Physics in Medicine & Biology

PAPER • OPEN ACCESS

Quantitative *in-vivo* full-waveform ultrasound tomography workflow integrating reflection imaging and resolution analysis

To cite this article: I E Ulrich *et al* 2026 *Phys. Med. Biol.* **71** 135039

IOP Publishing Phys. Med. Biol. 71 (2026) 135039 https://doi.org/10.1088/1361-6560/ae7cd6

Physics in Medicine & Biology

OPEN ACCESS

RECEIVED
10 December 2024

REVISED
29 May 2026

ACCEPTED FOR PUBLICATION
12 June 2026

PUBLISHED
10 July 2026

PAPER

Quantitative *in-vivo* full-waveform ultrasound tomography workflow integrating reflection imaging and resolution analysis

I E Ulrich[1,*], C Boehm[1], P Marty[1], S Noe[1], N Korta Martiartu[2,3], X L Dean-Ben[4,5], D Razansky[4,5] and A Fichtner[1]

1 Institute of Geophysics, ETH Zurich,Sonneggstrasse 5, 8092 Zurich, Switzerland
2 Institute of Applied Physics, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland
3 Institut Langevin, ESPCI Paris, PSL University, CNRS, 1 rue Jussieu, F-75005 Paris, France
4 Institute of Pharmacology and Toxicology and Institute for Biomedical Engineering, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
5 Institute for Biomedical Engineering, ETH Zurich,Wolfgang-Pauli-Strasse 27, 8049 Zurich, Switzerland
* Author to whom any correspondence should be addressed.

E-mail: ineselisa.ulrich@gmx.de

Keywords: full-waveform inversion, image reconstruction, *in-vivo* data, reflection imaging, resolution analysis, sound speed estimation, ultrasound tomography.

Supplementary material for this article is available online

Abstract

Objective. To demonstrate the potential of full-waveform inversion (FWI) for high-resolution ultrasonic imaging using *in-vivo* data. *Approach.* Acoustic FWI is applied to *in-vivo* measurements, accounting for the nonlinear relationship between the ultrasonic wavefield and model parameters. Two key components are investigated: (1) estimation of an effective source wavelet by inverting the source-time function using calibration data in water, and (2) reduction of nonlinearity and sensitivity to the initial model using a graph-space optimal transport misfit functional. A perturbation-based resolution analysis is employed to quantify local spatial smearing. *Main results.* The proposed approach enables improved spatial resolution in reconstructed models and allows accurate identification of anatomical features in *in-vivo* data. *Significance.* These results provide further evidence for the applicability of FWI to *in-vivo* ultrasonic imaging and highlight methodological components that influence reconstruction quality, offering insight into its potential for clinical applications.

1. Introduction

Medical ultrasound has long offered clinicians and researchers a unique tool for non-invasively interrogating tissue structure and function, making it a cornerstone e.g. of prenatal care, cardiology, and emergency medicine. More recently, transmission ultrasound tomography has been developed to provide quantitative images of tissue parameters such as sound speed and acoustic attenuation, with applications ranging from breast imaging and cancer screening (Yamashita *et al* 2024) to preclinical small-animal imaging (Lafci *et al* 2023). Clinical transmission ultrasound systems such as SoftVue and QT Ultrasound (Pratt *et al* 2007, Wiskin *et al* 2020) have demonstrated promising *in-vivo* imaging capabilities in humans. Alternative transmission configurations, including acoustic-mirror-based approaches, have also been investigated (Chintada *et al* 2022). The low operational costs of transmission ultrasound relative to alternative approaches, combined with the absence of ionizing radiation and the potential to implement portable devices, are poised to have high clinical impact.

Two approaches have been adopted to map internal tissue properties using recorded ultrasound waves, one may take two approaches: (1) focus on particular phases in the measured data, specifically the first arrivals in transmission or reflection, or (2) consider the entire wavefield. Approach (1) is the subject of conventional pulse-echo, or reflection ultrasound methods, where the travel time is computed using a linearized wave propagation model assuming single scattering. A prominent example is the delay-and-sum algorithm (Jensen *et al* 2006, Gemmeke *et al* 2017), which produces images of impedance contrasts within the tissue. Information on tissue parameters such as sound speed or attenuation may be obtained with ray tomography (Gemmeke *et al* 2017) that exploits the first arrivals of transmitted waves. In contrast to this, waveform-based imaging approaches aim to exploit the full time series of an ultrasound measurement to extract information on tissue parameters, hence the name full-waveform inversion (FWI). FWI is motivated by the fact that first arrivals constitute only a fraction of the measured signals with a large majority of data being contained in reflected, diffracted, and scattered waveforms encoding information on the tissue structure off the direct ray path.

In FWI, a model of the tissue properties is constructed via an iterative inversion process that harnesses numerical solutions to the wave equation to minimize their misfit with respect to the observed data (Tarantola 1984, Virieux and Operto 2009, Fichtner 2011). Taking into account sparse coverage, noise, attenuation, and realistic transducers, FWI can to achieve a spatial resolution close to the diffraction limit on the order of half the shortest wavelength (Devaney 1984). This improves upon ray tomography where the spatial resolution is limited by the width of the first Fresnel zone (Schatzberg and Devaney 1992, Cerveny 2001). The ability of FWI to reconstruct tissue properties such as sound speed or attenuation accurately from medical ultrasound data has been demonstrated for laboratory measurements (Korta Martiartu *et al* 2020), photoacoustic imaging (Huang *et al* 2013, Mitsuhashi *et al* 2017), soft tissue imaging (Pratt *et al* 2007, Wiskin *et al* 2020), *in vivo* human studies (Iuanow *et al* 2017, Littrup *et al* 2021), as well as transcranial ultrasound (Guasch *et al* 2020, Marty *et al* 2021, Mitcham *et al* 2023). However, *in vivo* imaging based on FWI has so far been challenged by the narrow bandwidth of common ultrasound transducers and the requirement for accurate starting models, both of which relate to the non-convex nature of the misfit characterized by local minima that are difficult to navigate using gradient-based optimization methods typically employed in FWI. Additionally, a central aspect of accurately simulating the ultrasonic wavefield is precise knowledge of the source-time function. While manufacturer-provided source-time functions can be used, in practice they may not capture the fine temporal details needed for optimal FWI reconstructions. As shown in the results, calibrating the source-time functions directly from the measured data can lead to noticeably improved image quality and resolution. Lastly, a resolution analysis in FWI is challenging since the data depend nonlinearly on the model. This implies that strategies to describe the posterior in linear inverse problems are not applicable to FWI. The objective of this study is therefore threefold: (1) to broaden the range of applicability of FWI for *in vivo* transmission ultrasound by choosing a more robust measure of misfit that can mitigate cycle skipping and allows for less informative starting models; (2) to increase the tomographic resolution by performing an auxiliary inversion for the source-time function; and (3) to propose a strategy for a resolution analysis that is applicable in practice without significantly adding to the computational costs.

We demonstrate the effectiveness of the proposed concepts with a time-domain acoustic FWI using an *in-vivo* ultrasound data set of a mouse. The acquisition device is a transmission-reflection optoacoustic ultrasound imaging platform (Lafci *et al* 2022), which allows for the joint imaging of optical absorption, ultrasound reflectivity, and acoustic properties. To overcome the tendency of FWI to stall in a local minimum, we follow Métivier *et al* (2019) and define the misfit in terms of graph-space optimal transport (GOT). To calibrate the source wavelet, we draw inspiration from seismic imaging (Groos *et al* 2014) and invert for each source-time function individually in an auxiliary linear inversion. This takes advantage of the water shot, which is routinely collected during data acquisition (DAQ). With this setup, the tissue sound speed map can be reconstructed with high accuracy, allowing for the differentiation between small-scale internal structures delineated in the corresponding reflection image. Lastly, we propose a practical method for resolution analysis to quantify the quality of the reconstructed images. For this, we infer curvature information from the Hessian of the misfit. With this, point-spread functions (PSFs) can be approximated that assess the local resolution lengths of reconstructed features. To render the proposed resolution analysis applicable in practice, we employ source encoding (Capdeville *et al* 2005, Krebs *et al* 2009, Wang *et al* 2015), a well-established technique to reduce the computational costs of forward and adjoint computations in FWI.

FWI-based mapping of tissue parameters *in vivo* is shown to result in enhanced resolution and contrast, thereby paving the way for a wider dissemination of ultrasound transmission tomography in preclinical and clinical settings.

2. Theory

2.1. Forward problem

The backbone of FWI is the accurate numerical modeling of ultrasound propagation. The aim is to estimate a model of the tissue properties $\mathbf{m}$, e.g. the sound speed $c(\mathbf{x})$ and mass density $\rho(\mathbf{x})$, that predicts a set of measurements of the acoustic pressure $p(\mathbf{x},t)$. We consider a setup where the domain of interest Ω consists of a water bath with the immersed mouse model enclosed by a transducer array, which delineates the boundary of the domain. For soft tissues, the governing equation for acoustic wave propagation is given by

$$\frac{1}{\rho(\mathbf{x})\,c^2(\mathbf{x})}\partial_t^2 p(\mathbf{x},t) - \nabla\cdot\left(\frac{1}{\rho(\mathbf{x})}\nabla p(\mathbf{x},t)\right) = \frac{1}{\rho(\mathbf{x})}f(\mathbf{x},t), \tag{1}$$

where $f(\mathbf{x},t)$ denotes the external sources. Neumann, Dirichlet or absorbing boundary conditions may be enforced along different parts of the domain boundary. To discretize (1) we use the spectral-element method Afanasiev *et al* (2019) and define $\mathbf{p} = [p(\mathbf{x},t)$ for $t\in[t_0,t_1]$ and $\mathbf{x}\in\Omega]$ as the pressure wavefield, $\mathbf{L}$ as the discrete wave operator in 2 dimensions, and let $\mathbf{f}$ represent a point source scaled by the inverse of the density at that point, which makes it the discrete counterpart to the scalar source-time function on the right-hand side of (1).

In a discrete setting, the pressure wavefield is observed at receiver positions $\mathbf{x}_r$, where $r = 1,2,\ldots,n_r$. The recorded signals $\mathbf{d} = (\mathcal{R}\mathbf{p})_r$ are extracted from the total field by application of a restriction operator $\mathcal{R}$. In practice, each receiver has its own impulse response that modifies the measured waveforms. In the present framework, this effect is absorbed into the estimated source-time function, which effectively accounts for the combined source and receiver system response, allowing the recorded data to be interpreted consistently in terms of pressure. We then write the discretized version of (1) compactly in terms of a matrix vector product

$$\mathbf{L}(\mathbf{m};\mathbf{x},t)\,\mathbf{p}(\mathbf{m}) = \mathbf{f}, \quad \mathbf{x}\in\Omega\subset\mathrm{R}^N,\ \ t\in[t_0,t_1]\subset\mathrm{R}. \tag{2}$$

Here, $\mathbf{L}$ is linear in the pressure wavefield $\mathbf{p}$ for a fixed set of model parameters $\mathbf{m} = [\rho(\mathbf{x});c(\mathbf{x})]$ and a given source $\mathbf{f}$. However, the parameter-to-solution map $\mathbf{m}\mapsto\mathbf{p}(\mathbf{m})$ is nonlinear.

