Full-wave Image Reconstruction in Transcranial Photoacoustic Computed Tomography using a Finite Element Method

Yilin Luo, Hsuan-Kai Huang, Karteekeya Sastry, Peng Hu, Xin Tong, Joseph Kuo, Yousuf Aborahama, Shuai Na, Umberto Villa, *Member, IEEE*, Mark. A. Anastasio, *Fellow, IEEE*, Lihong V. Wang, *Fellow, IEEE*

***Abstract*—Transcranial photoacoustic computed tomography presents challenges in human brain imaging due to skull-induced acoustic aberration. Existing full-wave image reconstruction methods rely on a unified elastic wave equation for skull shear and longitudinal wave propagation, therefore demanding substantial computational resources. We propose an efficient discrete imaging model based on finite element discretization. The elastic wave equation for solids is solely applied to the hard-tissue skull region, while the soft-tissue or coupling-medium region that dominates the simulation domain is modeled with the simpler acoustic wave equation for liquids. The solid-liquid interfaces are explicitly modeled with elastic-acoustic coupling. Furthermore, finite element discretization allows coarser, irregular meshes to conform to object geometry. These factors significantly reduce the linear system size by 20 times to facilitate accurate whole-brain simulations with improved speed. We derive a matched forward-adjoint operator pair based on the model to enable integration with various optimization algorithms. We validate the reconstruction framework through numerical simulations and phantom experiments.**

***Index Terms*—Domain Decomposition, Finite Element Method, Full-wave Reconstruction, Acoustic-elastic coupling, Transcranial imaging, Photoacoustic tomography, Skull deaberration**

I. INTRODUCTION

SKULL-induced acoustic aberration is one of the major obstacles in translating photoacoustic computed tomography (PACT) to noninvasive human brain imaging. The skull's presence introduces substantial distortions in the photoacoustic signal through attenuation, aberration, reverberation, and mode conversion [1]–[4]. Since the skull is a thick and solid medium where longitudinal and shear waves simultaneously propagate, we need to consider its elastic properties during image reconstruction. Neglecting shear waves can lead to image artifacts that significantly deteriorate the image quality [5].

Existing PACT reconstruction methods that account for the elastic properties of the skull can be broadly classified into three categories: ray-based, time reversal, and model-based optimization methods. Ray-based methods approximate the acoustic medium as homogeneous layers, backpropagating the waves in each layer individually [6]. These methods are computationally efficient but assume a single-layer homogeneous skull model without reverberation. Time reversal (TR) methods exploit the wave equation's TR property by reversing the recorded acoustic signals in time [7]–[9]. These techniques can accommodate arbitrary detection geometries and heterogeneous media. However, finite sampling or limited measurement view can render the inverse problem ill-posed. In such situations, model-based optimization methods are commonly employed [10]–[13]. Because they provide a flexible framework for regularization, these methods hold potential to mitigate the effects of data incompleteness or other physical factors.

In optimization-based reconstruction methods, it is often necessary to incorporate a forward and adjoint model to compute the gradients of the data fidelity terms. For transcranial PACT imaging, a discrete forward-adjoint pair based on the finite-difference time-domain (FDTD) method has been proposed and demonstrated notable deaberration effects for transcranial images [10], [14]. Nevertheless, the FDTD approach employs a uniform Cartesian grid across the entire simulation domain, which restricts its adaptability to irregular skull boundaries. Although increasing grid resolution can mitigate the issue, it comes at the expense of increased computational cost. More recently, a continuous forward-adjoint pair was introduced in the form of analytical partial differential equations [13]. While initially demonstrated within the framework of the pseudo-spectral time domain (PSTD) method, the continuous adjoint is independent of discretization. However, both the discrete adjoint in [14] and the continuous adjoint in [13] use a unified linear isotropic, lossy, and

This work was sponsored by the United States National Institutes of Health (NIH) grants U01 EB029823, R35 CA220436 (Outstanding Investigator Award), and R01 EB028277. *(Corresponding author: Lihong V. Wang.)*

Y. L., K. S., P. H., X. T., Y. A. and L. V. W. are with the Caltech Optical Imaging Laboratory, Andrew and Peggy Cherng Department of Medical Engineering and the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: yilinluo@caltech.edu; lvw@caltech.edu).

H. H., J. K, and M. M. A. are with the Computational Imaging Science Laboratory, Department of Electrical and Computer Engineering and Department of Bioengineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA.

U. V. is with Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712, USA.

S. N. is with the National Biomedical Imaging Center, Peking University, Beijing 100871, China.

heterogeneous stress-velocity elastic wave equation to model the entire domain, with the fluid region represented by a shear wave velocity of zero. In such unified models, velocity continuity is implicitly assumed throughout the entire domain [15]. Nonetheless, for the boundary between the inviscid fluid and solid, only the normal component of velocity remains continuous across the interfaces. Consequently, a unified model does not satisfy the correct interface condition and may compromise accuracy [16]. Moreover, the unified configuration is inefficient, as the majority of the simulation domain consists of soft tissue and coupling medium, which can be more efficiently modeled by the acoustic equation.

To address the limitations, we propose a discrete forward-adjoint operator pair for transcranial PACT based on the ssssfinite element method (FEM). The FEM is a well-established technique in seismology, known for its ability to provide accurate solutions given the complex earth structures and convenience for multiphysics simulation [17], [18]. In this work, we divide the simulation domain into two regions: a fluid region and a solid region. The solid region is characterized by the displacement elastic wave equation, while the fluid region is described by the acoustic equation, which requires fewer unknowns. By incorporating acoustic-elastic coupling, we explicitly consider the fluid-solid interfaces at the boundaries between the skull and the surrounding soft tissue. This domain decomposition approach not only enables more accurate modeling of the boundary but also fundamentally reduces the number of degrees of freedom (DOF) associated with each node. In addition, the FEM discretization allows the use of a coarser, unstructured, and variable-size mesh that adapts to the skull geometry, decreasing the number of nodes. These two aspects collectively lead to a remarkable reduction in the total number of unknowns in the linear system. This approach demonstrates the potential for enhancing reconstruction speed while maintaining accuracy.

Our approach is different from commercial FEM software in key three aspects: 1) To the best of our knowledge, no commercial FEM tools currently provide a numerically matched adjoint operator specifically for transcranial PACT image reconstruction; 2) Our method is based on an open-source package, thereby eliminating the need for costly subscriptions; 3) The open-source nature of our software enhances adaptability, allowing users to tailor the code for specific applications, unlike commercial products, which often have limited customization capabilities.

This paper is organized as follows: In Section II, we briefly describe the imaging physics and the related wave equations for transcranial PACT. Then we derive the explicit formulation of the discrete forward and adjoint operators. In Section III, we validate our implementation and demonstrate its feasibility through a comparative study with an established FDTD forward-adjoint operator pair [14]. To illustrate the practical applicability of our method, we apply it to phantom data. Finally, we end the paper with a conclusion and discussion of the merits and limitations of our method in Section IV.

II. Theory

The physics for transcranial photoacoustic wavefield propagation is described below in its continuous and discrete forms. Here, we treat the soft tissue and the coupling medium as a lossless, inviscid, and compressible fluid, and the skull as an isotropic, heterogeneous solid medium. We employ an empirical diffusive absorption model to account for the acoustic attenuation of the skull [14], [19]. This assumption is valid for the approximately monochromatic photoacoustic signals in the MHz frequency range for transcranial imaging [19]–[21]. Throughout the derivations, bold lowercase symbols represent vectors, while uppercase symbols denote tensors or matrices. While we present the derivations in two dimensions (2D) for simplicity, they can be readily extended to three-dimensional (3D) cases.

A. Transcranial photoacoustic wavefield propagation: continuous formulation

A typical simulation for transcranial PACT involves three regions (see Fig. 1): the fluid region Ω_f, representing soft tissue

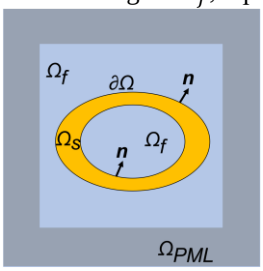

Fig. 1. **Domain decomposition in transcranial PACT simulation.** Ω_f represents the fluid region (blue) for soft tissue and coupling medium; Ω_s corresponds to the solid region (yellow) for skull; Ω_{PML} denotes the PML (gray); $\partial\Omega$ describes the solid-fluid interface for the solid-liquid boundary; $\boldsymbol{n}$ is the outward surface normal of the skull boundary.

and coupling medium; the solid region Ω_s, corresponding to the skull, and the perfectly matched layer (PML) Ω_{PML}, applied to the outer boundary of the whole domain for wave truncation. To facilitate accurate and efficient modeling, we adopt a domain decomposition approach that independently models the physics within each domain. Consequently, we formulate three sets of coupled equations: the acoustic equation governing the fluid domain, the elastic equation governing the solid domain, and the modified acoustic equation governing the PML.

In the fluid domain, Ω_f, the photoacoustic wavefield propagation is governed by the lossless acoustic equation [14]:

$$\frac{1}{c_f^2}\frac{\partial^2 p}{\partial t^2} - \nabla^2 p = 0, \tag{1}$$

subject to the initial conditions at location $\boldsymbol{r} \in \mathbb{R}^2$,

$$p(\boldsymbol{r}, t)|_{t=0} = p_0, \tag{2}$$

$$\frac{\partial p(\boldsymbol{r}, t)}{\partial t}|_{t=0} = 0, \tag{3}$$

where p is the acoustic pressure, c_f is the speed of sound (SOS) in the fluid region, and p_0 is the initial pressure distribution.

