively.

In a discrete setting, the above quantities described in all the sub equations of equation (A.9) may be replaced by their corresponding discrete approximations (see section 2.2).

For computing the gradient of the cost functional with respect to the acoustic and elastic properties of the skull described in equation (A.9), the fields associated with the forward elastic wave equation and the fields associated with the adjoint elastic wave equation must be available at the same time. When solving problems that are not memory intensive, the forward fields or state variables associated with the solution of the forward problem are stored in
Figure 13. The maximum intensity projections of the reconstructed 3D initial pressure distribution using the JR algorithm with a $Q = 5$ skull model. The first row of images correspond to images reconstructed using a TV regularization value of $\gamma = 0.02$, the second row of images correspond to images reconstructed using a TV regularization value of $\gamma = 0.05$ and the third row of images correspond to images reconstructed using a TV regularization value of $\gamma = 0.1$. Each column of images correspond to MIP projections along x-axis, y-axis, and z-axis, respectively.

Random access memory for all time points and its entire history is made accessible when the adjoint problem is solved for efficient gradient computation (Matthews 2017, Matthews et al 2018, Huang et al 2016). For large-scale problems, such as the 3D transcranial PACT problem, the storage requirement for gradient computation is severe. Hence, if the commonly employed strategy of storing forward state variables is adopted, it will require the algorithm to store the wavefield in the slowest levels of memory hierarchy such as a hard disk drive. This need for very large amounts of read/write operations performed on hard disk complicates the logistics.
**Algorithm 2.** Line search method for selecting step size $\hat{\delta}^{(k)}_\theta$.

**Input:** $\theta^{(k)}$, $\hat{\delta}^{(k)}_\theta$, $G^{(k)}_\theta$, $c^{(k)}_\theta$, $\eta > 0$.

**Output:** $\hat{\delta}^{(k)}_\theta$.

1: $\hat{\delta}^{(k)}_\theta \leftarrow \delta^{(k)}_\theta$
2: $\text{found} \leftarrow \text{false}$
3: $l \leftarrow 0$, where $l$ is the line search iteration number
4: while not $\text{found}$ do
5: $Q^{(k)} \leftarrow \left( \langle \text{prox}_{\delta^{(k)}_\theta} R (\theta^{(k)}_0 - \theta^{(k)}_0), G (\theta^{(k)}) \rangle + \frac{1}{2\delta^{(k)}_\theta} \| \text{prox}_{\delta^{(k)}_\theta} R (\theta^{(k)}_0 - \theta^{(k)}) \|^2_2 \right)^{-1}$
6: if $C \left( \text{prox}_{\delta^{(k)}_\theta} R (\theta^{(k)}_0 - \hat{\delta}^{(k)}_\theta G^{(k)}_\theta), c^{(k)}_\theta \right) \leq C \left( \theta^{(k)}_0, c^{(k)}_\theta \right) + Q^{(k)}$ then
7: $\text{found} \leftarrow \text{true}$
8: else
9: $\hat{\delta}^{(k)}_\theta \leftarrow \eta l \hat{\delta}^{(k)}_\theta$
10: end if
11: $l \leftarrow l + 1$
12: end while

of algorithm implementation and degrades the computational performance of the algorithm significantly. In order to circumvent the memory burden associated with storing large 3D volumes over all time, in this work the optimal binomial checkpointing method was adopted to compute the gradient (Symes 2007, Griewank and Walther 2000, Chen et al 2013)

