On the full-waveform inversion of Lamé parameters in semi-infinite solids in plane strain

Sezgin Kucukcoban[a], Heedong Goh[b], Loukas F. Kallivokas[b,c,*]

[a] *Stress Engineering Services, Inc., 13610 Westland East Blvd., Houston, TX, USA*
[b] *Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Austin, TX, USA*
[c] *The Institute for Computational Engineering & Sciences, The University of Texas at Austin, Austin, TX, USA*

A R T I C L E I N F O

A B S T R A C T

We are concerned with imaging the spatially varying Lamé parameters of semi-infinite, arbitrarily heterogeneous solids, when probed by elastic waves in the time domain. We use a full-waveform inversion approach to tackle the inverse medium problem, and seek the Lamé distributions that minimize the misfit between measured and computed responses, subject to the governing PDEs. As is commonly the case, the resulting inverted profiles of the second Lamé parameter (μ) are of better quality than those of the first (λ). To improve the resolution of both Lamé parameters, we discuss the use of three robustifying schemes, namely, source-frequency continuation, regularization factor continuation, and a search direction-biasing scheme. We demonstrate with numerical experiments the effect the schemes have on the inversion process and conclude with an application of the robustified full-waveform method to a challenging adaptation of the Marmousi2 model.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

The identification of parameters describing the behavior of physical or engineered systems based on their response to stimuli is of fundamental interest in various fields of science and engineering. In the context of solids, many of the early attempts focused on the vibrational response (natural frequencies, modes, strains, etc) to guide condition assessment and infer system parameters (e.g. Gladwell, 1984; Juang and Pappa, 1985; Salawu, 1997). More broadly, parameter identification is an inverse medium problem, where the spatially distributed medium properties are the unknown parameters to be identified using a process that accounts for the sources/input/stimuli and the response/output of the probed medium, while adhering to a mathematical model for describing its physical behavior.

To date, the most intensely pursued inverse medium problem pertains to the imaging of the subsurface, owing to the economic significance of the driving application (oil exploration). Early inversion attempts approached the subsurface as a single-parameter medium to simplify or constrain the inversion process. A reasonable next step was to model the subsurface as an elastic medium,

thus giving rise to an inverse medium problem where two spatially distributed parameters are sought. But, even under fully elastic conditions, joint inversion for the elastic properties of the subsurface, i.e., of an arbitrarily heterogeneous semi-infinite medium, remains, by and large, an open problem. While most developments focus on the geophysical probing application of deep earth formations, fundamentally the problem is that of a semi-infinite elastic solid probed with elastic waves: the applications range from damage detection in composite structures at various scales, to site characterization of the near-surface deposits, to skin cancer detection and elastography.

Many a technique have been developed in the pursuit of elastic solid subsurface imaging: migration velocity analysis approach (MVA) (Clément et al., 1995; Chavent and Jacewitz, 1995; Plessix et al., 1999) is a good example, based on a decomposition of the sought properties into slow-varying (background) and fast-varying components (reflectivity), where, however, each component is determined separately. The background component necessitates a travel-time inversion, whereas the reflectivity component requires a prestack migration, thus increasing the complexity and overall computational cost of the approach. More recently, full-waveform inversion (FWI) schemes (Bunks et al., 1995; Pratt et al., 1996; Shin and Min, 2006; Brenders and Pratt, 2007) have been developed, relying typically on a least-squares (or other norm) data fitting process, where simultaneous inversion, instead of the staggered MVA approach, is attempted for the elastic parameters. Examples

---

* Corresponding author at: Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Austin, TX, USA.
E-mail addresses: sezgin.kucukcoban@stress.com (S. Kucukcoban), heedong.goh@utexas.edu (H. Goh), loukas@mail.utexas.edu (L.F. Kallivokas).

include Pratt (1999), Pratt and Shipp (1999), Operto et al. (2004), Gao et al. (2007), and Choi et al. (2005, 2008) who explored full-waveform inversion using either real or synthetic data sets to reconstruct complicated subsurface structures. Both single parameter FWI (Bunks et al., 1995; Shin and Min, 2006; Bamberger et al., 1979; Epanomeritakis et al., 2008; Kang and Kallivokas, 2011; Métivier et al., 2014) and dual-parameter FWI (Clément et al., 1995; Tarantola, 1984; Kallivokas et al., 2013; Fathi et al., 2015) have been developed and continue to be refined.

As in all inverse problems, solution multiplicity is a challenge, and it remains a difficulty in full-waveform approaches as well (Symes, 2008). Typically, to alleviate the multiplicity, regularization techniques, such as Tikhonov (Tikhonov, 1963), or Total Variation (Rudin et al., 1992) are used. However, despite the adoption of regularization techniques, robust solutions to the joint inversion remain challenging. Furthermore, for elastic solids, one has several choices for defining the spatially-distributed parameter pair to invert for: the two Lamé parameters, Young's modulus and Poisson's ratio, bulk and shear moduli, P- and S-wave velocities or slownesses (Brossier et al., 2009; Virieux and Operto, 2009), etc. Irrespective of the choice of the pair of target inversion parameters, no single pair has been shown to result in superior resolution over another, under joint inversion conditions (Epanomeritakis, 2004).

In this paper, we are concerned with robustifying the application of full-waveform inversion, and discuss three schemes whose combination improves the resolution of the reconstructed profiles of the spatial distributions of the Lamé parameters in a surface-probed semi-infinite solid. Specifically, we discuss a search-direction biasing scheme, a regularization factor continuation scheme, and a source-frequency continuation scheme, henceforth referred to as RS1, RS2, and RS3, respectively. RS2 was first reported in Na and Kallivokas (2008), and used in combination with RS3 in Kang and Kallivokas (2011) for a 2D scalar full-waveform inversion problem, i.e., in order to invert for a single spatially distributed inversion parameter. While the combined use of RS2 and RS3 was sufficient to reduce the solution multiplicity in the scalar case, as we reported in Kang and Kallivokas (2011), the mere extension of their use to the vector wave case fails to produce acceptable profiles: typically, in the elastic case, one of the two elastic parameters is recovered more accurately than the other, as will be shown. To alleviate the resolution mismatch between the two elastic parameters, we discuss a search-direction biasing scheme (RS1), first introduced in Kucukcoban (2010), which, as will be shown, in combination with RS2 and RS3 lends robustness to the inversion. The three schemes were also used to robustify the inversion in 3D full-waveform inversion applications using synthetics (Fathi et al., 2015; Fathi, 2015) or field data (Fathi et al., 2016). There are two key differences between the 2D case treated herein and the 3D cases reported earlier: a. The 3D case results in equations of motion implicating four temporal orders (from zero to third order), whereas the 2D case engages only three orders. This is due to the treatment of the unbounded domain via Perfectly-Matched-Layers (PMLs), which force a rise in the temporal order in 3D compared to 2D, thus requiring specialized time integrators, which are not necessary in 2D. The 2D case, as will be shown, results in a classic second-order semi-discrete system that benefits from standard integrators. b. The 3D inversion case was treated within a discretize-then-optimize framework, where the discrete problem is side-imposed to a misfit functional; by contrast, here we follow an optimize-then-discretize approach, which is simpler and allows greater flexibility in the choice of the discretization scheme.

This article is organized as follows: we first review the modeling of the time-domain response of a semi-infinite elastic solid using a hybrid, symmetric, variational formulation for a PML-truncated plane-strain elastic medium. Next, we discuss the mathematical framework of the full-waveform inversion using the three robustifying schemes for the simultaneous reconstruction of the spatially-distributed Lamé parameters (λ and μ). We demonstrate the performance of the robustifying schemes in detail with numerical experiments, and conclude with an application involving the challenging Marmousi2 benchmark problem (Versteeg, 1994).

2. Problem definition

We are concerned with the reconstruction of the material properties of an arbitrarily heterogeneous two-dimensional semi-infinite medium using surficial measurements of its response to known dynamic excitations applied on its surface, as schematically shown in Fig. 1(a) and (b). To account for the unboundedness of the physical domain in a computational model, the domain is truncated through the introduction of PMLs, i.e., buffers surrounding the truncated domain within which outgoing waves are forced to rapidly decay. Then, the goal is to recover the spatial distribution of the Lamé parameters $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ in the PML-truncated domain (Fig. 1(b)). Our approach is based on the PDE-constrained optimization apparatus, and thus involves casting the forward problem as a constraint (Kallivokas et al., 2013). We review first the forward model for completeness; details can be found in Kucukcoban and Kallivokas (2013).

3. Full-waveform inversion

3.1. The forward problem

An efficient approach for forward wave simulations in 2D PML-truncated elastic media has been introduced in Kucukcoban and Kallivokas (2013). By contrast to preceding developments, Kucukcoban and Kallivokas (2013) combines a displacement-only formulation for the interior domain Ω^{ID}, with a mixed-field formulation for the PML buffer zone Ω^{PML}, where both displacements and stresses are unknowns: this hybrid approach is computationally optimal (See Kucukcoban and Kallivokas (2011) for a comprehensive review of PMLs).

The wave motion in the PML-truncated domain is governed by the system of PDEs and associated conditions shown below. Find $\mathbf{u} \equiv \mathbf{u}(\mathbf{x}, t)$ in $\Omega = \Omega^{\mathrm{ID}} \cup \Omega^{\mathrm{PML}}$ and $\mathbf{S} \equiv \mathbf{S}(\mathbf{x}, t)$ in Ω^{PML} such that:

$$\mathbf{div}\left\{\mu\left[\nabla\mathbf{u} + (\nabla\mathbf{u})^T\right] + \lambda(\mathrm{div}\mathbf{u})\mathcal{I}\right\} + \mathbf{f} = \rho\ddot{\mathbf{u}} \quad \text{in } \Omega^{\mathrm{ID}} \times \mathsf{J}, \quad (1a)$$

$$\mathbf{div}\left(\dot{\mathbf{S}}^T\tilde{\Lambda}_e + \mathbf{S}^T\tilde{\Lambda}_p\right) = \rho(a\ddot{\mathbf{u}} + b\dot{\mathbf{u}} + c\mathbf{u}) \quad \text{in } \Omega^{\mathrm{PML}} \times \mathsf{J}, \quad (1b)$$

$$\mathcal{D} : \left(a\ddot{\mathbf{S}} + b\dot{\mathbf{S}} + c\mathbf{S}\right) = \frac{1}{2}\left[(\nabla\dot{\mathbf{u}})\tilde{\Lambda}_e + \tilde{\Lambda}_e(\nabla\dot{\mathbf{u}})^T + (\nabla\mathbf{u})\tilde{\Lambda}_p \right. \\ \left. + \tilde{\Lambda}_p(\nabla\mathbf{u})^T\right] \quad \text{in } \Omega^{\mathrm{PML}} \times \mathsf{J}, \quad (1c)$$

subject to silent initial conditions, and the following boundary and interface conditions:

$$\left\{\mu\left[\nabla\mathbf{u} + (\nabla\mathbf{u})^T\right] + \lambda(\mathrm{div}\mathbf{u})\mathcal{I}\right\}\mathbf{n} = \mathbf{g}_n \quad \text{on } \Gamma_N^{\mathrm{ID}} \times \mathsf{J}, \quad (2a)$$

$$(\dot{\mathbf{S}}^T\tilde{\Lambda}_e + \mathbf{S}^T\tilde{\Lambda}_p)\mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_N^{\mathrm{PML}} \times \mathsf{J}, \quad (2b)$$

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma_D^{\mathrm{PML}} \times \mathsf{J}, \quad (2c)$$

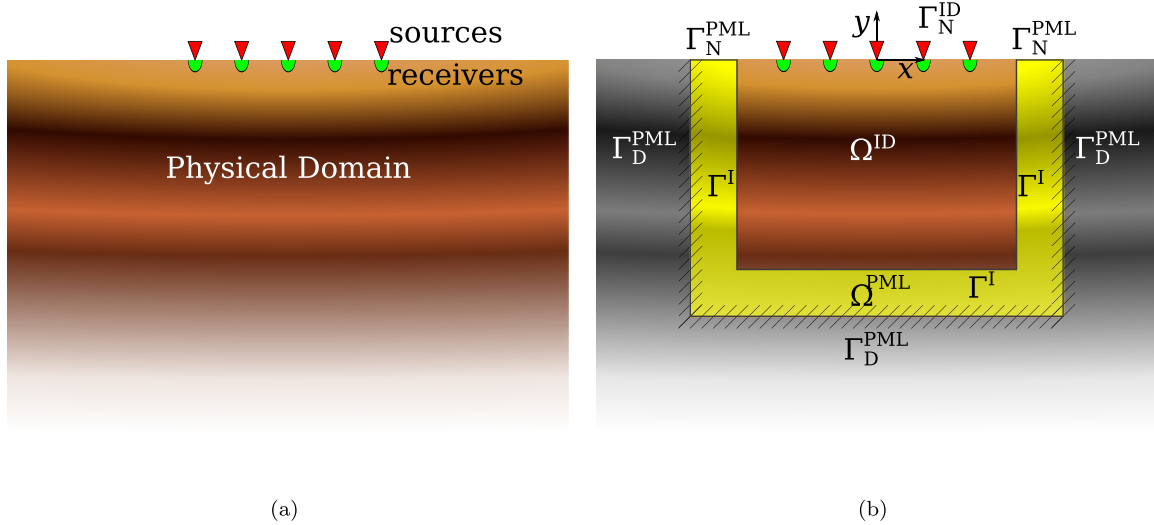$$\mathbf{u}^{\mathrm{ID}} = \mathbf{u}^{\mathrm{PML}} \quad \text{on } \Gamma^{\mathrm{I}} \times \mathsf{J}, \quad (2d)$$

Fig. 1. (a) A heterogeneous semi-infinite domain probed by surface sources; (b) configuration of a PML-truncated computational domain.

$$\left\{\mu\left[\nabla\mathbf{u}+(\nabla\mathbf{u})^T\right]+\lambda(\text{div}\mathbf{u})\mathcal{I}\right\}\mathbf{n}= -\left(\dot{\mathbf{S}}^T\tilde{\Lambda}_e+\mathbf{S}^T\tilde{\Lambda}_p\right)\mathbf{n}\quad\text{on }\Gamma^{\text{I}}\times\mathsf{J},\tag{2e}$$

where $\mathbf{u}$ is the displacement vector, $\dot{\mathbf{S}}$ is the stress tensor, $\mathcal{I}$ denotes the identity tensor, $\mathbf{f}$ is the load vector, and ρ is the mass density. $\mathcal{D}$ is the fourth order compliance tensor, a colon (:) denotes tensor inner product, and a dot ($\dot{\ }$) denotes differentiation with respect to time of the subtended function. $\tilde{\Lambda}_e$ and $\tilde{\Lambda}_p$ are the stretch tensors associated with evanescent and propagating waves, respectively. a, b, and c are products of certain elements of the stretch tensors (Kucukcoban and Kallivokas, 2013; 2011). Moreover, Γ^{I} is the interface boundary between the interior and PML domains. Ω is bounded by $\Gamma_D^{\text{PML}}\cup\Gamma_N^{\text{PML}}\cup\Gamma_N^{\text{ID}}$. Moreover, $\mathbf{g}_n$ denotes prescribed surface tractions on Γ_N^{ID}, and $\mathsf{J}=(0,T]$ denotes the time interval of interest. The hybrid approach couples two initially-uncoupled sets of governing equations via the continuity of displacements and tractions at the interface Γ^{I} between Ω^{ID} and Ω^{PML}.