In the solid domain, Ω_s, the wavefield can be described by the heterogeneous and isotropic linear elastic equation [22],

$$\rho_s \frac{\partial^2 \boldsymbol{u}}{\partial t^2} + \alpha \frac{\partial \boldsymbol{u}}{\partial t} = \nabla \cdot \left(\mathbb{C} : \frac{1}{2} (\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\dagger) \right), \tag{4}$$

where $\boldsymbol{u}$ is the displacement vector, ρ_s is the density of the solid material, $\mathbb{C}$ is the stiffness tensor, and α is the frequency-independent attenuation coefficient. The notation : and † denotes the inner product of two second-order tensors and matrix transpose, respectively.

For an isotropic material, the stiffness tensor reduces to the following expression in terms of the shear and longitudinal wave speeds, c_s and c_p, respectively,

$$C_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} \right), \tag{5}$$

where δ is the Kronecker delta. The Lamé constants λ and μ are related to the speed of sound as

$$\lambda = c_p^2 \rho_s - 2 c_s^2 \rho_s, \tag{6}$$

$$\mu = c_s^2 \rho_s. \tag{7}$$

Equations (1) and (4) are coupled through the boundary conditions along the interface $\partial\Omega$ [22]. The coupling is reflected in the continuity of the normal component of displacement acceleration from solid to liquid as

$$\boldsymbol{n} \cdot \frac{1}{\rho_f} \nabla p = -\boldsymbol{n} \cdot \frac{\partial^2 \boldsymbol{u}}{\partial t^2} \text{ on } \partial\Omega, \tag{8}$$

and the continuity of pressure from liquid to solid as

$$\left(\mathbb{C} : \frac{1}{2} (\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\dagger) \right) \cdot \boldsymbol{n} = -p\boldsymbol{n} \text{ on } \partial\Omega, \tag{9}$$

where $\boldsymbol{n}$ is the outward surface normal at the skull boundary.

In addition, we apply a PML at the outermost boundary of the whole domain, Ω_{PML}, to simulate the free-field condition by exponentially attenuating propagating waves according to [23]. The modified acoustic wave equations in the PML are given as

$$\frac{1}{c_f^2} \frac{\partial^2 p}{\partial t^2} + \chi \frac{\partial p}{\partial t} + \kappa p - \nabla^2 p - \nabla \cdot \boldsymbol{w} = 0, \text{ and} \tag{10}$$

$$\frac{\partial \boldsymbol{w}}{\partial t} + \boldsymbol{A}\boldsymbol{w} + \boldsymbol{B} \nabla p = \boldsymbol{0}, \tag{11}$$

where $\boldsymbol{w}$ is the vector-valued auxiliary variable. The definition of the PML-related properties $\boldsymbol{A}, \boldsymbol{B}, \chi, \kappa$ can be found in [23], and the attenuation terms related to their γ vanish in our 2D model. When $\boldsymbol{A}, \boldsymbol{B}, \chi, \kappa$ are set to zero, (10–11) simplify to the original acoustic wave equation in (1).

B. Transcranial photoacoustic propagation: discrete formulation

To solve the coupled problem with FEM, we need to derive the weak form for (1–11) by multiplying them with test functions and integrating by parts over the problem domain [24]. The variational form of the elastic wave equation after applying the Gauss-Green theorem is given as

$$\int_{\Omega_s} \rho_s \boldsymbol{\psi} \cdot \frac{\partial^2 \boldsymbol{u}}{\partial t^2} dS + \int_{\Omega_s} \alpha \boldsymbol{\psi} \cdot \frac{\partial \boldsymbol{u}}{\partial t} dS = -\int_{\Omega_s} \nabla \boldsymbol{\psi} : \mathbb{C} : \left(\frac{1}{2} (\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^\dagger) \right) dS + \oint_{\partial\Omega} \boldsymbol{\psi} \cdot p\boldsymbol{n}\, dl, \tag{12}$$

and the weak form of the modified acoustic wave equation is expressed as

$$\int_{\Omega_f} \phi \frac{1}{c_f^2} \frac{\partial^2 p}{\partial t^2} dS + \int_{\Omega_f} \chi \phi \frac{\partial p}{\partial t} dS + \int_{\Omega_f} \kappa \phi p\, dS + \int_{\Omega_f} \nabla \phi \nabla p\, dS + \oint_{\partial\Omega} \rho_f \phi \boldsymbol{n} \cdot \frac{\partial^2 \boldsymbol{u}}{\partial t^2} dl - \int_{\Omega_f} \phi \nabla \cdot \boldsymbol{w}\, dS = 0, \tag{13}$$

$$\int_{\Omega_f} \boldsymbol{\psi} \cdot \frac{\partial \boldsymbol{w}}{\partial t} dS + \int_{\Omega_f} \boldsymbol{\psi} \cdot (\boldsymbol{A}\boldsymbol{w})\, dS + \int_{\Omega_f} \boldsymbol{\psi} \cdot (\boldsymbol{B} \nabla p)\, dS = \boldsymbol{0}, \tag{14}$$

where $\boldsymbol{\psi} \in [H^1(\Omega_s)]^2$, $\phi \in H^1(\Omega_f)$ are arbitrary vector-valued and scalar test functions, respectively.

We spatially discretize the above weak forms (12–14) using the standard Galerkin method [25], [26]. Let N^u, N^p, and N^w specifiy the number of spatial nodes in the solid, fluid, and the PML domains, respectively. The discretization results in three coupled linear differential equations in the displacement vector, $\boldsymbol{u} \in \mathbb{R}^{2N^u}$, the pressure vector, $\boldsymbol{p} \in \mathbb{R}^{N^p}$ and the auxiliary vector for the PML, $\boldsymbol{w} \in \mathbb{R}^{2N^w}$:

$$\boldsymbol{M^u}\ddot{\boldsymbol{u}} + \boldsymbol{C^u}\dot{\boldsymbol{u}} + \boldsymbol{K^u}\boldsymbol{u} + \boldsymbol{R^u}\boldsymbol{p} = \boldsymbol{0}, \text{ in the solid domain,} \tag{15}$$

$$\boldsymbol{M^p}\ddot{\boldsymbol{p}} + \boldsymbol{R^p}\ddot{\boldsymbol{u}} + \boldsymbol{C^p}\dot{\boldsymbol{p}} + \boldsymbol{K^p}\boldsymbol{p} + \boldsymbol{E}\boldsymbol{w} = \boldsymbol{0}, \text{ in the fluid and PML domain,} \tag{16}$$

$$\boldsymbol{C^w}\dot{\boldsymbol{w}} + \boldsymbol{K^w}\boldsymbol{w} + \boldsymbol{F}\boldsymbol{p} = \boldsymbol{0}, \text{ in the PML domain,} \tag{17}$$

where $\boldsymbol{M}, \boldsymbol{C}, \boldsymbol{K}$ represent the mass matrix, the damping matrix, and the stiffness matrix, respectively. $\boldsymbol{M^u}, \boldsymbol{C^u}, \boldsymbol{K^u} \in \mathbb{R}^{2N^u \times 2N^u}$; $\boldsymbol{M^p}, \boldsymbol{C^p}, \boldsymbol{K^p} \in \mathbb{R}^{N^p \times N^p}$, $\boldsymbol{C^w}, \boldsymbol{K^w} \in \mathbb{R}^{2N^w \times 2N^w}$, and they are all symmetric positive definite [25], [27]. $\boldsymbol{R^u} \in \mathbb{R}^{2N^u \times N^p}$, $\boldsymbol{R^p} \in \mathbb{R}^{N^p \times 2N^u}$ are the coupling matrices for the fluid-solid interface. $\boldsymbol{E} \in \mathbb{R}^{N^p \times 2N^w}$, $\boldsymbol{F} \in \mathbb{R}^{2N^w \times N^p}$ are the matrices introduced by the PML. The notations of the single and double dots are used to denote the first and second time derivatives, respectively. For a detailed definition of the entries in the matrices, please refer to Appendix A.

Equations (15–17) can be reorganized into a single linear equation

$$\begin{bmatrix} \boldsymbol{M_u} & \boldsymbol{0}_{N^u\times N^u} & \boldsymbol{0}_{N^u\times N^u} \\ \boldsymbol{R_p} & \boldsymbol{M_p} & \boldsymbol{0}_{N^p\times N^p} \\ \boldsymbol{0}_{N^w\times N^w} & \boldsymbol{0}_{N^w\times N^w} & \boldsymbol{0}_{N^w\times N^w} \end{bmatrix} \begin{bmatrix} \ddot{\boldsymbol{u}} \\ \ddot{\boldsymbol{p}} \\ \ddot{\boldsymbol{w}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{C_u} & \boldsymbol{0}_{N^u\times N^u} & \boldsymbol{0}_{N^u\times N^u} \\ \boldsymbol{0}_{N^p\times N^p} & \boldsymbol{C_p} & \boldsymbol{0}_{N^p\times N^p} \\ \boldsymbol{0}_{N^w\times N^w} & \boldsymbol{0}_{N^w\times N^w} & \boldsymbol{C_w} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{u}} \\ \dot{\boldsymbol{p}} \\ \dot{\boldsymbol{w}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{K_u} & \boldsymbol{R_u} & \boldsymbol{0}_{N^u\times N^u} \\ \boldsymbol{0}_{N^p\times N^p} & \boldsymbol{K_p} & \boldsymbol{E} \\ \boldsymbol{0}_{N^w\times N^w} & \boldsymbol{F} & \boldsymbol{K_w} \end{bmatrix} \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{p} \\ \boldsymbol{w} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0}_{N^u\times 1} \\ \boldsymbol{0}_{N^p\times 1} \\ \boldsymbol{0}_{N^w\times 1} \end{bmatrix}, \tag{18}$$

and we simplify it to

$$\boldsymbol{M}\ddot{\boldsymbol{x}} + \boldsymbol{C}\dot{\boldsymbol{x}} + \boldsymbol{K}\boldsymbol{x} = \boldsymbol{0}, \tag{19}$$

with $\boldsymbol{M}, \boldsymbol{C}, \boldsymbol{K} \in \mathbb{R}^{N\times N}$, $\boldsymbol{x} \in \mathbb{R}^{N\times 1}$. N is the total number of DOFs in the entire domain, with $N = 2N^u + N^p + 2N^w$.

