ORIGINAL PAPER

On the inverse problem of soil profile reconstruction: a comparison of time-domain approaches

Seong-Won Na · Loukas F. Kallivokas

Abstract We discuss a one-dimensional inverse material profile reconstruction problem that arises in layered media underlain by a rigid bottom, when total wavefield surficial measurements are used to guide the reconstruction. To tackle the problem, we adopt the systematic framework of PDE-constrained optimization and construct an augmented misfit functional that is further endowed by a regularization scheme. We report on a comparison of spatial regularization schemes such as Tikhonov and total variation against a temporal scheme that treats the model parameters as time-dependent. We study numerically the effects of inexact initial estimates, data noise, and regularization parameter choices for all three schemes, and report inverted profiles for the modulus, and for simultaneous inversion of both the modulus and viscous damping. Our numerical experiments demonstrate comparable or superior performance of the time-dependent regularization over the Tikhonov and total variation schemes for both smooth and sharp target profiles, albeit at increased computational cost.

Keywords PDE-constrained optimization · Inverse problem · Material identification · Regularization

S.-W. Na · L. F. Kallivokas (✉)
Department of Civil, Architectural and
Environmental Engineering,
The University of Texas at Austin,
1 University Station, C1748, Austin,
TX 78712, USA
e-mail: loukas@mail.utexas.edu

1 Introduction

The reconstruction of the material profile of layered soils based on surficial measurements collected as either the response to dynamic loads imparted on the soil surface, or as seismic records, is central to geotechnical site characterization efforts. The same problem, albeit at considerably different length scales, is also of primary importance to seismic hazard mitigation efforts, to soil-structure-interaction problems, and to geophysics applications (e.g., discovery of hydrocarbon deposits). An often-used frequency-domain technique for geotechnical site investigations is based on the SASW method (spectral analysis of surface waves [1]), which relies on the analysis of surface Rayleigh waves for determining the dispersion curve that, in turn, leads to the shear wave velocity profile. The method however is limited in several ways, not the least of which is the difficulty with the reconstruction of the soil's dissipation characteristics, with the notable exception of the works by Rix et al. [2–4]. Despite SASW's and its variants' adoption and worldwide use, it is still unclear how the method can be modified so that it does not rely on a layered medium assumption, and thus account for arbitrary material heterogeneity.

In this article we discuss a partial-differential-equation (PDE)-constrained optimization approach for recovering the material profile of layered soils directly in the time-domain, including the spatial distribution of elastic modulus as well as of attenuation parameters; by and large the latter have, thus far, proved to be elusive to reconstruct. Though herein too the technical details are restricted to a horizontally layered medium, thus, effectively, giving rise to a one-dimensional problem, the methodology, in contrast to SASW, scales to heterogeneous media in higher spatial dimensions, and can make use of data collected directly in the time-domain.

We start with a conventional misfit functional that describes the difference between computed and measured response, and augment it via the weak imposition of the governing PDEs. Regularization terms are then added to the augmented functional in an attempt to alleviate solution multiplicity. The primary intent of this article is to compare classical approaches such as a Tikhonov [5], and a total variation (TV) scheme [6], against a time-dependent scheme (TD) [15] that treats, initially, the model parameters as time-dependent. Though computationally costly, the TD scheme appears to lend much needed algorithmic robustness, and typically leads to better profiles, without being too sensitive to regularization parameter choices. A partial motivation for the TD scheme stems from the fact that with both Tikhonov and TV schemes the quality of the solution depends highly on a user-selectable regularization parameter (see, e.g., Fig. 9 in [7]), and may not successfully reconstruct profiles for any value of the parameter, especially in the presence of sharp discontinuities or inexact initial estimates. Tadi in [15] reported that he used a TD scheme successfully to inverse problems in population dynamics, which, subsequently, led to his using it also in the reconstruction of the density profile of a one-dimensional rod [15]. Our own motivation for experimenting with a TD scheme stems from our interest in departing from spatial schemes, while still imposing some form of regularization; in this sense, the only other logical candidate was exploring regularization in the "remaining" direction, i.e., in time. As it will be numerically shown, the TD scheme is relatively insensitive to the regularization parameter, and, in most cases, leads to better profiles than spatial schemes.

The PDE-constrained approach followed herein was originally inspired by the optimization work in Navier–Stokes control problems [8], and the similar work on acoustic waves led by Ghattas and his collaborators [9,10]. In [9,10], the authors experimented successfully with material inversion in acoustics using TV regularization schemes. Overall, we are interested in recovering both sharp and smoothly varying profiles and treat the material properties as spatially piecewise continuous: as it will be seen, sharp discontinuities are still recoverable. The recovery of sharply varying properties (or, equivalently, of discontinuous PDE coefficients) has been addressed for elliptic problems (e.g., [7]), but the literature is thin on hyperbolic problems (but see, e.g., [11–13]). Alternatively, the problem could be posed over an arbitrarily predetermined number of layers, whose thicknesses and elastic properties serve as unknowns. For example, in [14] the authors used Love waves for probing, and recovered the (visco)elastic characteristics of the layers; still, the layered assumption remains an obstacle when generalizations that include lateral heterogeneity are of interest.

We choose a reduced-space approach to resolve the first-order optimality conditions associated with stationarity of the augmented functional, which, in turn, result in state, adjoint,
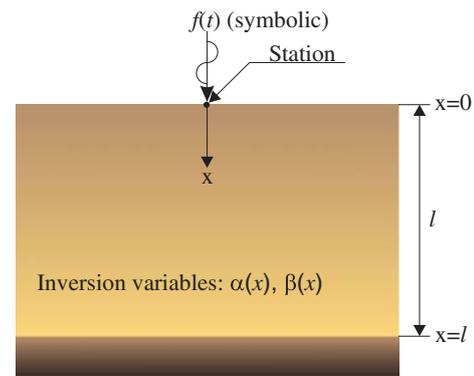
Fig. 1 Problem configuration: soil deposits on bedrock

and control problems (from a temporal character perspective, there result both initial- as well as final-value problems). We report on our numerical experiments that demonstrate robustness in the reconstruction of the material profiles, and the competitiveness of the TD regularization when compared with spatial schemes. We treat both single and dual-parameter systems (modulus and damping), and report on the sensitivity of the reconstructed profiles with respect to inexact estimates, data noise, and regularization parameter choices.

2 The forward problem

We consider the response of a layered (heterogeneous) medium (soil deposits) overlying a homogeneous halfspace (possibly made of rock) to surface excitation. We formally reduce the problem to a one-dimensional one by considering, for example, the case of compressional waves emanating from the surface of the soil due to a uniform excitation applied throughout the entire (two-dimensional) soil surface. Similar physical problems arise if one were to consider only shear waves in the same medium, or compressional waves in a rod [15]. Here, to fix ideas, we shall henceforth restrict the discussion to compressional waves, which allow the reduction of the problem to one dimension; ultimately, our target application is the three-dimensional inversion of highly heterogeneous deposits. In principle, the approach we discuss herein, can be applied to this more complex problem with only minor modifications to account for the higher spatial dimensionality. Let $u(x, t)$ denote the (scalar) displacement in the direction of the applied excitation (Fig. 1). Let l denote the depth of the soil deposits, and T the total observation period. Then, the strong form of the forward problem can be stated as:

Forward Problem
Find $u(x, t)$, such that:

$$\frac{\partial^2 u(x,t)}{\partial t^2} - \frac{\partial}{\partial x}\left(\alpha(x)\frac{\partial u(x,t)}{\partial x}\right) + \beta(x)\frac{\partial u(x,t)}{\partial t} = 0,$$
$$(x, t) \in (0, l) \times (0, T), \tag{1}$$

with

$$\alpha(0)\frac{\partial u(0, t)}{\partial x} = f(t), \tag{2}$$

$$u(l, t) = 0, \tag{3}$$

$$u(x, 0) = \frac{\partial u(x, 0)}{\partial t} = 0, \tag{4}$$

where x denotes location and t denotes time. In the above $\alpha(x)$ denotes the soil's modulus (e.g., $\lambda + 2\mu$ for compressional waves, with λ, μ denoting the Lamé constants), or the square of the wave propagation velocity. Throughout we assume that the material density is constant (a reasonable assumption in geotechnical site investigations); in particular, in (1), we assume, without loss of generality, that the density $\rho = 1$. Furthermore, in (1), $\beta(x)$ represents viscous damping (also normalized with respect to the density). The attenuation character of soil deposits is much more complex than simple viscous damping could ever capture. However, here the emphasis is on the inversion and profile reconstruction process; we expect to tackle more realistic attenuation models in the future. For simplicity, we consider the simple case in which the bottom of the layered medium is fixed (rigid) at $x = l$ (condition (3)). The extension to the more realistic case where the depth of the deposits is unknown and the computational domain need to be truncated will also be communicated in the future. We assume further that the system is initially at rest (condition (4)), and that the source excitation is at the origin (condition (2)). Whereas in the forward problem the excitation and the material distributions $\alpha(x)$ and $\beta(x)$ are known, in the inverse problem of interest herein, both $\alpha(x)$ and $\beta(x)$ are unknown; known, however, is the response $u(0, t), \forall t \in (0, T)$.

3 Inverse misfit problem

For the system defined by the forward problem given in (1)–(4), the inverse problem can be cast as a PDE-constrained optimization problem using a (least-squares) misfit functional, as in:

Minimize:

$$\mathcal{J} = \frac{1}{2} \int_0^T [u(0, t) - u_m(0, t)]^2 \ dt + \mathcal{R}_\alpha(\alpha) + \mathcal{R}_\beta(\beta), \tag{5}$$

subject to (1)–(4).

Here, $\mathcal{J}$ denotes the misfit functional in which u_m denotes the response measured at the surface and u is the computed response obtained for assumed model profiles $\alpha(x)$ and $\beta(x)$. The last two terms in (5) represent regularization terms for α and β, respectively. Their detailed expressions are discussed in the next section.

4 Regularization schemes

As in all inverse problems, the minimization problem (5) associated with the misfit term only is inherently ill-posed due to, at a minimum, the incomplete data set. One way to overcome or alleviate this difficulty is to impose additional constraints in an attempt to "regularize" the solution, if it at all exists. The choice of the regularization scheme is of paramount importance. We discuss candidate schemes:

4.1 Tikhonov regularization

One of the most widely adopted schemes is Tikhonov regularization [5]. A Tikhonov-type regularization enforces spatial smoothness on the model parameters. We experiment with the first-order Tikhonov regularization scheme, according to which the regularization term $\mathcal{R}_p^{Tk}$ is typically defined (in one-dimension), as:

$$\mathcal{R}_p^{Tk}(p) := \frac{R_p}{2} \int_0^l \left(\frac{dp(x)}{dx}\right)^2 \ dx, \tag{6}$$

where R_p is a regularization parameter that controls the amount of the penalty the regularization term (6) imposes on the functional (5), and p is a model parameter (e.g., elastic modulus). Clearly, the above Tikhonov scheme favors smooth profiles since the penalty term becomes smaller (modulo the regularization parameter) for smooth p distributions, whereas it increases with high (spatial) frequency perturbations of the model parameters. Therefore, the Tikhonov scheme works well for smooth target profiles, but is not well-suited to sharply varying target profiles. In addition, the Tikhonov scheme requires initial estimates which are quite close to the target, since the scheme precludes large perturbations from the initial guesses. Lastly, the regularization parameter R_p should be chosen with care, since, as will be shown, the solution is quite sensitive to the choice of the regularization parameter.

4.2 Total variation (TV) regularization

The total variation regularization term $\mathcal{R}_p^{TV}$ is defined as:

$$\mathcal{R}_p^{TV}(p) := R_p \int_0^l \sqrt{\left(\frac{dp(x)}{dx}\right)^2 + \epsilon} \ dx. \tag{7}$$

The scheme is similar to Tikhonov's in enforcing spatial smoothness of the model parameters, but exhibits better performance in dealing with discontinuities or sharp profile changes due to the imposition of a "lesser" penalty associated with discontinuities than the Tikhonov scheme. In the above, we have modified the standard TV form by adding a positive scalar ϵ to smoothen parameter functions that are

not strongly differentiable at discontinuities, as is typically done in TV implementations.

4.3 Time-dependent (TD) regularization

An alternative choice for the regularization term is to use time-derivatives of the model parameters [15]. To this end, we assume that $p \equiv p(x, t)$, thus violating the physical setting of the problem. Then, a possible form for the TD regularization term $\mathcal{R}_p^{\text{TD}}$ becomes:

$$\mathcal{R}_p^{\text{TD}}(p) := \frac{R_p}{2} \int_0^T \int_0^l \left(\frac{\partial p(x, t)}{\partial t} \right)^2 \, \mathrm{d}x \, \mathrm{d}t. \tag{8}$$

Even though the model parameter p is assumed to depend on both time and space, the minimization process enforces it to be independent of time: clearly, of all the possible trajectories $p(x, t)$ for times $t \in (0, T)$, the time-independent $p(x)$ is the one minimizing (8). To this end, we further impose that:

$$p(x, 0) = p_0, \quad \frac{\partial p(x, T)}{\partial t} = 0, \tag{9}$$

where p_0 is the initial estimate of the model parameter p. In other words, we force at final time $t = T$, the material property distribution to be chosen as time-independent. In contrast to spatially driven regularizations, such as Tikhonov or TV that filter out higher spatial frequencies of the model parameters, the time-dependent form (8), in principle, allows them. In this sense, and with an eye towards three-dimensional problems, where the spatial variability of the material parameters is greater, and may be unduly penalized by a spatial regularization scheme, the TD scheme shifts the emphasis from space to time. As stated, our motivation for experimenting with (8) stems from our interest in departing from spatial schemes, while still imposing some form of regularization; in this sense, the logical candidate was a temporal scheme. We readily acknowledge that in accepting (initially) time-dependent moduli, we are possibly widening the solution feasibility space in unintended ways. However, the numerical results are, thus far, promising and provide evidence that the TD scheme regularizes the solution in both space and time. To date, we know of no formal proof of the effects or the benefits afforded by the TD scheme, other than those betrayed by the numerical examples discussed herein.

