Full-waveform inversion of incoherent dynamic traction in a bounded 2D domain of scalar wave motions.

B. P. Guidio[1] and C. Jeong[2]

[1]Postdoctoral Researcher, School of Engineering and Technology, Central Michigan University, Mount Pleasant, MI 48859, USA

[2]Assistant Professor, School of Engineering and Technology, Central Michigan University, Mount Pleasant, MI 48859, USA (corresponding author). Email: jeong1c@cmich.edu

ABSTRACT

This paper presents a full-waveform inversion method for reconstructing the temporal and spatial distribution of unknown, incoherent dynamic traction in a heterogeneous, bounded solid domain from sparse, surficial responses. This work considers SH wave motions in a two-dimensional (2D) domain. The partial-differential-equation (PDE)-constrained optimization framework is employed to search a set of control parameters, by which a misfit between measured responses at sensors on the top surface induced by targeted traction and their computed counterparts induced by estimated traction is minimized. To mitigate the solution multiplicity of the presented inverse problem, we employ the Tikhonov (TN) regularization on the estimated traction function. We present the mathematical modeling and numerical implementation of both optimize-then-discretize (OTD) and discretize-then-optimize (DTO) approaches. The finite element method (FEM) is employed to obtain the numerical solutions of state and adjoint problems. Newton's method is utilized for estimating an optimal step length in combination with the conjugate-gradient scheme, calculating a desired search direction, throughout a minimization process.

Numerical results present that the complexity of a material profile in a domain increases the error between reconstructed traction and its target. Second, the OTD and DTO approaches lead

to the same inversion result. Third, when the sampling rate of the measurement is equal to the timestep for discretizing estimated traction, the ratio of the size of measurement data to the number of the control parameters can be as small as 1:12 in the presented work. Fourth, it is acceptable to tackle the presented inverse modeling of dynamic traction without the TN regularization. Fifth, the inversion performance is more compromised when the noise of a larger level is added to the measurement data, and using the TN regularization does not improve the inversion performance when noise is added to the measurement. Sixth, our minimizer suffers from solution multiplicity less when it identifies dynamic traction of lower frequency content than that of higher frequency content. The wave responses in a computational domain, induced by targeted traction and its reconstructed one, are in excellent agreement with each other. Thus, if the presented dynamic-input inversion algorithm is extended in realistic 3D settings, it could reconstruct seismic input motions in a truncated domain and, then, replay the wave responses in a computational domain.

INTRODUCTION

There is a need for estimating incident seismic wavefields in a soil-structure system from limited seismic measurement data because, by using the identified seismic inputs, engineers can reconstruct (i.e., replay) responses within structures and soils during an earthquake event. There have been two dominant, conventional methods for the purpose mentioned above: the one is deconvolution and the other is the inversion of a seismic source profile at a fault in a very large domain.

The deconvolution algorithm has been used for the identification of an incoming seismic wave signal into a soil column by using vibrational measurement on the ground surface. For instance, Mejia and Dawson (2006) have presented the deconvolution to compute a seismic input signal by using the SHAKE program (Schnabel 1972), which solves the 1D seismic wave propagation problem in a domain of a semi-infinite extent. Recently, there have been studies on the deconvolution of both vertical and horizontal components of surficial measurement data to identify the vertical and horizontal input wave motions (Poul and Zerva 2018a; Poul and Zerva 2018b). We note that, although the deconvolution has been widely used in geotechnical earthquake engineering, it is effective on individual soil columns only when soil stratification is horizontally uniform, and incoming seismic waves vertically propagate. Namely, when the soil property is arbitrarily heterogeneous (not horizontally layered), and incoming seismic waves, consisted of P, S, and/or surface waves, are highly incoherent (not vertically propagating), the deconvolution cannot effectively reconstruct the incoming waves.

On the other hand, there have been studies for inverting for seismic-source parameters (however simple or complicated an adopted seismic source model may be) at a hypocenter. For instance, Akcelik et al. (2002) presented a method to invert for a simplified seismic source time signal in a large 3D domain that includes a source at a fault. This method requires forward and inverse wave simulations of a very large domain. Upon characterizing the seismic source via inversion, then attention is typically turned on how to propagate the motion from the source to the surface, where the real interest is. However, in the large-scale source-inversion problem, the material properties of a large domain could be poorly characterized (on the other hand, the material properties could be better characterized by virtue of active wave source-based geotechnical characterization method (Fathi et al. 2016)).

The limitations of the two methods, mentioned above, necessitate developing an alternative method that can identify arbitrary, incoherent incoming seismic waves in a truncated 2D or 3D domain by using sparse seismic measurement. Such a potential method could serve as an alternative to the deconvolution, while bypassing all the complexities associated with the inversion of the source at the hypocenter and the subsequent propagation steps. Recently, Jeong and Seylabi (2018) presented prototype research that can reconstruct a seismic input signal propagating into a 1D truncated, heterogeneous, undamped solid system by using the partial differential equation (PDE)-constrained optimization method. Lloyd and Jeong (2018) also show that the PDE-constrained optimization can reconstruct the discretized parameters of moving vibrational body forces in both space and time in a 1D heterogeneous, linear, elastic, undamped solid by using the sparse measurement of wave motions. These works were cast into a minimization problem, where a misfit between a measured response(s) at a sensor(s) induced by a targeted wave source profile and a computed wave solution(s) induced by an estimated source profile is minimized, and the

PDE-constrained optimization scheme analytically evaluates the gradient of a misfit with respect to control parameters, which parameterize an estimated dynamic input function. Because of such an analytical nature, its computational efficiency of computing the gradient of a misfit with respect to control parameters does not depend on the number of them. Thus, it can update a large set of control parameters very efficiently. In addition, the PDE-constrained optimization can accommodate regularization to address the solution multiplicity of full-waveform inversion problems, typically caused by the sparsity of measurement, and to stabilize the convergence by penalizing an undesired aspect of an estimated profile while enhancing a selected feature (e.g., smoothness) of targeted profiles.

There had been a wide range of studies on elastodynamic inverse problems—e.g., full-waveform material inversion, inverse material design, full-waveform inverse-scattering, and source inversion—based on the PDE-constrained optimization as shown in the following literature review. Kang and Kallivokas (2010a) and Kang and Kallivokas (2011) examined the numerical algorithms to image the distributions of the scalar-wave speeds in one-dimensional and two-dimensional solids that are surrounded by Perfectly-Matched-Layers (PML), where waves are forced to decay and they are prevented from reflecting off the surrounding boundaries (Kang and Kallivokas 2010b; Kucukcoban and Kallivokas 2011; Kucukcoban and Kallivokas 2013; Fathi et al. 2015b; Poul and Zerva 2018c). Pakravan et al. (2016) devised a new methodology to probe the elastic and attenuating parameters of two-dimensional viscoelastic layered solids. Kallivokas et al. (2013), Fathi et al. (2015a), Fathi et al. (2016), and Kucukcoban et al. (2019) studied algorithms to invert for the Lamé parameters in two-dimensional and three-dimensional PML-truncated solid domains, and these computational studies made significant advancement in geotechnical site characterization using dynamic tests. Tran and McVay (2012) investigated the Gauss-Newton-based full waveform inversion approach to estimate the elastic modulus profile in a two-dimensional domain. Mashayekh et al. (2018) investigated a new methodology to estimate the mechanical properties of layered elastic or viscoelastic media by taking into account the dispersion relation of the layered medium. Tromp et al. (2008) and Zhu et al. (2017) investigated the adjoint-tomography geophysical inversion using the spectral element wave modeling in global- or regional-scale domains by using earthquake waves emitted from a seismic source of a known location and a known signal. Recently, Goh and Kallivokas (2019) investigated a new inverse metamaterial design method by minimizing the distance between the target and the computed group velocity profiles via a dispersion-constrained optimization method, and the method can be used for designing metamaterials, such as user-defined omnidirectional band gaps in an elastic medium. In addition, it had been shown that strong discontinuities within solids, such as the boundaries of voids, can be identified by using inverse modelings. Namely, Guzina et al. (2003) and Jeong et al. (2009) studied inverse scattering algorithms using the PDE-constrained optimization, associated with the boundary element method (BEM) wave solver, taking advantages of the moving boundary concept and the total derivative (Petryk and Mroz 1986). Nguyen-Tuan et al. (2019) also made recent progress in the inverse scattering algorithm, using the PDE-constrained optimization associated with the level-set based extended finite element method (XFEM) solver, so as to identify the geometry of voids in a static-hydro-mechanical system. Both BEM and XFEM wave solvers can model the boundaries of the strong discontinuities and update their geometries without cumbersome remeshing during an inversion process as opposed to a conventional finite element method (FEM) wave solver, which should remesh a domain to update the boundaries' geometries (Jung et al. 2013). On the other hand, Aquino et al. (2019) devised a novel algorithm to detect debonded interfaces (i.e., interface cracks or incomplete weld bonds) in composite solids by using steady-state vibrational tests and a density function that characterizes the bonding at the interfaces in composite solids. Besides, the following studies have investigated the methods to identify dynamic input functions. Hasanov and Baysal (2014) studied an algorithm to detect the time-independent spatial load distributions of a dynamic source on a cantilever beam. Binder et al. (2015) also recover virtual, stationary wave sources at possible locations of structural anomalies using the adjoint equation approach. Walsh et al. (2013) reported the inverse problems for the identification of dynamic sources in acoustics and elastodynamics, employing a DTO approach. That is, the discrete form of the forward wave equation at each time step is imposed into a Lagrangian, and the adjoint equation and the gradient of the Lagrangian with respect to the source parameters are derived in discrete forms. In particular, Walsh et al. (2013) suggested that the DTO approach is more suitable than the OTD counterpart when the nonlinearity is considered in the forward problem because the discrete, linearized forward equation of every time step can be individually side-imposed into the Lagrangian. The PDE-constrained optimization has been also used for identifying optimal, non-moving wave source profiles that can focus wave energy to specific areas in solids (Tadi et al. 1996; Jeong et al. 2010; Jeong et al. 2015; Karve et al. 2015; Karve and Kallivokas 2015; Jeong and Kallivokas 2016; Karve et al. 2016).

Due to the aforementioned robustness and scalability of the PDE-constrained optimization for inverse problems, it is worth continuing to investigate it so as to identify incoherent seismic input motions in a multi-dimensional domain from sparse seismic measurement data. This research reconstructs the spatial and temporal distributions of incoherent dynamic traction on a boundary of a heterogeneous, bounded, undamped solid system of anti-plane motions by using the PDE-constrained optimization method.

Our numerical experiments show that the presented dynamic-input identification approach can successfully identify unknown targeted traction without knowing any information about the target in heterogeneous domains. Our parametric studies investigate the performance of the presented inverse modeling with respect to the complexity of material heterogeneity in a domain, the number of sensors, the regularization intensity factor, the optimization modeling type (i.e., OTD versus DTO), the noise level in measurement data, and the traction signal type (i.e., a high-frequency Ricker wavelet versus a low-frequency realistic seismic signal).

PROBLEM DEFINITION

This study is aimed at reconstructing the spatial and temporal distributions of dynamic traction on a boundary of an undamped solid by using measured wave responses at sparsely-distributed sensors on the top surface of the solid (see Fig. 1). The geometries and the material properties of the solid are assumed to be known in advance.

The governing equation

The governing equation for the SH wave propagation in the undamped solid domain is (we omit to show the spatial and/or temporal dependency of the variables):

$$\nabla \cdot (G\nabla u) - \rho \frac{\partial^2 u}{\partial t^2} = 0, \quad \text{on } \Omega \times \mathsf{J}, \tag{1}$$

where $u = u(x, y, t)$ denotes the displacement field in the anti-plane (z) direction of the wave motion of a solid particle (i.e., SH wave motion); x, y, and t denote horizontal and vertical coordinates and time; $G(x, y)$ and $\rho(x, y)$ denote the shear modulus and the mass density of the solid; Ω denotes the domain, and $\mathsf{J} = (0, T]$ is the time interval of interest. The solid is subject to a traction-free condition on the top surface (Γ_t) and dynamic shear stress on the bottom surface (Γ_b):

$$G\frac{\partial u}{\partial y}(x, 0, t) = 0, \quad 0 \le x \le L, \tag{2}$$

$$G\frac{\partial u}{\partial y}(x, D, t) = F(x, t), \quad 0 \le x \le L, \tag{3}$$

where D is the y-coordinate of Γ_b, and $F(x, t)$ denotes the dynamic shear stress applied on Γ_b. The solid is constrained by fixed boundary conditions on the left (Γ_l) and right (Γ_r) boundaries:

$$u(0, y, t) = 0, \quad D < y < 0, \tag{4}$$

$$u(L, y, t) = 0, \quad D < y < 0. \tag{5}$$

where L is the x-coordinate of Γ_r. The governing wave physics is also subject to zero initial-value conditions:

$$u(x, y, 0) = 0, \tag{6}$$

$$\frac{\partial u}{\partial t}(x, y, 0) = 0. \tag{7}$$

We note that this work considers a 2D bounded domain as a prototype for the seismic-input inversion problem in the multi-dimensional setting. Continuing this work, we will investigate the seismic-input inversion in a 2D/3D unbounded (truncated) domain.

Parameterization of an estimated dynamic traction function

We discretize an estimated dynamic traction function, $F(x,t)$, over space and time as:

$$F(x,t) = \sum_{k=1}^{N_x} \sum_{j=1}^{N_t} \Phi_k(x)\phi_j(t)F_{kj}, \tag{8}$$

where $\Phi_k(x)$ denotes the k-th component of a vector of global basis functions used for the spatial discretization of $F(x,t)$; $\phi_j(t)$ denotes the j-th component of a vector of global basis functions used for the temporal discretization of $F(x,t)$; F_{kj} denotes the discretized value of $F(x,t)$ at each discrete location x_k and time t_j; and N_x and N_t denote the numbers of discretization points over space and time, respectively. Although the shape functions to construct $\Phi_k(x)$ and $\phi_j(t)$ can be of any low order, linear shape functions are used for both of them in the presented inverse modeling. The sizes of the temporal and spatial discretization are set to be, respectively, the time-step size (Δt) of the forward time integration and the element size (Δx) of a mesh on Γ_b for a forward wave solver. The presented inverse modeling is aimed at reconstructing the set of control parameters F_{kj}, of which corresponding wave responses in the domain are consistent with the measurement on Γ_t.

INVERSE MODELING—THE OPTIMIZE-THEN-DISCRETIZE (OTD) APPROACH

This section presents the OTD modeling for identifying the temporal and spatial distributions of unknown traction $F(x,t)$ based on measured wave responses on the top surface of the solid. First, this section presents the mathematical modeling of deriving the first-order optimality conditions in a continuous form. Second, this section shows the discrete forms of the state and adjoint equations and the gradient of the objective functional with respect to the control parameters.

The objective functional

We cast the presented inverse problem into a minimization problem, where we seek the values of control parameters (i.e., F_{kj} in (8) for all k and j) that correspond to a minimum (either global or local one) of an objective functional:

$$\mathcal{L} = \int_0^T \sum_{i=1}^{N_s} (u_{\mathrm{m}i} - u_i)^2 \, \mathrm{d}t + \mathcal{R}^{\mathrm{TN}}, \tag{9}$$

where $u_{\mathrm{m}i}$ denotes the displacement field of the measured wave response at the i-th sensor induced by a targeted $F(x,t)$; u_i denotes the computed counterpart due to an estimated $F(x,t)$, which is constructed by estimated control parameters; and N_s denotes the number of sensors. In this computational study, $u_{\mathrm{m}i}$ is synthetically created by using a forward wave solver with a pseudo target of $F(x,t)$. The first term of (9) is a misfit between $u_{\mathrm{m}i}$ and u_i. We speculate that the misfit functional is quadratic and convex (please see the Appendix III) so that the first-order optimality conditions of the Lagrangian functional (11) would be sufficient for obtaining the inversion solution. However, due to the sparsity of measurement data, we hypothesize that the considered inverse problem would suffer from the solution multiplicity, and the regularization could be effective for improving the convergence of the inversion solution to a targeted profile. To test the hypothesis, we employ the second term of (9), $\mathcal{R}^{\mathrm{TN}}$, which denotes the Tikhonov (TN) regularization term:

$$\mathcal{R}^{\mathrm{TN}} = \frac{R}{2} \int_0^T \int_{\Gamma_\mathrm{b}} \left(\frac{\partial F(x,t)}{\partial x}\right)^2 + \left(\frac{\partial F(x,t)}{\partial t}\right)^2 \mathrm{d}\Gamma \, \mathrm{d}t, \tag{10}$$

where R is the regularization factor, which adjusts the amount of penalty on the derivative of $F(x,t)$. By minimizing the regularization term $\mathcal{R}^{\mathrm{TN}}$ along with the misfit, we attempt to minimize the discontinuity of $F(x,t)$ and smooth it while mitigating the solution multiplicity of the presented inverse problem. It is well known that, when the material inversion is performed, the TN regularization on a material profile overly smooths the discontinuity at the interface of layered media. Thus, the TN regularization is suited for identifying a smooth material profile while the total variation (TV) is used for enhancing the discontinuity of the material profile. Meanwhile, for the seismic input inversion, a typical time signal of a seismic input motion should be a smooth function because the high-frequency content of a discontinuous signal cannot be retained along the propagation path from a seismic source to a near-surface domain due to attenuation. Thus, this work used the TN regularization under a hypothesis that using the TN regularization can improve the uniqueness of the inversion solution while smoothing the estimated traction function. This hypothesis is tested in Example 4 shown in the later section 'Numerical Experiments'.

Lagrangian functional

By imposing the governing equation (1) and the Neumann boundary condition (3) onto the side of the objective functional via Lagrange multipliers, a Lagrangian functional is built as:

$$
\begin{aligned}
\mathcal{A} = & \int_0^T \sum_{i=1}^{N_s} (u_{\mathrm{m}i} - u_i)^2 \,\mathrm{d}t + \mathcal{R}^{\mathrm{TN}} \\
& + \int_0^T \int_\Omega \lambda \left[\nabla \cdot (G \nabla u) - \rho \frac{\partial^2 u}{\partial t^2} \right] \mathrm{d}\Omega \,\mathrm{d}t \\
& + \int_0^T \int_{\Gamma_\mathrm{b}} \lambda_F \left[G \frac{\partial u}{\partial y} - F(x,t) \right] \mathrm{d}\Gamma \,\mathrm{d}t,
\end{aligned} \tag{11}
$$

where $\lambda = \lambda(x, y, t)$ and $\lambda_F = \lambda_F(x, t)$ are the Lagrange multipliers. Note that the boundary condition and initial conditions are implicitly imposed in (11): they are not shown in (11) but used for the derivation of the adjoint and control equations. The first-order optimality conditions of the Lagrangian functional lead to state, adjoint, and control equations. The satisfaction of these equations leads to an optimal solution, corresponding to the minimal value of the objective functional.

The first-order optimality conditions

The first-order optimality conditions of the Lagrangian functional $\mathcal{A}$ require the vanishing variations of $\mathcal{A}$ with respect to the state variable $u(x, y, t)$, the Lagrange variables $\lambda(x, y, t)$ and $\lambda_F(x, t)$, and the control parameter $\xi = F_{kj}$. Such vanishing conditions lead to a triad of state, adjoint, and control equations (Lions 1971):

$$\delta_{\lambda,\lambda_F} \mathcal{A} = 0 : \quad \text{The first condition (state problem),} \tag{12}$$

$$\delta_u \mathcal{A} = 0 : \quad \text{The second condition (adjoint problem),} \tag{13}$$

$$\delta_\xi \mathcal{A} = 0 : \quad \text{The third condition (control problem).} \tag{14}$$

For the first condition, the variation of $\mathcal{A}$ with respect to the Lagrange variables $\lambda(x, y, t)$ and $\lambda_F(x, t)$ vanishes when the state problem—the original governing wave equation (1) and its associated boundary and initial-value conditions—is satisfied. Our inverse modeling procedure automatically satisfies it by numerically solving the state problem for estimated control parameters.

As the second condition, the variation of $\mathcal{A}$ with respect to the state variable $u(x, y, t)$ should vanish. Such a vanishing variational condition leads to the following adjoint equation (see the derivation of the adjoint problem in Appendix):

$$\nabla \cdot (G\nabla\lambda) - \rho\frac{\partial^2\lambda}{\partial t^2} = 2(u_{\mathrm{m}} - u)\sum_{i=1}^{N_{\mathrm{s}}}\Delta(x - x_i, y - y_i), \quad \text{on } \Omega \times [0, T), \tag{15}$$

where $\Delta(x - x_i, y - y_i)$ is the Dirac delta. In the adjoint PDE, the difference between u_{m} and u serves as the time signal of a point wave source at each location of a sensor. It is noteworthy that the strong form of the adjoint PDE is derived from the weak-form like equation (77). Although the strong form of the adjoint PDE is weakly satisfied, the FEM solution of the adjoint PDE fully satisfies the weak-form like equation (77) and, thus, satisfies the second condition of the first-order optimality condition. On the other hand, if the Lagrangian functional is built by imposing the discrete form of the state PDE, the aforementioned issue does not arise because the corresponding discrete form of the adjoint problem fully satisfies the second condition of the first-order optimality condition. To study the latter aspect, this paper also presents the discretize-then-optimize (DTO) counterpart in the later section 'Inverse Modeling—the discretize–then-optimize (DTO) approach', and our numerical experiments tests its inversion performance in Example 2 in the later section 'Numerical Experiments'.