Next, we discretize in time using the implicit Newmark-beta time stepping with $\gamma = \frac{1}{2}, \beta = \frac{1}{4}$ for unconditional stability [28]. At the $(i+1)^{\text{th}}$ time step ($i = 0, 1, \ldots, T-1$), the field can be updated as

$$\boldsymbol{M}\ddot{\boldsymbol{x}}^{i+1} + \boldsymbol{C}\dot{\boldsymbol{x}}^{i+1} + \boldsymbol{K}\boldsymbol{x}^{i+1} = \boldsymbol{0}, \tag{20}$$

$$\dot{\boldsymbol{x}}^{i+1} = \dot{\boldsymbol{x}}^{i} + \frac{1}{2}\Delta t\ddot{\boldsymbol{x}}^{i} + \frac{1}{2}\Delta t\ddot{\boldsymbol{x}}^{i+1}, \tag{21}$$

$$\boldsymbol{x}^{i+1} = \boldsymbol{x}^{i} + \Delta t\dot{\boldsymbol{x}}^{i} + \frac{1}{4}\Delta t^2\ddot{\boldsymbol{x}}^{i} + \frac{1}{4}\Delta t^2\ddot{\boldsymbol{x}}^{i+1}. \tag{22}$$

The preceding three equations can be merged into the following matrix form:

$$\begin{bmatrix} \boldsymbol{M} & \boldsymbol{C} & \boldsymbol{K} \\ -\frac{1}{2}\Delta t\boldsymbol{I}_{N\times N} & \boldsymbol{I}_{N\times N} & \boldsymbol{0}_{N\times N} \\ -\frac{1}{4}\Delta t^2\boldsymbol{I}_{N\times N} & \boldsymbol{0}_{N\times N} & \boldsymbol{I}_{N\times N} \end{bmatrix} \begin{bmatrix} \ddot{\boldsymbol{x}}^{i+1} \\ \dot{\boldsymbol{x}}^{i+1} \\ \boldsymbol{x}^{i+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0}_{N\times N} & \boldsymbol{0}_{N\times N} & \boldsymbol{0}_{N\times N} \\ \frac{1}{2}\Delta t\boldsymbol{I}_{N\times N} & \boldsymbol{I}_{N\times N} & \boldsymbol{0}_{N\times N} \\ \frac{1}{4}\Delta t^2\boldsymbol{I}_{N\times N} & \Delta t\boldsymbol{I}_{N\times N} & \boldsymbol{I}_{N\times N} \end{bmatrix} \begin{bmatrix} \ddot{\boldsymbol{x}}^{i} \\ \dot{\boldsymbol{x}}^{i} \\ \boldsymbol{x}^{i} \end{bmatrix}, \tag{23}$$

with $\dot{\boldsymbol{x}}^{i+1}, \ddot{\boldsymbol{x}}^{i+1}$ being the first and second temporal derivatives of $\boldsymbol{x}^{i+1}$, and they will be solved simultaneously with $\boldsymbol{x}^{i+1}$.

We rewrite (23) into

$$\boldsymbol{W}\boldsymbol{m}^{i+1} = \boldsymbol{Q}\boldsymbol{m}^{i}, \tag{24}$$

where $\boldsymbol{W}, \boldsymbol{Q} \in \mathbb{R}^{3N\times 3N}, \boldsymbol{m}^i \in \mathbb{R}^{3N\times 1}$. Therefore, the original differential equations have been converted into algebraic equations using FEM.

Since $\boldsymbol{W}$ is invertible ($\det(\boldsymbol{W}) > 0$), we can solve for the current time step to reach

$$\boldsymbol{m}^{i+1} = \boldsymbol{W}^{-1}\boldsymbol{Q}\boldsymbol{m}^{i}, \tag{25}$$

where $\boldsymbol{W}^{-1}\boldsymbol{Q}$ is essentially the propagation matrix. Note that we do not explicitly solve $\boldsymbol{W}^{-1}\boldsymbol{Q}$; instead, we decompose it into smaller recursive steps through block elimination, as elaborated in the subsequent text.

The photoacoustic wavefield variables can be propagated forward in time from $t = 0$ to $t = (T-1)\Delta t$ as

$$\begin{bmatrix} \boldsymbol{m}^0 \\ \boldsymbol{m}^1 \\ \vdots \\ \boldsymbol{m}^{T-2} \\ \boldsymbol{m}^{T-1} \end{bmatrix} = \boldsymbol{P}_{T-1}\boldsymbol{P}_{T-2}\cdots\boldsymbol{P}_1 \begin{bmatrix} \boldsymbol{m}^0 \\ \boldsymbol{0}_{3N\times 1} \\ \vdots \\ \vdots \\ \boldsymbol{0}_{3N\times 1} \end{bmatrix}, \tag{26}$$

where $\boldsymbol{P}_i$ has a block structure with identities in the first i diagonal blocks, and $\boldsymbol{W}^{-1}\boldsymbol{Q}$ in the i^{th} block row and $(i-1)^{th}$ block column, i.e.,

$$\boldsymbol{P}_i = \begin{bmatrix} \boldsymbol{I}_{3N\times 3N} & \cdots \boldsymbol{0}_{3N\times 3N} & \boldsymbol{0}_{3N\times 3N} & \boldsymbol{0}_{3N\times(T-i-1)\cdot 3N} \\ \vdots & \ddots \quad \vdots & \vdots & \vdots \\ \boldsymbol{0}_{3N\times 3N} & \cdots \boldsymbol{I}_{3N\times 3N} & \vdots & \vdots \\ \boldsymbol{0}_{3N\times 3N} & \cdots \boldsymbol{W}^{-1}\boldsymbol{Q} & \boldsymbol{0}_{3N\times 3N} & \boldsymbol{0}_{3N\times(T-i-1)\cdot 3N} \\ \boldsymbol{0}_{(T-i-1)\cdot 3N\times 3N} \cdots & \cdots & \boldsymbol{0}_{(T-i-1)\cdot 3N\times 3N} & \boldsymbol{0}_{(T-i-1)\cdot 3N\times(T-i-1)\cdot 3N} \end{bmatrix} \in \mathbb{R}^{3TN\times 3TN}, \tag{27}$$

with $i = 1, 2, \ldots T-1$.

From the initial condition in (2), we can map the initial pressure $\boldsymbol{p}_0 \in \mathbb{R}^{N^p\times 1}$ to $\boldsymbol{m}^0$ using

$$\begin{bmatrix} \boldsymbol{m}^0 \\ \boldsymbol{0}_{3N\times 1} \\ \vdots \\ \vdots \\ \boldsymbol{0}_{3N\times 1} \end{bmatrix} = \boldsymbol{P}_0\boldsymbol{p}_0, \tag{28}$$

where

$$\boldsymbol{P}_0 = \begin{bmatrix} \boldsymbol{\tau} \\ \boldsymbol{0}_{3N\times N^p} \\ \vdots \\ \boldsymbol{0}_{3N\times N^p} \end{bmatrix} \in \mathbb{R}^{3TN\times N^p}, \boldsymbol{\tau} = \begin{bmatrix} \boldsymbol{0}_{N\times N^p} \\ \boldsymbol{0}_{N\times N^p} \\ \boldsymbol{0}_{N^u\times N^p} \\ \boldsymbol{I}_{N^p\times N^p} \\ \boldsymbol{0}_{N^w\times N^p} \end{bmatrix} \in \mathbb{R}^{3N\times N^p}. \tag{29}$$

Suppose we have L transducers to record the acoustic signals, we relate the measured data $\hat{\boldsymbol{p}} \in \mathbb{R}^{LT\times 3TN}$ to the computed field quantities via

$$\hat{\boldsymbol{p}} = \begin{bmatrix} \hat{\boldsymbol{p}}^0 \\ \hat{\boldsymbol{p}}^1 \\ \vdots \\ \hat{\boldsymbol{p}}^{T-1} \end{bmatrix} = \boldsymbol{S}\begin{bmatrix} \boldsymbol{m}^0 \\ \boldsymbol{m}^1 \\ \vdots \\ \boldsymbol{m}^{T-1} \end{bmatrix} = \boldsymbol{S}\boldsymbol{P}_{T-1}\cdots\boldsymbol{P}_1\begin{bmatrix} \boldsymbol{m}^0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \boldsymbol{S}\boldsymbol{P}_{T-1}\cdots\boldsymbol{P}_1\boldsymbol{P}_0\boldsymbol{p}_0, \tag{30}$$

with the sampling matrix defined as

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{\Theta} & \boldsymbol{0}_{L\times 3N} & \cdots & \boldsymbol{0}_{L\times 3N} \\ \boldsymbol{0}_{L\times 3N} & \boldsymbol{\Theta} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{0}_{L\times 3N} \\ \boldsymbol{0}_{L\times 3N} & \cdots & \boldsymbol{0}_{L\times 3N} & \boldsymbol{\Theta} \end{bmatrix} \in \mathbb{R}^{LT\times 3TN}, \tag{31}$$

$$\boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{s}_1 \\ \boldsymbol{s}_2 \\ \vdots \\ \boldsymbol{s}_L \end{bmatrix} \in \mathbb{R}^{L\times 3N}, \tag{32}$$

$$\boldsymbol{s}_l = [\boldsymbol{0}_{1\times N}, \boldsymbol{0}_{1\times N}, \boldsymbol{0}_{1\times N^u}, \boldsymbol{R}_l, \boldsymbol{0}_{1\times N^w}] \in \mathbb{R}^{1\times 3N}, \tag{33}$$

where $\boldsymbol{R}_l \in \mathbb{R}^{1\times N_p}$ is the weighting vector that relates the interpolated value at the l^{th} transducer location ($l = 0, 1, \ldots, L$) to the neighboring DOFs.