**Appendix B. Tuning of TV regularization parameter**

Maximum intensity projection (MIP) visualizations of 3D images reconstructed by use of the JR method with different choices of regularization parameter are shown in figure 13. The 3D MIP images correspond to an estimate of $p_0$ reconstructed by use of the JR method from the pressure data generated corresponding to the $Q = 5$ skull model. Here, inverse crime with respect to the dimension of the skull model occurred, even though the forward data was corrupted with additive white Gaussian noise prior to the reconstruction. Each column of images represent the MIP of the reconstructed initial pressure distribution along the $x$-, $y$- and $z$-axis, respectively. Each row of images in figure 13 were reconstructed with a different TV regularization parameter value $\gamma$; the chosen range of regularization parameters resulted in images that varied in appearance from being under-smoothed (top row) to over-smoothed (bottom row). The reconstructed images corresponding to different regularization parameters display the expected trade-off between variance and spatial resolution. For the JR results obtained for the $Q = 5$ skull model, it was observed that the optimum TV regularization parameter that gave us the minimum the RMSE between original initial pressure distribution and the reconstructed initial pressure distribution was $\gamma = 0.05$. Hence, it was observed that the value of the TV regularization parameter that gives the minimum RMSE value also resulted in images that possessed a potentially useful trade-off between variance and spatial resolution. This was observed for all the images reconstructed in this study that employed TV regularization.
Algorithm 3. Line search method for selecting step size $\delta_{\text{dp}}^{(k)}$.

**Input:** $p_0^{(k+1)}$, $\delta_0^{(k)}$, $G_{ep}^{(k)}$, $c_p^{(k)}$, $l_{\text{max}}$.

**Output:** $\delta_{\text{dp}}^{(k)}$.

1: $\delta_{\text{dp}}^{(k)} ← \delta_0^{(k)}$.
2: found ← false, flagup ← false.
3: $C_0 ← C(p_0^{(k+1)}, c_{p}^{(k)})$.
4: $l ← 0$ (line search iteration number).
5: while not found and $l ≤ l_{\text{max}}$ do
6:   $C_{\text{imp}} ← C(p_0^{(k+1)}, p_{ep}^{(k)}(c_p^{(k)} - \delta_{\text{dp}}^{(k)} G_{ep}^{(k)}), c_{p}^{(k)})$.
7:   if $l = 0$ then
8:     if $C_{\text{imp}} < C_0$ then
9:       flagup ← true.
10:      $\delta_{\text{dp}}^{(k)} ← 2\delta_{\text{dp}}^{(k)}$
11:     else
12:       $C_0 ← C_{\text{imp}}$
13:     end if
14:   else
15:     if flagup then
16:       if $C_{\text{imp}} < C_0$ then
17:         $\delta_{\text{dp}}^{(k)} ← \frac{\delta_{\text{dp}}^{(k)}}{2}$
18:       else
19:         $C_0 ← C_{\text{imp}}$
20:     end if
21:   else
22:     if $C_{\text{imp}} > C_0$ then
23:       $\delta_{\text{dp}}^{(k)} ← \frac{\delta_{\text{dp}}^{(k)}}{2}$
24:     else
25:       found ← true.
26:     end if
27:     end if
28:   end if
29: end while
30: $l ← l + 1$.

Appendix C. Linesearch method

The bi-directional line search procedure, detailed in algorithm 2, was employed to select a scalar step size for the initial pressure distribution. In addition, the bi-directional line search procedure detailed in algorithm 3 was employed to select a scalar step size for the low-dimensional parameterized representation of the spatial distribution of the elastic properties of the medium. As the bidirectional line search is memory efficient, the computation of the step size did not require the use of slow memory or disk memory.
ORCID iDs