The adjoint PDE is also subject to the following boundary conditions:

$$\begin{aligned} &\lambda(0, y, t) = \lambda(L, y, t) = 0, \quad D < y < 0, \\ &\frac{\partial\lambda}{\partial y}(x, 0, t) = \frac{\partial\lambda}{\partial y}(x, D, t) = 0, \quad 0 \le x \le L, \end{aligned} \tag{16}$$

and the following final-value conditions:

$$\lambda(x,y,T) = 0,$$
$$\frac{\partial \lambda}{\partial t}(x,y,T) = 0. \tag{17}$$

The third condition states that the variation of $\mathcal{A}$ with respect to a scalar-valued control parameter $\xi = F_{kj}$ should vanish. The vanishing variation condition leads to the following control equation (see the derivation of the control problem in Appendix):

$$\begin{aligned}
\delta_\xi \mathcal{A} &= \frac{\partial \mathcal{A}}{\partial \xi} \\
&= -\int_0^T \int_{\Gamma_\mathrm{b}} \lambda \frac{\partial}{\partial \xi}(F(x,t))\, \mathrm{d}\Gamma\, \mathrm{d}t + \frac{\partial \mathcal{R}^{\mathrm{TN}}}{\partial \xi} \\
&= \underbrace{-\int_0^T \int_{\Gamma_\mathrm{b}} \lambda \Phi_k(x)\phi_j(t)\, \mathrm{d}\Gamma\, \mathrm{d}t - R\int_0^T \int_{\Gamma_\mathrm{b}} \left(\frac{\partial^2 F(x,t)}{\partial x^2} + \frac{\partial^2 F(x,t)}{\partial t^2}\right)\Phi_k(x)\phi_j(t)\, \mathrm{d}\Gamma\, \mathrm{d}t}_{\text{the closed form of } \frac{\partial \mathcal{A}}{\partial \xi} = \frac{\partial \mathcal{L}}{\partial \xi}} = 0.
\end{aligned} \tag{18}$$

Note that $\delta_\xi \mathcal{A} = \frac{\partial \mathcal{A}}{\partial \xi}$ is the derivative of $\mathcal{A}$ with respect to a control parameter $\xi = F_{kj}$. Since the side-imposed terms in $\mathcal{A}$ vanish, $\frac{\partial \mathcal{A}}{\partial \xi}$ is equivalent to $\frac{\partial \mathcal{L}}{\partial \xi}$, which constitutes a gradient vector $\nabla_{\boldsymbol{\xi}} \mathcal{L}$, where $\boldsymbol{\xi}$ is a vector of all the control parameters. The control equation (18) implies that $\nabla_{\boldsymbol{\xi}} \mathcal{L}$ at any estimated values of $\boldsymbol{\xi}$ can be evaluated in a semi-analytical manner by using its closed form once the solutions of state and adjoint problems are computed.

Finite element solution of the state problem

To find $u(t) \in U$ for all $v \in \mathcal{V}$, we cast the weak form of the state problem as:

$$\int_\Omega \nabla v \cdot (G \nabla u)\, \mathrm{d}\Omega + \int_\Omega v\rho \frac{\partial^2 u}{\partial t^2}\, \mathrm{d}\Omega = -\int_{\Gamma_\mathrm{b}} vF(x,t)\, \mathrm{d}\Gamma, \tag{19}$$

where v denotes a test function. The function spaces for a scalar-valued u and v are defined as:

$$\begin{aligned} U &= \{u : u \in H^1, \quad u\big|_{\Gamma_l,\Gamma_r} = 0\} \\ \mathcal{V} &= \{v : v \in H^1, \quad v\big|_{\Gamma_l,\Gamma_r} = 0\} \end{aligned} \tag{20}$$

To resolve the weak form numerically, we use the standard finite-element approximation. We approximate the test and trial functions, respectively, as:

$$v(x, y) \simeq \mathbf{v}^T \boldsymbol{\psi}(x, y), \quad u(x, y, t) \simeq \boldsymbol{\psi}(x, y)^T \mathbf{u}(t), \tag{21}$$

where $\boldsymbol{\psi}(x, y)$ denotes a vector of global basis functions constructed by shape functions of each finite element mesh in the domain, and $\mathbf{u}(t)$ denotes a vector of nodal solutions of the state problem. Then, (19) reduces to the following discrete form:

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = \mathbf{F}(t), \tag{22}$$

where $(\ddot{\ })$ denotes the second-order derivative of its subtended variable with respect to t; $\mathbf{M}$ denotes a global mass matrix; $\mathbf{K}$ denotes a global stiffness matrix; $\mathbf{F}(t)$ denotes a global force vector. They are defined as:

$$\begin{aligned} \mathbf{K} &= \int_\Omega G\left(\frac{\partial \boldsymbol{\psi}}{\partial x}\frac{\partial \boldsymbol{\psi}^T}{\partial x} + \frac{\partial \boldsymbol{\psi}}{\partial y}\frac{\partial \boldsymbol{\psi}^T}{\partial y}\right) \mathrm{d}\Omega, \\ \mathbf{M} &= \int_\Omega \rho \boldsymbol{\psi}\boldsymbol{\psi}^T \, \mathrm{d}\Omega, \\ \mathbf{F}(t) &= -\int_{\Gamma_\mathrm{b}} \boldsymbol{\psi} F(x, t) \, \mathrm{d}\Gamma. \end{aligned} \tag{23}$$

Finite element solution of the adjoint problem

To find $\lambda(t) \in \Lambda$ for all $v \in \mathcal{V}$, the weak form of the adjoint equation (15) is obtained as:

$$\int_\Omega \nabla v \cdot (G \nabla \lambda)\, \mathrm{d}\Omega + \int_\Omega v \rho \frac{\partial^2 \lambda}{\partial t^2}\, \mathrm{d}\Omega = -\sum_{i=1}^{N_\mathrm{s}} 2 v(x_i, y_i)(u_\mathrm{m}(x_i, y_i, t) - u(x_i, y_i, t)). \tag{24}$$

The function space for a scalar-valued λ is defined as:

$$\Lambda = \{\lambda : \lambda \in H^1, \quad \lambda\big|_{\Gamma_l, \Gamma_r} = 0\} \tag{25}$$

The test and trial functions are approximated as follows:

$$v(x, y) \simeq \mathbf{v}^T \boldsymbol{\psi}(x, y), \quad \lambda(x, y, t) \simeq \boldsymbol{\psi}(x, y)^T \boldsymbol{\lambda}(t), \tag{26}$$

where $\boldsymbol{\lambda}(t)$ is a vector of the nodal adjoint solution. The weak form of the adjoint problem changes to the following time-dependent discrete form:

$$\mathbf{M}\ddot{\boldsymbol{\lambda}}(t) + \mathbf{K}\boldsymbol{\lambda}(t) = \mathbf{F}_\mathrm{adj}(t), \tag{27}$$

where $\mathbf{F}_\mathrm{adj}(t)$ is defined as:

$$\mathbf{F}_\mathrm{adj}(t) = 2 \sum_{i=1}^{N_\mathrm{s}} (\boldsymbol{\psi}(x_i, y_i) u(x_i, y_i, t) - \boldsymbol{\psi}(x_i, y_i) u_\mathrm{m}(x_i, y_i, t)). \tag{28}$$

We note that the specific forms of the matrices $\mathbf{K}$ and $\mathbf{M}$ in the discrete form of the adjoint problem in (27) are identical to those for the state problem in (23).

Time integration

We solve the time-dependent discrete form of the state problem in (22) by using the implicit Newmark time integration (i.e., average-acceleration scheme). We omit to show the detail of the forward time integration procedure of the state problem.

On the other hand, for the time-dependent discrete form of the adjoint problem in (27), the time integration begins from the final time $t = T$ and ends at the initial time $t = 0$. That is, the backward time integration begins with the final-value conditions, $\boldsymbol{\lambda}(T) = 0$, $\dot{\boldsymbol{\lambda}}(T) = 0$, and

$$\ddot{\boldsymbol{\lambda}}(T) = \mathbf{M}^{-1}\mathbf{F}_{\text{adj}}(T). \tag{29}$$

For the backward time-marching procedure from $t = T$, the following approximations are used:

$$\dot{\boldsymbol{\lambda}}_n = \dot{\boldsymbol{\lambda}}_{n+1} - \frac{\Delta t}{2}\ddot{\boldsymbol{\lambda}}_n - \frac{\Delta t}{2}\ddot{\boldsymbol{\lambda}}_{n+1}, \tag{30}$$

$$\ddot{\boldsymbol{\lambda}}_n = \frac{4}{(\Delta t)^2}(\boldsymbol{\lambda}_{n+1} - \boldsymbol{\lambda}_n) - \frac{4}{\Delta t}\dot{\boldsymbol{\lambda}}_n - \ddot{\boldsymbol{\lambda}}_{n+1}. \tag{31}$$

Substituting (30) into (31) yields:

$$\ddot{\boldsymbol{\lambda}}_n = -\frac{4}{(\Delta t)^2}(\boldsymbol{\lambda}_{n+1} - \boldsymbol{\lambda}_n) + \frac{4}{\Delta t}\dot{\boldsymbol{\lambda}}_{n+1} - \ddot{\boldsymbol{\lambda}}_{n+1}. \tag{32}$$

By inserting (32) into the discrete form of the adjoint equation (27), we obtain the following equation that will be used for computing $\boldsymbol{\lambda}_n$ at every n-th time step:

$$\boldsymbol{\lambda}_n = \left[\mathbf{K} + \frac{4}{(\Delta t)^2}\mathbf{M}\right]^{-1}\left\{\mathbf{F}_{\text{adj}}(t_n) + \mathbf{M}\left(\frac{4}{(\Delta t)^2}\boldsymbol{\lambda}_{n+1} - \frac{4}{\Delta t}\dot{\boldsymbol{\lambda}}_{n+1} + \ddot{\boldsymbol{\lambda}}_{n+1}\right)\right\}. \tag{33}$$

Once we obtain $\boldsymbol{\lambda}_n$ from (33), we obtain $\ddot{\boldsymbol{\lambda}}_n$ from (32) and, in turn, $\dot{\boldsymbol{\lambda}}_n$ from (30).

The discrete form of the gradient

The gradient of the objective functional $\mathcal{L}$ with respect to a scalar variable ξ can now be numerically computed as:

$$\nabla_{(\xi=F_{kj})}\mathcal{L} = -\Delta x\Delta t \times \lambda(x_k, D, t_j) - R \times \Delta x\Delta t \times \left[\frac{\partial^2 F(x,t)}{\partial x^2} + \frac{\partial^2 F(x,t)}{\partial t^2}\right]_{\text{at}\, x_k, t_j}. \tag{34}$$

Here, this works uses the aforementioned FEM solution of the adjoint problem for evaluating the gradient in (34) under the OTD approach.

INVERSE MODELING—THE DISCRETIZE–THEN-OPTIMIZE (DTO) APPROACH

This section presents the inverse modeling based on the DTO approach. The Lagrangian functional is built by imposing the discrete form of the state problem, using the discrete adjoint variable, into the objective functional in the discrete form. The first-order optimality conditions are derived in the discrete form. The time-integration implementation of the state and adjoint problems are already embedded in their discrete forms.

The discrete objective functional

The discrete-form counterpart of the objective functional (9) is given by:

$$\hat{\mathcal{L}} = (\hat{\mathbf{u}}_\mathrm{m} - \hat{\mathbf{u}})^\mathrm{T}\, \overline{\mathbf{B}}\, (\hat{\mathbf{u}}_\mathrm{m} - \hat{\mathbf{u}}) + \frac{R}{2}\, \hat{\mathbf{F}}^\mathrm{T} \mathbf{R}\, \hat{\mathbf{F}}, \tag{35}$$

where $\hat{\mathbf{u}} = [\mathbf{u}_0\, \dot{\mathbf{u}}_0\, \ddot{\mathbf{u}}_0\, \mathbf{u}_1\, \dot{\mathbf{u}}_1\, \ddot{\mathbf{u}}_1 \ldots \mathbf{u}_\mathrm{N}\, \dot{\mathbf{u}}_\mathrm{N}\, \ddot{\mathbf{u}}_\mathrm{N}]^T$ corresponds to the space-time discretization of $u(x, y, t)$ for $(x, y) \in \Omega$ and $t \in [0, T]$, induced by an estimated $F(x, t)$ (N is the number of time steps, and $\mathbf{u}_i$ are the spatial degrees of freedom at the i-th time step); $\hat{\mathbf{u}}_m$ is the space-time discretization of $u_\mathrm{m}(x, t)$ induced by a targeted $F(x, t)$; and $\overline{\mathbf{B}}$ is a block diagonal matrix, determined as $\overline{\mathbf{B}} = \Delta t\mathbf{B}$ on the diagonal, where $\mathbf{B}$ is a square matrix that is zero everywhere except on the diagonals that correspond to a degree of freedom for which measured data are available; R is the regularization factor; $\hat{\mathbf{F}} = [0\, 0\, \mathbf{F}_0\, \mathbf{F}_1\, 0\, 0 \ldots \mathbf{F}_\mathrm{N}\, 0\, 0]^T$ is a global force vector corresponding to all the time steps—that is, the discrete control parameter F_{kj} are populated in $\hat{\mathbf{F}}$; and $\mathbf{R}$ is the matrix corresponding to the discretization scheme used for the regularization terms defined as:

$$\mathbf{R} = \int_0^T \int_{\Gamma_\mathrm{b}} \left(\frac{\partial \boldsymbol{w}(x, 0, t)}{\partial x} \frac{\partial \boldsymbol{w}^T(x, 0, t)}{\partial x} + \frac{\partial \boldsymbol{w}(x, 0, t)}{\partial t} \frac{\partial \boldsymbol{w}^T(x, 0, t)}{\partial t} \right) \mathrm{d}\Gamma\, \mathrm{d}t, \tag{36}$$

where $\boldsymbol{w}(x, y, t)$ denotes a vector of global basis functions, in both space and time, constructed by shape functions of each finite element mesh in the domain and the shape functions over the time.

That is, an estimated traction function $F(x,t)$ can be discretized as:

$$F(x,t) = \boldsymbol{w}^T(x, y=0, t)\hat{\mathbf{F}}. \tag{37}$$

The discrete Lagrangian functional

The Lagrangian functional corresponding to (35) is built by imposing the discrete form of the state problem using the discrete adjoint variable:

$$\hat{\mathscr{A}} = (\hat{\mathbf{u}}_{\mathrm{m}} - \hat{\mathbf{u}})^{\mathrm{T}}\,\overline{\mathbf{B}}\,(\hat{\mathbf{u}}_{\mathrm{m}} - \hat{\mathbf{u}}) + \frac{R}{2}\,\hat{\mathbf{F}}^{\mathrm{T}}\mathbf{R}\,\hat{\mathbf{F}} + \hat{\boldsymbol{\lambda}}^T(\mathbf{Q}\hat{\mathbf{u}} - \hat{\mathbf{F}}), \tag{38}$$

where $\hat{\boldsymbol{\lambda}} = [\boldsymbol{\lambda}_0\,\dot{\boldsymbol{\lambda}}_0\,\ddot{\boldsymbol{\lambda}}_0\,\boldsymbol{\lambda}_1\,\dot{\boldsymbol{\lambda}}_1\,\ddot{\boldsymbol{\lambda}}_1 \ldots \boldsymbol{\lambda}_{\mathrm{N}}\,\dot{\boldsymbol{\lambda}}_{\mathrm{N}}\,\ddot{\boldsymbol{\lambda}}_{\mathrm{N}}]^T$ is the discrete (space-time) Lagrange multiplier that enforces the discrete forward problem as a constraint; and $\mathbf{Q}$ is the discrete forward operator defined as:

$$\mathbf{Q} = \begin{bmatrix}
\mathbf{I} & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \mathbf{I} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{K} & 0 & \mathbf{M} & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{L}_1 & \mathbf{L}_2 & \mathbf{L}_3 & \mathbf{Keff} & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 \\
a_1\mathbf{I} & \mathbf{I} & 0 & -a_1\mathbf{I} & \mathbf{I} & 0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 \\
a_0\mathbf{I} & a_2\mathbf{I} & \mathbf{I} & -a_0\mathbf{I} & 0 & \mathbf{I} & \ldots & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & \ldots & \mathbf{L}_1 & \mathbf{L}_2 & \mathbf{L}_3 & \mathbf{Keff} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \ldots & a_1\mathbf{I} & \mathbf{I} & 0 & -a_1\mathbf{I} & \mathbf{I} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \ldots & a_0\mathbf{I} & a_2\mathbf{I} & \mathbf{I} & -a_0\mathbf{I} & 0 & \mathbf{I}
\end{bmatrix}, \tag{39}$$

where:

$$a_0 = \frac{4}{(\Delta t)^2},\ a_1 = \frac{2}{\Delta t},\ a_2 = \frac{4}{\Delta t}, \tag{40}$$

$$\mathbf{Keff} = a_0\mathbf{M} + \mathbf{K}, \tag{41}$$

$$\mathbf{L}_1 = -a_0\mathbf{M},\ \mathbf{L}_2 = -a_2\mathbf{M},\ \mathbf{L}_3 = -\mathbf{M}. \tag{42}$$

The first-order optimality condition in the DTO modeling

The discrete optimality conditions of (35) require that the variations of $\hat{\mathcal{A}}$ with respect to $\hat{\boldsymbol{\lambda}}$, $\hat{\mathbf{u}}$ and $\hat{\mathbf{F}}$ vanish. The first condition, taking the variation with respect to $\hat{\boldsymbol{\lambda}}$, recovers the discrete form of the state equation:

$$\frac{\partial \hat{\mathcal{A}}}{\partial \hat{\boldsymbol{\lambda}}} = \mathbf{Q}\hat{\mathbf{u}} - \hat{\mathbf{F}} = 0. \tag{43}$$

For the second condition, the variation of $\hat{\mathcal{A}}$ with respect to $\hat{\mathbf{u}}$ should vanish:

$$\frac{\partial \hat{\mathcal{A}}}{\partial \hat{\mathbf{u}}} = \mathbf{Q}^T \hat{\boldsymbol{\lambda}} + 2\,\overline{\mathbf{B}}\,(\hat{\mathbf{u}}_\mathrm{m} - \hat{\mathbf{u}}) = 0. \tag{44}$$

Equation (44) represents the discrete adjoint equation. Since it involves the transpose of $\mathbf{Q}$, we solve it by marching backwards in time. For example, from the last two rows of (39), we obtain the final conditions:

$$\ddot{\boldsymbol{\lambda}}_N = 0, \tag{45}$$

$$\dot{\boldsymbol{\lambda}}_N = 0, \tag{46}$$

respectively; and the third row from the bottom yields:

$$\mathbf{Keff}^T \boldsymbol{\lambda}_N = 2\Delta t\mathbf{B}\,(\mathbf{u}_N - \mathbf{u}_{\mathrm{m}N}) + a_1\dot{\boldsymbol{\lambda}}_N + a_0\ddot{\boldsymbol{\lambda}}_N, \tag{47}$$

which can be solved for $\boldsymbol{\lambda}_N$. For time steps $n = N-1,\ N-2, \ldots,\ 1$, we first update $\ddot{\boldsymbol{\lambda}}_n$ and $\dot{\boldsymbol{\lambda}}_n$ as the following:

$$\ddot{\boldsymbol{\lambda}}_n = \mathbf{M}^T \boldsymbol{\lambda}_{n+1} - \ddot{\boldsymbol{\lambda}}_{n+1}, \tag{48}$$

$$\dot{\boldsymbol{\lambda}}_n = a_2 \mathbf{M}^T \boldsymbol{\lambda}_{n+1} - \dot{\boldsymbol{\lambda}}_{n+1} - a_2 \ddot{\boldsymbol{\lambda}}_{n+1}, \tag{49}$$

and, then, solve the following:

$$\begin{aligned} \mathbf{Keff}^T \boldsymbol{\lambda}_n =& 2\Delta t \mathbf{B} \left(\mathbf{u}_n - \mathbf{u}_{\mathrm{m}n}\right) + a_1 \dot{\boldsymbol{\lambda}}_n + a_0 \ddot{\boldsymbol{\lambda}}_n \\ &+ a_0 \mathbf{M}^T \boldsymbol{\lambda}_{n+1} - a_1 \dot{\boldsymbol{\lambda}}_{n+1} - a_0 \ddot{\boldsymbol{\lambda}}_{n+1}, \end{aligned} \tag{50}$$

Finally, the first three rows of (15) result in the following equations. First, we solve:

$$\mathbf{M}^T \ddot{\boldsymbol{\lambda}}_0 = \mathbf{M}^T \boldsymbol{\lambda}_1 - \ddot{\boldsymbol{\lambda}}_1, \tag{51}$$

and, then, update $\dot{\boldsymbol{\lambda}}_0$ and $\boldsymbol{\lambda}_0$ as the following:

$$\dot{\boldsymbol{\lambda}}_0 = a_2 \mathbf{M}^T \boldsymbol{\lambda}_1 - \dot{\boldsymbol{\lambda}}_1 - a_2 \ddot{\boldsymbol{\lambda}}_1, \tag{52}$$

$$\boldsymbol{\lambda}_0 = -\mathbf{K}^T \ddot{\boldsymbol{\lambda}}_0 + a_0 \mathbf{M}^T \boldsymbol{\lambda}_1 - a_1 \dot{\boldsymbol{\lambda}}_1 - a_0 \ddot{\boldsymbol{\lambda}}_1 + \Delta t \mathbf{B} \left(\mathbf{u}_0 - \mathbf{u}_{\mathrm{m}0}\right). \tag{53}$$

We note that the backward adjoint time integration in the DTO approach differs from that shown in its counterpart of the OTD approach.

The third condition states that the variation of $\mathcal{A}$ with respect to $\hat{\mathbf{F}}$ should vanish:

$$\frac{\partial \hat{\mathcal{A}}}{\partial \hat{\mathbf{F}}} = R\, \mathbf{R}\, \hat{\mathbf{F}} - \hat{\boldsymbol{\lambda}} = 0, \tag{54}$$

which represents the discrete control equation and implies that $\frac{\partial \hat{\mathcal{A}}}{\partial \xi} = \frac{\partial \hat{\mathcal{L}}}{\partial \xi}$ is the component of the vector:

$$R\,\mathbf{R}\,\hat{\mathbf{F}} - \hat{\lambda}, \tag{55}$$

at its row corresponding to $\xi = F_{kj}$.