Upon discretization, we obtain the following *semi-discrete form* using standard finite element approximants:

$$\mathbf{M}\ddot{\mathbf{d}}+\mathbf{C}\dot{\mathbf{d}}+\mathbf{K}\mathbf{d}=\mathbf{F},\tag{3}$$

where $\mathbf{M}$, $\mathbf{C}$, $\mathbf{K}$ are symmetric system matrices, $\mathbf{d}$ is the vector of nodal unknowns (comprising displacements in $\Omega^{\text{ID}}\cup\Omega^{\text{PML}}$, and stresses only in Ω^{PML}), and $\mathbf{F}$ is the vector of applied forces. To resolve the second-order time integration in (3), we employ the classical Newmark-β approach (Hughes, 1987), using the average acceleration scheme.

3.2. The inverse problem

We consider the sources and the response collected at receivers on the solid's top surface as known. We formulate the inverse problem initially as a misfit minimization problem, where the misfit is defined as the difference between the measured response at the receivers and a computed response, where the latter is obtained using trial distributions of the material parameters $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$.

Referring to Fig. 1(b), let Γ_m denote the part of the surface Γ_N^{ID} occupied by measuring stations (receivers), and let $\mathbf{u}_m(\mathbf{x},t)$ denote the measured displacement response to a known excitation. Let N_r denote the total number of receivers on the surface, and let $\mathbf{u}(\mathbf{x},t)$

be the computed response corresponding to a trial material profile. Then, the misfit least-squares minimization problem can be cast as:

$$\min_{\lambda,\mu}\mathcal{F}:=\frac{1}{2}\sum_{j=1}^{N_r}\int_0^T\int_{\Gamma_m}(\mathbf{u}-\mathbf{u}_m)^2\,\delta(\mathbf{x}-\mathbf{x}_j)\,\mathrm{d}\Gamma_m dt+\mathcal{R}(\lambda,\mu),\tag{4}$$

subject to (1a–1c) and (2a–2e). $\mathcal{F}$ denotes the objective functional, comprising the misfit functional augmented by the regularization functional. In (4), $\mathcal{R}$ could either adhere to a Tikhonov (TN) scheme (Tikhonov, 1963),

$$\mathcal{R}(\lambda,\mu)=\frac{R_\lambda}{2}\int_\Omega\nabla\lambda\cdot\nabla\lambda\ \mathrm{d}\Omega+\frac{R_\mu}{2}\int_\Omega\nabla\mu\cdot\nabla\mu\ \mathrm{d}\Omega,\tag{5}$$

or to a Total Variation (TV) (Rudin et al., 1992) scheme:

$$\mathcal{R}(\lambda,\mu)=R_\lambda\int_\Omega(\nabla\lambda\cdot\nabla\lambda+\epsilon)^{\frac{1}{2}}\,\mathrm{d}\Omega+R_\mu\int_\Omega(\nabla\mu\cdot\nabla\mu+\epsilon)^{\frac{1}{2}}\,\mathrm{d}\Omega.\tag{6}$$

R_λ and R_μ are user-defined regularization factors for λ and μ, respectively. We cast the constrained optimization problem (4) with the aid of a Lagrangian $\mathcal{L}$, whereby the misfit functional $\mathcal{F}$ is augmented with the side-imposition of the governing PDEs and boundary conditions via Lagrange multipliers, per:

$$\mathcal{L}\big(\mathbf{u},\mathbf{S},\boldsymbol{\theta_{u1}},\boldsymbol{\theta_{u2}},\boldsymbol{\theta_s},\boldsymbol{\theta_{b1}},\boldsymbol{\theta_{b2}},\lambda,\mu\big)$$
$$=\frac{1}{2}\sum_{j=1}^{N_r}\int_0^T\int_{\Gamma_m}(\mathbf{u}-\mathbf{u}_m)^2\,\mathrm{d}\Gamma_m dt\ +\ \mathcal{R}(\lambda,\mu)$$
$$+\int_{\Omega^{\text{ID}}}\int_0^T\boldsymbol{\theta_{u1}}\cdot\Big[\mathbf{div}\big\{\mu\big[\nabla\mathbf{u}+(\nabla\mathbf{u})^T\big]+\lambda(\text{div}\mathbf{u})\mathcal{I}\big\}$$
$$+\mathbf{f}-\rho\ddot{\mathbf{u}}\Big]\,dt\mathrm{d}\Omega$$
$$+\int_{\Omega^{\text{PML}}}\int_0^T\boldsymbol{\theta_{u2}}\cdot\Big[\mathbf{div}(\dot{\mathbf{S}}^T\tilde{\Lambda}_e+\mathbf{S}^T\tilde{\Lambda}_p)-\rho(a\ddot{\mathbf{u}}+b\dot{\mathbf{u}}+c\mathbf{u})\Big]\,dt\mathrm{d}\Omega$$
$$+\int_{\Omega^{\text{PML}}}\int_0^T\boldsymbol{\theta_s}:\Big\{\mathcal{D}:(a\ddot{\mathbf{S}}+b\dot{\mathbf{S}}+c\mathbf{S})-\frac{1}{2}\big[(\nabla\dot{\mathbf{u}})\tilde{\Lambda}_e+\tilde{\Lambda}_e(\nabla\dot{\mathbf{u}})^T$$
$$+(\nabla\mathbf{u})\tilde{\Lambda}_p+\tilde{\Lambda}_p(\nabla\mathbf{u})^T\big]\Big\}\,dt\mathrm{d}\Omega$$
$$+\int_{\Gamma_N^{\text{ID}}}\int_0^T\boldsymbol{\theta_{b1}}\cdot\Big\{\big(\mu\big[\nabla\mathbf{u}+(\nabla\mathbf{u})^T\big]+\lambda(\text{div}\mathbf{u})\mathcal{I}\big)\mathbf{n}-\mathbf{g}_n\Big\}\,dt\mathrm{d}\Gamma$$
$$+\int_{\Gamma_N^{\text{PML}}}\int_0^T\boldsymbol{\theta_{b2}}\cdot\Big[\big(\dot{\mathbf{S}}^T\tilde{\Lambda}_e+\mathbf{S}^T\tilde{\Lambda}_p\big)\mathbf{n}\Big]\,dt\mathrm{d}\Gamma\tag{7}$$

with

$$\mathbf{u} = \mathbf{0} \qquad \text{on } \Gamma_D^{\text{PML}} \times \mathsf{J}, \qquad (8a)$$

$$\mathbf{u}^+ = \mathbf{u}^- \qquad \text{on } \Gamma^{\text{I}} \times \mathsf{J}, \qquad (8b)$$

$$\left\{ \mu \left[\nabla\mathbf{u} + (\nabla\mathbf{u})^T \right] + \lambda (\text{div}\mathbf{u})\mathcal{I} \right\}\mathbf{n} = -(\dot{\mathbf{S}}^T \tilde{\Lambda}_e + \mathbf{S}^T \tilde{\Lambda}_p)\mathbf{n} \qquad \text{on } \Gamma^{\text{I}} \times \mathsf{J}, \qquad (8c)$$

$$\mathbf{u}(\mathbf{x}, 0) = \mathbf{0}, \qquad \dot{\mathbf{u}}(\mathbf{x}, 0) = \mathbf{0} \qquad \text{in } \Omega, \qquad (8d)$$

$$\mathbf{S}(\mathbf{x}, 0) = \mathbf{0}, \qquad \dot{\mathbf{S}}(\mathbf{x}, 0) = \mathbf{0} \qquad \text{in } \Omega. \qquad (8e)$$

In the above, $\{\mathbf{u}, \mathbf{S}\}$ are the *state variables* (s), $\{\boldsymbol{\theta_{u1}}, \boldsymbol{\theta_{u2}}, \boldsymbol{\theta_s}, \boldsymbol{\theta_{b1}}, \boldsymbol{\theta_{b2}}\}$ are the *adjoint variables* (m), and $\{\lambda, \mu\}$ are the *control variables* (c). Thus, the inverse medium problem is cast as a PDE-constrained least-squares misfit optimization problem, similar to Epanomeritakis et al. (2008), Kang (2010), Kucukcoban (2010), Kang and Kallivokas (2011), Kallivokas et al. (2013), and Fathi et al. (2015). Next, we seek to satisfy the stationarity of $\mathcal{L}$ by requiring that the first variations of $\mathcal{L}$ vanish, i.e.,

$$\nabla \mathcal{L} = \begin{bmatrix} \delta_{\text{m}}\mathcal{L} \\ \delta_{\text{s}}\mathcal{L} \\ \delta_{\text{c}}\mathcal{L} \end{bmatrix} = \mathbf{0}. \qquad (9)$$

The 1st optimality condition ($\delta_{\text{m}}\mathcal{L}$), results in the *state (or forward) problem*, identical to the one given by (1a–2e). The state problem is solved using the hybrid approach described in Section 3.1. The 2nd optimality condition ($\delta_{\text{s}}\mathcal{L}$) results in the following *adjoint problem*:

$$\mathbf{div}\left[\mu(\nabla\boldsymbol{\theta_{u1}} + \nabla\boldsymbol{\theta_{u1}}^T) + \lambda \, \text{div}\boldsymbol{\theta_{u1}} \, \mathcal{I} \right] = \rho\ddot{\boldsymbol{\theta_{u1}}} \quad \text{in } \Omega^{\text{ID}} \times \mathsf{J}, \quad (10a)$$

$$\mathbf{div}\left(-\dot{\boldsymbol{\theta_s}}\tilde{\Lambda}_e + \boldsymbol{\theta_s}\tilde{\Lambda}_p \right) = \rho\left(a\ddot{\boldsymbol{\theta_{u2}}} - b\dot{\boldsymbol{\theta_{u2}}} + c\boldsymbol{\theta_{u2}} \right) \qquad \text{in } \Omega^{\text{PML}} \times \mathsf{J}, \qquad (10b)$$

$$\mathcal{D} : \left(a\ddot{\boldsymbol{\theta_s}} - b\dot{\boldsymbol{\theta_s}} + c\boldsymbol{\theta_s} \right) = -\tilde{\Lambda}_e\left(\nabla\dot{\boldsymbol{\theta_{u2}}} \right)^T + \tilde{\Lambda}_p\left(\nabla\boldsymbol{\theta_{u2}} \right)^T \quad \text{in } \Omega^{\text{PML}} \times \mathsf{J}, \qquad (10c)$$

subject to

$$\left\{ \mu\left(\nabla\boldsymbol{\theta_{u1}} + \nabla\boldsymbol{\theta_{u1}}^T \right) + \lambda \, \text{div}\boldsymbol{\theta_{u1}} \, \mathcal{I} \right\}\mathbf{n} = \sum_{j=1}^{N_r} (\mathbf{u} - \mathbf{u}_m)\delta(\mathbf{x} - \mathbf{x}_j)$$
$$\text{on } \Gamma_m \times \mathsf{J}, \qquad (11a)$$

$$\left(-\dot{\boldsymbol{\theta_s}}\tilde{\Lambda}_e + \boldsymbol{\theta_s}\tilde{\Lambda}_p \right)\mathbf{n} = \mathbf{0} \qquad \text{on } \Gamma_N^{\text{PML}} \times \mathsf{J}, \qquad (11b)$$

$$\boldsymbol{\theta_{u2}} = \mathbf{0} \qquad \text{on } \Gamma_D^{\text{PML}} \times \mathsf{J}, \qquad (11c)$$

$$\boldsymbol{\theta_{u1}} = \boldsymbol{\theta_{u2}} \qquad \text{on } \Gamma^{\text{I}} \times \mathsf{J}, \qquad (11d)$$

$$\left\{ \mu\left(\nabla\boldsymbol{\theta_{u1}} + \nabla\boldsymbol{\theta_{u1}}^T \right) + \lambda \, \text{div}\boldsymbol{\theta_{u1}} \, \mathcal{I} \right\}\mathbf{n}$$
$$= -\left(-\dot{\boldsymbol{\theta_s}}\tilde{\Lambda}_e + \boldsymbol{\theta_s}\tilde{\Lambda}_p \right)\mathbf{n} \qquad \text{on } \Gamma^{\text{I}} \times \mathsf{J}, \qquad (11e)$$

$$\boldsymbol{\theta_{u1}}(\mathbf{x}, T) = \mathbf{0}, \qquad \dot{\boldsymbol{\theta_{u1}}}(\mathbf{x}, T) = \mathbf{0} \qquad \text{in } \Omega^{\text{ID}}, \qquad (11f)$$

$$\boldsymbol{\theta_{u2}}(\mathbf{x}, T) = \mathbf{0}, \qquad \dot{\boldsymbol{\theta_{u2}}}(\mathbf{x}, T) = \mathbf{0} \qquad \text{in } \Omega^{\text{PML}}, \qquad (11g)$$

$$\boldsymbol{\theta_s}(\mathbf{x}, T) = \mathbf{0}, \qquad \dot{\boldsymbol{\theta_s}}(\mathbf{x}, T) = \mathbf{0} \qquad \text{in } \Omega^{\text{PML}}. \qquad (11h)$$

We note that, as usual, the adjoint problem is a *final-value* problem: it is driven by the misfit. We also note that the operators implicated in the adjoint PDEs are identical to the state operators, modulo the sign reversal for the terms implicating first-order time derivatives. By construction, the adjoint equations are also hybrid and PML-endowed, with $(\boldsymbol{\theta_{u1}}, \boldsymbol{\theta_{u2}})$ and $\boldsymbol{\theta_s}$ playing a role analogous to $\mathbf{u}$ and $\mathbf{S}$ of the state problem, respectively.

