On the adjoint operator in photoacoustic tomography

You may also be interested in:

Photoacoustic tomography in absorbing acoustic media using time reversal
Bradley E Treeby, Edward Z Zhang and B T Cox

A gradient-based method for quantitative photoacoustic tomography using the radiative transfer equation
T Saratoon, T Tarvainen, B T Cox et al.

Compressed sensing and sparsity in photoacoustic tomography
Markus Haltmeier, Thomas Berer, Sunghwan Moon et al.

Photoacoustic tomography in a rectangular reflecting cavity
L Kunyansky, B Holman and B T Cox

Photoacoustic tomography with a reverberant cavity
B T Cox, S R Arridge and P C Beard

A Bayesian approach to spectral quantitative photoacoustic tomography
A Pulkkinen, B T Cox, S R Arridge et al.

Single-stage reconstruction algorithm for quantitative photoacoustic tomography
Markus Haltmeier, Lukas Neumann and Simon Rabanser

On the adjoint operator in photoacoustic tomography

Simon R Arridge[1], Marta M Betcke[1], Ben T Cox[2], Felix Lucka[1] and Brad E Treeby[2]

[1] Department of Computer Science, University College London, WC1E 6BT London, UK
[2] Department of Medical Physics and Bioengineering, University College London, WC1E 6BT London, UK

E-mail: f.lucka@ucl.ac.uk

CrossMark

Abstract

Photoacoustic tomography (PAT) is an emerging biomedical *imaging from coupled physics* technique, in which the image contrast is due to optical absorption, but the information is carried to the surface of the tissue as ultrasound pulses. Many algorithms and formulae for PAT image reconstruction have been proposed for the case when a complete data set is available. In many practical imaging scenarios, however, it is not possible to obtain the full data, or the data may be sub-sampled for faster data acquisition. In such cases, image reconstruction algorithms that can incorporate prior knowledge to ameliorate the loss of data are required. Hence, recently there has been an increased interest in using *variational image reconstruction*. A crucial ingredient for the application of these techniques is the adjoint of the PAT forward operator, which is described in this article from physical, theoretical and numerical perspectives. First, a simple mathematical derivation of the adjoint of the PAT forward operator in the continuous framework is presented. Then, an efficient numerical implementation of the adjoint using a *k*-space time domain wave propagation model is described and illustrated in the context of variational PAT image reconstruction, on both 2D and 3D examples including inhomogeneous sound speed. The principal advantage of this analytical adjoint over an algebraic adjoint (obtained by taking the direct adjoint of the particular numerical forward scheme used) is that it can be implemented using currently available fast wave propagation solvers.

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

1. Introduction

Photoacoustic tomography (*PAT*) is a biomedical imaging modality that has come to pro-
minence over the past two decades due to its ability to provide images of soft tissue based on
optical absorption with high spatial resolution. Optical absorption contrast is desirable
because imaging at multiple wavelengths can in principle provide spectroscopic (chemical)
information on the absorbing molecules (chromophores). However, due to the highly scat-
tering nature of most biological tissues, imaging purely with light—light in, light out—can
only be achieved with high spatial resolution in the vicinity of the tissue surface; beyond that
the achievable image resolution falls off quickly with depth. PAT overcomes this problem by
generating broadband ultrasonic pulses at the regions of optical absorption which then carry
the information about the optical contrast to the tissue surface without being scattered sig-
nificantly. PAT is increasingly widely used in preclinical studies to image small animal
anatomy, pathology and physiology [36, 37, 40] and PAT's potential as a clinical technique is
also being actively explored [20, 30, 41]. The spectroscopic information PAT provides (ie.
differentiating endogenous chromophores or contrast agents based on their absorption
spectra) opens up the possibility of molecular and genomic imaging, and even for functional
imaging by monitoring changes in blood oxygenation and flow. For further references on the
experimental aspects and applications of PAT see the review papers [3, 26, 35].

1.1. The physics of photoacoustic tomography

PAT is based on the principle that acoustic waves can be generated by the absorption of
modulated electromagnetic radiation. When the radiation is in the visible to near-infrared
spectrum this is known as the *photoacoustic* or *optoacoustic* effect. The term *thermoacoustic*
is used when the radiation is in the microwave region. In PAT experiments, it is common to
use pulses of a few nanoseconds duration, such as are generated by a Q-switched laser, to
illuminate the tissue. The light will be scattered within the tissue and absorbed by any
chromophores present. For the photoacoustic effect to take place, the energy must be sub-
sequently thermalised (as opposed to emitted radiatively) and, if it occurs fast enough, this
will give rise to a pressure increase localised to the regions of absorption. These regions of
raised pressure and temperature will initiate acoustic and thermal waves as the tissue is elastic
and thermally conducting. There are two relevant timescales to be considered. First, for the
pressure and temperature fluctuations to be decoupled, the thermal relaxation time must be
much longer than the characteristic timescale for the acoustic propagation. This condition is
known as *thermal confinement*, and it must be satisfied if the acoustic pressure is to be
modelled using a wave equation, rather than a set of coupled equations in temperature and
pressure. The second condition, known as *stress* or *pressure confinement*, is that the char-
acteristic timescale for the acoustic propagation, which will depend on the sound speed and

the size of the absorbing regions, must be much slower than the time taken to deposit the optical energy as heat. This is essentially an isochoric condition requiring there to be no change is density of the tissue during heat deposition. When this condition is satisfied, the acoustic propagation can be modelled as an initial value problem with the *initial (acoustic) pressure distribution* as the initial condition. The second initial condition—that the time rate of the change of the pressure is zero—derives from the isochoric condition through the assumption of zero initial particle velocity.

In PAT, measurements of the induced broadband acoustic pulses are recorded at the surface of the tissue by an array of ultrasound sensors. In order to form a PAT image, it is necessary to solve an inverse initial value problem by inferring an initial acoustic pressure distribution from measured acoustic time series. This image reconstruction step is the subject of this paper. A possible subsequent image reconstruction step, not considered here, is to recover images of the optical properties of the tissue. This requires a model of light transport and is sometimes known as *quantitative PAT* [10]. The solution of the acoustic inversion examined here provides the data for this optical inversion.

1.2. Challenges of photoacoustic tomography

Image reconstruction in PAT calls for the robust inversion of the linear system

$$f = \mathcal{A}p_0 + \varepsilon, \tag{1}$$

where p_0 represents the initial pressure distribution we wish to recover, f represents the measured acoustic pressure time series, ε an additional measurement error, and the linear operator A models both the ultrasonic wave propagation in the tissue and the measurement process (we will discuss the details of this modeling in section 2.1). Solving (1) can be done in several ways:

 i. Assuming homogeneous acoustic properties and the absence of acoustic absorption, the measured time series can be related to the initial pressure distribution via the spherical mean Radon transform. In this case, the methods of integral geometry can be used to derive direct, explicit inversion formulae for certain sensor geometries, such as e.g. spherical arrays. See [21] and the references therein for an overview of such methods.
 ii. In more general scenarios, a direct inverse operator corresponding to a forward model of the wave equation can be used. One commonly used example is given by *time reversal* (*TR*) approaches [12, 15, 38, 39]: The measured time series at the boundary is used in reversed temporal order as a time dependent Dirichlet boundary condition for the backward wave equation and the pressure field at $t = 0$ is used as an estimate of the initial pressure distribution. With the development of fast numerical methods that are capable of computing full 3D wave propagation with high spatial and temporal resolution, TR and its enhanced variants [27, 28] became an attractive choice in applications [16, 18, 23].
 iii. In particular in scenarios with incomplete or sub-sampled data, numerical models of wave propagation can be used within the *variational image reconstruction* framework to find a regularized least-squares solution of (1) by solving the optimization problem

$$p_{\mathrm{rec}} = \operatorname*{argmin}_{p_0} \left\{ \frac{1}{2} \|\mathcal{A}p_0 - f\|_2^2 + \lambda \mathcal{J}(p_0) \right\}. \tag{2}$$

Here, $\lambda > 0$ is a *regularization parameter* and $\mathcal{J}(p_0)$ is a suitable *regularization functional* that aims to encode *a priori* knowledge about the true solution, p_0.