5 PDE-constrained optimization approach

5.1 Augmented functional

To reconstruct the material profile we seek to minimize (5) subject to the governing PDE and the boundary and initial conditions given by (1)–(4). To this end, we first recast the

problem as an unconstrained optimization problem by defining an augmented functional based on (5), where now the governing PDE and the boundary/initial conditions have been imposed (added) via Lagrange multipliers as side constraints (notice, only Neumann-type conditions need to be added as part of the side constraints; essential conditions are explicitly enforced). We then seek to satisfy the first-order optimality conditions. We discuss first the case of the TD regularization; the Tikhonov and TV schemes are simpler. The details follow: we define the augmented functional as:

$$
\begin{aligned}
&\mathcal{A}(u, \lambda, \alpha, \beta) \\
&= \frac{1}{2} \int_0^T [u(0, t) - u_m(0, t)]^2 \, \mathrm{d}t \\
&\quad + \frac{R_\alpha}{2} \int_0^T \int_0^l \left(\frac{\partial \alpha}{\partial t} \right)^2 \, \mathrm{d}x \, \mathrm{d}t + \frac{R_\beta}{2} \int_0^T \int_0^l \left(\frac{\partial \beta}{\partial t} \right)^2 \, \mathrm{d}x \, \mathrm{d}t \\
&\quad + \int_0^T \int_0^l \lambda \left\{ \frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x} \left(\alpha \frac{\partial u}{\partial x} \right) + \beta \frac{\partial u}{\partial t} \right\} \, \mathrm{d}x \, \mathrm{d}t \\
&\quad - \int_0^T \lambda(0, t) \left[\alpha(0, t) \frac{\partial u(0, t)}{\partial x} - f(t) \right] \, \mathrm{d}t \\
&\quad + \int_0^l \lambda(x, 0) \frac{\partial u(x, 0)}{\partial t} \, \mathrm{d}x,
\end{aligned}
\tag{10}
$$

where $\lambda(x, t)$ is the Lagrange multiplier (or adjoint variable), and, wherever appropriate, functional dependence has been dropped for brevity. Notice that the originally spatially dependent parameters $\alpha(x)$ and $\beta(x)$ have been modified to $\alpha(x, t)$ and $\beta(x, t)$ to account for the temporal dependence, per the TD regularization scheme. For the Tikhonov scheme, the augmented functional may be defined in a similar way:

$$
\begin{aligned}
&\mathcal{A}(u, \lambda, \alpha, \beta) \\
&= \frac{1}{2} \int_0^T [u(0, t) - u_m(0, t)]^2 \, \mathrm{d}t \\
&\quad + \frac{R_\alpha}{2} \int_0^l \left(\frac{d\alpha}{dx} \right)^2 \, \mathrm{d}x + \frac{R_\beta}{2} \int_0^l \left(\frac{d\beta}{dx} \right)^2 \, \mathrm{d}x \\
&\quad + \int_0^T \int_0^l \lambda \left\{ \frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x} \left(\alpha \frac{\partial u}{\partial x} \right) + \beta \frac{\partial u}{\partial t} \right\} \, \mathrm{d}x \, \mathrm{d}t \\
&\quad - \int_0^T \lambda(0, t) \left[\alpha(0) \frac{\partial u(0, t)}{\partial x} - f(t) \right] \, \mathrm{d}t \\
&\quad + \int_0^l \lambda(x, 0) \frac{\partial u(x, 0)}{\partial t} \, \mathrm{d}x.
\end{aligned}
\tag{11}
$$

In the TV case, the augmented functional can be obtained by replacing the regularization terms in (11) with the corresponding terms given by (7).

Next, the first-order optimality conditions are obtained from the variation of the augmented functional with respect to the state variable u, the adjoint variable λ, and the model

parameters α and β:

$$\begin{Bmatrix} \delta_\lambda \mathcal{A} \\ \delta_u \mathcal{A} \\ \delta_\alpha \mathcal{A} \\ \delta_\beta \mathcal{A} \end{Bmatrix} = \mathbf{0}. \tag{12}$$

5.2 The first-order optimality conditions

We seek quadruplets $(u, \lambda, \alpha, \beta)$ for which the augmented functional $\mathcal{A}$ becomes stationary. Accordingly:

5.2.1 First optimality condition

$$\delta_\lambda \mathcal{A} = \int_0^T \int_0^l \delta\lambda \left\{ \frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x}\left(\alpha\frac{\partial u}{\partial x}\right) + \beta\frac{\partial u}{\partial t} \right\} dx\, dt$$
$$- \int_0^T \delta\lambda(0, t)\left[\alpha(0, t)\frac{\partial u(0, t)}{\partial x} - f(t) \right] dt$$
$$+ \int_0^l \delta\lambda(x, 0)\frac{\partial u(x, 0)}{\partial t} dx = 0, \tag{13}$$

where $\delta\lambda$ denotes an arbitrary variation of λ. By taking into account the explicitly imposed homogeneous essential boundary $u(l, t) = 0$, and the initial condition $u(x, 0) = 0$, we recover the state problem:

State problem

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x}\left(\alpha\frac{\partial u}{\partial x}\right) + \beta\frac{\partial u}{\partial t} = 0, \ (x, t) \in (0, l) \times (0, T),$$
$$\tag{14}$$
$$\alpha(0, t)\frac{\partial u(0, t)}{\partial x} = f(t), \tag{15}$$
$$u(l, t) = 0, \tag{16}$$
$$u(x, 0) = \frac{\partial u(x, 0)}{\partial t} = 0. \tag{17}$$

Clearly the state problem is identical to the forward problem given by (1)–(4).[1]

5.2.2 Second optimality condition

The variation of the augmented functional with respect to the state variable u yields the second optimality condition.

[1] When the TD regularization scheme is used, then $\alpha(x) \equiv \alpha(x, t)$, and $\beta(x) \equiv \beta(x, t)$.

Accordingly:

$$\delta_u \mathcal{A} = \int_0^T [u(0, t) - u_m(0, t)]\, \delta u(0, t)\, dt$$
$$+ \int_0^T \int_0^l \lambda \left\{ \frac{\partial^2 \delta u}{\partial t^2} - \frac{\partial}{\partial x}\left(\alpha\frac{\partial \delta u}{\partial x}\right) + \beta\frac{\partial \delta u}{\partial t} \right\} dx\, dt$$
$$- \int_0^T \lambda(0, t)\alpha(0, t)\frac{\partial \delta u(0, t)}{\partial x} dt$$
$$+ \int_0^l \lambda(x, 0)\frac{\partial \delta u(x, 0)}{\partial t} dx. \tag{18}$$

By integrating by parts, and taking into account the boundary and initial conditions there results:

$$\delta_u \mathcal{A} = \int_0^T \int_0^l \delta u \left\{ \frac{\partial^2 \lambda}{\partial t^2} - \frac{\partial}{\partial x}\left(\alpha\frac{\partial \lambda}{\partial x}\right) - \frac{\partial}{\partial t}(\beta\lambda) \right\} dx\, dt$$
$$+ \int_0^T \delta u(0, t)\left\{ [u(0, t) - u_m(0, t)] - \alpha(0, t)\frac{\partial \lambda(0, t)}{\partial x} \right\} dt$$
$$- \int_0^T \lambda(l, t)\alpha(l, t)\frac{\partial \delta u(l, t)}{\partial x} dt$$
$$+ \int_0^l \lambda(x, T)\left\{ \frac{\partial \delta u(x, T)}{\partial t} + \beta(x, T)\delta u(x, T) \right\} dx$$
$$- \int_0^l \frac{\partial \lambda(x, T)}{\partial t}\delta u(x, T)\, dx. \tag{19}$$

Since δu is arbitrary, by setting $\delta_u \mathcal{A} = 0$ the following adjoint problem ensues:
Adjoint problem

$$\frac{\partial^2 \lambda}{\partial t^2} - \frac{\partial}{\partial x}\left(\alpha\frac{\partial \lambda}{\partial x}\right) - \frac{\partial}{\partial t}(\beta\lambda) = 0,$$
$$(x, t) \in (0, l) \times (0, T), \tag{20}$$
$$\alpha(0, t)\frac{\partial \lambda(0, t)}{\partial x} = [u(0, t) - u_m(0, t)], \tag{21}$$
$$\lambda(l, t) = 0, \tag{22}$$
$$\lambda(x, T) = \frac{\partial \lambda(x, T)}{\partial t} = 0. \tag{23}$$

We remark that the adjoint problem is similar to the state problem, with, however, two important differences: first, the right-hand-side of (21), i.e., the source term, depends on the misfit between the computed and measured values of the state variable. Secondly, by virtue of (23), the adjoint problem is a final-value problem; in addition, the damping term's sign has changed. If the Tikhonov or TV scheme were used, the same state and adjoint problem would have been obtained, with, naturally, α and β being only spatially dependent.

5.2.3 Third optimality condition

We obtain the third condition as the variation of the augmented functional with respect to α; there results

$$
\begin{aligned}
\delta_\alpha \mathcal{A} = {} & R_\alpha \int_0^T \int_0^l \frac{\partial \alpha}{\partial t} \frac{\partial \delta \alpha}{\partial t} \, dx \, dt \\
& - \int_0^T \int_0^l \lambda \frac{\partial}{\partial x} \left(\delta \alpha \frac{\partial u}{\partial x} \right) \, dx \, dt \\
& - \int_0^T \lambda(0, t) \frac{\partial u(0, t)}{\partial x} \delta \alpha(0, t) \, dt \\
= {} & R_\alpha \int_0^T \int_0^l \left[\frac{\partial}{\partial t} \left(\frac{\partial \alpha}{\partial t} \delta \alpha \right) - \frac{\partial^2 \alpha}{\partial t^2} \delta \alpha \right] dx \, dt \\
& - \int_0^T \int_0^l \left[\frac{\partial}{\partial x} \left(\lambda \delta \alpha \frac{\partial u}{\partial x} \right) - \delta \alpha \frac{\partial \lambda}{\partial x} \frac{\partial u}{\partial x} \right] dx \, dt \\
& - \int_0^T \lambda(0, t) \frac{\partial u(0, t)}{\partial x} \delta \alpha(0, t) \, dt \\
= {} & R_\alpha \int_0^l \left[\frac{\partial \alpha(x, T)}{\partial t} \delta \alpha(x, T) - \frac{\partial \alpha(x, 0)}{\partial t} \delta \alpha(x, 0) \right] dx \\
& - R_\alpha \int_0^T \int_0^l \frac{\partial^2 \alpha}{\partial t^2} \delta \alpha \, dx \, dt \\
& - \int_0^T \left[\lambda(l, t) \delta \alpha(l, t) \frac{\partial u(l, t)}{\partial x} \right. \\
& \qquad\quad \left. - \lambda(0, t) \delta \alpha(0, t) \frac{\partial u(0, t)}{\partial x} \right] dt \\
& + \int_0^T \int_0^l \delta \alpha \frac{\partial \lambda}{\partial x} \frac{\partial u}{\partial x} \, dx \, dt \\
& - \int_0^T \lambda(0, t) \frac{\partial u(0, t)}{\partial x} \delta \alpha(0, t) \, dt = 0. \qquad (24)
\end{aligned}
$$

By taking into account that $\delta \alpha(x, 0) = 0$, $\frac{\partial \alpha(x, T)}{\partial t} = 0$ and $\lambda(l, t) = 0$ (per (9)), (24) reduces to:

$$
\delta_\alpha \mathcal{A} = \int_0^T \int_0^l \left\{ -R_\alpha \frac{\partial^2 \alpha}{\partial t^2} + \frac{\partial \lambda}{\partial x} \frac{\partial u}{\partial x} \right\} \delta \alpha \, dx \, dt = 0. \quad (25)
$$

Similarly, the variation of the augmented functional with respect to α, in the Tikhonov case, results in:

$$
\delta_\alpha \mathcal{A} = \int_0^l \left[-R_\alpha \frac{d^2 \alpha}{dx^2} + \int_0^T \frac{\partial \lambda}{\partial x} \frac{\partial u}{\partial x} \, dt \right] \delta \alpha \, dx = 0, \quad (26)
$$

whereas, in the TV case, results in:

$$
\delta_\alpha \mathcal{A} = \int_0^l \left[-R_\alpha \epsilon \left(\frac{d\alpha}{dx} \right)^2 \left\{ \left(\frac{d\alpha}{dx} \right)^2 + \epsilon \right\}^{-\frac{3}{2}} \right. \\
\left. + \int_0^T \frac{\partial \lambda}{\partial x} \frac{\partial u}{\partial x} \, dt \right] \delta \alpha \, dx = 0. \quad (27)
$$

5.2.4 Fourth optimality condition

The last condition can be obtained similarly to the third one, by taking the variation of the augmented functional with respect to β; there results:

$$
\begin{aligned}
\delta_\beta \mathcal{A} = {} & R_\beta \int_0^T \int_0^l \frac{\partial \beta}{\partial t} \frac{\partial \delta \beta}{\partial t} \, dx \, dt + \int_0^T \int_0^l \lambda \delta \beta \frac{\partial u}{\partial t} \, dx \, dt \\
= {} & R_\beta \int_0^T \int_0^l \left[\frac{\partial}{\partial t} \left(\frac{\partial \beta}{\partial t} \delta \beta \right) - \frac{\partial^2 \beta}{\partial t^2} \delta \beta \right] dx \, dt \\
& + \int_0^T \int_0^l \lambda \frac{\partial u}{\partial t} \delta \beta \, dx \, dt \\
= {} & R_\beta \int_0^l \left[\frac{\partial \beta(x, T)}{\partial t} \delta \beta(x, T) - \frac{\partial \beta(x, 0)}{\partial t} \delta \beta(x, 0) \right] dx \\
& - \int_0^T \int_0^l R_\beta \frac{\partial^2 \beta}{\partial t^2} \delta \beta \, dx \, dt \\
& + \int_0^T \int_0^l \lambda \frac{\partial u}{\partial t} \delta \beta \, dx \, dt = 0. \qquad (28)
\end{aligned}
$$

By taking into account that $\delta \beta(x, 0) = 0$, $\frac{\partial \beta(x, T)}{\partial t} = 0$, (28) yields:

$$
\delta_\beta \mathcal{A} = \int_0^T \int_0^l \left\{ -R_\beta \frac{\partial^2 \beta}{\partial t^2} + \lambda \frac{\partial u}{\partial t} \right\} \delta \beta \, dx \, dt = 0. \quad (29)
$$

If the Tikhonov regularization were used, instead of (29), the last condition would have read:

$$
\delta_\beta \mathcal{A} = \int_0^l \left\{ -R_\beta \frac{d^2 \beta}{dx^2} + \int_0^T \lambda \frac{\partial u}{\partial t} \, dt \right\} \delta \beta \, dx = 0. \quad (30)
$$

Lastly, in the TV case, the last condition becomes:

$$
\delta_\beta \mathcal{A} = \int_0^l \left[-R_\beta \epsilon \left(\frac{d\beta}{dx} \right)^2 \left\{ \left(\frac{d\beta}{dx} \right)^2 + \epsilon \right\}^{-\frac{3}{2}} \right. \\
\left. + \int_0^T \lambda \frac{\partial u}{\partial t} \, dt \right] \delta \beta \, dx = 0. \quad (31)
$$

Conditions (25) and (29) constitute the *control problem* for the TD regularization case, whereas (26) and (30) correspond to the Tikhonov case, and (27) and (31) to the TV case. The control equations are used in order to update the model parameters. Obviously, they are satisfied only for the true profiles.