2.2. FWI

FWI is formulated as a nonlinear, local constrained optimization procedure with the objective to minimize the waveform misfit χ between simulated data $\mathbf{d}(\mathbf{m}) = (\mathcal{R}\mathbf{p}(\mathbf{m}))_r$ and the measured data $\mathbf{d}^{\mathrm{obs}}$. To efficiently compute gradients of χ with respect to $\mathbf{m}$, FWI uses the adjoint method (Tarantola 1988, Pratt 1999, Tromp *et al* 2004). To construct an augmented objective function ψ, a Lagrange multiplier $\mathbf{q}$ can be used to include the constraint, which is the forward equation in (2):

$$\psi(\mathbf{m},\mathbf{p},\mathbf{q}) = \chi\left((\mathcal{R}\mathbf{p})_r;\mathbf{d}^{\mathrm{obs}}\right) - \mathbf{q}^T(\mathbf{L}\mathbf{p}-\mathbf{f}). \tag{3}$$

Equation (3) is minimized by forcing all partial derivatives to zero. This provides formulations for the gradient in a perturbation direction $\delta\mathbf{m}$,

$$\nabla_{\mathbf{m}}\chi(\mathbf{m})\,\delta\mathbf{m} = -\mathbf{q}^T\nabla_{\mathbf{m}}\mathbf{L}\mathbf{p}\delta\mathbf{m}, \tag{4}$$

and the adjoint equation

$$\mathbf{L}^T\mathbf{q} = \nabla_{\mathbf{p}}\chi, \tag{5}$$

where $\mathbf{q}$ plays now the role of the adjoint wavefield, $\mathbf{L}^T$ is the adjoint wave operator and $\nabla_{\mathbf{p}}\chi$ is the adjoint source. To compute the gradient in (4), one therefore needs two wavefield simulations: one to solve the forward equation (2) for a given $\mathbf{m}$ and one to solve the adjoint equation in (5), propagating the data misfit backward in time simultaneously from all receivers. To find the descent direction $\delta\mathbf{m}$, we use a quasi-Newton optimization technique based on trust-region L-BFGS (Boehm *et al* 2018) where the misfit functional $\chi(\mathbf{m})$ is approximated quadratically and optimized in the vicinity of the current model, the trust region. Within the L-BFGS approximation, the action of the inverse of the approximate Hessian on a vector can be calculated, which we will exploit during the resolution analysis in section 2.6.

2.3. Optimal transport misfit functional

The shape of the misfit landscape is controlled by the choice of χ. In FWI, χ is commonly defined as the L_2 distance between observed and measured data (Tarantola 1988, Pratt *et al* 1998), which is sensitive to amplitude mismatches and inherently dominated by large-amplitude waveforms, amplifying the effects of large differences between predicted and measured data. In addition, valuable information contained in time shifts of low-amplitude waves tends to be under-represented in the L_2-misfit. This makes the L_2-misfit functional sensitive to the nonlinearities with respect to the model parameters and may lead to cycle skipping (Virieux and Operto 2009, Warner and Guasch 2016). Unlike the conventional L_2-norm, GOT quantifies waveform differences in a joint time–amplitude space, making it less sensitive to phase shifts and better suited to avoid local minima during inversion. The amplitude-time samples of individual signals are thereby mapped to discrete point clouds, representing the nodes in a graph. The distance between individual points of the two sets now has two components: amplitude and time. The aim of GOT is to find the optimal transport plan that matches points between point clouds pair-wise such that the total cost is minimized. The cost is defined as the cumulative space-time distance across all pairs, where the space component is the amplitude difference and the time component represents the relative temporal offset between individual points. An effective measure for the transport cost from one point cloud to another is the Wasserstein metric, which, in contrast to L_2, takes into account the geometry of the space where the point clouds are defined. The resulting mapping is then computed by solving an auxiliary optimization problem for the computation of the misfit, reformulated as a linear sum assignment problem (LSAP) (Boehm *et al* 2022). The solution to the LSAP defines a point-wise mapping between the measured and simulated signals, indicating how each sample in one signal corresponds to a sample in the other. Mathematically, this correspondence enables an important feature of GOT: by adjusting the relative weighting of the amplitude and time components in the cost function, the misfit functional can be made convex with respect to the point-wise temporal offsets. This mapping, however, is not static, but can be tuned by weighting the distance in amplitude versus the distance in time. The tuning parameter that controls this weighting therefore defines the convexity of the misfit functional. Figure 1(a) shows how different choices of the tuning parameter widen or steepen the valley of attraction, comparing a signal of observed data to time-shifted versions of a simulated signal, taken from the *in-vivo* FWI inversion results presented in section 3.3. The minimum indicates that the simulated signal must be shifted by around 1.5 μs to optimally fit the observed signal, as indicated by the line [1] and seen in figure 1(b). A 5.2 μs offset produces a signal that is out of phase with the observed data, as indicated by the line [2] and figure 1(c). Figures 1(d) and (e) display the point-to-point mapping between the signals, where the arrows indicate how individual samples are associated according to GOT. A small maximum time shift of 0.4 μs emphasizes amplitude mismatches, effectively reproducing an L_2 misfit. Increasing the maximum expected time shift allows points to be correctly matched in time, aligning time-shifted peaks and troughs with their counterparts in the observed signal.

Computationally, GOT is more demanding compared to least-squares or cross-correlation misfit functionals. Métivier *et al* indicate an increase of 6% in computational cost of GOT compared to a least-squares misfit functional, based on numerical data. This value should be understood as an order-of-magnitude estimate rather than a general result, since the actual cost depends on factors such as the number of samples and the maximum expected time shift. For real data with time-shifted and distorted waveforms, it is therefore attractive to first start the inversion at relatively low frequencies with a GOT misfit functional, using a large maximum expected time shift (assuming the starting model is rather far from the optimal model), and then subsequently decrease the tuning parameter during the inversion. Once sufficient progress has been made, one may switch to a least-squares misfit functional for the final iterations to speed up convergence. We adopt this strategy for our *in-vivo* FWI in section 3.3.

2.4. Reflection imaging using reverse-time migration (RTM)

In medical ultrasound, reflection images are commonly acquired via beamforming techniques using, for instance, delay-and-sum (see figure 2(b)). In most practical implementations, the time delays are computed using simplified propagation assumptions that do not fully account for finite-frequency wave effects, making these methods efficient for obtaining structural images from high-frequency data (>1 MHz) (Stotzka *et al* 2005). In this study, we apply RTM (Levin 1984, Roy *et al* 2016), which is a back-propagation technique leveraging numerical solutions to the wave equation. Both, RTM and FWI build on the same imaging condition that correlates the forward and adjoint wavefield (Chattopadhyay and McMechan 2008, Dai and Schuster 2013) to generate an image $I(\mathbf{x})$ of either the reflectivity distribution (RTM) or the model parameter values (FWI). For RTM, the adjoint wavefield is equal to the time-reversed measured wavefield $q(\mathbf{x}, T-t)$ propagated from the receivers. The image $I(\mathbf{x})$ is then