While variational image reconstruction in PAT has recently been shown to yield superior results compared to TR solutions [5, 17], a potential challenge to its application is that solving (2) by *first order optimization* methods (see, e.g., [8] for an overview) requires an efficient numerical scheme for the application of the *adjoint* PAT operator, $\mathcal{A}^*$. One approach is to first discretise the forward operator and then to find its algebraic adjoint, for example by explicit reversal of the computational steps of the forward scheme [17]. The adjoint obtained by this *discretise-then-adjoint* approach however is not the adjoint of the continuous operator but of the particular numerical scheme. Therefore, in this paper we revisit the mathematical modelling of PAT to derive the general analytical form of the PAT adjoint. The explicit analytical form of the adjoint allows us new theoretical insights, in particular, it clarifies the relation between the adjoint and TR. The analytical adjoint is independent of discretization, i.e., both the forward and adjoint equations can be discretised with any suitable discretization scheme and pre-existing high performance codes can be used to apply the forward and adjoint operators. The remainder of the paper is structured as follows: In section 2, we revisit the mathematical modelling of the forward problem in PAT and derive an explicit form of the adjoint PAT operator. Section 3 recapitulates how the adjoint operator can be used within a variational scheme. The utility of *adjoint-then-discretise* paradigm is evaluated in section 4: after a direct comparison to the *discretise-then-adjoint* paradigm in [17], variational methods are compared to approaches based on TR for two image reconstruction problems including sound-speed inhomogeneity. In section 5, we summarize our work and discuss its relations to other approaches. An efficient implementation of the adjoint operator using a *k-space pseudospectral wave propagation model* is given in appendix B.

2. Theory

2.1. Mathematical modeling

Wave propagation. Acoustic wave propagation through a compressible fluid is usually modelled mathematically by linearising the equations of fluid dynamics derived from conservation laws and an equation of state. Under the condition that the acoustic particle velocity is much slower than the sound speed, an acoustic wave in tissue can then be modelled by the system of first order equations [33]:

$$\frac{\partial}{\partial t}\mathbf{u}(\mathbf{x}, t) = -\frac{1}{\rho_0(\mathbf{x})}\nabla p(\mathbf{x}, t) \qquad \text{(momentum conservation)} \quad (3a)$$

$$\frac{\partial}{\partial t}\rho(\mathbf{x}, t) = -\rho_0(\mathbf{x})\nabla \cdot \mathbf{u}(\mathbf{x}, t) - \mathbf{u}(\mathbf{x}, t) \cdot \nabla\rho_0(\mathbf{x}) + s(\mathbf{x}, t) \quad \text{(mass conservation)} \quad (3b)$$

$$p(\mathbf{x}, t) = c_0^2(\mathbf{x})(\rho(\mathbf{x}, t) + \mathbf{d} \cdot \nabla\rho_0(\mathbf{x})), \qquad \text{(pressure—density relation)} \quad (3c)$$

where p and ρ are the acoustic pressure and acoustic density fluctuations, $\mathbf{u}$ is the acoustic particle velocity (the time derivative of the acoustic displacement $\mathbf{d}$), s is a mass source term, and ρ_0 and c_0 are the ambient density and sound speed respectively. Both ρ_0 and c_0 are positive and bounded from above and below. The system (3a)–(3c) can be combined into the second-order wave equation for a heterogeneous medium:

$$\Box^2 p(\mathbf{x}, t) := \left(\frac{1}{c_0^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \rho_0(\mathbf{x})\nabla \cdot \left(\frac{1}{\rho_0(\mathbf{x})}\nabla\right)\right)p(\mathbf{x}, t) = \frac{\partial}{\partial t}s(\mathbf{x}, t), \quad (4)$$

where we introduced the inhomogenous d'Alembert operator $\Box^2$.

Initial and boundary conditions. As described in section 1.1, p_0 can be introduced as the initial value of the pressure rather than as a source term. Also, assuming the particle velocity **u** at $t = 0$ vanishes, (3a)–(3c) suggests that the time derivative of p should be initialized with zero. With these assumptions, the forward problem of PAT in d dimensions is stated as

$$\Box^2 p(\mathbf{x}, t) = 0, \qquad \text{in} \quad \mathbb{R}^d \times (0, \infty), \tag{5a}$$

$$p(\mathbf{x}, 0) = p_0(\mathbf{x}), \quad \text{in} \quad \mathbb{R}^d, \tag{5b}$$

$$\frac{\partial p}{\partial t}(\mathbf{x}, 0) = 0, \qquad \text{in} \quad \mathbb{R}^d. \tag{5c}$$

Because in practice it is possible to time-gate out any reflections from the experimental equipment, the propagation can be modelled as though it occurs in an unbounded domain i.e., no explicit boundary conditions are required.

Measurement. We deliberately model the measurement of the pressure field $p(\mathbf{x}, t)$ in two steps: Firstly, each element of the acoustic sensor array has only limited access to the pressure field: It only detects the field over a small, but finite, volume of space, and the measurement continues only for a finite time T. To signify this, we introduce the *window* function $\omega(\mathbf{x}, t) \in C_0^\infty(\Gamma \times (0, T))$, which maps $p(\mathbf{x}, t)$ to the field accessible by the sensor array: $g(\mathbf{x}, t) = \omega(\mathbf{x}, t)p(\mathbf{x}, t)$. Here, Γ is a d-dim, open, bounded set with non-zero measure that contains the cumulated volume of all sensor elements. The sound speed on Γ is assumed to be constant. In a second step, each sensor element transforms the accessible spatio-temporal part of the pressure field into a sequence of measured values. This process can be modelled by a measurement operator $\mathcal{M}$ which maps $g(\mathbf{x}, t) \in C_0^\infty(\Gamma \times (0, T))$ to the data $f \in \mathbb{R}^L$, $L \in \mathbb{N}$. Each type of ultrasonic detector elements commonly used for PAT (see [3, 24] for comprehensive reviews) leads to a different particular form of $\mathcal{M}$. For this work, we only assume that have access to the adjoint operator $\mathcal{M}^*$.

2.2. Derivation of the adjoint PAT operator

In this section, we derive the adjoint operator $\mathcal{A}^*$ to the forward operator

$$\mathcal{A} : C_0^\infty(\mathbb{R}^d) \to \mathbb{R}^L \tag{6a}$$

$$\mathcal{A}[p_0](\mathbf{x}, t) = \mathcal{M}\, \omega(\mathbf{x}, t)p(\mathbf{x}, t), \tag{6b}$$

where $p(\mathbf{x}, t)$ is the time dependent solution of the PAT forward problem (5a)–(5c). The assumption $p_0 \in C_0^\infty(\mathbb{R}^d)$, is justified by the short heat diffusion always present in the thermalisation process generating the initial pressure (see section 1.1). As we assume that $\mathcal{M}^*$ is given, $\mathcal{A}^* = \mathcal{P}^*\mathcal{M}^*$, where $\mathcal{P}^* : C_0^\infty(\Gamma \times (0, T)) \to C_0^\infty(\mathbb{R}^d)$ is the adjoint of

$$\mathcal{P} : C_0^\infty(\mathbb{R}^d) \to C_0^\infty(\Gamma \times (0, T)) \tag{7a}$$

$$\mathcal{P}[p_0](\mathbf{x}, t) = \omega(\mathbf{x}, t)p(\mathbf{x}, t), \qquad \text{in} \quad \Gamma \times (0, T), \tag{7b}$$

with respect to the generic L_2 bilinear form in $C_0^\infty(\Gamma \times (0, T))$ and $C_0^\infty(\mathbb{R}^d)$, respectively. By definition, this adjoint $\mathcal{P}^*$ has to satisfy the equality

$$\underbrace{\int_0^T \int_\Gamma g(\mathbf{x}, t)\mathcal{P}[p_0](\mathbf{x}, t)\mathrm{d}\mathbf{x}\mathrm{d}t}_{(\Diamond)} = \int_{\mathbb{R}^d} p_0(\mathbf{x})\mathcal{P}^*[g](\mathbf{x})\mathrm{d}\mathbf{x} \tag{8}$$

for any $p_0 \in C_0^\infty(\mathbb{R}^d)$, $g \in C_0^\infty(\Gamma \times (0, T))$. Therefore, we start with the left hand side $(\Diamond)$ of (8) and we seek to reformulate it in terms of an operator acting on the function g. The

solution of the initial value problem on an unbounded domain (5*a*)–(5*c*) can be written as

$$p(\mathbf{x}, t) = \frac{\partial}{\partial t} \int_{\mathbb{R}^d} G(\mathbf{x}, t|\mathbf{x}', 0) p_0(\mathbf{x}') \mathrm{d}\mathbf{x}', \tag{9}$$

where $G(\mathbf{x}, t|\mathbf{x}', t')$ is the causal Green's function of the homogeneous problem corresponding to (4) [2]:

$$\Box^2 G(\mathbf{x}, t|\mathbf{x}', t') = \delta_\mathrm{d}(\mathbf{x} - \mathbf{x}')\delta(t - t'). \tag{10}$$