6 Inversion process

In order to satisfy the first-order optimality conditions, we adopt a reduced-space method. Notice that, in principle, the state problem (14)–(17), the adjoint problem (20)–(23), and the control conditions (25) and (29) (or (26) and (30), or

(27) and (31)) can be solved as a coupled problem (by using a full-space method). However, the computational cost per iteration increases, given, upon discretization, the resulting matrix sizes. We remark that for the solution of the state and the adjoint problem (either as a coupled system or individually) any numerical scheme may be used (finite differences, finite elements, etc.). In contrast to a full-space method, here we opt for a reduced-space method that maps the optimization problem to the space of the model parameters (α and β), whereby eliminating the state and adjoint variables. We start by solving the state problem (14)–(17) to obtain the state variable u, given an estimate of the model parameters, thereby satisfying the first condition $\delta_\lambda \mathcal{A} = 0$. Then, we solve the adjoint problem using the state variable computed in the first step, to obtain the Lagrange multiplier λ that satisfies the second condition $\delta_u \mathcal{A} = 0$. To solve both the state and adjoint problems, we employ conventional finite elements. Then, there remains to seek to update the model parameters, α and β, so that the third and fourth conditions be satisfied. We use the control equations to iteratively provide updates to the model parameters.

6.1 State and adjoint semi-discrete forms

In order to satisfy the first condition for an assumed set of inversion variables, we use a standard Galerkin approach to solve the state problem (14)–(17). Accordingly, the weak form can be obtained by multiplying the state Eq.(14) by an appropriate test function $\upsilon(x)$ (with $\upsilon(l) = 0$) and integrating over the entire domain. Using integration by parts, there results:

$$
\int_0^l \left[\frac{\partial^2 u(x,t)}{\partial t^2} \upsilon(x) + \alpha(x,t) \frac{\partial u(x,t)}{\partial x} \frac{\partial \upsilon(x)}{\partial x} \right.
$$
$$
\left. + \beta(x,t) \frac{\partial u(x,t)}{\partial t} \upsilon(x) \right] dx = -\upsilon(0) f(t), \tag{32}
$$

where the boundary conditions have been taken into account. With a similar process, where $q(x)$ is now used as a test function, we obtain the weak form of the adjoint problem:

$$
\int_0^l \left[\frac{\partial^2 \lambda(x,t)}{\partial t^2} q(x) + \alpha(x,t) \frac{\partial \lambda(x,t)}{\partial x} \frac{\partial q(x)}{\partial x} \right.
$$
$$
\left. - q(x) \frac{\partial}{\partial t} (\beta(x,t)\lambda(x,t)) \right] dx
$$
$$
= q(0) [u_m(0,t) - u(0,t)]. \tag{33}
$$

We introduce next standard polynomial approximations for the trial functions of the state $u(x,t)$ and the adjoint $\lambda(x,t)$,

and their respective test functions $\upsilon(x)$, and $q(x)$; let:

$$
u(x,t) = \sum_{i=1}^N u_i(t)\phi_i(x), \quad \upsilon(x) = \sum_{i=1}^N \upsilon_i \phi_i(x), \tag{34}
$$

$$
\lambda(x,t) = \sum_{i=1}^N \lambda_i(t)\phi_i(x), \quad q(x) = \sum_{i=1}^N q_i \phi_i(x), \tag{35}
$$

where N is the number of nodal points, ϕ are basis functions, and $u_i, \lambda_i, \upsilon_i, q_i$ denote nodal quantities. Then the semi-discrete forms of the state and adjoint problems can be cast as:

$$
\mathbf{M} \frac{\partial^2 \mathbf{u}(t)}{\partial t^2} + \mathbf{K}(t) \mathbf{u}(t) + \mathbf{C}(t) \frac{\partial \mathbf{u}(t)}{\partial t} = \mathbf{F}(t), \tag{36}
$$

$$
\mathbf{M} \frac{\partial^2 \boldsymbol{\lambda}(t)}{\partial t^2} + [\mathbf{K}(t) + \mathbf{Q}(t)] \boldsymbol{\lambda}(t) - \mathbf{C}(t) \frac{\partial \boldsymbol{\lambda}(t)}{\partial t} = \mathbf{G}(t), \tag{37}
$$

where:

$$
M_{ij} = \int_0^l \phi_i \phi_j \, dx, \quad K_{ij} = \int_0^l \alpha(x,t) \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial x} \, dx, \tag{38}
$$

$$
C_{ij} = \int_0^l \beta(x,t)\phi_i\phi_j \, dx, \quad Q_{ij} = -\int_0^l \frac{\partial \beta(x,t)}{\partial t} \phi_i\phi_j \, dx, \tag{39}
$$

$$
F_i = -f(t)\delta_{i1}, \quad G_i = [u_m(0,t) - u(0,t)]\delta_{i1}. \tag{40}
$$

In the above, δ_{i1} denotes the Kronecker delta, $\boldsymbol{u}$ and $\boldsymbol{\lambda}$ are the vectors of the nodal state and adjoint variables, respectively, and customary notation has been used for the matrices. Notice that, whereas the mass matrix $\boldsymbol{M}$ is independent of time, the stiffness $\boldsymbol{K}$, and damping matrices $\boldsymbol{C}$ and $\boldsymbol{Q}$ depend on time, due to the presence of the (assumed) time-dependent moduli.

6.2 Temporal discretization

To arrive at a solution first for the state variable and then for the adjoint variable, the semi-discrete forms (36) and (37) need next be discretized in time. We note that, whereas (36) is an initial value problem for which $\boldsymbol{u}(0) = \frac{\partial \boldsymbol{u}}{\partial t}(0) = \boldsymbol{0}$, (37) is a final-value problem for which $\boldsymbol{\lambda}(T) = \frac{\partial \boldsymbol{\lambda}}{\partial t}(T) = \boldsymbol{0}$. In the case of the TD regularization scheme, the time-dependent matrices $\boldsymbol{K}(t)$, $\boldsymbol{C}(t)$, and $\boldsymbol{Q}(t)$ need to be appropriately treated. Their temporal dependence stems from the moduli, which in turn, need also be discretized in both space and time. Accordingly, let:

$$
\alpha(x,t) = \sum_{j=1}^N a_j(t)\varphi_j(x), \quad \beta(x,t) = \sum_{j=1}^N b_j(t)\varphi_j(x), \tag{41}
$$

in which φ_j are basis functions, and a_j and b_j denote nodal values of α and β, respectively. At the first iteration we start with the initially guessed model parameters, which we enforce to be constant in time. Upon updating the model

parameters (see Sect. 6.3), there result time-dependent moduli $\alpha(x, t)$ and $\beta(x, t)$. Using those, one could formally proceed by updating $K(t)$, $C(t)$, and $Q(t)$, per (38)–(39); we refer to this approach as Scheme TD-I. Clearly, this approach is computationally costly, as it entails the evaluation of various matrices on a per time-step basis. Alternatively, one could approximate the temporal dependence of the moduli by constant values (per inversion iteration); candidate choices include:

$$a_j(t) \simeq \langle a_j(t) \rangle, \quad \text{or} \quad a_j(t) \simeq a_j(T), \quad \forall j = 1, \ldots, N,$$
(42)

where the former expression refers to the mean value of $a_j(t)$ over the period $(0, T)$, and the latter expression refers to its final value (similarly, for the coefficients of β, b_j). We opted for the second of (42) (piecewise constant in space, as well). Thus, effectively, over an element e, (41) can be rewritten as:

$$\alpha(x, t)|_e \simeq a_e(T), \quad \beta(x, t)|_e \simeq b_e(T).$$
(43)

Consequently, the element matrices k_e, c_e, and q_e corresponding to K, C, and Q in (38)–(39), respectively, are modified to now read:

$$k_e = a_e(T) \int_e \frac{\partial \boldsymbol{\phi}}{\partial x} \frac{\partial \boldsymbol{\phi}^T}{\partial x} \, dx, \quad c_e = b_e(T) \int_e \boldsymbol{\phi} \boldsymbol{\phi}^T dx, \quad (44)$$

and $q_e = 0$. We refer to this second approach as Scheme TD-II. Next, standard time integration schemes can be used: here, we opted for Newmark's average-acceleration scheme. Through the numerical experiments, we tested both the time-dependent matrices (Scheme TD-I), and the time-independent matrices (Scheme TD-II) cases in solving the state and adjoint problems; we observed only minor differences in the final estimated model parameters.

6.3 Model parameter updates

By solving the state and adjoint problems, the state variable u and the adjoint variable λ satisfying the first and second conditions, respectively, are obtained. Then, the problem is reduced to a minimization problem with respect to the model parameters α and β. Here, notice that the variations with respect to the model parameters of the augmented functional, $\delta_\alpha \mathcal{A}$ and $\delta_\beta \mathcal{A}$, are tantamount to the gradient components of the misfit functional $\nabla_\alpha \mathcal{J}$ and $\nabla_\beta \mathcal{J}$, since the side constraints in the augmented functional (10) have already vanished owing to the solution of the state problem. Then, what remains to be done is to provide the mechanism for updating the model parameters: this can be directly accomplished via the control equations derived for the TD, Tikhonov, or TV regularizations. We outline the details below.

6.3.1 Time-dependent (TD) regularization

The control Eq. (25) yields:

$$\frac{\partial^2 \alpha(x, t)}{\partial t^2} = \frac{1}{R_\alpha} \frac{\partial \lambda(x, t)}{\partial x} \frac{\partial u(x, t)}{\partial x}.$$
(45)

The right-hand-side of (45) can be readily computed, once u and λ have been obtained. Then, the update for $\alpha(x, t)$, on a per element basis, can be computed by integrating (45) as shown below (superscripts to α and a indicate new and previous values between inversion iterations):

$$\alpha_e^{(k+1)}(x, t) = a_e^{(k+1)}(t) = a_e^{(k)}(T)$$
$$- \frac{t}{R_\alpha} \int_0^T \frac{\partial \boldsymbol{\phi}^T}{\partial x} \boldsymbol{\lambda}(\tau) \quad \frac{\partial \boldsymbol{\phi}^T}{\partial x} \boldsymbol{u}(\tau) \, d\tau$$
$$+ \frac{1}{R_\alpha} \int_0^t \int_0^s \frac{\partial \boldsymbol{\phi}^T}{\partial x} \boldsymbol{\lambda}(\tau) \quad \frac{\partial \boldsymbol{\phi}^T}{\partial x} \boldsymbol{u}(\tau) \, d\tau \, ds,$$
(46)

where $\boldsymbol{\phi}$, $\boldsymbol{\lambda}$, and $\boldsymbol{u}$ are restricted to element e, and we have enforced conditions (9) so that:

$$a_e^{(k+1)}(0) = a_e^{(k)}(T), \quad \frac{\partial a_e^{(k+1)}}{\partial t}(T) = 0.$$
(47)

Similarly, for the β updates: we use control Eq. (29) to obtain:

$$\frac{\partial^2 \beta(x, t)}{\partial t^2} = \frac{1}{R_\beta} \lambda(x, t) \frac{\partial u(x, t)}{\partial t},$$
(48)

and therefore, by analogy to (46):

$$\beta_e^{(k+1)}(x, t) = b_e^{(k+1)}(t) = b_e^{(k)}(T)$$
$$- \frac{t}{R_\beta} \int_0^T \boldsymbol{\phi}^T(x) \boldsymbol{\lambda}(\tau) \quad \boldsymbol{\phi}^T(x) \frac{\partial \boldsymbol{u}(\tau)}{\partial \tau} \, d\tau$$
$$+ \frac{1}{R_\beta} \int_0^t \int_0^s \boldsymbol{\phi}^T(x) \boldsymbol{\lambda}(\tau) \boldsymbol{\phi}^T(x) \frac{\partial \boldsymbol{u}(\tau)}{\partial \tau} \, d\tau \, ds.$$
(49)

Notice that, as evidenced by the above relations, the use of large values for the regularization parameters does not distort the misfit information, since the regularization terms vanish as the time-derivatives of the inversion variables become or approach zero. However, as it can be seen from (45) and (48), the accelerations of the time-dependent coefficients of the model parameters are inversely proportional to the regularization parameters R_α and R_β. Therefore, the use of larger regularization parameters will force the convergence rate to be slower. It is, thus, beneficial to use smaller regularization parameter values (as long as the resulting optimization problem converges). The entire inversion process with the TD regularization scheme is summarized in the following Algorithm 1, where trivial steps have been omitted. To accelerate convergence it is possible to use a line search scheme that optimizes the step length: accordingly, in (46) and (49),

Algorithm 1 Inversion algorithm using TD regularization (Scheme TD-I or TD-II)

1: Choose R_α, R_β
2: Set k=0
3: Set initial guess of inversion variables, $a_e^{(k)}$ and $b_e^{(k)}$
4: Set convergence tolerance TOL
5: Set misfit = TOL + 1
6: **while** (misfit > TOL) **do**
7:   Solve the state problem to obtain u
8:   Solve the adjoint problem to obtain λ
9:   Compute $\frac{\partial^2 \boldsymbol{a}}{\partial t^2}$ using:
$$\frac{\partial^2 a_e^{(k)}}{\partial t^2} = \frac{1}{R_\alpha} \frac{\partial \lambda}{\partial x} \frac{\partial u}{\partial x}\bigg|_e$$
10:   Compute $\frac{\partial^2 \boldsymbol{b}}{\partial t^2}$ using:
$$\frac{\partial^2 b_e^{(k)}}{\partial t^2} = \frac{1}{R_\beta} \lambda \frac{\partial u}{\partial t}\bigg|_e$$
11:   Update the model parameters $\boldsymbol{\alpha}$ using:
$$\boldsymbol{a}^{(k+1)}(t) = \boldsymbol{a}^{(k)}(T) - t \int_0^T \frac{\partial^2 \boldsymbol{a}(\tau)}{\partial \tau} \, \mathrm{d}\tau +$$
$$\int_0^t \int_0^s \frac{\partial^2 \boldsymbol{a}(\tau)}{\partial \tau} \, \mathrm{d}\tau \, \mathrm{d}s$$
12:   Update the model parameters $\boldsymbol{\beta}$ using:
$$\boldsymbol{b}^{(k+1)}(t) = \boldsymbol{b}^{(k)}(T) - t \int_0^T \frac{\partial^2 \boldsymbol{b}(\tau)}{\partial \tau} \, \mathrm{d}\tau +$$
$$\int_0^t \int_0^s \frac{\partial^2 \boldsymbol{b}(\tau)}{\partial \tau} \, \mathrm{d}\tau \, \mathrm{d}s$$
13:   k=k+1
14: **end while**
15: Save final estimates $\boldsymbol{\alpha}(x, t)$ and $\boldsymbol{\beta}(x, t)$

Algorithm 2 Inversion algorithm using TD regularization (accelerated version)