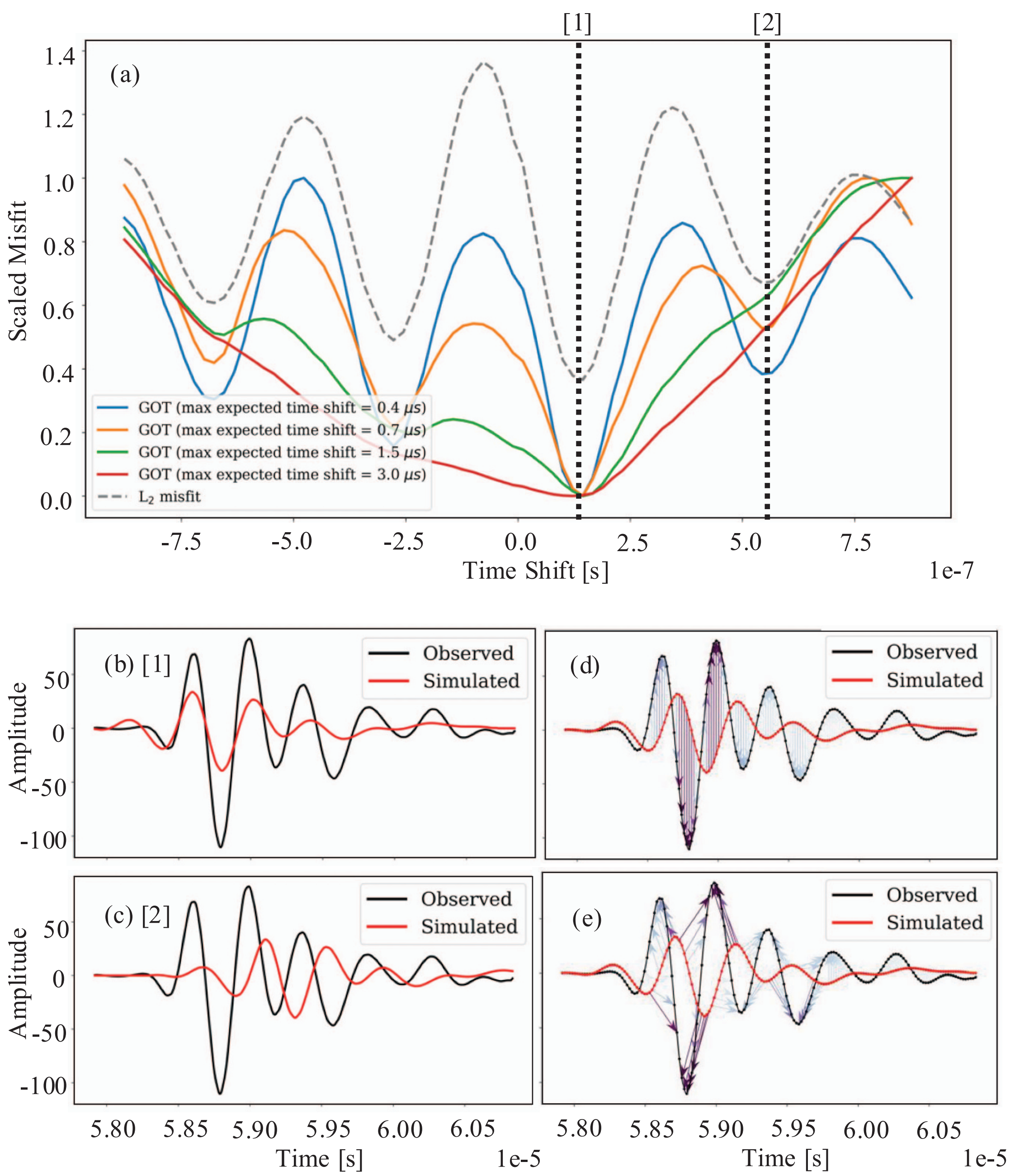

Figure 1. (a) Optimal transport misfit for two time-shifted signals using different values for the tuning parameter (maximum expected time shift). Choosing a larger maximum time shift for the GOT misfit functional eliminates local minima and widens the valley of attraction. (b) and (c) show one signal of simulated and observed data, with the simulated trace shifted to match the observed signal in (b) which shows the observed and simulated signals at the vertical dotted line denoting the time shift at 1.25 s or deliberately shifted out of phase in (c), which is linked to the vertical line at approximately 5.5 s. (d) and (e) show the point cloud matching for two time-shifted signals using different values for the tuning parameter. The color of the arrows indicates the cost of the assignment in graph space, with darker colors corresponding to higher costs. The arrows illustrate the mapping between points of the simulated and observed signals used for computing the GOT misfit and its corresponding adjoint source. A small tuning parameter (d) emphasizes amplitude differences and reproduces the behavior of an L_2 misfit. The tuning parameter corresponds here to the maximum expected time shift chosen for the blue curve in figure 1(a). Increasing the maximum expected time shift (e) reduces the cost of moving points in time, allowing time-shifted peaks and troughs to align correctly with their counterparts in the observed signal.

formed by summing the contributions from all sources, where s denotes the source index and $p_s(\mathbf{x}, t)$ and $q_s(\mathbf{x}, T-t)$ are the forward and adjoint wavefields associated with source s, respectively,

$$I(\mathbf{x}) = \sum_s \sum_{t=0}^{T} \partial_t^2 p_s(\mathbf{x}, t)\, q_s(\mathbf{x}, T-t)\,. \tag{6}$$

This allows us to generate a reflectivity image with the same setup that is used for the sound speed inversion by solely changing the definition of the misfit functional χ in (5). Douma et al (2010) show that when the first arrival is removed, the background medium is smooth and when defining χ to be the least-squares misfit between the forward propagated simulated wavefield and the adjoint wavefield, the image $I(\mathbf{x})$ is given by the impedance kernel.

2.5. Source-wavelet estimation

A crucial ingredient of FWI is an accurate description of the source wavelet used to generate the simulated data (Cueto *et al* 2021). To account for individual transducer characteristics, we perform a source inversion prior to each frequency band during the multiscale inversion, an implementation initially described in Ulrich *et al* (2024). We take inspiration from seismic imaging where inverting for the source mechanism is often necessary to account for the complex nature of seismic sources (Pratt 1999, Groos *et al* 2014). Formulated as a least-squares problem, the source inversion performs a deconvolution of the measured data to eliminate the influence of the effective system response, using its more stable formulation known as a water-level deconvolution (Langston 1979). In our framework, the optimized correction filter represents an effective source signature estimated jointly over all receivers for a given source and frequency band with receiver effects being partially absorbed into the effective correction filter through the global least-squares fit. We hereby profit from the typically available reference water shot, where the recorded signals are approximated by the convolution of this effective source-time function with the homogeneous Green's function and the noise spectrum. Our approach to source wavelet estimation is similar to the source wavelet optimization via Wiener filtering as described in Robins *et al* (2023).

To formulate the auxiliary inverse problem, we let $\hat{\mathbf{d}}^k_{\text{obs}}$, $k = 1, \ldots, n_r$, denote the discrete Fourier transforms of measured time-domain signals at the k-th receiver. Starting from an initial estimate of the source wavelet $\mathbf{f}$, a set of simulations for k receivers are generated and transformed to their frequency domain equivalent $\hat{\mathbf{d}}^k_{\text{sim}}$. We seek a correction filter $\mathbf{c}$ that, when applied to the initial wavelet, results in a corrected source wavelet $\hat{\mathbf{f}}_{\text{opt}} = (c_i \hat{f}_i)^l_{i=1}$ for each frequency l. The simulated data generated by the corrected source $\hat{\mathbf{f}}_{\text{opt}}$ has then an improved fit to the measured data $\hat{\mathbf{d}}^k_{\text{obs}}$ in a least-squares sense. This in turn implies that the vector of optimized coefficients $\mathbf{c}$ is obtained as the solution of the regularized least-squares problem

$$\min_{\mathbf{c}} \sum_{k=1}^{n_r} w_k^2 \|\hat{\mathbf{d}}^k_{\text{obs}} - \mathbf{c}\hat{\mathbf{d}}^k_{\text{sim}}\|^2 + \alpha\, n_r \|\mathbf{c}\|^2, \tag{7}$$

with $\alpha > 0$ being a regularization parameter. Prior knowledge on individual transducer characteristics, such as the opening angle, the distance between the emitters, as well as prior knowledge on a potentially transducer-dependent signal-to-noise ratio can be adapted by changing the weights $\mathbf{w} = (w_k)$. Let W be the average power of the coefficients $\hat{\mathbf{d}}^k_{\text{sim}}$ defined as the weighted, normalized L_2 norm of the simulated signals

$$W = \frac{1}{n_r} \sum_{k=1}^{n_r} w_k^2 \|\hat{\mathbf{d}}^k_{\text{sim}}\|^2. \tag{8}$$

We then obtain the coefficients solving (7) at each discrete frequency ω_l by using a thresholding parameter $\varepsilon > 0$ and defining $\alpha = \varepsilon^2 W$, resulting in

$$c_l = \left(n_r W \varepsilon^2 + \sum_{k=1}^{n_r} w_k^2 |\hat{d}^{k,l}_{\text{sim}}|^2 \right)^{-1} \left(\sum_{k=1}^{n_r} w_k^2 \left(\hat{d}^{k,l}_{\text{sim}} \right)^* \hat{d}^{k,l}_{\text{obs}} \right). \tag{9}$$

2.6. Resolution analysis for FWI

Ideally, any tomographic image should be complemented by a statement about its resolution. In contrast to linear ray-based techniques, there exists no closed-form solution that allows us to compare prior and posterior information contents for nonlinear waveform inversion (Ulrich *et al* 2023). We can, however, exploit curvature information of the misfit functional to learn more about the characteristics of the minimum. We hereby build on the second-order approximation of the misfit within L-BFGS, where the Hessian $\mathbf{H}$ describes the change of the misfit when the optimal model $\tilde{\boldsymbol{m}}$ is slightly perturbed to $\tilde{\boldsymbol{m}} + \delta\mathbf{m}$.

In the image domain, the information contained in the Hessian therefore encodes inter-parameter blurring and position- and direction-dependent resolution lengths. For realistic-sized problems, the computation of the Hessian matrix is, however, often not possible. Instead, the change in sensitivity of the misfit with respect to a particular perturbation can be analyzed by approximating Hessian-vector products with gradient information in the vicinity of the optimal model

$$\mathbf{H}(\tilde{\boldsymbol{m}})\,\delta\mathbf{m} \approx \nabla\chi(\tilde{\boldsymbol{m}} + \delta\mathbf{m}) - \nabla\chi(\tilde{\boldsymbol{m}}), \tag{10}$$

where the gradient itself is calculated efficiently with first-order adjoints (Fichtner and Trampert 2011, Fichtner and Leeuwen 2015). The perturbations $\delta\mathbf{m}$ are artificially inserted in the inverted model and may, for instance, be a collection of n point perturbations, probing different regions of the reconstruction. In this case (10) provides a proxy for the PSF, describing how the individual point perturbations are smeared in space. If $\tilde{\boldsymbol{m}}$ is the optimal model, then the second term on the right-hand side of (10) vanishes due to the optimality condition. This implies that given the reconstructed model lies in the vicinity of the optimum, the Hessian-vector product can be approximated by computing the gradient of the misfit functional evaluated at the perturbed model $\nabla\chi(\tilde{\boldsymbol{m}}+\delta\mathbf{m})$. If the reconstructed model was equal to the optimal model, all features would be perfectly resolved with a resolution limit constrained by the wavelength. In this case, the application of the Hessian would have no effect on the model perturbations and the gradient would recover the original size of the perturbation within the error bounds.

In a time-domain inversion framework, the calculation of gradients has computational costs equivalent to an entire FWI iteration. To make the proposed resolution analysis tractable for realistic problems, we exploit the linearity of the wave equation with respect to the sources by a source-stacking strategy, referred to as source-encoding, based on time shifts (Capdeville *et al* 2005, Krebs *et al* 2009, Matharu and Sacchi 2018), where individual source contributions are superimposed to simulate multiple sources within a single forward calculation. These encoded sources are motivated by the fact that for explicit time-domain FWI, the computational cost is proportional to the number of sources. To construct one stacked source the estimated source-time functions belonging to a particular frequency band are individually shifted in time and stacked into one time series. Analogously, the stacked data are generated by superimposing the measured traces after applying the same time shifts used for the corresponding source-time functions. Assigning a different time shift to each source contribution reduces interference between overlapping wavefields, commonly referred to as cross-talk. The latter is further suppressed by averaging gradients computed from multiple independent realizations of random time shifts. The resulting wavefield therefore represents data that would have been acquired from a spatially distributed set of sources activated at different onset times and, if applicable, with slightly different estimated source-time functions. Generally, source stacking finds application during the inversion workflow to save computation time during gradient calculation. In practice, it is often difficult to overcome the cross-talk, which means that more sophisticated source stacking methods need to be used to avoid mapping artifacts (Tromp and Bachmann 2019). As we are only interested in specific parts of the gradient to derive proxies for the resolution length, we consider the simple source stacking method via time-shifting sufficient.

3. Results

3.1. The *in-vivo* data set

The data were obtained using an optoacoustic ultrasound imaging platform (Lafci *et al* 2022) designed to capture cross-sectional images through the abdomen of a mouse. The aperture consists of 512 ultrasonic transducers arranged on two half circles with a radius of 40 mm enclosing a water bath. A schematic sketch of the setup is given in figure 2(a). Our goal is to optimally combine transmission and reflection information to obtain a high-resolution model of the sound speed that approaches the theoretical resolution limit of half a minimum wavelength (Virieux and Operto 2009). Consequently, applying a bandpass filter in the range of 750 kHz–2.5 MHz to the data would allow the resolution of structural features on the order of 0.3 mm, which is comparable to the resolution of MRI devices (Ranger *et al* 2012). In section 3.5, we will show that the effective resolution length for realistic, noisy data is closer to one wavelength. This is, however, still in the submillimeter range and therefore accurate enough to reconstruct prominent features within the mouse such as the highly reflective vertebrae, measuring approximately 2 mm in diameter as seen in the reflectivity image of the cross-section in figure 2(b). This image has been obtained with a delay-and-sum algorithm by considering, for each source, 32 receivers to the left and right, which predominantly record backscattered signals.