Owing to the similarity of the operators implicated in the state and adjoint problems, we obtain the following *semi-discrete form*, where the superscript 'ad' stands for the *adjoint* problem.

$$\mathbf{M}^{\text{ad}}\ddot{\mathbf{d}}^{\text{ad}} + \mathbf{C}^{\text{ad}}\dot{\mathbf{d}}^{\text{ad}} + \mathbf{K}^{\text{ad}}\mathbf{d}^{\text{ad}} = \mathbf{F}^{\text{ad}}, \qquad (12)$$

where $\mathbf{M}^{\text{ad}}$, $\mathbf{C}^{\text{ad}}$, $\mathbf{K}^{\text{ad}}$ are the mass-like, damping-like, stiffness-like system matrices, $\mathbf{d}^{\text{ad}}$ is the vector of nodal unknowns (comprising $\boldsymbol{\theta_u}$ and $\boldsymbol{\theta_s}$), and $\mathbf{F}^{\text{ad}}$ is the load vector. The adjoint problem matrices are the *same* as those resulting from the state problem, modulo a sign reversal for the damping matrix; that is

$$\mathbf{M}^{\text{ad}} = \mathbf{M}, \quad \mathbf{C}^{\text{ad}} = -\mathbf{C}, \quad \mathbf{K}^{\text{ad}} = \mathbf{K}. \qquad (13)$$

The above equalities lead to cost-effective implementations since the assembly process for the adjoint problem is avoided. To resolve the time integration, we employ a Newmark-β-like scheme in which time marching is reversed (Kucukcoban, 2010). The 3rd optimality condition ($\delta_{\text{c}}\mathcal{L}$) results in the following two boundary-value control problems:

λ-*control problem*:

$$-R_\lambda \Delta\lambda - \int_0^T (\text{div}\boldsymbol{\theta_{u1}})(\text{div}\mathbf{u}) \, dt = 0 \qquad \text{in } \Omega^{\text{ID}}, \qquad (14a)$$

subject to

$$\int_{\Gamma_N^{\text{ID}}} \nabla\lambda \cdot \mathbf{n} \, d\Gamma = 0 \qquad \text{on } \Gamma_N^{\text{ID}}, \qquad (14b)$$

$$R_\lambda \nabla\lambda \cdot \mathbf{n} = -\int_0^T \boldsymbol{\theta_{u1}} \cdot (\text{div}\mathbf{u})\mathbf{n} \, dt \qquad \text{on } \Gamma^{\text{I}}. \qquad (14c)$$

μ-*control problem*:

$$-R_\mu \Delta\mu - \int_0^T \nabla\boldsymbol{\theta_{u1}} : \left(\nabla\mathbf{u} + \nabla\mathbf{u}^T \right) dt = 0 \qquad \text{in } \Omega^{\text{ID}}, \qquad (15a)$$

subject to

$$\int_{\Gamma_N^{\text{ID}}} \nabla\mu \cdot \mathbf{n} \, d\Gamma = 0 \qquad \text{on } \Gamma_N^{\text{ID}}, \qquad (15b)$$

$$R_\mu \nabla\mu \cdot \mathbf{n} = -\int_0^T \boldsymbol{\theta_{u1}} \cdot \left(\nabla\mathbf{u} + \nabla\mathbf{u}^T \right)\mathbf{n} \, dt \qquad \text{on } \Gamma^{\text{I}}. \qquad (15c)$$

In writing (14a–14c) and (15a–15c), we adopted the TN scheme for regularizing the solutions. If the TV regularization were to be used instead, the first terms in (14a) and (15a) should be replaced by:

$$-R_\lambda \, \nabla \cdot \left[(\nabla\lambda \cdot \nabla\lambda + \epsilon)^{-\frac{1}{2}} \nabla\lambda \right]$$
$$-\int_0^T (\text{div}\boldsymbol{\theta_{u1}})(\text{div}\mathbf{u}) \, dt = 0 \qquad \text{in } \Omega^{\text{ID}}, \qquad (16)$$

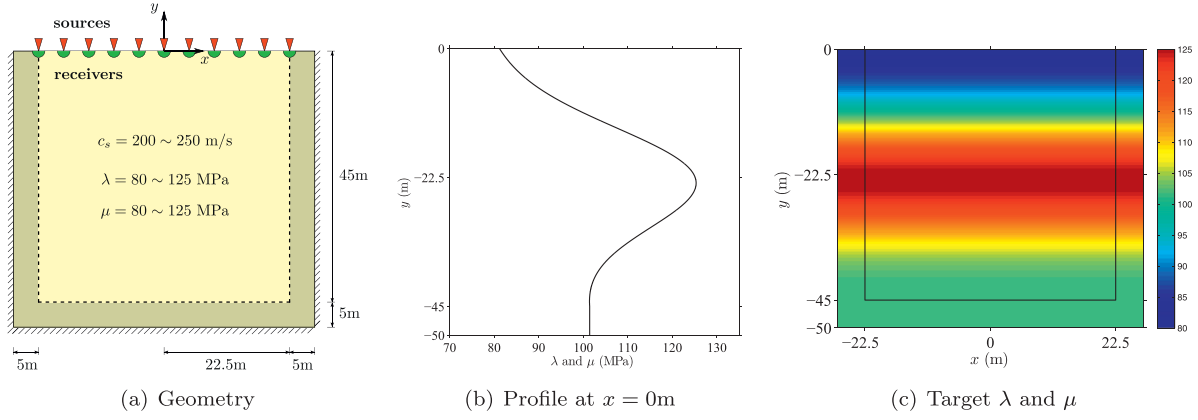(a) Geometry   (b) Profile at $x = 0$m   (c) Target λ and μ

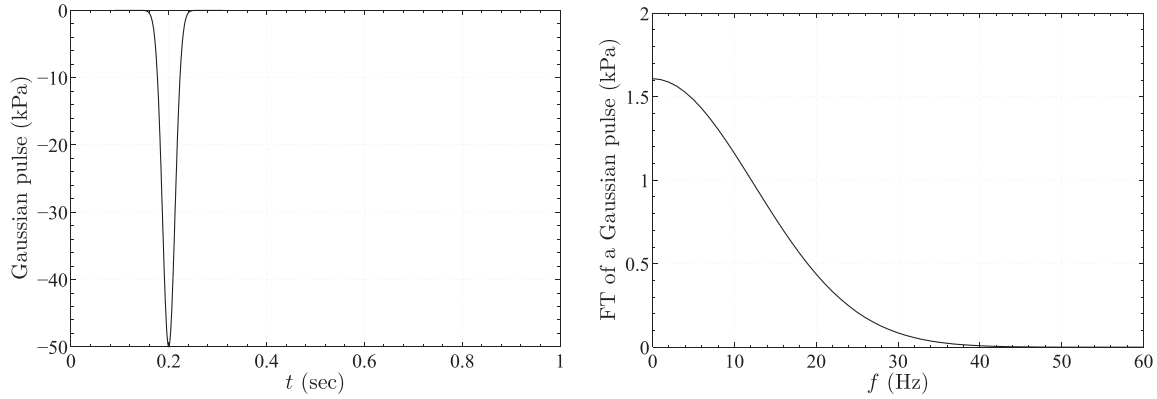**Fig. 2.** A PML-truncated semi-infinite domain in two dimensions.

Fig. 3. Excitation time signal and its Fourier spectrum.

$$- R_\mu \, \nabla \cdot \left[\left(\nabla\mu \cdot \nabla\mu + \epsilon \right)^{-\frac{1}{2}} \nabla\mu \right]$$

$$- \int_0^T \nabla\boldsymbol{\theta_{u1}} : \left(\nabla\mathbf{u} + \nabla\mathbf{u}^T \right) dt = 0 \qquad \text{in } \Omega^{\text{ID}}. \qquad (17)$$

3.2.1. The inversion process

Similar to Epanomeritakis et al. (2008) and Kang and Kallivokas (2011), a *reduced-space method* is used to solve for the stationarity of $\mathcal{L}$, in which the coupled system of PDEs are solved in the reduced space of the control variables. The procedure is iterative: we start with an assumed initial spatial distribution of the control parameters (λ and μ) and solve the state problem (1a–2e) for the state variables $\mathbf{u}$ and $\mathbf{S}$. Then, we solve the adjoint problem (10a–11h) to obtain the adjoint variables $\boldsymbol{\theta_u}$ and $\boldsymbol{\theta_s}$ based on the state solutions. Then, we satisfy the first and second optimality conditions, that is, the gradient of the Lagrangian becomes

$$\nabla\mathcal{L} = \begin{bmatrix} 0 \\ 0 \\ \delta_c\mathcal{L} \end{bmatrix} \qquad \Rightarrow \qquad \nabla_c\mathcal{L} = \delta_c\mathcal{L}. \qquad (18)$$

The term ($\delta_c\mathcal{L}$) is the reduced gradient ($\nabla_c\mathcal{L}$), and can be incorporated into a gradient-based scheme to update the control/material parameters.

3.2.2. Material parameter updates

The first and second optimality conditions are automatically satisfied once the state and adjoint problems are solved. However, the third optimality condition is satisfied exactly only for the true/target material profile. Other profiles result in non-vanishing control equations: (14a) and (15a) for the TN case or (16) and (17) for the TV case. We iteratively update the control parameters

(λ and μ) so that the misfit between the measured and computed responses reduces to a preset tolerance, thereby allowing the third optimality condition to be satisfied. To this end, we use a conjugate gradient method with inexact line search and make use of the information stored in the reduced gradient ($\nabla_c\mathcal{L}$).

The continuous forms of the associated reduced gradients can be expressed, in the TN case, as

$$\nabla_\lambda\mathcal{L} = -R_\lambda\Delta\lambda - \int_0^T (\text{div}\boldsymbol{\theta_{u1}})(\text{div}\mathbf{u}) \, dt \qquad \text{in } \Omega^{\text{ID}}, \qquad (19)$$

$$\nabla_\mu\mathcal{L} = -R_\mu\Delta\mu - \int_0^T \nabla\boldsymbol{\theta_{u1}} : \left(\nabla\mathbf{u} + \nabla\mathbf{u}^T \right) dt \qquad \text{in } \Omega^{\text{ID}}, \qquad (20)$$

whereas, in the TV case, they become

$$\nabla_\lambda\mathcal{L} = -R_\lambda \, \nabla \cdot \left[\left(\nabla\lambda \cdot \nabla\lambda + \epsilon \right)^{-\frac{1}{2}} \nabla\lambda \right] - \int_0^T (\text{div}\boldsymbol{\theta_{u1}})(\text{div}\mathbf{u}) \, dt, \qquad (21)$$

$$\nabla_\mu\mathcal{L} = -R_\mu \, \nabla \cdot \left[\left(\nabla\mu \cdot \nabla\mu + \epsilon \right)^{-\frac{1}{2}} \nabla\mu \right]$$

$$- \int_0^T \nabla\boldsymbol{\theta_{u1}} : \left(\nabla\mathbf{u} + \nabla\mathbf{u}^T \right) dt. \qquad (22)$$

In the above, the reduced gradients are restricted to Ω^{ID}. In the implementation, $\nabla_\lambda\mathcal{L}$ and $\nabla_\mu\mathcal{L}$ are evaluated at each nodal point.