Implementation of the regularization term in the gradient

This subsection shows the detail of evaluating the regularization term, $R\,\mathbf{R}\,\hat{\mathbf{F}}$, in (55). The non-zero contribution of an element (shown in Fig. 2) in the space in terms of x and t to $\mathbf{R}$ matrix is:

$$\mathbf{R}^e = \Delta x \Delta t \begin{bmatrix} (\frac{1}{3}\frac{1}{(\Delta x)^2} + \frac{1}{3}\frac{1}{(\Delta t)^2}) & (-\frac{1}{3}\frac{1}{(\Delta x)^2} + \frac{1}{6}\frac{1}{(\Delta t)^2}) & (-\frac{1}{6}\frac{1}{(\Delta x)^2} - \frac{1}{6}\frac{1}{(\Delta t)^2}) & (\frac{1}{6}\frac{1}{(\Delta x)^2} - \frac{1}{3}\frac{1}{(\Delta t)^2}) \\ (-\frac{1}{3}\frac{1}{(\Delta x)^2} + \frac{1}{6}\frac{1}{(\Delta t)^2}) & (\frac{1}{3}\frac{1}{(\Delta x)^2} + \frac{1}{3}\frac{1}{(\Delta t)^2}) & (\frac{1}{6}\frac{1}{(\Delta x)^2} - \frac{1}{3}\frac{1}{(\Delta t)^2}) & (-\frac{1}{6}\frac{1}{(\Delta x)^2} - \frac{1}{6}\frac{1}{(\Delta t)^2}) \\ (-\frac{1}{6}\frac{1}{(\Delta x)^2} - \frac{1}{6}\frac{1}{(\Delta t)^2}) & (\frac{1}{6}\frac{1}{(\Delta x)^2} - \frac{1}{3}\frac{1}{(\Delta t)^2}) & (\frac{1}{3}\frac{1}{(\Delta x)^2} + \frac{1}{3}\frac{1}{(\Delta t)^2}) & (-\frac{1}{3}\frac{1}{(\Delta x)^2} + \frac{1}{6}\frac{1}{(\Delta t)^2}) \\ (\frac{1}{6}\frac{1}{(\Delta x)^2} - \frac{1}{3}\frac{1}{(\Delta t)^2}) & (-\frac{1}{6}\frac{1}{(\Delta x)^2} - \frac{1}{6}\frac{1}{(\Delta t)^2}) & (-\frac{1}{3}\frac{1}{(\Delta x)^2} + \frac{1}{6}\frac{1}{(\Delta t)^2}) & (\frac{1}{3}\frac{1}{(\Delta x)^2} + \frac{1}{3}\frac{1}{(\Delta t)^2}) \end{bmatrix}. \tag{56}$$

Accordingly, the four elements surrounding F_{kj} (see Fig. 2) contribute to the 9×9 submatrix (i.e., $\mathbf{R}^{4\text{E}}$) of $\mathbf{R}$. Then, the component of $R\,\mathbf{R}\,\hat{\mathbf{F}}$ corresponding to F_{kj} can be computed as:

$$\frac{\partial \mathcal{R}^{\text{TN}}}{\partial F_{kj}} = R(\mathbf{R}^{4\text{E}}_{\text{5-th row}})\mathbf{F}^{4\text{E}}, \tag{57}$$

where the 5-th row of $\mathbf{R}^{4E}$ (i.e., $\mathbf{R}^{4E}_{\text{5-th row}}$) is:

$$\mathbf{R}^{4E}_{\text{5-th row}} = \Delta x \Delta t \begin{bmatrix} (-\frac{1}{6(\Delta x)^2} - \frac{1}{6(\Delta t)^2}) \\ (\frac{1}{3(\Delta x)^2} - \frac{2}{3(\Delta t)^2}) \\ (-\frac{1}{6(\Delta x)^2} - \frac{1}{6(\Delta t)^2}) \\ (-\frac{2}{3(\Delta x)^2} + \frac{1}{3(\Delta t)^2}) \\ (\frac{4}{3(\Delta x)^2} + \frac{4}{3(\Delta t)^2}) \\ (-\frac{2}{3(\Delta x)^2} + \frac{1}{3(\Delta t)^2}) \\ (-\frac{1}{6(\Delta x)^2} - \frac{1}{6(\Delta t)^2}) \\ (\frac{1}{3(\Delta x)^2} - \frac{2}{3(\Delta t)^2}) \\ (-\frac{1}{6(\Delta x)^2} - \frac{1}{6(\Delta t)^2}). \end{bmatrix}^T , \tag{58}$$

and $\mathbf{F}^{4E}$ is defined as:

$$\mathbf{F}^{4E} = \begin{bmatrix} F_{(k-1)(j-1)} \\ F_{(k-1)j} \\ F_{(k-1)(j+1)} \\ F_{k(j-1)} \\ F_{kj} \\ F_{k(j+1)} \\ F_{(k+1)(j-1)} \\ F_{(k+1)j} \\ F_{(k+1)(j+1)} \end{bmatrix} . \tag{59}$$

Thus, under the DTO approach, (57) can be implemented as:

$$
\begin{aligned}
\frac{\partial \mathcal{R}^{\mathrm{TN}}}{\partial F_{kj}}_{\mathrm{DTO}} =& -R \times \Delta x \Delta t \times \\
& \left\{ \frac{1}{6} \left(\frac{F_{(k-1)(j-1)} - 2F_{(k-1)j} + F_{(k-1)(j+1)}}{(\Delta x)^2} \right) + \frac{4}{6} \left(\frac{F_{k(j-1)} - 2F_{kj} + F_{k(j+1)}}{(\Delta x)^2} \right) \right. \\
& + \frac{1}{6} \left(\frac{F_{(k+1)(j-1)} - 2F_{(k+1)j} + F_{(k+1)(j+1)}}{(\Delta x)^2} \right) + \frac{1}{6} \left(\frac{F_{(k-1)(j-1)} - 2F_{k(j-1)} + F_{(k+)(j-1)}}{(\Delta t)^2} \right) \\
& \left. + \frac{4}{6} \left(\frac{F_{(k-1)j} - 2F_{kj} + F_{(k+1)j}}{(\Delta t)^2} \right) + \frac{1}{6} \left(\frac{F_{(k-1)(j+1)} - 2F_{k(j+1)} + F_{(k+1)(j+1)}}{(\Delta t)^2} \right) \right\}, \qquad (60)
\end{aligned}
$$

while, its counterpart in the OTD approach can be implemented as:

$$
\frac{\partial \mathcal{R}^{\mathrm{TN}}}{\partial F_{kj}}_{\mathrm{OTD}} = -R \times \Delta x \Delta t \times \left\{ \left(\frac{F_{k(j-1)} - 2F_{kj} + F_{k(j+1)}}{(\Delta x)^2} \right) + \left(\frac{F_{(k-1)j} - 2F_{kj} + F_{(k+1)j}}{(\Delta t)^2} \right) \right\}, \qquad (61)
$$

which corresponds to (34).

NUMERICAL IMPLEMENTATION OF THE INVERSION PROCESS

By utilizing the semi-analytically evaluated gradient vector $\nabla_{\xi}\mathcal{L}$, this work iteratively updates a set of estimated control parameters by using the gradient-based minimization scheme as follows:

(a) First, we compute synthetic measured data u_{m} at sensors by using pseudo-target traction $F(x,t)$.

(b) Then, $u(x,t)$ is obtained by using estimated $F(x,t)$ that is constructed by estimated control parameters $\boldsymbol{\xi}$.

(c) The adjoint problem is, then, solved by using the solution of the state problem in the previous step.

(d) The gradient of the objective functional, $\nabla_{\xi}\mathcal{L}$, is evaluated.

(e) Finally, the gradient-based minimization scheme updates the estimated control parameters $\boldsymbol{\xi}$ via the conjugate-gradient method and the Newton's method. The conjugate-gradient method determines the best search direction, and the Newton's method determines an optimal step length.

The numerical optimizer repeats the above steps (b) to (e) and iteratively solve for the control parameters that satisfy the vanishing control equation. A set of these steps is counted as an inversion iteration. The detailed procedure of the numerical optimizer is summarized in Algorithm 1.

Algorithm 1 Minimization Algorithm

Set TOL=10^{-10}
Build $\mathbf{M}$ and $\mathbf{K}$ matrices for forward and adjoint wave solvers.
Compute u_{m} by using a forward wave solver with target traction $F(x,t)$.
Set an iteration index $s=1$ and initial control parameters ($\boldsymbol{\xi}_{(s=1)}=0$) and compute $\mathcal{L}_{(s=1)}$.
while ($\mathcal{L}_{(s+1)} > \text{TOL} \times \mathcal{L}_{(1)}$ and $s < 10^3$) **do**
    Compute $u(x,y,t)$ by solving the discrete form of the forward problem using $\boldsymbol{\xi}_{(s)}$.
    Solve the discrete form of the adjoint problem.
    Compute the components of a gradient vector, $\mathbf{g}_{(s)} = \nabla_{\xi}\mathcal{L}_{(s)}$.
    Compute an optimal search direction $\mathbf{d}_{(s)}$ by using the conjugate-gradient scheme.
    Compute an optimal step length $h_{(s)}$ by using the Newton's method.
    Update estimated control parameters as $\boldsymbol{\xi}_{(s+1)} = \boldsymbol{\xi}_{(s)} + \mathbf{d}_{(s)} h_{(s)}$ and compute $\mathcal{L}_{(s+1)}$.
    $s \leftarrow s+1$
end while

Conjugate gradient

In every s-th inversion iteration, the gradient vector is computed as $\mathbf{g}_{(s)} = \nabla_{\xi}\mathcal{L}_{(s)}$. By using $\mathbf{g}_{(s)}$, the search direction $\mathbf{d}_{(s)}$ is computed by using the conjugate-gradient scheme (Fletcher and Reeves 1951):

$$\begin{aligned} \mathbf{d}_{(s)} &= -\mathbf{g}_{(s)} && (s=0 \text{ and every } m \text{ (e.g., } m=5)), \\ \mathbf{d}_{(s)} &= -\mathbf{g}_{(s)} + \frac{\mathbf{g}_{(s)}\cdot\mathbf{g}_{(s)}}{\mathbf{g}_{(s-1)}\cdot\mathbf{g}_{(s-1)}}\mathbf{d}_{(s-1)} && (s \geq 1). \end{aligned} \tag{62}$$

The $\mathbf{g}_{(s)}$ is reset to be equal to $-\mathbf{g}_{(s)}$ at every m inversion iteration in order to eliminate the progressively-accumulated error in the search direction (Kang and Kallivokas 2010a). In the presented numerical experiments, we used $m=5$.

Adaptively-calculated regularization factor

The value of the regularization factor R in the above gradient (18) determines the extent, to which the penalty is imposed on the oscillation of the spatial and temporal variation of $F(x,t)$. If R is too large, the estimated traction profile may remain too smooth. If R is too small, the numerical optimizer could suffer from the solution multiplicity. To determine the optimal value of R during the inversion process, this work adopts the regularization factor continuation scheme (Kang and Kallivokas 2010a). To this end, we decompose $\mathcal{L}$ into $\mathcal{L}_m$ and $\mathcal{R}^{\text{TN}} = R\mathcal{L}_R$, which are defined as:

$$\begin{aligned} \mathcal{L}_m &= \int_0^T \sum_{i=1}^{N_\text{s}} (u_{\text{m}i} - u_i)^2 \,\mathrm{d}t, \\ \mathcal{L}_R &= \frac{1}{2}\int_0^T \int_{\Gamma_\text{b}} \left(\frac{\partial F(x,t)}{\partial x}\right)^2 + \left(\frac{\partial F(x,t)}{\partial t}\right)^2 \mathrm{d}\Gamma\,\mathrm{d}t. \end{aligned} \tag{63}$$

Accordingly, $\nabla_{(\xi=F_{kj})}\mathcal{L}$ in (18) can be decomposed into the following two in the OTD approach:

$$\nabla_{(\xi=F_{kj})}\mathcal{L}_m = -\int_0^T \int_{\Gamma_\text{b}} \lambda\Phi_k(x)\phi_j(t)\,\mathrm{d}\Gamma\,\mathrm{d}t, \tag{64}$$

$$R\nabla_{(\xi=F_{kj})}\mathcal{L}_R = -R\int_0^T \int_{\Gamma_\text{b}} \left(\frac{\partial^2 F(x,t)}{\partial x^2} + \frac{\partial^2 F(x,t)}{\partial t^2}\right)\Phi_k(x)\phi_j(t)\,\mathrm{d}\Gamma\,\mathrm{d}t. \tag{65}$$

or the following two in the DTO approach: $\nabla_{(\xi=F_{kj})}\hat{\mathcal{L}}_m$ and $R\nabla_{(\xi=F_{kj})}\hat{\mathcal{L}}_R$, which are the components of the vectors, respectively, $-\hat{\lambda}$ and $R\mathbf{R}\,\hat{\mathbf{F}}$ in (55), at their rows corresponding to $\xi = F_{kj}$. Here, Kang and Kallivokas (2010a) suggested imposing the following inequality:

$$R|\nabla_\xi \mathcal{L}_R| < |\nabla_\xi \mathcal{L}_m|, \quad \text{or} \quad R < \frac{|\nabla_\xi \mathcal{L}_m|}{|\nabla_\xi \mathcal{L}_R|}. \tag{66}$$

Thus, in each inversion iteration, R can be set as:

$$R = I_R \frac{|\nabla_\xi \mathcal{L}_m|}{|\nabla_\xi \mathcal{L}_R|}, \tag{67}$$

where I_R denotes the regularization intensity factor. Since Kang and Kallivokas (2010a) heuristically found that the value of I_R should be $0 < I_R < 0.5$ for the material inversion, we tested the performance of the presented inversion with respect to I_R in the numerical experiments as well.

Updating the estimated control parameters

Starting with an initial guess for the control parameter vector $\boldsymbol{\xi}$, which is comprised of F_{kj}, the estimate of $\boldsymbol{\xi}$ can be updated iteratively as:

$$\boldsymbol{\xi}_{(s+1)} = \boldsymbol{\xi}_{(s)} + h_{(s)}\mathbf{d}_{(s)}, \tag{68}$$

where $h_{(s)}$ is a scalar-valued step size and $\mathbf{d}_{(s)}$ is a search-direction vector computed by the conjugate-gradient scheme, and s is the iteration index in the optimization process. For each iteration, s, the Newton's method (Lloyd and Jeong 2018) is used to determine the optimal step size, $h_{(s)}$. Namely, in each s-th iteration in the numerical optimization process, there are up to 4 sub-iterations (r is the sub-iteration index). We begin from an initially estimated $h_{s_{(r=1)}}$, and, for the next sub-iteration of $(r+1)$ up to r of 4, we update $h_{s_{(r+1)}}$ as the following:

$$h_{s_{(r+1)}} = h_{s_r} - \left(\frac{\mathcal{L}'(h_{s_r})_{(s+1)}}{\mathcal{L}''(h_{s_r})_{(s+1)}}\right) = h_{s_r} - \frac{\left(\frac{\mathcal{L}(h_{s_r}+\eta)_{(s+1)} - \mathcal{L}(h_{s_r}-\eta)_{(s+1)}}{2\eta}\right)}{\left(\frac{\mathcal{L}(h_{s_r}+\eta)_{(s+1)} - 2\mathcal{L}(h_{s_r})_{(s+1)} + \mathcal{L}(h_{s_r}-\eta)_{(s+1)}}{\eta^2}\right)} \tag{69}$$

where $\mathcal{L}(h_{s_r})_{(s+1)}$ is the objective functional given the updated $\boldsymbol{\xi}$ in (68) using h_{s_r}. For each sub-iteration, r, values for the first and second derivatives of the objective functional with respect to h_s—i.e., $\mathcal{L}'(h_{s_r})_{(s+1)}$ and $\mathcal{L}''(h_{s_r})_{(s+1)}$—are determined numerically via central difference approximations. In the very right hand side term of (69), $\mathcal{L}(h_{s_r} \pm \eta)_{(s+1)}$ is $\mathcal{L}_{(s+1)}$ evaluated when $\boldsymbol{\xi}_{(s+1)} = \boldsymbol{\xi}_s + (h_{s_r} \pm \eta)\mathbf{d}_s$.

NUMERICAL EXPERIMENTS

This section shows a set of numerical examples, investigating the performance of the presented inverse modeling with respect to various factors. In all the examples, we consider a square-shaped solid domain, of which extent is 60 m × 60 m. To avoid an inverse crime, we compute the synthetic measured data u_{m} using an element size set as 0.5 m, while an element size of 1 m is used in the FEM solvers for obtaining state and adjoint solutions in the computational domain. The same time step of 0.001 s is used in the forward and inversion problems. The solid has a uniform mass density ρ of 1000 kg/m^3.

Two targeted dynamic tractions, $F_1(x,t)$ and $F_2(x,t)$, are considered in this section. The targeted $F_1(x,t)$, of which amplitude changes over space and time, is shown in Fig. 6(a). Its peak amplitude is 200 N/m^2, and the total observation duration T_1 is 1.5 s in the inversion simulation of $F_1(x,t)$. The time-dependent value of the targeted $F_1(x,t)$ at any specific value of x is a Ricker wavelet (for instance, see $F_1(30,t)$ in Fig. 3(a)) with its central frequency of 10 Hz (see the frequency contents of $F_1(30,t)$ in Fig. 3(b)).

The targeted $F_2(x,t)$ is shown in Fig. 19(a), and its amplitude changes over space and time. Fig. 3(c) shows the time-dependent value of the targeted $F_2(30,t)$ and Fig. 3(d) shows its frequency contents. The signal is a part of a recorded ground motion signal during the 1994 Northridge earthquake from the Pacific Earthquake Engineering Research Center (PEER) ground motion database (PEER 2000). The total observation duration T_2 is 6 s for the inversion simulation to identify the targeted $F_2(30,t)$.

For the inversion process, we discretize estimated $F_1(x,t)$ by using 91,500 control parameters—i.e., 61 (over space) $\times$ 1500 (over time)—and $F_2(x,t)$ by using 366,000 control parameters—i.e., 61 (over space) $\times$ 6000 (over time). The temporal and spatial intervals for the discretization are 0.001 s and 1 m, respectively, for both forces. Initially-estimated values of all the control parameters are zero. Our numerical optimizer iteratively updates the values of all the control parameters by using the presented inverse modeling procedure.

In what follows, six examples of numerical experiments are presented. The first example is focused on the performance of inverting for $F_1(x,t)$ with respect to the complexity of the material profile in the domain. To this end, we consider the following material profiles:

- Material profile 1: A homogeneous solid with the wave speed of V_s = 250 m/s,
- Material profile 2: A 2-layered solid with 1 inclusion as in Fig. 4(a) with wave speeds of V_{s_1}

= 300 m/s, V_{s_2} = 350 m/s, and V_{s_3} = 200 m/s,

- Material profile 3: A 3-layered solid with 3 inclusions as in Fig. 4(b) with wave speeds of V_{s_1} = 400 m/s, V_{s_2} = 450 m/s, V_{s_3} = 300 m/s, V_{s_4} = 350 m/s, V_{s_5} = 200 m/s, and V_{s_6} = 250 m/s.

The second example compares the performance of identifying $F_1(x,t)$ by using the OTD method versus the DTO method. The third example tests the inversion performance of identifying $F_1(x,t)$ with respect to the number of sensors on the top surface of the domain. The fourth example examines the convergence of the estimated $F_1(x,t)$ into the target with respect to the regularization intensity factor, I_R. The fifth example is focused on the performance of inverting for $F_1(x,t)$ with respect to the noise level. Lastly, the sixth example shows the capability of our inverse modeling to reconstruct $F_2(x,t)$, which is a realistic seismic signal as opposed to $F_1(x,t)$. Each example considers the result data of multiple cases, of which input parameters are summarized in Table 1. For the sake of assessing the accuracy to reconstruct $F(x,t)$ in the numerical results, the following error norm between estimated $F(x,t)$ and its target is used:

$$\mathcal{E} = \frac{\int_0^T \int_\Omega |F(x,t)_{\text{target}} - F(x,t)_{\text{estimate}}|^2 \, \mathrm{d}\Omega \, \mathrm{d}t}{\int_0^T \int_\Omega |F(x,t)_{\text{target}}|^2 \, \mathrm{d}\Omega \, \mathrm{d}t} \times 100[\%]. \tag{70}$$

Before our parametric studies on the inversion performance, we verify the theoretical derivation of the adjoint and control equations and the numerical implementation. That is, the gradients obtained by our semi-analytical OTD and DTO approaches are compared with that from the finite difference (FD) approach. In this verification, we used the targeted $F_1(x,t)$, a heterogeneous domain of the material profile 3, shown in Fig. 4(b), only one sensor in the center of Γ_t, I_R of 0, and the noise level of 0%. In order to reduce the computational cost of the FD approach, the spatial distribution of $F_1(x,t)$ is set to be uniform, and, thus, only the temporal variation of $F_1(x,t)$ is considered. Fig. 5 shows excellent agreement among the normalized gradients (i.e., $\nabla_\xi \mathcal{L}/|\nabla_\xi \mathcal{L}|$) that are calculated by using the OTD approach, the DTO approach, and the FD approach, respectively, at the first inversion iteration. Thus, our theoretical derivation and numerical implementation of the presented inverse modeling are trustable.

Example 1: Investigating the inversion performance with respect to the material profile complexity.

In this example, we test the performance of the presented inverse modeling with respect to the complexity of the material profile by using the results of Cases 1 to 3. We used the OTD approach, 30 sensors, I_R of 0.5, and the noise level of 0% in these Cases.

Fig. 6(b,c,d) show reconstructed dynamic traction for Cases 1, 2, and 3, respectively. Figs. 7 and 8 show that the values of $\mathcal{L}$ and $\mathcal{E}$, in general, decrease as the number of iterations increases. Fig. 7 depicts that $\mathcal{L}$ shows a sawtooth behavior over iterations. We suggest that it occurs because of the penalty that is imposed by the regularization term on the derivative of $F_1(x,t)$. Namely, as seen in the later Examples 4 and 6, when I_R is equal to zero, the sawtooth behavior of $\mathcal{L}$ does not occur. Besides, Fig. 8 shows that the more complex the material profile is, the higher terminal value of $\mathcal{E}$ is obtained. Fig. 9 shows that the wave responses of u_{m} due to the targeted $F_1(x,t)$ match those of u due to the reconstructed one at two sensors in Cases 1-3. It implies that our numerical minimizer is very effective in minimizing the misfit: there is a very small difference, of the scale of 10^{-13} to 10^{-12}, among the terminal values of $\mathcal{L}$ in Fig. 7 for Cases 1 to 3. Even though u match u_{m} at the end of the inversion simulation, our optimizer results in a higher terminal value of $\mathcal{E}$ for a more complex material profile. To explain this aspect, we note that Lloyd and Jeong (2018) reported that a more heterogeneous background solid leads to a larger terminal value of an error between a targeted moving, body force-typed wave source function and its estimated counterpart in a 1D solid setting. It is because, the more heterogeneous the material property of a domain is, the waves reverberate more inside the domain. That is, because of the reverberation, as seen in Fig. 6, the reconstructed traction has a stronger noise-like behavior, leading to a higher final value of $\mathcal{E}$ when a more heterogeneous material profile is used. Thus, in a more heterogeneous solid setting, the inversion solver is less likely to converge towards a targeted traction function.