Finally, we reach the discrete forward imaging model for transcranial PACT:

$$\hat{\boldsymbol{p}} = \boldsymbol{S}\boldsymbol{P}_{T-1}\cdots\boldsymbol{P}_1\boldsymbol{P}_0\boldsymbol{p}_0 = \boldsymbol{H}\boldsymbol{p}_0. \tag{34}$$

The explicit form of $\boldsymbol{H}^\dagger$ is thus given as

$$\boldsymbol{H}^\dagger = \boldsymbol{P}_0^\dagger\boldsymbol{P}_1^\dagger\cdots\boldsymbol{P}_{T-1}^\dagger\boldsymbol{S}^\dagger. \tag{35}$$

C. Implementation of the forward and adjoint operators

For the implementation of the forward and adjoint operators, we use the open-source C++ finite-element library deal.ii [29]. Our choice of this library is motivated by its support for parallelization using multiple threads and multiple processors. This capability is crucial for the whole-brain simulation, where a massive number of nodes are required.

Similar to the forward propagation, the action of the discrete adjoint operator, $\boldsymbol{p}_{\text{adj}} = \boldsymbol{H}^\dagger\hat{\boldsymbol{p}}$, can be explicitly decomposed into recursive backward steps as

$$\boldsymbol{m}^{T-1} = \boldsymbol{\Theta}^\dagger\hat{\boldsymbol{p}}^{T-1} \tag{36}$$

$$\boldsymbol{m}^{i-1} = \boldsymbol{\Theta}^\dagger\hat{\boldsymbol{p}}^{i-1} + (\boldsymbol{W}^{-1}\boldsymbol{Q})^\dagger\boldsymbol{m}^i, i = 1, 2, \cdots, T-1, \tag{37}$$

$$\boldsymbol{p}_{\text{adj}} = \boldsymbol{\tau}^\dagger\boldsymbol{m}^0. \tag{38}$$

The update in (36) can be written in terms of quantities similar to (20-22)

$$\ddot{\boldsymbol{x}}^{i-1} = \boldsymbol{M}^\dagger(\boldsymbol{V}^\dagger\boldsymbol{x}_z^i) - \ddot{\boldsymbol{x}}^i, \tag{39}$$

$$\dot{\boldsymbol{x}}^{i-1} = -(\boldsymbol{C} + \Delta t\boldsymbol{K})^\dagger(\boldsymbol{V}^\dagger\boldsymbol{x}_z^i) + \dot{\boldsymbol{x}}^i + \Delta t\boldsymbol{x}^i, \tag{40}$$

$$\boldsymbol{x}^{i-1} = \boldsymbol{x}^i - \boldsymbol{K}^\dagger(\boldsymbol{V}^\dagger\boldsymbol{x}_z^i) + \sum_{l=1}^{L}\boldsymbol{R}_l^\dagger\hat{\boldsymbol{p}}_l^{i-1}, \tag{41}$$

where

$$\boldsymbol{V} = \left(\boldsymbol{M} + \frac{\Delta t}{2}\boldsymbol{I} + \frac{\Delta t^2}{4}\boldsymbol{I}\right)^{-1}, \tag{42}$$

$$\boldsymbol{x}_z^i = \ddot{\boldsymbol{x}}^i + \frac{\Delta t}{2}\dot{\boldsymbol{x}}^i + \frac{\Delta t^2}{4}\boldsymbol{x}^i \tag{43}$$

are the auxiliary matrix and vector, respectively.

Only one linear algebraic equation of size N needs to be solved within each forward and backward time step. The detailed solution procedure is explained in Appendix B.

D. Image reconstruction using the forward and adjoint operators

For PACT transcranial image reconstruction, the goal is to estimate $\boldsymbol{p}_0$ given the measured photoacoustic data $\boldsymbol{p}_{\text{m}}$ and the forward operator $\boldsymbol{H}$ in (34). Since we have obtained the numerically matched adjoint operator $\boldsymbol{H}^\dagger$, we can directly apply it as a reconstruction operator to the measured data [10], [11]. This results in $\boldsymbol{p}_{\text{adj}} = \boldsymbol{H}^\dagger\boldsymbol{p}_{\text{m}}$, effectively generating a reconstructed image. Furthermore, we can integrate the forward-adjoint operator pair into an iterative reconstruction algorithm by solving the optimization problem:

$$\boldsymbol{p}_0^* = \underset{\boldsymbol{p}_0\geq 0}{\text{argmin}}\|\boldsymbol{p}_{\text{m}} - \boldsymbol{H}\boldsymbol{p}_0\|_2^2 + \gamma\text{R}(\boldsymbol{p}_0)\,. \tag{44}$$

The first term on the right-hand side represents the data fidelity term corresponding to a least squares functional. $\text{R}(\boldsymbol{p}_0)$ denotes a regularization term reflecting prior knowledge on $\boldsymbol{p}_{\text{opt}}$, while γ serves as the regularization parameter controlling the regularization weight. Note that explicitly computing the action of the adjoint operator is essential for calculating the gradient of the data fidelity term.

III. Results

In this section, we demonstrate our proposed forward-adjoint pair through both numerical simulations and phantom experiments. For all of the results below, the meshes are generated by the widely-used commercial FEM software, COMSOL multiphysics [30], and subsequently imported into our customized FEM solver. For spatial discretization, we use the second-order Lagrange finite elements.

A. Validation of the FEM forward and adjoint operator

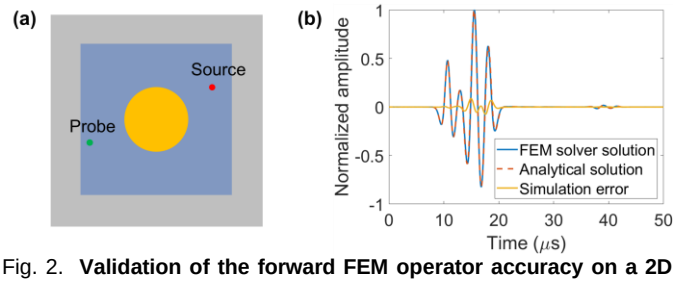

Fig. 2. **Validation of the forward FEM operator accuracy on a 2D scattering problem.** (a) Illustration of the validation setup. An infinite line source generates an acoustic field that is subsequently scattered by an elastic cylinder. (b) Comparison between the analytical solution and the numerical solution obtained from our customized solver at the receiver position.

We validate the accuracy of our 2D FEM forward operator by comparing our FEM simulation result with the analytical solution provided in [31] for the scattering of cylindrical acoustic waves by an elastic cylinder, as shown in Fig. 2(a). The setup involves a lossless fluid domain (blue), wherein an infinite Ricker wavelet line source (red) with a peak frequency of $f_c = \frac{1}{3}$ MHz generates an acoustic field [32]. The acoustic field interacts with a lossy elastic medium (yellow), and the resulting scattered field is recorded by a point probe (green). The acoustic properties of the medium are defined as follows: $c_f = 1500$ m/s, $\rho_f = 1000$ kg/m^3, $c_p = 3000$ m/s , $c_s = 1500$ m/s, $\alpha = 0.75$/μs, $\rho_s = 1850$ kg/m^3. For our FEM simulation,

we consider a computational region of size 30 mm × 30 mm, with a 4.5 mm-thick PML applied in all directions to minimize boundary reflections. A spatial discretization of 5 elements per wavelength (EPW, number of elements per minimum wavelength within the frequency band) and a time step of $\Delta t =$ 20 ns are employed. Fig. 2(b) shows the excellent agreement between the solution obtained from our customized FEM forward solver and the analytical solution. To assess the accuracy, we calculate the L2 relative error with respect to the analytical solution at the sensor position. The L2 relative error is defined as

$$\text{L2 relative error} = \sqrt{\frac{\|\boldsymbol{p}_a - \boldsymbol{p}_s\|_2^2}{\|\boldsymbol{p}_a\|_2^2}} \tag{45}$$

where $\boldsymbol{p_a}$ is the analytical solution, and $\boldsymbol{p_s}$ is the solution obtained from a simulation. The relative L2 error amounts to only 10% over the entire waveform in a duration of 50 µs. This validation confirms the high accuracy of our FEM forward operator implementation.

We also validate the implementation of the adjoint operator. Although the adjoint operator conceptually corresponds to the transpose of the forward operator, the matrices $\boldsymbol{H}^\dagger$ and $\boldsymbol{H}$ are prohibitively large to compute in a single step, making direct verification difficult. Therefore, we conduct validation using a dot-product test, which is a well-established routine for assessing the numerical adjointness between the forward and adjoint operators [33]. This test involves verifying the identity of the inner product $\langle \boldsymbol{Hp}, \hat{\boldsymbol{p}} \rangle = \langle \boldsymbol{H}^\dagger \hat{\boldsymbol{p}}, \boldsymbol{p} \rangle$, which arises from the associative property of linear algebra. Our implementation demonstrates agreement between the left and right sides of the equation up to the machine precision (i.e., 15 digits), thus confirming the accuracy of our implementation.

B. Comparative feasibility study with FDTD

To assess the feasibility of FEM for time-dependent problems, we conduct a comparative analysis of our FEM solver with the widely adopted FDTD method. Our analysis contains two aspects: computation time and image reconstruction quality, which are evaluated on simulated datasets. The FDTD forward-adjoint pair in this study is implemented according to [14] using the Python FDTD library Devito [34].

1) Comparison of computation time for the forward operator

We compare the computation time required to achieve the desired level of accuracy for the two methods. Note that the disparity in computation time is not solely influenced by the numerical schemes (FEM vs FDTD), but also by the details of the implementation. Thus, the primary objective of this comparison is not to directly evaluate the computation speed, but to offer insights into the practical usability of FEM.