3.2. Pre-processing and data selection

Figure 2(c) shows a full view of the bandpassed filtered data (750 kHz–2.5 MHz) with an annotation of important wavefield components that determine the data selection entering the inversion. The reported frequency ranges denote the lower and upper corner frequencies of a zero-phase Butterworth bandpass filter. First, between 0 and 7 μs, 30–40 μs and after 80 μs, distinct bands appear that originate from deterministic changes in the DAQ system's impedance, affecting all channels simultaneously. Further, we note linear wave fronts that coincide with the first arrival for approximately the first and last fifth of the receivers. We expect these events to be related to interface waves propagating in the transducer ring that

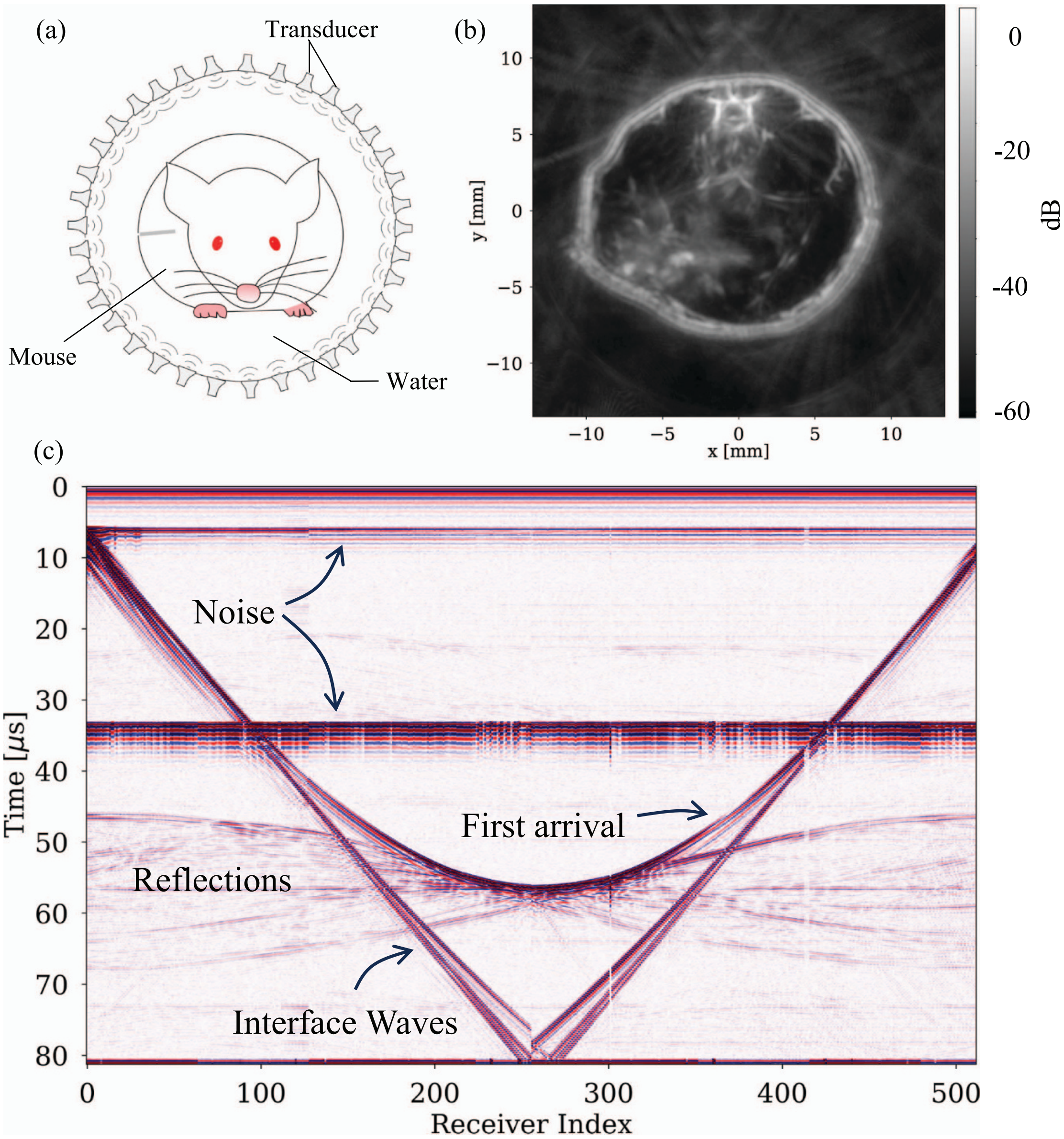

Figure 2. (a) A schematic sketch (not to scale) of the imaging system. The mouse is anesthetized and immersed into a water bath during data acquisition. (b) Reflectivity image of the cross-sectional slice acquired with the delay-and-sum method. (c) Entire view of the bandpassed filtered data (750 kHz–2.5 MHz) for one source, with traces stacked along the x-axis.

houses the individual transducer elements. To avoid mapping these effects into the reconstructed model a top mute is applied for all receivers up to 40 μs. The selected muting windows were chosen conservatively to suppress only clearly identifiable acquisition-related artifacts while preserving the physical wavefield used in the inversion. Although this reduces the length of the signals used during the inversion, all the information constraining the model parameters is contained in the waves recorded after 40 μs.

3.3. Inversion setup and results

Once the data are windowed, we start the FWI workflow, which consists of: (1) an initial model derived from a straight-ray inversion, intentionally kept smooth and low in spatial detail to provide a sufficiently accurate starting model for FWI, (2) an auxiliary inversion to recover individual source-time functions, (3) a multiscale scheme, where the frequency content is successively increased in three steps starting with a frequency band at 750 kHz–1.5 MHz and ending with 750 kHz–2.5 MHz, and (4) the implementation of the GOT to define the misfit measure. To reduce the computational load inherent to time domain approaches for applications with a large number of sources, batches of 64 sources are used per iteration and exchanged throughout the inversion such that full coverage of the domain is ensured. Since the aperture is circular, we choose the batches systematically to consist of every 8th transducer and shift this selection by one transducer per iteration. To calculate one model update and hence one iteration at the

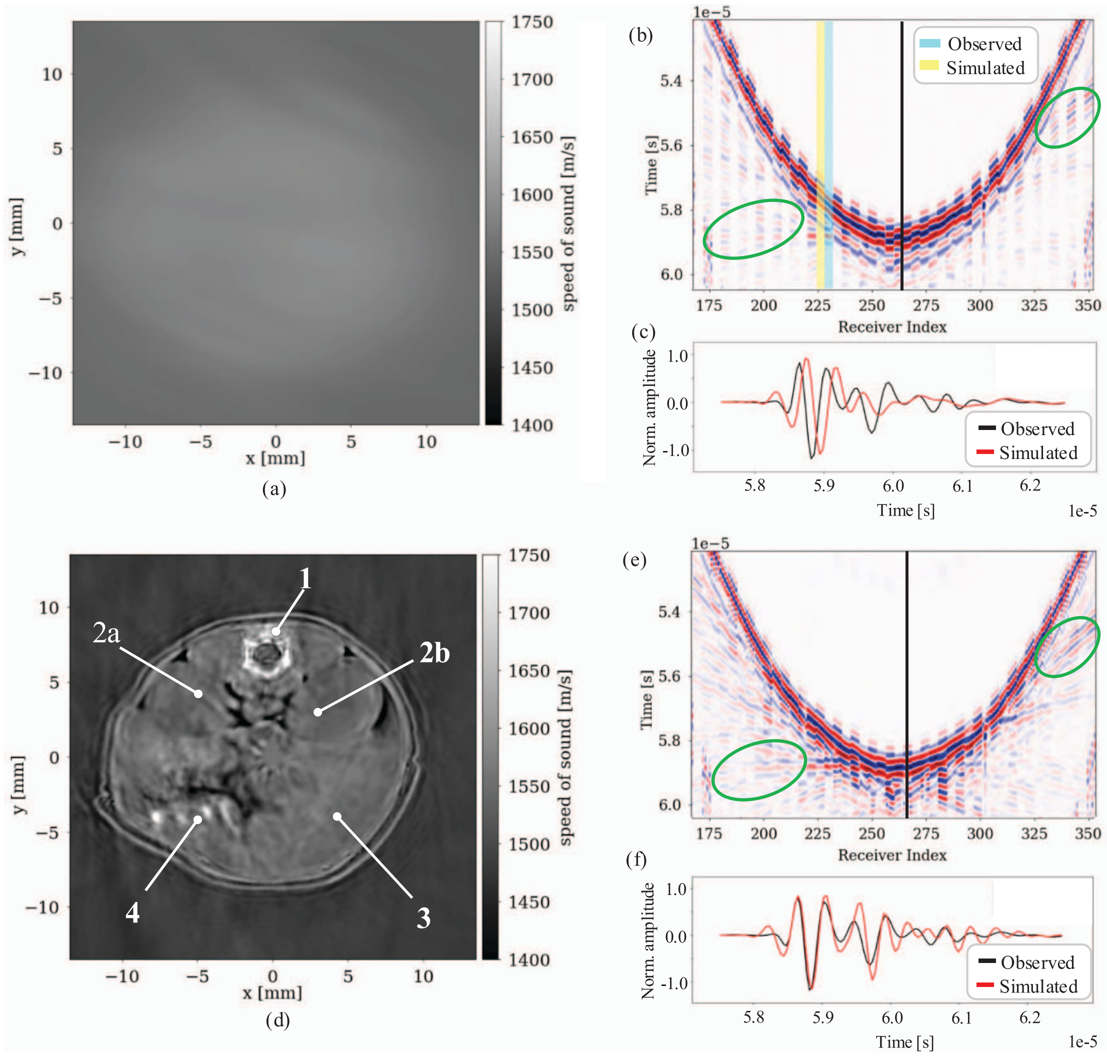

Figure 3. (a) Initial model derived from a straight-ray reconstruction. (b) Interwoven A-scan gather of transmission data showing 10 traces from the measured data at the frequency band 750 kHz–2.5 MHz and 10 traces from the initial model alternatingly. (c) Waveform fit for the receiver marked by the black line in (b). (d) Final model for the frequency band 750 kHz–2.5 MHz. Reconstructed features are (1) the vertebral column, (2a) right kidney, (2b) left kidney, (3) spleen and (4) liver. (e) Same as for (b), but comparing the measured data at the final frequency band 750 kHz–2.5 MHz and the data from the final model displayed in (d). (f) same as for (c) but for the final fit. Green circles in (b) and (e) mark areas where the filled in reflections are well visible.

last frequency band (including frequencies up to 2.5 MHz) using 64 sources and the aforementioned data selection takes about 30 min on an HPC cluster using 72 GPUs (Nvidia Tesla P100). Across all frequency bands, a total of 330 iterations were performed.