Fig. 10 shows the snapshots of the wave responses in the entire computational domain, induced by (a) the targeted $F_1(x,t)$ and (b) its reconstructed counterpart in Case 3. Both responses are in excellent agreement with each other even though the terminal value of $\mathcal{E}$ in Case 3 is the largest among the Cases 1-3. Thus, the presented dynamic-input inversion algorithm could be further developed for reconstructing seismic input motions and, then, “replaying" the corresponding wave responses in a truncated computational domain during seismic events.

Example 2: Comparison between the inversion performances of the OTD and DTO approaches

This example compares the performance of the OTD approach with that of the DTO counterpart by using the results of Cases 3 and 4, in both of which we reconstruct the targeted traction $F_1(x,t)$ by using the material profile 3, I_R of 0.5, and the noise level of 0%. Fig. 11 shows the targeted $F_1(x,t)$ and its reconstructed counterparts obtained by using the OTD and DTO approaches, in Cases 3 and 4, respectively. We could not visually noice the difference between the two reconstructed ones in Fig. 11. Thus, we suggest that there is no difference between the final reconstructed traction function obtained by using the OTD approach and that by using the DTO approach in this 2D SH wave work. Fig. 12 also shows that the terminal values of $\mathcal{L}$ and $\mathcal{E}$, when using the OTD approach, are in excellent agreement with those using the DTO counterpart. However, a sawtooth behavior of $\mathcal{L}$ occurs more when we use the DTO approach than the OTD approach. As mentioned earlier, the sawtooth behavior occurs because of the regularization. Namely, as shown in (60) and (61), the gradient of the regularization term in the OTD approach is implemented differently from the DTO counterpart so that their corresponding behaviors of $\mathcal{L}$ differ from each other.

Example 3: Investigating the inversion performance with respect to the number of sensors

In this example, we study the inversion performance with respect to the number of sensors on the top surface. We considered Cases 3, 5, 6, and 7, which use 30, 15, 10, and 5 sensors, respectively, and the material profile 3, I_R of 0.5, and the noise level of 0%. Fig. 13 shows that, although increasing the number of sensors decreases the terminal value of $\mathcal{E}$ of the inversion, using 5, 10, 15, and 30 sensors gives rise to the terminal values of $\mathcal{E}$ of the almost same magnitude, i.e., 25 to 26% with a slight difference of only up to 1% from each other. Thus, we suggest that when the sampling rate of the measurement is equal to the timestep for discretizing $F_1(x,t)$, the ratio of the size of measurement data to the number of the control parameters can be as small as 1:12—e.g., (5 sensors × 1500 timesteps):(61 nodes × 1500 timesteps)—in the presented example.

Example 4: Investigating the inversion performance with respect to I_R

This example studies the accuracy of the inversion with respect to the regularization intensity factor, I_R. We used 15 sensors on the top surface, the material profile 3, and the noise level of 0% for all the cases in this example. In addition to the Case 5 in Example 3, where I_R of 0.5 is used, we considered Cases 8-11, which use I_R of 1.0, 0.1, 0.01, and 0.0, respectively.

As can be seen in Fig. 14, the terminal value of $\mathcal{L}$ for a large value of I_R (e.g., $I_R = 1.0$) is larger than those for smaller values of I_R. We also note that, when I_R is 0.0, the sawtooth behavior of $\mathcal{L}$ does not occur in Fig. 14. Furthermore, Fig. 15 shows that the terminal value of $\mathcal{E}$ for I_R of 1.0 is 54.9 % and is about twice higher than those (about 25 %) for $0 \leq I_R \leq 0.5$.

The inversion performance, in terms of $\mathcal{E}$ over the iterations, is relatively unstable for I_R of 1.0, while it is stable for $0 \leq I_R \leq 0.5$. That is, using I_R of 1.0 increases $\mathcal{E}$ after the 180-th iteration. Fig. 16 compares the final estimated traction function, when I_R is 1.0, with that for I_R of 0.0. Fig. 16 shows that the noise-like behavior is more severe in the reconstructed traction for I_R of 1.0 than that for I_R of 0.0. We also note that when I_R of 0.0 was used, $\mathcal{E}$ is the smallest (24.65%) in this example as shown in Fig. 15.

We discuss why using I_R of 0.0 leads to as a small terminal value of $\mathcal{E}$ as I_R greater than 0.0 in the following. As discussed in the previous 1D work of the seismic-input inversion (Jeong and Seylabi 2018), the proposed dynamic-traction inversion is naturally equipped with the smoothing effect even without using the TN regularization. Namely, the FEM solver naturally filters out high frequencies of the estimated traction function and smooths its temporal variation even without the regularization method. Similarly, it filters out the high-wavelength content of the spatial variation of an estimated traction function. Because of such inherent low-pass filtering of the FEM solver, as shown in our presented numerical results, even when TN regularization is not used, the inversion solver smooths an estimated traction function.

Example 5: Investigating the inversion performance with respect to the noise level

In this example, we focus on examining the performance of inverting $F_1(x,t)$ with respect to the noise level of random noise that is added to u_m prior to the inversion. We used the OTD approach, the material profile 3, 15 sensors, I_R of 0.0, and examined the noise level of 0%, 1%, 2%, and 3%, which correspond to Cases 11-14, respectively. Fig. 17 shows that, the larger the noise level is, the larger the terminal values of $\mathcal{L}$ and $\mathcal{E}$ are obtained when no regularization is used.

Fig. 18 shows the inversion performance with respect to I_R (Cases 15-17 in addition to the Case 13) when we use the material profile 3, 15 sensors, and the noise level of 2%. We note that I_R of 0.01, 0.1, and 0,5 did not make any difference in the terminal value of $\mathcal{E}$ compare to that for I_R of 0.0. Thus, we suggest that using the TN regularization does not improve the inversion performance in the presented work when u_m contains noises.

Example 6: Examining the feasibility of the presented inverse modeling to reconstruct a realistic seismic signal $F_2(x,t)$

In this example, we focus on examining the feasibility of inverting for a realistic seismic signal $F_2(x,t)$. We used the OTD approach, the material profile 3, 15 sensors, I_R of 0.0, and the noise level of 0%. Fig. 19 shows the excellent agreement between the targeted and reconstructed dynamic tractions, $F_2(x,t)$, in Case 18. Fig. 20 and 21 show the values of $\mathcal{L}$ and $\mathcal{E}$, respectively, over iterations, and $\mathcal{L}$ decreases without the sawtooth behavior because I_R of 0.0 is used. Overall, $F_2(x,t)$ has much lower frequency content than $F_1(x,t)$ so that the wave responses induced by $F_2(x,t)$ are less complex than those by $F_1(x,t)$ (see the wave responses in Fig. 22 and compare them with those in Fig. 10). Thus, our minimizer suffers from solution multiplicity less severely when it identifies $F_2(x,t)$ than $F_1(x,t)$. Accordingly, the terminal value of $\mathcal{E}$, 3.86%, for reconstructing $F_2(x,t)$ in Case 18 is much smaller than its counterpart, 29.76%, of reconstructing $F_1(x,t)$ in Case 11. Fig. 22 shows the snapshots of the wave responses in the entire computational domain, induced by (a) the targeted $F_2(x,t)$ and (b) its reconstructed counterpart in Case 18. In general, both responses are in excellent agreement with each other.

CONCLUSION

We present the mathematical modeling and numerical implementation of a new inversion process for identifying the spatial and temporal distributions of dynamic traction applied on a boundary of a 2D solid domain with SH scalar wave motions. We tackle the inverse problem by using a gradient-based minimization scheme. The gradient of an objective functional is evaluated semi-analytically by using the adjoint solution. We present both OTD and DTO methods, each of which resolves the adjoint problem differently from each other.

Numerical results show the following findings of the performance of this new inversion method. First, the complexity of the material profile in a domain increases the error between the reconstructed traction and its target. The more heterogeneous the material property of a domain is, the waves reverberate more inside the domain. Because of the reverberation, the reconstructed traction has a stronger noise-like behavior, leading to a higher terminal value of $\mathcal{E}$ when a more heterogeneous material profile is used. Thus, in a more heterogeneous solid setting, the inversion solver is less likely to converge towards a targeted traction function. Second, the OTD and DTO methods lead to the same inversion results, but the distribution of $\mathcal{L}$ over the iterations shows a more significant sawtooth behavior when we use the DTO method than the OTD method. Third, when the sampling rate of the measurement is equal to the timestep for discretizing $F(x,t)$, the ratio of the size of measurement data to the number of the control parameters can be as small as 1:12 in the presented work. Fourth, the regularization intensity I_R should not be too large: for instance, I_R is recommended to be smaller than and equal to 0.5. We also note that the terminal value of $\mathcal{E}$, when I_R of 0 is used, is as small as those when using $0.01 \leq I_R \leq 0.5$. Thus, it is acceptable to tackle the presented inverse modeling of dynamic tractions without the regularization. Fifth, the terminal values of $\mathcal{L}$ and $\mathcal{E}$ increase as the noise of a larger level is added to u_{m}, and using the TN regularization does not improve the inversion performance when noise is added to u_{m}. Sixth, our minimizer suffers from solution multiplicity less when it identifies dynamic traction of lower frequency content than that of higher frequency content.

As shown in the numerical results, the wave responses in the entire computational domain, induced by the targeted traction and the reconstructed one, are in excellent agreement with each other in the presented highly-reverberating domain. Thus, if the presented dynamic-input inversion algorithm is extended into realistic 3D settings (see below for the details of the extension), it could allow engineers to reconstruct incident seismic motions and, then, to replay the wave responses in a 3D truncated domain.

Future extensions

The presented domain does not fully represent a realistic one, which should be truncated by a wave-absorbing boundary condition (WABC) and subject to remote seismic excitation. Thus, we will extend this work as follows. A large extent of the computational domain will be truncated by using a WABC. Then, one can invert for seismic input motions in a truncated 2D/3D domain in the following two possible methods.

First, an estimated incident seismic wave can be modeled as an equivalent traction function on a WABC, e.g., dashpot (Lysmer and Kuhlemeyer 1969). The estimated traction function will be discretized over space and time, and all the discretized values will be control parameters to be identified. For instance, an x-directional traction function, $F_x(x, y, t)$, on a face that is perpendicular to the z-axis of a WABC will be discretized over space (x and y) and time (t). Then, by using the presented traction-inversion approach, we can reconstruct the traction on the surfaces of a WABC.

Second, the Domain Reduction Method (DRM) will be featured in a forward wave solver. Bielak and Christiano (1984) and Bielak et al. (2003) had developed the DRM, by which free-field wave motions are applied, as a dynamic input, along a fictitious boundary (also known as a DRM boundary) enclosed by the WABC. The DRM has been widely used for modeling wave behaviors of truncated solid domains subject to remote seismic excitations (Paolucci and Pitilakis 2007; Tripe et al. 2013; Jeremić et al. 2013; Rahnema et al. 2016; Poursartip et al. 2017; Zhang et al. 2019). Thus, the extension of the presented method will be aiming at reconstructing free-field seismic input motions at a DRM boundary. That is, we will spatially and temporally discretize estimated incident seismic wavefield functions at the two boundary surfaces of a single-element DRM buffer layer of the domain. For instance, we will discretize an x-component incident-wavefield function at the horizontal boundary of a DRM buffer layer, i.e., $u^0_{b_x}(x, y, t)$ or $u^0_{e_x}(x, y, t)$—the subscripts b and e denote the boundaries of the DRM layer neighboring the interior and exterior domains, respectively. Next, we will reconstruct the spatial and temporal distributions of the estimated incident seismic wavefield functions.

DATA AVAILABILITY STATEMENT

Some or all data, models, or code generated or used during the study are available from the corresponding author by request.

- MATLAB code (.m format) of the presented inverse modeling that contains the optimization solver and the forward and adjoint wave solvers.
- MATLAB datasets (.mat format) of the presented numerical results (Cases 1 to 18).

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Award No. CMMI-1855406. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

APPENDIX I. DERIVATION OF THE ADJOINT PROBLEM

The variation of $\mathcal{A}$ with respect to u should vanish as:

$$\delta_u \mathcal{A} = \underbrace{\int_0^T \delta_u \sum_{i=1}^{N_\mathrm{s}} (u_{\mathrm{m}i} - u_i)^2 \,\mathrm{d}t}_{a} + \underbrace{\int_0^T \int_\Omega \lambda \nabla \cdot (G \nabla \delta u) \,\mathrm{d}\Omega \,\mathrm{d}t}_{b} - \underbrace{\int_0^T \int_\Omega \lambda \rho \frac{\partial^2 \delta u}{\partial t^2} \,\mathrm{d}\Omega \,\mathrm{d}t}_{c}$$
$$+ \underbrace{\int_0^T \int_{\Gamma_\mathrm{b}} \lambda_F G \frac{\partial \delta u}{\partial y} \,\mathrm{d}\Gamma \,\mathrm{d}t}_{d} - \underbrace{\int_0^T \int_{\Gamma_\mathrm{b}} \lambda_F F(x,t) \,\mathrm{d}\Gamma \,\mathrm{d}t}_{=0} = 0. \tag{71}$$

Part a in (71) can be written as:

$$\begin{aligned} a &= \int_0^T \delta_u \sum_{i=1}^{N_\mathrm{s}} (u_{\mathrm{m}i} - u_i)^2 \,\mathrm{d}t = -\int_0^T \sum_{i=1}^{N_\mathrm{s}} 2(u_{\mathrm{m}i} - u_i) \delta u_i \,\mathrm{d}t \\ &= -\int_0^T \int_\Omega 2(u_\mathrm{m} - u) \delta u \sum_{i=1}^{N_\mathrm{s}} \Delta(x - x_i, y - y_i) \,\mathrm{d}\Omega \,\mathrm{d}t, \end{aligned} \tag{72}$$

where $\Delta(x - x_i, y - y_i)$ is the Dirac delta function. Integrating part b in (71) by parts over space twice leads to:

$$\begin{aligned} b &= \int_0^T \int_\Omega \lambda \nabla \cdot (G \nabla \delta u) \,\mathrm{d}\Omega \,\mathrm{d}t \\ &= \int_0^T \int_\Omega \nabla \cdot (\lambda G \nabla \delta u) \,\mathrm{d}\Omega \,\mathrm{d}t - \int_0^T \int_\Omega (\nabla \lambda \cdot G \nabla \delta u) \,\mathrm{d}\Omega \,\mathrm{d}t \\ &= \int_0^T \int_\Omega \nabla \cdot (\lambda G \nabla \delta u) \,\mathrm{d}\Omega \,\mathrm{d}t + \int_0^T \int_\Omega \delta u \nabla \cdot (G \nabla \lambda) \,\mathrm{d}\Omega \,\mathrm{d}t - \int_0^T \int_\Omega \nabla \cdot (\delta u G \nabla \lambda) \,\mathrm{d}\Omega \,\mathrm{d}t. \end{aligned} \tag{73}$$

Due to the Divergence Theorem, (73) becomes:

$$
\begin{aligned}
b =& \int_0^T \int_\Gamma \lambda G \frac{\partial \delta u}{\partial n} \,\mathrm{d}\Gamma \,\mathrm{d}t + \int_0^T \int_\Omega \delta u \nabla \cdot (G \nabla \lambda) \,\mathrm{d}\Omega \,\mathrm{d}t - \int_0^T \int_\Gamma \delta u G \frac{\partial \lambda}{\partial n} \,\mathrm{d}\Gamma \,\mathrm{d}t \\
=& \int_0^T \left[\int_{\Gamma_\mathrm{t}} \lambda G \underbrace{\frac{\partial \delta u}{\partial n}}_{=0} \,\mathrm{d}\Gamma + \int_{\Gamma_\mathrm{b}} \lambda G \frac{\partial \delta u}{\partial n} \,\mathrm{d}\Gamma + \int_{\Gamma_\mathrm{l,r}} \lambda G \frac{\partial \delta u}{\partial n} \,\mathrm{d}\Gamma \right] \mathrm{d}t \\
&+ \int_0^T \int_\Omega \delta u \nabla \cdot (G \nabla \lambda) \,\mathrm{d}\Omega \,\mathrm{d}t - \int_0^T \left[\int_{\Gamma_\mathrm{t,b}} \delta u G \frac{\partial \lambda}{\partial n} \,\mathrm{d}\Gamma + \int_{\Gamma_\mathrm{l,r}} \underbrace{\delta u}_{=0} G \frac{\partial \lambda}{\partial n} \,\mathrm{d}\Gamma \right] \mathrm{d}t \\
=& - \int_0^T \int_{\Gamma_\mathrm{b}} \lambda G \frac{\partial \delta u}{\partial y} \,\mathrm{d}\Gamma \,\mathrm{d}t + \int_0^T \int_{\Gamma_\mathrm{l,r}} \lambda G \frac{\partial \delta u}{\partial n} \,\mathrm{d}\Gamma \,\mathrm{d}t \\
&+ \int_0^T \int_\Omega \delta u \nabla \cdot (G \nabla \lambda) \,\mathrm{d}\Omega \,\mathrm{d}t - \int_0^T \int_{\Gamma_\mathrm{t,b}} \delta u G \frac{\partial \lambda}{\partial n} \,\mathrm{d}\Gamma \,\mathrm{d}t.
\end{aligned} \tag{74}
$$

where $\frac{\partial (\cdot)}{\partial n}$ denotes a directional derivative of a variable in the direction of an outward unit normal vector $\mathbf{n}$ on Γ. Integrating c in (71) by parts over time twice leads to:

$$
\begin{aligned}
c =& - \int_0^T \int_\Omega \rho \lambda \frac{\partial^2 \delta u}{\partial t^2} \,\mathrm{d}\Omega \,\mathrm{d}t \\
=& - \int_\Omega \rho \left[\lambda \frac{\partial \delta u}{\partial t} \right]_0^T \mathrm{d}\Omega + \int_0^T \int_\Omega \rho \frac{\partial \lambda}{\partial t} \frac{\partial \delta u}{\partial t} \,\mathrm{d}\Omega \,\mathrm{d}t \\
=& - \int_\Omega \rho \left[\lambda \frac{\partial \delta u}{\partial t} \right]_0^T \mathrm{d}\Omega + \int_\Omega \rho \left[\frac{\partial \lambda}{\partial t} \delta u \right]_0^T \mathrm{d}\Omega - \int_0^T \int_\Omega \rho \frac{\partial^2 \lambda}{\partial t^2} \delta u \,\mathrm{d}\Omega \,\mathrm{d}t,
\end{aligned} \tag{75}
$$

Because $\frac{\partial \delta u}{\partial t}(t=0)$ and $\delta u(t=0)$ vanish, (75) becomes:

$$
c = - \int_\Omega \left[\rho \lambda \frac{\partial \delta u}{\partial t} \right]_T \mathrm{d}\Omega + \int_\Omega \left[\rho \frac{\partial \lambda}{\partial t} \delta u \right]_T \mathrm{d}\Omega - \int_0^T \int_\Omega \rho \frac{\partial^2 \lambda}{\partial t^2} \delta u \,\mathrm{d}\Omega \,\mathrm{d}t. \tag{76}
$$

Due to (72), (74), and (76), (71) can be written as:

$$
\begin{aligned}
\delta_u \mathcal{A} = & \int_0^T \int_{\Gamma_\mathrm{b}} (\lambda_F - \lambda) G \frac{\partial \delta u}{\partial y} \,\mathrm{d}\Gamma \,\mathrm{d}t \\
& + \int_0^T \int_{\Gamma_\mathrm{l,r}} \lambda G \frac{\partial \delta u}{\partial n} \,\mathrm{d}\Gamma \,\mathrm{d}t - \int_0^T \int_{\Gamma_\mathrm{t,b}} \delta u G \frac{\partial \lambda}{\partial n} \,\mathrm{d}\Gamma \,\mathrm{d}t \\
& - \int_\Omega \left[\rho \lambda \frac{\partial \delta u}{\partial t} \right]_T \mathrm{d}\Omega + \int_\Omega \left[\rho \frac{\partial \lambda}{\partial t} \delta u \right]_T \mathrm{d}\Omega \\
& + \int_0^T \int_\Omega \delta u \left[-2(u_\mathrm{m} - u) \sum_{i=1}^{N_\mathrm{s}} \Delta(x - x_i, y - y_i) + \nabla \cdot (G \nabla \lambda) - \rho \frac{\partial^2 \lambda}{\partial t^2} \right] \mathrm{d}\Omega \,\mathrm{d}t = 0. \qquad (77)
\end{aligned}
$$

$\delta_u \mathcal{A} = 0$ in (77) is satisfied when we satisfy the adjoint problem in (15) to (17).

