

Sobolev space norm regularized full waveform inversion for ultrasound computed tomography

Panpan Li(李盼盼), Yubing Li(李玉冰), Chang Su(苏畅), Zeyuan Dong(董则元), and Weijun Lin(林伟军)
Citation: Chin. Phys. B, 2025, 34 (**5**): 054301. DOI: 10.1088/1674-1056/adbbc0
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Multi-parameter ultrasound imaging for musculoskeletal tissues based on a physics informed generative adversarial network

Pengxin Wang(王鹏鑫), Heyu Ma(马贺雨), Tianyu Liu(刘天宇), Chengcheng Liu(刘成成), Dan Li(李旦), and Dean Ta(他得安)
Chin. Phys. B, 2025, 34 (**4**): 044301. DOI: 10.1088/1674-1056/adb390

Quantitative ultrasound brain imaging with multiscale deconvolutional waveform inversion

Yu-Bing Li(李玉冰), Jian Wang(王建), Chang Su(苏畅), Wei-Jun Lin(林伟军), Xiu-Ming Wang(王秀明), and Yi Luo(骆毅)
Chin. Phys. B, 2023, 32 (**1**): 014303. DOI: 10.1088/1674-1056/ac6dad

High-resolution bone microstructure imaging based on ultrasonic frequency-domain full-waveform inversion

Yifang Li(李义方), Qinzhen Shi(石勤振), Ying Li(李颖), Xiaojun Song(宋小军), Chengcheng Liu(刘成成), Dean Ta(他得安), and Weiqi Wang(王威琪)
Chin. Phys. B, 2021, 30 (**1**): 014302. DOI: 10.1088/1674-1056/abc7aa

Enhancement of photoacoustic tomography in the tissue with speed-of-sound variance using ultrasound computed tomography

Cheng Ren-Xiang (程任翔), Tao Chao (陶超), Liu Xiao-Jun (刘晓峻)
Chin. Phys. B, 2015, 24 (**11**): 114301. DOI: 10.1088/1674-1056/24/11/114301

Sobolev space norm regularized full waveform inversion for ultrasound computed tomography

Panpan Li(李盼盼)[1,2], Yubing Li(李玉冰)[1,2,†], Chang Su(苏畅)[1,2], Zeyuan Dong(董则元)[1,2], and Weijun Lin(林伟军)[1,2,‡]

[1] *Institute of Acoustics, Chinese Academy of Sciences, Beijing* 100190, *China*
[2] *University of the Chinese Academy of Sciences, Beijing* 100049, *China*

(Received 31 December 2024; revised manuscript received 6 February 2025; accepted manuscript online 3 March 2025)

Full waveform inversion (FWI) is a complex data fitting process based on full wavefield modeling, aiming to quantitatively reconstruct unknown model parameters from partial waveform data with high-resolution. However, this process is highly nonlinear and ill-posed, therefore achieving high-resolution imaging of complex biological tissues within a limited number of iterations remains challenging. We propose a multiscale frequency–domain full waveform inversion (FDFWI) framework for ultrasound computed tomography (USCT) imaging of biological tissues, which innovatively incorporates Sobolev space norm regularization for enhancement of prior information. Specifically, we investigate the effect of different types of hyperparameter on the imaging quality, during which the regularization weight is dynamically adapted based on the ratio of the regularization term to the data fidelity term. This strategy reduces reliance on predefined hyperparameters, ensuring robust inversion performance. The inversion results from both numerical and experimental tests (i.e., numerical breast, thigh, and *ex vivo* pork-belly tissue) demonstrate the effectiveness of our regularized FWI strategy. These findings will contribute to the application of the FWI technique in quantitative imaging based on USCT and make USCT possible to be another high-resolution imaging method after x-ray computed tomography and magnetic resonance imaging.

Keywords: full waveform inversion, Sobolev space norm regularization, ultrasound computed tomography

PACS: 43.60.Lq, 43.80.Qf, 43.35.Wa, 87.63.dh

DOI: 10.1088/1674-1056/adbbc0 **CSTR:** 32038.14.CPB.adbbc0

1. Introduction

High-resolution images of biological tissues are of great significance for clinical examination, diagnosis, and decision-making in diseases such as breast cancer,[1] osteoporosis, fracture,[2,3] Alzheimer's disease, and head trauma.[4–7] The primary techniques for tissue imaging include x-ray computed tomography (X-CT), magnetic resonance imaging (MRI), and ultrasound imaging. Among these, X-CT offers high-resolution images by scanning the body layer by layer using x-ray beams and sensitive detectors, with image reconstruction performed computationally.[1] However, it involves harmful ionizing radiation, raising concerns about the induction of cancer.[5] So, X-CT demands a careful check of risk-benefit balance in clinical practice. Alternatively, MRI generates both structural and functional images of the human body with superior soft tissue contrast, involving no ionizing radiations. However, patients often find the long and noisy examinations in a confined tubular space uncomfortable, though "open" MRI designs could alleviate this. In addition, implants or other immovable metallic objects in the body may pose risks and prevent some patients from safely undergoing MRI examinations.[8]

Ultrasound is a non-invasive, radiation-free, cost-effective, and versatile imaging technique, playing a crucial role in biological tissue imaging.[9] It is especially favored in cases requiring dynamic observation and real-time evaluation of the human body.[10,11] As a widely used device, handheld pulse-echo B-mode ultrasound typically uses linear or convex array probes to generate structural images by capturing the back-scattered reflection signals.[12] However, its spatial resolution and field of view are limited, and image quality can be influenced by the operator's skill and experience, often lacking standardization in ultrasound imaging.[1,5,13]

Recently, ultrasound computed tomography (USCT) has shown significant development potential. It utilizes a type of surrounding ultrasound transducer array (e.g., circular array) for full-azimuth acquisition to reveal the internal structures and acoustic properties of the detection target.[1,2,4,5,14] USCT collects vast amount of data, associated with various kinds of waves such as transmission, reflection, diffraction, and scattering. This wealth of information makes USCT highly promising for high-resolution imaging of complex structures, including the breast, limbs, and brain. There have been various imaging methods developed for USCT, including the ray-based methods, like delay-and-sum imaging (DAS)[15,16] and time-of-flight tomography (TOFT),[17] as well as wave-based

†Corresponding author. E-mail: liyubing@mail.ioa.ac.cn
‡Corresponding author. E-mail: linwj@mail.ioa.ac.cn

© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

methods like diffraction tomography (DT).[18,19] In practical imaging applications, ray-based methods typically rely on ray assumptions and do not fully utilize all available ultrasound data, leading to suboptimal levels of resolution.[1,20] DT methods based on Born or Rytov approximations achieve higher resolution imaging compared to ray-based approaches. However, their validity relies on the assumption of weak scattering, where the acoustic contrast is relatively low.[1,5]

To address these limitations, full waveform inversion (FWI), which leverages full wave information and optimization algorithms, has been increasingly incorporated into USCT to achieve high-resolution imaging.[2,4,5,21–24] It minimizes the difference between simulated and observed data through an iterative process, during which acoustic parameters (e.g., sound speed, density, impedance) are reconstructed, leading to a high-resolution quantitative model for the medium.[22–25] The key elements of FWI include an efficient forward simulation solver and a local optimization method, so that gradients and Hessian operators can be efficiently estimated. However, FWI is a highly nonlinear and ill-posed inverse problem, where multiple models can fit the observed data.[22,26,27] In certain conditions — such as when the initial model is far from the true solution, there is a lack of low-frequency data, or the data is noisy — FWI may fail to prevent the misfit function from converging to an irrelevant local minimum value, resulting in suboptimal solutions.

To alleviate this problem, the multiscale strategy is commonly employed in time-domain full waveform inversion (TDFWI).[28] This approach incrementally introduces shorter wavelength information into the parameter space by applying successive low-pass filters to the data. However, time-domain methods require calculations at each time step, leading to significant demands on both computational resources and time cost during the simulation step, especially when processing long time series or large-scale models. To meet the requirements of clinical applications, we focus our research on frequency–domain full waveform inversion (FDFWI).[29,30] For the frequency–domain method, wave propagation is solved for individual frequency components, making FWI perform at specific frequency points. FDFWI offers two significant advantages: (i) it effectively utilizes the information redundancy by selecting a finite set of frequency points;[31] (ii) it naturally accommodates the multiscale inversion strategy, using the inversion result from low-frequency points as the initial model for high-frequency points and progressively transitioning to higher frequency components.[2,25,26] This might mitigate the high computational overburden associated with time-domain inversion.

Although the multiscale inversion strategy effectively mitigates convergence to local minima, FWI remains challenged by the noise and underdetermined issues, especially in regions with complex and strongly contrasting physical parameters. This can be improved by incorporating appropriate regularization terms.[2,26,27,32] As a commonly used method, Tikhonov regularization promotes smoothness in the model by adding a smoothing term to the objective function, ensuring continuity of the model during the inversion process.[27,32,33] However, when aiming for high-resolution imaging, only using Tikhonov regularization may lead to undesirable effects like excessive smoothing, causing a loss of resolution and fine details. To overcome this problem, researchers have explored various other regularization methods. For example, total variation (TV) regularization is introduced as a non-smooth constraint that enhances the sparsity of image gradients, making it particularly suitable for models with significant contrast by preserving distinct features like edges.[34–36] However, it may exhibit the staircase effect when applied to regions with certain slanted angles. Additionally, the hybrid scheme combining Tikhonov and TV regularizations has also been investigated to improve inversion performance particularly in scenarios featuring smooth backgrounds and high-contrast edges.[2,37,38] While this approach achieves a balance between smoothness and edge preservation, its linear weighting scheme requires careful tuning of the relative contributions of each term. Moreover, it lacks the flexibility to adjust the balance between smoothness and edge enhancement at different regions of the model, which may limit its adaptability in complex imaging scenarios.