1: Choose θ, ρ, μ, R_α, R_β
2: Set k=0
3: Set initial guess of inversion variables, $a_e^{(k)}$ and $b_e^{(k)}$
4: Set $\mathbf{p}(t) = \{\mathbf{a}(t) \ \mathbf{b}(t)\}^T$
5: Set convergence tolerance TOL
6: Set misfit = TOL + 1
7: **while** (misfit > TOL) **do**
8:   Solve the state problem to obtain u
9:   Solve the adjoint problem to obtain λ
10:   Compute $\frac{\partial^2 \boldsymbol{a}}{\partial t^2}$ using:
$$\frac{\partial^2 a_e^{(k)}}{\partial t^2} = \frac{1}{R_\alpha} \frac{\partial \lambda}{\partial x} \frac{\partial u}{\partial x}\bigg|_e$$
11:   Compute $\ddot{\boldsymbol{b}}$ using:
$$\frac{\partial^2 b_e^{(k)}}{\partial t^2} = \frac{1}{R_\beta} \lambda \frac{\partial u}{\partial t}\bigg|_e$$
12:   Compute the search direction $\boldsymbol{d}_k(t) = \begin{Bmatrix} \boldsymbol{d}_\alpha(t) \\ \boldsymbol{d}_\beta(t) \end{Bmatrix}$
   where
$$\boldsymbol{d}_\alpha(t) = -t \int_0^T \frac{\partial^2 \boldsymbol{a}(\tau)}{\partial \tau^2} \, \mathrm{d}\tau + \int_0^t \int_0^s \frac{\partial^2 \boldsymbol{a}(\tau)}{\partial \tau^2} \, \mathrm{d}\tau \, \mathrm{d}s,$$
$$\boldsymbol{d}_\beta(t) = -t \int_0^T \frac{\partial^2 \boldsymbol{b}(\tau)}{\partial \tau^2} \, \mathrm{d}\tau + \int_0^t \int_0^s \frac{\partial^2 \boldsymbol{b}(\tau)}{\partial \tau^2} \, \mathrm{d}\tau \, \mathrm{d}s.$$
13:   **while** ($\mathcal{J}(\mathbf{p}_k + \theta_k \mathbf{d}_k) > \mathcal{J}(\mathbf{p}_k) + \mu \theta_k \mathbf{p}_k \cdot \nabla \mathcal{J}(\mathbf{p}_k)$) **do**
14:     $\theta \leftarrow \rho\theta$
15:   **end while**
16:   Update the estimates $\mathbf{p}_{k+1}(t) = \mathbf{p}_k(T) + \theta_k \mathbf{d}_k(t)$
17:   k=k+1
18: **end while**
19: Save final estimates $\mathbf{p}(T)$

---

the integrals play the role of the search direction, while the reciprocal of the regularization parameter plays the role of the step length. Therefore, by combining any gradient-based scheme with a line search, we could accelerate the convergence rate. We employ the steepest descent method with an inexact line search scheme; the approach is summarized in Algorithm 2 below.

6.3.2 Tikhonov or TV regularization

Using the Tikhonov regularization, the inversion process is similar to Algorithm 2; the primary difference is in the computation of the search direction. Again, by solving the state and adjoint problems, we obtain the state variable u and the Lagrange multiplier λ, which satisfy the first and second optimality conditions, under given estimates of the model parameters. Then, the remaining third and fourth conditions given by (26) and (30), or by (27) and (31), respectively, provide, essentially, the first-order variations (or components of the gradient) of the misfit functional with respect to the model parameters. Using these first-order variations, we can adapt appropriately any gradient-based optimization scheme. Here we opt again for the steepest descent method combined with an inexact line search. In other words, we take the search direction as the negative of the first-order variations. The

entire inversion process using either the Tikhonov or the TV regularization is summarized in Algorithm 3.

7 Numerical experiments

We study the efficiency and robustness of the algorithms described in the previous sections via a series of numerical experiments. For simplicity, we consider one-dimensional problems in which several layers lie over the rigid bottom. Through the numerical experiments, we test the performance of the studied algorithms for both smooth and sharp target profiles. In addition, we study the effect of the regularization parameter, and the algorithmic performance against noisy data and different initial estimates. In all cases, we use synthetic data produced in a manner that avoids committing classical inverse crimes. That is, we obtain the measured response numerically by using the (exact) target profile, the prescribed source signal, and a very fine mesh that resolves more than adequately the physics of the problem (e.g., the mesh density is such that there are approximately 15–20 points corresponding to the smallest wavelength present in the source signal). The mesh density we use in the inversion process is always different than the density used to obtain the synthetic data.

Algorithm 3 Inversion using Tikhonov or TV regularization

1: Choose θ, ρ, μ, R_α, R_β
2: Set k=0
3: Set initial guess of inversion variables, $a_e^{(k)}$ and $b_e^{(k)}$
4: Set $\mathbf{p} = \{\mathbf{a} \ \mathbf{b}\}^T$
5: **while** (misfit > TOL) **do**
6:    Solve the state problem and obtain u
7:    Solve the adjoint problem and obtain λ
8:    Compute the search direction $\mathbf{d}_k = \begin{Bmatrix} d_\alpha \\ d_\beta \end{Bmatrix}$
      where, for Tikhonov regularization,

$$d_{\alpha,e} = R_\alpha \left.\frac{d^2\alpha}{dx^2}\right|_e - \int_0^T \left.\frac{\partial\lambda}{\partial x}\frac{\partial u}{\partial x}\right|_e dt,$$

$$d_{\beta,e} = R_\beta \left.\frac{d^2\beta}{dx^2}\right|_e - \int_0^T \left.\lambda\frac{\partial u}{\partial t}\right|_e dt,$$

      and for TV regularization,

$$d_{\alpha,e}=R_\alpha\epsilon\left(\frac{d\alpha}{dx}\right)^2\left[\left(\frac{d\alpha}{dx}\right)^2+\epsilon\right]^{-\frac{3}{2}}\Bigg|_e-\int_0^T\left.\frac{\partial\lambda}{\partial x}\frac{\partial u}{\partial x}\right|_e dt,$$

$$d_{\beta,e}=R_\beta\epsilon\left(\frac{d\beta}{dx}\right)^2\left[\left(\frac{d\beta}{dx}\right)^2+\epsilon\right]^{-\frac{3}{2}}\Bigg|_e-\int_0^T\left.\lambda\frac{\partial u}{\partial t}\right|_e dt.$$

9:    **while** ($\mathcal{J}(\mathbf{p}_k + \theta_k\mathbf{d}_k) > \mathcal{J}(\mathbf{p}_k) + \mu\theta_k\mathbf{p}_k \cdot \nabla\mathcal{J}(\mathbf{p}_k)$) **do**
10:      $\theta \leftarrow \rho\theta$
11:   **end while**
12:   Update the estimates $\mathbf{p}_{k+1} = \mathbf{p}_k + \theta_k\mathbf{d}_k$
13:   k=k+1
14: **end while**
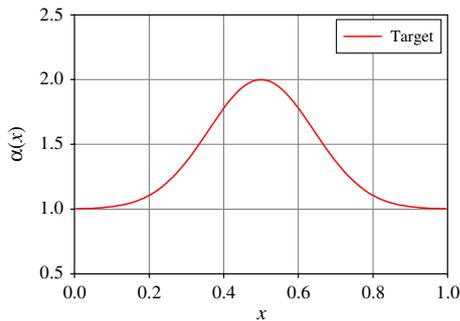
Fig. 2 A smooth $\alpha(x)$ target profile

7.1 On smooth profiles

As a first example problem, we consider a simple case, in which only the modulus (or wave velocity) is unknown. We consider $l = 1$, and thus $x \in [0, 1]$. The target modulus profile is a Gaussian bell-like distribution (Fig. 2):

$$\alpha(x) = \left[1 + \exp\left(-\frac{(x - 0.5)^2}{0.04}\right)\right]. \tag{50}$$

The source excitation is a rapidly decaying pulse-like signal given below by (51). Both the signal and its Fourier transform are depicted in Fig. 3a, b. For the given target profile and source excitation, the measured (synthetic) data $u_m(0, t)$ are shown in Fig. 3c.

$$f(t) = \exp\left[-\frac{(t - 0.1)^2}{0.0025}\right]. \tag{51}$$

Based on the given measured data, we inverted for the modulus profile via the TD regularization schemes (Algorithms 1,2), and via the Tikhonov and TV regularization schemes (Algorithm 3). In the time-dependent case we tested both Schemes TD-I and TD-II.

7.1.1 Regularization parameter effects

We begin by choosing an initial distribution for the modulus that is constant throughout the entire domain and observation period, that is, we set the initial guess to $\alpha(x, t) = 1.0$. We study the effect the magnitude of the regularization parameters has on the performance of the regularization schemes we considered. For the TD regularization case, we treat the same problem by two ways: first, we use the time-dependent matrices in solving the state and adjoint problems (Scheme TD-I), and second, we solve them with the modified matrices given by (44) (Scheme TD-II). Figure 4 a, b, d, e summarizes the α profiles obtained using the TD regularization schemes for different regularization parameters $R_\alpha = 0.1$ and 0.01 (the results are reported for the same level of misfit error tolerance, set at 2.0×10^{-7}). In addition, the differences of

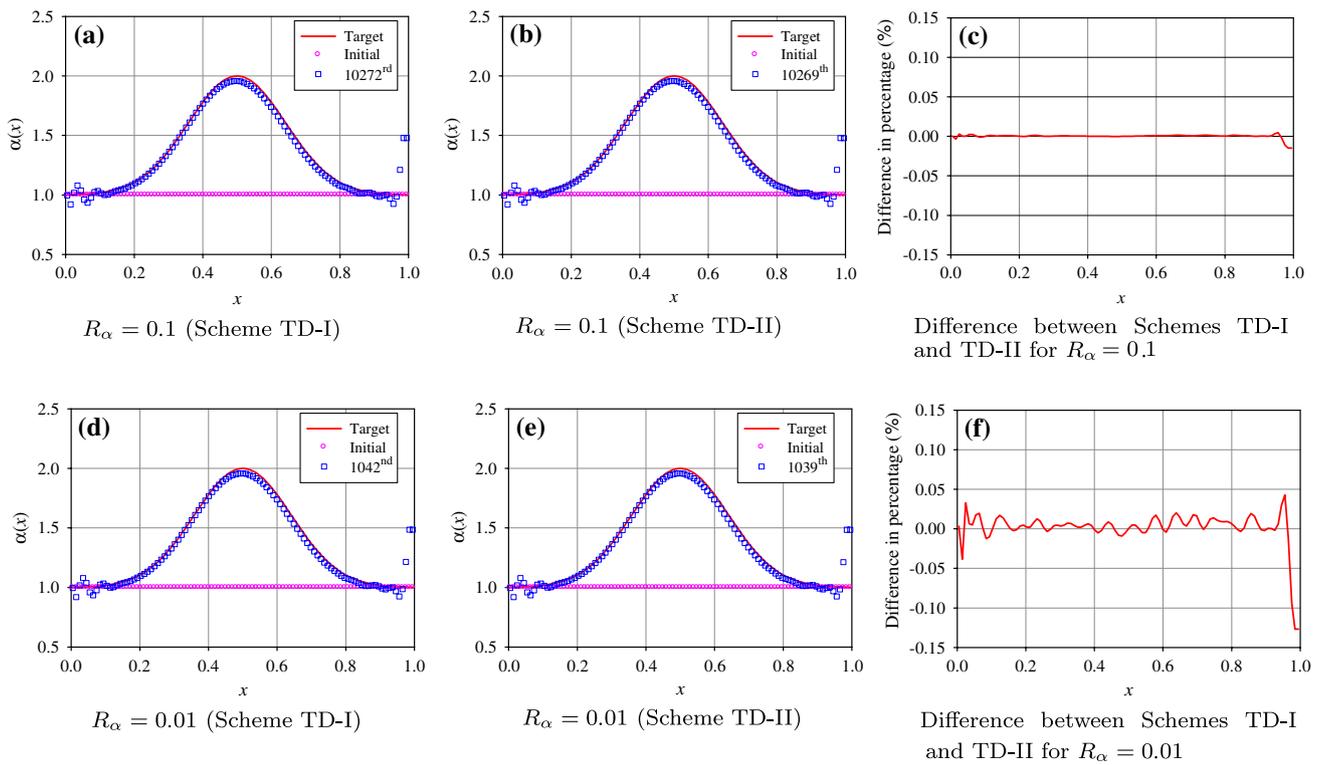**Fig. 3** Source signal, its frequency content, and surface data

Fig. 4 Target, initial, and estimated profile of $\alpha(x)$ using the TD regularization schemes

the estimated $\alpha(x)$ obtained using Scheme TD-I and Scheme TD-II are shown in Fig. 4c, f. Here, the differences $D(x)$ are expressed in percentage as:

$$D(x) = \frac{\alpha_{TD-II}(x) - \alpha_{TD-I}(x)}{\alpha_{TD-I}(x)} \times 100 \ (\%), \qquad (52)$$

where $\alpha_{TD-I}(x)$ and $\alpha_{TD-II}(x)$ are the final estimated $\alpha(x)$ using Scheme TD-I and Scheme TD-II, respectively. As it can be seen from Fig. 4, only minor differences are observed between the two schemes. For both cases, the final estimates are almost identical (the differences are less than 0.15%), and there is no significant difference in the required number of iterations. In addition, for both cases, the solution converges to the target profile, regardless of the value of the regularization parameter, albeit at higher computational cost associated with higher parameter choices (notice that, here, we employed Algorithm 1 in which the line search scheme is not implemented, and, as a result, the regularization parameter remains constant throughout the process). Naturally, smaller parameter values fail to achieve the penalty intent of the regularization term: for example in this first problem, whereas for $R_\alpha = 0.01$ satisfactory results were obtained, for $R_\alpha = 0.001$ we observed divergence and increased misfit errors. Figure 5a summarizes the convergence patterns of the misfit error for the considered values of the regularization parameters in the TD-II case, whereas Fig. 5b, c depicts the

corresponding convergence patterns for the modulus. We remark that the convergence patterns using Scheme TD-I are nearly identical to those obtained using TD-II.

