

PAPER

Adaptive eigenspace method for inverse scattering problems in the frequency domain

To cite this article: Marcus J Grote *et al* 2017 *Inverse Problems* **33** 025006

View the [article online](#) for updates and enhancements.

You may also like

- [Quantum walks induced by Dirichlet random walks on infinite trees](#)
Yusuke Higuchi and Etsuo Segawa
- [Revisiting Riemannian geometry-based EEG decoding through approximate joint diagonalization](#)
Fotis P Kalaganis, Nikos A Laskaris, Vangelis P Oikonomou et al.
- [Adaptive spectral decompositions for inverse medium problems](#)
Daniel H Baffet, Marcus J Grote and Jet Hoe Tang

Adaptive eigenspace method for inverse scattering problems in the frequency domain

Marcus J Grote, Marie Kray and Uri Nahum

Department of Mathematics and Computer Science, University of Basel,
Basel, Switzerland

E-mail: marcus.grote@unibas.ch

Received 6 July 2016

Accepted for publication 2 December 2016

Published 9 January 2017



CrossMark

Abstract

A nonlinear optimization method is proposed for the solution of inverse scattering problems in the frequency domain, when the scattered field is governed by the Helmholtz equation. The time-harmonic inverse medium problem is formulated as a PDE-constrained optimization problem and solved by an inexact truncated Newton-type iteration. Instead of a grid-based discrete representation, the unknown wave speed is projected to a particular finite-dimensional basis of eigenfunctions, which is iteratively adapted during the optimization. Truncating the adaptive eigenspace (AE) basis at a (small and slowly increasing) finite number of eigenfunctions effectively introduces regularization into the inversion and thus avoids the need for standard Tikhonov-type regularization. Both analytical and numerical evidence underpins the accuracy of the AE representation. Numerical experiments demonstrate the efficiency and robustness to missing or noisy data of the resulting adaptive eigenspace inversion method.

Keywords: inverse medium problem, Helmholtz equation, PDE constrained optimization, adaptive eigenspace inversion, regularization, full waveform inversion

(Some figures may appear in colour only in the online journal)

1. Introduction

Inverse scattering problems occur in a wide range of applications such as radar and sonar technology, non-destructive testing, geophysical exploration or medical imaging. By illuminating an unknown body, the scatterer, with waves of various directions or wavelengths, one attempts to obtain information about that body from the scattered waves recorded at some distance. In

the inverse medium problem, the scatterer is a penetrable, bounded inhomogeneity inside the medium characterized by one or several varying physical parameters and the inverse problem consists in estimating these parameters from scattering data. Typical inverse medium problems include oil and gas exploration [35] in geophysics or breast tumor detection [23] in medical imaging. Numerical methods for the solution of inverse scattering problems essentially fall into either of two classes: qualitative and quantitative methods.

Qualitative methods [3] generally require little *a priori* knowledge about the scatterer and permit to estimate its location and shape quite efficiently. Examples are the MUSIC (multiple signal classification) algorithm [6, 19], the DORT (decomposition of the time reversal operator) method [4, 27], the linear sampling method [7, 15], the probe method [26], and the factorization method [19, 20]. These methods provide a criterion for deciding whether any point lies inside or outside the scatterer only on the basis of far-field measurements of the scattered field. Hence, they effectively determine the support of the scatterer and thus permit to quickly detect material defects or obstacles, but do not provide any further quantitative information about its physical characteristics such as the local sound speed.

Quantitative methods typically reformulate the inverse scattering problem as a PDE-constrained optimization problem, where the unknown physical parameters, the control u , are determined by minimizing an appropriate objective functional [34]. The objective functional, $L[y, u]$, measures the misfit between the simulated wave field, the state y , and the true scattered data, the observations y^{obs} . Both *optimize-then-discretize* or *discretize-then-optimize* strategies are common [13, 18]. To tackle the ill-posedness of the inverse problem, a penalization term is usually included in the objective functional, such as standard Tikhonov regularization [2, 37].

At least two strategies are then available: the *full-space* approach, where u and y are sought simultaneously, and the *reduced-space* approach, where y is eliminated from the objective functional as $L[u] = L[y(u), u]$, taking advantage of the linearity of the underlying wave equation [12, 13, 22, 36]. Then, standard Newton or quasi-Newton methods from nonlinear optimization can be applied [25, 17]. For inverse medium problems, however, the exact solution of the Newton equations at each iteration may be prohibitively expensive due to the very large number of (unknown) parameters. In recent years, inexact truncated Newton methods [10, 11, 24], where at each (outer) iteration the (quasi-)Newton equations are solved by using only a few (inner) Krylov subspace iterations, have proved particularly effective for large-scale inverse medium problems [22, 31].