Sobolev space norm regularization offers a novel approach, that naturally combines the features of TV and Tikhonov regularizations. This method aims to maintain a balance between low and high wavenumbers in the model, particularly during the early stages of inversion. Originally developed for salt-body (a typical high-impedance imaging target in geophysics) inversion, it has been shown to effectively recover the shape and contrast of salt bodies, even in cases where the full velocity distribution is too complex to resolve.[32]

In this work, we propose to develop a multiscale frequency–domain FWI framework incorporating Sobolev space norm regularization for USCT imaging of biological tissues. The method adapts the regularization coefficient dynamically at each inversion stage, based on the ratio between the regularization term and the data fidelity term, reducing the reliance on the predefined hyperparameter. This adaptive strategy ensures flexibility and enhances the applicability of the method for complex imaging scenarios in the frequency domain. The rest of this paper is organized as follows. Section 2 presents the theoretical framework and specific technical details of the proposed method. Section 3 outlines the setup of the models and data. Section 4 demonstrates the imaging results of both numerical and experimental tests to showcase the effectiveness of our proposed approach. Finally, Section 5 provides an anal-

ysis and discussion of the practical advantages and limitations, as well as potential directions for future improvements.

2. Methods

In this section, we begin with a brief introduction for FWI methodology for USCT. Specifically, the convergent Born series (CBS) is utilized as the forward solver to solve the Helmholtz equation. Additionally, we adopt the Sobolev space norm regularization to improve the resolution and reliability of biological tissue imaging inversion. In the last, error evaluation metrics are introduced for a quantitatively assessment of the inversion results.

2.1. Full waveform inversion

The FWI reconstructs the distribution of medium parameters (namely the model) based on the full wavefield information. The process begins with an initial guess of the model, which is used as input to a forward solver that simulates the wavefield by solving the wave equation (i.e., the Helmholtz equation in this work). A misfit function, typically characterized by the $L2$ norm, is used to measure the discrepancy between the simulated and actual observed wavefields at receivers.[39,40] The gradient of this misfit function with respect to the model parameters is then computed, and the model is iteratively updated using local optimization algorithms (e.g., gradient decent and line search).[41,42] This iterative process continues until the misfit function is minimized, yielding an optimal parameter model that best explains the observed data.

In FDFWI, we adopt a multiscale strategy, where the inversion is performed sequentially from low to high frequencies, refining the model step by step. Each inversion iteration focuses exclusively on a single frequency point, which simplifies the problem and enhances stability. This hierarchical approach leverages low-frequency data to reconstruct large-scale features and progressively incorporates higher frequencies to resolve finer details. The following outlines the specific derivation process.

2.1.1. Misfit function

The misfit function E_{fwi} of FWI in frequency–domain reads

$$E_{\text{fwi}}(\boldsymbol{m}) = \frac{1}{2}\sum_{s}\left\|\boldsymbol{R}\boldsymbol{u}_s[\boldsymbol{m}] - \boldsymbol{d}_s^{\text{obs}}\right\|^2, \tag{1}$$

subject to

$$\left(\omega^2 m^2(\boldsymbol{x}) + \nabla^2\right) u_s(\boldsymbol{x}) = -f_s(\boldsymbol{x}), \tag{2}$$

where $\boldsymbol{d}_s^{\text{obs}} \in \mathbb{C}^{N_{\text{r}}\times 1}$ is the vector of observed data at the receiver positions $\boldsymbol{x}_{\text{r}}$. The superscript s denotes the s-th source. $\boldsymbol{u}_s \in \mathbb{C}^{N_x\times 1}$ denotes the simulated wavefield associated with the slowness model $\boldsymbol{m} \in \mathbb{R}^{N_x\times 1}$ (i.e., the reverse of speed of sound) and the angular frequency ω. Note that $\boldsymbol{u}_s$ and $\boldsymbol{m}$ are the respective vector form of u_s and m, which both depend on the spatial position $\boldsymbol{x} \in \mathbb{C}^{N_x\times 1}$. Masking matrix $\boldsymbol{R} \in \mathbb{R}^{N_{\text{r}}\times N_x}$ is used for extracting the data at the receiver locations, effectively converting the observed wavefield to the receiver positions (e.g., for a ring array in USCT). The right-hand side of Eq. (2) defines the source term $f_s(\boldsymbol{x}) = \delta(\boldsymbol{x}, \boldsymbol{x}_s) I_s$, consisting of source intensity I_s and the source position vector $\boldsymbol{x}_s \in \mathbb{C}^{N_s\times 1}$. N_s, N_{r}, and N_x represent the number of sources, receivers, and spatial points, respectively.

In this part, we have defined the optimization problem for FWI at a single frequency. To implement the multiscale strategy, the optimization results from the lower frequency are used as input for the inverse process at the higher frequency, iterating over frequencies in sequence.

2.1.2. Model parameter update

We use the following equation to update the model parameters (i.e., slowness)

$$\boldsymbol{m}_{k+1} = \boldsymbol{m}_k + \alpha_k \boldsymbol{D}_k, \tag{3}$$

where the subscript k represents the k-th iteration, and α_k and $\boldsymbol{D}_k$ denote the step size and decent direction of the iteration, respectively. In this work, we perform the line search method that satisfies the Wolf condition to determine the step size.[22] The descent direction vector $\boldsymbol{D}_k$ is expressed as

$$\boldsymbol{D}_k = -\boldsymbol{H}(\boldsymbol{m}_k)\nabla_{\boldsymbol{m}_k} E_{\text{fwi}}(\boldsymbol{m}_k), \tag{4}$$

where $\boldsymbol{H}(\boldsymbol{m}_k) \in \mathbb{R}^{N_x\times N_x}$ is a preconditioner approximating the inverse Hessian matrix, and $\nabla_{\boldsymbol{m}} E_{\text{fwi}}$ is the gradient of the misfit function E_{fwi} with respect to the model parameter $\boldsymbol{m}$. Note that the full Hessian matrix is prohibitively expensive to compute in most of the scenarios such that approximation is essential. The explicit expression and associated derivation process are detailed in the next section.

2.1.3. Local optimization algorithm

We use the adjoint-state method[43] to compute the gradient $\nabla_{\boldsymbol{m}} E_{\text{fwi}}$. Equation (2) can be simplified as follows:

$$\boldsymbol{A}\boldsymbol{u}_s = \boldsymbol{f}_s, \tag{5}$$

where $\boldsymbol{A} \in \mathbb{R}^{N_x\times N_x}$ defines the impedance matrix associated with $\omega^2 m^2(\boldsymbol{x}) + \nabla^2$. By differentiating both sides of Eq. (5) with respect to model parameter $\boldsymbol{m}$, we obtain the Fréchet derivative matrix, reading

$$\frac{\partial \boldsymbol{u}_s}{\partial \boldsymbol{m}} = -\boldsymbol{A}^{-1}\frac{\partial \boldsymbol{A}}{\partial \boldsymbol{m}}\boldsymbol{u}_s. \tag{6}$$

The partial derivative of misfit function with respect to parameter $\boldsymbol{m}$ leads to

$$\nabla_{\boldsymbol{m}} E_{\text{fwi}} = \frac{\partial E_{\text{fwi}}}{\partial \boldsymbol{m}}$$

$$= -\sum_s \mathrm{Re}\left[\boldsymbol{u}_s^{\dagger}\left(\frac{\partial \boldsymbol{A}}{\partial \boldsymbol{m}}\right)^{\dagger}\left(\boldsymbol{A}^{-1}\right)^{\dagger}\boldsymbol{R}^{\dagger} \times \left(\boldsymbol{R}\boldsymbol{u}_s - \boldsymbol{d}_s^{\mathrm{obs}}\right)\right], \quad (7)$$

where Re and † represent the real part and the transpose conjugate, respectively. The term $\boldsymbol{R}^{\dagger}\left(\boldsymbol{R}\boldsymbol{u}_s - \boldsymbol{d}_s^{\mathrm{obs}}\right)$ maps the data residual back to the observation boundaries, acting as the adjoint source $\boldsymbol{g}_s \in \mathbb{C}^{N_x\times 1}$ for backpropagation.[21,23,42] The column of $\boldsymbol{A}^{-1}$ is related to the Green's functions for unit impulse sources located at each grid of the model. Due to the spatial reciprocity of the Green's functions, $\boldsymbol{A}^{-1}$ is symmetric. Therefore, $\left(\boldsymbol{A}^{-1}\right)^{\dagger}$ can be replaced by $\boldsymbol{A}^{-1}$, and the product $\boldsymbol{A}^{-1}\boldsymbol{g}_s$ leads to the back-propagated adjoint wave-field which is denoted as $\boldsymbol{q}_s = \boldsymbol{A}^{-1}\boldsymbol{g}_s = \left(\boldsymbol{A}^{-1}\right)^{\dagger}\boldsymbol{R}^{\dagger}\left(\boldsymbol{R}\boldsymbol{u}_s - \boldsymbol{d}_s^{\mathrm{obs}}\right)$, $\boldsymbol{q}_s \in \mathbb{C}^{N_x\times 1}$. This simplification is expressed as follows:

$$\nabla_{\boldsymbol{m}} E_{\mathrm{fwi}} = \frac{\partial E_{\mathrm{fwi}}}{\partial \boldsymbol{m}} = -\sum_s \mathrm{Re}\left[\boldsymbol{u}_s^{\dagger}\left(\frac{\partial \boldsymbol{A}}{\partial \boldsymbol{m}}\right)^{\dagger}\boldsymbol{q}_s\right], \quad (8)$$

where $\partial \boldsymbol{A}/\partial \boldsymbol{m}$ reads $2\omega^2\mathrm{diag}(\boldsymbol{m})$, denoting diagonal matrix with the values of the slowness model $m(\boldsymbol{x})$ on the diagonal. Given the derived gradient, the approximating inverse Hessian $\boldsymbol{H}$ can be obtained by the L-BFGS method.