APPENDIX II. DERIVATION OF THE CONTROL EQUATION

The variation of $\mathcal{A}$ with respect to ξ should vanish as:

$$\begin{aligned}
\delta_\xi \mathcal{A} = \frac{\partial \mathcal{A}}{\partial \xi} = & \underbrace{-\int_0^T \sum_{i=1}^{N_s} \left[2(u_{mi} - u_i) \left(\frac{\partial u_i}{\partial \xi} \right) \right] \mathrm{d}t}_{e} + \underbrace{\int_0^T \int_\Omega \frac{\partial \lambda}{\partial \xi} \left[\nabla \cdot (G \nabla u) - \rho \frac{\partial^2 u}{\partial t^2} \right] \mathrm{d}\Omega \, dt}_{=0} \\
& + \underbrace{\int_0^T \int_{\Gamma_b} \frac{\partial \lambda_F}{\partial \xi} \left[G \frac{\partial u}{\partial y} - F(x,t) \right] \mathrm{d}\Gamma \, \mathrm{d}t}_{=0} + \underbrace{\int_0^T \int_\Omega \lambda \nabla \cdot \left(G \nabla \left(\frac{\partial u}{\partial \xi} \right) \right) \mathrm{d}\Omega \, \mathrm{d}t}_{f} \\
& \underbrace{- \int_0^T \int_\Omega \rho \lambda \frac{\partial^2}{\partial t^2} \left(\frac{\partial u}{\partial \xi} \right) \mathrm{d}\Omega \, \mathrm{d}t}_{g} \\
& + \int_0^T \int_{\Gamma_b} \lambda_F G \frac{\partial}{\partial y} \left(\frac{\partial u}{\partial \xi} \right) \mathrm{d}\Gamma \, \mathrm{d}t - \int_0^T \int_{\Gamma_b} \lambda_F \frac{\partial}{\partial \xi} (F(x,t)) \, \mathrm{d}\Gamma \, \mathrm{d}t \\
& + \frac{\partial \mathcal{R}^{\mathrm{TN}}}{\partial \xi} = 0.
\end{aligned} \tag{78}$$

Part e in (78) can be written as:

$$e = -\int_0^T \int_\Omega 2(u_m - u) \left(\frac{\partial u}{\partial \xi} \right) \sum_{i=1}^{N_s} \Delta(x - x_i, y - y_i) \, \mathrm{d}\Omega \, \mathrm{d}t. \tag{79}$$

Integrating f in (78) by parts twice over space leads to:

$$\begin{aligned}
f = & -\int_0^T \int_{\Gamma_b} \lambda G \frac{\partial}{\partial y} \left(\frac{\partial u}{\partial \xi} \right) \mathrm{d}\Gamma \, \mathrm{d}t + \int_0^T \int_{\Gamma_{l,r}} \lambda G \frac{\partial}{\partial n} \left(\frac{\partial u}{\partial \xi} \right) \mathrm{d}\Gamma \, \mathrm{d}t \\
& + \int_0^T \int_\Omega \left(\frac{\partial u}{\partial \xi} \right) \nabla \cdot (G \nabla \lambda) \, \mathrm{d}\Omega \, \mathrm{d}t - \int_0^T \int_{\Gamma_{t,b}} \left(\frac{\partial u}{\partial \xi} \right) G \frac{\partial \lambda}{\partial n} \, \mathrm{d}\Gamma \, \mathrm{d}t.
\end{aligned} \tag{80}$$

By integrating g in (78) by parts over time twice and knowing that $\frac{\partial}{\partial t} \frac{\partial u}{\partial \xi}(t = 0)$ and $\frac{\partial u}{\partial \xi}(t = 0)$ vanish, we rewrite g in (78) as:

$$g = -\int_\Omega \left[\rho \lambda \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial \xi} \right) \right]_T \mathrm{d}\Omega + \int_\Omega \left[\rho \frac{\partial \lambda}{\partial t} \left(\frac{\partial u}{\partial \xi} \right) \right]_T \mathrm{d}\Omega - \int_0^T \int_\Omega \rho \frac{\partial^2 \lambda}{\partial t^2} \left(\frac{\partial u}{\partial \xi} \right) \mathrm{d}\Omega \, \mathrm{d}t. \tag{81}$$

Due to (79), (80), and (81), (78) becomes:

$$
\begin{aligned}
\delta_\xi \mathcal{A} = &\int_0^T \int_{\Gamma_\mathrm{b}} \underbrace{(\lambda_F - \lambda)}_{=0} G \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial \xi}\right) \mathrm{d}\Gamma \,\mathrm{d}t \\
&+ \int_0^T \int_{\Gamma_\mathrm{l,r}} \underbrace{\lambda}_{=0} G \frac{\partial}{\partial n}\left(\frac{\partial u}{\partial \xi}\right) \mathrm{d}\Gamma \,\mathrm{d}t - \int_0^T \int_{\Gamma_\mathrm{t,b}} \left(\frac{\partial u}{\partial \xi}\right) G \underbrace{\frac{\partial \lambda}{\partial n}}_{=0} \mathrm{d}\Gamma \,\mathrm{d}t \\
&- \int_\Omega \left[\rho \underbrace{\lambda}_{=0} \frac{\partial}{\partial t}\left(\frac{\partial u}{\partial \xi}\right)\right]_T \mathrm{d}\Omega + \int_\Omega \left[\rho \underbrace{\frac{\partial \lambda}{\partial t}}_{=0} \left(\frac{\partial u}{\partial \xi}\right)\right]_T \mathrm{d}\Omega \\
&+ \int_0^T \int_\Omega \frac{\partial u}{\partial \xi} \underbrace{\left[-2(u_\mathrm{m} - u) \sum_{i=1}^{N_\mathrm{s}} \Delta(x - x_i, y - y_i) + \nabla \cdot (G \nabla \lambda) - \rho \frac{\partial^2 \lambda}{\partial t^2}\right]}_{=0} \mathrm{d}\Omega \,\mathrm{d}t \\
&- \int_0^T \int_{\Gamma_\mathrm{b}} \lambda_F \frac{\partial}{\partial \xi}(F(x,t)) \,\mathrm{d}\Gamma \,\mathrm{d}t + \frac{\partial \mathcal{R}^\mathrm{TN}}{\partial \xi} = 0.
\end{aligned} \tag{82}
$$

Eq. (82) reduces to:

$$
\delta_\xi \mathcal{A} = -\int_0^T \int_{\Gamma_\mathrm{b}} \lambda \frac{\partial}{\partial \xi}(F(x,t)) \,\mathrm{d}\Gamma \,\mathrm{d}t + \frac{\partial \mathcal{R}^\mathrm{TN}}{\partial \xi} = 0. \tag{83}
$$

The first term of (83) is:

$$
-\int_0^T \int_{\Gamma_\mathrm{b}} \lambda \frac{\partial}{\partial \xi}(F(x,t)) \,\mathrm{d}\Gamma \,\mathrm{d}t = \underbrace{-\int_0^T \int_{\Gamma_\mathrm{b}} \lambda \Phi_k(x) \phi_j(t) \,\mathrm{d}\Gamma \,\mathrm{d}t}_{\text{the first term of (18)}}, \tag{84}
$$

In addition, the second term in (83) is:

$$
\begin{aligned}
\frac{\partial \mathcal{R}^{\mathrm{TN}}}{\partial \xi} =& \frac{\partial}{\partial \xi}\left(\frac{R}{2}\int_0^T \int_{\Gamma_\mathrm{b}} \left(\frac{\partial F(x,t)}{\partial x}\right)^2 + \left(\frac{\partial F(x,t)}{\partial t}\right)^2 \mathrm{d}\Gamma\,\mathrm{d}t\right) \\
=& R\int_0^T \int_{\Gamma_\mathrm{b}} \left(\frac{\partial F}{\partial x}\frac{\partial \hat{F}}{\partial x}\right) + \left(\frac{\partial F}{\partial t}\frac{\partial \hat{F}}{\partial t}\right) \mathrm{d}\Gamma\,\mathrm{d}dt \\
=& \underbrace{\left[R\int_0^T \frac{\partial F}{\partial x}\hat{F}\,\mathrm{d}t\right]_0^L}_{=0} - R\int_0^T \int_{\Gamma_\mathrm{b}} \frac{\partial^2 F}{\partial x^2}\hat{F}\,\mathrm{d}\Gamma\,\mathrm{d}t \\
&+ \underbrace{\left[R\int_{\Gamma_\mathrm{b}} \frac{\partial F}{\partial t}\hat{F}\,\mathrm{d}\Gamma\right]_0^T}_{=0} - R\int_0^T \int_{\Gamma_\mathrm{b}} \frac{\partial^2 F}{\partial t^2}\hat{F}\,\mathrm{d}\Gamma\,\mathrm{d}t \\
=& \underbrace{-R\int_0^T \int_{\Gamma_\mathrm{b}} \left(\frac{\partial^2 F(x,t)}{\partial x^2} + \frac{\partial^2 F(x,t)}{\partial t^2}\right)\Phi_k(x)\phi_j(t)\,\mathrm{d}\Gamma\,\mathrm{d}t,}_{\text{the second term of (18)}}
\end{aligned}
\tag{85}
$$

where $\hat{F}$ is defined as:

$$
\hat{F} = \frac{\partial F(x,t)}{\partial \xi} = \Phi_k(x)\phi_j(t), \tag{86}
$$

and we enforce that:

$$
\begin{aligned}
\frac{\partial F}{\partial x} &= 0, \quad \text{at } x = 0, L, \\
\frac{\partial F}{\partial t} &= 0, \quad \text{at } t = 0, T.
\end{aligned}
\tag{87}
$$

APPENDIX III. ON THE QUADRATIC AND CONVEX OBJECTIVE FUNCTIONAL

In this section, we prove that the objective functional is quadratic and convex. From the compact form of the state problem, $\hat{\mathbf{u}}$ can be defined as:

$$\hat{\mathbf{u}} = \mathbf{Q}^{-1}\hat{\mathbf{F}}. \tag{88}$$

Due to (88), the discrete objective functional ($\hat{\mathcal{L}}$) can be written as:

$$\begin{aligned}
\hat{\mathcal{L}} &= (\hat{\mathbf{u}}_{\mathrm{m}} - \hat{\mathbf{u}})^{\mathrm{T}}\,\overline{\mathbf{B}}\,(\hat{\mathbf{u}}_{\mathrm{m}} - \hat{\mathbf{u}}) \\
&= (\hat{\mathbf{u}}_{\mathrm{m}} - \mathbf{Q}^{-1}\hat{\mathbf{F}})^{\mathrm{T}}\,\overline{\mathbf{B}}\,(\hat{\mathbf{u}}_{\mathrm{m}} - \mathbf{Q}^{-1}\hat{\mathbf{F}}) \\
&= (\hat{\mathbf{u}}_{\mathrm{m}}^{\mathrm{T}} - \hat{\mathbf{F}}^{\mathrm{T}}(\mathbf{Q}^{-1})^{\mathrm{T}})\,\overline{\mathbf{B}}\,(\hat{\mathbf{u}}_{\mathrm{m}} - \mathbf{Q}^{-1}\hat{\mathbf{F}}) \\
&= \underbrace{\hat{\mathbf{u}}_{\mathrm{m}}^{\mathrm{T}}\,\overline{\mathbf{B}}\,\hat{\mathbf{u}}_{\mathrm{m}}}_{a}\,\underbrace{-\,\hat{\mathbf{F}}^{\mathrm{T}}(\mathbf{Q}^{-1})^{\mathrm{T}}\,\overline{\mathbf{B}}\,\hat{\mathbf{u}}_{\mathrm{m}}}_{b}\,\underbrace{-\,\hat{\mathbf{u}}_{\mathrm{m}}^{\mathrm{T}}\,\overline{\mathbf{B}}\,\mathbf{Q}^{-1}\hat{\mathbf{F}}}_{c}\,\underbrace{+\,\hat{\mathbf{F}}^{\mathrm{T}}(\mathbf{Q}^{-1})^{\mathrm{T}}\,\overline{\mathbf{B}}\,\mathbf{Q}^{-1}\hat{\mathbf{F}}}_{d},
\end{aligned} \tag{89}$$

where the part a is a constant, and the parts b and c are linear functions so that their Hessians vanish. The part d in (89) is a quadratic function. Therefore, proving that the part d in (89) is convex will show that $\hat{\mathcal{L}}$ is convex.

To this end, we consider the following simple example—a homogeneous square-shaped solid domain, of which extent is 60 m × 60 m with its shear wave speed of 400 m/s and mass density of 1000 kg/m^3. The shear stress is applied on the bottom surface, and one sensor is placed in the middle of top surface (see Fig. 23). The solid is constrained by fixed boundary conditions on the left and right boundaries, and the element size is 30 m. Therefore, in this case, there are only three degrees of freedom in the discretized domain. We consider 2 time steps, where $\Delta t = 0.1$ s. In such a case, the matrices $\mathbf{M}$, $\mathbf{K}$, and $\mathbf{Q}$ are the followings:

$$\mathbf{M} = \begin{bmatrix} 199980 & 99990 & 0 \\ 99990 & 399960 & 99990 \\ 0 & 99990 & 199980 \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} 213344000 & -53344000 & 0 \\ -53344000 & 426688000 & -53344000 \\ 0 & -53344000 & 213344000 \end{bmatrix}, \tag{90}$$

and

$$\mathbf{Q} = \begin{bmatrix} \mathbf{I} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & 0 & 0 & 0 \\ \mathbf{K} & 0 & \mathbf{M} & 0 & 0 & 0 \\ \mathbf{L}_1 & \mathbf{L}_2 & \mathbf{L}_3 & \mathbf{Keff} & 0 & 0 \\ a_1\mathbf{I} & \mathbf{I} & 0 & -a_1\mathbf{I} & \mathbf{I} & 0 \\ a_0\mathbf{I} & a_2\mathbf{I} & \mathbf{I} & -a_0\mathbf{I} & 0 & \mathbf{I} \end{bmatrix}, \tag{91}$$

where:

$$a_0 = \frac{4}{(\Delta t)^2}, \;\; a_1 = \frac{2}{\Delta t}, \;\; a_2 = \frac{4}{\Delta t}, \tag{92}$$

$$\mathbf{Keff} = a_0\mathbf{M} + \mathbf{K}, \tag{93}$$

$$\mathbf{L}_1 = -a_0\mathbf{M}, \;\; \mathbf{L}_2 = -a_2\mathbf{M}, \;\; \mathbf{L}_3 = -\mathbf{M}. \tag{94}$$

The vector $\hat{\mathbf{F}}$ is built as:

$$\hat{\mathbf{F}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{F_0} \\ \mathbf{F_1} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{F_0} = \begin{bmatrix} F_{30,0} \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{F_1} = \begin{bmatrix} F_{30,1} \\ 0 \\ 0 \end{bmatrix}, \tag{95}$$

where $F_{30,0}$ and $F_{30,1}$ are the components of a force vector at the node of $x = 30$ m and at the timesteps of $t = 0$ s and $t = 0.1$ s, respectively. In addition, the block diagonal matrix $\overline{\mathbf{B}}$ is defined as:

$$\overline{\mathbf{B}} = \Delta t \begin{bmatrix} \mathbf{B_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{B_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{B_1} = \mathbf{B_2} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{96}$$

Therefore, in this case, due to (90)-(96), the part d in (89) can be written as:

$$\begin{aligned} d &= \hat{\mathbf{F}}^{\mathrm{T}} (\mathbf{Q}^{-1})^{\mathrm{T}} \overline{\mathbf{B}}\, \mathbf{Q}^{-1} \hat{\mathbf{F}} \\ &= 1.25 \times 10^{-24} F_{30,0}^2 + 1.25 \times 10^{-24} F_{30,1}^2 + 2.5 \times 10^{-24} F_{30,0} F_{30,1}. \end{aligned} \tag{97}$$

As shown in (97), the part d in (89) is a quadratic function in terms of $F_{30,0}$ and $F_{30,1}$.

We note that the part d in (89) is convex if the eigenvalues of its Hessian matrix are non-negative (greater than or equal to 0). Here, the Hessian matrix of the part d in (89) is:

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 d}{\partial F_{30,0}^2} & \frac{\partial^2 d}{\partial F_{30,0} \partial F_{30,1}} \\ \frac{\partial^2 d}{\partial F_{30,1} \partial F_{30,0}} & \frac{\partial^2 d}{\partial F_{30,1}^2} \end{bmatrix} = \begin{bmatrix} 2.5 \times 10^{-24} & 2.5 \times 10^{-24} \\ 2.5 \times 10^{-24} & 2.5 \times 10^{-24} \end{bmatrix}, \tag{98}$$

of which eigenvalues are obtained as:

$$\mathbf{|I^{*}}\lambda \textbf{ - H|} = \begin{vmatrix} \lambda - 2.5 \times 10^{-24} & 2.5 \times 10^{-24} \\ 2.5 \times 10^{-24} & \lambda - 2.5 \times 10^{-24} \end{vmatrix} = 0. \tag{99}$$

Namely, the eigenvalues are:

$$\lambda_1 = 0, \quad \lambda_2 = 5 \times 10^{-24}. \tag{100}$$

Both eigenvalues are non-negative so that the part d in (89) is convex, and, therefore, the objective functional $\hat{\mathcal{L}}$ is also convex.

REFERENCES

Akcelik, V., Biros, G., and Ghattas, O. (2002). “Parallel multiscale Gauss-Newton-Krylov methods for inverse wave propagation.” *Supercomputing, ACM/IEEE 2002 Conference*, IEEE, 41–41.

Aquino, W., Bunting, G., Miller, S. T., and Walsh, T. F. (2019). “A gradient-based optimization approach for the detection of partially connected surfaces using vibration tests.” *Computer Methods in Applied Mechanics and Engineering*, 345, 323 – 335.

Bielak, J. and Christiano, P. (1984). “On the effective seismic input for non-linear soil-structure interaction systems.” *Earthquake Engineering and Structural Dynamics*, 12, 107–119.

Bielak, J., Loukakis, K., Hisada, Y., and Yoshimura, C. (2003). “Domain reduction method for three-dimensional earthquake modeling in localized regions, Part I: Theory.” *Bulletin of the Seismological Society of America*, 93(2), 817–824.

Binder, F., Schöpfer, F., and Schuster, T. (2015). “Defect localization in fibre-reinforced composites by computing external volume forces from surface sensor measurements.” *Inverse Problems*, 31(2), 025006.

Fathi, A., Kallivokas, L. F., and Poursartip, B. (2015a). “Full-waveform inversion in three-dimensional PML-truncated elastic media.” *Computer Methods in Applied Mechanics and Engineering*, 296, 39–72.

Fathi, A., Poursartip, B., and Kallivokas, L. F. (2015b). “Time-domain hybrid formulations for wave simulations in three-dimensional PML-truncated heterogeneous media.” *International Journal for Numerical Methods in Engineering*, 101(3), 165–198.

Fathi, A., Poursartip, B., Stokoe II, K. H., and Kallivokas, L. F. (2016). “Three-dimensional P-and S-wave velocity profiling of geotechnical sites using full-waveform inversion driven by field data.” *Soil Dynamics and Earthquake Engineering*, 87, 63–81.

Fletcher, R. and Reeves, C. M. (1951). “Function minimization by conjugate gradients.” *Computer Journal*, 7, 149–154.

Goh, H. and Kallivokas, L. F. (2019). “Group velocity-driven inverse metamaterial design.” *Journal of Engineering Mechanics*, 145(12), 04019094.

Guzina, B. B., Fata, S. N., and Bonnet, M. (2003). “On the stress-wave imaging of cavities in a semi-infinite solid.” *International Journal of Solids and Structures*, 40(6), 1505–1523.

Hasanov, A. and Baysal, O. (2014). “Identification of an unknown spatial load distribution in a vibrating cantilevered beam from final overdetermination.” *Journal of Inverse and Ill-posed Problems*, 23(1), 85–102.

Jeong, C., Kallivokas, L., Kucukcoban, S., Deng, W., and Fathi, A. (2015). “Maximization of wave motion within a hydrocarbon reservoir for wave-based enhanced oil recovery.” *Journal of Petroleum Science and Engineering*, 129, 205–220.

Jeong, C. and Kallivokas, L. F. (2016). “An inverse-source problem for maximization of pore-fluid oscillation within poroelastic formations.” *Inverse Problems in Science and Engineering*.

Jeong, C., Kallivokas, L. F., Huh, C., and Lake, L. W. (2010). “Optimization of sources for focusing wave energy in targeted formations.” *Journal of Geophysics and Engineering*, 7, 242–256.

Jeong, C., Na, S.-W., and Kallivokas, L. F. (2009). “Near-surface localization and shape identification of a scatterer embedded in a halfplane using scalar waves.” *Journal of Computational Acoustics*, 17(03), 277–308.

Jeong, C. and Seylabi, E. E. (2018). “Seismic input motion identification in a heterogeneous halfspace.” *Journal of Engineering Mechanics*, 144(8), 04018070.

Jeremić, B., Tafazzoli, N., Ancheta, T., Orbović, N., and Blahoianu, A. (2013). “Seismic behavior of NPP structures subjected to realistic 3D, inclined seismic motions, in variable layered soil/rock, on surface or embedded foundations.” *Nuclear Engineering and Design*, 265, 85–94.

Jung, J., Jeong, C., and Taciroglu, E. (2013). “Identification of a scatterer embedded in elastic heterogeneous media using dynamic XFEM.” *Computer Methods in Applied Mechanics and Engineering*, 259, 50–63.

Kallivokas, L. F., Fathi, A., Kucukcoban, S., Stokoe, K. H., Bielak, J., and Ghattas, O. (2013). “Site characterization using full waveform inversion.” *Soil Dynamics and Earthquake Engineering*, 47, 62–82.

Kang, J. W. and Kallivokas, L. F. (2010a). “The inverse medium problem in 1D pml-truncated heterogeneous semi-infinite domains." *Inverse Problems in Science and Engineering*, 18(6), 759–786.

Kang, J. W. and Kallivokas, L. F. (2010b). "Mixed unsplit-field perfectly matched layers for transient simulations of scalar waves in heterogeneous domains." *Computational Geosciences*, 14(4), 623–648.

Kang, J. W. and Kallivokas, L. F. (2011). "The inverse medium problem in heterogeneous PML-truncated domains using scalar probing waves." *Computer Methods in Applied Mechanics and Engineering*, 200(1), 265 – 283.

Karve, P. M. and Kallivokas, L. F. (2015). "Wave energy focusing to subsurface poroelastic formations to promote oil mobilization." *Geophysical Journal International*, 202(1), 119–141.

Karve, P. M., Kallivokas, L. F., and Manuel, L. (2016). "A framework for assessing the uncertainty in wave energy delivery to targeted subsurface formations." *Journal of Applied Geophysics*, 125, 26–36.

Karve, P. M., Kucukcoban, S., and Kallivokas, L. F. (2015). "On an inverse source problem for enhanced oil recovery by wave motion maximization in reservoirs." *Computational Geosciences*, 19(1), 233–256.

Kucukcoban, S., Goh, H., and Kallivokas, L. F. (2019). "On the full-waveform inversion of lame parameters in semi-infinite solids in plane strain." *International Journal of Solids and Structures*, 164, 104 – 119.

Kucukcoban, S. and Kallivokas, L. F. (2011). "Mixed perfectly-matched-layers for direct transient analysis in 2D elastic heterogeneous media." *Computer Methods in Applied Mechanics and Engineering*, 200, 57–76.

Kucukcoban, S. and Kallivokas, L. F. (2013). "A symmetric hybrid formulation for transient wave simulations in PML-truncated heterogeneous media." *Wave Motion*, 50, 57 – 79.

Lions, J. (1971). *Optimal Control of Systems Governed by Partial Differential Equations*. Springer-Verlag.

Lloyd, F. and Jeong, C. (2018). "Adjoint equation-based inverse-source modeling to reconstruct moving acoustic sources in a 1D heterogeneous solid." *Journal of Engineering Mechanics*, 144(9), 04018089.