We cast the iterative evolution of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ in the context of the conjugate gradient method (Fletcher and Reeves, 1964) as

$$\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + \alpha_k^\lambda \mathbf{d}_k^\lambda, \;\; \boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k + \alpha_k^\mu \mathbf{d}_k^\mu, \qquad (23)$$

(a) Target λ

(b) Target μ

(c) λ (TN)

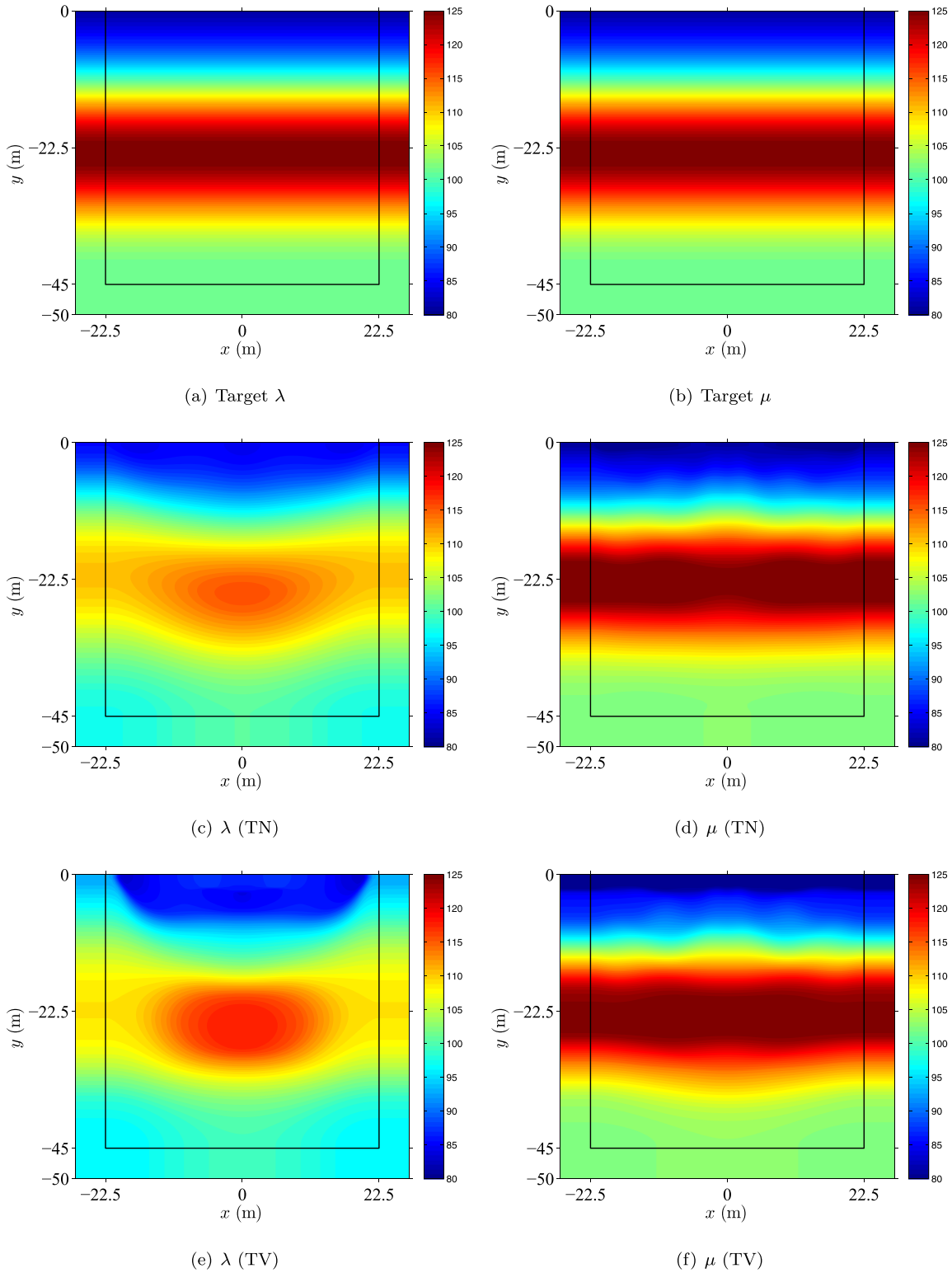(d) μ (TN)

(e) λ (TV)

(f) μ (TV)

Fig. 4. Simultaneous inversion for the Lamé parameters without robustifying schemes; convergence after 278 iterations (TN), and 280 iterations (TV).

where $k = 1, 2, \ldots, (\alpha_k^\lambda, \alpha_k^\mu)$ are the step lengths for λ and μ, and $(\mathbf{d}_k^\lambda, \mathbf{d}_k^\mu)$ are the associated search directions. An Armijo-like line search for both $\boldsymbol{\lambda}_k$ and $\boldsymbol{\mu}_k$ is used. However, instead of enforcing sufficient decrease on the Lagrangian $\mathcal{L}$, we favor the use of the objective functional $\mathcal{F}$ instead, simply because the side-imposed PDEs vanish in a weak sense and minimizing $\mathcal{L}$ becomes nearly equivalent to minimizing $\mathcal{F}$ (we have found this to be a more robust choice than enforcing the Armijo condition on the Lagrangian). Thus, we force sufficient decrease in the objective functional $\mathcal{F}$ by

$$\mathcal{F}\left(\boldsymbol{\lambda}_{k+1}, \boldsymbol{\mu}_{k+1}\right) - \mathcal{F}\left(\boldsymbol{\lambda}_k, \boldsymbol{\mu}_k\right) \leq \min\left\{\delta^\lambda \alpha_k^\lambda (\mathbf{g}_k^\lambda)^T \mathbf{d}_k^\lambda, \ \delta^\mu \alpha_k^\mu (\mathbf{g}_k^\mu)^T \mathbf{d}_k^\mu\right\},$$

(24)

where α_k is chosen to be the largest of $\{\alpha, \alpha\rho, \alpha\rho^2, \ldots\}$, where $\alpha > 0$ is an initial step length, $\rho \in (0, 1)$ is a contraction factor, and $\delta \in (0, 1)$ is a constant parameter (in practice, δ is chosen

(a) λ profile at $x = 0$m   (b) λ profile at $x = 7$m   (c) λ profile at $x = 14$m

(d) μ profile at $x = 0$m   (e) μ profile at $x = 7$m   (f) μ profile at $x = 14$m
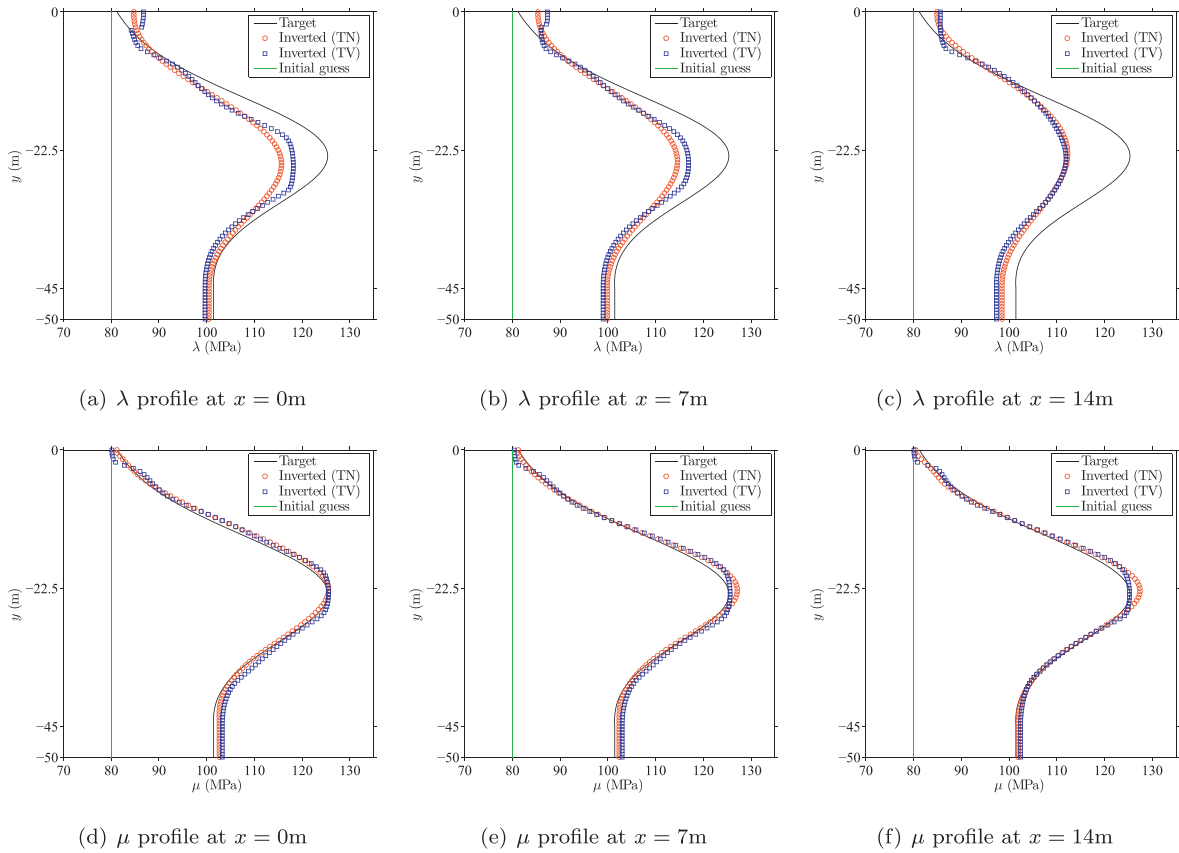
Fig. 5. Cross-sectional profiles for simultaneously inverted Lamé parameters without robustifying schemes; convergence after 278 iterations (TN), and 280 iterations (TV).
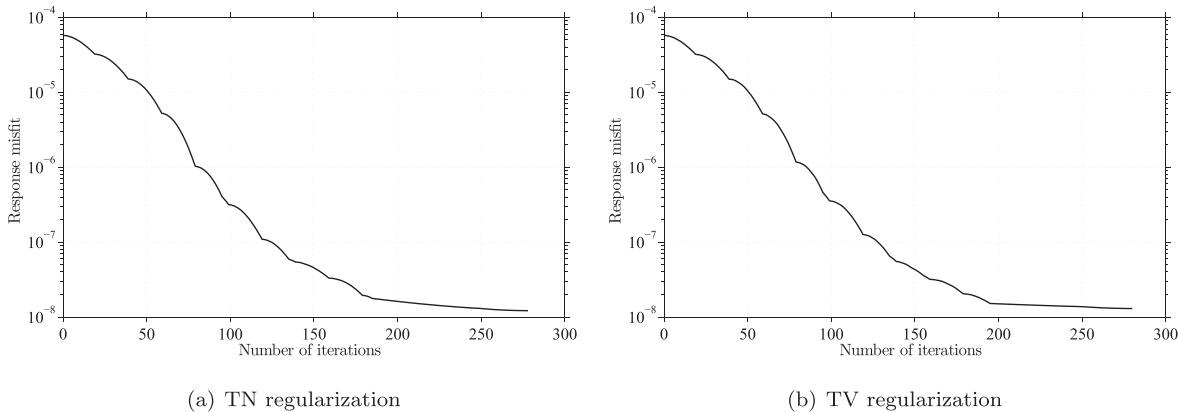

(a) TN regularization   (b) TV regularization

Fig. 6. Variation of response misfit during simultaneous inversion process without robustifying schemes.

to be quite small, e.g. $\delta^\lambda = \delta^\mu = 10^{-10}$). $\mathbf{g}_k^\lambda$ and $\mathbf{g}_k^\mu$ are the discrete reduced gradients at $\boldsymbol{\lambda}_k$ and $\boldsymbol{\mu}_k$, respectively. If the inequality (24) is not satisfied, a *backtracking* procedure is followed by contracting the associated step size (α_k^λ and/or α_k^μ) by ρ until the sufficient decrease condition is satisfied. Specifically, the step sizes for $\boldsymbol{\lambda}_k$ and $\boldsymbol{\mu}_k$ are determined independently since each has its own search direction. Once we achieve adequate reductions in the objective functional (4), we update the material property vectors $\boldsymbol{\lambda}_k$ and $\boldsymbol{\mu}_k$, and progress to the next iteration. We summarize the inversion process in Algorithm 1 in Appendix A.

4. Robustifying schemes

As noted, imposing regularization over the control parameters helps only partially to alleviate the ill-posedness and associated

solution multiplicity in joint inversion. To further assist the optimizer, we discuss next three robustifying schemes.

4.1. Biased search directions (RS1)

Most simultaneous inversion schemes for the elastic properties, irrespective of the choice of parameters, i.e., whether cast in terms of the pair of P- and S-wave velocities, the slownesses, or bulk and shear moduli, etc., result in reconstructions of dissimilar quality for one of the two parameters. This is indeed the case here as well with the pair of our choice – the Lamé parameters. A plausible explanation is this: the updates for μ are driven by gradient operators in the μ-control problem (see time integral in either (20) or (22)), whereas, by contrast, the updates on λ are based on divergence operators in the λ-control problem ((19) or (21)). The

(a) Target λ

(b) Target μ

(c) λ (TN)

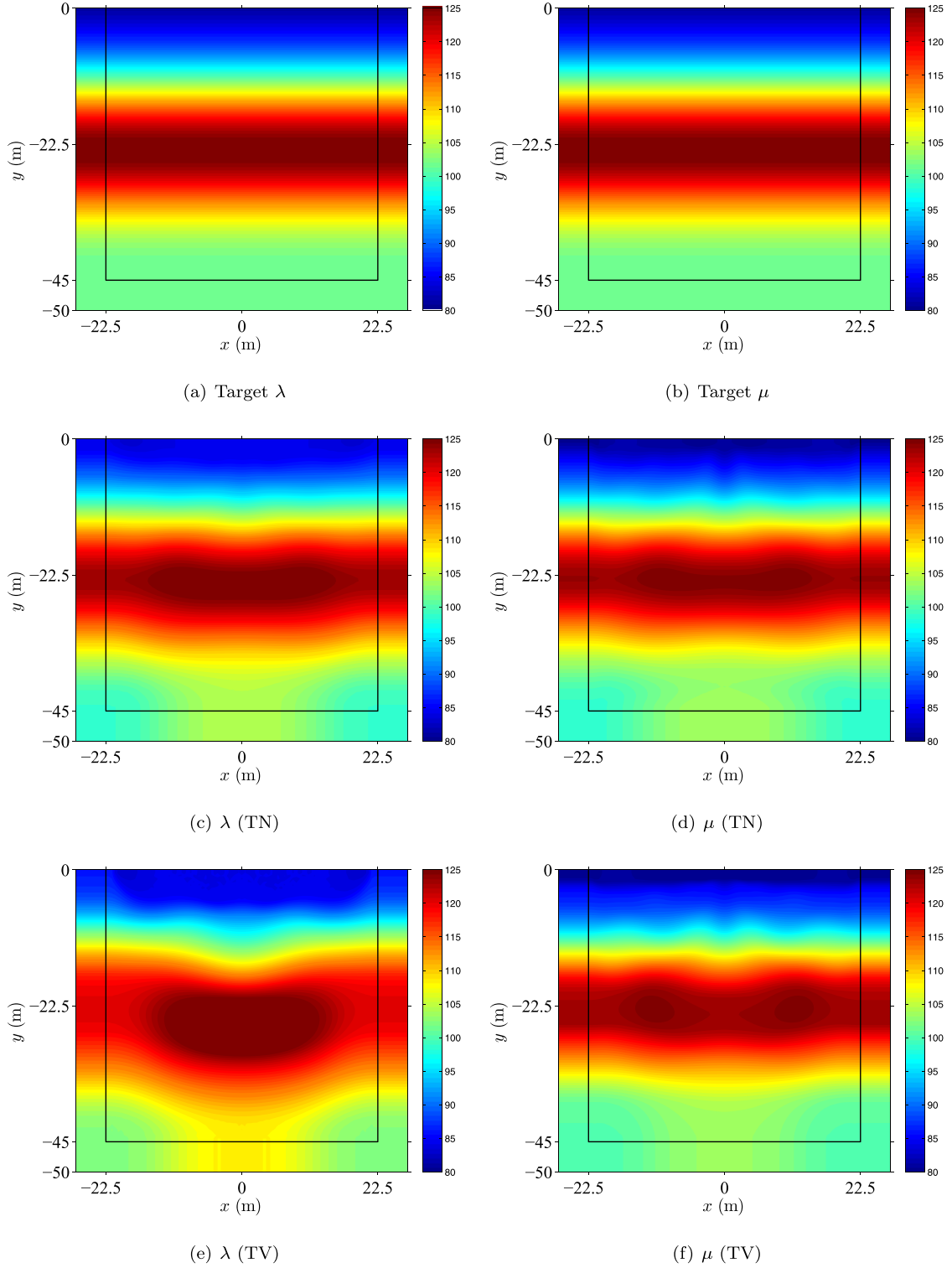(d) μ (TN)