2.1.4. Forward solver

Instead of the commonly used direct approach (e.g., LU decomposition associated with finite difference discretization) for solving the frequency–domain wave equation, we employ the iterative CBS method as the forward modeling kernel. This choice offers an optimal balance among computational efficiency, accuracy, scalability, and memory usage. A key advantage of the CBS method lies in its memory efficiency. Let degree of freedom (DOF) represent the total number of grid points in the computational domain. In the LU decomposition approach, memory requirements scale with the square of DOF, even though the underlying matrix may be sparse. In contrast, the memory requirement of the CBS method is almost linearly proportional to DOF. This significant memory advantage makes the CBS method particularly well-suited for large-scale wave-field simulations.

Specifically, CBS is an improved Born series (BS) method by introducing a well-designed preconditioner. Notably, the CBS method is capable of accurately solving the Helmholtz equations for structures of arbitrary size and with high scattering potentials, ensuring convergence to precise solutions. This iterative framework offers a robust approach for complex wave-field computations. The core implementation of the iterative process is

$$u_s^{(k+1)}(\boldsymbol{x}) = u_s^{(k)}(\boldsymbol{x}) - \frac{\mathrm{i}}{\epsilon}V(\boldsymbol{x})\Big(u_s^{(k)}(\boldsymbol{x}) - \mathcal{F}^{-1} \times \left[g_0(\boldsymbol{p})\mathcal{F}\left[V(\boldsymbol{x})u_s^{(k)}(\boldsymbol{x}) + f_s(\boldsymbol{x})\right]\right]\Big), \quad (9)$$

where $V(\boldsymbol{x}) \equiv \omega^2(m^2(\boldsymbol{x}) - m_0^2) - \mathrm{i}\epsilon$ with m_0 the background slowness value, i the imaginary unit, and ϵ an infinitesimally small constant. $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the forward and inverse Fourier transforms, respectively. The Green's function is defined as $g_0(\boldsymbol{p}) = 1/(|\boldsymbol{p}|^2 - \omega^2 m_0^2 - \mathrm{i}\epsilon)$ with $\boldsymbol{p}$ the wavenumber related to $\boldsymbol{x}$. The readers are referred to Osnabrugge *et al.*[44] for more details.

2.2. Sobolev space norm regularization

Regularization is a widely used technique in solving inverse problems to control the smoothness or complexity of solutions, preventing overfitting or instability. Common methods include Tikhonov regularization, which smooths models to enhance numerical stability, and TV regularization, which promotes sharp edge information. According to Ref. [32], a generalized regularization approach, the Sobolev space norm (also refered to as W_p^1) was proposed, which blends TV and Tikhonov regularizations by introducing an index to modulate the spatial gradient of $m(\boldsymbol{x})$. In this work, we will choose W_p^1 as our regularization term E_{reg}, scaled by a hyperparameter coefficient β. Consequently, the total misfit function E is reformulated as

$$E(\boldsymbol{m}) = E_{\mathrm{fwi}}(\boldsymbol{m}) + \beta E_{\mathrm{reg}}(\boldsymbol{m}), \quad (10)$$

where E_{fwi} and E_{reg} represent the data fidelity term and the regularization term, respectively. The E_{reg} can be defined as

$$E_{\mathrm{reg}}(\boldsymbol{m}) = E_{\mathrm{reg}}(m) = \int (\nabla_{\boldsymbol{x}} m \cdot \nabla_{\boldsymbol{x}} m + \varepsilon)^{p/2}\,\mathrm{d}\boldsymbol{x}, \quad (11)$$

where ε is a small positive constant added to avoid divergence when calculating the gradient of the regularization term, and the spatial gradient of $m(\boldsymbol{x})$ is given by $\nabla_{\boldsymbol{x}} m = [\partial m/\partial x, \partial m/\partial y]^{\dagger}$ assuming $\boldsymbol{x} = [x, y]^{\dagger}$ in two dimensions. For simplicity, we express Eq. (11) in operator form rather than matrix form. The gradient of E_{reg} with respect to $\boldsymbol{m}$ is

$$\frac{\partial E_{\mathrm{reg}}}{\partial \boldsymbol{m}} = \frac{\partial E_{\mathrm{reg}}}{\partial m(\boldsymbol{x})} = -p\nabla_{\boldsymbol{x}} \cdot \frac{\nabla_{\boldsymbol{x}} m}{(\nabla_{\boldsymbol{x}} m \cdot \nabla_{\boldsymbol{x}} m + \varepsilon)^{1-p/2}}. \quad (12)$$

By substituting $p = 1$ into Eq. (11), we obtain

$$E_{\mathrm{reg}} = \int (\nabla_{\boldsymbol{x}} m \cdot \nabla_{\boldsymbol{x}} m + \varepsilon)^{1/2}\,\mathrm{d}\boldsymbol{x}, \quad (13)$$

which can be simplified as

$$E_{\mathrm{reg}} = \int |\nabla_{\boldsymbol{x}} m|\,\mathrm{d}\boldsymbol{x}. \quad (14)$$

This is precisely the definition form of TV regularization. By substituting $p = 2$ into the formula, we obtain

$$E_{\mathrm{reg}} = \int (\nabla_{\boldsymbol{x}} m \cdot \nabla_{\boldsymbol{x}} m + \varepsilon)\,\mathrm{d}\boldsymbol{x}, \quad (15)$$

which can be approximated as

$$E_{\text{reg}} = \int |\nabla_{\boldsymbol{x}} m|^2 \, \mathrm{d}\boldsymbol{x}, \tag{16}$$

transforming into Tikhonov regularization form. Figure 1 shows the dependence of W_p^1 on the local parameter gradient $\nabla_{\boldsymbol{x}} m$ and p.

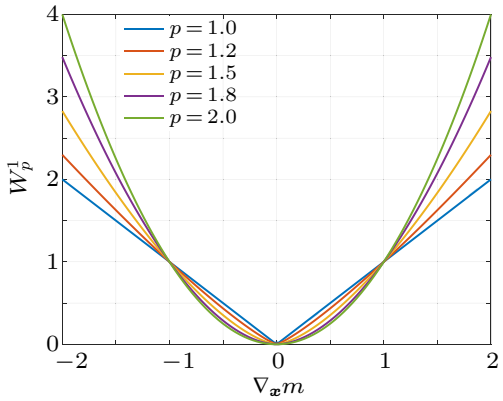


Fig. 1. The variations of the W_p^1 regularization at different values of p and local parameter $\nabla_{\boldsymbol{x}} m$, with the parameter $\varepsilon = 10^{-7}$.

Through the modification of p, W_p^1 regularization can flexibly switch between TV and Tikhonov regularization. When p is between 1 and 2, W_p^1 regularization combines the characteristics of TV and Tikhonov regularization.

2.3. Error evaluation metrics

To quantitatively assess the inversion results, we calculate three commonly used image quality metrics: peak signal-to-noise ratio (PSNR), root mean square error (RMSE), and structural similarity index (SSIM). These metrics are widely applied in image compression, restoration, and reconstruction. PSNR quantifies the difference between the ground truth and reconstructed images using the formula

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{(\max(\boldsymbol{m}_{\text{true}}))^2}{\text{MSE}} \right), \tag{17}$$

where $\max(\boldsymbol{m}_{\text{true}})$ is the maximum possible pixel value in the ground truth image or model. MSE is the mean squared error, which is calculated as

$$\text{MSE} = \frac{1}{N_x} \|\boldsymbol{m}_{\text{true}} - \boldsymbol{m}_{\text{inv}}\|_2^2, \tag{18}$$

where $\boldsymbol{m}_{\text{inv}}$ is the reconstructed image or model. A higher PSNR value indicates better quality (i.e., closer to the ground truth). Alternatively, RMSE is calculated as

$$\text{RMSE} = \sqrt{\text{MSE}}. \tag{19}$$

A smaller RMSE value indicates a smaller difference between the reconstructed image and the ground truth, reflecting higher model accuracy. SSIM is used to measure the structural similarity between two images. It considers the luminance, contrast, and structural information of the images. The definition of SSIM is given below:

$$\begin{aligned} &\text{SSIM}(\boldsymbol{m}_{\text{true}}, \boldsymbol{m}_{\text{inv}}) \\ &= \frac{(2\mu_{\text{true}}\mu_{\text{inv}} + C_1)(2\sigma_{\text{cross}} + C_2)}{(\mu_{\text{true}}^2 + \mu_{\text{inv}}^2 + C_1)(\sigma_{\text{true}}^2 + \sigma_{\text{inv}}^2 + C_2)}, \end{aligned} \tag{20}$$

where μ_{true} and μ_{inv} are the mean luminances of the images $\boldsymbol{m}_{\text{true}}$ and $\boldsymbol{m}_{\text{inv}}$, respectively. σ_{true}^2 and σ_{inv}^2 are the variances of the images $\boldsymbol{m}_{\text{true}}$ and $\boldsymbol{m}_{\text{inv}}$ respectively, indicating the contrast. σ_{cross} is the covariance between images $\boldsymbol{m}_{\text{true}}$ and $\boldsymbol{m}_{\text{inv}}$, representing the structural information. $C_1 = (\gamma_1 L)^2$ and $C_2 = (\gamma_2 L)^2$ are constants used to stabilize the division with small values, where L is the dynamic range of the image, and γ_1 and γ_2 are small constants. The range of SSIM values is in between [0,1]. When SSIM = 1, the two images are completely identical without any noticeable distortion. When SSIM = 0, these two images are completely different.[45]

3. Setup

This section details the generation of numerical observed data for breast and thigh tissues and the collection of experimental observed data for pork-belly tissue. All tests are related to the ring-array based USCT system.