Next, we consider a Tikhonov regularization scheme using different values for the regularization parameters; the resulting estimated profiles are shown in Fig. 6. We continue the iteration process until the misfit error converges. As it can be seen in Fig. 6a, when $R_\alpha = 0.001$, the solution does not converge to the target profile. However, as smaller values of the regularization parameter are adopted, the solution tends closer to the target profile (Fig. 6b, c). We failed to obtain a solution as close to the target profile (even when $R_\alpha = 0$) as that obtained using the TD regularization. We remark that one could improve on the Tikhonov case results by, for example, using a source with a higher-frequency content, or a continuation scheme on the regularization parameter, or a number of other approaches (e.g., L-curve) that aim at the intelligent choice of the regularization parameter. We have not implemented such schemes for either the Tikhonov, TV, or TD cases (in principle, all schemes stand to benefit). Nevertheless, the results shown thus far suggest that the TD regularization performs better than Tikhonov, for blind, yet broad, choices of the regularization parameter. Figure 7a shows the misfit error, whereas Fig. 7b, d depicts the convergence patterns for the modulus, in the Tikhonov regularization case. The results obtained when the TV scheme
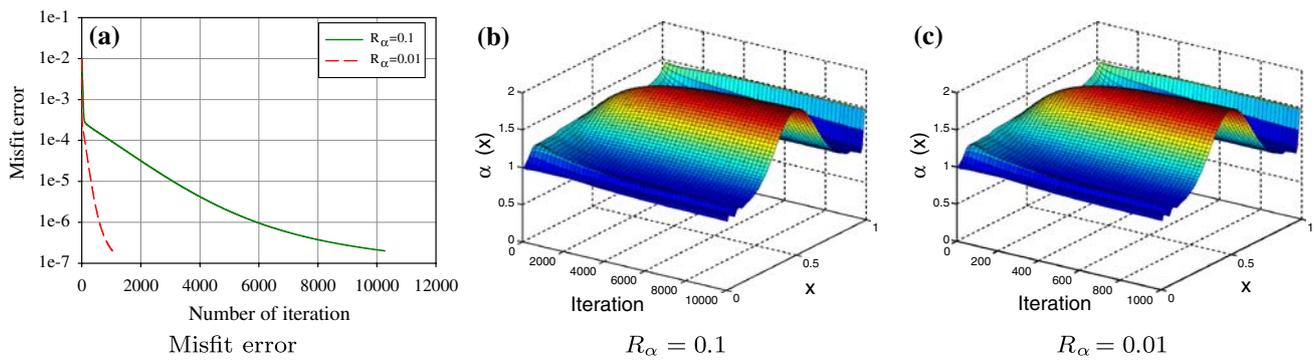
Fig. 5 Misfit error and convergence patterns of $\alpha(x)$ using the TD regularization scheme (Scheme TD-II)
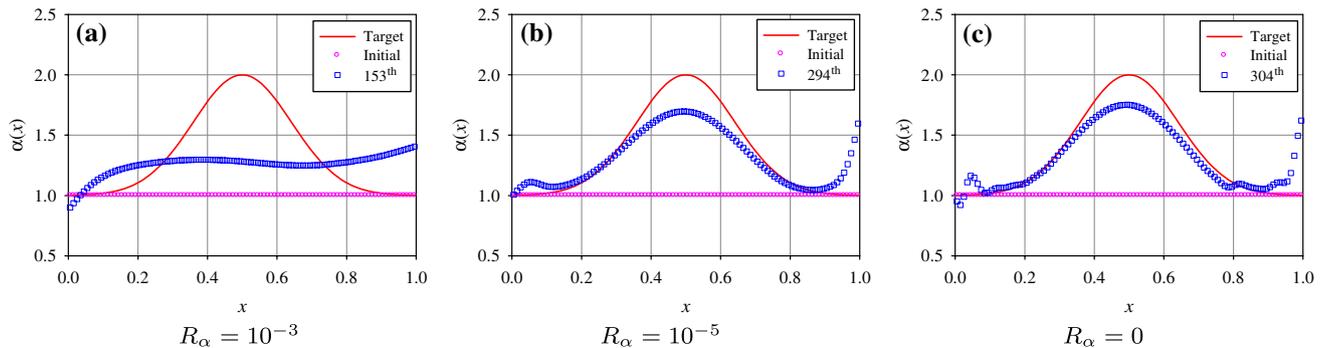

Fig. 6 Target, initial and estimated profile of $\alpha(x)$ using the Tikhonov regularization scheme

was used are shown in Fig. 8: no major qualitative differences are observed between the TV and Tikhonov schemes for this smooth profile case.

7.1.2 Initial estimates effects

Even though, in the preceding problem, our initial guess for the modulus distribution is a constant one throughout the domain, it is still close to the modulus values at the ends of the domain. To explore the algorithmic performance when the initial guess, while still constant, is not close to the end values, we seek to reconstruct the profile starting with $\alpha(x, t) = 0.7$. We test again using different regularization parameters, $R_\alpha = 1.0, 0.1, 0.01$. We set the tolerance to 2.0×10^{-7}. For the TD regularization scheme, we report only the case in which the modified time-independent matrices (Scheme TD-II) are used in solving the state and adjoint problems, given the comparative performance reported in the preceding section. As it can be seen in Fig. 9, the reconstructed profile converged to the target when the TD regularization was used, regardless of the values of the regularization parameter R_α (for the range considered). The effect of the regularization parameter on the rate of convergence is shown in Fig. 9d.

Next, we used the Tikhonov regularization scheme, exploring various values for the regularization parameter, in

an attempt to reconstruct the profile when the initial guess was again set at $\alpha(x, t) = 0.7$. In contrast to the case where the initial guess was closer to the target profile, here all attempts failed to converge as depicted in Fig. 10. The performance of the TV scheme was equally problematic, as shown in Fig. 11.

7.1.3 Noise effects

One of the practical difficulties arising in inversion is associated with the presence of noise in the measured response (and/or on the source) that typically result in contaminated data. It is thus of interest to study the performance of the algorithms in the presence of noisy data. The problem parameters are the same as those of the first test case, that is, a smooth target profile with an initial guess at $\alpha(x, t) = 1$. However, now noise-contaminated data are used. To artificially inject noise in the measured data, we use Gaussian noise (GN) having standard deviation of 1, 5, and 10% with respect to the maximum amplitude of the measured data. The noise-contaminated data are shown in Fig. 12.

The reconstructed profiles obtained using the time-dependent regularization scheme are shown in Fig. 13. For all cases, $R_\alpha = 0.01$ is employed. As can be seen in the figure, the

Fig. 7 Convergence pattern of estimated profile of $\alpha(x)$ using the Tikhonov regularization scheme

Misfit error

$R_\alpha = 10^{-3}$

$R_\alpha = 10^{-5}$

$R_\alpha = 0$

$R_\alpha = 10^{-3}$
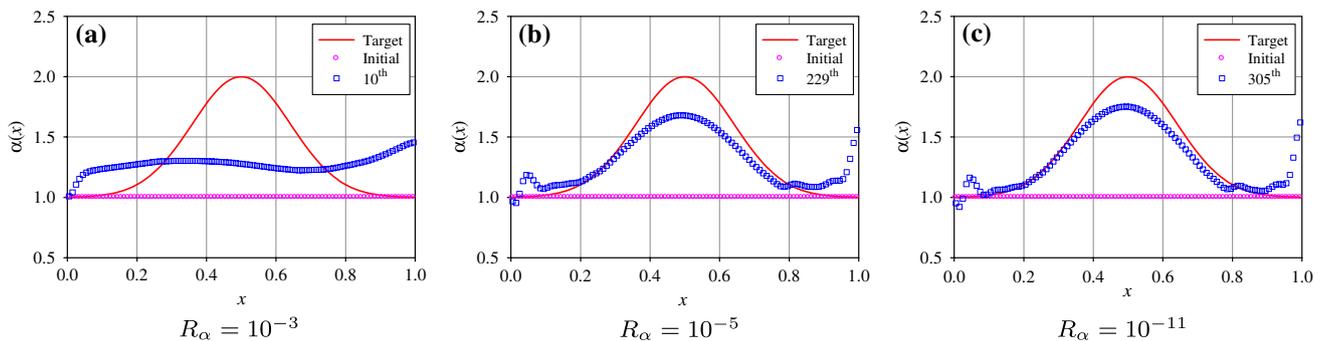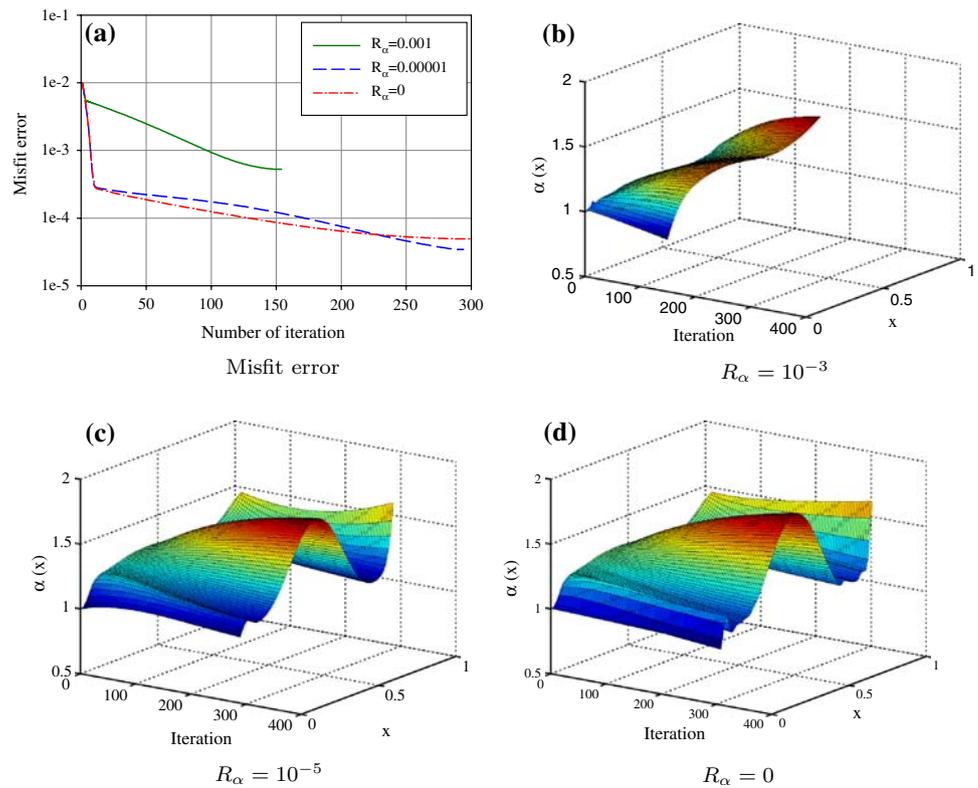
$R_\alpha = 10^{-5}$

$R_\alpha = 10^{-11}$

Fig. 8 Target, initial, and estimated profile of $\alpha(x)$ using the TV regularization scheme

solution converged satisfactorily to the target in all cases. However, as the level of the noise in the signal increases, the estimations show higher level of fluctuations due to the effects of the noise. If the regularization parameter increases, similar results to the above can be obtained, albeit at a slower convergence rate: Fig. 13d summarizes the convergence of the misfit error for data contaminated with different noise levels.

Next, we consider again a Tikhonov scheme and study the performance for a single case with 10% Gaussian noise. We adopt different values of the regularization parameter. The inverted profiles are shown in Fig. 14. Figure 14a depicts the reconstructed profile when no regularization is used ($R_\alpha =$

0); interestingly, despite the mild oscillations, the profile is certainly competitive to, or even better than, the one produced by the TD regularization (see Fig. 13c). With $R_\alpha = 10^{-5}$, the reconstructed profile becomes smoother, albeit somewhat deviating from the target (Fig. 14b) at the neighborhood of the peak. As the regularization parameter increases, further deviation (or failure) is observed (e.g., Fig. 14c), since the Tikhonov regularization begins to weigh heavily on the inversion process.

When the TV scheme is used, as shown in Fig. 15, the solutions are more oscillatory than the Tikhonov case and, in this instance, we could not observe any superior performance over the TD or the Tikhonov schemes.