The spatial discretization of the parameter u on a standard finite difference (FD) or finite element (FE) mesh typically results in a huge number of (unknown) nodal values, in fact increasingly so, as the mesh is refined. For time-dependent scattering problems, the adaptive eigenspace inversion (AEI) method [8, 9] recently achieved a significant reduction in the number of parameters by projecting u to a finite-dimensional basis of eigenfunctions of a particular elliptic differential operator, which is iteratively adapted during the nonlinear optimization.

Here, we propose the AEI method for the inverse medium problem in the frequency domain, where the scattered field is governed by the Helmholtz equation. In section 2, we describe our AEI approach, which combines state-of-the-art techniques from large-scale nonlinear optimization, such as inexact truncated Newton-like methods and frequency stepping [1, 5, 22], with an adaptive eigenspace representation of u for regularization. Next, in section 3, we present both analytical and numerical evidence which underpins the remarkable accuracy of our particular choice of basis functions. In particular, we show how adapting the dimension of the eigenspace basis effectively builds regularization into the inversion. Finally, we subject our AEI method in section 4 to a series of numerical tests that demonstrate not only its accuracy and robustness with respect to missing or noisy data, but also its versatility by combining it with a sample averaging approximation [14].

2. Adaptive eigenspace inversion (AEI)

In this section, we present the adaptive eigenspace inversion (AEI) method for inverse medium problems in the frequency domain, when the scattered wave field is governed by the Helmholtz equation. First, we formulate the inverse medium problem for the squared wave speed $u(x)$ as a PDE-constrained optimization problem. Next, we introduce the adaptive eigenspace representation for u , the distinguishing feature of the AEI method. Finally we describe in detail the full AEI algorithm where both the dimension and the basis functions themselves are adapted during the optimization process.

2.1. Inverse medium problem

We consider a time-harmonic scattering problem in unbounded space from a penetrable inhomogeneity located inside a bounded convex domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$. Outside Ω , the wave speed is known and may vary. Inside Ω , the scattered field y satisfies the Helmholtz equation together with a Sommerfeld-type radiation condition at the boundary $\Gamma = \partial\Omega$, for simplicity:

$$\begin{cases} -\omega^2 y - \nabla \cdot (c^2(x) \nabla y) = f(x), & \text{in } \Omega, \\ \frac{\partial y}{\partial n} - ik(x)y = 0, & \text{on } \Gamma. \end{cases} \quad (1)$$

Here $\omega > 0$ denotes the constant frequency, $c(x) > 0$ the wave speed and $k(x) = \omega/c(x)$ the wave number.

Next, we perform N_s illuminations of the medium inside Ω with source terms $f = f_\ell$, $\ell = 1, \dots, N_s$, and denote by y_ℓ the corresponding (unique) solutions of (1). Given the measurements y_ℓ^{obs} on Γ , or part of it, we seek to reconstruct the (unknown) squared wave speed $u = c^2$ inside Ω such that every solution y_ℓ of (1) with $f = f_\ell$ coincides at Γ with the measurements y_ℓ^{obs} , $\ell = 1, \dots, N_s$. In doing so, we assume that the wave speed c is known on the boundary Γ .

To solve the inverse medium problem, we now formulate it as a PDE-constrained optimization problem and thus seek a minimizer u of the standard data misfit functional

$$L[u] = \frac{1}{2} \sum_{\ell=1}^{N_s} \|y_\ell(u) - y_\ell^{\text{obs}}\|_{L^2(\Gamma)}^2. \quad (2)$$

Then, the inverse problem reads:

$$\begin{aligned} &\text{Find } u \in V = L^\infty(\Omega) \text{ such that} \\ &\quad u = \underset{v \in V}{\operatorname{argmin}} L[v], \\ &\text{with } y_\ell(u) \text{ satisfying (1) for } f = f_\ell, \ell = 1, \dots, N_s. \end{aligned} \quad (3)$$

Since the inverse problem (3) is generally ill-posed, Tikhonov regularization term is typically added for stability. Instead, we shall incorporate regularization by restricting u to a finite-dimensional subspace V_K , adaptively determined during the optimization.