3.1. Numerical test I: Breast model setup

In order to set up the numerical model for breasts, we use the clinical magnetic resonance angiography (MRA) optical and acoustic breast phantom database approved by IRB and collected by Washington University School of Medicine in St. Louis. We select a coronal plane as the breast numerical model, where each voxel value corresponds to a specific tissue type: background 0, fibroglandular tissue 2, fat 3, skin 4, and blood vessels 5. According to Ref. [46], we assign corresponding sound speed values to different tissue types in the model. Table 1 shows the sound speed distribution of the numerical breast model. To generate the synthetic data as ground truth observation, 512 array elements are uniformly distributed on a circumference with a diameter of 22 cm (as shown by the white dots in Fig. 2). The ground truth map of the real sound speed model is also displayed in the Fig. 2. Each element functions as both a source and a receiver. Data acquisition is performed using full matrix capture (FMC) mode, where each array element sequentially emits a pulse signal while all elements participate in receiving. A Gaussian pulse with a center frequency of 0.75 MHz is used as the wavelet for the sound source. Figures 3(a) and 3(b) show the waveform distribution of the sound source in both the time and frequency domains. For simplicity, the emission source in this study is

modeled as an ideal point source in space. By using a time-domain two-dimensional (2D) *k*-wave forward solver to simulate the acquisition procedure for constant-density acoustic modeling,[47] we obtain a synthetic observed dataset of size 7788×512×512 (time samplings × transmitting numbers × receiving numbers), with a sampling interval of 27.3 ns over a duration of 212 μs. The spatial domain consists of 1601×1601 grid points with a step size of 0.015 cm, covering a computational domain of 24 cm×24 cm.

Table 1. Sound speed for the ground truth of the numerical breast model.

Tissue type	Sound speed range (m/s)
Background (water)	1500
Blood vessel	1584
Skin	1650
Fat	1470
Fibroglandular	1515

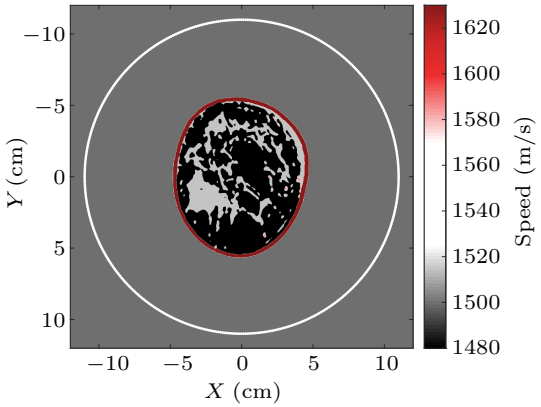


Fig. 2. The ground truth sound speed map of the numerical breast model and the distribution of 512 sound sources (indicated by the white dots).

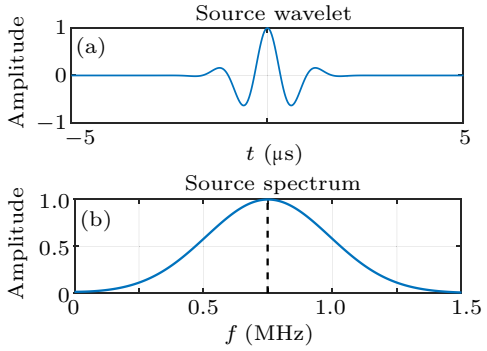


Fig. 3. The source wavelet and its spectrum used in *k*-wave simulations. (a) The time–domain source wavelet signal. (b) The spectrum of the wavelet, with a central frequency of 0.75 MHz.

Figure 4(a) illustrates the observed data received by all array elements when the 1st element acts as the transmitter. Various wave types, including reflected, transmitted, and refracted waves, can be clearly identified. Figure 4(b) illustrates the frequency–domain data for all transmitting and receiving elements at the frequency of 0.75 MHz. In order to evaluate the robustness of the proposed algorithm under noisy conditions, we add Gaussian white noise to the synthetic observed data, with a signal-to-noise ratio (SNR) of 15 dB, to mimic the practical environment. Figures 5(a) and 5(b) present the comparison of the signals emitted by the 1st element and received by the 256th element under two conditions: without noise and with noise.

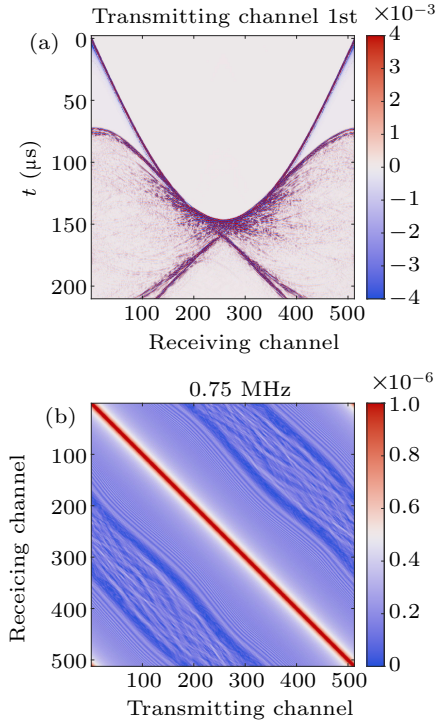


Fig. 4. Partial time–domain and frequency–domain observed data of breast model. (a) Time–domain signals received by all array elements after the transmitting of the 1st element. (b) Frequency–domain data (the amplitude map) of each transmitting and receiving array element at the 0.75-MHz frequency point.

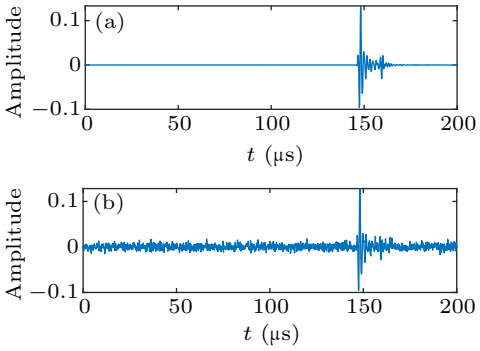


Fig. 5. The time–domain observed data of the numerical breast model: (a) Noise-free signals received by the 256th element after the transmitting of the 1st element. (b) Signals with added Gaussian white noise under the same conditions.

3.2. Numerical test II: Thigh model setup

The tissue model for thigh is derived from the CT dataset of the Visual Human Program at the National Library of Medicine in the United States. A two-dimensional (2D) cross-section is selected from this dataset with different tissue types (e.g., skin, fat, muscle, and femur). These tissues are empirically mapped to specific sound speed values based on CT images to construct a ground truth reference, as shown in Fig. 6

and Table 2. Then, for the simulation, the settings are consistent with those of the breast model, with the only difference being the center frequency of the emitting sound source. Here we set this frequency as 0.6 MHz to ensure sufficient low-frequency information and to account for the reduced penetration capability of high-frequency ultrasound in bone tissue.

Table 2. Sound speed for the ground truth of the numerical thigh model.

Tissue type	Sound speed range (m/s)
Background	1500
Skin	∼ 1600–1720
Fat	∼ 1382–1468
Muscle	∼ 1580–1596
Femur	∼ 2000–3121

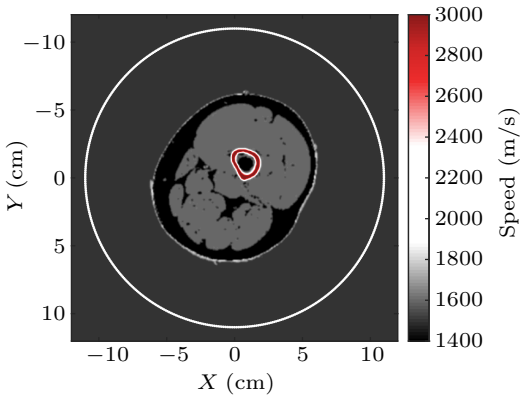


Fig. 6. The ground truth sound speed map of the numerical thigh model and the distribution of 512 sound sources (indicated by the white dots).

The time–domain and frequency–domain waveforms of the source are shown in Figs. 7(a) and 7(b), respectively. The size of the synthetic observed dataset is 16041×512×512 (time samplings×transmitting numbers×receiving numbers), with a sampling interval of 14.08 ns and a sampling duration of 225.9 μs.

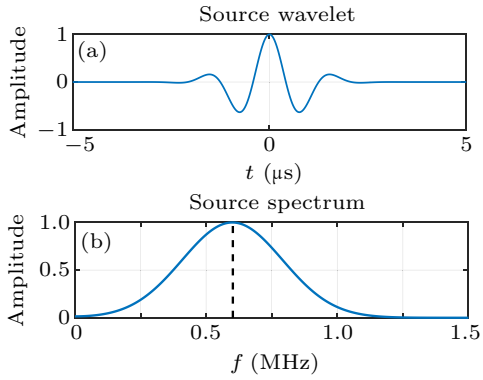


Fig. 7. The source wavelet and its spectrum used in *k*-wave simulations: (a) the time–domain source wavelet signal, (b) the spectrum of the wavelet, with a central frequency of 0.6 MHz.

We also display the observed data after the transmitting of the 1st element in Fig. 8(a), revealing more reflection data compared to the breast model, owing to the bone tissue. Meanwhile, figure 8(b) presents the frequency–domain observed data of each transmitting and receiving array element at the 0.6-MHz frequency point. For the robustness, similar to the breast model, a signal-to-noise ratio (SNR) of 15 dB is also involved.