Mark A Anastasio © https://orcid.org/0000-0002-3192-4172

References

Acosta S 2019 Inverse Problems 35 115013
Alterman Z and F K Jr 1968 Bull. Seismol. Soc. Am. 58 367–98
Anastasio M A, Jin Zhang J, Xiaochuan Pan X, Yu Zou Y, Geng Ku G and Wang L V 2005 IEEE Trans. Med. Imaging 24 199–210
Anastasio M A, Jin Zhang J, Sidky E Y, Yu Zou Y, Xia D and Xiaochuan Pan X 2005 IEEE Trans. Med. Imaging 24 1100–12
Aubry J-F, Tanter M, Pernot M, Thomas J-L and Fink M 2003 J. Acoust. Soc. Am. 113 84–93
Auperrin A, Delille R, Lesueur D, Bruyère K, Masson C and Dražetic P 2014 J. Biomech. 47 1180–5
Beck A and Teboulle M 2009a IEEE Trans. Image Process. 18 2419–34
Beck A and Teboulle M 2009b SIAM J. Imaging Sci. 2 183–202
Boore D M 1972 Methods in Computational Physics ed B Bolt (New York: Academic)
Burgholzer P, Matt G J, Haltmeier M and Paltauf G 2007 Phys. Rev. E 75 046706
Chambolle A, Caselles V, Cremers D, Novaga M and Pock T 2010 Theoretical Foundations and Numerical Methods for Sparse Recovery 9 p 227
Chang Y, Kim N and Stenfelt S 2016 J. Acoust. Soc. Am. 140 1635–51
Chen G, Wang X, Wang J, Zhao Z, Nie Z-P and Liu Q H 2013 J. Acoust. Soc. Am. 140 46 41–57
Cheong W, Prahl S and Welch A 2003 IEEE J. Quantum Electron. 26 1086–99
Colton D L 2013 New Analytic and Geometric Methods in Inverse Problems: Lectures Given at the EMS Summer School and Conf. Held in Edinburgh (Scotland 2000) (Berlin: Springer) p 283
2015 CUDA C Programming Guide Technical Report PG-02829-001_v7.5 (NVIDIA Corporation)
Estrada H, Gottschalk S, Reiss M, Neuschmelting V, Goldbrunner R and Razansky D 2018 Ultrasound Med. Biol. 44 2388–92
Fabien-Ouellet G, Glaouguen E and Giroux B 2016 78th EAGE Conf. and Exhibition 2016
Finch S K P D and Rakesh 2004 SIAM J. Math. Anal. 35 1213–40
Fry F J and Barger J E 1978 J. Acoust. Soc. Am. 63 1576–90
Fry F J and Barger J E 1978 J. Acoust. Soc. Am. 63 1576–90
Griewank A and Walther A 2000 ACM Trans. Math. Softw. 26 19–45
Grün H, Nuster R, Paltauf G, Haltmeier M and Burgholzer P 2008 Photons Plus Ultrasound: Imaging and Sensing 2008: the Ninth Conf. on Biomedical Thermoacoustics, Optoacoustics, and Acousto-Optics (Bellingham, WA: International Society for Optics and Photonics) 6856 685620
Hickmann K S 2010 (https://ir.library.oregonstate.edu/concern/graduate_thesis_or_dissertations/f848s492)
Hristova Y, Kuchment P and Nguyen L 2008 Inverse Problems 24 055006
Huang C, Nie L, Schoonover R W, Guo Z, Schirra C O, Anastasio M A and Wang L V 2012 J. Biomed. Opt. 17 066016
Huang C, Nie L, Wang L V and Anastasio M A 2013 IEEE. Trans. Med. Imaging 32 1097–110
Huang C, Wang K, Nie L, Wang L V and Anastasio M A 2013 IEEE Trans. Med. Imaging 32 1097–110
Huang C, Wang K, Schoonover R W, Wang L V and Anastasio M A 2016 IEEE Trans. Comput. Imaging 2 136–49
Javaherian A and Holman S 2018 Inverse Problems 34 085003
Jin X, Li C and Wang L V 2008 Med. Phys. 35 3205–14
Jones W, Jirtle R, Rafal M and Schaeffer D 1980 Radiat. Oncol. Biol. Phys. 6 681–7
Kirsch A and Scherzer O 2012 SIAM J. Appl. Math. 72 1508–23
Kruger R A, Reinecke D R and Kruger G A 1999 Med. Phys. 26 1832–7
Kunyansky L A 2007 Inverse Problems 23 373–83
Liu H and Uhlmann G 2015 Inverse Problems 31 105005
Madariaga R, Olsen K and Archuleta R 1998 Bull. Seismol. Soc. Am. 88 1182–97
Matthews T P 2017
Matthews T P, Poudel J, Li L, Wang L V and Anastasio M A 2018 SIAM J. Imaging Sci. 11 1560–88
McElhaney J H, Fogle J L, Melvin J W, Haynes R R, Roberts V L and Alem N M 1970 J. Biomech. 3 495–511
Mitsuhashi K, Poudel J, Matthews T P, Garcia-Uribe A, Wang L V and Anastasio M A 2017 *SIAM J. Imaging Sci.* **10** 2022–48