Figure 3(a) displays the starting model and (d) shows the final inverted sound speed model with labeled reconstructed features such as the vertebral column (1), the kidneys (2), the spleen (3) and the liver (4). The quality of the reconstruction becomes also apparent upon comparison of figure 3(b) with figure 3(e). These figures show interleaved A-scan gathers, where four traces of the measured data filtered at the last frequency band 750 kHz–2.5 MHz are plotted in alternance with four traces of simulated data though the initial model in (a) or the model at the final iteration in (d). Not only have the wavefronts been adjusted correctly in most cases, but also the reflections are properly included, which is highlighted by the green circles. Both of these features are also visualized in the waveform fit in figures 3(c) and (f) respectively, which displays the trace marked by the black line in the A-scan gathers.

We have previously stressed the importance of calibrating the source wavelet to improve the waveform fit. This increase in accuracy directly translates into improved spatial resolution of the reconstructed models, as illustrated in figure 4. Comparing the model obtained by using an estimated source wavelet, derived by averaging the first-arrival signals recorded in the reference water shot, in figure 4(a), to the model that utilizes the calibrated source-time functions in figure 4(b), one notes an increase in image sharpness as well as a decrease of spatial artifacts, for instance the vertical line to the left in figure 4(a).

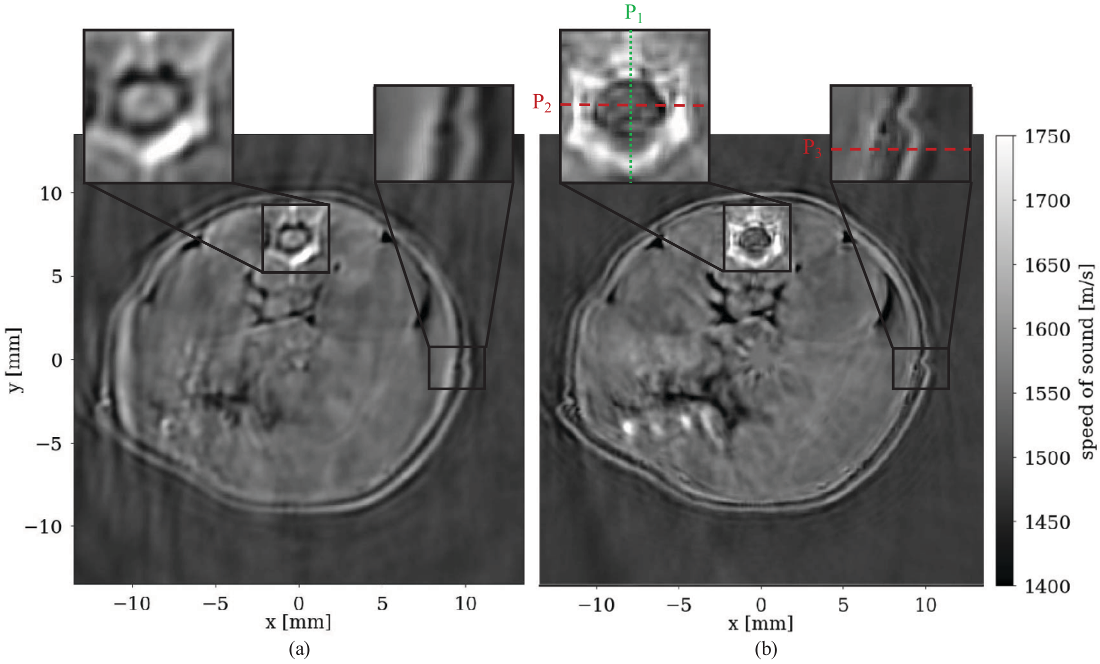

Figure 4. (a) The inverted sound speed model with an approximated source-time function and (b) the inverted model obtained when using the calibrated source-time function from section 2.5. Profiles P_1, P_2 and P_3 are shown in figure 6.

3.4. Multistage inversion for reflection and transmission tomography

Images that capture accurate structural information are in practice as valuable as quantitative images, which motivates the implementation of a multistage inversion that reconstructs not only sound speed images, but also reflection images with RTM. When a homogeneous background medium is employed, Korta Martiartu *et al* (2019) have shown that the forward operators in RTM and delay-and-sum are equivalent, assuming the same spatial and temporal discretizations have been used. Figure 5(a) shows the impedance kernel obtained with RTM for a homogeneous background model. Although the visualization differs from the reflectivity image reconstructed with standard delay-and-sum in figure 2(b), both images can be compared with respect to the spatial localization and sharpness of the dominant structural contrasts. We note here that RTM relies on backpropagating the measured wavefield in the domain and correlating it with the forward propagated wavefield to image impedance contrasts, ideally for every source–receiver combination. This differs from delay-and-sum beamforming, which reconstructs reflectors by applying kinematic delays to the recorded signals. In both approaches, direct-wave energy may interfere with the imaging of reflections if not sufficiently separated in time, and is therefore muted prior to imaging. In order to comply with our previous FWI setup, we use the same batch sampling approach as described in section 3.3 also for the reconstruction of the RTM images.

The benefit of a more accurate physical wave propagation model reveals itself when feeding a heterogeneous, more accurate sound speed distribution as the background model to the RTM. This warrants that the backpropagated wavefield only interferes constructively with the forward wavefield at those locations that mark actual impedance contrasts within the medium. High-frequency variations in the sound speed model can lead to misalignment in interference patterns, resulting in inaccurate imaging of reflectors (Davydenko and Verschuur 2021). To ensure that constructive interference only occurs at actual scatterer locations, two preprocessing steps are often applied: (1) muting the direct wave, which does not contain any reflection information and (2) ensuring that the background model is accurate, but not too detailed. The latter suggests either the use of a low-frequency reconstruction, or a smoothed version of the reconstructed model as a prior for RTM. Although it might be surprising that a smooth prior can lead to increased reflector contrast in the reflection image, it is a consequence of RTM being based on backpropagation. A detailed model will generate many events that need to be mapped by the backpropagation operator. Obtaining such an exact operator is difficult if not impossible and any additional events that cannot be mapped correctly will appear as artifacts such as multiples. A smooth model without sharp boundaries, however, does not generate a large number of reflection events besides the first reflection arrival. Therefore, the corresponding backpropagation operator for a smooth model will

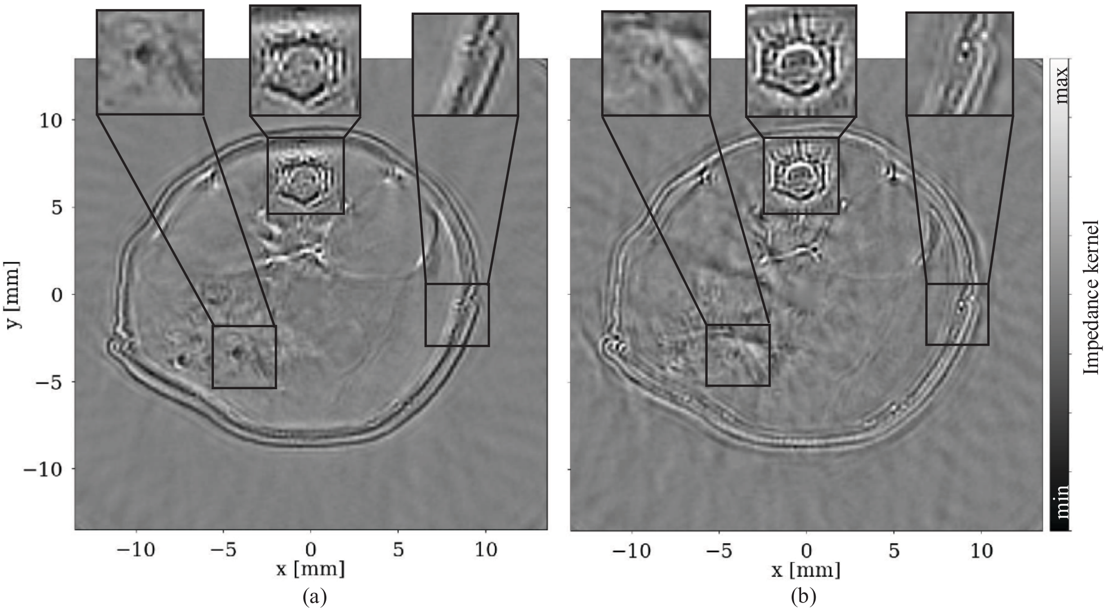

Figure 5. (a) The impedance kernel, a reflection image, obtained with RTM using a homogeneous background model; (b) the impedance kernel obtained with the reconstructed sound speed distribution shown in figure 3(d) as the background model.

primarily contain the initial arrival belonging to the first reflected wave, which is what we mostly rely on in reflection imaging.

Following this argumentation, we applied a diffusion-based smoothing operator to the final model in figure 3(d), corresponding to a characteristic smoothing length of half the wavelength at 2.5 MHz in soft tissue (approximately 0.6 mm). Here, the smoothing length refers to the spatial scale over which variations in the sound speed model are attenuated. Feeding this smoothed reconstruction as a background model to the RTM results in figure 5(b). Using the FWI-derived velocity model produces observable changes in the RTM image compared to the homogeneous background, for instance in the contrast distribution and the appearance of internal structures such as the vertebrae and liver region. These differences illustrate the effect of the background model on the imaging condition, without implying that one image is definitively sharper than the other.

3.5. Resolution analysis

To assess the spatial resolution of the sound speed reconstruction, we propose a twofold approach: first, we analyze reconstructed features directly in pixel space. Second, we investigate the resolution in model space by characterizing the spatial blurring of point perturbations. We start with the profile plots P_1 and P_2 across the vertebrae in figure 6, which demonstrate that the osseous vertebral column with sound speed values around 1940 $\mathrm{m\,s^{-1}}$ is clearly separated from the lower sound speed values of the surrounding soft tissue around 1540 $\mathrm{m\,s^{-1}}$ as well as from the vertebral canal with sound speed values around 1520 $\mathrm{m\,s^{-1}}$. To measure the spatial resolution of an isolated feature, we focus on a zoomed-in region of the skin layer and determine the full width at half maximum (FWHM) value of profile P_3 to be 0.28 mm, which falls within the range of values reported for epidermal thickness in mice (Sabino *et al* 2016).