Substituting $\mathcal{P}$, (7*b*), and $p(\mathbf{x}, t)$, (9) into ($\Diamond$) we obtain

$$(\Diamond) = \int_0^T \int_\Gamma g(\mathbf{x}, t)\omega(\mathbf{x}, t)\frac{\partial}{\partial t} \int_{\mathbb{R}^d} G(\mathbf{x}, t|\mathbf{x}', 0) p_0(\mathbf{x}') \mathrm{d}\mathbf{x}'\mathrm{d}\mathbf{x}\mathrm{d}t. \tag{11}$$

Shifting the time derivative of G into the integral and rearranging the order of integration, (11) becomes

$$(\Diamond) = \int_{\mathbb{R}^d} p_0(\mathbf{x}') \int_\Gamma \underbrace{\int_0^T g(\mathbf{x}, t)\omega(\mathbf{x}, t)\frac{\partial}{\partial t} G(\mathbf{x}, t|\mathbf{x}', 0) \mathrm{d}t}_{(\star)} \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}'. \tag{12}$$

Performing integration by parts in t on $(\star)$, we further obtain

$$(\Diamond) = -\int_{\mathbb{R}^d} p_0(\mathbf{x}') \int_\Gamma \int_0^T G(\mathbf{x}, t|\mathbf{x}', 0)\frac{\partial}{\partial t}(g(\mathbf{x}, t)\omega(\mathbf{x}, t))\mathrm{d}t\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}' \tag{13a}$$

$$+\int_{\mathbb{R}^d} p_0(\mathbf{x}') \int_\Gamma G(\mathbf{x}, T|\mathbf{x}', 0)g(\mathbf{x}, T)\omega(\mathbf{x}, T) - G(\mathbf{x}, 0|\mathbf{x}', 0)g(\mathbf{x}, 0)\omega(\mathbf{x}, 0)\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}'. \tag{13b}$$

The window function $\omega \in C_0^\infty(\Gamma \times (0, T))$ models which spatio-temporal domain is accessible to the sensor array (see section 2.1). As the detection time is restricted, $\omega(\cdot, 0) = \omega(\cdot, T) = 0$, and the second term (13*b*) vanishes. The next important step is to use the basic physical invariances of the Green's function [2],

$$G(\mathbf{x}, t|\mathbf{x}', t') = G(\mathbf{x}', -t'|\mathbf{x}, -t), \qquad \text{(time reversal invariance)} \tag{14a}$$

$$G(\mathbf{x}, t|\mathbf{x}', t') = G(\mathbf{x}, t + T|\mathbf{x}', t' + T), \qquad \text{(time translation invariance)} \tag{14b}$$

to rewrite (13*a*) as

$$(\Diamond) = -\int_{\mathbb{R}^d} p_0(\mathbf{x}') \int_\Gamma \int_0^T G(\mathbf{x}', T|\mathbf{x}, T - t)\frac{\partial}{\partial t}(g(\mathbf{x}, t)\omega(\mathbf{x}, t))\mathrm{d}t\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}'. \tag{15}$$

Applying a transformation of variables $t \leftarrow T - t$, (15) becomes

$$(\Diamond) = \int_{\mathbb{R}^d} p_0(\mathbf{x}') \int_\Gamma \int_0^T G(\mathbf{x}', T|\mathbf{x}, t)\frac{\partial}{\partial t}(g(\mathbf{x}, T - t)\omega(\mathbf{x}, T - t)\mathrm{d}t\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}'.$$

If we define

$$\mathcal{P}^*[g](\mathbf{x}') = \int_\Gamma \int_0^T G(\mathbf{x}', T|\mathbf{x}, t)\frac{\partial}{\partial t}(g(\mathbf{x}, T - t)\omega(\mathbf{x}, T - t))\mathrm{d}t\mathrm{d}\mathbf{x}, \tag{16}$$

and compare with (8), we found the adjoint operator $\mathcal{P}^* : C_0^\infty(\Gamma \times (0, T)) \to C_0^\infty(\mathbb{R}^d)$ (note that $\mathcal{P}^*[g](\mathbf{x}')$ is compactly supported due to the compact spatial support of ω, the finite propagation time and the upper bound on $c_0(\mathbf{x})$). With the above derivation we can see that the adjoint can be defined through $\mathcal{P}^*[g](\mathbf{x}') := q(\mathbf{x}', T)$, i.e. $\mathcal{P}^*$ maps g to the function $q(\mathbf{x}', t)$ evaluated at $t = T$, where q is given by the following wave equation:

$$\Box^2 q(\mathbf{x}', t) = \begin{cases} \dfrac{\partial}{\partial t}(g(\mathbf{x}', T - t)\omega(\mathbf{x}', T - t)) & \text{in } \Gamma \times (0, T) \\ 0 & \text{everywhere else} \end{cases} \tag{17a}$$

$$q(\mathbf{x}', 0) = 0 \qquad\qquad \text{in} \quad \mathbb{R}^d, \tag{17b}$$

$$\frac{\partial q}{\partial t}(\mathbf{x}', 0) = 0 \qquad\qquad \text{in} \quad \mathbb{R}^d. \tag{17c}$$

One way to compare the adjoint operator defined by the above system (17a)–(17c) to time reversal approaches (see section 1.2), is to choose a closed, sufficiently smooth $d - 1$ dimensional surface Θ inside the sensor volume Γ, on which we can impose a boundary condition. $\mathcal{P}^\triangleleft$ can then be expressed as mapping g to the solution $q(\mathbf{x}', t)$ of the following wave equation,

$$\Box^2 q(\mathbf{x}', t) = 0 \qquad\qquad \text{in} \quad \mathbb{R}^d \times (0, \infty) \backslash \Theta \times (0, T), \tag{18a}$$

$$q(\mathbf{x}', t) = g(\mathbf{x}', T - t)\omega(\mathbf{x}', T - t) \qquad \text{on} \quad \Theta \times (0, T), \tag{18b}$$

$$q(\mathbf{x}', 0) = 0 \qquad\qquad \text{in} \quad \mathbb{R}^d \backslash \Theta, \tag{18c}$$

$$\frac{\partial q}{\partial t}(\mathbf{x}', 0) = 0 \qquad\qquad \text{in} \quad \mathbb{R}^d, \tag{18d}$$

again, evaluated at $t = T$: $\mathcal{P}^\triangleleft[g](\mathbf{x}') \coloneqq q(\mathbf{x}', T)$. The main difference between time reversal and the adjoint is therefore given by how the time reversed pressure in the sensor element, $g(\mathbf{x}', T - t)\omega(\mathbf{x}', T - t)$, is introduced into the wave equation (4): In the adjoint (17a)–(17c), it enters as the mass source term $s(\mathbf{x}, t)$ while in time reversal (18a)–(18d), it enters as an explicit constraint on a sub-set.

3. Methods and implementation

3.1. Numerical wave propagation model

The adjoint model derived above could be implemented numerically using any of the standard techniques for modelling wave propagation in the time domain, so long as they allow to include a time-varying mass source term. One possible realization of such a scheme is given by the *k*-space time domain method [11, 25, 34], which is already widely used for forward modelling in photoacoustics, e.g. in the *k*-Wave Matlab Toolbox [31]. In appendix A, we give a full technical description of how to use *k*-Wave to implement discretised versions of the forward operator $\mathcal{A} = \mathcal{MP}$ (6a)–(6b), the adjoint operator $\mathcal{A}^* = \mathcal{P}^*\mathcal{M}^*$ (16) and the time reversal operator $\mathcal{A}^\triangleleft = \mathcal{P}^\triangleleft\mathcal{M}^*$ (18a)–(18d), which we will denote by $A \in \mathbb{R}^{L \times N}$, $A^* \in \mathbb{R}^{N \times L}$ and $A^\triangleleft \in \mathbb{R}^{N \times L}$, respectively. Note that for the discretisation, we restrict the spatial domain to a *d*-orthotope Ω (*bounding box*) with $\Gamma \subset \Omega$, which is enclosed by a *Perfectly Matched Layer* (*PML*) to model wave propagation on an unbounded domain.