To conduct this investigation, we apply both algorithms to the same 2D scattering problem as described in Section III. A. We allow the spatiotemporal step sizes of both algorithms to vary, ensuring that each method operates with its maximum step size. We measure the computational time by conducting three simulation runs for each configuration on a single thread of an Intel Xeon E5 processor.

Table I presents three examples where the solution accuracy is comparable for both methods. Points per wavelength (PPW), which refers to the number of grid points per minimum wavelength within the frequency band, describes the size of the FDTD mesh. Meanwhile, DOFs represent the total number of unknowns in the resulting formulation. At 3 EPW for FEM and 32 PPW for FDTD, respectively, the two approaches exhibit similar performance in terms of time and accuracy. However, the FEM requires significantly lower mesh density and about 20 times fewer DOFs. This advantage is attributed to the adaptability of unstructured mesh to irregular geometries and the domain-decomposed formulation. Notably, as the desired accuracy level increases, the FEM approach becomes asymptotically faster due to the more efficient representation of the geometry with a coarser mesh (Fig. 3). This comparison demonstrates that our FEM approach is computationally feasible in 2D for time-dependent problems.

Method	Similar L2 relative error	PPW/ EPW	# DOFs	Averaged computation time (s)
FEM	13.54%	3	31934	9.62
FDTD	13.04%	32	592900	12.70
FEM	9.59%	5	88306	40.25
FDTD	9.02%	64	2365444	120.94
FEM	3.25%	7	171374	153.39
FDTD	3.39%	128	9449476	1084.37

TABLE I. **Comparison of FEM with FDTD on the accuracy and computation time for forward simulation.** Three examples of comparable solution accuracies for both methods are presented.

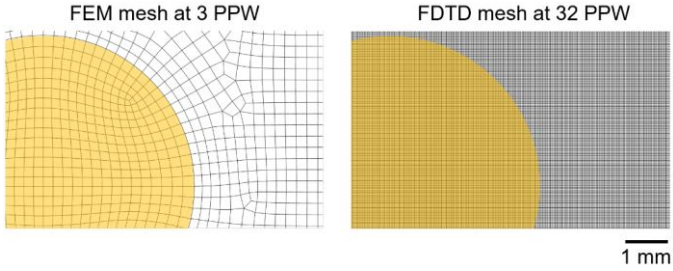

Fig. 3. **Computation mesh used in the FEM and FDTD simulations to achieve a similar solution accuracy.** Yellow regions denote the geometry of the solid object.

2) Comparison of iterative transcranial image reconstruction of simulated data

The value of the forward-adjoint pair for transcranial PACT lies in its ability to seamlessly integrate with various iterative optimization frameworks. In the preceding sections, we have established the validity and feasibility of our developed forward and adjoint operators. Here, our focus shifts toward the ultimate goal of image reconstruction, for which these operators are devised. We present a comparative analysis of the reconstructed images using both the FEM and FDTD approaches. As the direct adjoint image may exhibit variations based on the specific operator formulation, our comparison is centered on the iteratively reconstructed images of simulated noiseless pressure measurements. Once again, we emphasize that this comparison serves solely to demonstrate the feasibility of our FEM method

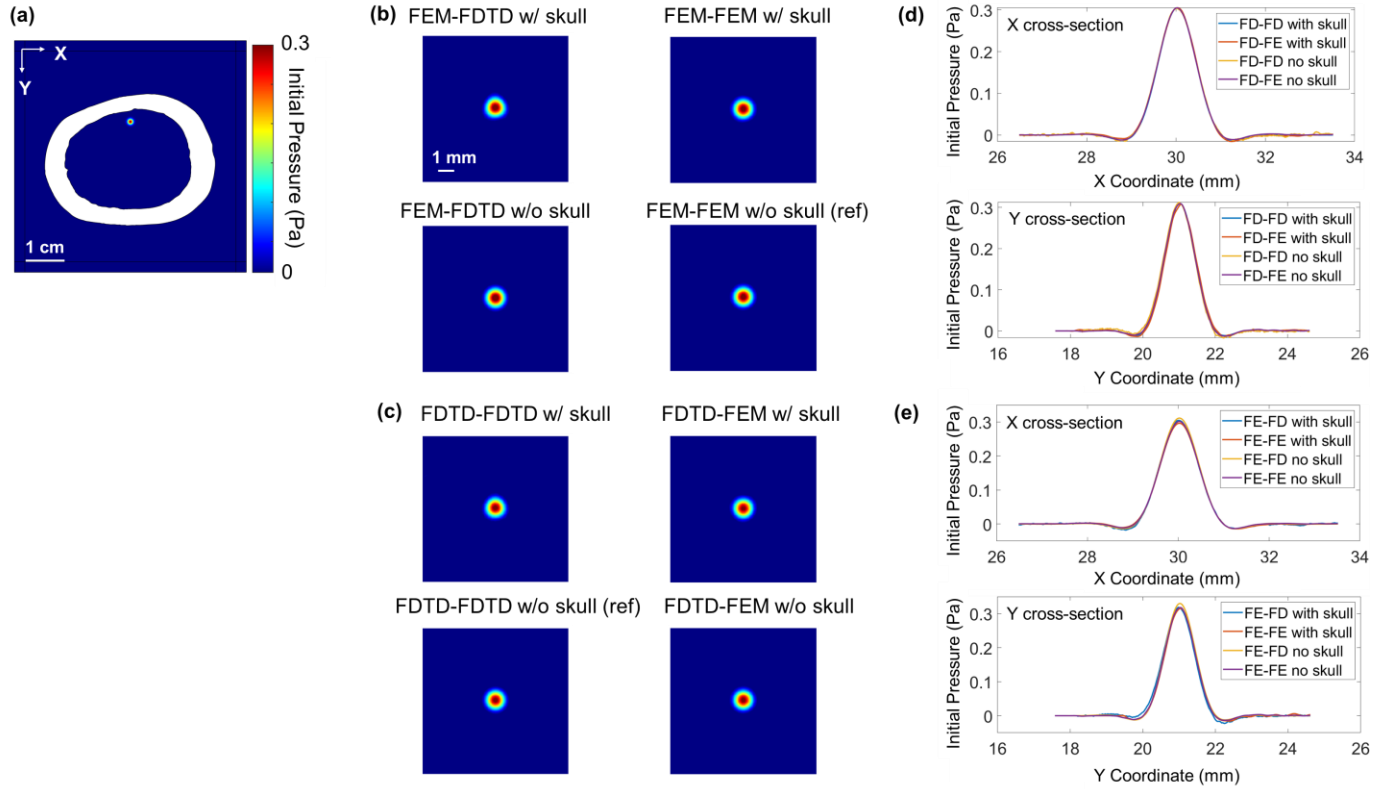

Fig. 4. **Iterative image reconstruction comparison with FDTD solver.** (a) Initial pressure distribution and the demagnified skull geometry. (b, c) PSFs reconstructed using either FEM or FDTD from the (b) FEM and (c) FDTD forward data. The labels use an A-B format where A represents the discretization method (FEM or FDTD) used to generate the data and B denotes the discretization method used for the inversion. (d, e) PSFs along (d) X and (e) Y across the center of the point target in (b) and (c). FE, FEM method; FD, FDTD method; w/ skull: forward data generated with skull present in simulation; w/o skull: forward data generated without skull present in simulation.

	FWHM X (mm)	FWHM Y (mm)	Background STD	SSIM	PSNR
FEM-FDTD w/skull	**0.950**	0.962	9.52e-4	0.9664	51.77
FEM-FEM w/skull	0.959	**0.953**	**8.54e-4**	**0.9836**	**56.30**
FEM-FDTD w/o skull	0.946	0.954	1.38e-5	0.9917	53.37
FEM-FEM w/o skull (ref)	0.951	0.953	6.63e-4		
FDTD-FDTD w/skull	0.955	0.948	1.20e-3	0.9787	53.30
FDTD-FEM w/skull	**0.959**	**0.960**	**9.35e-4**	**0.9860**	**56.47**
FDTD-FDTD w/o skull (ref)	0.964	0.963	1.04e-4		
FDTD-FEM w/o skull	0.955	0.953	6.79e-4	0.9919	59.60

TABLE II. **Quantitative comparison of reconstructed images.** The quantification error is assumed to be half of the interpolated grid size, in this case, 0.012 mm.

and should not be construed as a rigorous evaluation of the superiority of either method.

Fig. 4(a) depicts the simulation setup for transcranial PACT. The setup involves a skull-shaped elastic medium immersed in a fluid medium. We derive the skull boundaries from the segmentation of a skull slice obtained using X-ray computed tomography (CT) and then apply a four times demagnification to reduce computation time. The inner and outer boundaries are demagnified with slightly different ratios to maintain an approximate skull thickness of 6 mm for sufficient acoustic aberration [1]. The computational region spans 60 mm × 60 mm and incorporates a 5 mm PML layer in all directions. The material properties remain consistent with those described in Subsection III. B. 1. We low-pass filter the recorded waveform up to 2 MHz prior to inversion to simulate the limited bandwidth of transducers. To assess the reconstruction quality of the inversion methods, we analyze the point spread function (PSF). For this purpose, we select an object size smaller than the maximum supported wavelength. Specifically, we use a 2D Gaussian with a full width at half maximum (FWHM) of 0.4 mm for the initial pressure distribution.