The second approach to estimate resolution lengths is model based and follows the approach presented in section 2.6. This model-based resolution analysis is intended to inform and potentially guide aspects of the decision-making process. For instance, the appearance of suspicious masses in the reconstruction may lead to questions regarding the resolution of the associated reconstructed features. It may then be desirable to assess the local resolution of these features efficiently and without substantial computational cost. To probe specific features within the reconstructions, we start by inserting a set of 30 % low-velocity perturbations onto the reconstructed model, as shown in figure 7(a). These perturbations have a diameter of 0.25 mm, which is on the order of half a wavelength at 2.5 MHz. Using this perturbed model, Hessian-vector products for five encoded sources instead of the entire 512 original sources are calculated and summed, which results in the image shown in figure 7(b). Following (10), the Hessian-vector product indicates the direction of a single-iteration model update and is considered as a conservative estimate of the PSF (Fichtner and Leeuwen 2015). Given that the reconstructed model

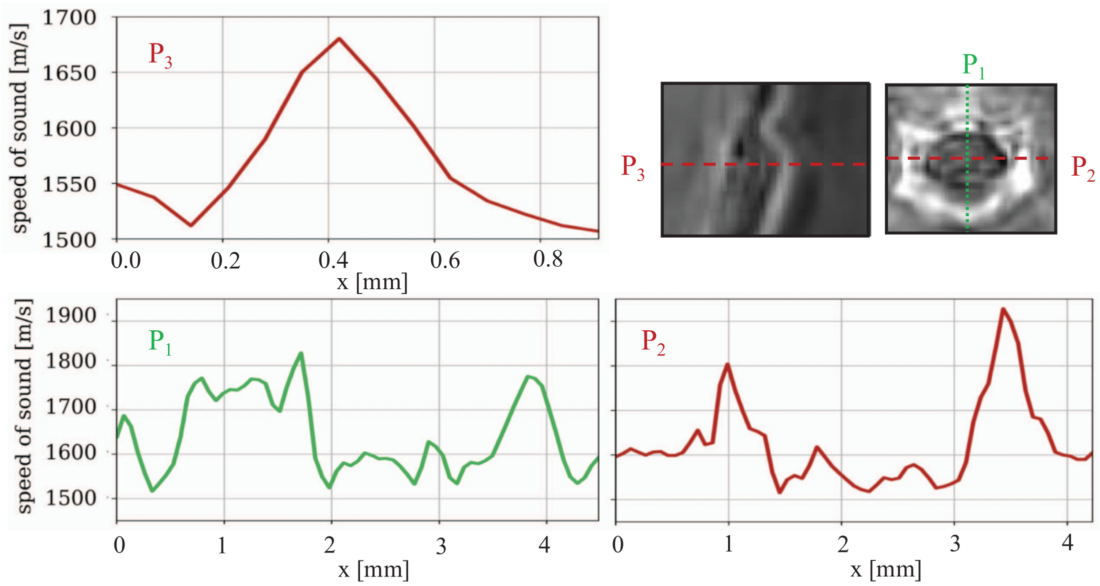

Figure 6. Profile plots across the zoomed-in region on the vertebral column and profile plot across the skin layer from figure 4.

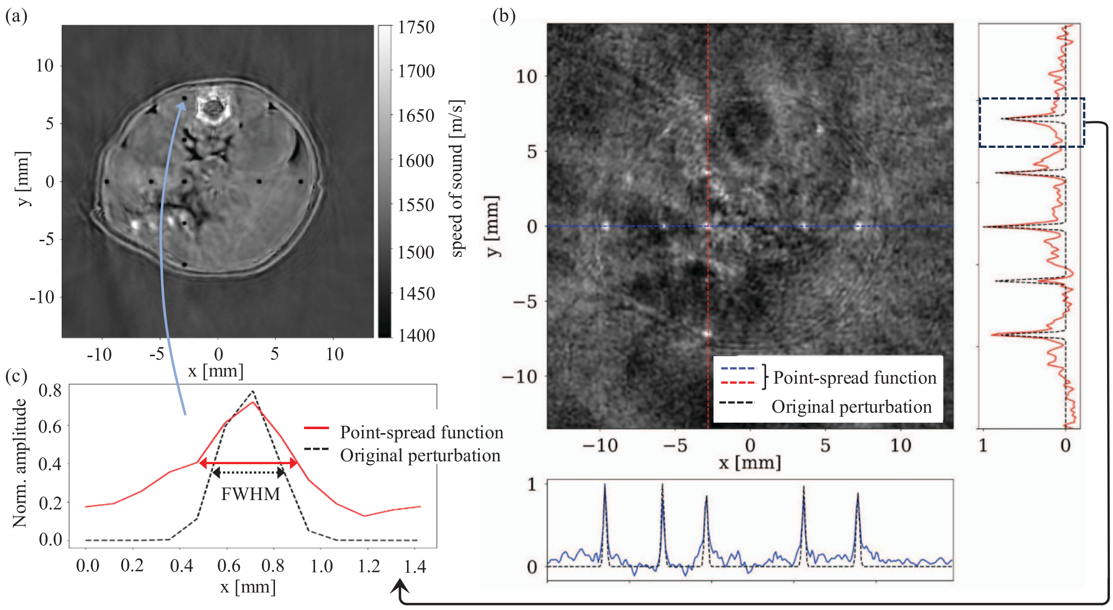

Figure 7. (a) Point perturbations with a diameter of half a wavelength and a strength of 30 % from the surrounding values inserted in the FWI reconstructed model; (b) the gradient of the misfit with respect to the model parameters calculated from 5 encoded sources. The perturbations from (a) are clearly discernible. The profile plots show that the magnitude of the imaged perturbation is space dependent, which indicates that not all areas are resolved equally well; (c) a zoom on a PSF with the FWHM measure indicated.

lies in the vicinity of the optimal model, (10) can be approximated by one gradient computation using the same codes as for the initial FWI. In the context of source encoding, this implies that one needs to compute as many gradients as encoded sources per perturbed model. These gradients are then summed to reduce cross-talk noise introduced by the encoded sources. Since we use five encoded sources, we sum five encoded gradients to obtain the image in figure 7(b). The imprint of the model in the gradient in figure 7(b) is clearly discernible, and the inserted low velocity perturbations appear expectedly as bright spots, indicating that the sound speed at these points needs to be increased to fit the measurements. Taking the profiles along the vertical (red line) and horizontal (blue line) perturbations and comparing them to their original size (black line) provides a proxy for the spread of the individual peak, which may, for instance, be calculated using the FWHM as indicated in figure 7(c). In this case, the FWHM of the original perturbation is 0.29 mm, while the FWHM of the PSF is 0.56 mm. Averaging the

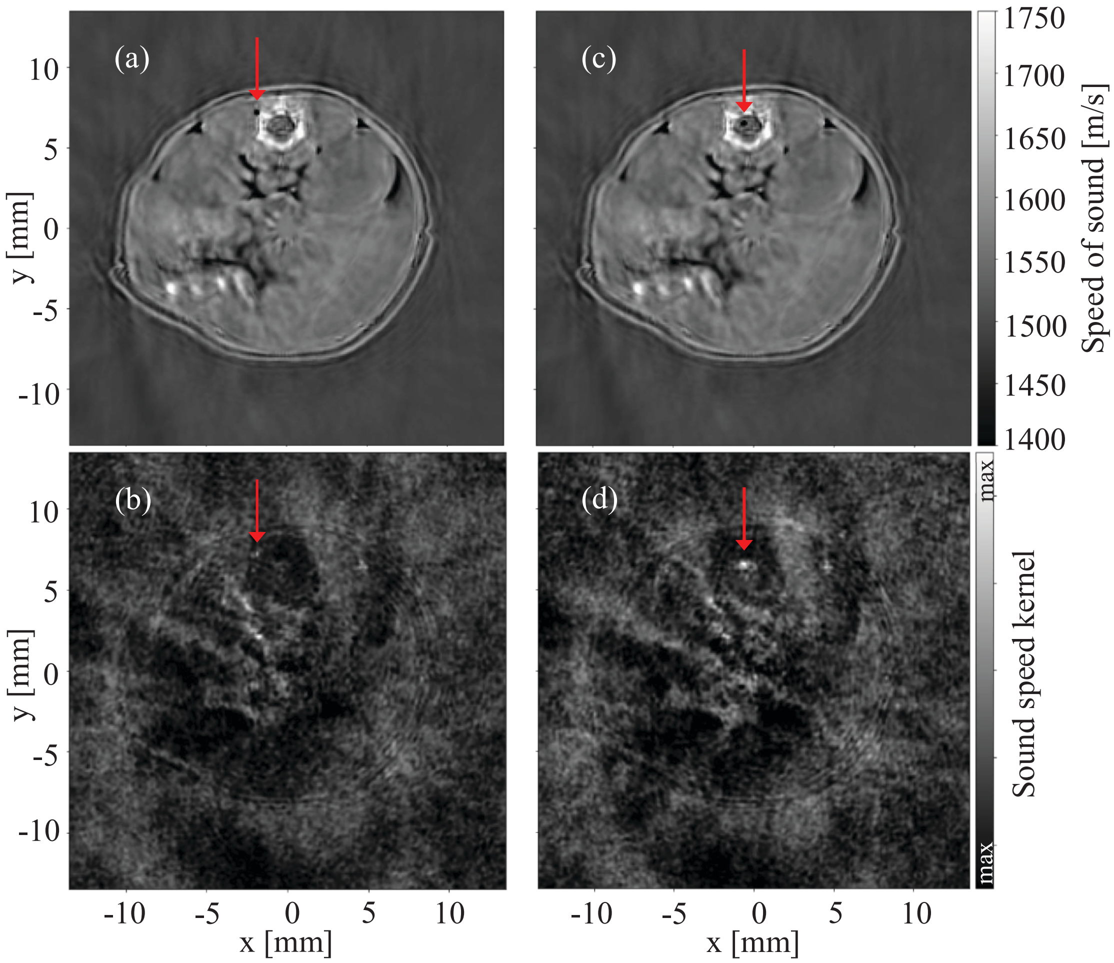

Figure 8. Point perturbations of various size, inserted in regions of the model with a large sound speed difference between the perturbation and the background medium. (a) Low sound speed perturbation with a diameter of half a wavelength, inserted at the edge of the vertebrae, a high sound speed scatterer. (b) In the corresponding gradient, the perturbation is not discernible. (c) Low sound speed perturbation with a diameter of half a wavelength inside the vertebral canal. (d) Due to the reduced sound speed difference between the background and the perturbation, the latter is clearly visible in the gradient.