3.2. Image reconstruction methods

We will now recapitulate some image reconstruction methods that rely on the discrete PAT operators described in the previous section. The first two elementary methods are obtained by directly applying the time reversal and the adjoint operator to the data:

(TR) $$p_{\mathrm{TR}} := A^{\triangleleft} f \tag{19}$$

(BP) $$p_{\mathrm{BP}} := A^{*} f. \tag{20}$$

In [27, 28], an enhancement of standard time reversal was proposed that takes the form of a Neumann series approach: The k-th iterate is obtained as

(iTR) $$p_{\mathrm{iTR}}^{k} = \sum_{j=0}^{k} K^{j} A^{\triangleleft} f, \qquad \text{with} \quad K = I_{N} - A^{\triangleleft} A, \tag{21}$$

where I_{N} is the identity in $\mathbb{R}^{N}$. A simple reformulation of (21) leads to an iterative image reconstruction algorithm:

$$p_{\mathrm{iTR}}^{k+1} = \sum_{j=0}^{k+1} K^{j} A^{\triangleleft} f = K \sum_{j=0}^{k} K^{j} A^{\triangleleft} f + A^{\triangleleft} f = K p_{\mathrm{iTR}}^{k} + A^{\triangleleft} f = (I_{N} - A^{\triangleleft} A) p_{\mathrm{iTR}}^{k} + A^{\triangleleft} f$$

$$\implies \quad p_{\mathrm{iTR}}^{k+1} = p_{\mathrm{iTR}}^{k} - A^{\triangleleft} (A p_{\mathrm{iTR}}^{k} - f). \tag{22}$$

Note that in [28], a modified time reversal operator is used that replaces (18c) by a harmonic extension of the pressure field $g(\mathbf{x}', T)$ to $\mathbb{R}^{d}$, which is particularly important if not all waves have left the domain, i.e., $p(\mathbf{x}, T) \neq 0$ for $\mathbf{x} \in \Gamma$. However, in the scenarios we consider in the numerical simulations, T is chosen large enough so that the power of the residual waves is well below the noise level of the sensors. We therefore neglect the harmonic extension in this study.

Next, we discuss iterative image reconstruction methods based on a variational model (2). The simplest case is given by omitting $\mathcal{J}(p)$, i.e., we are just looking for the least-squares solution

(LS) $$p_{\mathrm{LS}} := \operatorname*{argmin}_{p} \left\{ \tfrac{1}{2} \|A p - f\|_{2}^{2} \right\}. \tag{23}$$

In principle, p_{LS} is given as the solution of the normal equations, which we could iteratively solve with customized variants of the *conjugate gradient* scheme, such as the *conjugate gradient least-squares* (*CGLS*) algorithm [6]. Equipped with a suitable stopping criterion, it yields a well known and understood *regularization strategy* for solving (1), i.e., a scheme that is robust to noise and ill-conditioning (see [9, 14] and the references in [1]). However, for a better comparison to the previously presented iTR, we solve (23) by a *first order method* [8]. We only use the gradient of the cost function at p,

$$\nabla \tfrac{1}{2} \|A p - f\|_{2}^{2} = A^{*} (A p - f), \tag{24}$$

in a simple *gradient descent algorithm*: For a given initialization p^{0}, the iteration is given as

$$p_{\mathrm{LS}}^{k+1} = p_{\mathrm{LS}}^{k} - \eta A^{*} (A p_{\mathrm{LS}}^{k} - f), \tag{25}$$

which is very similar to (22): $A^{\triangleleft}$ has been replaced by A^{*} and we introduced a step size η.

As a slight modification of the above, we can incorporate the a-priori knowledge $p_{0} \geqslant 0$ and solve the positivity-constrained least-squares problem

(LS +) $$p_{\mathrm{LS+}} := \operatorname*{argmin}_{p \geqslant 0} \left\{ \tfrac{1}{2} \|A p - f\|_{2}^{2} \right\}, \tag{26}$$

by modifying (25) to a *gradient descent re-projection algorithm*:

$$p_{\mathrm{LS+}}^{k+1} = \Pi_{+} (p_{\mathrm{LS_+}}^{k} - \eta A^{*} (A p_{\mathrm{LS_+}}^{k} - f)), \tag{27}$$

where Π_{+} is a projection onto $\mathbb{R}_{+}^{N}$.

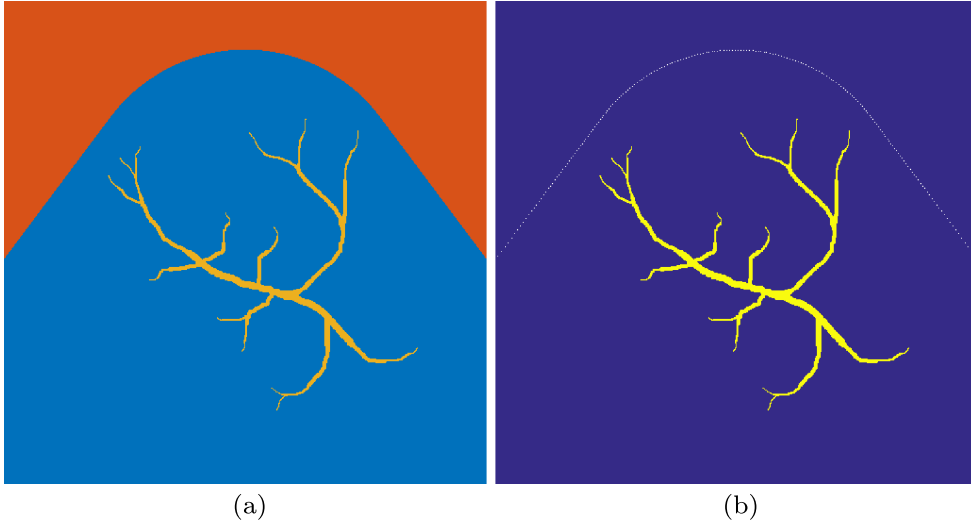(a)   (b)

Figure 1. 2D PAT scenario (II) used in numerical studies. The spatial resolution is $N = 512^2$ plus 24 pixels of PML layer in all directions. (a) The scenario cosists of three materials: material A (covering the top part of the domain): $c = 1500$, $\rho_0 = 1000$, material B (parabolic-like part): $c = 1400$, $\rho_0 = 1200$, material C (vessel-like part): $c = 1560$, $\rho_0 = 800$. (b) Ground truth p_0 and sensor configuration (white pixels, 200 sensors at the interface between materials A and B.)

Given the similarity between (22) and (25), it seems natural to define a positivity-constrained version of iterative time reversal as

$$p_{\text{iTR}+}^{k+1} = \Pi_+(p_{\text{iTR}_+}^k - A^\triangleleft(Ap_{\text{iTR}+}^k - f)). \tag{28}$$

The next variational method employs the *total variation* (*TV*) energy [7] as a regularizing functional:

$$(\text{TV}+) \qquad p_{\text{TV}+} := \underset{p \geqslant 0}{\text{argmin}} \left\{ \frac{1}{2} \|Ap - f\|_2^2 + \lambda \text{TV}(p) \right\}, \tag{29}$$

where $\lambda > 0$ is the regularization parameter and $\text{TV}(p)$ is the ℓ_1 norm of the amplitude of the gradient field of p (the details of the implementation are given in the appendix). We can solve (29), by modifying (27) from a projected to a *proximal gradient descent*:

$$p_{\text{TV}+}^{k+1} = \text{prox}_{\eta\lambda, \text{TV}+}(p_{\text{TV}_+}^k - \eta A^*(Ap_{\text{TV}+}^k - f)), \tag{30}$$

where the *proximal operator*

$$\text{prox}_{\alpha, \text{TV}+}(y) := \underset{x \geqslant 0}{\text{argmin}} \left\{ \frac{1}{2} \|x - y\|_2^2 + \alpha \text{TV}(x) \right\} \tag{31}$$

is given as a positivity-constrained total variation denoising problem. For our purposes, solving (31) by the algorithm described in [4] is sufficiently fast.

We chose the schemes (25), (27) and (30) because they are easy and intuitive to explain and provide an interesting connection to iterative time-reversal. In all the above algorithms, $\eta \in (0, 2/\theta)$ with θ being the largest singular value of A^*A, ensures the convergence to a

minimizer of the corresponding optimization problems, see [8]. For a given experimental PAT setting, θ only has to be computed once, which can be done by a simple power iteration.

4. Computational studies

4.1. Computational scenarios

To illustrate our results, we use two simple settings:

   I. A 3D scenario: $\Omega = [0, 1]^3$ discretised by $N = N_x^3$ isotropic voxels. The sound speed and ambient density are assumed to be homogeneous with $c = 1500 \, (\mathrm{m/s})$, $\rho_0 = 1000 \, (\mathrm{kg \ m^{-3}})$ and we add N_{PML} voxels of PML at all sides of the cube. The pressure is measured in the plane $x = 0$, which corresponds to the top layer of voxels. This situation is, e.g., encountered in planar Fabry-Pérot (FP) photoacoustic scanners [18, 22, 23, 42].
  II. A 2D scenario: $\Omega = [0, 1]^2$ with inhomogeneous medium properties and an irregular sensor geometry, see figure 1. The medium properties vary in the range encountered, e.g., in the human breast [19].

As our derivation in section 2 is independent of the concrete model for $\mathcal{M}$, we use simple point-sampling in space and time in both scenarios.