Lysmer, J. and Kuhlemeyer, R. L. (1969). "Finite dynamic model for infinite media." *Journal of Engineering Mechanics Division, ASCE*, 95(EM4), 859–877.

Mashayekh, H., Kallivokas, L. F., and Tassoulas, J. L. (2018). "Parameter estimation in layered media using dispersion-constrained inversion." *Journal of Engineering Mechanics*, 144(11), 04018099.

Mejia, L. and Dawson, E. (2006). "Earthquake deconvolution for FLAC." *FLAC and Numerical*.

Nguyen-Tuan, L., Nanthakumar, S. S., and Lahmer, T. (2019). "Identification of multiple flaws in dams using inverse analysis based on hydro-mechanical XFEM and level sets." *Computers and Geotechnics*, 110, 211 – 221.

Pakravan, A., Kang, J. W., and Newtson, C. M. (2016). "A Gauss-Newton full-waveform inversion for material profile reconstruction in viscoelastic semi-infinite solid media." *Inverse Problems in Science and Engineering*, 24(3), 393–421.

Paolucci, R. and Pitilakis, K. (2007). *Seismic Risk Assessment of Underground Structures under Transient Ground Deformations*. Springer Netherlands, Dordrecht, 433–459.

PEER (2000). "Pacific Earthquake Engineering Research Center (PEER) ground motion database, <http://ngawest2.berkeley.edu/site>.

Petryk, H. and Mroz, Z. (1986). "Time derivatives of integrals and functionals defined on varying volume and surface domains." *Archives of Mechanics*.

Poul, M. K. and Zerva, A. (2018a). "Efficient time-domain deconvolution of seismic ground motions using the equivalent-linear method for soil-structure interaction analyses." *Soil Dynamics and Earthquake Engineering*, 112, 138 – 151.

Poul, M. K. and Zerva, A. (2018b). "Nonlinear dynamic response of concrete gravity dams considering the deconvolution process." *Soil Dynamics and Earthquake Engineering*, 109, 324 – 338.

Poul, M. K. and Zerva, A. (2018c). "Time-domain PML formulation for modeling viscoelastic waves with Rayleigh-type damping in an unbounded domain: Theory and application in ABAQUS." *Finite Elements in Analysis and Design*, 152, 1 – 16.

Poursartip, B., Fathi, A., and Kallivokas, L. F. (2017). "Seismic wave amplification by topographic features: A parametric study." *Soil Dynamics and Earthquake Engineering*, 92, 503–527.

Rahnema, H., Mohasseb, S., and JavidSharifi, B. (2016). "2-D soil-structure interaction in time domain by the SBFEM and two non-linear soil models." *Soil Dynamics and Earthquake Engineering*, 88, 152–175.

Schnabel, P. B. (1972). "SHAKE a computer program for earthquake response analysisi of horizontally layered sites." *EERC report*, 72–12.

Tadi, M., Rabitz, H., Sik, K. Y., Askar, A., Prevost, J. H., and McManus, J. B. (1996). "Interior energy focusing within an elasto-plastic material." *International Journal of Solids and Structures*, 33(13), 1891–1901.

Tran, K. T. and McVay, M. (2012). "Site characterization using Gauss-Newton inversion of 2-D full seismic waveform in the time domain." *Soil Dynamics and Earthquake Engineering*, 43, 16–24.

Tripe, R., Kontoe, S., and Wong, T. K. C. (2013). "Slope topography effects on ground motion in the presence of deep soil layers." *Soil Dynamics and Earthquake Engineering*, 50, 72–84.

Tromp, J., Komattisch, D., and Liu, Q. (2008). "Spectral-element and adjoint methods in seismology." *Communications in Computational Physics*, 3(1), 1–32.

Walsh, T., Aquino, W., and Ross, M. (2013). "Source Identification in Acoustics and Structural Mechanics using SIERRA/SD." *Report no.*, Sandia National Laboratories, <https://prod-ng.sandia.gov/techlib-noauth/access-control.cgi/2013/132689.pdf>.

Zhang, W., Seylabi, E. E., and Taciroglu, E. (2019). "An ABAQUS toolbox for soil-structure interaction analysis." *Computers and Geotechnics*, 114, 103143.

Zhu, H., Komatitsch, D., and Tromp, J. (2017). "Radial anisotropy of the North American upper mantle based on adjoint tomography with USArray." *Geophysical Journal International*, 211(1), 349–377.

List of Tables

TABLE 1. Summary of all cases.

Case number	Material profile	Force profile	Approach	# of sensors	I_R	Noise level [%]
1	1	1	OTD	30	0.5	0
2	2	1	OTD	30	0.5	0
3	3	1	OTD	30	0.5	0
4	3	1	DTO	30	0.5	0
5	3	1	OTD	15	0.5	0
6	3	1	OTD	10	0.5	0
7	3	1	OTD	5	0.5	0
8	3	1	OTD	15	1.0	0
9	3	1	OTD	15	0.1	0
10	3	1	OTD	15	0.01	0
11	3	1	OTD	15	0.0	0
12	3	1	OTD	15	0.0	1
13	3	1	OTD	15	0.0	2
14	3	1	OTD	15	0.0	3
15	3	1	OTD	15	0.01	2
16	3	1	OTD	15	0.1	2
17	3	1	OTD	15	0.5	2
18	3	2	OTD	15	0.0	0

List of Figures

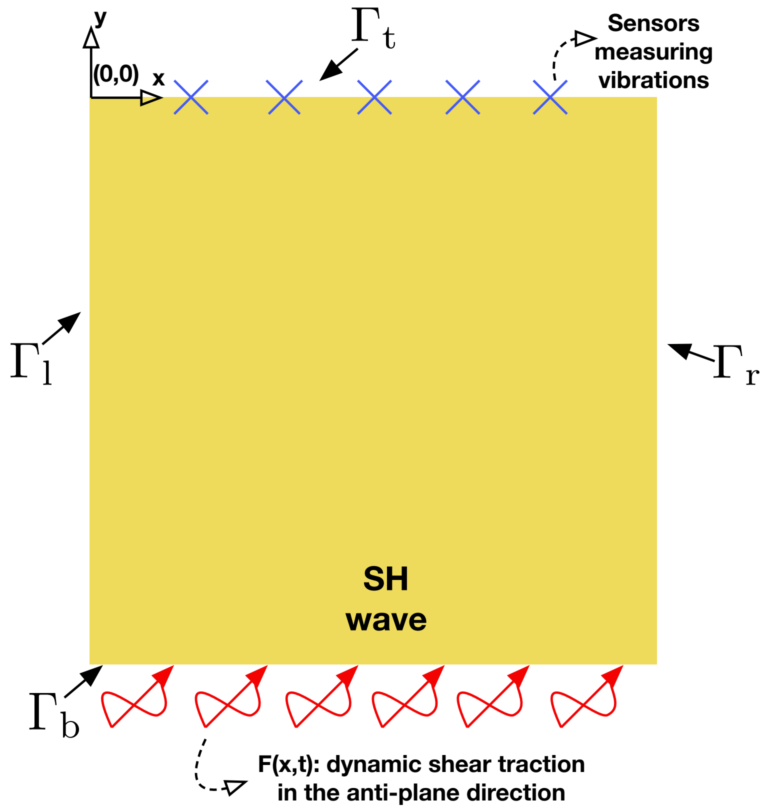


Fig. 1. Problem setting.

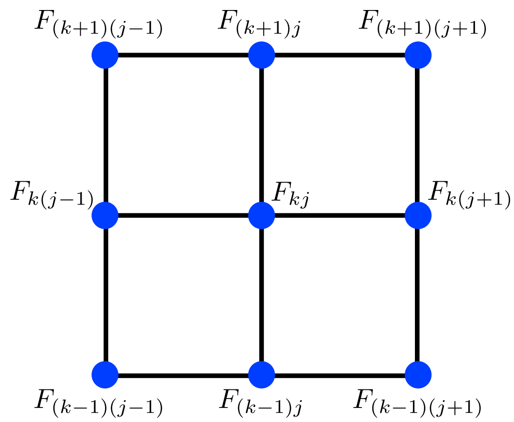


Fig. 2. F_{kj} is surrounded by four elements in the space in terms of x and t: the horizontal and vertical axes represent, respectively, the x coordinate and time t.

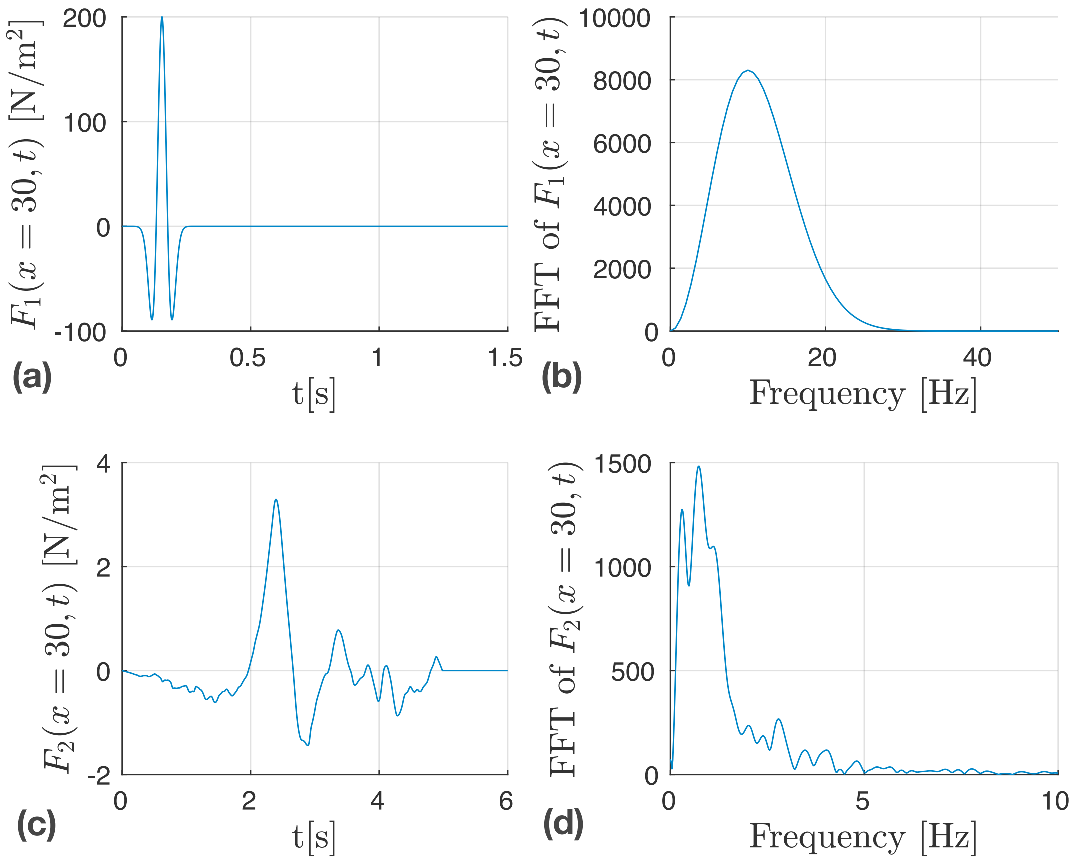


Fig. 3. (a) The time signal of $F_1(x = 30, t)$; (b) the amplitude of Fourier Transform of $F_1(x = 30, t)$; (c) the time signal of $F_2(x = 30, t)$; and (d) the FFT of $F_2(x = 30, t)$.

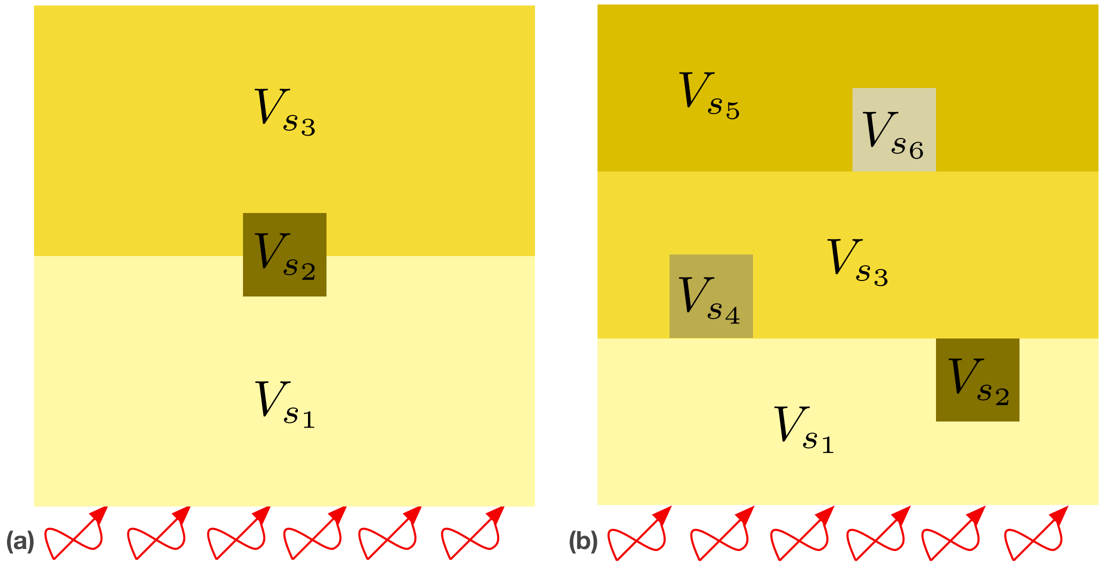


Fig. 4. Heterogeneous solids: (a) Material profile 2; and (b) Material profile 3.

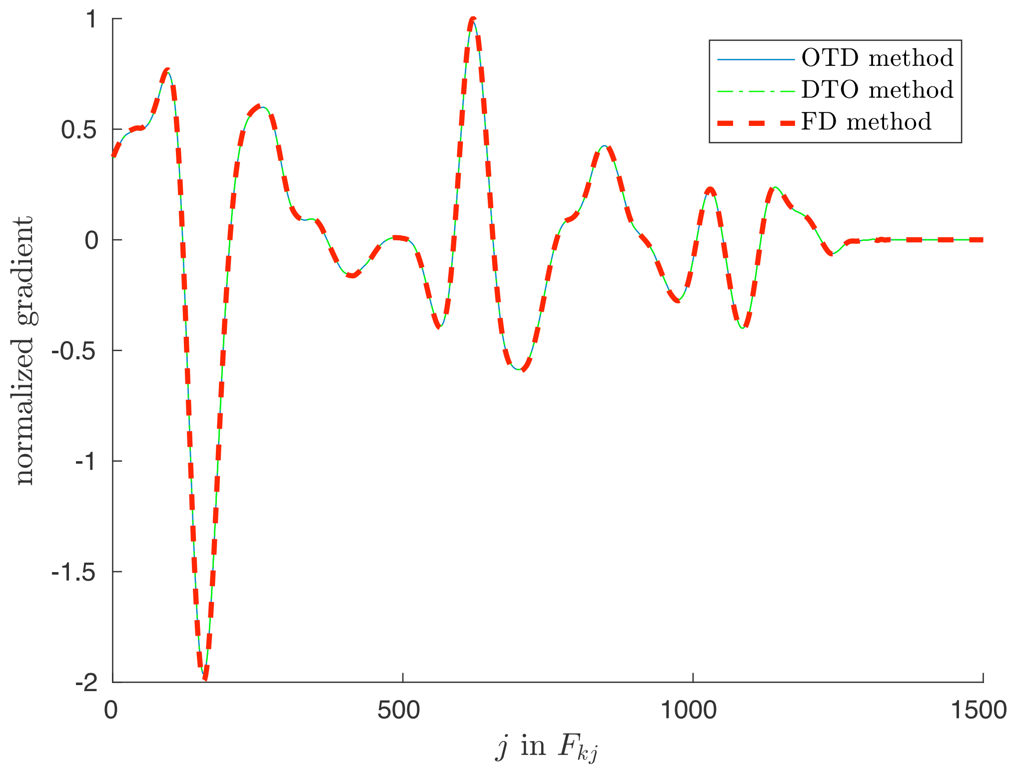


Fig. 5. Comparison between the gradients generated by the OTD approach, DTO approach and the FD approximation.

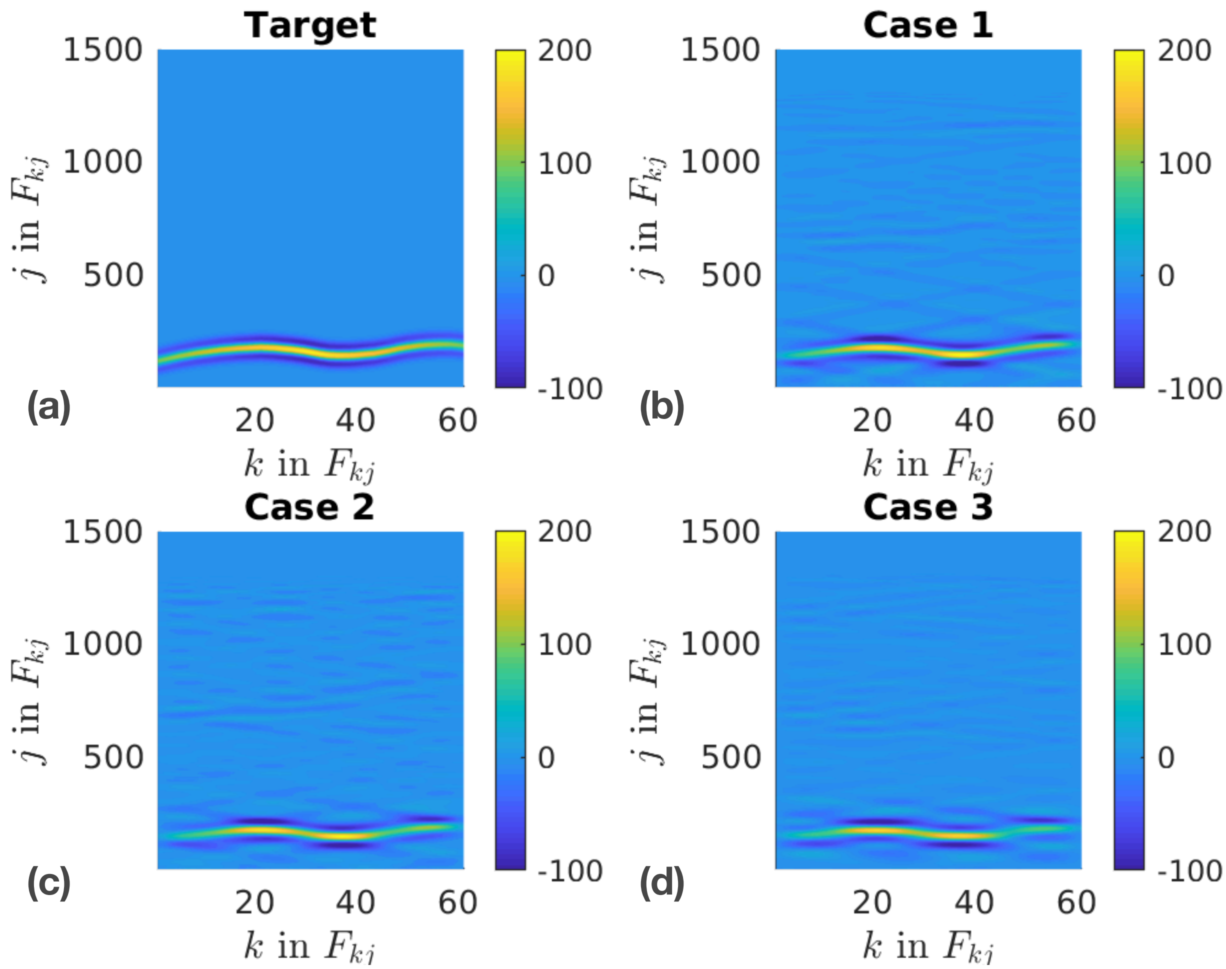


Fig. 6. (a) Target and (b-d) Reconstructed $F_1(x,t)$ for Cases 1-3 at the 1000th iteration. Horizontal and vertical axis in the contour plot represent, respectively, the numbering of the discretized points over space (x) and time (t) of the distribution of $F_1(x,t)$.

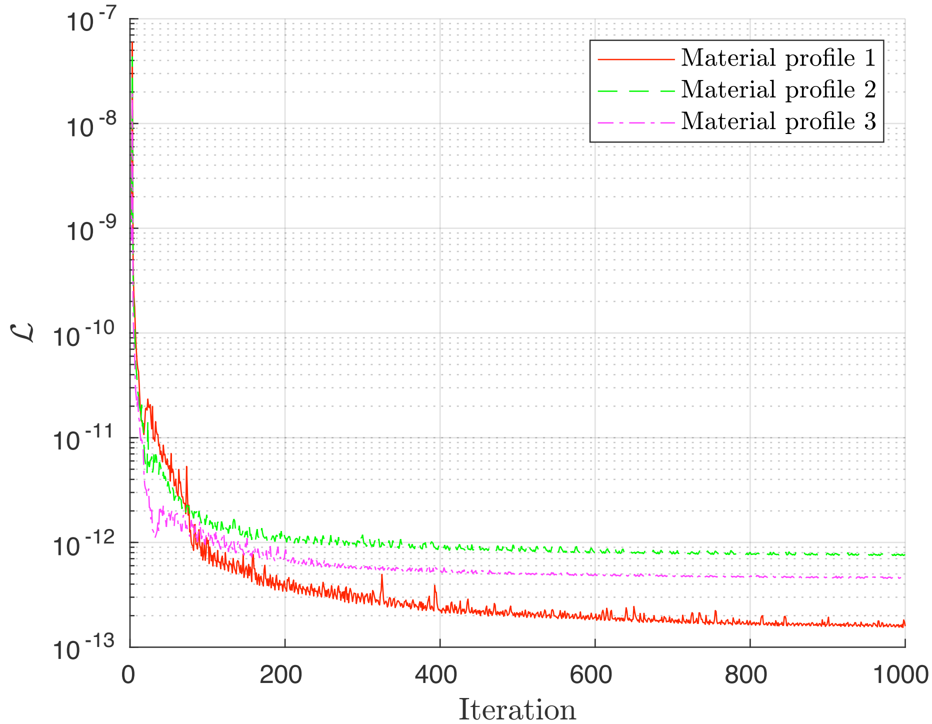


Fig. 7. Example 1 - Objective function, $\mathcal{L}$, versus iterations with respect to the material profiles.