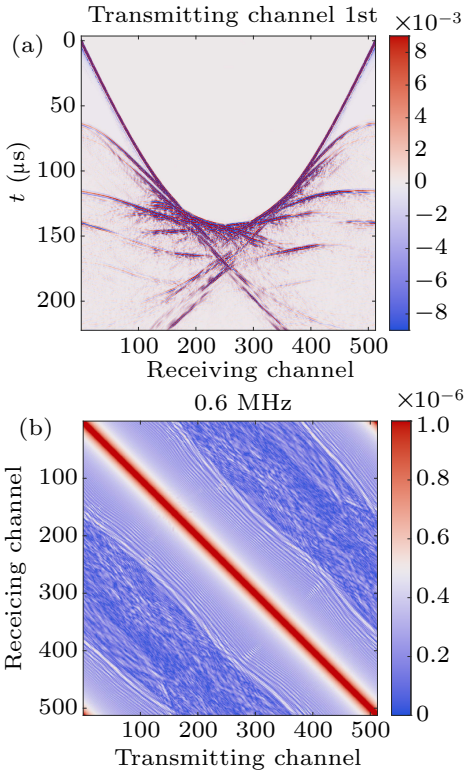


Fig. 8. Partial time–domain and frequency–domain observed data of numerical thigh model: (a) time–domain signals received by all array elements after the transmitting of the 1st element, (b) frequency–domain data (the amplitude map) of each transmitting and receiving array element at the 0.6-MHz frequency point.

Figures 9(a) and 9(b) compare the signals emitted by the 1st element and received by the 256th element under two conditions: without noise and with noise.

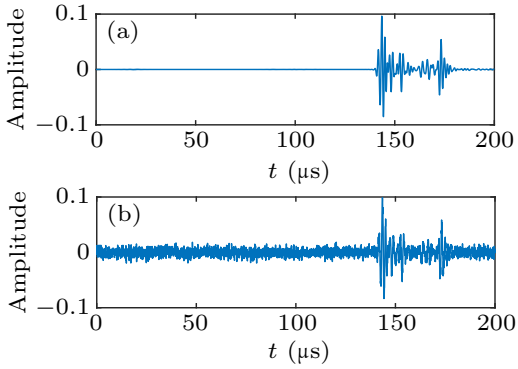


Fig. 9. The time–domain observed data of the numerical thigh model: (a) noise-free signals received by the 256th element after the transmitting of the 1st element, (b) signals with added Gaussian white noise under the same conditions.

3.3. *Ex vivo* experimental test: Pork-belly tissue setup

In this section, we give the detailed physical experiment setup. In the experimental test, we likewise acquire observed data using an ultrasonic ring array. The USCT acquisition system consists of a physical ring array of 512 piezoelectric trans-

ducers and a series of modules responsible for signal transmission, acquisition, storage and display. The entire ring array possesses an internal diameter of 22 cm. It consists of four identical arc-shaped arrays, with a spacing of one element size between them, approximately 1.2 mm. The schematic diagram of the distribution of element positions in physical ring array is shown by the red dots in Fig. 10(b). We select a segment of *ex vivo* pork-belly tissue as our imaging subject, approximately 44 centimeters in length, 7 centimeters in width, and roughly 3 centimeters in thickness.

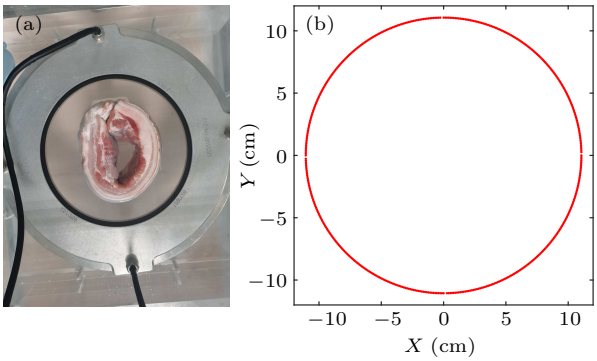


Fig. 10. The relative position between the pork-belly and the ring array, as well as the location of the array elements within the computational domain: (a) the ultrasonic ring array and *ex vivo* pork-belly tissue, (b) the distribution of 512 sound sources, represented by red dots, along with the size of the computational domain.

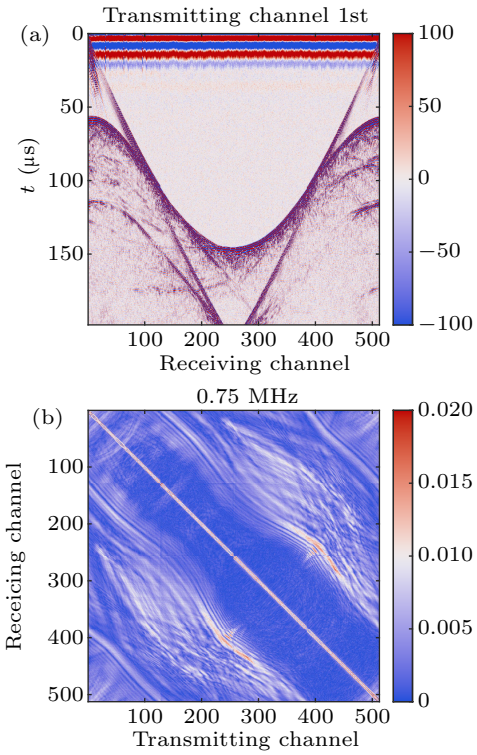


Fig. 11. Partial time–domain and frequency–domain observed data of pork-belly tissue: (a) time–domain signals received by all array elements when the 1st element is transmitting, (b) frequency–domain data of each transmitting and receiving array element at the 0.75-MHz frequency point.

During the acquisition of observed data, we still employ the FMC mode. Both the ring array and the imaging subject are fixed and submerged in water, as depicted in Fig. 10(a). We have set the transmission voltage to 75 V and selected 0.75 MHz as the central frequency of the emission pulse signal. The size of received data is 4955×512×512 (time samplings × transmitting numbers × receiving numbers), with a sampling interval of 40 ns and a sampling duration of 198 μs.

Figure 11(a) presents the observed data received by all array elements after the transmitting of the 1st element and figure 11(b) displays the frequency domain observed data of each transmitting and receiving array element at the 0.75-MHz frequency point. In Fig. 11(a), we can see obvious noise signals prior to the transmitted wave. These noises mainly come from electronic crosstalk, multiplexer switching, inherent thermal noise, quantization noise, etc. in the USCT system. To mitigate these issues, after acquiring the raw time–domain observed data, we apply a bandpass filter to remove the background noise. Subsequently, based on the time-of-arrival information and a one-sided Gaussian window, we eliminate such noises before the transmitted wave.[24]

4. Result

4.1. Numerical test I: Breast model

During the inversion process of the breast model, we set the frequency range from 0.25 MHz to 1.2 MHz, with a frequency interval of 50 kHz at lower frequencies and 100 kHz at higher frequencies. Each frequency point is iterated 5 times in the inversion process. To mitigate the near-field effects, we have excluded approximately 15% of the received data closest to each source in whole study. More specifically, in the process of using CBS for wavefield calculation, we set 601×601 grid points in the spatial domain, with a total computational domain size of 24 cm×24 cm. Figures 12(a) and 13(a) show the same inversion results without regularization ($p = 0$). The skin edge and internal tissues (fat, fibroglandular, blood vessels) of the breast model, are reconstructed, accurately capturing the macroscopic morphology with minimal artifacts. However, compared to the ground truth model in Fig. 2, the overall image appears slightly blurred with lower resolution, more pronounced background noise and artifacts near the inner side of the skin. Furthermore, as shown in the line profile comparison in Fig. 14, when $p = 0$, the reconstructed sound speed is generally underestimated, with a significant deviation in the skin's sound speed compared to the true values.

To refine these issues, we perform the inversion with W_p^1 regularization in the following. While in the regularization term, the selection of hyperparameters including the regularization coefficient β and the p-value, requires extensive testing. Moreover, the appropriate hyperparameters will change when the models are altered. Here, we will perform some trials to find the proper hyperparameters. In the

inversion process, we set a specific value χ_0 for the ratio $\chi = \beta \left\|\nabla_m E_{\text{reg}}\right\|_2 / \left\|\nabla_m E_{\text{fwi}}\right\|_2$. While χ will vary due to the updating of the model, we reset χ to χ_0 at the beginning of inversion for each frequency point by automatically adjusting the regularization coefficient β. Specifically, we set χ_0 to 10% and 15%, and the corresponding results are shown in Figs. 12 and 13, respectively. Additionally, we test various values for p (1.0, 1.2, 1.5, 1.8, and 2.0), and the corresponding inversion results are shown in panels (b)–(f) of Figs. 12 and 13, respectively.

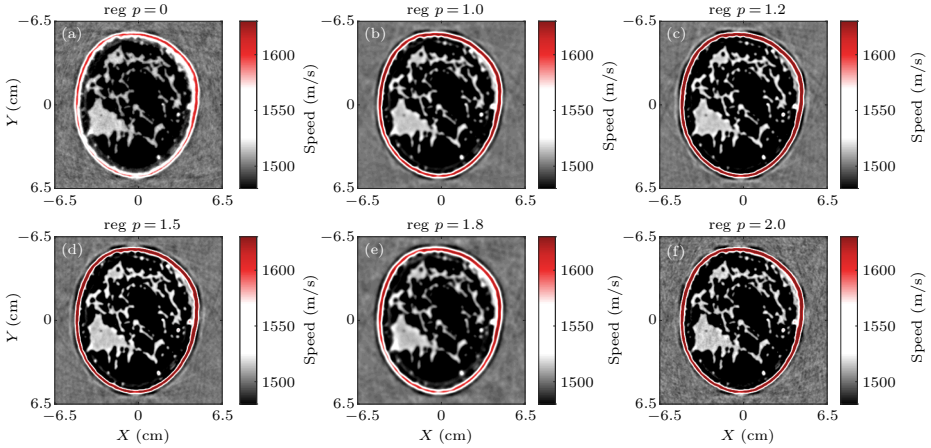


Fig. 12. The reconstruction results of sound speed in the breast model, with the ratio χ_0 set to 10%, are shown in panels (a)–(f), corresponding to the reconstruction results for $p = 0$, 1.0, 1.2, 1.5, 1.8, 2.0, respectively.