FWHM values for all five vertical perturbations suggests that the size of the perturbation has increased by approximately 30 % in the PSFs compared to the original perturbation. Linking this to the theoretical, idealized noise-free resolution limit of FWI of half a wavelength, we find that for our real and thus noisy data, the spatial resolution is closer to one wavelength. Additionally, comparing amplitudes provides further information on the local resolution. The spread and the amplitude of the point perturbations within the perturbed gradient varies for each location and is especially affected in the vicinity of scatterers, such as the top distortion near the point inserted inside the high velocity structure in the liver.

Owing to the time shift introduced to construct the encoded sources, the encoded simulations are approximately 10 times longer than an individual simulation, however, since we only use 5 encoded sources the computation time for this resolution analysis with encoded sources is still reduced by a factor 10, compared to using the original sources.

To evaluate the effect of a large sound speed difference between the background and the perturbation, we test two scenarios in figure 8, showing the perturbed model in the upper row and the corresponding gradient below. The most prominent scatterer of the model is the vertebral column, a bone structure with reconstructed sound speed values ranging from 1510 m s^{-1} to 1810 m s^{-1} (P1) and 1510 m s^{-1} to 1940 m s^{-1} (P2), with mean values around 1770 m s^{-1} and 1800 m s^{-1}, respectively, and standard deviations of approximately 80–100 m s^{-1}. The surrounding soft tissue has a mean sound speed of 1540 m s^{-1} (std. dev. 15 m s^{-1}), while the vertebral canal shows a mean of 1520 m s^{-1} (std. dev. 10–15 m s^{-1}). Here, waves are strongly scattered by the nearby vertebral column, which acts as a dominant local scatterer. Consequently, the small 0.25 mm perturbation (as seen in figure 8(a)) produces only a small additional effect in the gradient near this structure, as shown in figure 8(b). In contrast, placing the same 0.25 mm perturbation inside the vertebral canal, where the local scattering is weaker, leads to a more clearly visible imprint in the gradient, as shown in figure 8(d). This distinction is entirely due to the relative position of the perturbation with respect to strong scatterers, rather than changes in the velocity contrast of the vertebra itself.

Approximating resolution lengths via gradient perturbation using source stacking provides a conservative estimate of resolution lengths and may serve as a quick, minimum quality check and can be readily automated. For instance, when detecting a suspicious mass, a clinician may test the resolution at this location by placing a PSF. Knowing the sound speed of tumors, one can then calculate how large the inclusion must be to appear truthfully in the image.

4. Conclusions

In this study, we have demonstrated that FWI can be successfully applied to *in-vivo* ultrasound data, treating long-standing challenges related to nonlinearity and computational cost. By combining three essential elements—a GOT misfit to mitigate local minima, source-time function calibration to accurately reproduce the ultrasonic wavefield, and an efficient resolution analysis based on PSFs, we obtained quantitative sound speed maps that resolve soft tissue and bone structures at sub-millimeter scales. These results show that FWI, even within an acoustic approximation, yields sound speed values consistent with literature (Duck 1990) and reveals anatomical detail that conventional ultrasound imaging struggles to capture.

The workflow is also readily refined to probe smaller features by (1) iterating on the perturbed models to improve the sharpness of the point perturbations, (2) increasing the number of encoded wavefields, and (3) using more sophisticated source-encoding methods to minimize cross-talk, for instance by using phase shifts (Bachmann and Tromp 2020). A limitation of the proposed resolution analysis is that the insertion of 2D point perturbations cannot represent off-plane scattering, which is present in real 3D raw data. We therefore emphasize that the presented resolution analysis provides conservative resolution proxies.

Although the inversion scheme uses the acoustic wave equation and is thus, strictly speaking, only correct for soft tissues with negligible elasticity, the reconstructed shape and sound speed values of the vertebral column on the order of 1940 $\mathrm{m\,s^{-1}}$ are significantly higher than values for soft tissue, as expected for the sound speed in bones (Hasgall *et al* 2012). This provides a potential new perspective for ultrasound computed tomography: orthopedic imaging, which has been previously investigated by other authors, e.g. Fincke *et al* (2022) and Zhou *et al* (2023). Presently, to our best knowledge, there exists no USCT system in clinical practice for orthopedic imaging due to the difficulty of handling the high impedance contrast between bone and soft tissue. This difficulty primarily stems from the fact that, in many conventional ultrasound tomography approaches, the forward problem is often approximated as linear for weak contrast structures, which reduces the influence of diffraction effects when perturbations in sound speed are small relative to the background. In the presence of bone, this linear approximation is inadequate and produces artifacts. Wiskin *et al* (2020) have shown the capability of transmission ultrasound tomography to produce high-resolution images in the presence of bone when using a wave-based imaging approach. In this case, the modeling accounts for both the nonlinearity of the problem and all wave phenomena. The reconstruction of the vertebral column in this study further highlights the potential of wave-based ultrasound tomography for orthopedic imaging, with applications for instance in sports medicine or in the context of osteoporosis treatment.

Beyond the technical advances, our study highlights how high-resolution FWI reconstructions can directly benefit medical imaging practice. Accurate sound speed distributions enable sharper reflection imaging, while resolution tests provide clinicians with a tool to assess the reliability of features of interest. Such capabilities open new perspectives for medical ultrasound, ranging from improved soft-tissue characterization to applications in orthopedic imaging, where differentiating bone and surrounding tissue remains a challenge for conventional approaches.

Taken together, our findings suggest that FWI has the potential to bridge the gap between quantitative tomography and clinically usable ultrasound imaging. With continued progress in computational efficiency and integration into portable platforms, the methodology presented here could contribute to bringing wave-equation-based ultrasound imaging closer to everyday clinical reality.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://berkanlafci.github.io/pyruct/_documentation/data.html#ultrasound-data.

Estimated vs calibrated source available at https://doi.org/10.1088/1361-6560/ae7cd6/data1.

Inverted source time functions available at https://doi.org/10.1088/1361-6560/ae7cd6/data2.

GOT vs L2 misfit available at https://doi.org/10.1088/1361-6560/ae7cd6/data3.

Raw time series available at https://doi.org/10.1088/1361-6560/ae7cd6/data4.

Appendix. FWI parameters

In the following, we list some important technical details of the time-domain multiscale FWI described in section 3.3. At the lowest frequency band, comprising frequencies between 750 kHz and 1.5 MHz, the spectral element mesh has 246 015 elements and a time step of 1.38×10^{-8} is used. The total domain length of the mesh is equal to 0.13 m. For the simulation we use three elements per wavelength and, tensor order 1 mesh and absorbing boundaries with a width of five wavelengths. Since higher frequencies require smaller spatial sampling, the number of elements in the mesh for the second frequency band increases to 651 248 elements and for the third frequency band to 1261 128 elements. In order to remove short-wavelength structure that cannot be represented by the current frequency band and to handle large orders of differences in the gradient magnitude in the vicinity of the sources and the rest of the domain, we use gradient pre-conditioning in the form of smoothing applied to both, the sound speed and density gradients. For the lowest frequency band from 750 kHz to 1.5 MHz, we start with a smoothing length of 0.0005 m equal to approximately half a wavelength assuming a sound speed of 1540 $\frac{m}{s}$, and then decrease the smoothing length to 0.0002 m for the second frequency band from 750 kHz to 2.1 MHz, and lastly to 0.00 015 m for the band 750 kHz–2.5 MHz.

Figure A1(a) shows the reconstruction at the last iteration (iteration 128) of the first frequency band, plotted on a grid with 1587^2 pixels, which results in a resolution per pixel of $\frac{0.13\,\mathrm{m}}{1587} \approx 8.1916 \times 10^{-5}\,\mathrm{m} = 0.0819\,\mathrm{mm}$. The last model of the second frequency band (iteration 121) and the last model of the third frequency band (iteration 81) are shown in figures A1(b) and (c), both plotted on the same grid as for the first frequency band. We note that for the last model of the second frequency band shows already a good level of detail, which means that we expect this model to be closer to the global minimum. This allowed us to switch from the GOT misfit to an L_2 misfit for the last frequency band, which reduces the computation time per iteration significantly as described in section 2.3.

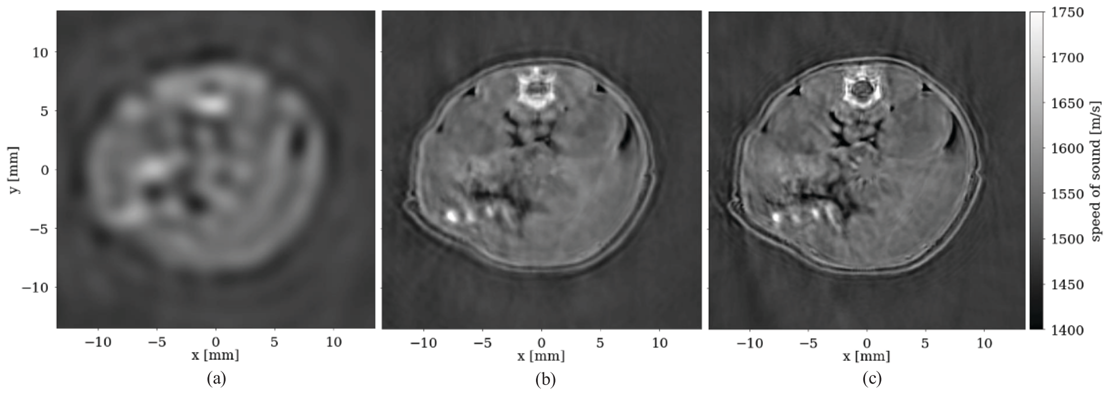

Figure A1. Reconstructions of the last model at frequency band (a) 750 kHz–1.5 MHz, (b) 750 kHz–2.1 MHz, and (c) 750 kHz–2.5 MHz.

Ethical approval

All animal experiments were conducted in accordance with the Swiss Federal Act on Animal Protection and approved by the Cantonal Veterinary Office Zurich (ZH062/2019).