To solve (3), we shall consider standard Newton or Quasi-Newton methods, which require the gradient of $L[u]$. For a direction p , it is given by

$$\langle \nabla L[u], p \rangle = \sum_{\ell=1}^{N_s} \left(\int_{\Omega} (\nabla y_\ell(u) \cdot \nabla \bar{z}_\ell(u)) p \right), \quad (4)$$

where each Lagrange multiplier z_ℓ , $\ell = 1, \dots, N_s$, solves the adjoint problem

$$\begin{cases} -\omega^2 z_\ell - \nabla \cdot (u(x) \nabla z_\ell) = 0, & \text{in } \Omega, \\ \frac{\partial z_\ell}{\partial n} + ik(x) z_\ell = (y_\ell - y_\ell^{\text{obs}}), & \text{on } \Gamma. \end{cases} \quad (5)$$

For the solution of (3), we can choose between the *optimize-then-discretize* or the *discretize-then-optimize* approaches. In the former case, we solve the inverse problem (2)–(4), whereas in the latter case, we consider the discrete functional

$$L_h[u] = \frac{1}{2} \sum_{\ell=1}^{N_s} \|P y_\ell(u) - y_\ell^{\text{obs}}\|_2^2, \quad (6)$$

together with its corresponding discrete gradient

$$\nabla L_h[u] = \sum_{\ell=1}^{N_s} \left(\frac{\partial y_\ell}{\partial u} \right)^\top P^\top (P y_\ell(u) - y_\ell^{\text{obs}}), \quad (7)$$

where P denotes a projection matrix from the underlying discrete computational subspace to the observations. We will consider either approach indifferently and in doing so demonstrate that our approach does not rely on any particular choice of discretization.

2.2. Adaptive eigenspace basis

Instead of a standard nodal basis (FD, FE grid-based discretization), we shall use a basis of (global) eigenfunctions $\{\phi_m\}_{m \geq 1}$ to represent the parameter $u(x)$ as

$$u(x) = u_0(x) + \sum_{m \geq 1} \beta_m \phi_m(x). \quad (8)$$

Here the ‘background’ $u_0(x) \in H^1(\Omega)$ solves the elliptic problem:

$$\begin{cases} -\nabla \cdot (\mu(x) \nabla u_0(x)) = 0, & \forall x \in \Omega, \\ u_0(x) = c^2(x), & \forall x \in \Gamma, \end{cases} \quad (9)$$

where $\mu(x)$ is defined by

$$\mu(x) = \frac{1}{\max\{|\nabla u(x)|, \varepsilon\}}, \quad \forall x \in \Omega, \varepsilon > 0. \quad (10)$$

The parameter $\varepsilon > 0$, which ensures that the denominator of μ does not vanish, is typically set to a very small value such as $\varepsilon = 10^{-6}$. While the primary role of u_0 is to accommodate the (known) inhomogeneous boundary values of u , we shall demonstrate in section 3, that u_0 also captures much of the behavior of u in the interior.

Following [8], we choose for the functions ϕ_m the orthonormal basis of eigenfunctions $\phi_m \in H_0^1(\Omega)$ of the elliptic operator,

$$\begin{cases} -\nabla \cdot (\mu(x) \nabla \phi_m(x)) = \lambda_m \phi_m(x), & \forall x \in \Omega, \\ \phi_m(x) = 0, & \forall x \in \Gamma, \end{cases} \quad (11)$$

with corresponding eigenvalues $0 < \lambda_m \leq \lambda_{m+1}$, $m \geq 1$. Clearly, at higher λ_m , the eigenfunctions ϕ_m in (11) will be increasingly oscillatory.

In section 3, we shall provide analytical and numerical evidence which underpins the remarkable accuracy of this basis for representing any given $u(x)$. In our AEI approach, the eigenfunctions $\{\phi_m\}_m$ and the background u_0 are repeatedly recomputed as the underlying control $u(x)$ varies during the optimization. Hence, we call $\{u_0\} \cup \{\phi_m\}_{m \geq 1}$ an adapted eigenspace (AE) basis.

Since $u(x)$ is precisely the quantity we seek, and thus unknown, we always use in (10) the value from the previous optimization step. At the first step, when no information about u inside Ω is available yet, we simply set $\mu(x) \equiv 1$. Then, u_0 is a harmonic prolongation of c^2 from Γ into Ω while the basis $\{\phi_m\}_{m \geq 1}$ simply corresponds to the eigenfunctions of the Laplacian operator in Ω .