The inversion process is often hindered by a significant pitfall known as the inverse crime, where the utilization of the same model to generate and invert the synthetic data results in [35]. While we do not commit the inverse crime here, we prevent bias towards any method by obtaining the measured data $\hat{\boldsymbol{p}}$ using an exceedingly high-resolution grid for both the FEM and FDTD methods. The FEM forward data is generated by COMSOL with a 10-EPW mesh, and the FDTD forward data is generated by Devito with a 128-PPW mesh. Subsequently, we reconstruct images from these accurate forward data using

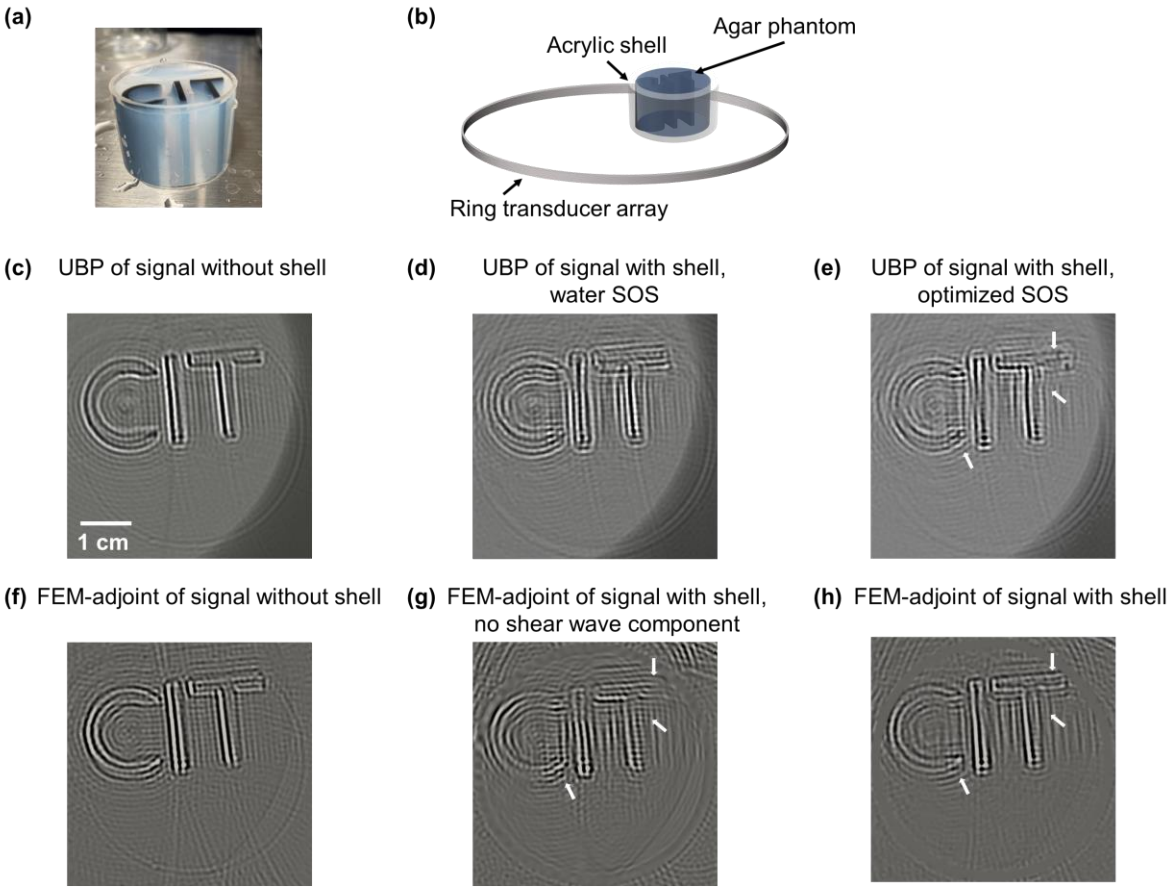

Fig. 5. **Phantom adjoint reconstruction using the customized FEM forward-operator pair.** (a) Photograph of the z-invariant agar phantom contained in the acrylic shell. (b) Schematic of the experimental setup. (c) UBP reconstructed images of the phantom without acrylic shell. UBP reconstructed images of the phantom with shell using water SOS (d) and manually optimized SOS (e). (f) FEM adjoint of the aberration-free data. FEM adjoint of the aberrated data with fluid inhomogeneity modeling (g) and solid inhomogeneity modeling (h).

practical grid sizes. Specifically, we employ meshes of 3 EPW for FEM and 32 PPW for FDTD, as these configurations yield comparable results in Table I. For the iterative inversion process, we adopt the accelerated gradient descent method [36] for both frameworks. The objective function is the L2 norm between the measured and predicted data, and the iteration terminates when it stops decreasing. The reference image is obtained by reconstructing measurements without the presence of a skull using the same inversion algorithm.

Figs. 4(b–e) display reconstructed PSF and their cross-sections at the center using forward data generated by COMSOL and FDTD. Upon visual inspection, there are negligible differences observed among the different reconstructions.

To quantitatively compare the reconstructions, we interpolate the FEM solution to the same Cartesian grid prior to further analysis. We present the results in Table II, which includes metrics such as FWHM, background standard deviation (STD), structural similarity index (SSIM), and peak signal-to-noise ratio (PSNR). The reconstructions that exhibit better performance in the presence of a skull are highlighted in bold. While both schemes exhibit comparable overall performance, FEM reconstructions yield better results in terms of background STD, SSIM, and PSNR. This case study demonstrates the viability and feasibility of our FEM method for accurate iterative image reconstruction.

C. Phantom result

Having established the validity and feasibility of our method, we proceed to apply the FEM adjoint to the experiment data. As illustrated in Fig. 5(a), our phantom consists of a z-invariant, CIT-shaped absorbing structure (made of 4% agar mixed with black ink) embedded in transparent agar. These structures are enclosed within a cylindrical acrylic shell, which introduces acoustic heterogeneities and solid-induced mode conversions. The cylinder has inner and outer diameters of 4.2 mm and 4.5 mm, respectively, and a height of 3 cm. We set the material properties of the acrylic shell as $c_s = 1400$ m/s, $c_p = 2746$ m/s, $\alpha = 0.1$/μs, and $\rho_s = 1178.2$ kg/m^3. We assume the water and agar phantom to be lossless and homogeneous, with properties of $c_f = 1487$ m/s and $\rho_f = 1000$ kg/m^3.

As depicted in Fig. 5(b), our experimental data are acquired from a 512-element unfocused full-ring transducer array, with a central frequency of 2.25 MHz and a one-way bandwidth of 95% [35]. To simulate the frequency and bandwidth of a brain imaging system [1], we low-pass filter the measured data to 1.5 MHz. The aberration-free data, acquired without the acrylic shell, serve as the reference. Since both the absorbing target and the detection array are approximately z-invariant, we employ the 2D FEM model for the cylindrical wave propagation in this experiment. The computational region spans 23 cm × 23 cm, with a grid size of 0.3 mm.

The reconstructed images are shown in Fig 5(c-g). Fig. 5(c) displays the reference image reconstructed with the aberration-free data using the typical universal back projection (UBP) method [36]. Fig. 5(d) and (e) show the original and manually optimized single speed-of-sound (SOS) UBP reconstruction of the aberrated data, respectively. In Fig. 5(f–h), we demonstrate the direct FEM adjoint image of the aberration-free data, aberrated data with fluid inhomogeneity modeling, and aberrated data with solid inhomogeneity modeling, respectively. Since our model inherently accounts for the acoustic attenuation and mode conversions in the elastic medium, the FEM adjoint naturally compensates for the solid-induced aberrations in image reconstruction. As a result, the FEM adjoint image notably recovers several original features compared with the optimized UBP image, as indicated by the white arrows.

IV. Conclusion and Discussion

In this paper, we present a novel discrete forward-adjoint operator pair based on the finite element method (FEM) for transcranial PACT. Instead of using a unified elastic wave equation throughout the whole domain, we divide the simulation region into fluid and solid regions, and explicitly model the solid-liquid interfaces with coupled elastic-acoustic equations. This FEM discretization and domain-decomposed formulation reduces the number of unknowns by 20 times in the discretized linear system, which has the potential for more efficient computation while maintaining accuracy. We validate the framework through numerical simulations and comparison with a well-established FDTD method. We also demonstrate the capability of the adjoint operator to effectively mitigate solid-induced aberrations in experimental phantom data.

By utilizing unstructured meshes that adapt to the skull geometry, our finite element discretization significantly reduces computational resources for spatial discretization. However, the nearly 20-fold reduction in DOFs in our FEM approach results not only from the accommodation of a coarser mesh but, more importantly, from the explicit decomposed modeling of the fluid and solid domains. This approach enables us to assign only 1 DOF per node in the fluid domain, 2 DOFs in the solid domain, and 3 DOFs in the PML domain. In contrast, the unified elastic wave equation proposed in [14] imposes 13 DOFs per spatial node encompassing all domains. The advantage of DOF reduction becomes more pronounced in 3D simulations, considering that most of the simulation region will be fluid, which needs only 1 DOF per node. In comparison, FDTD requires 27 DOFs per node uniformly throughout the entire region, as elaborated in the appendix of [14].

While our paper focuses on the 2D full-wave transcranial PACT reconstruction using FEM, extending the method to 3D simulations presents a significant computational challenge. The challenge arises from the exponential growth of the large sparse system of equations generated during 3D finite element assembly, which necessitates computational acceleration strategies. To address this problem, iterative solvers [37], [38] should be used for solving the sparse system, as direct solvers [39], [40] become infeasible in the 3D simulations. Additionally, we suggest implementing the mass lumping technique for explicit time stepping [17] to simplify system matrix inversion. Finally, parallelization techniques such as multithreading, GPU acceleration, and multiprocessor utilization can be implemented to further enhance the algorithm's efficiency.

We have observed applications employing the spectral element method (SEM) for transcranial ultrasound imaging [41], [42]. SEM, as a variant of Galerkin-based FEM using Lagrange basis functions and Gauss-Lobatto-Legendre (GLL) quadrature rules, offers the advantage of generating a diagonal mass matrix, thus significantly alleviating computational demands. While SEM may encounter challenges in modeling complex geometries such as the skull [43-44], it presents an interesting direction to explore for computational efficiency in the context of transcranial PACT. Our theoretical derivations for the forward-adjoint pair in transcranial PACT remain the same under the SEM framework, requiring only the substitution of SEM-specific basis functions and quadrature rules.