4.2. Precision assessment

The discretised operators A and A^* are not fully adjoint to each other anymore, which means that if we define

$$\chi[A, B](x, y) := |\langle Ax, y \rangle - \langle x, By \rangle|, \quad A \in R^{L \times N}, B \in R^{N \times L}, x \in R^N, y \in R^L, \quad (32)$$

then $\chi[A, A^*](x, y) > 0$ for most x, y. In this section, we investigate this error by relating it to other sources of error such as the modeling error caused by the PML or numerical errors due to limited precision and examine its influence on inverse reconstructions. As a 'baseline' to validate the *adjoint-then-discretize* approach described in this work, we will use the *discretize-then-adjoint* approach described in [17], which relies on the explicit adjoining of all the steps in the $\boldsymbol{k}$-Wave iteration (see appendix A). Hence, for the algebraic adjoint operator $\bar{A}^*$ obtained this way, $\chi[A, \bar{A}^*]$ should be 0 up to the numerical precision.

To test this, we computed (32) for 100 random realizations of x and y in scenario (I). The results are summarized in table 1. To interpret the results correctly, one has to bear in mind that the PML is known to be the largest limitation of numerical accuracy of the $\boldsymbol{k}$-Wave iteration [11], effectively preventing the use of double precision to improve numerical accuracy. Therefore, the most relevant results in table 1 are the ones for $n = 128$, PML $=16$ and single precision. In this setting, the statistics of the error (32) are comparable for A^* and $\bar{A}^*$. The results for double precision have to be regarded as a sanity check: While the error (32) for $\bar{A}^*$ is, as expected, smaller than for A^*, double precision is not used in practical computations. Nonetheless, we investigated the effects of this higher numerical precision in more detail in a second study using scenario (II). For this, we compare the iterates of scheme (27) to compute (26): In the first case, we use $\bar{A}^*$ as the adjoint operator, and in the second case, we use A^*. The corresponding iterates will be denoted by p_{alg}^k and p_{ana}^k, respectively. Figure 2 shows how the relative error between p_{alg}^k and p_{ana}^k develops with k. While the relative ℓ_2 error remains almost constant, the relative ℓ_∞ error grows slowly but does not indicate any instability.

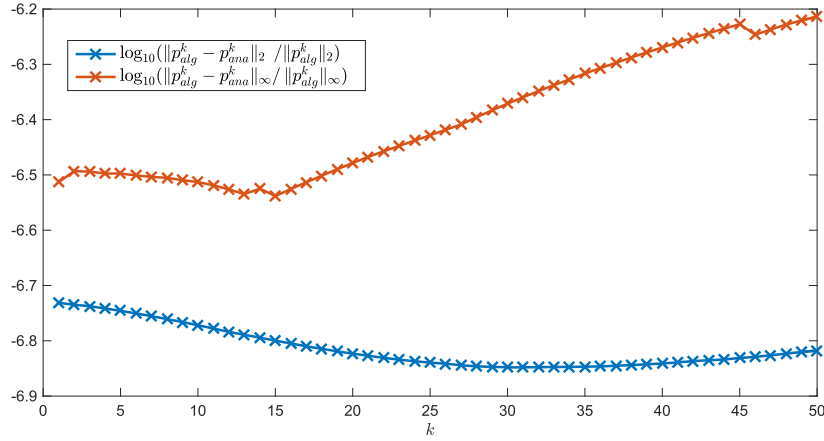**Figure 2.** Relative ℓ_2 and ℓ_∞ errors between the iterates of scheme (27) to compute (26) using either A^* or $\bar{A}^*$ as the adjoint operator. The corresponding iterates are denoted by p_{ana}^k and p_{alg}^k, respectively.

Table 2 completes our accuracy studies with the statistics of the ℓ_2 and ℓ_∞ norm of the distance $A^*y - \bar{A}^*y$ for 100 random realizations of y. We can see that in both norms the maximum/median discrepancy slightly decreases with decreasing N. However, we also see that this decrease can be compensated for by increasing the PML size. Furthermore, we cannot observe a difference between single and double precision in this measure anymore.

Summarising, the studies indicate that the proposed implementation of the analytical adjoint A^* is sufficiently close to the algebraic adjoint to be confidently used in variational reconstruction schemes.

4.3. Inverse reconstructions

Figure 3 shows the results of the different inverse methods applied to scenario (I) using $N = 256^3$ voxels and a sphere as the ground truth for p_0. Figure 4 shows the corresponding results for scenario (II), see, figure 1. The *peak-signal-to-noise ratio* (*PSNR*) of the reconstructed solutions was computed as:

$$\mathrm{PSNR}(p, q) = 10 \log_{10}\left(\frac{N}{\|\tilde{p} - \tilde{q}\|_2^2}\right), \qquad \text{where} \quad \tilde{p} = \mathrm{thres}\left(\frac{p}{\|p\|_\infty}, 0.01\right), \tag{33}$$

$$\tilde{q} \quad \text{accordingly}, \qquad \text{and} \qquad \mathrm{thres}(x, \alpha) = \begin{cases} x & \text{if } x \geqslant \alpha \\ 0 & \text{else} \end{cases} \tag{34}$$

All computations have been performed with single precision on a GPU. For all iterative methods, $K = 100$ iterations have been used and $\eta = 1.8/\theta$ was used as a step size. All results have been projected to $\mathbb{R}_+^N$ as a post-processing step.

In line with the theoretical prediction (see (17a)–(17c) and (18a)–(18d)) and the numerical implementation (see appendix B), the differences between BP and TR are rather subtle. A particular reason for this is that the measurement time T was chosen large enough, see [5]. However, figure 3 shows that the limited-view geometry of scenario (I) leads to differences in the decay of intensity with depth and the appearance of the well-known circular limited-view artefacts [13]. The differences between the iterative approaches based on time reversal or

Table 1. Statistics of the error $\chi[A, B](x, y)$, $B = A^*$ (*adjoint-then-discretize*) and $B = \bar{A}^*$ (*discretize-then-adjoint*), for 100 random realization of x and y. Displayed are maximum/median of $\log_{10}$ applied to the data.

single precision					
A^* (*adjoint-then-discretize*)			$\bar{A}^*$ (*discretize-then-adjoint*)		
	PML = 8	PML = 16		PML = 8	PML = 16
$N = 32^3$	$-2.86/-4.55$	$-3.45/-4.99$	$N = 32^3$	$-3.45/-5.14$	$-3.29/-5.14$
$N = 64^3$	$-3.12/-4.44$	$-2.92/-4.60$	$N = 64^3$	$-3.30/-4.68$	$-2.98/-4.63$
$N = 128^3$	$-2.00/-4.20$	$-1.78/-4.26$	$N = 128^3$	$-2.15/-4.22$	$-1.82/-4.24$

double precision					
A^* (*adjoint-then-discretize*)			$\bar{A}^*$ (*discretize-then-adjoint*)		
	PML = 8	PML = 16		PML = 8	PML = 16
$N = 32^3$	$-2.79/-4.58$	$-3.44/-5.15$	$N = 32^3$	$-12.44/-13.69$	$-12.12/-13.87$
$N = 64^3$	$-3.02/-4.72$	$-3.56/-5.28$	$N = 64^3$	$-11.41/-13.47$	$-11.75/-13.40$
$N = 128^3$	$-2.44/-4.77$	$-2.79/-5.46$	$N = 128^3$	$-10.72/-12.92$	$-10.97/-12.90$

(a) Ground truth p_0 (b) TR, PNSR = 15.71   (c) iTR, PSNR = 18.00   (d) iTR+, PSNR = 20.97

(e) TV+, λ = 0.05, PSNR = 23.57   (f) BP, PSNR = 15.75   (g) LS, PSNR = 18.072   (h) LS+, PSNR = 21.16
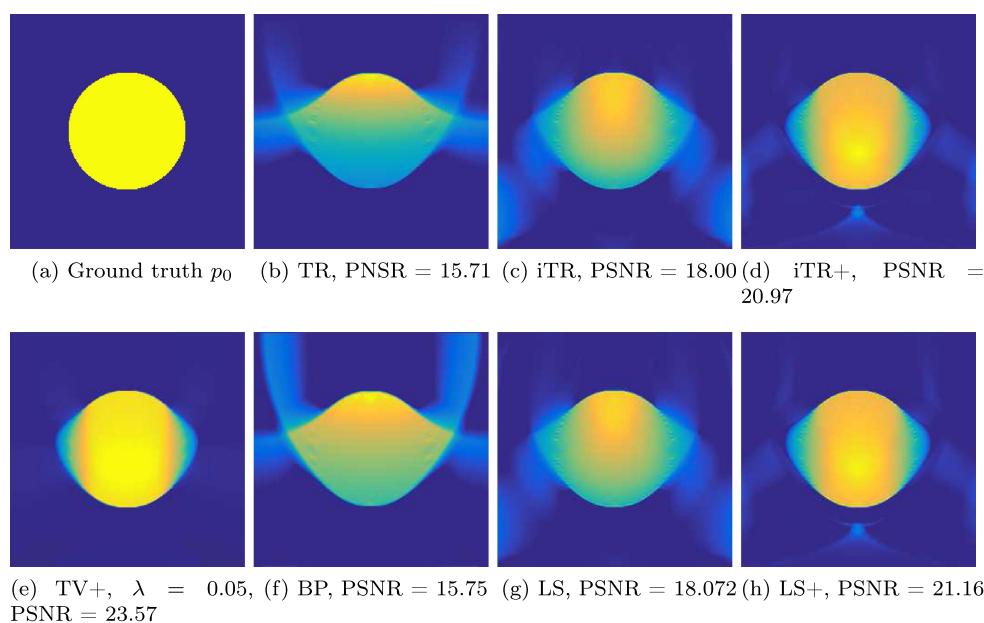
Figure 3. 2D slices at $y = 128$ through the 3D phantom p_0 and reconstructions thereof within scenario (I). The colorbar in each figure is scaled individually.

least-squares minimization are even more subtle than between plain TR and BP. The results rather show that including *a priori* information such as positivity or total variation constraints leads to a far more substantial improvement in terms of reconstruction quality.