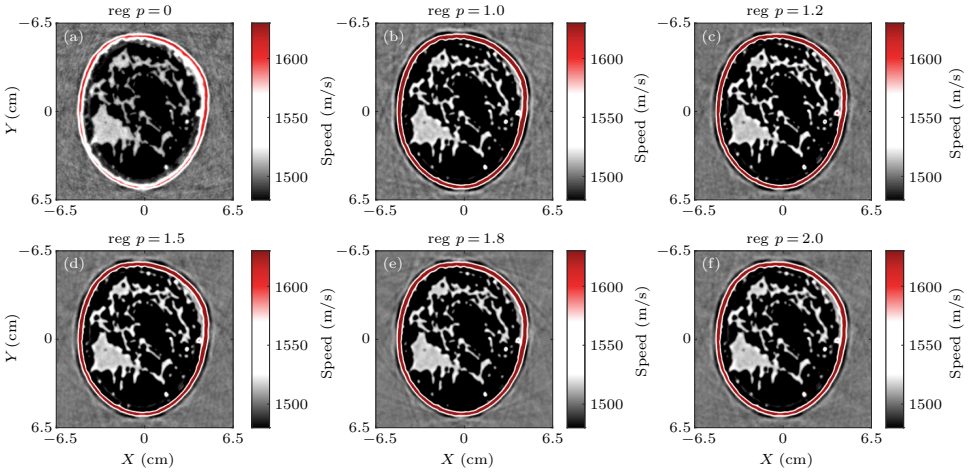


Fig. 13. The reconstruction results of sound speed in the breast model, with the ratio χ_0 set to 15%, are shown in panels (a)–(f), corresponding to the reconstruction results for $p = 0$, 1.0, 1.2, 1.5, 1.8, 2.0, respectively.

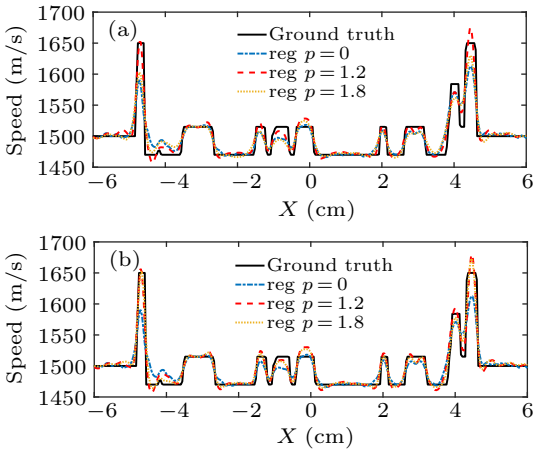


Fig. 14. The partial plots of the sound speed of a horizontal cutting line located at $Y = 0.2$ cm: panels (a) represents the situations $\chi_0 = 10\%$ and panel (b) represents the situations $\chi_0 = 15\%$.

The results demonstrate that the introduction of W_p^1 regularization term leads to varying degrees of improvement in reconstruction quality. When the χ_0 is set to 10%, the reconstructed breast model with $p = 1.2$, $p = 1.5$, and $p = 2.0$ exhibits significantly better visual quality, higher resolution, and fewer artifacts near the skin, all within the same color scale range. However, for $p = 2.0$, background noise outside the model remains unmitigated. As shown in Table 3, quantitative error evaluation indicators, including PSNR, SSIM, and RMSE, further validate these results. While $p = 2.0$ yields relatively higher PSNR and SSIM values, indicating better visual quality, it also produces the highest RMSE among the tested regularization terms. This suggests a larger overall numerical discrepancy between the reconstructed sound speed model and the ground truth, despite its improved visual appearance.

Table 3. Error evaluation of the reconstructed breast model in the rectangular region of interest, with the ratio χ_0 set to 10% and 15%, and for different values of p. The PSNR, SSIM, and RMSE error values are compared between the reconstructed model and the true model.

Error evaluation	FWI	FWI+$W^1_{1.0}$	FWI+$W^1_{1.2}$	FWI+$W^1_{1.5}$	FWI+$W^1_{1.8}$	FWI+$W^1_{2.0}$
PSNR (dB) 10%	37.603	40.187	41.926	41.667	38.799	42.000
SSIM 10%	0.814	0.911	0.948	0.943	0.869	0.947
RMSE 10%	21.744	16.148	13.219	13.618	18.946	20.517
PSNR (dB) 15%	37.603	41.589	41.753	41.637	41.988	41.580
SSIM 15%	0.814	0.948	0.946	0.943	0.950	0.942
RMSE 15%	21.744	13.742	13.485	13.666	13.124	20.486

When $\chi_0 = 15\%$, the results of all tested nonzero p show significant improvements in reconstruction quality compared to $p = 0$, which indicates that appropriately tuning the ratio χ not only enhances reconstruction quality but also simplifies the selection of the p-value. Among them, the result with $p = 1.8$ shows the best performance with PSNR (41.988 dB), SSIM (0.950), and RMSE (13.124). However, when $p = 2.0$, the situation is similar to the case of $p = 2.0$ when $\chi_0 = 10\%$, where the RMSE is relatively higher even the visual quality is decent.

4.2. Numerical test II: Thigh model

We apply the same inversion strategy to the case of the numerical thigh, with the difference being that the frequency range is set from 0.2 MHz to 1.0 MHz, and the number of iterations per frequency point is 8. Figures 15(a) and 16(a) show

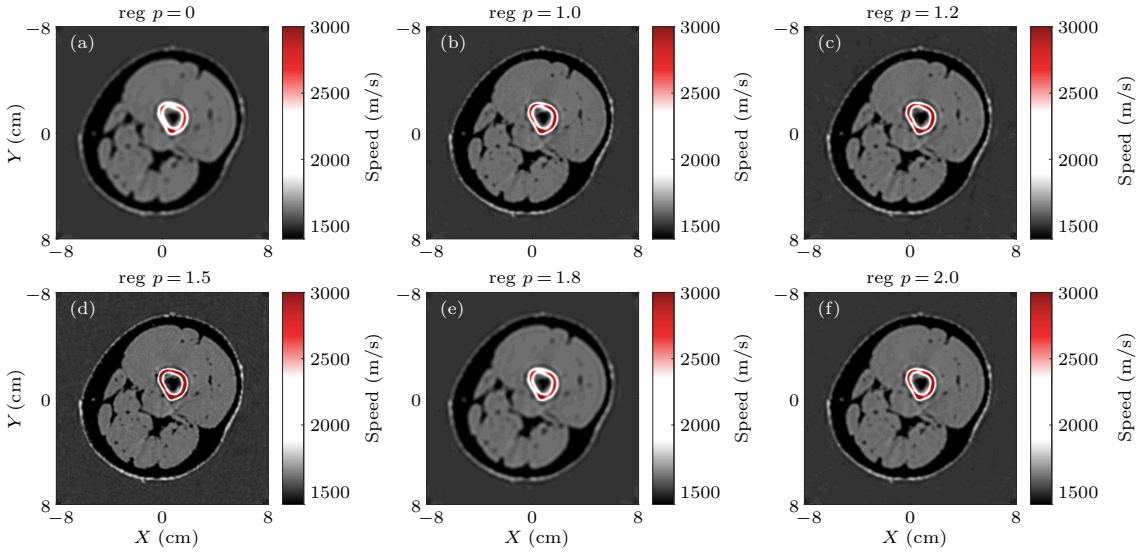


Fig. 15. The reconstruction results of sound speed in the thigh model, with the ratio χ_0 set to 10%, are shown in panels (a)–(f), corresponding to the reconstruction results for $p = 0$, 1.0, 1.2, 1.5, 1.8, 2.0, respectively.

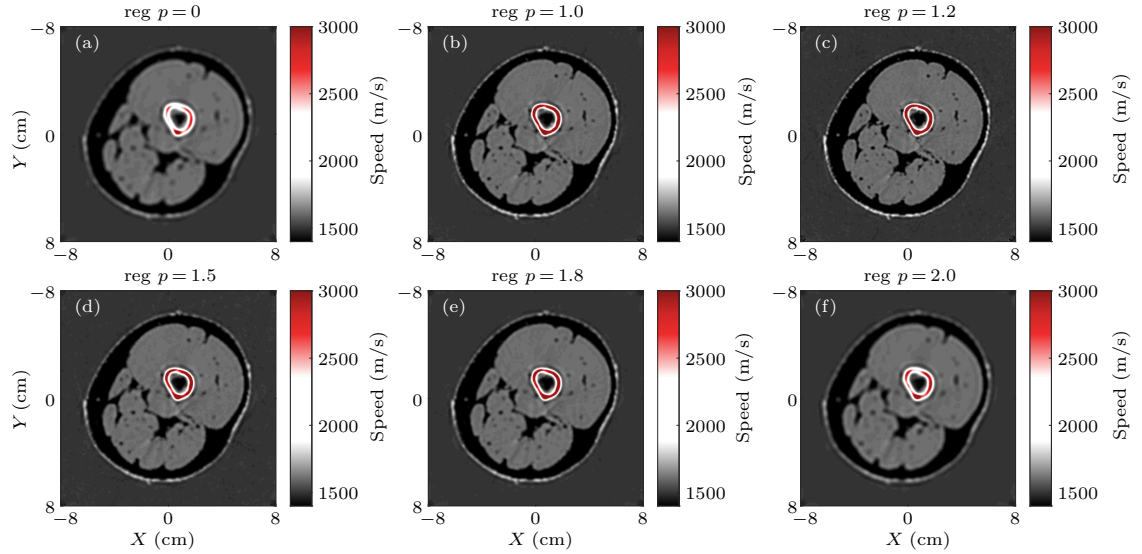


Fig. 16. The reconstruction results of sound speed in the thigh model, with the ratio χ_0 set to 15%, are shown in panels (a)–(f), corresponding to the reconstruction results for $p = 0$, 1.0, 1.2, 1.5, 1.8, 2.0, respectively.

the same sound speed inversion results without regularization ($p = 0$). From the figures, it can be seen that the skin, fat, muscle, and bone of the thigh have been reconstructed. However, compared to the true model, the details in the muscle region are somewhat blurred, and the reconstruction of the cortical bone is uneven, showing discrepancies with the true model, as shown in the line profile comparison in Fig. 17.