(e) λ (TV)

(f) μ (TV)

Fig. 7. Simultaneous inversion for the Lamé parameters using the *biased*-directions RS1 scheme; convergence after 70 iterations (TN), and 180 iterations (TV).

gradient operator incorporates richer information on the spatial variability of $\mathbf{u}(\mathbf{x}, t)$ than the smoothing divergence operator. As a result, μ evolves towards the target profile faster than λ. We propose a scheme where the search directions of λ are biased by the search directions of μ during the early stages of the inversion process. The key idea behind this is to let λ know about the spatial structure of μ and update itself with a similar pattern; after all, the physics dictate that both λ and μ profiles should look

similar: since the spatial update pattern of μ is embedded in a n-dimensional unit search direction vector, we use a weighted average of unit λ- and μ-search directions for the evolution of λ. To implement, an extra step is inserted in Algorithm 1 after line 12 to define the new biased search direction:

$$\mathbf{d}_k^\lambda \leftarrow \left\| \mathbf{d}_k^\lambda \right\| \left[W \frac{\mathbf{d}_k^\mu}{\left\| \mathbf{d}_k^\mu \right\|} + (1 - W) \frac{\mathbf{d}_k^\lambda}{\left\| \mathbf{d}_k^\lambda \right\|} \right] \tag{25}$$

(a) λ profile at $x = 0$m　　　　(b) λ profile at $x = 7$m　　　　(c) λ profile at $x = 14$m

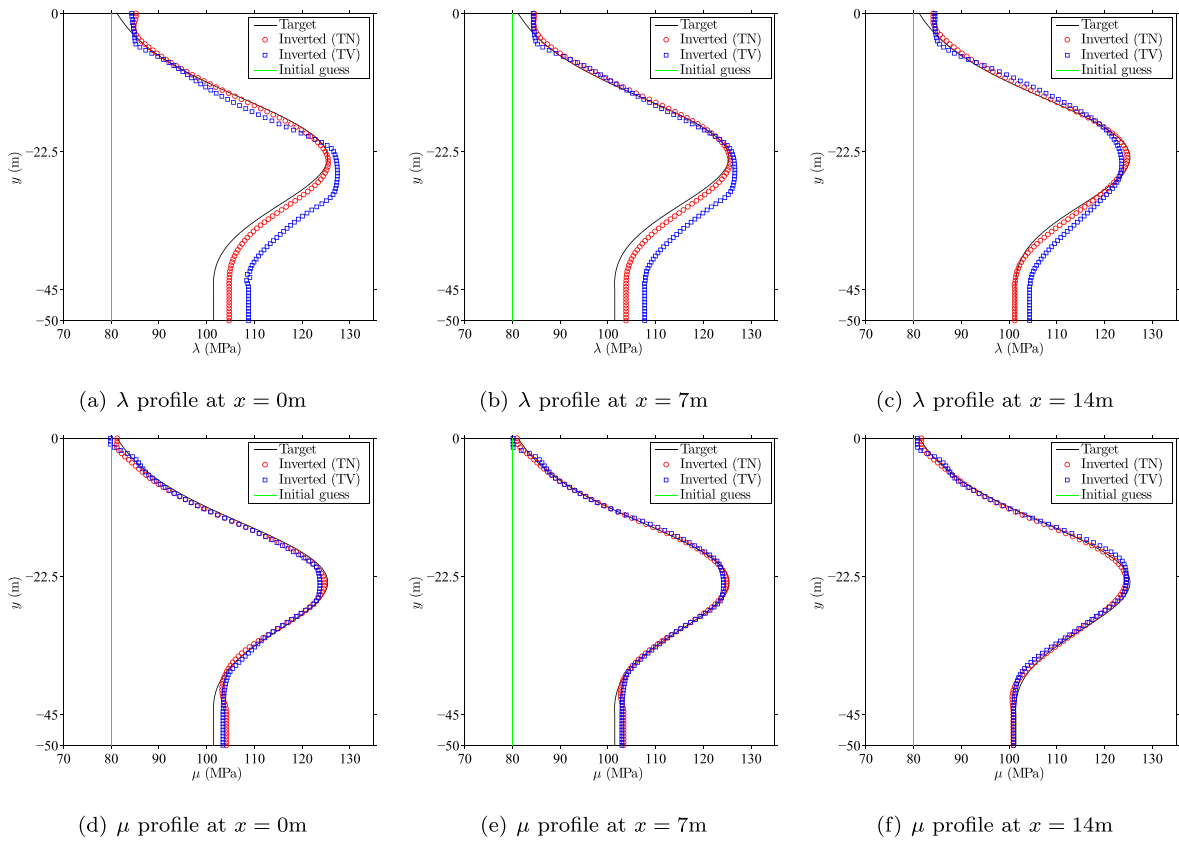(d) μ profile at $x = 0$m　　　　(e) μ profile at $x = 7$m　　　　(f) μ profile at $x = 14$m

Fig. 8. Cross-sectional profiles obtained using the *biased*-directions RS1 scheme.
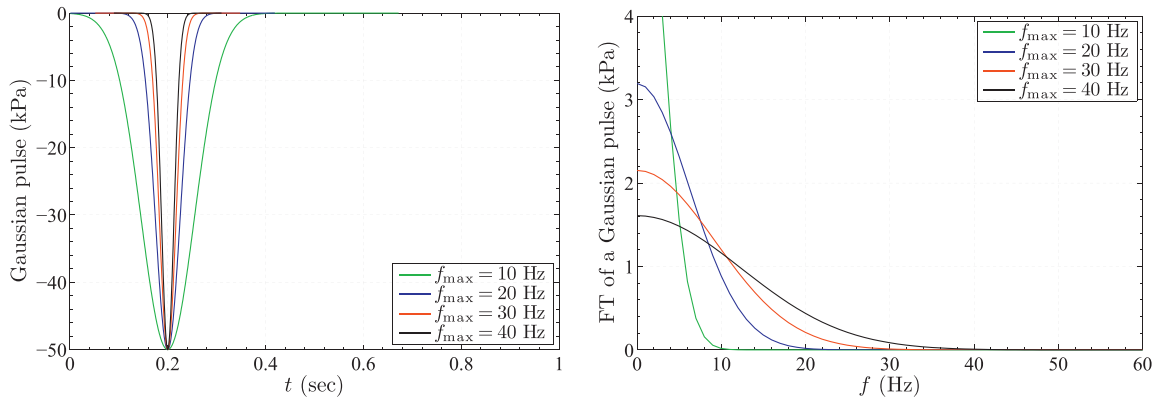
Fig. 9. Gaussian pulses and their Fourier spectra.
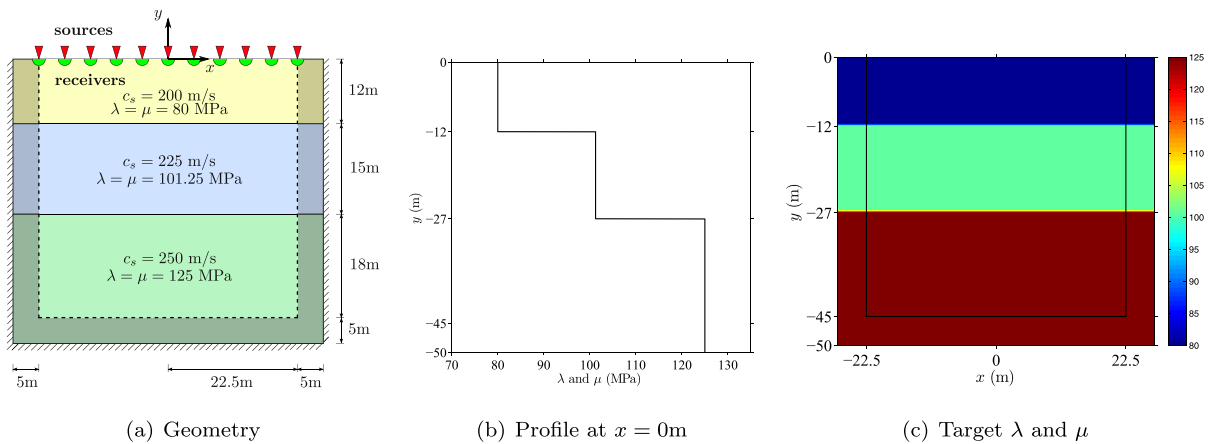

(a) Geometry　　　　(b) Profile at $x = 0$m　　　　(c) Target λ and μ

Fig. 10. A PML-truncated horizontally-layered semi-infinite domain in two dimensions.

(a) λ ($f_{\max} = 10$Hz)

(b) μ ($f_{\max} = 10$Hz)

(c) λ ($f_{\max} = 20$Hz)

(d) μ ($f_{\max} = 20$Hz)

(e) λ ($f_{\max} = 30$Hz)

(f) μ ($f_{\max} = 30$Hz)

(g) λ ($f_{\max} = 40$Hz)
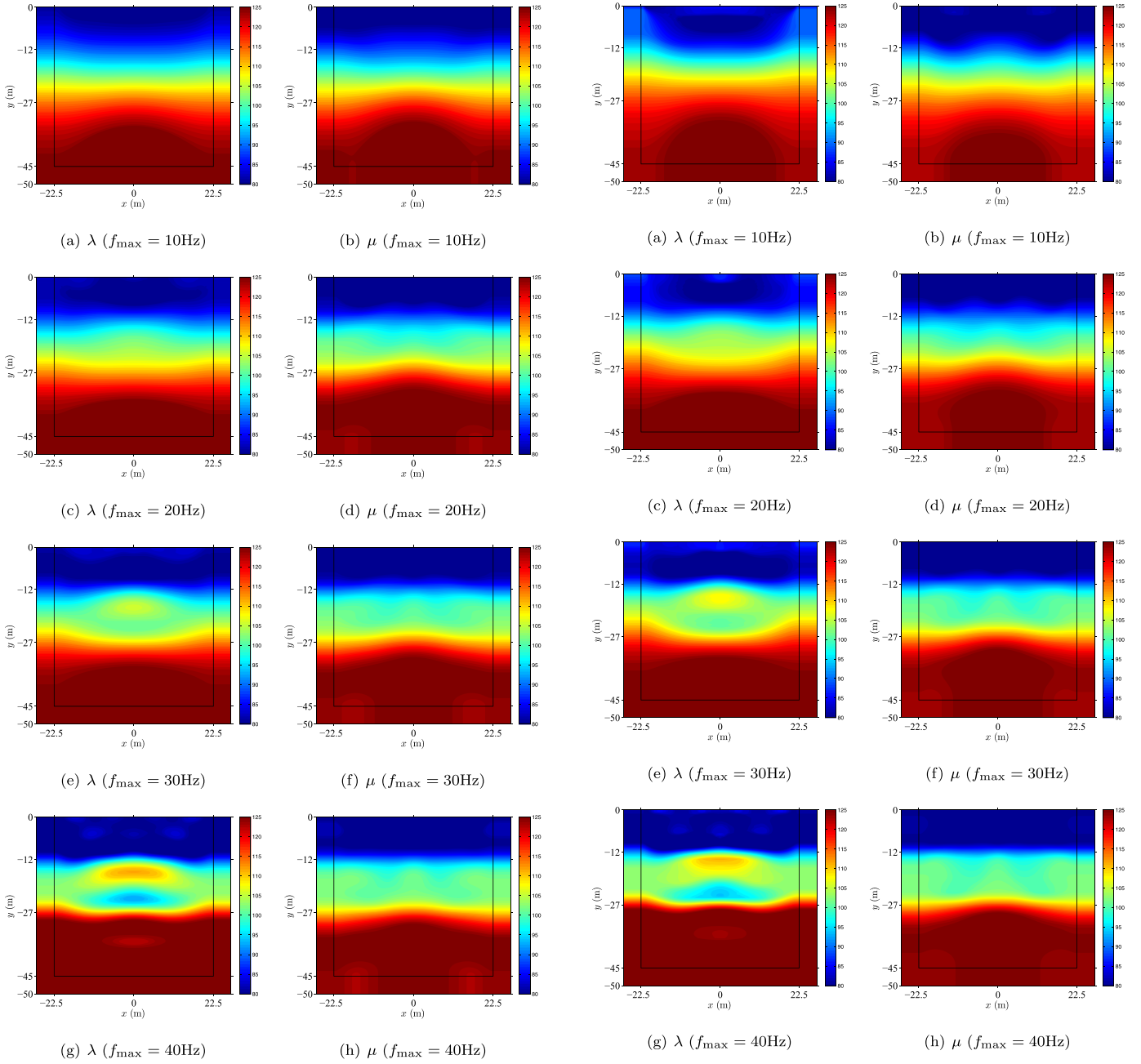
(h) μ ($f_{\max} = 40$Hz)

Fig. 11. Simultaneously inverted Lamé parameters using TN regularization; 10 Hz source (105 iterations); 20 Hz source (124 iterations); 30 Hz source (145 iterations); and 40 Hz source (189 iterations).