Although our approach offers an accurate and efficient representation of the physics of transcranial PACT, it requires precise prior knowledge of the spatial distribution of SOS within the imaging domain. In practical implementation, we can infer the geometry and acoustic properties of the skull from adjunct CT scans, a method commonly employed to estimate the heterogeneities in density, absorption, and SOS of the skull [45]–[47]. The position of the skull during transcranial imaging can be derived from a co-registration of the CT image and an adjunct ultrasound skull boundary measurement. The remaining mismatch in the model can be further alleviated through the adoption of a joint reconstruction framework for both the skull's properties and the initial pressure distribution [48]–[51].

Appendix A: Definition of Assembled FEM Matrices

We denote the i^{th} shape function in the solid and PML domains as $\boldsymbol{\psi_i}$, the i^{th} shape function in the fluid domain as ϕ_i. The finite elements in the fluid, solid, and PML domain are represented by $\Omega_{e,f}$, $\Omega_{e,s}$, and $\Omega_{e,PML}$. The entries in the element matrices are defined as

$$M_{ij}^{u} = \int_{\Omega_{e,s}} \boldsymbol{\rho_s} \boldsymbol{\psi}_j \cdot \boldsymbol{\psi}_i \, dS \,, \tag{46}$$

$$K_{ij}^{u} = \int_{\Omega_{e,s}} \left(\lambda(\nabla \cdot \boldsymbol{\psi}_i) \cdot \left(\nabla \cdot \boldsymbol{\psi}_j\right)\right) dS + \int_{\Omega_{e,s}} \left((\mu \nabla \boldsymbol{\psi}_i) \cdot \left(\nabla \boldsymbol{\psi}_j + (\nabla \boldsymbol{\psi}_j)^{\dagger}\right)\right) dS \,, \tag{47}$$

$$C_{ij}^{u} = \int_{\Omega_{e,s}} \alpha \boldsymbol{\psi}_i \cdot \boldsymbol{\psi}_j dS \,, \tag{48}$$

$$R_{ij}^{u} = -\oint_{\partial\Omega_{e,s}} \boldsymbol{\psi}_i \cdot (\phi_j \boldsymbol{n_j}) \, dl, \tag{49}$$

$$M_{ij}^{p} = \int_{\Omega_{e,f}} \phi_i \cdot \frac{1}{c_f^2} \phi_j dS \,, \tag{50}$$

$$K_{ij}^{p} = \int_{\Omega_{e,f}} \left(\nabla\phi_i \cdot \nabla\phi_j + \phi_i \cdot \kappa\phi_j\right) dS\,, \tag{51}$$

$$C_{ij}^{p} = \int_{\Omega_{e,f}} \phi_i \cdot \chi\phi_j dS\,, \tag{52}$$

$$R_{ij}^{p} = \oint_{\partial\Omega_{e,s}} \phi_i \cdot \left(\rho_f \boldsymbol{n_j} \cdot \boldsymbol{\psi}_j\right) dl, \tag{53}$$

$$E_{ij} = \int_{\Omega_{e,PML}} \nabla\phi_i \cdot \boldsymbol{\psi}_j dS\,, \tag{54}$$

$$F_{ij} = \int_{\Omega_{e,PML}} \boldsymbol{\psi}_i \cdot \boldsymbol{B}\nabla\phi_j dS\,, \tag{55}$$

$$C_{ij}^{w} = \int_{\Omega_{e,PML}} \boldsymbol{\psi}_i \cdot \boldsymbol{\psi}_j dS\,, \tag{56}$$

$$K_{ij}^{w} = \int_{\Omega_{e,PML}} \boldsymbol{\psi}_i \cdot \boldsymbol{A}\boldsymbol{\psi}_j dS\,. \tag{57}$$

APPENDIX B: SOLUTION PROCEDURE IN THE FORWARD AND ADJOINT STEP

The forward stepping in (20–22) can be solved using the following procedure according to [28]

Solve for predictors:

$$\tilde{\boldsymbol{x}}^{i+1} = \boldsymbol{x}^{i} + \Delta t\dot{\boldsymbol{x}}^{i} + \frac{1}{4}\Delta t^2\ddot{\boldsymbol{x}}^{i}, \tag{58}$$

$$\tilde{\dot{\boldsymbol{x}}}^{i+1} = \dot{\boldsymbol{x}}^{i} + \frac{1}{2}\Delta t\ddot{\boldsymbol{x}}^{i}. \tag{59}$$

Solve the linear equation:

$$\left(\boldsymbol{M} + \frac{1}{2}\Delta t\boldsymbol{C} + \frac{1}{4}\Delta t^2\boldsymbol{K}\right)\ddot{\boldsymbol{x}}^{i+1} = -\boldsymbol{C}\tilde{\dot{\boldsymbol{x}}}^{i+1} - \boldsymbol{K}\tilde{\boldsymbol{x}}^{i+1}. \tag{60}$$

Update state variables:

$$\boldsymbol{x}^{i+1} = \tilde{\boldsymbol{x}}^{i+1} + \frac{1}{4}\Delta t^2\ddot{\boldsymbol{x}}^{i+1}, \tag{61}$$

$$\dot{\boldsymbol{x}}^{i+1} = \tilde{\dot{\boldsymbol{x}}}^{i+1} + \frac{1}{2}\Delta t\ddot{\boldsymbol{x}}^{i+1}. \tag{62}$$

The adjoint stepping can be calculated using the following procedure

Calculate auxiliary vector

$$\boldsymbol{x}_z^{i} = \ddot{\boldsymbol{x}}^{i} + \frac{\Delta t}{2}\dot{\boldsymbol{x}}^{i} + \frac{\Delta t^2}{4}\boldsymbol{x}^{i}. \tag{63}$$

Let $\boldsymbol{V}^{\dagger}\boldsymbol{x}_z^{i} = \boldsymbol{y}_z^{i}$, we solve $\boldsymbol{y}^{i}$ with this linear equation

$$\left(\boldsymbol{M} + \frac{\Delta t}{2}\boldsymbol{I} + \frac{\Delta t^2}{4}\boldsymbol{I}\right)^{\dagger}\boldsymbol{y}_z^{i} = \boldsymbol{x}_z^{i}. \tag{64}$$

Finally, we update the variables through (36–38), which we rewrite here as

$$\ddot{\boldsymbol{x}}^{i-1} = \boldsymbol{M}^{\dagger}\boldsymbol{y}_z^{i} - \ddot{\boldsymbol{x}}^{i}, \tag{65}$$

$$\dot{\boldsymbol{x}}^{i-1} = -(\boldsymbol{C} + \Delta t\boldsymbol{K})^{\dagger}\boldsymbol{y}_z^{i} + \dot{\boldsymbol{x}}^{i} + \Delta t\boldsymbol{x}^{i}, \tag{66}$$

$$\boldsymbol{x}^{i-1} = \boldsymbol{x}^{i} - \boldsymbol{K}^{\dagger}\boldsymbol{y}_z^{i} + \sum_{l=1}^{L}\boldsymbol{R}_l^{\dagger}\hat{\boldsymbol{p}}_l^{i-1}\,. \tag{67}$$

The matrix $\left(\boldsymbol{M} + \frac{\Delta t}{2}\boldsymbol{I} + \frac{\Delta t^2}{4}\boldsymbol{I}\right)$ requires assembly only once prior to conducting the forward and adjoint calculations. Within each forward and adjoint step, only one linear equation, i.e., (58) and (62) need to be solved.

ACKNOWLEDGMENT

L.V.W. has a financial interest in Microphotoacoustics, Inc., CalPACT, LLC, and Union Photoacoustic Technologies, Ltd., which, however, did not support this work.

REFERENCES

[1] S. Na and L. V. Wang, "Photoacoustic computed tomography for functional human brain imaging [Invited]," *Biomed. Opt. Express*, vol. 12, no. 7, pp. 4056-4083, 2021.

[2] B. Liang, S. Wang, F. Shen, Q. H. Liu, Y. Gong, and J. Yao, "Acoustic impact of the human skull on transcranial photoacoustic imaging," *Biomedical optics express*, vol. 12, no. 3, pp. 1512–1528, 2021.

[3] S. Na et al., "Massively parallel functional photoacoustic computed tomography of the human brain," *Nat Biomed Eng*, no.5, pp. 584-592, 2021.

[4] B. Liang, W. Liu, Q. Zhan, M. Li, M. Zhuang, Q.H. Liu, J. Yao, "Impacts of the murine skull on high-frequency transcranial photoacoustic brain imaging," *J Biophotonics*, vol. 12, no. 7, p. e201800466, 2019.

[5] R.W. Schoonover, L.V. Wang, M.A. Anastasio, "Numerical investigation of the effects of shear waves in transcranial photoacoustic tomography with a planar geometry," *J Biomed Opt*, vol. 17, no. 6, p. 061215, 2012.

[6] S. Na, X. Yuan, L. Lin, J. Isla, D. Garrett, L.V. Wang, "Transcranial photoacoustic computed tomography based on a layered back-projection method," *Photoacoustics*, p. 100213, 2020.

[7] C. Huang *et al*, "Aberration correction for transcranial photoacoustic tomography of primates employing adjunct image data," *J Biomed Opt,* vol. 17, no. 6, p. 066016, 2012.

[8] B.E. Treeby, E.Z. Zhang, B.T. Cox, "Photoacoustic tomography in absorbing acoustic media using time reversal," *Inverse Problems*, vol. 26, no. 11, pp. 115003, 2010.