ORCID iDs

I E Ulrich 0000-0003-0444-4998
S Noe 0000-0002-2622-1843
N Korta Martiartu 0000-0002-7970-7900

References

Afanasiev M, Boehm C, van Driel M, Krischer L, Rietmann M, May D A, Knepley M G and Fichtner A 2019 Modular and flexible spectral-element waveform modelling in two and three dimensions *Geophys. J. Int.* **216** 1675–92

Bachmann E and Tromp J 2020 Source encoding for viscoacoustic ultrasound computed tomography *J. Acoust. Soc. Am.* **147** 3221–35

Boehm C, Korta-Martiartu N, Vinard N, Balic I J and Fichtner A 2018 Time-domain spectral-element ultrasound waveform tomography using a stochastic quasi-Newton method *SPIE Med. Imaging 2018* vol 10580 (SPIE) pp 92–100

Boehm C, Krischer L, Ulrich I, Marty P, Afanasiev M and Fichtner A 2022 *Using Optimal Transport to Mitigate Cycle-Skipping in Ultrasound Computed Tomography* vol 12038 (SPIE) p 1203809

Capdeville Y, Gung Y and Romanowicz B 2005 Towards global earth tomography using the spectral element method: a technique based on source stacking *Geophys. J. Int.* **162** 541–54

Cerveny V 2001 *Seismic Ray Theory* (Cambridge University Press)

Chattopadhyay S and McMechan G A 2008 Imaging conditions for prestack reverse-time migration *Geophysics* **73** S81–S89

Chintada B R, Rau R and Goksel O 2022 Spectral ultrasound imaging of speed-of-sound and attenuation using an acoustic mirror *Front. Phys.* **10** 860725

Cueto C, Cudeiro J, Agudo O C, Guasch L and Tang M-X 2021 Spatial response identification for flexible and accurate ultrasound transducer calibration and its application to brain imaging *IEEE Trans. Ultrason., Ferroelectr. Freq. Control* **68** 143–53

Dai W and Schuster G T 2013 Plane-wave least-squares reverse-time migration *Geophysics* **78** S165–77

Davydenko M and Verschuur E 2021 Full wavefield least-squares reverse time migration *Geophysics* **86** WC67–WC74

Devaney A J 1984 Geophysical diffraction tomography *IEEE Trans. Geosci. Remote Sens.* **GE-22** 3–13

Douma H, Yingst D, Vasconcelos I and Tromp J 2010 On the connection between artifact filtering in reverse-time migration and adjoint tomography *Geophysics* **75** S219–23

Duck F A 1990 *Physical Properties of Tissue: A Comprehensive Reference Book* (Academic)

Fichtner A 2011 Full Seismic Waveform Modelling and Inversion

Fichtner A and Leeuwen T v 2015 Resolution analysis by random probing *J. Geophys. Res.* **120** 5549–73

Fichtner A and Trampert J 2011 Resolution analysis in full waveform inversion *Geophys. J. Int.* **187** 1604–24

Fincke J, Zhang X, Shin B, Ely G and Anthony B W 2022 Quantitative sound speed imaging of cortical bone and soft tissue: results from observational data sets *IEEE Trans. Med. Imaging* **41** 502–14

Gemmeke H, Hopp T, Zapf M, Kaiser C and Ruiter N V 2017 3d ultrasound computer tomography: hardware setup, reconstruction methods and first clinical results *Nucl. Instrum. Methods Phys. Res. A* **873** 59–65

Groos L, Schäfer M, Forbriger T and Bohlen T 2014 The role of attenuation in 2d full-waveform inversion of shallow-seismic body and Rayleigh waves *Geophysics* **79** R247–61

Guasch L, Calderon Agudo O, Tang M-X, Nachev P and Warner M 2020 Full-waveform inversion imaging of the human brain *npj Digi. Med.* **3** 28

Hasgall P, Neufeld E, Gosselin M, Klingenböck A and Kuster N 2012 It'is database for thermal and electromagnetic parameters of biological tissues

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

Iuanow E, Smith K, Obuchowski N A, Bullen J and Glock J C 2017 Accuracy of cyst versus solid diagnosis in the breast using quantitative transmission ultrasound *Acad. Radiol.* **24** 1148–53

Jensen J A, Nikolov S I, Gammelmark K L and Pedersen M H 2006 Synthetic aperture ultrasound imaging *Ultrasonics* **44** e5–e15

Korta Martiartu N, Boehm C and Fichtner A 2020 3-D wave-equation-based finite-frequency tomography for ultrasound computed tomography *IEEE Trans. Ultrason., Ferroelectr. Freq. Control* **67** 1332–43

Korta Martiartu N, Boehm C, Hapla V, Maurer H, Balic I J and Fichtner A 2019 Optimal experimental design for joint reflection-transmission ultrasound breast imaging: From ray- to wave-based methods *J. Acoust. Soc. Am.* **146** 1252–64

Krebs J R, Anderson J E, Hinkley D, Neelamani R, Lee S, Baumstein A and Lacasse M-D 2009 Fast full-wavefield seismic inversion using encoded sources *Geophysics* **74** WCC177–WCC188

Lafci B 2022 pyruct (available at: https://berkanlafci.github.io/pyruct/_documentation/data.html#ultrasound-data)

Lafci B *et al* 2023 Multimodal assessment of non-alcoholic fatty liver disease with transmission-reflection optoacoustic ultrasound *Theranostics* **13** 4217–28

Lafci B, Robin J, Dean-Ben X L and Razansky D 2022 Expediting image acquisition in reflection ultrasound computed tomography *IEEE Trans. Ultrason., Ferroelectr. Freq. Control* **69** 2837–48

Langston C A 1979 Structure under Mount Rainier, Washington, inferred from teleseismic body waves *J. Geophys. Res.* **84** 4749–62

Levin S A 1984 Principle of reverse-time migration *Geophysics* **49** 581–3

Littrup P J, Duric N, Schmidt S, Li C, Roy O, Bey-Knight L, Larsen L H and Yamashita M 2021 Multicenter study of whole breast stiffness imaging by ultrasound tomography (SoftVue) for characterization of breast tissues and masses *J. Clin. Med.* **10** 5528

Marty P, Boehm C and Fichtner A 2021 *Acoustoelastic Full-Waveform Inversion for Transcranial Ultrasound Computed Tomography* vol 11602 (SPIE) p 1160211

Matharu G and Sacchi M D 2018 Source encoding in multiparameter full waveform inversion *Geophys. J. Int.* **214** 792–810

Métivier L, Brossier R, Mérigot Q and Oudet E 2019 A graph space optimal transport distance as a generalization of lp distances: application to a seismic imaging inverse problem *Inverse Problems* **35** 085001

Mitcham T, Ali R, Schartz D, Singh M, Bender M and Duric N 2023 Feasibility of imaging ischemic stroke through the skull using ultrasound tomography *2023 IEEE Int. Ultrasonics Symp. (IUS)* pp 1–3

Mitsuhashi K, Poudel J, Matthews T P, Garcia-Uribe A, Wang L V and Anastasio M A 2017 A forward-adjoint operator pair based on the elastic wave equation for use in transcranial photoacoustic computed tomography *SIAM J. Imaging Sci.* **10** 2022–48

Pratt R G 1999 Seismic waveform inversion in the frequency domain, part 1: theory and verification in a physical scale model *Geophysics* **64** 888–901

Pratt R G, Shin C and Hick G J 1998 Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion *Geophys. J. Int.* **133** 341–62

Pratt R, Huang L, Duric N and Littrup P 2007 Sound-speed and attenuation imaging of breast tissue using waveform tomography of transmission ultrasound data *Proc. SPIE - The Int. Society for Optical Engineering* p 6510

Ranger B, Littrup P, Duric N, Chandiwala-Mody P, Li C, Schmit S and Lupinacci J 2012 Breast ultrasound tomography versus MRI for clinical display of anatomy and tumor rendering: preliminary results *Am. J. Roentgenol.* **198** 233–9

Robins T C, Cueto C, Cudeiro J, Bates O, Agudo O C, Strong G, Guasch L, Warner M and Tang M-X 2023 Dual-probe transcranial full-waveform inversion: a brain phantom feasibility study *Ultrasound Med. Biol.* **49** 2302–15

Roy O, Zuberi M A H, Pratt R G and Duric N 2016 *Ultrasound Breast Imaging Using Frequency Domain Reverse Time Migration* vol 9790 (SPIE) p 97900B

Sabino C P, Deana A M, Yoshimura T M, da Silva D F, França C M, Hamblin M R and Ribeiro M S 2016 The optical properties of mouse skin in the visible and near infrared spectral regions *J. Photochem. Photobiol.* **160** 72–78

Schatzberg A and Devaney A J 1992 Super-resolution in diffraction tomography *Inverse Problems* **8** 149

Stotzka R, Ruiter N, Mueller T, Liu R and Gemmeke H 2005 High resolution image reconstruction in ultrasound computer tomography using deconvolution *Medical Imaging 2005: Ultrasonic Imaging and Signal Processing* vol 5750 (SPIE) pp 315–25

Tarantola A 1984 Inversion of seismic reflection data in the acoustic approximation *Geophysics* **49** 1259–66

Tarantola A 1988 Theoretical background for the inversion of seismic waveforms including elasticity and attenuation *Pure Appl. Geophys.* **128** 365–99

Tromp J and Bachmann E 2019 Source encoding for adjoint tomography *Geophys. J. Int.* **218** 2019–44

Tromp J, Tape C and Liu Q 2004 Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels *Geophys. J. Int.* **160** 195–216

Ulrich I E, Boehm C, Marty P, Martiartu N K, Lafci B, Dean-Ben X L, Razansky D and Fichtner A 2024 Waveform inversion with calibrated source-time functions for improving in-vivo ultrasound computed tomography *Medical Imaging 2024: Ultrasonic Imaging and Tomography* vol 12932, ed C Boehm and N Bottenus (International Society for Optics and Photonics, SPIE)

Ulrich I E, Noe S, Boehm C, Martiartu N K, Lafci B, Dean-Ben X L, Razansky D and Fichtner A 2023 Full-waveform inversion with resolution proxies for in-vivo ultrasound computed tomography *2023 IEEE Int. Ultrasonics Symp. (IUS)* pp 1–4

Virieux J and Operto S 2009 An overview of full-waveform inversion in exploration geophysics *Geophysics* **74** 1–26

Wang K, Matthews T, Anis F, Li C, Duric N and Anastasio M A 2015 Waveform inversion with source encoding for breast sound speed reconstruction in ultrasound computed tomography *IEEE Trans. Ultrason., Ferroelectr. Freq. Control* **62** 475–93

Warner M and Guasch L 2016 Adaptive waveform inversion: theory *Geophysics* **81** R429–45

Wiskin J, Malik B, Pirshafiey N and Klock J 2020 Limited view reconstructions with transmission ultrasound tomography: clinical implications and quantitative accuracy 32

Yamashita M W, Larsen L H, Perez J, Edwards A V, Papaioannou J and Jiang Y 2024 Comparison of mammography and mammography with supplemental whole-breast us tomography for cancer detection in patients with dense breasts *Radiology* **311** e231680

Zhou C, Xu K and Ta D 2023 Frequency-domain full-waveform inversion-based musculoskeletal ultrasound computed tomography *J. Acoust. Soc. Am.* **154** 279–94