(a) λ ($f_{\max} = 10$Hz)

(b) μ ($f_{\max} = 10$Hz)

(c) λ ($f_{\max} = 20$Hz)

(d) μ ($f_{\max} = 20$Hz)

(e) λ ($f_{\max} = 30$Hz)

(f) μ ($f_{\max} = 30$Hz)

(g) λ ($f_{\max} = 40$Hz)

(h) μ ($f_{\max} = 40$Hz)

Fig. 12. Simultaneously inverted Lamé parameters using TV regularization; 10 Hz source (145 iterations); 20 Hz source (180 iterations); 30 Hz source (205 iterations); and 40 Hz source (255 iterations).

Initially, we assign full weight ($W = 1$) to μ, but as iterations evolve, the weight is reduced down to zero ($W = 0$), thereby letting λ evolve on its own. Once the updates on μ vanish numerically (μ has converged), we fix μ, and continue the inversion process with λ only.

4.2. Regularization factor continuation (RS2)

The amount of penalty placed on the gradients via the regularization factors R_λ and R_μ is critical, since the inversion process depends on the reduced gradients. Though there are various developments for choosing the regularization factor intelligently (e.g., L-curve (Hansen and O'Leary, 1993; McCarthy, 2003)), a fixed value for R_λ and R_μ can also be used. Here, we favor adjusting the reg-

ularization factor dynamically so that it penalizes high-frequency material oscillations initially, but as iteration grow, it relaxes the imposed penalty in order to be able to refine the target reconstructions. To address the manner by which the regularization factors are continuously updated, we first recast the continuous form of the reduced gradients (19–22) as

$$\nabla_i \mathcal{L} = - R_i \, \mathbb{K} \, - \, \mathbb{F}, \tag{26}$$

where the subscript i refers to either of the two control/material parameters λ or μ, $\mathbb{K}$ corresponds to the part of the reduced gradient stemming from the regularization, and $\mathbb{F}$ corresponds to the part that originates in the side-imposed PDEs. Once discretized, $\mathbb{K}$ and $\mathbb{F}$ denote n-dimensional vectors with n being the number of nodal unknowns in Ω^{ID}. Representing each vector as a product of
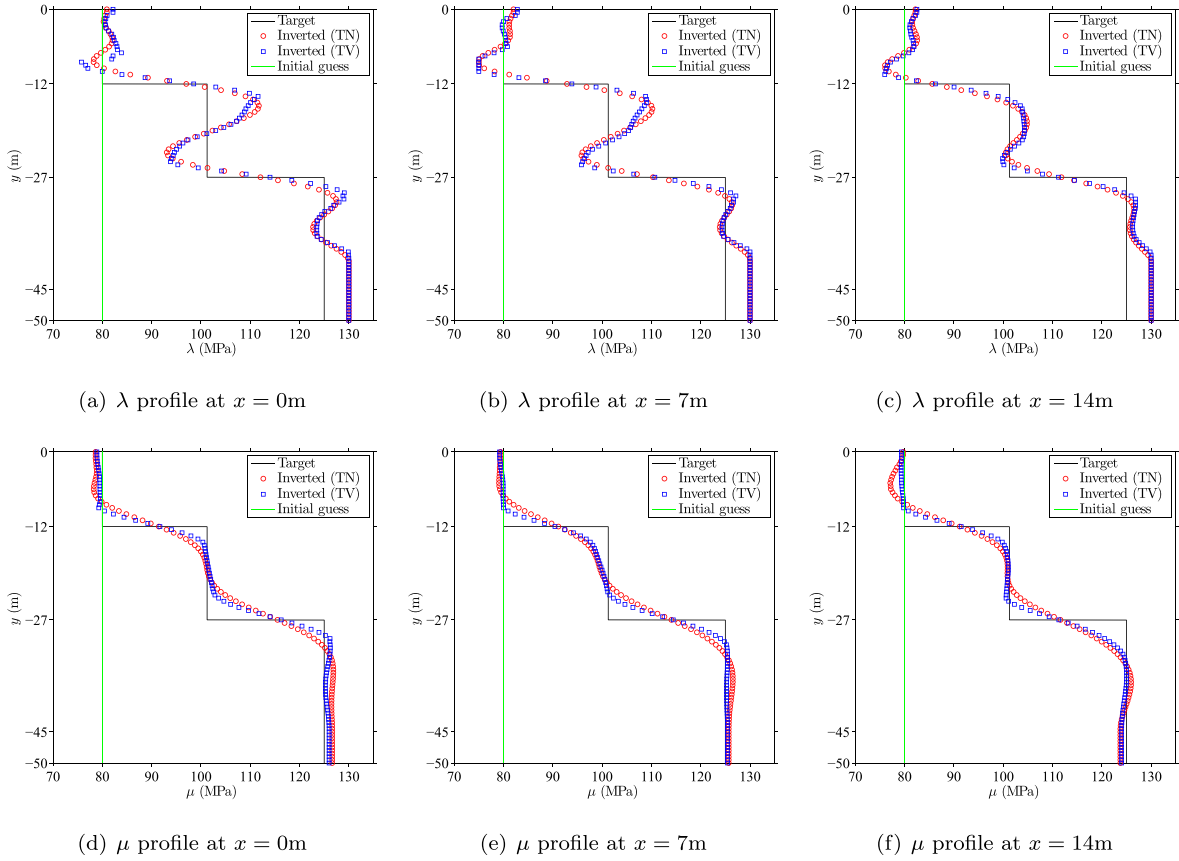
(a) λ profile at $x = 0$m

(b) λ profile at $x = 7$m

(c) λ profile at $x = 14$m

(d) μ profile at $x = 0$m

(e) μ profile at $x = 7$m

(f) μ profile at $x = 14$m

Fig. 13. Cross-sectional profiles for simultaneously inverted Lamé parameters.
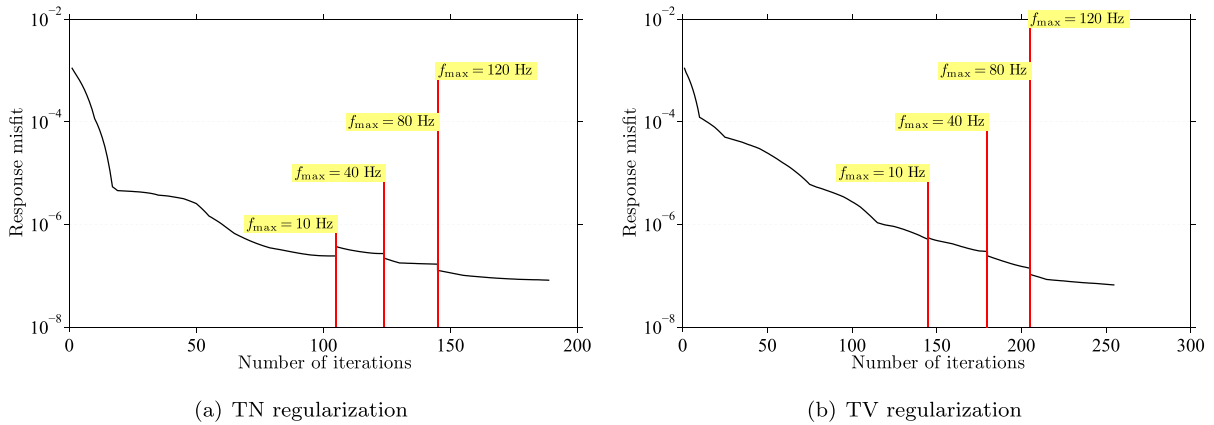

(a) TN regularization

(b) TV regularization

Fig. 14. Variation of response misfit during the simultaneous inversion endowed with the source-frequency continuation scheme.

its magnitude by the associated unit vector, yields

$$(\nabla_i \mathcal{L})_k = -\|\mathbb{F}\| \left[R_i \ \frac{\|\mathbb{K}\|}{\|\mathbb{F}\|} \mathbf{n}_{\mathbb{K}} \ + \ \mathbf{n}_{\mathbb{F}} \right], \tag{27}$$

where $\|\cdots\|$ denotes the Euclidean norm, and $\mathbf{n}_j$ is the unit vector along j. Eq. (27) shows explicitly that the discrete reduced gradient, which drives the conjugate gradient method, is the weighted average of the gradient information coming from the regularization term ($\mathbb{K}$), and the side-imposed PDEs ($\mathbb{F}$). Depending on the weight, i.e., the factor in front of $\mathbf{n}_{\mathbb{K}}$, we decide on what should control the search direction: having a weight that is too small produces high-frequency fluctuations, and the inversion process suffers from solution multiplicity. By contrast, placing too much weight on the regularization smoothens the reconstructed mate-

rial profile by over-penalizing the gradients. Therefore, $\mathbf{n}_{\mathbb{K}}$ needs to be scaled continuously so that it regularizes $\mathbf{n}_{\mathbb{F}}$ without hindering the overall search direction. This can be achieved by forcing $\mathbf{n}_{\mathbb{K}}$ to compete with $\mathbf{n}_{\mathbb{F}}$ throughout the entire inversion process, that is,

$$R_i \ \frac{\|\mathbb{K}\|}{\|\mathbb{F}\|} = \wp, \tag{28}$$

where $\wp$ $(0 \le \wp \le 1)$ is a tuning parameter. The above can be recast to choose R_i as

$$R_i = \wp \ \frac{\|\mathbb{F}\|}{\|\mathbb{K}\|}, \qquad 0 \le \wp \le 1, \tag{29}$$

at each iteration. By this continuation scheme, the regularization factor can take a large value at the beginning of the inversion pro-
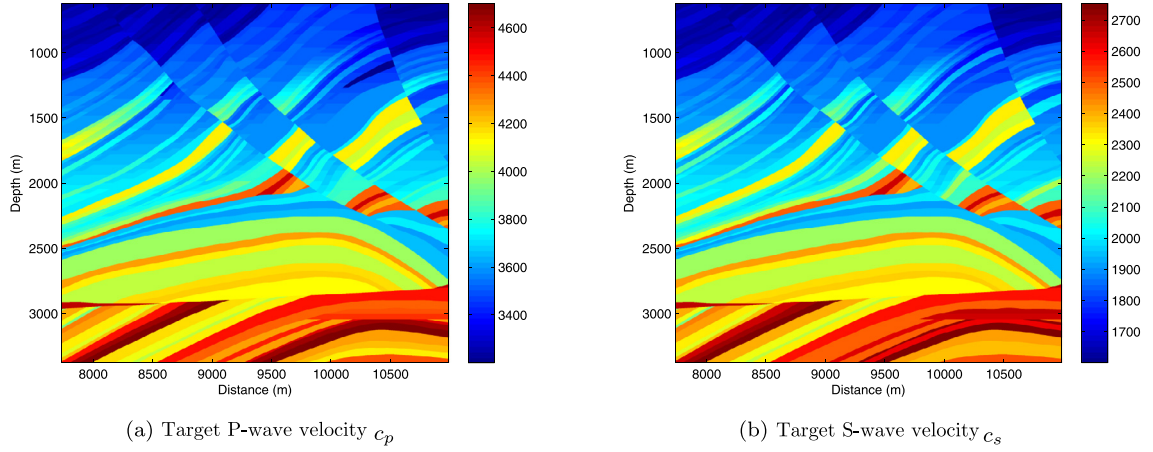
(a) Target P-wave velocity c_p

(b) Target S-wave velocity c_s

Fig. 15. Truncated and scaled computational model of the Marmousi2 problem.

(a) c_p ($f_{max} = 1.5$Hz)

(b) c_s ($f_{max} = 1.5$Hz)

(c) c_p ($f_{max} = 5$Hz)

(d) c_s ($f_{max} = 5$Hz)

(e) c_p ($f_{max} = 10$Hz)

(f) c_s ($f_{max} = 10$Hz)

(g) c_p ($f_{max} = 20$Hz)

(h) c_s ($f_{max} = 20$Hz)
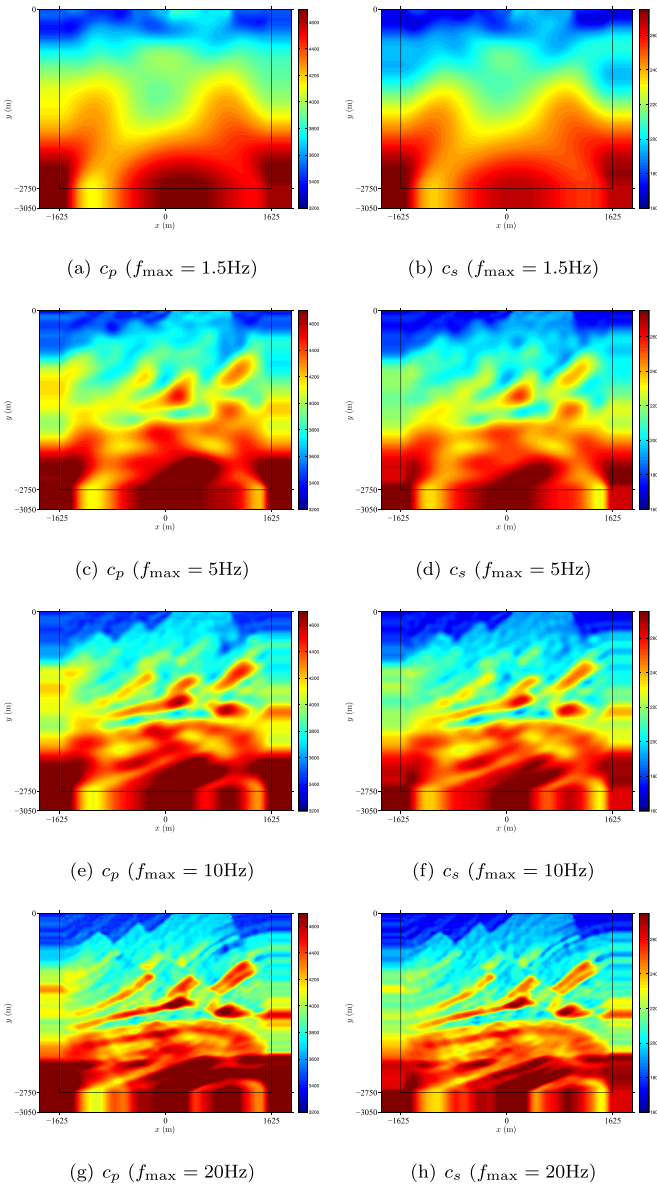
Fig. 16. c_p and c_s - obtained from simultaneously inverted Lamé parameters using TV regularization; 1.5 Hz source (490 iterations); 5 Hz source (1170 iterations); 10 Hz source (920 iterations); 20 Hz source (1590 iterations).

cess (to narrow down the initial feasibility space of the solution), and it is continuously reduced as iterations progress so as to enable refinement of the reconstructions.