Moczo P, Kristek J, Galis M, Pakzad P and Balazovjech M 2007 *Acta Phys. Slovaca* **57** 177–406

Motherway J A, Verschueren P, Van der Perre G, Vander Sloten J and Gilchrist M D 2009 *J. Biomech.* **42** 2129–35

Nie L, Guo Z and Wang L V 2011 *J. Biomed. Opt.* **16** 076005

Norton S J 1999 *J. Acoust. Soc. Am.* **106** 2653–60

Oraevsky A and Karabutov A A 2003 *Biomedical Photonics Handbook* ed T Vo-Dinh (Boca Raton, FL: CRC Press)

O’Donoghue B and Candes E 2015 *Found. Comput. Math.* **15** 715–32

Plessix R-E 2006 *Geophys. J. Int.* **167** 495–503

Poudel J, Lou Y and Anastasio M A 2019 *Phys. Med. Biol.* **64** 14TR01

Poudel J, Matthews T P, Li L, Anastasio M A and Wang L V 2017 *J. Biomed. Opt.* **22** 041018

Poudel J, Matthews T P, Mitsuhashi K, Garcia-Uribe A, Wang L V and Anastasio M A 2017 *Medical Imaging 2017: Ultrasonic Imaging and Tomography* vol 10139 (International Society for Optics and Photonics) p 101390C

Poudel J, Na S, Wang L V and Anastasio M A 2020 *Phys. Med. Biol.* **65** 055009

Schoonover R W and Anastasio M A 2011 *J. Opt. Soc. Am. A* **28** 2091–9

Schoonover R W, Wang L V and Anastasio M A 2012 *J. Biomed. Opt.* **17** 061215

Stefanov P and Uhlmann G 2012 arXiv:1211.6217

Stefanov P and Uhlmann G 2013 *Trans. Am. Math. Soc.* **365** 5737–58

Symes W W 2007 *Geophysics* **72** SM213–21

Szabo T L 1994 *J. Acoust. Soc. Am.* **96** 491–500

Treeby B E and Cox B T 2014 *J. Acoust. Soc. Am.* **136** 1499–510

Treeby B E, Zhang E Z and Cox B T 2010 *Inverse Problems* **26** 115003

Tromp J, Tape C and Liu Q 2005 *Geophys. J. Int.* **160** 195–216

Virieux J 1986 *Geophysics* **51** 899–901

Wang L V and Wu H i 2012 *Biomedical Optics: Principles and Imaging* (New York: Wiley)

White P J, Clement G T and Hynynen K 2006 *Ultrasonic Med. Biol.* **32** 1085–96

Xu M and Wang L V 2006a *Rev. Sci. Instrum.* **77** 041101

Xu Y and Wang L V 2006b *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **53** 542–8

Xu Z, Zhu Q and Wang L V 2011 *J. Biomed. Opt.* **16** 066020

Yang P, Brossier R, Métivier L and Virieux J 2016 *Geophysics* **81** R349–62

Yang X and Wang L V 2008 *J. Biomed. Opt.* **13** 044009

Zhang J, Wang K, Yang Y and Anastasio M A 2008 *Photons Plus Ultrason: Imaging and Sensing 2008: the Ninth Conf. on Biomedical Thermacoustics, Optoacoustics, and Acousto-Optics* (Bellingham, WA: International Society for Optics and Photonics) 6856 68561F