Fig. 9 Target, initial and estimated profile of $\alpha(x)$, and misfit error using the TD-II regularization scheme
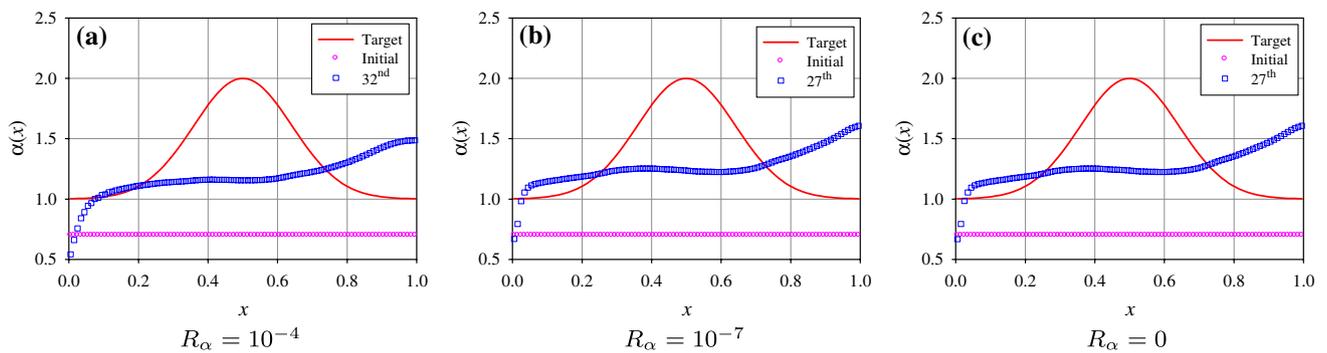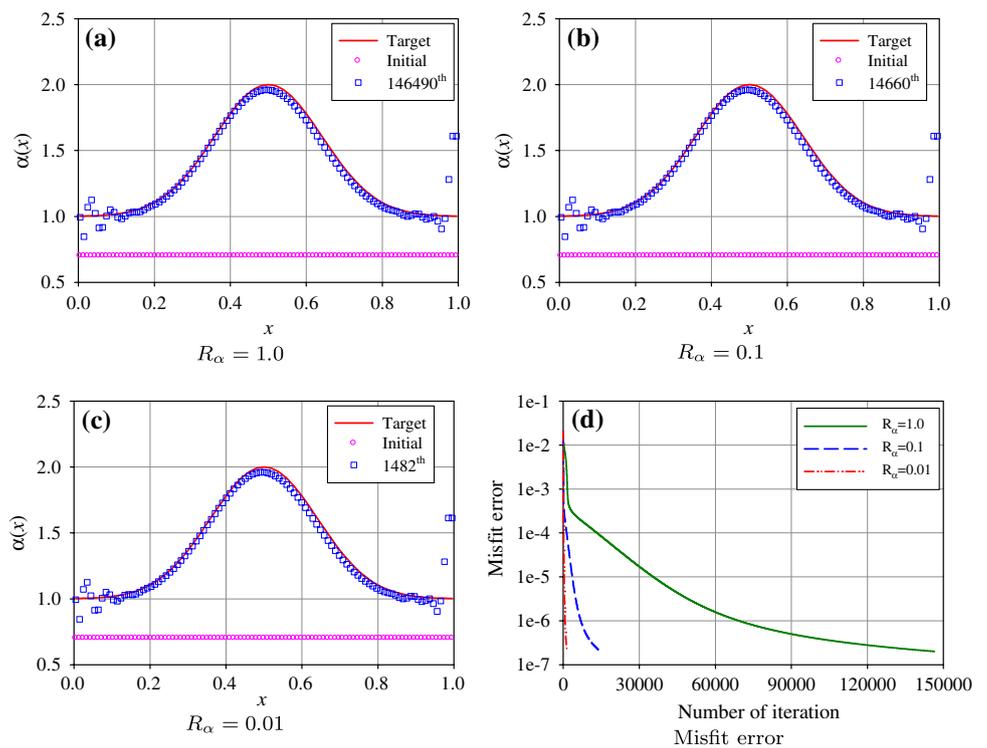
Fig. 10 Target, initial and estimated profile of $\alpha(x)$ using the Tikhonov regularization scheme
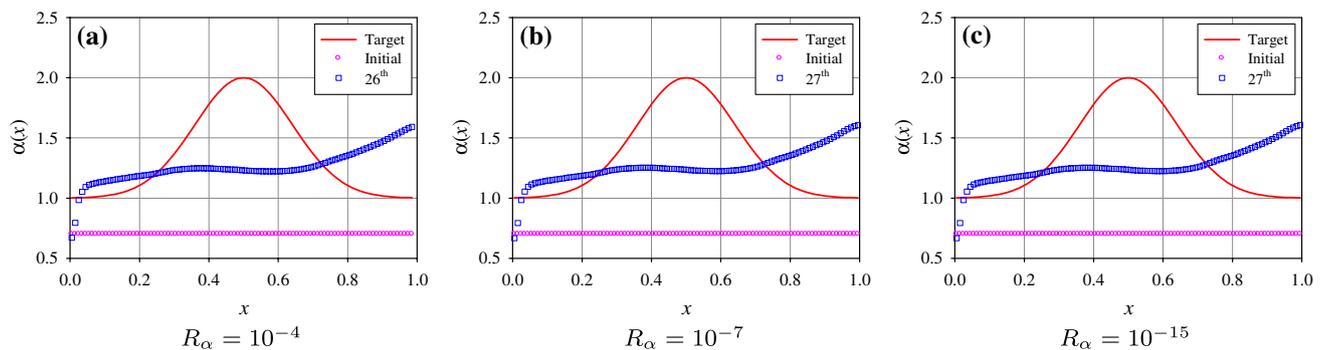
Fig. 11 Target, initial and estimated profile of $\alpha(x)$ using the TV regularization scheme
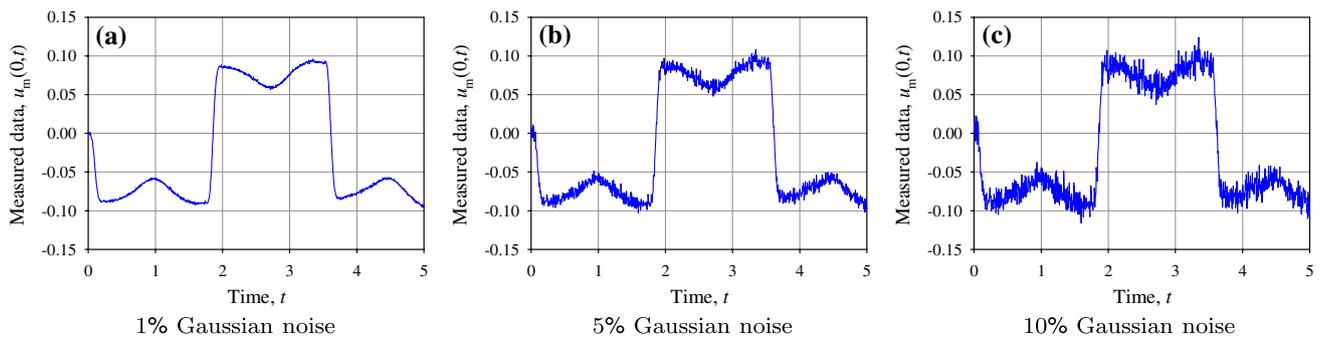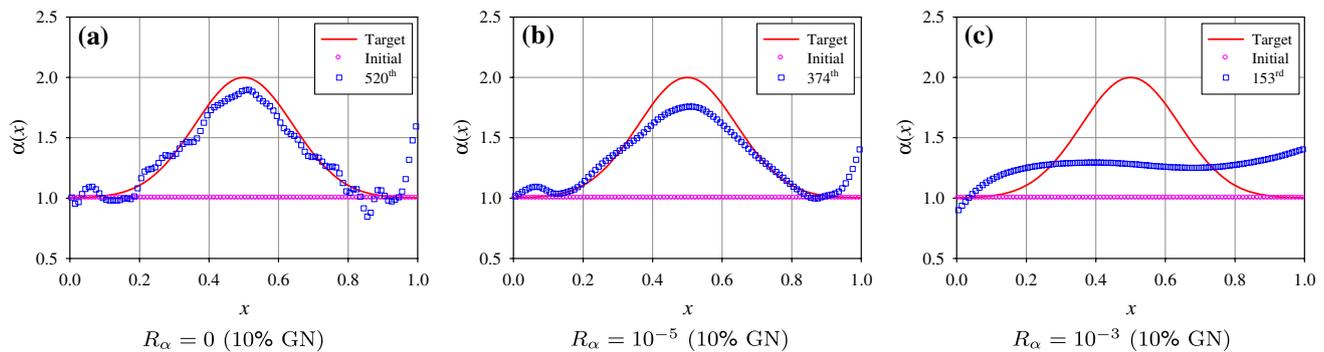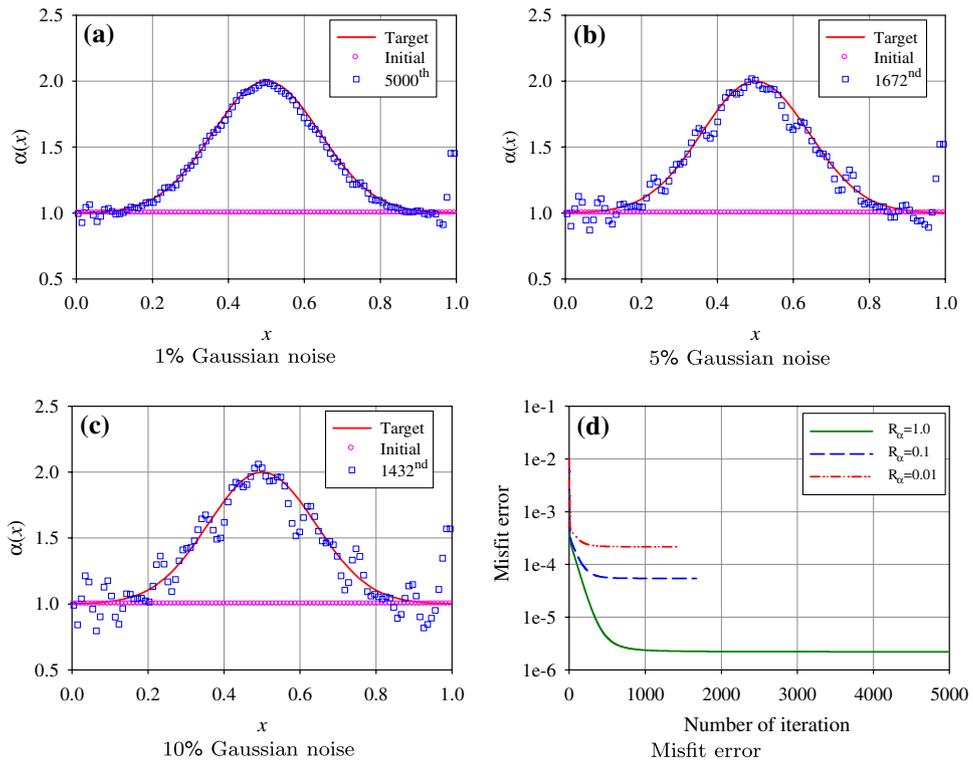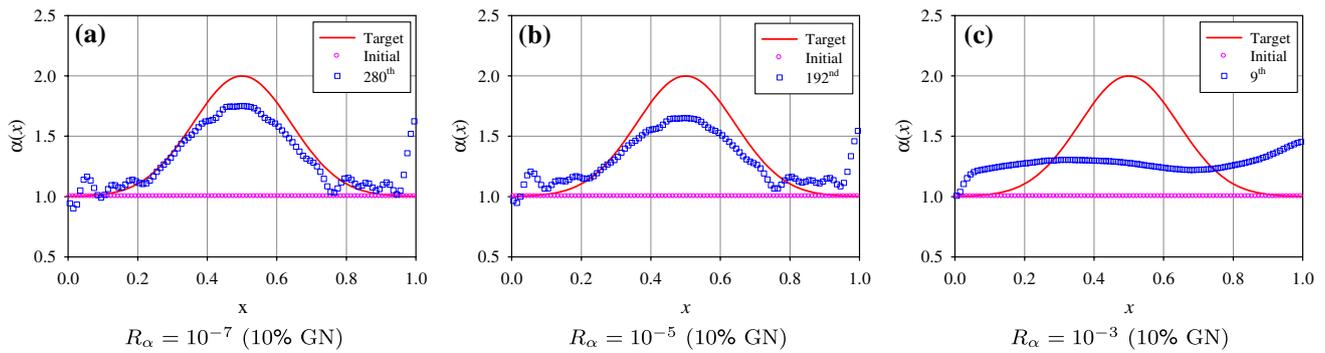
Fig. 12 Measured data contaminated by noise

Fig. 13 Target, initial and estimated profile of $\alpha(x)$ using the TD regularization scheme

Fig. 14 Target, initial and estimated profile of $\alpha(x)$ using the Tikhonov regularization scheme

Fig. 15 Target, initial and estimated profile of $\alpha(x)$ using the TV regularization scheme

7.2 On sharp profiles

In the next series of numerical experiments we study the performance of the algorithms on sharply changing profiles. We start again by considering $\alpha(x)$ as the sole profile to invert for, and adopt the following step form for it:

$$\alpha(x) = \begin{cases} 1.0 & 0.0 \leq x < 0.3 \\ 2.0 & 0.3 \leq x < 0.7 \\ 1.0 & 0.7 \leq x < 1.0 \end{cases} . \tag{53}$$

We choose $\alpha(x, t) = 1.0$ as the initial guess and reconstruct the profile using a TD, a Tikhonov, and a TV regularization scheme. We also study the performance of each scheme with noise-contaminated data. For each scheme, we tested the performance using 0, 1, 5 and 10% Gaussian noise. In Fig. 16, the measured noise-contaminated data are shown.

The estimated profiles of $\alpha(x)$ using the TD regularization scheme are shown together with the target and initial profiles in Figs. 17 and 18. In particular, Fig. 17a–c are the converged profiles for different regularization parameters and noise-free data. Figure 18a–c depict the reconstructed profiles obtained using the noise-contaminated data; here we chose to report results only for $R_\alpha = 0.01$. The results are deemed satisfactory in all cases.

Next, the estimated profiles of $\alpha(x)$ using the Tikhonov regularization scheme are shown together with the target and initial profiles in Figs. 19 and 20. In contrast to the TD regularization, here we observe that the solutions depend on the choice of the regularization parameter. In addition, the results using the TV scheme are depicted in Figs. 21 and 22: TV's performance is similar to the Tikhonov case. For 10% noise the Tikhonov case appears superior to the TD case (compare Fig. 18c against Fig. 20c).

We seek to reconstruct the same profile starting with $\alpha(x, t) = 0.7$ to explore the algorithmic performance when the initial guess is not close to the end value. We again invert using both a TD and a Tikhonov regularization scheme and, for each scheme, we tested the performance using data contaminated with 0, 1, 5 and 10% Gaussian noise.

The estimated profiles of $\alpha(x)$ using the TD regularization scheme are shown together with the target and initial profiles in Figs. 23 and 24. Similarly to the previous case, it can be seen that the profiles converge to the target profile in all cases, but the convergence rate depends on the magnitude of the regularization parameter (Fig. 23). In addition, the reconstructed profiles using noisy data are depicted in Fig. 24; here we chose to report results only for $R_\alpha = 0.01$. The results are again deemed satisfactory in all cases. In contrast, Figs. 25
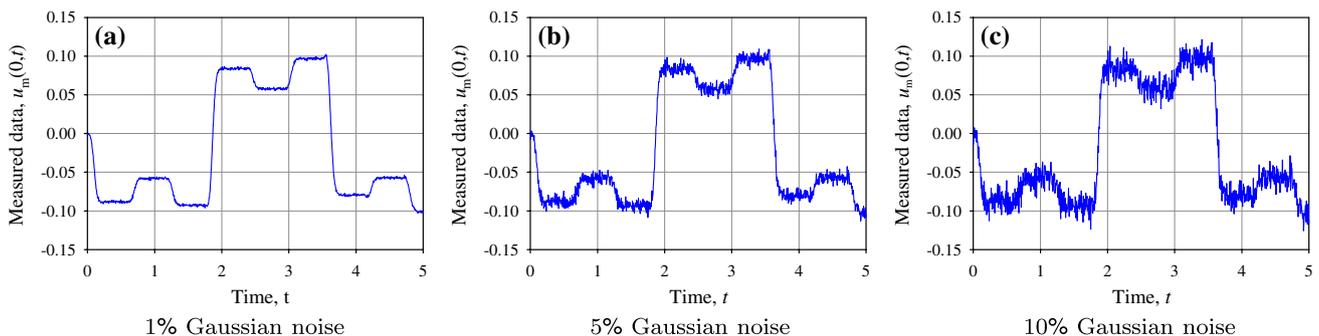

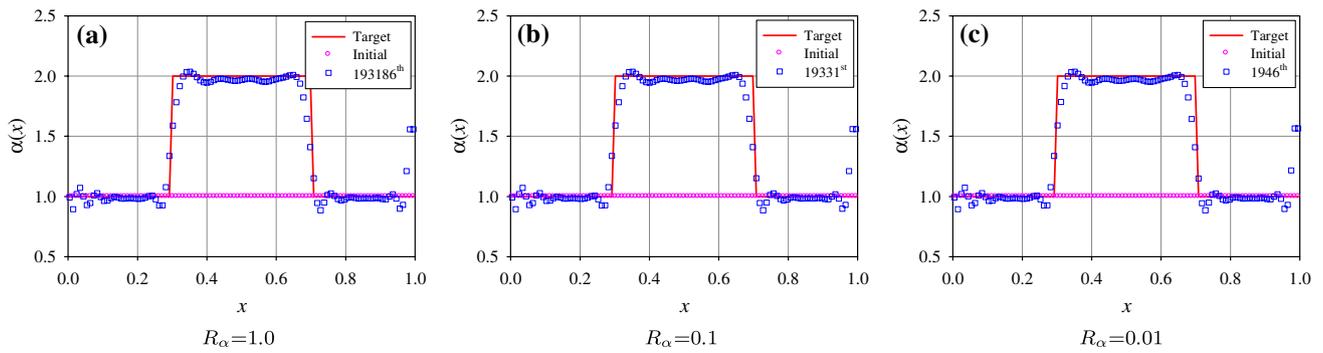**Fig. 16** Measured data contaminated by different levels of noise