Table 2. Statistics of the ℓ_2 and ℓ_∞ norm of $A^*y - \bar{A}^*y$ for 100 random realization of y (*adjoint-then-discretize* versus *discretize-then-adjoint*). Displayed are maximum/ median of $\log_{10}$ applied to the data.

single precision					
ℓ_2	PML $= 8$	PML $= 16$	ℓ_∞	PML $= 8$	PML $= 16$
$N = 32^3$	$-3.10/-3.13$	$-3.79/-3.84$	$N = 32^3$	$-4.21/-4.37$	$-4.91/-5.03$
$N = 64^3$	$-2.74/-2.76$	$-3.48/-3.50$	$N = 64^3$	$-4.20/-4.32$	$-4.83/-4.97$
$N = 128^3$	$-2.35/-2.36$	$-3.12/-3.13$	$N = 128^3$	$-4.14/-4.26$	$-4.78/-4.89$

double precision					
ℓ_2	PML $= 8$	PML $= 16$	ℓ_∞	PML $= 8$	PML $= 16$
$N = 32^3$	$-3.10/-3.13$	$-3.79/-3.84$	$N = 32^3$	$-4.21/-4.37$	$-4.90/-5.04$
$N = 64^3$	$-2.74/-2.77$	$-3.49/-3.51$	$N = 64^3$	$-4.20/-4.32$	$-4.83/-4.97$
$N = 128^3$	$-2.35/-2.36$	$-3.17/-3.18$	$N = 128^3$	$-4.14/-4.26$	$-4.79/-4.89$

5. Discussion, conclusions and outlook

In this work, we examined the adjoint of the continuous PAT operator (6*a*)–(6*b*) from different perspectives. We first described how the adjoint is connected to the solution of a wave equation with a time-varying mass source term (17*a*)–(17*c*) (see section 2). The simple proof was based on two main concepts:

- Our mathematical model (see section 2.1) of the measurement process explicitly accounts for the fact that the sensor array can only access the pressure field within a finite spatial volume and within a finite time interval. This was modelled by the introduction of a window function, which causes the boundary terms arising from the integration by parts in time to vanish in a central step in the proof, see (13*b*).
- The use of the Green's function representation of the solution of the wave equation (4), which reflects the temporal invariances of the underlying physical system, see (14*a*)–(14*b*).

We would like to remark that our setting is more general in terms of choice of Γ and T than the standard one, which assumes measurements on the boundary of the domain. In this standard setting, the boundary terms arising from the integration by parts in (13*b*) only vanish under the assumption that the support of p_0 is contained in the interior of the domain enclosed by the measurement surface and that the time T is large enough for all waves to leave this domain.

The explicit form of the continuous adjoint not only allows for an easy theoretical comparison to time-reversal (18*a*)–(18*d*), it also enables its numerical implementation with any existing wave propagation code that is able to include a time-varying mass source. As a particular example thereof, we examined an implementation based on the *k*-space pseudospectral time domain method (see sections 3 and 4):

- We demonstrated that our *adjoint-then-discretize* approach is sufficiently precise compared to the *discretize-then-adjoint* approach [17] that has to be derived for each specific wave propagation code separately and yields a discrete, purely algebraic, characterization of the adjoint only.
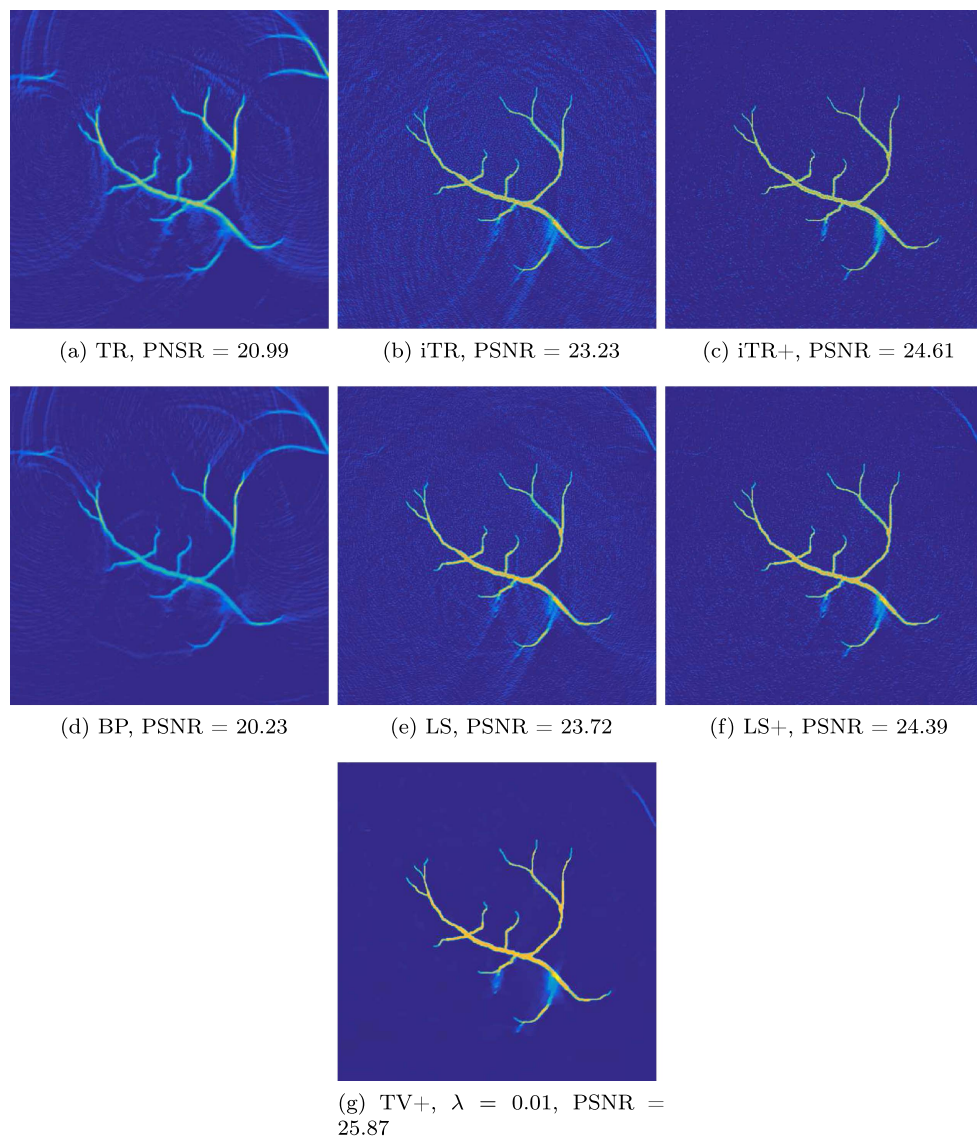
(a) TR, PNSR = 20.99   (b) iTR, PSNR = 23.23   (c) iTR+, PSNR = 24.61

(d) BP, PSNR = 20.23   (e) LS, PSNR = 23.72   (f) LS+, PSNR = 24.39

(g) TV+, λ = 0.01, PSNR = 25.87

Figure 4. Results of the different image reconstruction techniques applied to scenario (II), see figure 1. The colorbar in each figure is scaled individually.

• We illustrated how the adjoint can be used to solve variational image reconstruction problems (2) that incorporate a-priori constraints on the initial pressure.

An alternative description and implementation of the adjoint operator for PAT was recently presented in [5]: in comparison to our approach, the forward problem is modelled in a more restrictive way and with a different motivation, which, by using the weak formulation approach (as opposed to the integral formulation approach we take here) leads to the characterization of the adjoint as a wave transmission problem. The latter was solved by a specific

BEM-FEM approach, which is computationally intensive and the numerical studies focus on short recording times T with full boundary measurements.