4.3. Source-frequency continuation (RS3)

We start the probing with a signal having a low-frequency content, and increase the content frequency in subsequent iterations, until convergence. The key idea is that a low-frequency excitation typically allows for a rough resolution of the material profile, whereas an excitation with higher-frequency components fine-tunes the profile. In practice, we need only a few probing signals to arrive at a converged profile. Starting with the lowest source frequency, the inversion process is let to converge to a profile that captures the overall spatial variability of the material parameters in a rather crude way. Next, we feed this coarse profile as an initial guess to the inversion process driven by the next higher frequency in the sequence, and let again the optimizer to arrive at a converged profile. The process is repeated for all probing signals.

5. Numerical experiments

To test the robustifying schemes, we discuss next numerical experiments, involving arbitrarily heterogeneous hosts and synthetic data. The first example is a fictitious medium that has material properties varying smoothly with depth. We use the first example to discuss the effect of the biased search directions scheme (RS1). Example 2 involves a horizontally-layered medium, and we use it to discuss the effect of the source-frequency continuation scheme (RS3) in the presence of sharp layer interfaces. Lastly, Example 3 tackles the challenging Marmousi2 (Martin et al., 2006) problem, which has been used extensively in the literature as a benchmark. The regularization factor continuation scheme (RS2) is used in all examples.

5.1. Example 1 - effect of RS1+RS2

We consider first a heterogeneous half-plane with constant density $\rho = 2000$kg/m^3 and constant Poisson's ratio $\nu = 0.25$ (both assumptions are physically consistent for near-surface deposits). We reduce the half-plane, through truncation, to a 45 m × 45 m computational domain, surrounded on its sides and bottom by a 5 m-thick PML, as shown in Fig. 2(a). The material profile varies smoothly with depth; specifically, we define the spatial variation
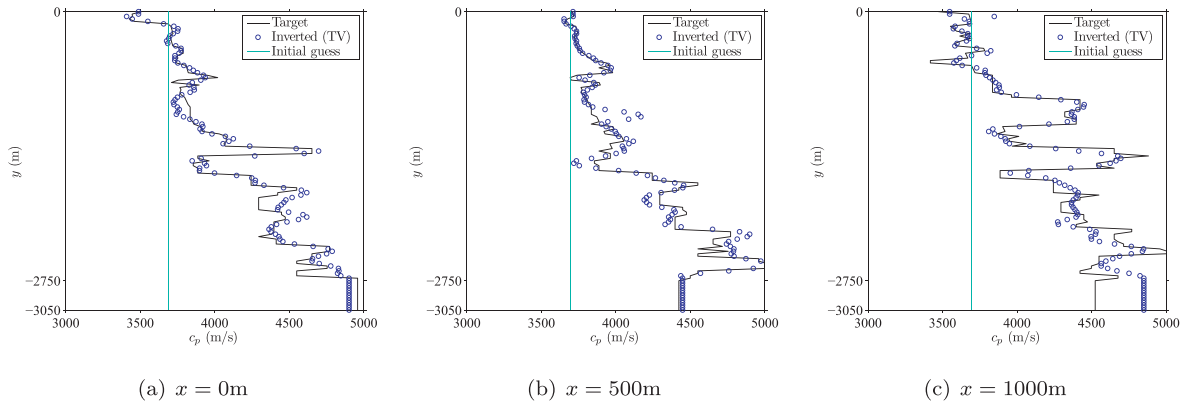
(a) $x = 0$m (b) $x = 500$m (c) $x = 1000$m

Fig. 17. c_p profiles obtained from simultaneously inverted Lamé parameters.

of the Lamé parameters as

$$\lambda(y) = \mu(y) = 80 + 0.45|y| + 35 \exp\left[-\frac{(|y| - 22.5)^2}{150}\right] \text{ (MPa)},$$

$$(30)$$

The material interfaces were extended horizontally into the PML, thereby, avoiding sudden material changes at the interface between the PML and the regular domain. The PML and regular domains were discretized by square 0.5 m × 0.5 m 8-noded serendipity elements. We used a quadratic-quadratic pair for the displacements and stresses, whereas the material properties λ and μ were approximated linearly. The discretization resulted in a 10-cell-thick PML with a quadratic attenuation profile.

To recover the target material profiles depicted in Fig. 2(c), we apply Gaussian pulse point loads at every grid point (uniformly distributed with 0.25m spacing) on the surface of the regular domain, with a maximum frequency of $f_r = 40$ Hz and amplitude of 50 kPa (Fig. 3). A Gaussian pulse is defined as $T_p(t) = -50 \exp\left[-(t - \bar{\mu})^2/\bar{\sigma}^2\right]$, where $\bar{\mu}$ and $\bar{\sigma}$ denote the mean and deviation of the Gaussian pulse. The receivers that measure the displacement response $\mathbf{u}(\mathbf{x}, t)$ were also located at every grid point on the surface resulting in 179 receiver locations. Using a time step of 0.001 seconds, we let the forward problem run for 1.0 second under the target material profile, but using a different (refined) mesh, in order to obtain the synthetic data at the sensor locations.

First, we attempt blind simultaneous inversions following the procedure described in Algorithm 1 but without using the RS1 (or RS3). We start the inversion process with a homogeneous profile that has both λ and μ set to 80 MPa. The top row in Fig. 4 depicts the target profiles, whereas the center and bottom rows correspond to *simultaneous* two-parameter inversion results obtained by using TN and TV regularization, respectively. We observe that μ is recovered satisfactorily, but λ is not. The same observation can be made from the cross-sectional profiles (Fig. 5). Fig. 6 shows the variation of the response misfit during the inversion process.

Next, we apply the biased search-direction scheme, RS1 (25). We start with the same initial profile as previously used for blind inversion. The result is shown in Fig. 7. We remark that the spatial variability of μ is captured much more accurately than the spatial variability of λ. As evidenced also from the cross-sectional profiles depicted in Fig. 8, the inversion scheme did an excellent job in capturing the smoothly-varying target μ profile and quite satisfactorily reconstructed the λ profile, a marked improvement over the results shown in Fig. 4(c) or (e).

5.2. Example 2 - effect of RS1+RS2+RS3

In this example, we use a source-frequency continuation scheme (RS3) according to which a few time signals with different frequency content are used to probe the domain, thereby creating a set of, seemingly, uncoupled inversion problems. We start the inversion process with a low-frequency source and feed the converged reconstructed λ and μ profiles as initial guesses to the problem excited with a higher-frequency source. This procedure was repeated as many times as the number of independent sources probing the domain: here, we considered four different Gaussian pulses with maximum frequencies ranging from 10 Hz to 40 Hz, as shown in Fig. 9.

As shown in Fig. 10(a), we consider next a 45 m × 45 m layered medium surrounded by 5 m-thick PML on its sides and bottom with constant density $\rho = 2000$kg/m^3 and constant Poisson's ratio $\nu = 0.25$. We define the spatial variation of the Lamé parameters as

$$\lambda(y) = \mu(y) = \begin{cases} 80 \text{ MPa,} & \text{for } -12 \text{ m} \leq y \leq 0 \text{ m,} \\ 101.25 \text{ MPa,} & \text{for } -27 \text{ m} \leq y < -12 \text{ m,} \\ 125 \text{ MPa,} & \text{for } -50 \text{ m} \leq y < -27 \text{ m.} \end{cases} \quad (31)$$

As in Example 1, the material interfaces were extended horizontally into the PML. The PML and domain parameterization (mesh, PML thickness, etc.) are the same as in Example 1.

Following the inversion procedure outlined in Algorithm 1 endowed with RS1 and RS2, we reconstruct the material profiles shown in Figs. 11 and 12 using the TN and TV regularization schemes, respectively. In each figure, the first row displays the reconstructed profiles when the domain was probed by a source signal with $f_{max} = 10$ Hz. These profiles were fed as an initial guess to the next problem in the sequence, where the probing source now increased to $f_{max} = 20$ Hz. Upon inversion, we obtained the profiles presented in the second row. Continuing in this manner, we obtained the converged reconstructed profiles for λ and μ shown in the last row corresponding to a source with $f_{max} = 40$Hz. The effect of source-frequency continuation is clearly visible: increasing the source frequency results in refinement of the recovered profiles. Here, the refinement is in the form of localization and sharpening of the layer interfaces, which has been achieved quite satisfactorily under both TN and TV regularization schemes.

Fig. 13 compares the inverted profiles with the target profile at the $x = 0$m, 7m, and 14m cross-sectional lines of the domain. The agreement is excellent for the μ profiles, whereas the λ profiles seem to fluctuate around the target even with the proposed biased-directions inversion algorithm. We are inclined to blame the objective functional being less-sensitive to λ, since varying λ has a mild effect on the misfit provided that it captures, in an average
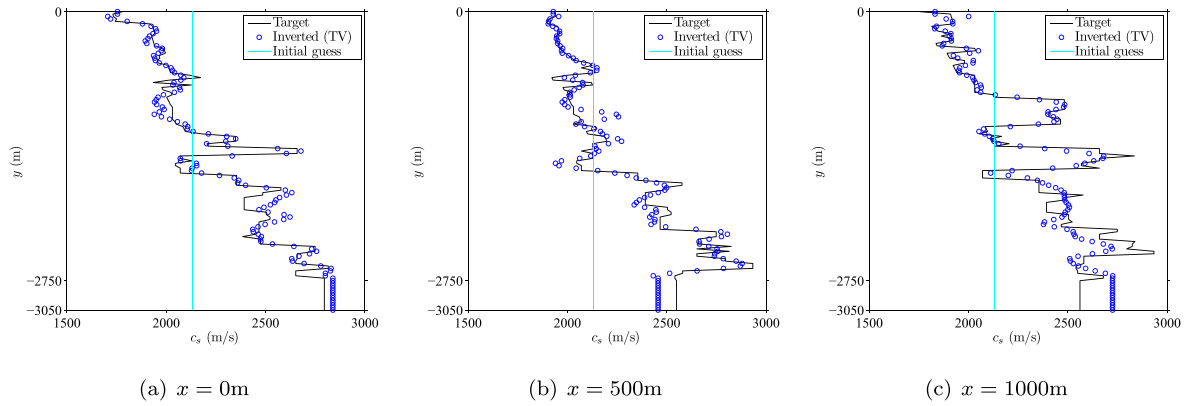
Fig. 18. c_s profiles obtained from simultaneously inverted Lamé parameters.

sense, the target profile. As stated in the earlier discussions, the TV regularization scheme is expected to perform better than the TN regularization in recovering the *sharply*-varying target profiles, and indeed, this is the case here, especially with the reconstructed μ profiles. One may sharpen the layer interfaces further by increasing the source-frequency content.

In Fig. 14, the response misfit was reduced from its initial value of 0.134×10^{-2} down to 0.819×10^{-7} and 0.662×10^{-7} for the TN and TV regularizations, respectively; this corresponds to about 0.006% and 0.005% of the initial misfit, respectively. Since the TV regularization captures the sharp layer interfaces better, the final value of the corresponding misfit is lower than the one attained by the TN regularization.

5.3. Example 3 - Marmousi2 model

We try next to recover the velocity profiles of the highly-heterogeneous Marmousi2 model (Martin et al., 2006) shown in Fig. 15. It is an elastic extension to the original acoustic Marmousi model defined in Versteeg (1994). The synthetic data for the Marmousi2 model are created for a 17km × 3.5km region. Here, we opted for a reduced and scaled version of the Marmousi2 model to reduce the computational cost. To this end, we truncated the original Marmousi2 domain down to a 3250 m × 2750 m domain, as shown in Fig. 15, and scaled the wave velocities to be in the range:

$$3200 \text{ m/s} \leq c_p \leq 4700 \text{ m/s} \quad \text{and} \quad 1600 \text{ m/s} \leq c_s \leq 2752 \text{ m/s}, \tag{32}$$

where c_p and c_s denote the P-wave and S-wave velocities, respectively.nd The truncated domain is obtained by extracting a box between 7736.25 m and 10986.25 m (horizontal) and 626.25 m and 3376.25 m (depth) of the original Marmousi2 model. The density is assumed to be constant and taken as $\rho = 2500$ kg/m^3.

As in previous examples, the material interfaces were extended horizontally into the PML. The PML and interior domains were discretized by quadrilateral elements with an element size of 25 m. We used linear elements with low probing frequencies to reduce the per-iteration time cost, and then, switched to quadratic elements with high probing frequencies to resolve the source loading better. We used a linear-linear (quadratic-quadratic) pair for the displacements and stresses, and linear approximations for the material properties λ and μ. The discretization resulted in a 14-cell-thick PML with a quadratic attenuation profile.