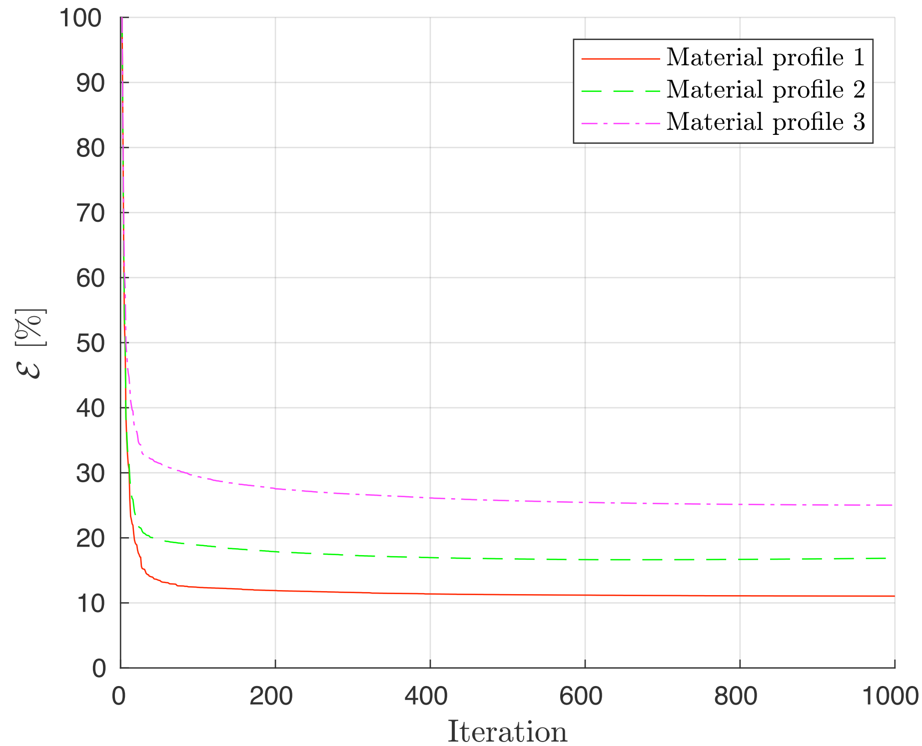


Fig. 8. Example 1 - Error, $\mathcal{E}$, versus iterations with respect to the material profiles.

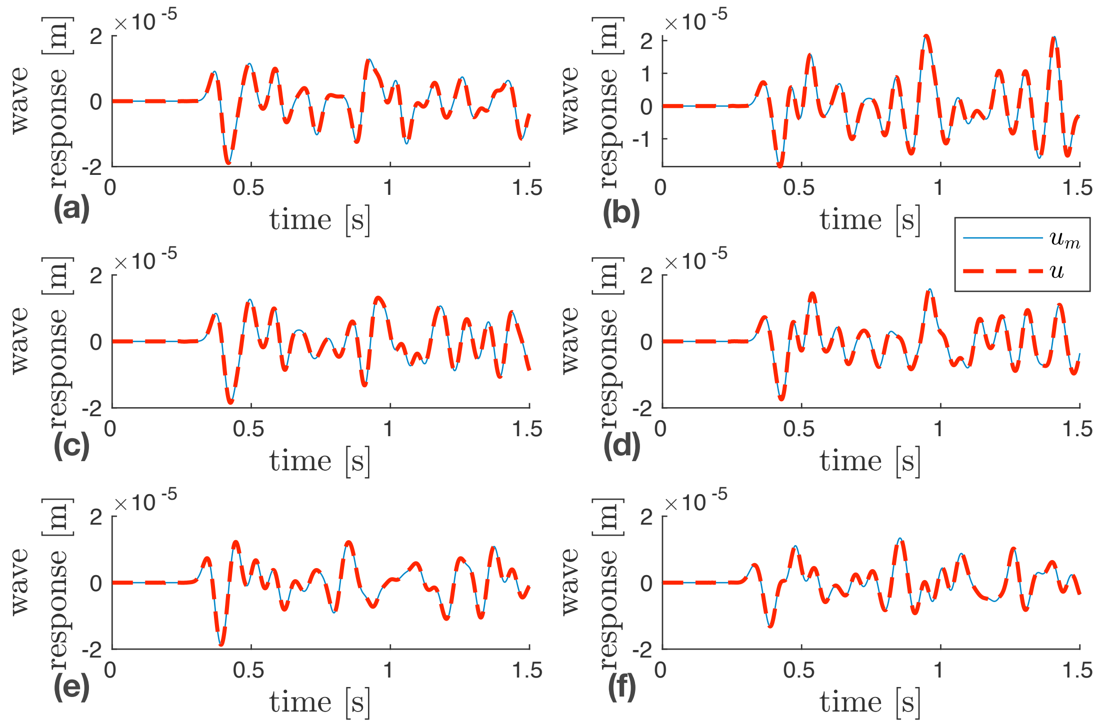
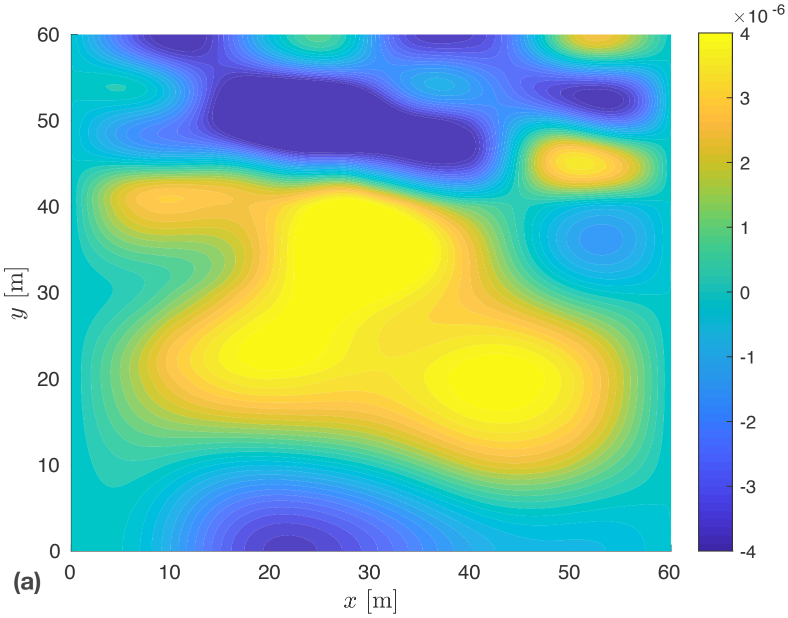


Fig. 9. u_m and u, at sensors placed on the top surface. (a-b) measured at x = 20 m and x = 40 m, respectively, for Case 1. (c-d) measured at x = 20 m and x = 40 m, respectively, for Case 2. (e-f) measured at x = 10 m and x = 20 m, respectively, for Case 3.



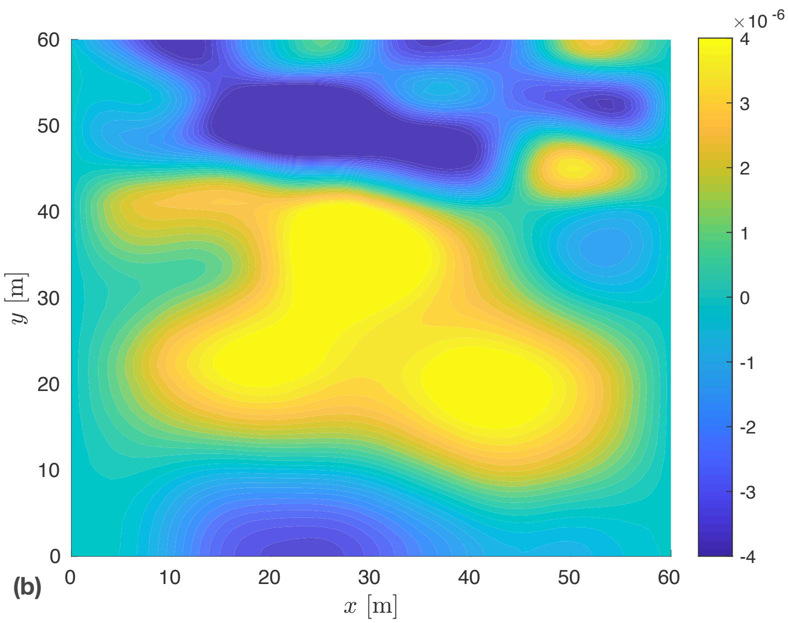


Fig. 10. Wave responses, $u(x, y, t)$, at 1.5 seconds due to (a) target traction and (b) reconstructed traction in Case 3.

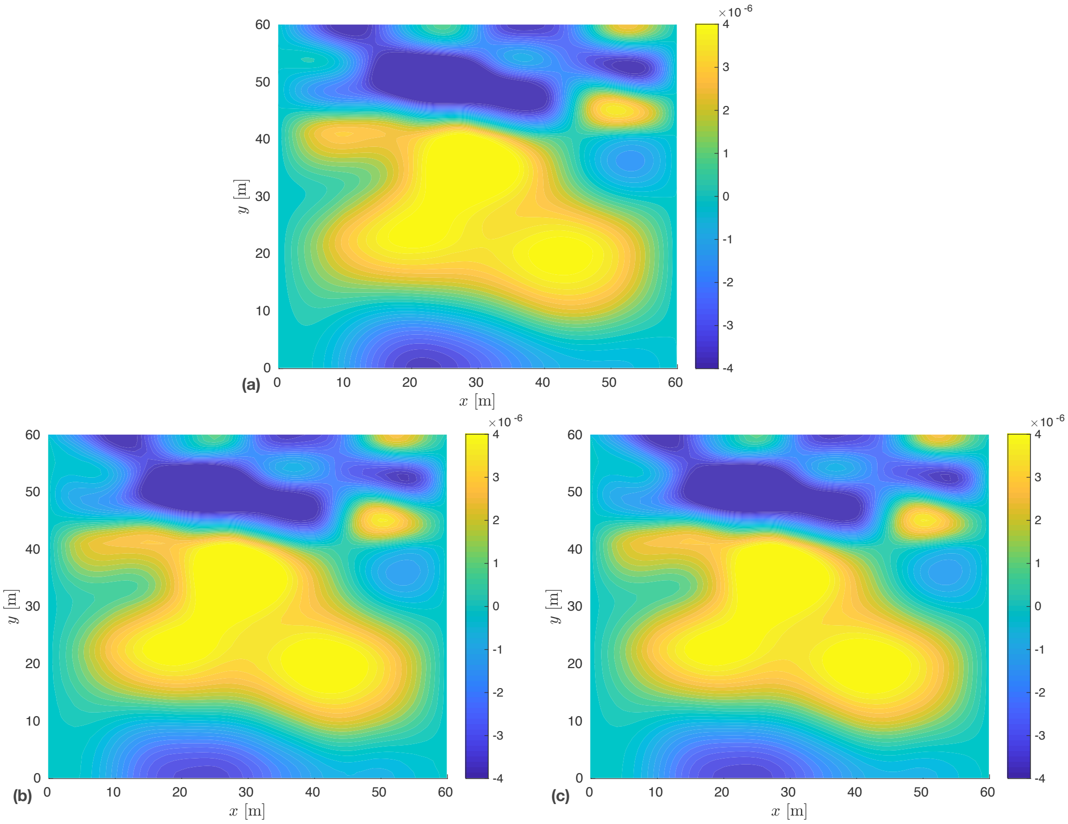
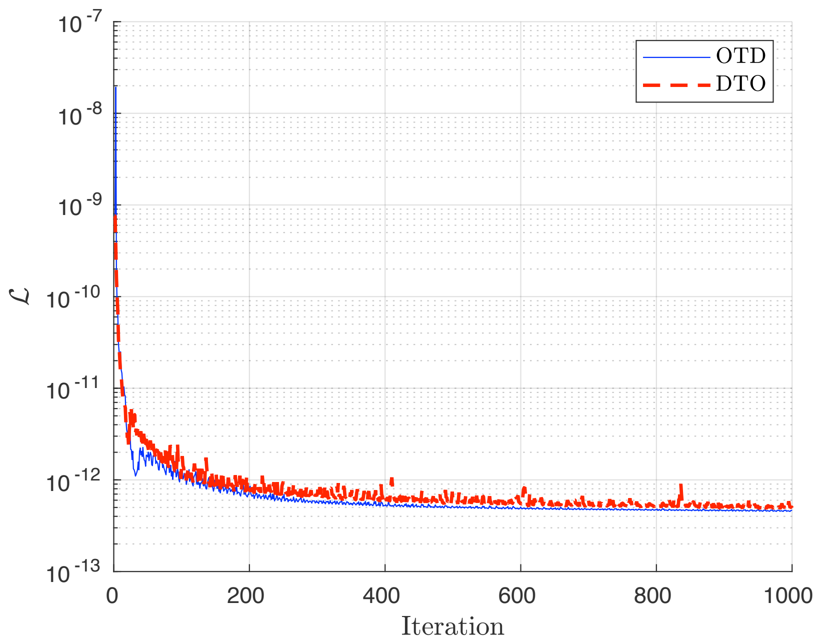
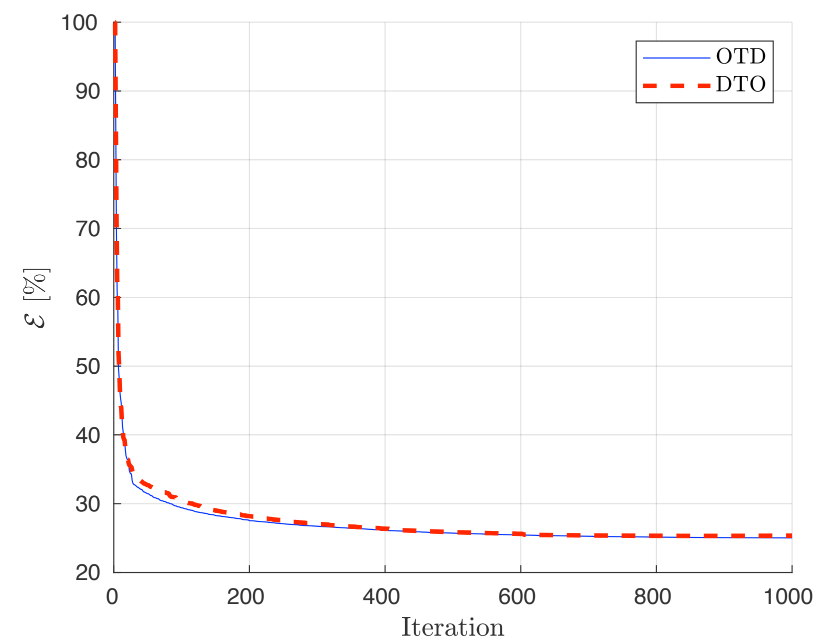


Fig. 11. (a) Targeted $F_1(x,t)$ and (b,c) Reconstructed $F_1(x,t)$ for the OTD and DTO approaches, respectively, at the 1000th iteration.

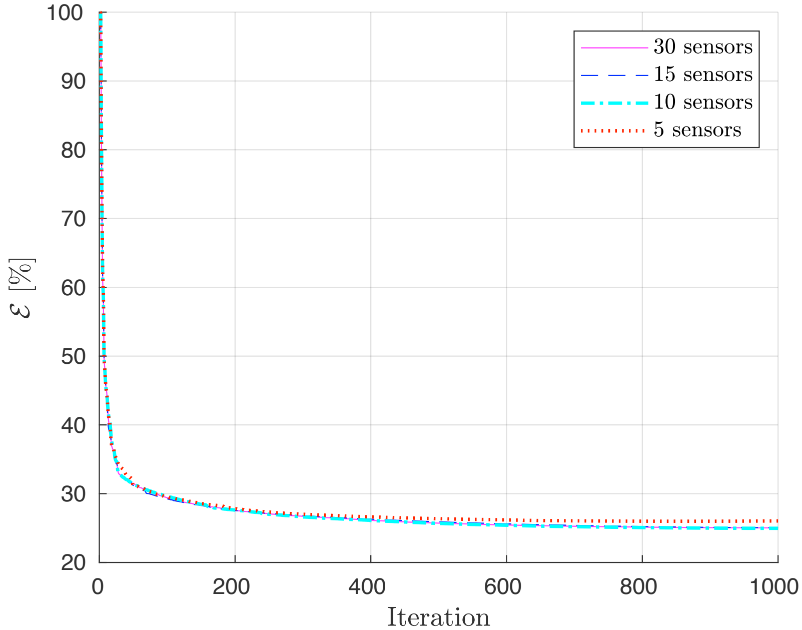


(a) Misfit, $\mathcal{L}$, versus iterations

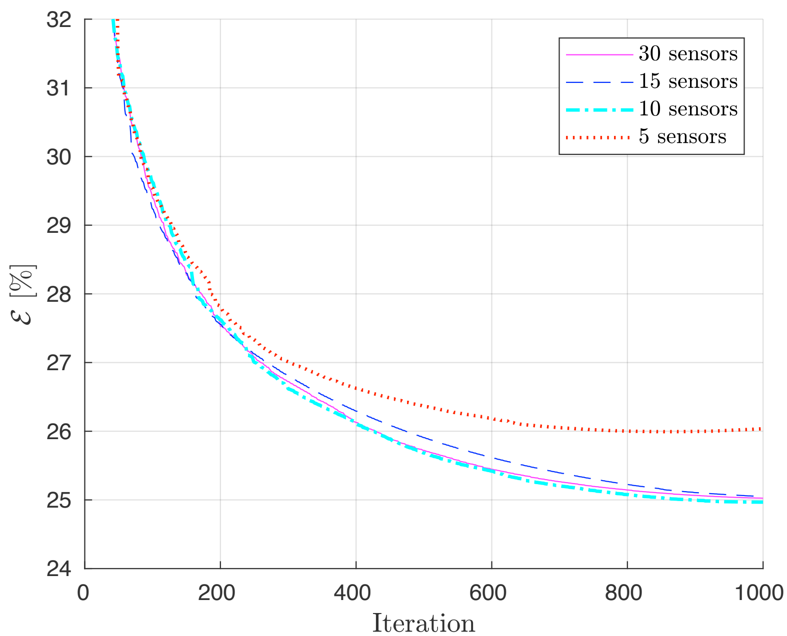


(b) Error, $\mathcal{E}$, versus iterations

Fig. 12. Example 2 - $\mathcal{L}$ and $\mathcal{E}$ with respect to the inverse approach.



(a) Overall scale



(b) Close-up scale

Fig. 13. Example 3 - Error, $\mathcal{E}$, versus iterations with respect to the number of sensors.

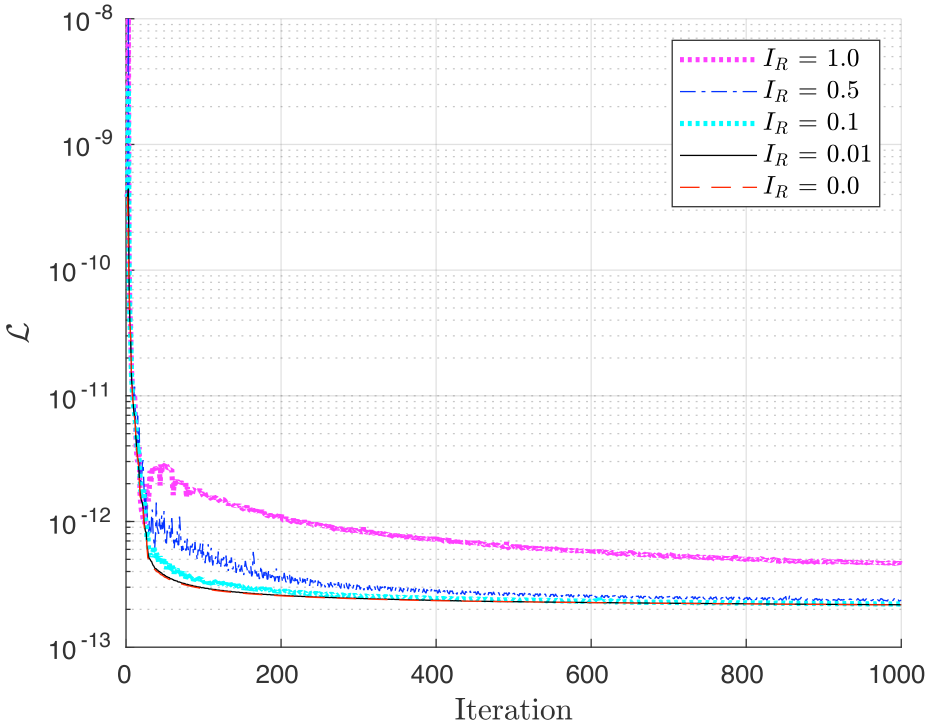
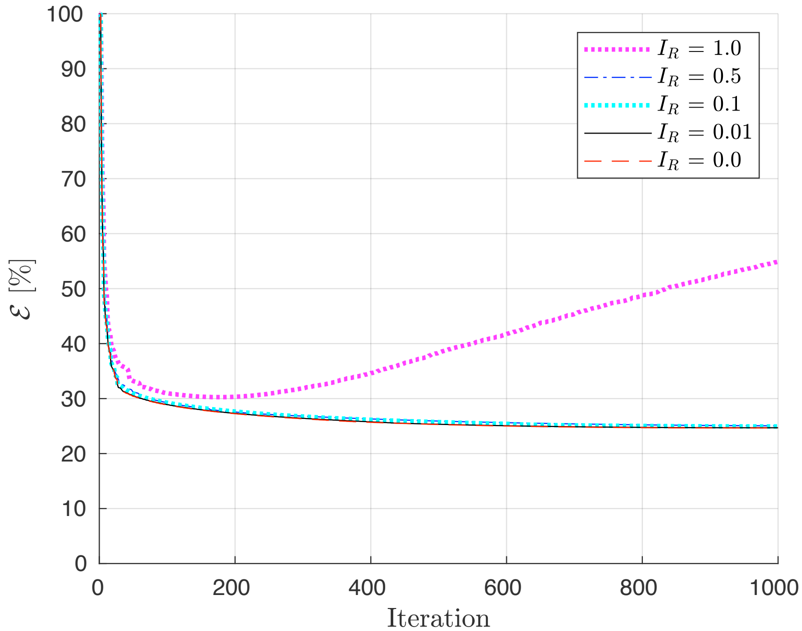
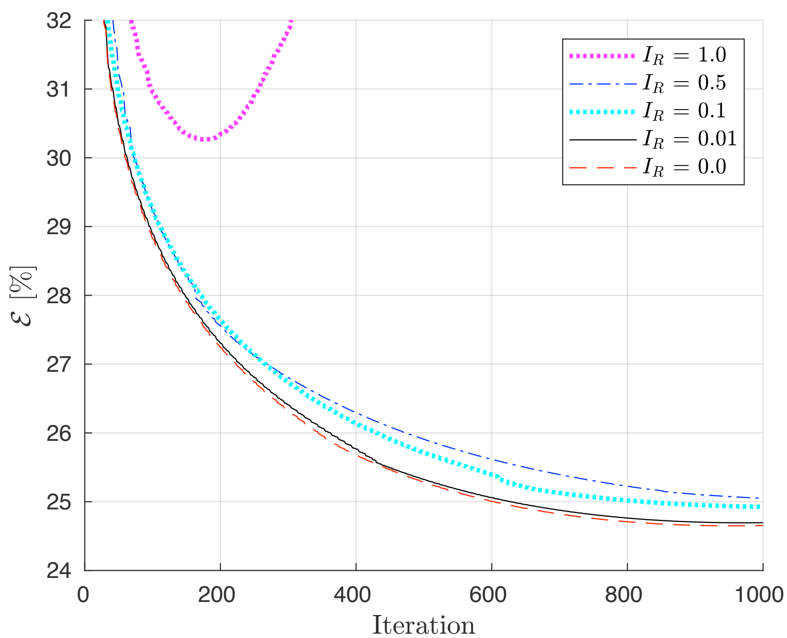


Fig. 14. Example 4 - Objective functional, $\mathcal{L}$, versus iterations with respect to I_R.