Finally, these future directions of research are directly related to this work:

- The inverse reconstructions (see section 3.2 and 4.3) revealed interesting connections and differences between iterative methods based on time reversal and the adjoint operator. This topic has to be examined in more detail, both from theoretical and numerical perspectives: As (17a)–(17c) and (18a)–(18d) differ in the way the time-reversed pressure is introduced at the sensor elements, the effect of different sensor geometries, targets in close proximity to the sensors, artefacts caused by limited sensor coverage [13] and short recording times, T, need to be examined. Ultimately, this is not only a mathematical question: The different artefacts caused by time reversal and adjoint operator lead to different deficiencies in in-vivo images, each causing problems for the interpretation of those images.
- The optimization schemes (25), (27) and (30) were only used in this work because they are easy to describe and directly relate to iterative time-reversal. However, we note here that they are computationally rather inefficient and were not tuned for optimal performance. Therefore, the number of iterations used in the numerical examples does not provide a good indicator of the expected computational cost. More sophisticated algorithms for solving the corresponding optimization problems [8] will be examined in forthcoming work, in particular in the context of more challenging large-scale models encountered in many experimental data scenarios.
- Code to reproduce the results of this article will be provided as example scripts for a PAT image reconstruction toolbox that will be released to complement the k-`Wave` toolbox (see appendix A).

Acknowledgments

Appendix A. *k*-space time domain method

This appendix gives a technical overview of the k-space time domain scheme used for solution of the system of equations (3a)–(3c), as it is implemented in k-`Wave`[3] [31], as well as a description of how to discretise the continuous operators defined in section 2. The k-space time domain method is a collocation scheme that interpolates between the collocation points using a truncated Fourier series. This allows field gradients to be calculated efficiently using the *fast Fourier transform* (*FFT*). The 'k-space' in the name refers to a correction κ applied in the wavenumber domain to account for the finite difference approximation made to the time derivative. In the case of a homogeneous medium, this correction is exact, and no error arises from the temporal discretisation scheme [29]. For acoustically heterogeneous media, the k-space correction still reduces the errors, e.g., those due to numerical dispersion. To solve the system (3a)–(3c) using this scheme, it is first necessary to embed the spatial domain Ω (for simplicity, we assume for the spatial dimension $d = 3$ in the following description) into a

---

[3] For more details, see the k-`Wave` manual on www.k-Wave.org.

rectangular volume, which is then discretised by a regular grid of N collocation points. Then, a *Perfectly Matched Layer* (*PML*) absorbing boundary is wrapped around the box to damp outgoing waves without reflecting them, mimicking free-space propagation. The measurement time interval [0, T] is discretised by $t_n = n\Delta t$, $n = 0, \ldots, N_t$, $\Delta t = T/N_t$. The discrete pressure at iteration step n is denoted by p^n and the ξ-component of the discrete particle velocity by u_ξ^n. Here, ξ stands for one of the spatial components x, y, z of these vector fields. The terms containing $\nabla \rho_0(\mathbf{x})$ in (3a)–(3c) cancel out, and are therefore not included in the discrete scheme, to save unnecessary calculations [33]. The scalar 'density' is split, unphysically, into three parts, ρ_ξ^n, to allow anisotropic absorption to be included within the PML. The ambient density will be denoted by $\tilde{\rho}_0$ and the sound speed by c_0^2. The *k-space derivative* and the *k-space operator* κ are defined as

$$\frac{\partial}{\partial \xi} g := \mathbb{F}^{-1}\{i k_\xi \kappa \mathbb{F}\{g\}\}; \qquad \kappa = \mathrm{sinc}\left(\frac{c_{\mathrm{ref}}\Delta t}{2}\sqrt{k_x^2 + k_y^2 + k_z^2}\right), \qquad (\text{A.1})$$

where $k_\xi \in \mathbb{R}^N$ is a the discrete wavevector in ξ direction and all multiplications between N-dimensional vectors are understood as componentwise. c_{ref} is a homogeneous reference sound speed, chosen to ensure stability, e.g., $c_{\mathrm{ref}} = \max c_0(x)$. Staggering in time is included by interleaving the gradient and updates steps, as shown below. Grid-staggering is incorporated into the calculation of the gradients; a spatial translation of $\pm \Delta \xi$ is included as

$$\frac{\partial^{\pm}}{\partial \xi} g := \mathbb{F}^{-1}\{i k_\xi \kappa e^{\pm i k_\xi \Delta \xi/2} \mathbb{F}\{g\}\}. \qquad (\text{A.2})$$

The PML is implemented by multiplication operators Λ_ξ and Λ_ξ^s on the normal and staggered grid, respectively. Finally, we define $W : \mathbb{R}^N \longrightarrow \mathbb{R}^{N_\Gamma}$ to discretise the window function ω (see section 2.1).

*k***-Wave** *Iteration* Set $p^{-1}, u_\xi^{-3/2}$ and ρ_ξ^{-1} to zero (for all $\xi \in [x, y, z]$). Start the iteration at $n = -1$, and iterate until $N_t - 1$ ($t_0 = 0$ and $t_{N_t} = T$):

$$u_\xi^{n+1/2} = \Lambda_\xi^s\left(\Lambda_\xi^s u_\xi^{n-1/2} - \frac{\Delta t}{\tilde{\rho}_0}\frac{\partial^+}{\partial \xi}p^n\right) \qquad \text{(momentum conservation)} \qquad (\text{A.3})$$

$$\rho_\xi^{n+1} = \Lambda_\xi\left(\Lambda_\xi \rho_\xi^n - \Delta t \tilde{\rho}_0 \frac{\partial^-}{\partial \xi}u_\xi^{n+1/2}\right) + \Delta t\, s^{n+1/2} \qquad \text{(mass conservation)} \qquad (\text{A.4})$$

$$p^{n+1} = c_0^2(\rho_x^{n+1} + \rho_y^{n+1} + \rho_z^{n+1}) \qquad \text{(pressure–density relation)} \qquad (\text{A.5})$$

$$g^{n+1} = W p^{n+1} \qquad \text{(extract pressure)} \qquad (\text{A.6})$$

Note that equations (A.3) and (A.4) are really each three equations, for $\xi \in [x, y, z]$ separately.

Appendix B. Pseudospectral implementation of the PAT operators

Using the *k*-Wave iteration, the forward operator $\mathcal{P}$ (7a)–(7b) and the adjoint operator $\mathcal{P}^*$ (17a)–(17c) can be implemented as follows:

$\mathcal{P}$ is implemented by setting

$$s^{n+1/2} = \frac{1}{2c_0^2 d \Delta t} \begin{cases} Q p_0 & \text{for} \quad n = -1, 0 \\ 0 & \text{else} \end{cases} \tag{B.1}$$

and the result is given by $f = Mg$. Here, M is a discrete implementation of the measurement operator $\mathcal{M}$, i.e., it maps the discrete pressure time series in the sensor elements, $g \in \mathbb{R}^{N_t N_\Gamma}$, to the data $f \in \mathbb{R}^L$. In the simulation studies, M is just an identity matrix. In (B.1), Q is a smoothing operator introduced to eliminate the high spatial frequencies of p_0 to avoid unintended oscillations in the field being introduced by the band-limited interpolant (see section 2.8 in the $\textbf{\textit{k}}$-Wave manual and [32]).

$\mathcal{P}^*$ is implemented by setting

$$s^{n+1/2} = \frac{\tilde{\rho}_0}{2d \Delta t} W^* \begin{cases} \tilde{g}^{N_t} & \text{for} \quad n = -1 \\ \tilde{g}^{N_t-n+1} + \tilde{g}^{N_t-n} & \text{for} \quad n = 0, \ldots, N_t - 2 \\ \tilde{g}^1 + 2\tilde{g}^0 & \text{for} \quad n = N_t - 1 \end{cases} \tag{B.2}$$

with $\tilde{g} := M^* f$ and omitting (A.6). The result is given by $Q p^{N_t}/(c_0^2 \tilde{\rho}_0)$, where the smoothing by the self-adjoint Q has to be included to obtain adjointness with the forward implementation. The rescaling of p_t^N and $\tilde{g}$ by c_0^2 and $\tilde{\rho}_0$ is necessary since we are actually solving the adjoint of the first order system (3a)–(3c) instead of the second order system (4): p and ρ are essentially the same variables up to a scaling by c_0^2. In the first order adjoint system, the scaling is transferred from ρ to p. The same holds for $\rho_0(\mathbf{x})$, which is absent from the second order system.