Fig. 17 Target, initial and estimated profile of $\alpha(x)$ using the TD regularization scheme; noise-free data
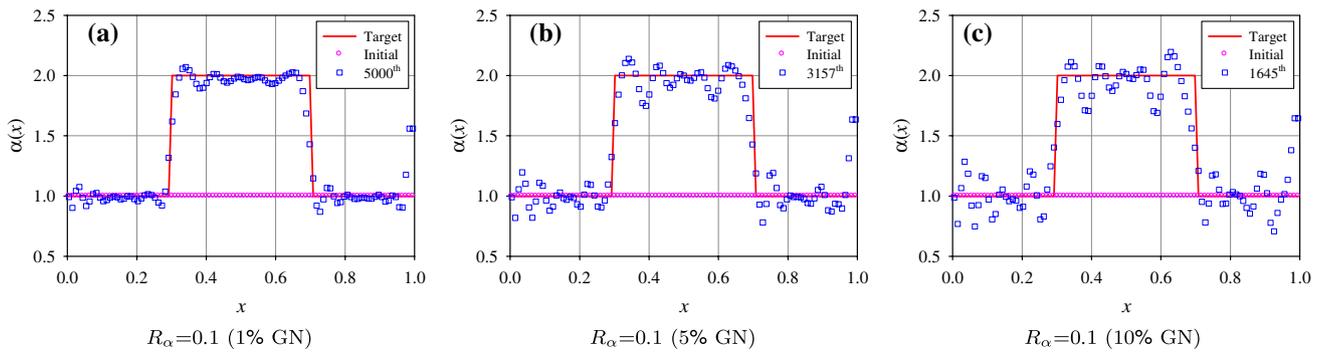

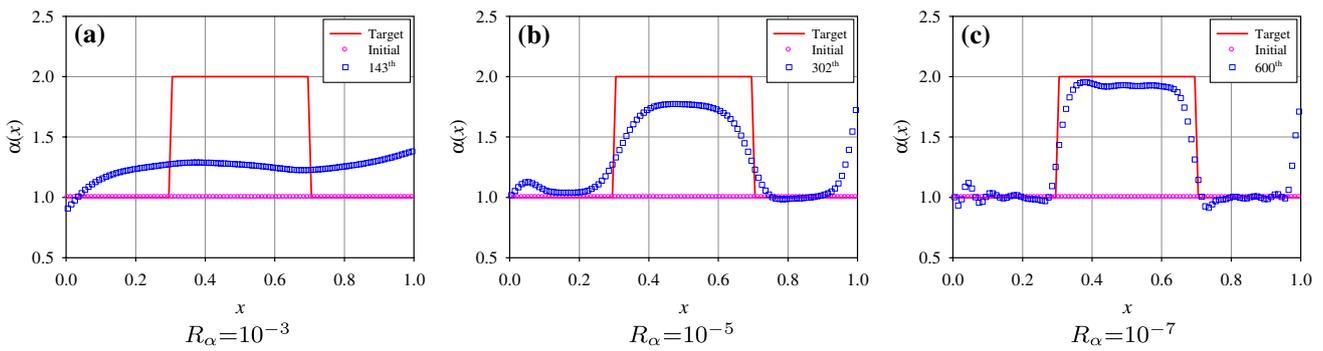**Fig. 18** Target, initial and estimated profile of $\alpha(x)$ using the TD regularization scheme; noisy data

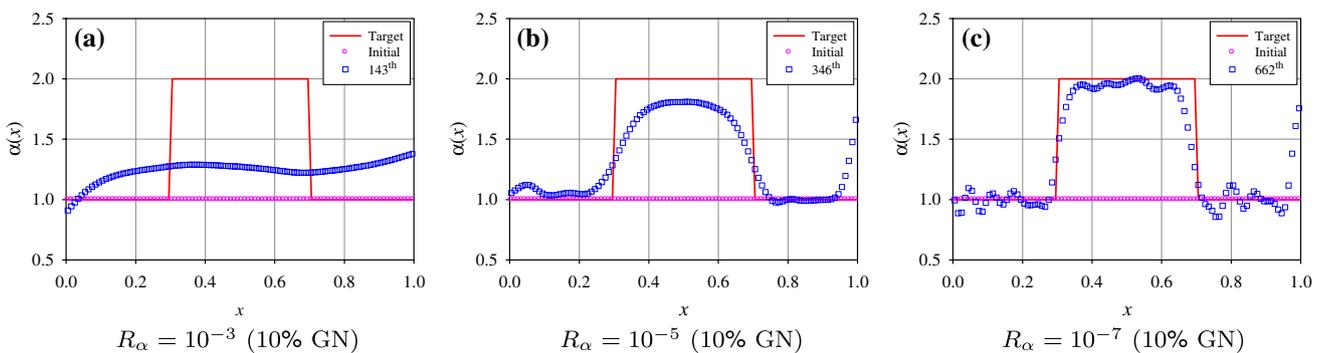**Fig. 19** Target, initial and estimated profile of $\alpha(x)$ using the Tikhonov regularization; noise-free data

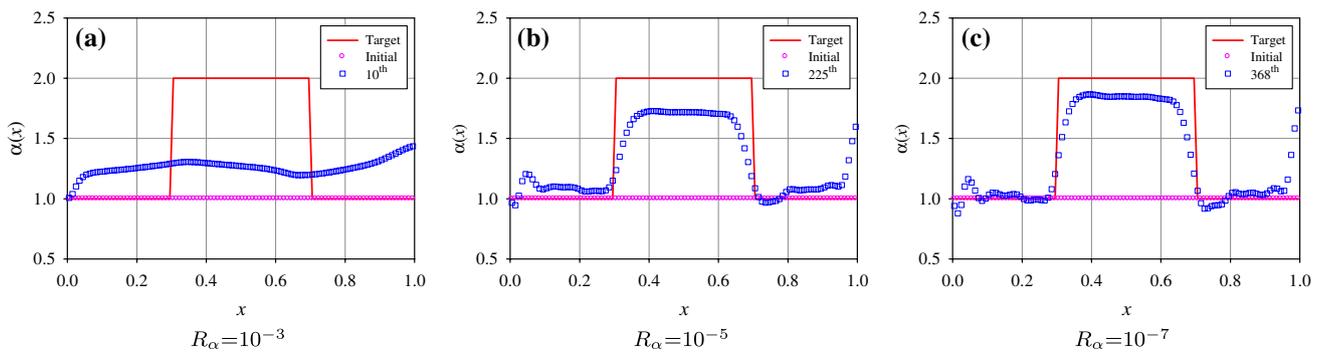**Fig. 20** Target, initial and estimated profile of $\alpha(x)$ using the Tikhonov regularization; noisy data

Fig. 21 Target, initial and estimated profile of $\alpha(x)$ using the TV regularization; noise-free data
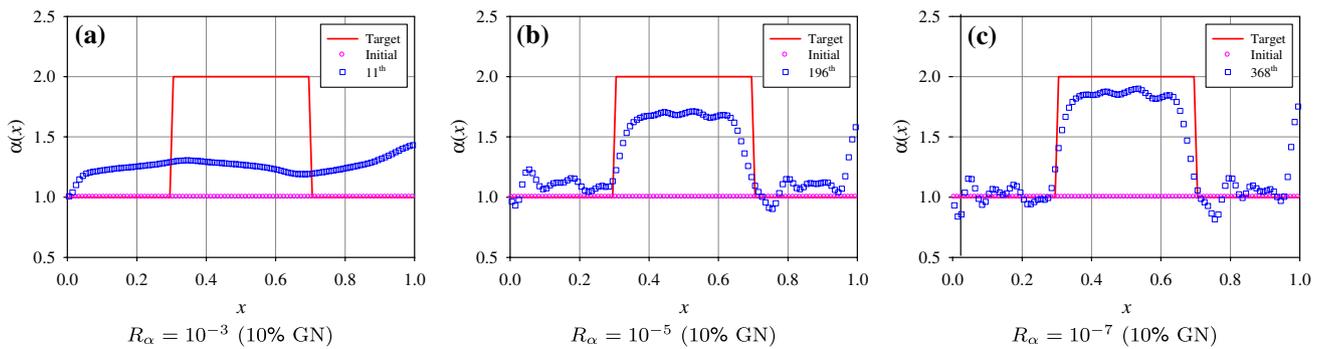

Fig. 22 Target, initial and estimated profile of $\alpha(x)$ using the TV regularization; noisy data

and 26 depict attempts to reconstruct the profile using the Tikhonov and TV schemes, respectively, which failed in all cases.

7.3 Simultaneous inversion of modulus and damping

We report next on the performance of the algorithm with the TD regularization when inverting simultaneously for both the modulus α, and the attenuation metric β. We tested two different profiles. In Case I, we choose a sharp modulus profile (multiple layers of different moduli), combined with a smooth linearly varying damping profile. Case II pertains to sharply varying profiles for both the modulus and damping. The target profiles for Case I are:

$$
\alpha(x) = \begin{cases} 1.5 & \text{for } 0.0 \le x < 0.25 \\ 2.0 & \text{for } 0.25 \le x < 0.5 \\ 1.0 & \text{for } 0.5 \le x < 0.75 \\ 2.5 & \text{for } 0.75 \le x < 1.0 \end{cases}, \tag{54}
$$

$$
\beta(x) = 1.0 - 0.5x, \tag{55}
$$

and for Case II:

$$
\begin{aligned}
\alpha(x) &= 1.5, & \beta(x) &= 0.6 & \text{for } 0.00 \le x < 0.25, \\
\alpha(x) &= 2.0, & \beta(x) &= 0.8 & \text{for } 0.25 \le x < 0.50, \\
\alpha(x) &= 1.0, & \beta(x) &= 0.4 & \text{for } 0.50 \le x < 0.75, \\
\alpha(x) &= 2.5, & \beta(x) &= 1.0 & \text{for } 0.75 \le x < 1.00.
\end{aligned} \tag{56}
$$

We choose constant $\alpha(x, t) = 1.2$ and $\beta(x, t) = 0.5$ as initial guesses for Case I, and $\alpha(x, t) = 1.2$ and $\beta(x, t) = 1.0$ for Case II. The profiles are reconstructed using both noise-free data, and data contaminated with 5% Gaussian noise. The data for each case are shown in Fig. 27; the regularization parameters were set at $R_\alpha = 0.01$ and $R_\beta = 0.01$.

The estimated profiles are shown together with the target and initial profiles in Figs. 28 and 29. As it can be seen, the performance is quite satisfactory; notice that the described process allows, in essence, the recovery of the number of layers, the material composition of each layer, and the thickness (or depth) of each layer, without having to explicitly declare them as model parameters.[2]

---

[2] It is possible to improve on the damping profile (β), by a judicious choice of R_β relative to R_α; the specifics escape the scope of this article.
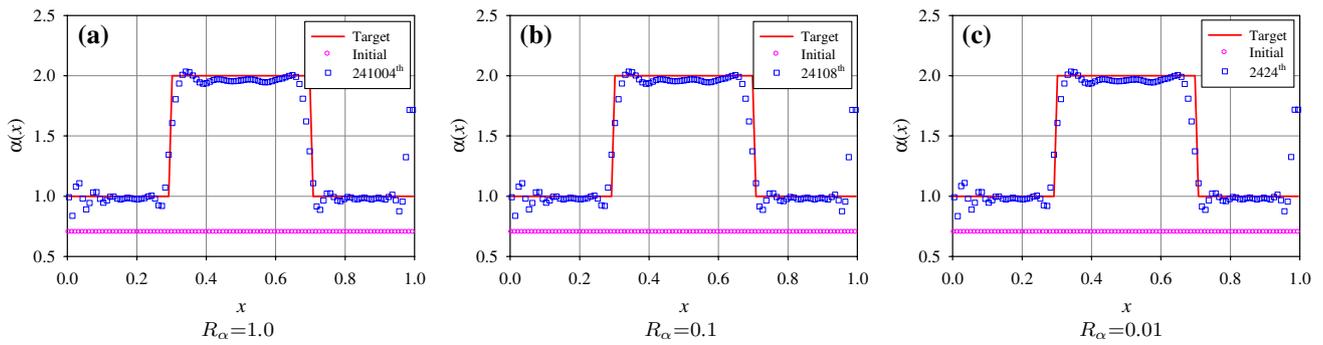
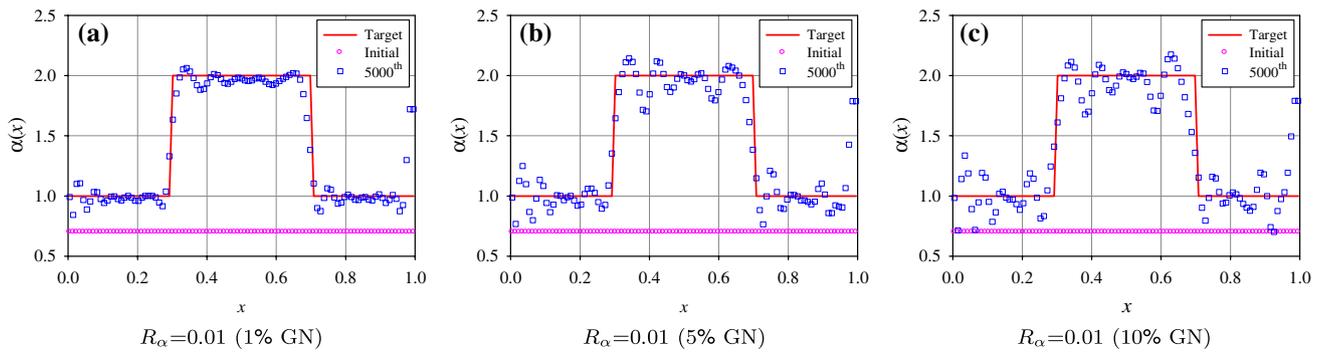**Fig. 23** Target, initial and estimated profile of $\alpha(x)$ using the TD regularization scheme; noise-free data