[9] Y. Hristova, P. Kuchment, L. Nguyen, "Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media," *Inverse Problems*, vol. 24, no. 5, p. 055006. 2008.

[10] J. Poudel, S. Na, L. V. Wang, and M. A. Anastasio, "Iterative image reconstruction in transcranial photoacoustic tomography based on the elastic wave equation," *Phys. Med. Biol.,* vol. 65, no. 5, p. 055009, 2020.

[11] C. Huang, K. Wang, L. Nie, L. V. Wang, and M. A. Anastasio, "Full-Wave Iterative Image Reconstruction in Photoacoustic Tomography With Acoustically Inhomogeneous Media," *IEEE Transactions on Medical Imaging*, vol. 32, no. 6, pp. 1097–1110, 2013.

[12] J. Zhu *et al*., "Mitigating the limited view problem in photoacoustic tomography for a planar detection geometry by regularised iterative reconstruction," *IEEE Transactions on Medical Imaging*, pp. 1–1, 2023.

[13] A. Javaherian, S. Holman, "A continuous adjoint for photo-acoustic tomography of the brain," *Inverse Problems*, vol. 34, no. 8, p. 085003, 2018.

[14] K. Mitsuhashi, J. Poudel, T.P. Matthews, A. Garcia-Uribe, L.V. Wang, M.A. Anastasio, "A Forward-Adjoint Operator Pair Based on the Elastic Wave Equation for Use in Transcranial Photoacoustic Computed Tomography," *SIAM J. Imaging Sci*., vol. 10, no. 4, pp. 2022–2048, 2017.

[15] J. Zhang, "Wave propagation across fluid—solid interfaces: a grid method approach," *Geophysical Journal International*, vol. 159, no. 1, pp. 240–252, 2004.

[16] J.M. Carcione, C. Bagaini, J. Ba, E. Wang, A. Vesnaver, "Waves at fluid–solid interfaces: explicit versus implicit formulation of the boundary condition," *Geophysical Journal International*, vol. 215, no. 1, pp. 37–48, 2018.

[17] J. Virieux, H. Calandra, R.-É. Plessix, "A review of the spectral, pseudo‐spectral, finite‐difference and finite‐element modelling techniques for geophysical imaging," *Geophysical Prospecting*, vol. 59, no. Modelling Methods for Geophysical Imaging: Trends and Perspectives, pp. 794–813, 2011.

[18] D. Peter et al., "Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes," *Geophysical Journal International*, vol. 186, no. 2, pp. 721–739, 2011.

[19] G. Pinton, J. F. Aubry, E. Bossy, M. Muller, M. Pernot, and M. Tanter, "Attenuation, scattering, and absorption of ultrasound in the skull bone," *Medical Physics*, vol. 39, no. 1, pp. 299–307, 2012.

[20] M. Sasso, G. Haïat, Y. Yamato, S. Naili, M. Matsukawa, "Frequency Dependence of Ultrasonic Attenuation in Bovine Cortical Bone: An In Vitro Study," *Ultrasound in Medicine & Biology*, vol. 33, no. 12, pp. 1933–1942, 2007.

[21] S. Chaffaï, F. Padilla, G. Berger, P. Laugier, "In vitro measurement of the frequency-dependent attenuation in cancellous bone between 0.2 and 2 MHz," *The Journal of the Acoustical Society of America*, vol. 108, no. 3, pp. 1281–1289, 2000.

[22] R. Ohayon, "Vibrations of Fluid-Structuré Coupled Systems," in *The finite element method in the 1990's: A Book Dedicated to O.C.* Zienkiewicz, E. Oñate, J. Periaux, and A. Samuelsson, Eds., Berlin, Heidelberg: Springer, 1991, pp. 357–366.

[23] B. Kaltenbacher, M. Kaltenbacher, and I. Sim, "A modified and stable version of a perfectly matched layer technique for the 3-d second order wave equation in time domain with an application to aeroacoustics," *Journal of Computational Physics*, vol. 235, pp. 407–422, 2013.

[24] D. Komatitsch, C. Barnes, J. Tromp, "Wave propagation near a fluid-solid interface: A spectral-element approach," *Geophysics*, vol. 65, no. 2, pp. 623–631, 1999.

[25] K. J. Bathe, *Finite Element Procedures*, 2nd Edition. MA: Watertown, 2021.

[26] D. Keyes, A. Ecer, N. Satofuka, P. Fox, J. Periaux, *Parallel Computational Fluid Dynamics '99: Towards Teraflops, Optimization and Novel Formulations.* Elsevier, 2000.

[27] O. C. Zienkiewicz, R. L. Taylor, and J. Z. Zhu, *The finite element method: its basis and fundamentals*, Seventh edition. Amsterdam: Elsevier, Butterworth-Heinemann, 2013.

[28] N. M. Newmark, "A Method of Computation for Structural Dynamics," *Journal of the Engineering Mechanics Division*, vol. 85, no. 3, pp. 67–94, 1959.

[29] D. Arndt *et al*, "The deal.II finite element library: Design, features, and insights," *Computers & Mathematics with Applications*, vol. 81, pp. 407–422, 2021.

[30] "COMSOL Multiphysics® v5.6." COMSOL AB, Stockholm, Sweden, 2020. [Online]. Available: www.comsol.com.

[31] N. V. Larin and L. A. Tolokonnikov, "The scattering of a plane sound wave by an elastic cylinder with a discrete-layered covering," *Journal of Applied Mathematics and Mechanics*, vol. 79, no. 2, pp. 164–169, 2015.

[32] Y. Wang, "Frequencies of the Ricker wavelet," *Geophysics*, vol. 80, no. 2, pp. A31–A37, 2015.

[33] J. Claerbout, R Abma, *Earth Soundings Analysis: Processing Versus Inversion*, London: Blackwell Scientific Publications, 1992.

[34] M. Louboutin et al., "Devito (v3.1.0): an embedded domain-specific language for finite differences and geophysical exploration," *Geoscientific Model Development*, vol. 12, no. 3, pp. 1165–1187, 2019.

[35] A. Wirgin, "The inverse crime." arXiv, Jan. 28, 2004. doi: 10.48550/arXiv.math-ph/0401050.

[36] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, 1st ed. Springer Publishing Company, Incorporated, 2014.

[37] A. Brandt, S. McCormick, and J. Ruge, "Algebraic Multigrid (AMG) for Sparse Matrix Equations," *Sparsity and its Applications*, pp. 257–284, 1985.

[38] M. R. Hestenes and E. Stiefel, "Methods of Conjugate Gradients for Solving Linear Systems", *Journal of research of the National Bureau of Standards*, vol 49, no 6, pp. 409–436, 1952.

[39] T. A. Davis, "Algorithm 832: UMFPACK V4.3---an unsymmetric-pattern multifrontal method," *ACM Trans. Math. Softw.,* vol. 30, no. 2, pp. 196–199, 2004.

[40] P. R. Amestoy, I. S. Duff, and J. Y. L'Excellent, "Multifrontal parallel distributed symmetric and unsymmetric solvers," *Computer Methods in Applied Mechanics and Engineering*, vol. 184, no. 2, pp. 501–520, 2000.

[41] P. Marty, C. Boehm, C. Paverd, M. Rominger and A. Fichtner, "Full-waveform ultrasound modeling of soft tissue-bone interactions using conforming hexahedral meshes." *SPIE Medical Imaging 2022: Physics of Medical Imaging*, vol. 12031, 2022.

[42] P. Marty, C. Boehm and A. Fichtner, "Acoustoelastic full-waveform inversion for transcranial ultrasound computed tomography." *SPIE Medical Imaging 2021: Ultrasonic Imaging and Tomography*, vol. 11602. 2021.

[43] E Chaljub *et al*. "Spectral-element analysis in seismology." *Advances in geophysics*, vol 48, pp. 365-419, 2017.

[44] Y. Liu, J. Teng, H. Lan, X. Si, and X. Ma. "A comparative study of finite element and spectral element methods in seismic wavefield modeling." *Geophysics,* vol 79, no. 2, pp. 91-104, 2014.

[45] F. Marquet *et al*, "Non-invasive transcranial ultrasound therapy based on a 3D CT scan: protocol validation and in vitro results," *Physics in Medicine & Biology*, vol. 54, no. 9, pp. 2597, 2019.

[46] R. M. Jones, A. O. Meaghan, and K. Hynynen, "Transcranial passive acoustic mapping with hemispherical sparse arrays using CT-based skull-specific aberration corrections: a simulation study," *Physics in Medicine & Biology*, vol. 58, no. 14, pp. 4981, 2013.

[47] J. F. Aubry, M. Tanter, M. Pernot, J. L. Thomas, M. Fink, "Experimental demonstration of noninvasive transskull adaptive focusing based on prior computed tomography scans," *The Journal of the Acoustical Society of America*, vol. 113, no. 1, pp. 84-93, 2003.

[48] T. P. Matthews, J. Poudel, L. Li, L. V. Wang, and M. A. Anastasio, "Parameterized joint reconstruction of the initial pressure and sound speed distributions for photoacoustic computed tomography." *SIAM journal on imaging sciences*, vol. 11, no. 2, pp. 1560-1588, 2018.

[49] H. Liu, and U. Gunther, "Determining both sound speed and internal source in thermo-and photo-acoustic tomography." *Inverse Problems*, vol. 31, no. 10, pp. 105005, 2015.

[50] A. Kirsch, S. Otmar, "Simultaneous reconstructions of absorption density and wave speed with photoacoustic measurements." *SIAM Journal on Applied Mathematics*, vol. 72, no. 5, pp. 1508-1523, 2023.

[51] Y. Zhen, Q. Zhang, H. Jiang, "Simultaneous reconstruction of acoustic and optical properties of heterogeneous media by quantitative photoacoustic tomography." *Optics express*, vol. 14, no. 15, pp. 6749-6754, 2006.