It is interesting to compare the implementation of the adjoint with the implementation of TR where the time-reversed pressure field in the sensor elements, $\tilde{g}^{N_t-n+1}$, is introduced as a *Dirichlet source*: In the $\textbf{\textit{k}}$-Wave iteration, (A.4) is replaced by

$$\rho_\xi^{n+1} = (I_N - WW^*)\Lambda_\xi \left(\Lambda_\xi \rho_\xi^n - \Delta t \tilde{\rho}_0 \frac{\partial^-}{\partial \xi} \tilde{u}^{n+1/2} \right) + \frac{1}{dc_0^2} W^* \tilde{g}^{N_t-n}, \tag{B.3}$$

and step (A.6) is omitted. The result is given by $Q p^T$. In (B.3), the multiplication with $(I_N - WW^*)$ sets the density at the sensor locations to 0.

Appendix C. Discrete total variation functional

If we index the pixels of a 2D image $p \in \mathbb{R}^N$, with $N = N_x \times N_y$ by (i, j), $i = 1, \ldots, N_x$, $j = 1, \ldots, N_y$, a commonly used discretization of the total variation seminorm with Neumann boundary conditions relying on forward finite differences is given by

$$\mathrm{TV}(p) = \sum_{(i,j)}^N \sqrt{(p_{(i+1,j)} - p_{(i,j)})^2 + (p_{(i,j+1)} - p_{(i,j)})^2}, \tag{C.1}$$

where $p_{(N_x+1,j)} := p_{(N_x,j)}$, $p_{(i,N_y+1)} := p_{(i,N_y)}$. The implementation in 3D works in the same way.

References

[1] Arridge S R, Betcke M M and Harhanen L 2014 Iterated preconditioned lsqr method for inverse problems on unstructured grids *Inverse Problems* **30** 075009
[2] Barton G 1989 Elements of Green's functions and propagation *Potentials, Diffusion, and Waves* (Oxford: Oxford University Press)
[3] Beard P 2011 Biomedical photoacoustic imaging *Interface Focus* **1** 602–31

[4]   Beck A and Teboulle M 2009 Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems *IEEE Trans. Image Process.* **18** 2419–34

[5]   Belhachmi Z, Glatz T and Scherzer O 2016 A direct method for photoacoustic tomography with inhomogeneous sound speed *Inverse Problems* **32** 045005

[6]   Björck Å 1996 *Numerical Methods for Least Squares Problems* (Philadelphia, PA: SIAM)

[7]   Burger M and Osher S 2013 A Guide to the TV Zoo *Level Set and PDE Based Reconstruction Methods in Imaging* (*Lecture Notes in Mathematics*) (Berlin: Springer) 1–70

[8]   Burger M, Sawatzky A and Steidl G 2014 First order algorithms in variational image processing (arXiv:1412.4237)

[9]   Calvetti D, Reichel L and Shuibi A 2003 Enriched Krylov subspace methods for ill-posed problems *Linear Algebr. Appl.* **362** 257–73

[10] Cox B, Laufer J G, Arridge S R and Beard P C 2012 Quantitative spectroscopic photoacoustic imaging: a review *J. Biomed. Opt.* **17** 061202

[11] Cox B T, Kara S, Arridge S R and Beard P C 2007 k-space propagation models for acoustically heterogeneous media: application to biomedical photoacoustics *J. Acoust. Soc. Am.* **121** 3453–64

[12] Finch D and Patch S K 2004 Determining a function from its mean values over a family of spheres *SIAM J. Math. Anal.* **35** 1213–40

[13] Frikel J and Quinto E T 2015 Artifacts in incomplete data tomography with applications to photoacoustic tomography and sonar *SIAM J. Appl. Math.* **75** 703–25

[14] Hanke-Bourgeois M 1995 *Conjugate Gradient Type Methods for Ill-Posed Problems* (*Research Notes in Mathematics Series*) (Essex: Longman Scientific & Technical)

[15] Hristova Y, Kuchment P and Nguyen L 2008 Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media *Inverse Problems* **24** 055006

[16] Huang C, Nie L, Schoonover R W, Guo Z, Schirra C O, Anastasio M A and Wang L V 2012 Aberration correction for transcranial photoacoustic tomography of primates employing adjunct image data *J. Biomed.Opt.* **17** 066016

[17] Huang C, Wang K, Nie L, Wang L and Anastasio M 2013 Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media *IEEE Trans. Med. Imaging* **32** 1097–110

[18] Jathoul A P *et al* 2015 Deep *in vivo* photoacoustic imaging of mammalian tissues using a tyrosinase-based genetic reporter *Nat. Photon.* **9** 239–46

[19] Jin X and Wang L V 2006 Thermoacoustic tomography with correction for acoustic speed variations *Phys. Med. Biol.* **51** 6437

[20] Kruger R A, Lam R B, Reinecke D R, Del Rio S P and Doyle R P 2010 Photoacoustic angiography of the breast *Med. Phys.* **37** 6096–100

[21] Kuchment P and Kunyansky L 2011 Mathematics of photoacoustic and thermoacoustic tomography *Handbook of Mathematical Methods in Imaging* ed O Scherzer (New York: Springer) pp 817–65

[22] Laufer J, Johnson P, Zhang E, Treeby B, Cox B, Pedley B and Beard P 2012 In vivo preclinical photoacoustic imaging of tumor vasculature development and therapy *J. Biomed. Opt.* **17** 056016

[23] Laufer J, Norris F, Cleary J, Zhang E, Treeby B, Cox B, Johnson P, Scambler P, Lythgoe M and Beard P 2012 In vivo photoacoustic imaging of mouse embryos *J. Biomed. Opt.* **17** 061220

[24] Lutzweiler C and Razansky D 2013 Optoacoustic imaging and tomography: reconstruction approaches and outstanding challenges in image performance and quantification *Sensors* **13** 7345

[25] Mast T, Souriau L, Liu D- L, Tabei M, Nachman A and Waag R 2001 A k-space method for large-scale models of wave propagation in tissue *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control* **48** 341–54

[26] Nie L and Chen X 2014 Structural and functional photoacoustic molecular tomography aided by emerging contrast agents *Chem. Soc. Rev.* **43** 7132–70

[27] Qian J, Stefanov P, Uhlmann G and Zhao H 2011 An efficient neumann series—based algorithm for thermoacoustic and photoacoustic tomography with variable sound speed *SIAM J. Imaging Sci.* **4** 850–83

[28] Stefanov P and Uhlmann G 2009 Thermoacoustic tomography with variable sound speed *Inverse Problems* **25** 075011

[29] Tabei M, Mast T D and Waag R C 2002 A k-space method for coupled first-order acoustic propagation equations *J. Acoust. Soc. Am.* **111** 53–63

[30] Taruttis A and Ntziachristos V 2015 Advances in real-time multispectral optoacoustic imaging and its applications *Nat. Photon.* **9** 219–27

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

[32] Treeby B E and Cox B T 2011 A k-space Green's function solution for acoustic initial value problems in homogeneous media with power law absorption *J. Acoust. Soc. Am.* **129** 3652–60

[33] Treeby B E, Jaros J, Rendell A P and Cox B T 2012 Modeling nonlinear ultrasound propagation in heterogeneous media with power law absorption using a k-space pseudospectral method *J. Acoust. Soc. Am.* **131** 4324–36

[34] Treeby B E, Zhang E Z and Cox B T 2010 Photoacoustic tomography in absorbing acoustic media using time reversal *Inverse Problems* **26** 115003

[35] Wang L V 2009 Multiscale photoacoustic microscopy and computed tomography *Nat. Photon.* **3** 503–9

[36] Wang X, Pang Y, Ku G, Xie X, Stoica G and Wang L V 2003 Noninvasive laser-induced photoacoustic tomography for structural and functional *in vivo* imaging of the brain *Nat. Biotechnol.* **21** 803–6

[37] Xia J and Wang L V 2014 Small-animal whole-body photoacoustic tomography: a review *IEEE Trans. Biomed. Eng.* **61** 1380–9

[38] Xu Y and Wang L V 2004 Application of time reversal to thermoacoustic tomography *Proc. SPIE* **5320** 257

[39] Xu Y and Wang L V 2004 Time reversal and its application to tomography with diffracting sources *Phys. Rev. Lett.* **92** 033902

[40] Yao J and Wang L V 2014 Photoacoustic Brain Imaging: from Microscopic to Macroscopic Scales *Neurophotonics* **1** 011003

[41] Zackrisson S, van de Ven S M W Y and Gambhir S S 2014 Light in and sound out: emerging translational strategies for photoacoustic imaging *Cancer Res.* **74** 979–1004

[42] Zhang E Z, Laufer J G, Pedley R B and Beard P C 2009 In vivo high-resolution 3D photoacoustic imaging of superficial vascular anatomy *Phys. Med. Biol.* **54** 1035