Note that to recover the target velocity profiles shown in Fig. 15, we still invert for the Lamé parameters and compute the associated wave velocities to display the reconstructed profiles. We use a source-frequency continuation scheme with six distinct
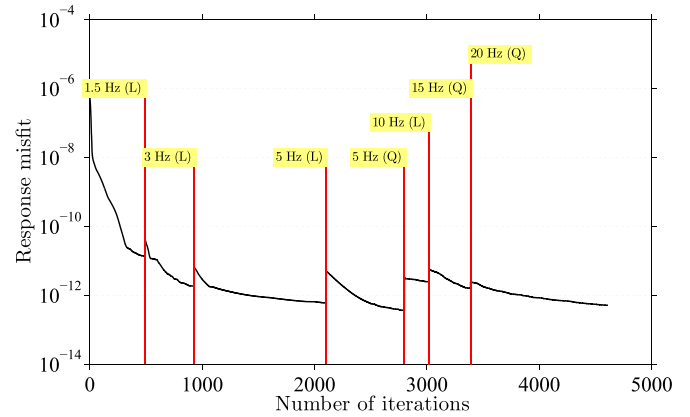

Fig. 19. Variation of response misfit during simultaneous inversion process with source-frequency continuation (L - linear elements, Q - quadratic elements).

Gaussian pulses: $f_{max} = 1.5$, 3, 5, 10, 15, and 20 Hz to probe the domain. The receivers were located at every grid point on the surface resulting in 259 receiver locations. Using a time step of 0.005 seconds, we let the forward problem run for 8.0 seconds under the target material profile, but use a different (refined) mesh to generate the synthetic data.

Here we opted to use the TV regularization only. We initiated the inversion process with a homogeneous medium that has $\lambda = \mu = 10000$ MPa, and used the time signal associated with the lowest source frequency ($f_{max} = 1.5$ Hz) to probe the domain. Accordingly, we reconstructed the material profiles shown in Fig. 16(a) and (b) corresponding to c_p and c_s, respectively. The reconstructed profiles capture the layering of the domain to some extent and were able to localize the high and low velocity regions. Though the quality seems rather poor, we note that the minimum wavelength in this setting was 1067 m. To improve the resolution, the source frequency is increased. To this end, we fed these profiles as an initial guess to the next problem, where the domain was excited by a Gaussian pulse with $f_{max} = 3$ Hz. Continuing along the same lines with the source-frequency continuation scheme, we increased the source frequency to $f_{max} = 5$ Hz and then to 10 Hz, concluding with 20 Hz Gaussian pulses, fed the previous reconstructed profiles as an initial guess, and obtained the inverted profiles shown in Fig. 16(c) and (d). The reconstructed profiles are depicted in Fig. 16(g) and (h), where further sharpening of the material interfaces has been achieved.

Figs. 17 and 18 compare the inverted profiles with the target profile at the $x = 0$ m, 500 m, and 1000 m cross-sectional lines of the domain. It is apparent from these figures that the ar-

bitrary heterogeneity is recovered satisfactorily. Further improvements around the material interfaces can be achieved by using a finer computational grid with an increased source frequency. Lastly, Fig. 19 depicts the variation of the response misfit with the number of iterations.

6. Conclusions

We discussed algorithmic improvements to a full-waveform-based inversion methodology for the reconstruction of the material properties of elastic solids when probed by stress waves. In particular, we focused on three robustifying schemes for alleviating the solution multiplicity associated with the underlying inverse problem and for improving parameter reconstruction. While the regularization factor continuation (RS2), and the source-frequency continuation scheme (RS3) assisted the optimizer in narrowing the initial solution space in a systematic manner, the search-direction biasing scheme (RS1) had the greatest effect in improving the resolution under joint inversion conditions. The numerical results highlighted the effects of the robustifying schemes, while the application of the developed inversion process to the highly heterogeneous and challenging Marmousi2 model resulted in a reconstructed profile in fairly good agreement with the target distribution.

Acknowledgment

Appendix A. Inversion process algorithm

The standard inversion process without robustifying schemes is summarized in Algorithm 1.

Algorithm 1 Inversion process for material profile reconstruction.

1: **procedure** MPR($\boldsymbol{\lambda}_o$, $\boldsymbol{\mu}_o$)   ▷ initial guesses
2:     Choose $\rho, \delta^\lambda, \delta^\mu$ ▷ e.g., $\rho = 0.5$, $\delta^\lambda = \delta^\mu = 10^{-10}$
3:     Set convergence tolerance tol ▷ e.g., $tol = 10^{-12}$
4:     $k \leftarrow 0$ ▷ k: iteration counter
5:     Set $\mathcal{F}_k \leftarrow tol + 1$ ▷ $\mathcal{F}_k$ denotes $\mathcal{F}(\boldsymbol{\lambda}_k, \boldsymbol{\mu}_k)$
6:     **while** $\{\mathcal{F}_k > tol\}$ **do**
7:        Solve the state problem for $(\mathbf{u}, \mathbf{S})$ ▷ eq. (1a-2e)
8:        Solve the adjoint problem for $(\boldsymbol{\theta}_u, \boldsymbol{\theta}_s)$ ▷ eq. (10a-11h)
9:        Evaluate the discrete reduced gradients
10:           $\mathbf{g}_k^\lambda \leftarrow (\nabla_\lambda \mathcal{L})_k$ ▷ eq. (19) for TN, eq. (21) for TV
11:           $\mathbf{g}_k^\mu \leftarrow (\nabla_\mu \mathcal{L})_k$ ▷ eq. (20) for TN, eq. (22) for TV
12:        Compute the search directions $\mathbf{d}_k^\lambda$ and $\mathbf{d}_k^\mu$ ▷ eq. (20)
13:        $\alpha^\lambda \leftarrow s^\lambda \, \text{mean}[\mathbf{d}_k^\lambda]^{-1}$ ▷ initial step size for λ
14:        $\alpha^\mu \leftarrow s^\mu \, \text{mean}[\mathbf{d}_k^\mu]^{-1}$ ▷ initial step size for μ
15:        **while** $\left\{ \mathcal{F}_{k+1} - \mathcal{F}_k > \min\left[\delta^\lambda \alpha^\lambda (\mathbf{g}_k^\lambda)^T \mathbf{d}_k^\lambda, \; \delta^\mu \alpha^\mu (\mathbf{g}_k^\mu)^T \mathbf{d}_k^\mu \right] \right\}$ **do**
16:           $\alpha^\lambda \leftarrow \rho \alpha^\lambda$ **if** $\left[\mathcal{F}_{k+1} - \mathcal{F}_k > \delta^\lambda \alpha^\lambda (\mathbf{g}_k^\lambda)^T \mathbf{d}_k^\lambda \right]$
17:           $\alpha^\mu \leftarrow \rho \alpha^\mu$ **if** $\left[\mathcal{F}_{k+1} - \mathcal{F}_k > \delta^\mu \alpha^\mu (\mathbf{g}_k^\mu)^T \mathbf{d}_k^\mu \right]$
18:        **end while**
19:        $\alpha_k^\lambda \leftarrow \alpha^\lambda$
20:        $\alpha_k^\mu \leftarrow \alpha^\mu$
21:        Update the material property vectors $\boldsymbol{\lambda}_k$ and $\boldsymbol{\mu}_k$ ▷ eq. (23)
22:        $k \leftarrow k + 1$
23:     **end while**
24: **end procedure**

References

Bamberger, A., Chavent, G., Lailly, P., 1979. About the stability of the inverse problem in 1-D wave equation - application to the interpretation of seismic profiles. Appl. Math. Optim. 5, 1–47.

Brenders, A.J., Pratt, R.G., 2007. Full waveform tomography for lithosperic imaging: results from a blind test in a realistic crustal model. Geophys. J. Int. 168, 133–151.

Brossier, R., Operto, S., Virieux, J., 2009. Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. Geophysics 74 (6), WCC105–WCC118.

Bunks, C., Saleck, F., Zaleski, S., Chavent, G., 1995. Multiscale seismic waveform inversion. Geophysics 60, 1457–1473.

Chavent, G., Jacewitz, C.A., 1995. Determination of background velocities by multiple migration fitting. Geophysics 60 (2), 476–490.

Choi, Y., Min, D., Shin, C., 2008. Frequency-domain elastic full waveform inversion using the new pseudo-hessian matrix - experience of elastic marmousi-2 synthetic data. Bull. Seismol. Soc. Am. 98 (5), 2402–2415.

Choi, Y., Shin, C., Min, D., Ha, T., 2005. Efficient calculation of the steepest descent direction for source-independent seismic waveform inversion: an amplitude approach. J. Comput. Phys. 208, 455–468.

Clément, F., Chavent, G., Gómez, S., 1995. Waveform inversion by MBTT formulation. In: Proceedings of the 3rd International Conference on Mathematical and Numerical Aspects of Wave Propagation, Philadelphia, PA, pp. 713–722.

Epanomeritakis, I., Akçelik, V., Ghattas, O., Bielak, J., 2008. A newton-CG method for large-scale three-dimensional elastic full-waveform seismic inversion. Inverse Probl. 24 (3), 34015–34040.

Epanomeritakis, I.K., 2004. Identification of elastic basin properties by large-scale inverse earthquake wave propagation. Carnegie Mellon University Ph.D. thesis.

Fathi, A., 2015. Full-waveform inversion in three-dimensional PML-truncated elastic media: theory, computations, and field experiments. The University of Texas at Austin Ph.D. thesis.

Fathi, A., Kallivokas, L.F., Poursartip, B., 2015. Full-waveform inversion in three-dimensional PML-truncated elastic media. Comput. Methods Appl. Mech. Eng. 296, 39–72.

Fathi, A., Poursartip, B., Stokoe, K.H., Kallivokas, L.F., 2016. Three-dimensional p- and s-wave velocity profiling of geotechnical sites using full-waveform inversion driven by field data. Soil Dyn. Earthquake Eng. 87, 63–81.

Fletcher, R., Reeves, C., 1964. Function minimization by conjugate gradients. Comput. J. 7 (2), 149–154.

Gao, F., Levander, A., Pratt, R.G., Zelt, C.A., Fradelizio, G.-L., 2007. Waveform tomography at a groundwater contamination site: surface reflection data. Geophysics 72, G45–G55.

Gladwell, G., 1984. The inverse problem for the vibrating beam. In: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 393. The Royal Society, pp. 277–295.

Hansen, P., O'Leary, D., 1993. The use of the l-curve in the regularization of discrete ill-posed problems. SIAM J. Sci. Comput. 14 (6), 1487–1503.

Hughes, T., 1987. The finite element method. Prentice-Hall, Englewood Cliffs, NJ.

Juang, J.-N., Pappa, R.S., 1985. An eigensystem realization algorithm for modal parameter identification and model reduction. J. Guid. 8 (5), 620–627.

Kallivokas, L., Fathi, A., Kucukcoban, S., Stokoe, K., Bielak, J., Ghattas, O., 2013. Site characterization using full waveform inversion. Soil Dyn. Earthquake Eng. 47 (Supplement C), 62–82. SI: Jos Manuel Rosset.

Kang, J.W., 2010. A mixed unsplit-field PML-based scheme for full waveform inversion in the time-domain using scalar waves. The University of Texas at Austin, Austin, TX, USA Ph.D. thesis.

Kang, J.W., Kallivokas, L.F., 2011. The inverse medium problem in heterogeneous PML-truncated domains using scalar probing waves. Comput. Methods Appl. Mech. Eng. 200 (1), 265–283.

Kucukcoban, S., 2010. The inverse medium problem in PML-truncated elastic media. The University of Texas at Austin Ph.D. thesis.

Kucukcoban, S., Kallivokas, L.F., 2011. Mixed perfectly-matched-layers for direct transient analysis in 2D elastic heterogeneous media. Comput. Methods Appl. Mech. Eng. 200 (1–4), 57–76.

Kucukcoban, S., Kallivokas, L.F., 2013. A symmetric hybrid formulation for transient wave simulations in PML-truncated heterogeneous media. Wave Motion 50 (1), 57–79.

Martin, G.S., Wiley, R., Marfurt, K.J., 2006. Marmousi2: an elastic upgrade for marmousi. Lead. Edge 156–166.

McCarthy, P., 2003. Direct analytic model of the l-curve for tikhonov regularization parameter selection. Inverse Probl. 19, 643–663.

Métivier, L., Bretaudeau, F., Brossier, R., Operto, S., Virieux, J., 2014. Full waveform inversion and the truncated newton method: quantitative imaging of complex subsurface structures. Geophys. Prospect. 62 (6), 1353–1375.

Na, S.-W., Kallivokas, L.F., 2008. Continuation schemes for shape detection in inverse acoustic scattering problems. Comput. Model. Eng. Sci. 35 (1), 73–90.

Operto, S., Ravaut, C., Improta, L., Virieux, J., Herrero, A., Dell'Aversana, P., 2004. Quantitative imaging of complex structures from dense wide-aperture seismic data by multiscale traveltime and waveform inversions: a case study. Geophys. Prospect. 52 (6), 625–651.

Plessix, R.-E., De Roeck, Y.-H., Chavent, G., 1999. Waveform inversion of reflection seismic data for kinematic parameters by local optimization. SIAM J. Sci. Comput. 20 (3), 1033–1052.

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

Pratt, R.G., Shipp, R.M., 1999. Seismic waveform inversion in the frequency domain, part 2: fault delineation in sediments using crosshole data. Geophysics 64 (3), 902–914.

Pratt, R.G., Song, Z.-M., Williamson, P.R., Warner, M.R., 1996. Two-dimensional velocity models from wide angle seismic data by waveform inversion. Geophys. J. Int. 124, 323–340.

Rudin, L., Osher, S., Fatemi, E., 1992. Nonlinear total variation based noise removal algorithms. Phys. D 60, 259–268.

Salawu, O., 1997. Detection of structural damage through changes in frequency: a review. Eng. Struct. 19 (9), 718–723.

Shin, C., Min, D.-J., 2006. Waveform inversion using a logarithmic wavefield. Geophysics 71, R31–R42.

Symes, W.W., 2008. Migration velocity analysis and waveform inversion. Geophys. Prospect. 56 (6), 765–790.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation. Geophysics 49, 1259–1266.

Tikhonov, A., 1963. Solution of incorrectly formulated problems and the regularization method. Soviet Math. Doklady 4, 1035–1038.

Versteeg, R., 1994. The marmousi experience: velocity model determination on a synthetic complex data set. Lead. Edge 13, 927–936.

Virieux, J., Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics. Geophysics 74 (6), WCC1–WCC26.