(a) Overall scale



(b) Close-up scale

Fig. 15. Example 4 - Error, $\mathcal{E}$, versus iterations with respect to I_R.

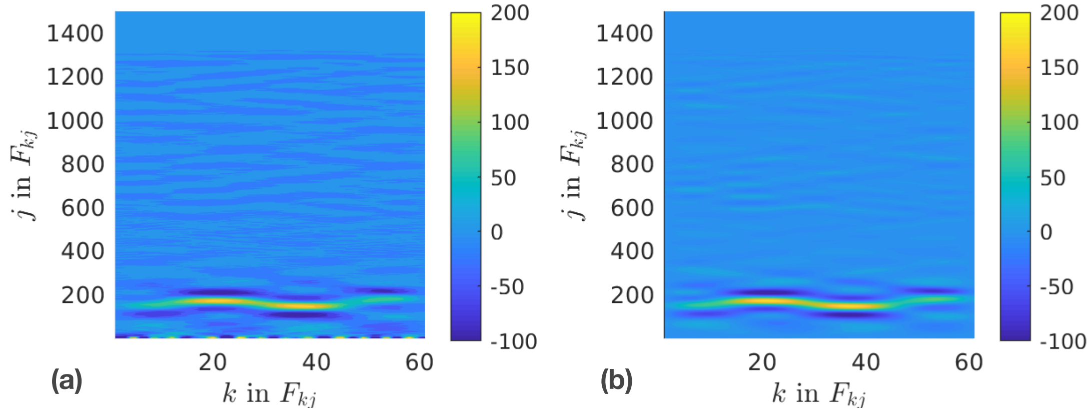
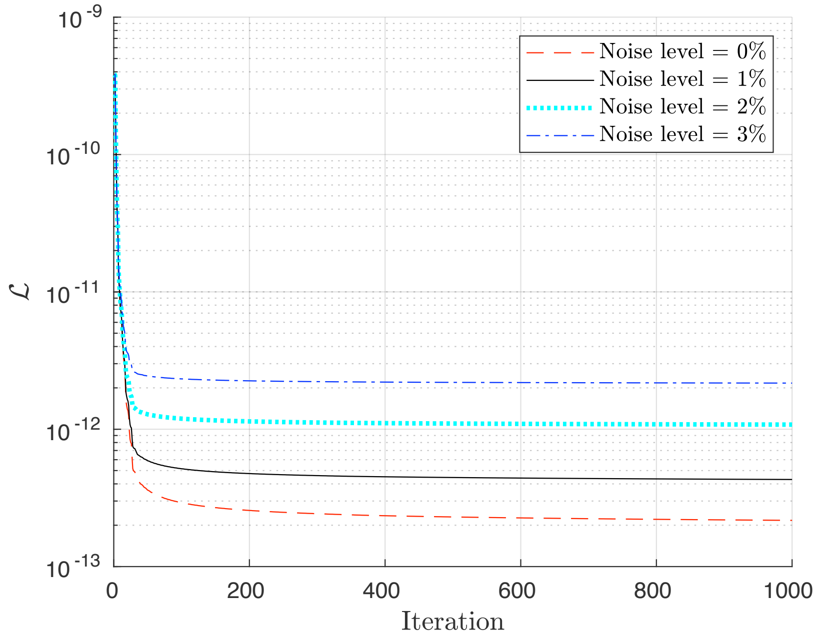
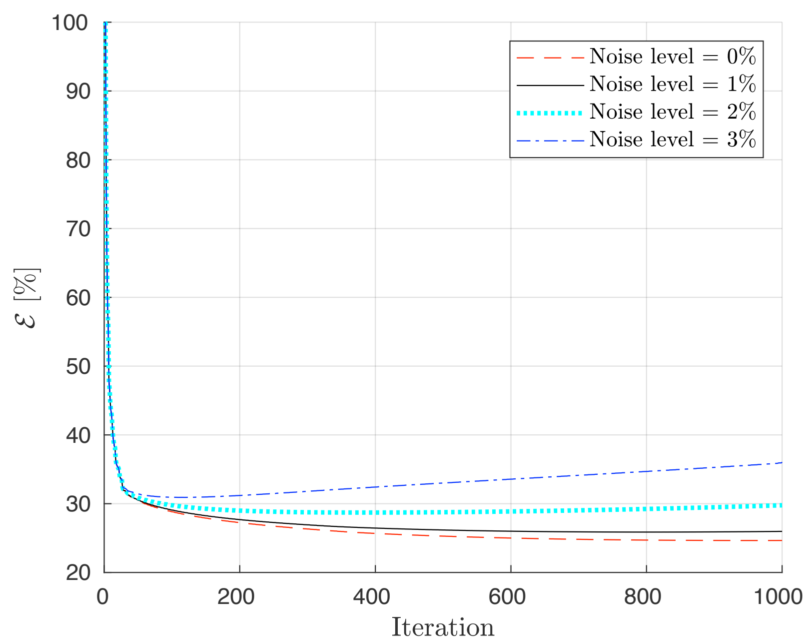


Fig. 16. Example 4 - Estimated traction $F_1(x,t)$ for (a) Case 8, where $I_R = 1.0$, and (b) Case 11, where $I_R = 0.0$

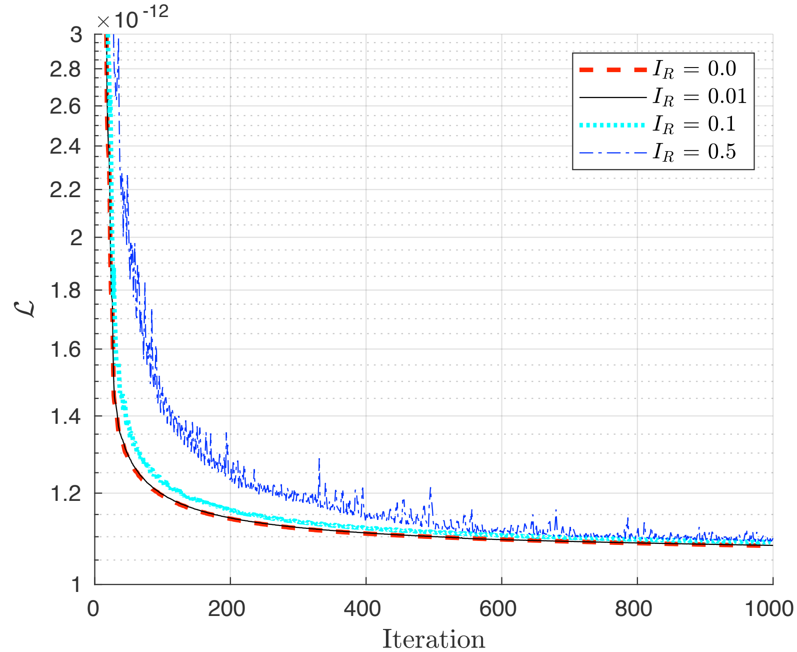


(a) Misfit, $\mathcal{L}$, versus iterations

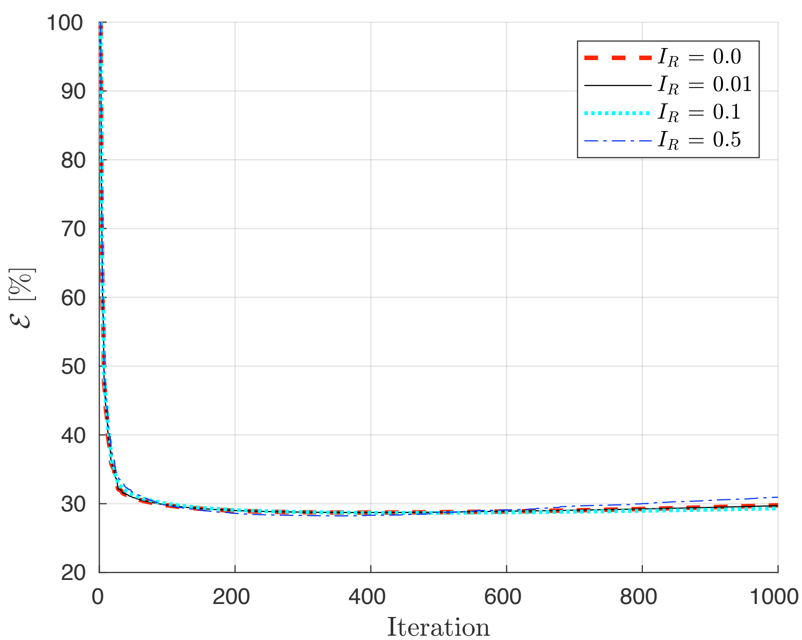


(b) Error, $\mathcal{E}$, versus iterations

Fig. 17. Example 5 - $\mathcal{L}$ and $\mathcal{E}$ with respect to the noise level.



(a) Misfit, $\mathcal{L}$, versus iterations



(b) Error, $\mathcal{E}$, versus iterations

Fig. 18. Example 5 - $\mathcal{L}$ and $\mathcal{E}$ for 2% of noise with respect of I_R.

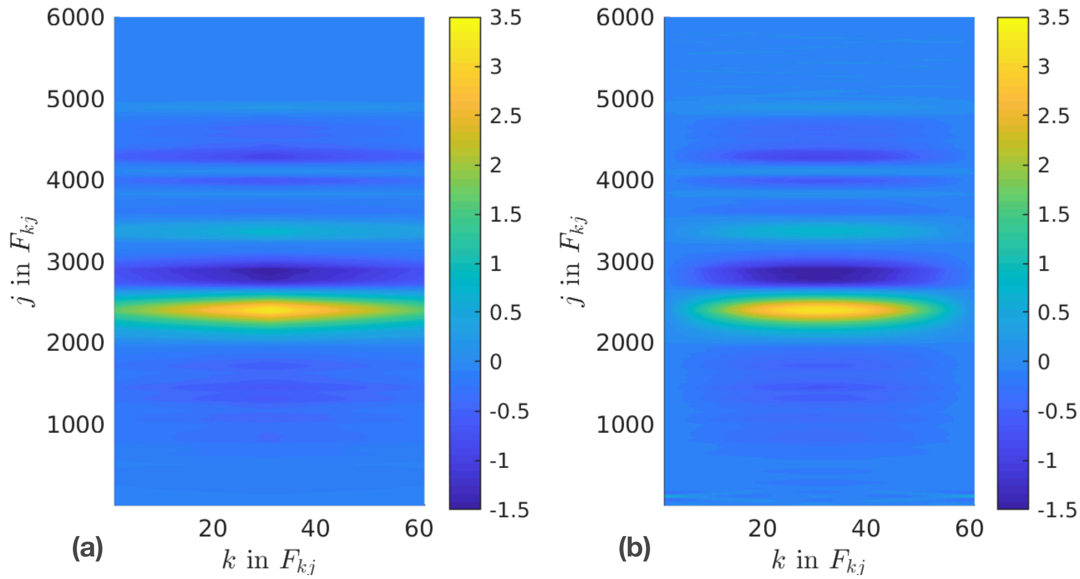


Fig. 19. Example 6 - (a) Target and (b) Reconstructed $F_2(x,t)$ for Case 18 at the 6000-th iteration.

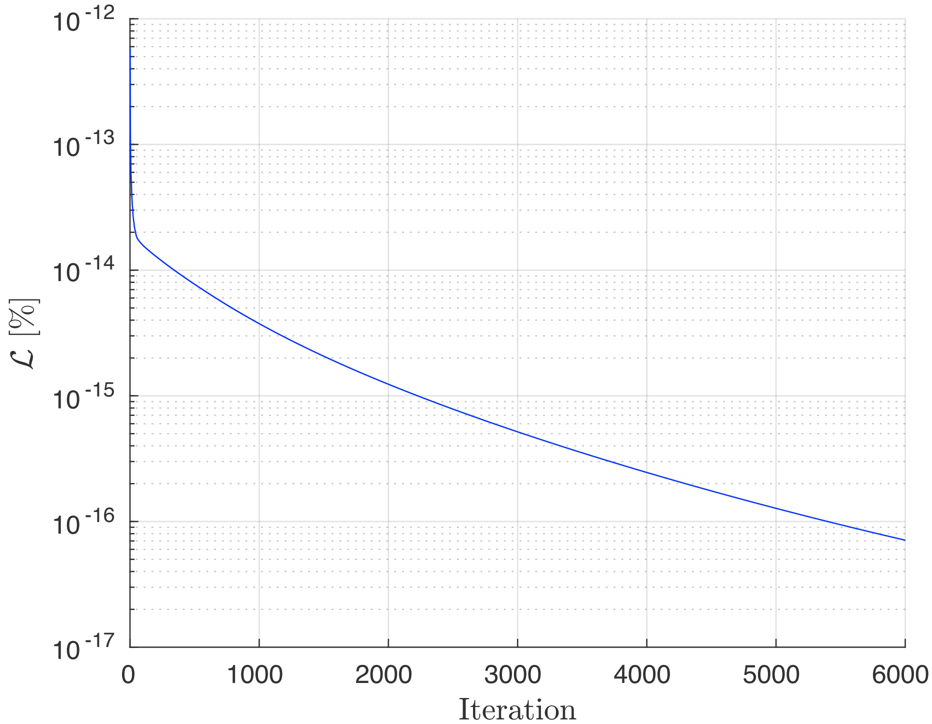


Fig. 20. Example 6 - Objective function, $\mathcal{L}$, versus iterations.

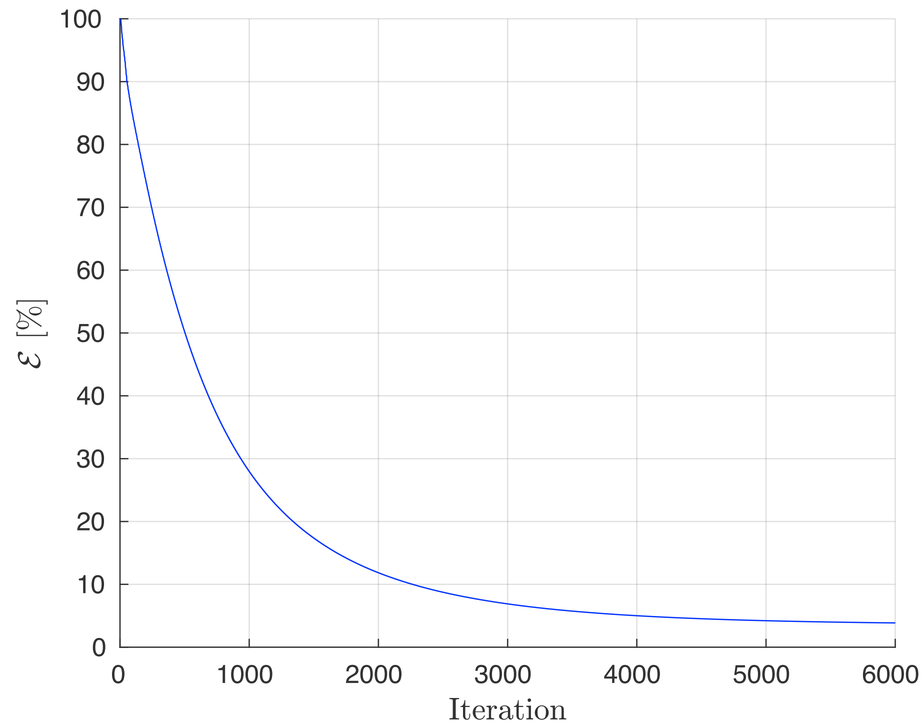
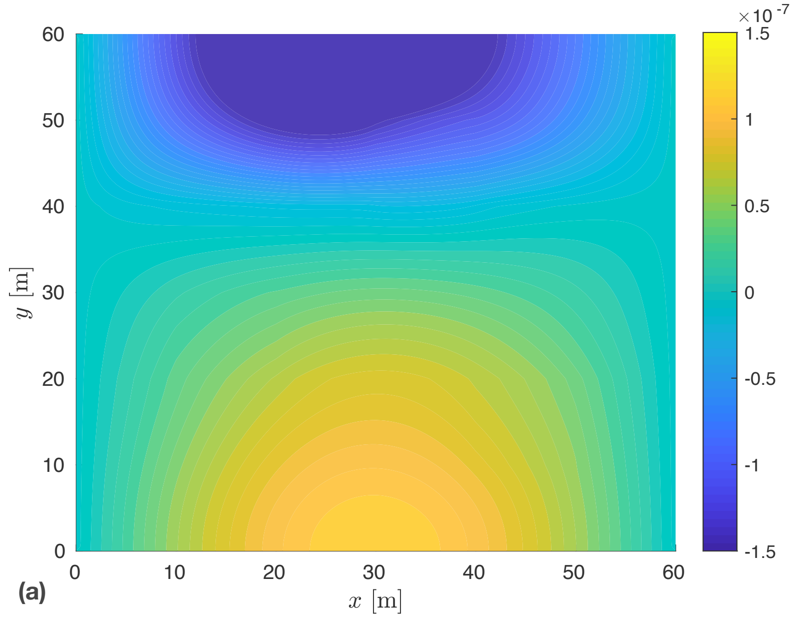


Fig. 21. Example 6 - Error, $\mathcal{E}$, versus iterations.



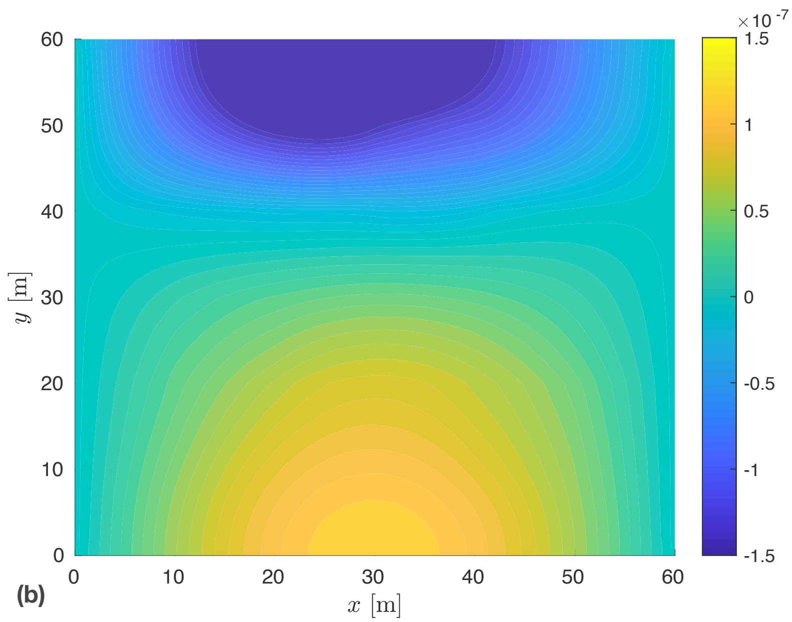


Fig. 22. Wave responses, $u(x, y, t)$, at 6.0 seconds due to (a) target load and (b) reconstructed load in Case 18.

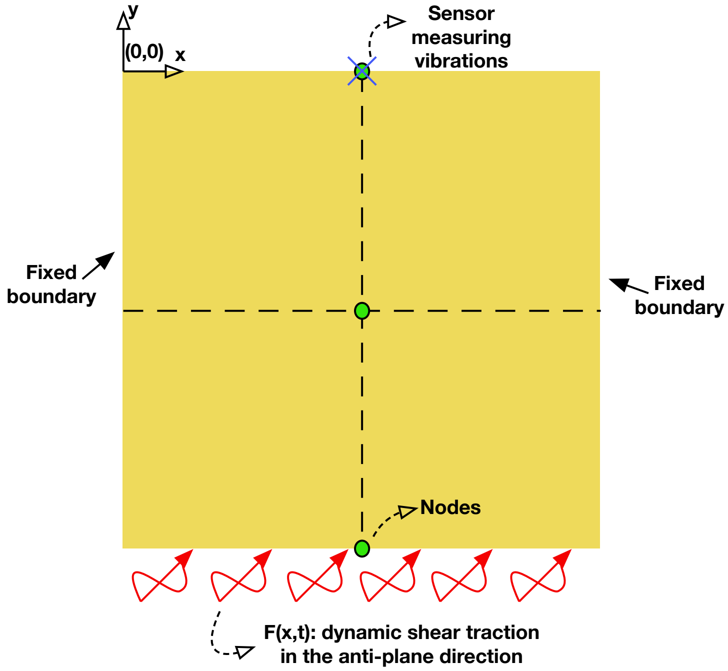


Fig. 23. Problem setting of Appendix III