Similarly, we perform the same regularization tests in the inversion process. Overall, the results show varying degrees of improvement in reconstruction quality. When $\chi_0 = 10\%$, the reconstructed thigh model with $p = 1.5$ shows significantly better visual quality, higher resolution, and sound speed of bone closer to the true model. These observations are supported by the profile comparison in Fig. 17 and the error evaluation indicators presented in Table 4. When $\chi_0 = 15\%$, all tested values of p (1.0, 1.2, 1.5, 1.8) almost have the consistent improvement in reconstruction quality compared to $p = 0$. The overall sound speed distribution improves, with clearer details in the muscle and more uniform sound speed in the cortical bone, making it more consistent with the true model. The line profiles in Fig. 17 and the error indicators in Table 4 further validate these inversion results.

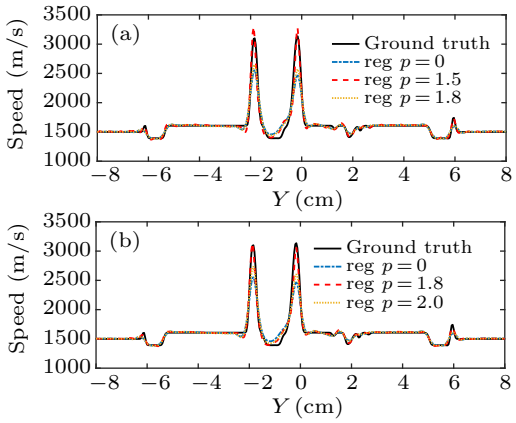


Fig. 17. The partial plots of the sound speed of a vertical cutting line located at $X = 1.2$ cm, panel (a) represents the situations $\chi_0 = 10\%$ and panel (b) represents the situations $\chi_0 = 15\%$.

Table 4. Error evaluation of the reconstructed thigh model in the rectangular region of interest, with the ratio χ_0 set to 10%, and 15%, and for different values of p. The PSNR, SSIM, and RMSE error values are compared between the reconstructed model and the true model.

Error evaluation	FWI	FWI+$W^1_{1.0}$	FWI+$W^1_{1.2}$	FWI+$W^1_{1.5}$	FWI+$W^1_{1.8}$	FWI+$W^1_{2.0}$
PSNR (dB) 10%	35.944	38.719	39.373	39.889	37.003	39.021
SSIM 10%	0.933	0.968	0.974	0.979	0.950	0.971
RMSE 10%	50.959	37.025	34.340	32.361	45.111	35.761
PSNR (dB) 15%	35.944	40.384	39.609	40.833	41.066	37.833
SSIM 15%	0.933	0.982	0.979	0.982	0.983	0.960
RMSE 15%	50.959	30.565	33.419	29.027	28.259	41.000

4.3. *Ex vivo* test: Pork-belly tissue

For the inversion of pork-belly tissue observed data, we employ the same inversion strategy as the numerical breast model. Figures 18(a) and 19(a) display the same results of the sound speed inversion without the regularization term. We can distinctly observe the textures of various tissues (such as skin, fat, and muscle), which correspond well with the actual structures of pork tissue. While, the overall resolution and contrast are not high enough, and the reconstructed sound speed of the skin is uneven. After implementing the regularization strategy, the reconstruction results have been improved, especially in the reconstruction of sound speed at the skin site, as shown in panels (b)–(f) of Figs. 18 and 19.

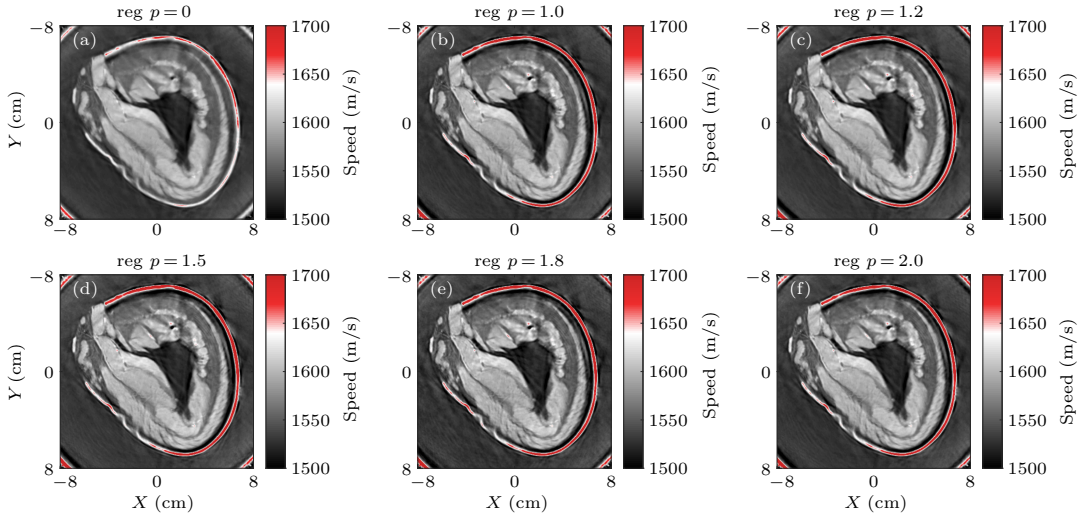


Fig. 18. The reconstruction results of sound speed for *ex vivo* pork-belly tissue, with the ratio χ_0 set to 10%, are shown in panels (a)–(f), corresponding to the reconstruction results for $p = 0$, 1.0, 1.2, 1.5, 1.8, 2.0, respectively.

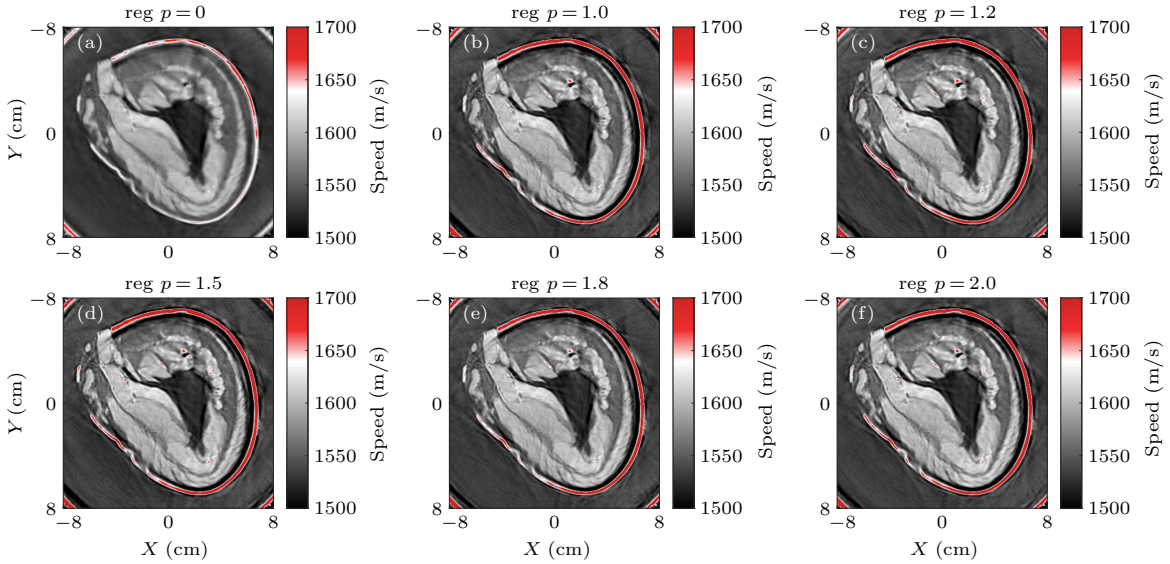


Fig. 19. The reconstruction results of sound speed for *ex vivo* pork-belly tissue, with the ratio χ_0 set to 15%, are shown in panels (a)–(f), corresponding to the reconstruction results for $p = 0$, 1.0, 1.2, 1.5, 1.8, 2.0, respectively.

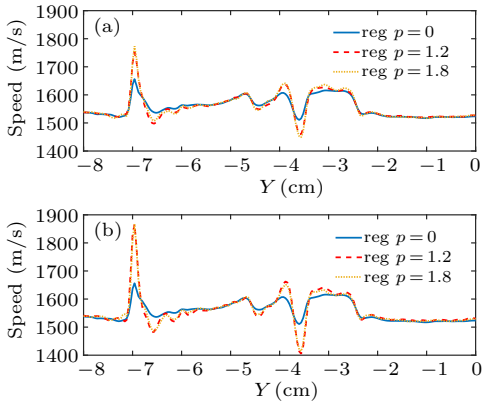


Fig. 20. The partial plots of the sound speed of a vertical cutting line located at $X = 1.32$ cm, panel (a) represents the situations $\chi_0 = 10\%$ and panel (b) represents the situations $\chi_0 = 15\%$.

We find that for both $\chi_0 = 10\%$ and $\chi_0 = 15\%$, in addition to the increase in sound speed at the skin, the details of muscles and fat are clearer, and the contrast is also improved for all tested values of p (1.0, 1.2, 1.5, 1.8, 2.0). These results show consistent improvements in reconstruction quality compared to $p = 0$. From the line profile plots in Fig. 20, significant changes of the 10% and 15% ratios are observed primarily in the sound speed of the skin region and areas with substantial variations in internal sound speed. However, when $\chi_0 = 15\%$, additional artifacts appear, particularly on the outer side of the skin even the internal details of the reconstructed tissue become more refined. This suggests a trade-off between enhancing detail resolution and controlling artifacts. Further fine-tuning of the regularization parameters is required to fully optimize the reconstruction quality across different tissue.