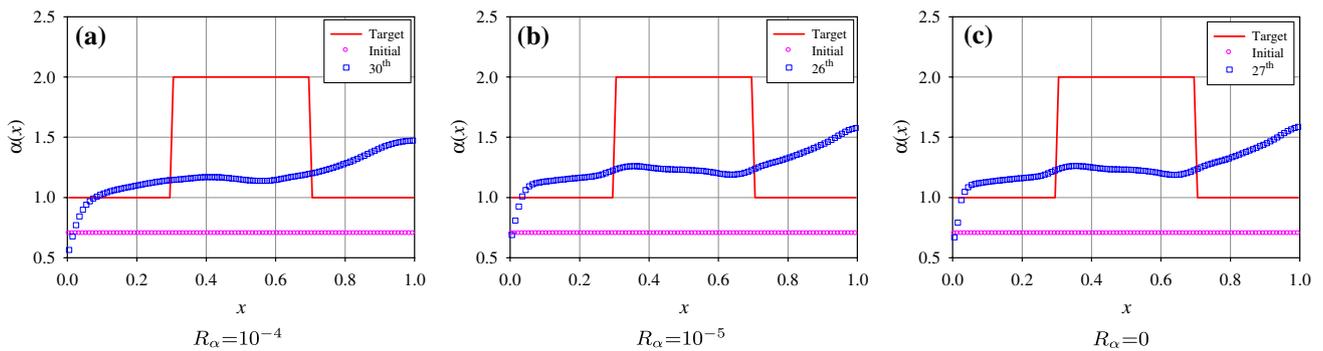**Fig. 24** Target, initial and estimated profile of $\alpha(x)$ using the TD regularization scheme; noisy data

Fig. 25 Target, initial and estimated profile of $\alpha(x)$ using the Tikhonov regularization scheme
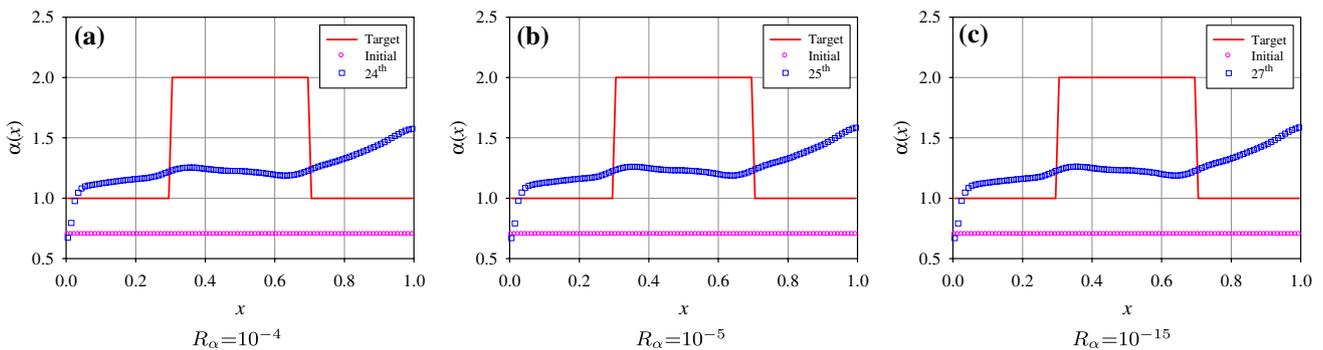

Fig. 26 Target, initial and estimated profile of $\alpha(x)$ using the TV regularization scheme
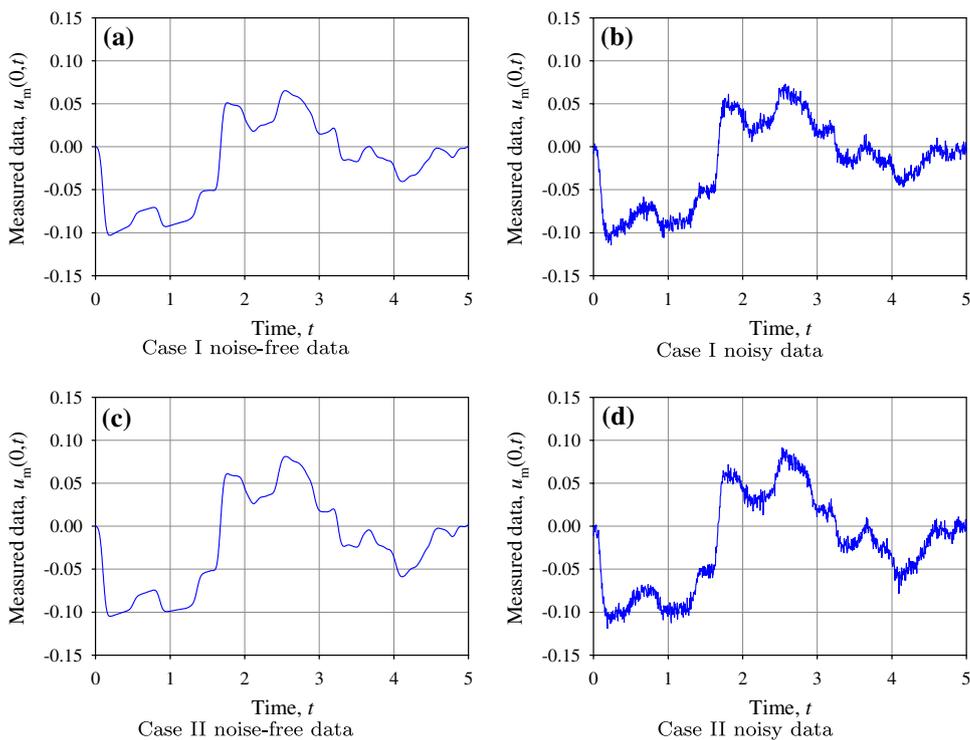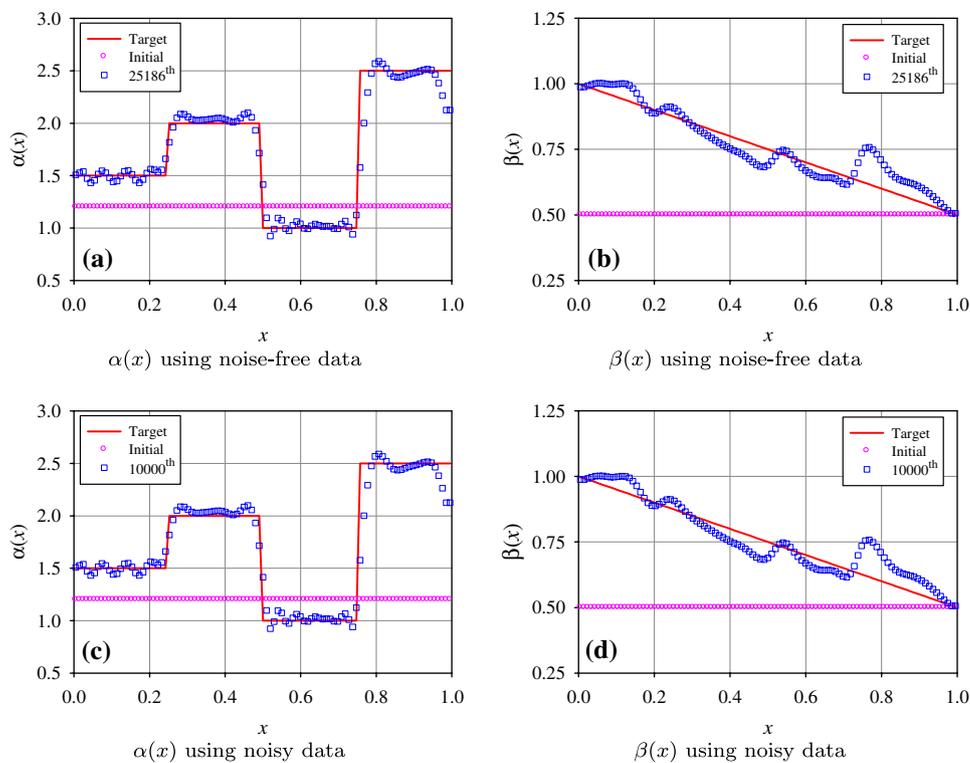
Fig. 27 Measured data

(a) Case I noise-free data

(b) Case I noisy data

(c) Case II noise-free data

(d) Case II noisy data

Fig. 28 Target, initial and estimated profiles of $\alpha(x)$ and $\beta(x)$ for Case I using the TD regularization scheme

(a) $\alpha(x)$ using noise-free data

(b) $\beta(x)$ using noise-free data

(c) $\alpha(x)$ using noisy data

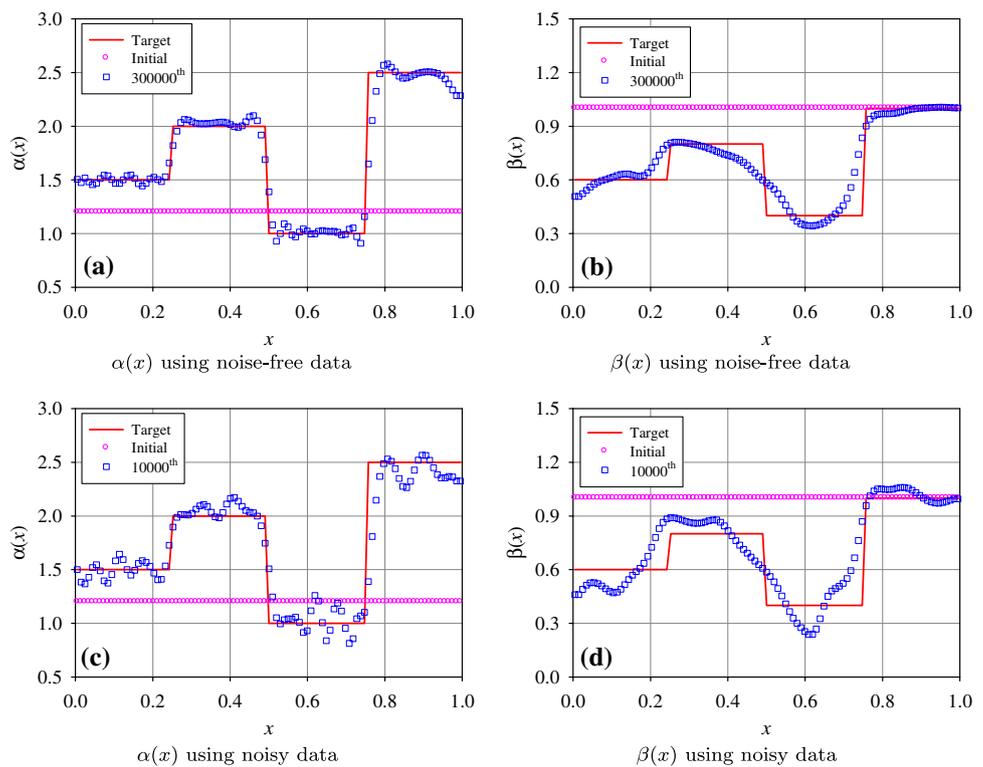(d) $\beta(x)$ using noisy data

8 Conclusions

We discussed a PDE-constrained optimization approach for reconstructing the material profile in a layered medium based on surficial measurements of its response to surface excitation directly in the time-domain. The primary focus was on the comparison, via numerical experiments, of spatial regularization schemes against a TD scheme. To this end, we

Fig. 29 Target, initial and estimated profiles of $\alpha(x)$ and $\beta(x)$ for Case II using the TD regularization scheme

(a) $\alpha(x)$ using noise-free data

(b) $\beta(x)$ using noise-free data

(c) $\alpha(x)$ using noisy data

(d) $\beta(x)$ using noisy data

tested the algorithmic performance for both smooth and sharp profiles, and discussed the effect the regularization parameter, noisy data, and initial estimates have on the inversion process. While the studies are not exhaustive, based on the reported numerical experiments, our observations are:

– The TD scheme can capture both sharp and smooth profiles, while the Tikhonov and TV regularizations exhibit difficulties when confronted with sharp profiles.
– The solutions obtained using the TD scheme appear less sensitive to the regularization parameter than the Tikhonov and TV schemes.
– The TD scheme appears more robust under inexact initial guesses when compared to the Tikhonov and TV schemes.
– In the presence of noisy data, the Tikhonov scheme seems to outperform the TD scheme only when initial estimates are close to the target profile.

Lastly, we also extended the development to the case of the simultaneous inversion for both modulus and damping and reported on the reconstruction of both smooth and sharp profiles. Overall, the TD approach appears robust (no failure to converge), albeit at a substantial computational cost over either Tikhonov or TV regularizations. Preliminary results in two dimensions with highly heterogeneous domains support similar conclusions. In our opinion, the TD scheme, despite its higher computational cost, is quite promising in lending robustness to inversion processes similar to the one discussed herein.

References

1. Stokoe KH, Wright SG, Bay JA, Roesset JM (1994) Characterization of geotechnical sites by SASW method. In: Woods RD (ed) Geophysical characterization of sites, pp 15–25
2. Rix GJ, Lai CG, Spang AW Jr (2000) In-Situ measurement of damping ratio using surface waves. J Geotech Geoenviron Eng 126:472–480
3. Rix GJ, Lai CG, Foti S (2001) In-Situ measurement of damping ratio using surface waves. Geotech Test J 24:350–358
4. Lai CG, Rix GJ, Foti S, Roma V (2002) Soil dynamics and earthquake engineering. Simultaneous Meas Inversion Surf Wave Dispersion Attenuation Curves 22:923–930
5. Tikhonov AN (1963) Solution of incorrectly formulated problems and the regularization method. Soviet Math Doklady 4:1035–1038
6. Rudin L, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. Physica D 60:259–268
7. Chan TF, Tai XC (2003) Identification of discontinuous coefficients in elliptic problems using total variation regularization. SIAM J Sci Comput 25(3):881–904
8. Biros G, Ghattas O (1999) Parallel Newton-Krylov Methods for PDE-constrained optimization. In: Proceedings of Supercomputing 1999, Portland, Oregon, USA
9. Akcelik V (2002) Multiscale Newton-Krylov methods for inverse acoustic wave propagation. Dissertation presented in partial

fulfillment of the requirements for the doctoral degree. Carnegie Mellon University, USA

10. Akcelik V, Biros G, Ghattas O (2002) Proceedings of the IEEE/ACM SC2002 Conference, Parallel multiscale Gauss-Newton-Krylov methods for inverse wave propagation, Baltimore, MD

11. Akcelik V, Bielak J, Biros G, Epanomeritakis I, Fernandez A, Ghattas O, Joong Kim E, Lopez J, O'Hallaron DR, Tu T, Urbanic J (2003) High-resolution forward and inverse earthquake modeling on terascale computers. In: Proceedings of ACM/IEEE SC2003, Phoenix, AZ, (2003)

12. Akcelik V, Bielak J, Biros G, Epanomeritakis I, Ghattas O, Kallivokas LF, Kim EJ (2004) Lecture Notes in Computer Science, Computational Science. A framework for online inversion-based 3D site characterization. Lecture Notes in Computer Science, Computational Science, vol 3038/2004, pp 717–724

13. Akcelik V, Biros G, Ghattas O, Hill J, Keyes D, van Bloemen Waanders B (2006) Frontiers of parallel computing. In: Heroux M, Raghaven P, Simon H (eds) SIAM, Parallel algorithms for PDE-constrained optimization

14. Guzina B, Madyarov AI (2005) On the spectral analysis of Love waves. Bull Seismol Soc Am 95(3):1150–1169

15. Tadi M (1997) Explicit method for inverse wave scattering in solids. Inverse Probl 13:509–521