5. Discussion, conclusion, and perspectives

The FWI technique has been developed into a powerful high-resolution quantitative imaging method in the field of ultrasound imaging. It breaks the limitations of conventional pulse echo ultrasound and TOFT, with the potential to develop a high-resolution quantitative imaging technique in a standardized manner, similar to MRI and x-ray CT.

Considering computational efficiency, hardware requirements, and imaging resolution, this study explores the FDFWI with Sobolev space norm regularization. The inverse problem for USCT-FWI is generally ill-posed and highly nonlinear such that the introduction of appropriate prior information by regularization is crucial to stabilize the inversion procedure. However, the weight of regularization shall be carefully balanced during the iterative process. Specifically, we investigate the effect of different types of hyperparameter (i.e., regularization ratio and p-value) on the imaging quality. Numerical and experimental tests demonstrate that the inversion results can be significantly improved under regularization with proper selection of hyperparameters. The results of this study indicate that the designed algorithm exhibits nice adaptability and robust features, even when applied to the experimental target.

Theoretically, increasing the number of iterations for each frequency point allows the model to refine gradually, achieving the same precision even without regularization. This is feasible under ideal conditions, such as low noise and high-quality observed data. However, this approach significantly increases computational costs and runtime. Given the nonlinear nature of FWI, excessive iterations can lead to overfitting and oscillations, particularly in complex models or noisy conditions, resulting in greater instability and error accumulation. Regularization effectively promotes faster and more stable convergence, especially in scenarios with complex models or limited low-frequency information. This study demonstrates that incorporating Sobolev norm regularization yields better inversion results with a similar computational overburden.

In the exploration of Sobolev norm regularization for

FWI, selecting appropriate regularization coefficients and the p-value requires extensive testing and fine-tuning. This challenge arises from the complexity of balancing data fidelity with regularization to mitigate overfitting and stabilize the inversion process. To address this issue, we test different ratios of the regularization term relative to the data fidelity term, and adaptively adjust the regularization coefficient to fix the ratio at the beginning of the inversion process for each frequency point. While when the ratio is too high, it may cause the regularization term to dominate the inversion process, leading to over-enhancement of boundary information. Thus, we have chosen it as 10% and 15% in this work. As for p-value, we test it in the range [1.0, 2.0], and we find that different inversion models may correspond to different optimal p-values. However, when the regularization ratio is large, such as in this study when the ratio is 15%, different p-values almost yield similarly decent inversion results. Therefore, the selection of the p-value can both depend on the model and the regularization term ratio, which makes the inversion process more stable.

In solving the Helmholtz equation using the CBS method, we simplify the medium by assuming uniform density, which deviates from the medium's actual characteristics. In reality, acoustic wave propagation and observed data are influenced by multiple parameters, including sound speed, density, impedance, attenuation, and elastic properties. Focusing solely on sound speed inversion at low frequencies is reasonable, as studies have shown that parameter coupling effects are minimal in this range, with sound speed dominating the inversion process. However, at higher frequencies, incorporating factors like density, attenuation, and elastic effects becomes essential to achieve finer imaging.[5] While this enhances accuracy, it also significantly increases computational demands. Consequently, developing advanced Helmholtz equation solvers that account for density, impedance, and other factors, and exploring multi-parameter joint inversion strategies for complex media, such as musculoskeletal tissues, remain critical challenges in FWI based medical imaging.

Acknowledgements

Project supported by the National Natural Science Foundation of China (Grant No. 12474461), the Basic and Frontier Exploration Project Independently Deployed by Institute of Acoustics, Chinese Academy of Sciences (Grant No. JCQY202402), and the Goal-Oriented Project Independently Deployed by Institute of Acoustics, Chinese Academy of Sciences (Grant No. MBDX202113).

References

[1] Huthwaite P and Simonetti F 2011 *J. Acoust. Soc. Am.* **130** 1721
[2] Fincke J, Zhang X, Shin B, Ely G and Anthony B W 2022 *IEEE Trans. Med. Imaging.* **41** 502
[3] Zhou C, Xu K and Ta D 2023 *J. Acoust. Soc. Am.* **154** 279
[4] Guasch L, Calderón Agudo O, Tang M X, Nachev P and Warner M 2020 *NPJ Digit. Med.* **3** 28
[5] Li Y B, Wang J, Su C, Lin W J, Wang X M and Luo Y 2023 *Chin. Phys. B* **32** 014303
[6] Na S and Wang L V 2021 *Biomed. Opt. Express* **12** 4056
[7] Pan Y, Qiang Y, Liang W, Huang W, Wang N, Wang X, Zhang Z, Qiu W and Zheng H 2024 *Ultrasonics* **143** 107405
[8] Mittendorff L, Young A and Sim J 2022 *J. Med. Radiat. Sci.* **69** 250
[9] Sanches J M, Laine A F and Suri J S 2012 *Ultrasound Imaging: Advances and Applications* (New York: Springer)
[10] Azhari H 2012 *Curr. Pharm. Biotechnol.* **13** 2104
[11] Iyer A, Sun Z, Lambeth K, Singh M, Cleveland C and Sharma N 2024 *IEEE Trans. Rob.* **40** 4322
[12] Martin K 2010 *Introduction to B-mode imaging* (in: Hoskins P R, Martin K, Thrush A eds.) *Diagnostic Ultrasound: Physics and Equipment* (Cambridge University Press) pp. 1–3
[13] Wiskin J W, Borup D T, Iuanow E, Klock J and Lenox M W 2017 *IEEE Trans. Ultrason. Ferroelectr. Freq. Control.* **64** 1161
[14] Duric N, Littrup P, Poulo L, Babkin A, Pevzner R, Holsapple E, Rama O and Glide C 2007 *Med. Phys.* **34** 773
[15] Qu X, Azuma T, Yogi T, Azuma S, Takeuchi H, Tamano S and Takagi S 2016 *J. Med. Ultrason.* **43** 461
[16] Perrot V, Polichetti M, Varray F and Garcia D 2021 *Ultrasonics* **111** 106309
[17] Bao Y and Jia J 2020 *IEEE Trans. Instrum. Meas.* **69** 974
[18] Devaney A 1982 *Ultrason. Imag.* **4** 336
[19] Simonetti F, Huang L, Duric N and Littrup P 2009 *Med. Phys.* **36** 2955
[20] Schuster G T 1996 *Geophys. J. Int.* **127** 427
[21] Wu X, Li Y, Su C, Li P, Wang X and Lin W 2023 *Ultrasonics* **132** 107004
[22] Virieux J and Operto S 2009 *Geophysics* **64** 74
[23] Pratt R, Huang L, Duric N and Littrup P 2007 *Proceedings* **6510** 65104
[24] Ali R, *et al.* 2024 *IEEE Trans. Med. Imaging.* **43** 2988
[25] Wu X, Li Y, Su C, Li P and Lin W 2025 *Ultrasonics* **147** 107505
[26] Li Y, Shi Q, Li Y, Song X, Liu C, Ta D and Wang W 2021 *Chin. Phys. B* **30** 014302
[27] Régo R C L, *et al.* 2019 *Proceedings of the* 16*th International Congress of the Brazilian Geophysical Society & Expogef* August 19–22, 2019, Rio de Janeiro, Brazil
[28] Bunks C, Saleck F M, Zaleski S and Chavent G 1995 *Geophysics.* **60** 1457
[29] Pratt R G 1999 *Geophysics* **64** 888
[30] Pratt R and Shipp R 1999 *Geophysics* **64** 902
[31] Sirgue L and Pratt R G 2004 *Geophysics* **69** 231
[32] Kazei V V, Kalita M and Alkhalifah T 2017 *Proceedings of the* 79*th EAGE Conference and Exhibition* 2017, June 12–15, 2017 Paris, France, pp. 1–5
[33] Tikhonov A N and Arsenin V Y 1977 *Solutions of ill-posed problems* (New York: John Wiley & Sons)
[34] Bertete-Aguirre H, Cherkaev E and Oristaglio M 2002 *Geophysical Journal International* **149** 499
[35] Aghamiry H, Gholami A and Operto S 2020 *Geophysics* **85** 116
[36] Aghamiry H S, Gholami A and Operto S 2021 *SIAM Journal on Imaging Sciences* **14** 58
[37] Aghamiry H S, Gholami A and Operto S 2018 *Proceedings of the SEG meeting,* October 18, 2018, Anaheim, California, USA pp. 1253–1257
[38] Agazade K, Gholami A and Aghamiry H S 2023 84*th EAGE Annual Conference & Exhibition*, June 2023, Vienna, Autria pp. 1–5
[39] Tarantola A 2005 *Inverse Problem Theory and Methods for Model Parameter Estimation* (Philadelphia: Society for Industrial and Applied Mathematics) pp. 81–96
[40] Wang Y 2017 *Seismic Inversion: Theory and Applications* (John Wiley & Sons, Ltd)
[41] Nocedal J and Wright S J 2006 *Numerical optimization* (New York: Springer) pp. 30–191
[42] Byrd R H, Lu P, Nocedal J and Zhu C 1995 *SIAM J. Sci. Comput.* **16** 1190
[43] Plessix R E 2006 *Geophys. J. Int.* **167** 495
[44] Osnabrugge G, Leedumrongwatthanakun S and Vellekoop I M 2016 *J. Comput. Phys.* **322** 113
[45] Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 *IEEE Trans. Image Process* **13** 600
[46] Lou Y, *et al.* 2017 *J. Biomed. Opt.* **22** 041015
[47] Treeby B E and Cox B T 2010 *J. Biomed. Opt.* **15** 021314