Remark 1. The elliptic operator in (9) and (11) essentially coincides with the gradient of the penalized total variation (TV) regularization term [32, 38],

$$R_{\text{TV}}(u) = \frac{1}{2} \int_{\Omega} \sqrt{|\nabla u|^2 + \varepsilon^2} \, dx,$$

given by

$$\begin{aligned} \nabla R_{\text{TV}}(u) &= -\nabla \cdot \left(\frac{1}{\sqrt{|\nabla u|^2 + \varepsilon^2}} \nabla u \right) \\ &\simeq -\nabla \cdot (\mu(x) \nabla u). \end{aligned}$$

Penalized TV-regularization is well-known in image processing for noise removal, while preserving sharp interfaces. Instead of adding a Tikhonov regularization term to the objective functional, the AEI approach projects u to the basis of eigenfunctions of the gradient of the penalized TV-regularization functional; hence, the AE basis inherits similar properties.

2.3. AEI algorithm

In practice, we truncate the infinite sum in (8) at a finite number $K \geq 1$:

$$u(x) = u_0(x) + \sum_{m=1}^K \beta_m \phi_m(x). \quad (12)$$

To keep both the memory requirements and the computational effort low, it is imperative to keep the number K of eigenfunctions minimal. The truncation of the eigenfunction expansion is also crucial for numerical stability, as it builds regularization into the AEI approach—see remark 1 above but also remark 3 below.

At higher frequencies, waves detect and carry more detailed information about the scatterer, yet the number of local minima of $L[u]$ may also increase. To minimize the chance of landing in a (false) local minimum, we also apply a standard frequency continuation procedure [1, 5]. First, we solve the inverse problem (3) at the lowest frequency ω_1 . Then we progressively increase $\omega = \omega_2, \dots, \omega_n$ while re-initializing the optimization at every ω_j from the previous lower frequency ω_{j-1} . In doing so, we assume that the measurements are available through a range of frequencies, for instance via Fourier transform of a time-dependent signal.

AEI Algorithm.

Input: initial guess $u = 1$, observations y_ℓ^{obs} . Output: u^* .

1. Choose $K \geq 1$ and compute $\{\phi_m\}_{m=1}^K$ from (11) and u_0 from (9) with $\mu \equiv 1$
2. Expand $u(x) = u_0(x) + \sum_{m=1}^K \beta_m \phi_m(x)$
3. For $\omega = \omega_1, \dots, \omega_n$
 - (a) Compute $L[u]$ and $\nabla L[u]$, set H
 - (b) STOP: if $\|\nabla L[u]\| \leq \text{Tol}$
 - i. Solve $Hp = -\nabla L[u]$
 - ii. Determine step size α and set $u := u + \alpha p$
 - iii. Update $L[u]$, $\nabla L[u]$ and H
 - (c) Set μ from (10) with ∇u
 - (d) Update K , compute $\{\phi_m\}_{m=1}^K$ from (11) and u_0 from (9)
 - (e) Expand $u(x) = u_0(x) + \sum_{m=1}^K \beta_m \phi_m(x)$
4. $u^* = u$

The AEI approach applies regardless of the underlying optimization method used. Here we consider truncated Newton-like methods [10, 24] and denote by H either the true Hessian or some approximation of it, depending on (Newton, BFGS, or Gauss–Newton) method used. In all cases the linear system in 3(b) is solved by a truncated CG-iteration with the Eisenstat–Walker criterion [11]. In 3(b)ii, the step size α of the search direction p is determined either by Armijo (Newton, Gauss–Newton) or Wolfe–Powell (BFGS) step-size control, depending on the underlying method [25]. In steps 1 and 3(d) of the AEI Algorithm, we compute the first K eigenfunctions in (11) by using a standard restarted Lanczos iteration [21].

3. Approximation properties of the AEI expansion

In the AEI method presented in section 2, the (unknown) parameter u is expanded as in (12) in the L^2 -orthogonal basis of eigenfunctions $\{\phi_m\}_{m=1,\dots,K}$ defined by (11) together with u_0 defined by (9). In this entire section, we shall assume that u is known and shall now provide some analytical and some numerical evidence which underpins the remarkable accuracy provided by our particular choice of u_0 and the AE basis.

3.1. One-dimensional case

In one space dimension, (9) reduces to

$$\begin{cases} -\frac{d}{dx}\left(\mu(x)\frac{d}{dx}u_0(x)\right) = 0 & \forall x \in (a, b), \\ u_0(a) = c^2(a), \quad u_0(b) = c^2(b), \end{cases} \quad (13)$$

where

$$\mu(x) = \frac{1}{\max\{|u'(x)|, \varepsilon\}}, \quad \forall x \in (a, b). \quad (14)$$

The following result is immediate.

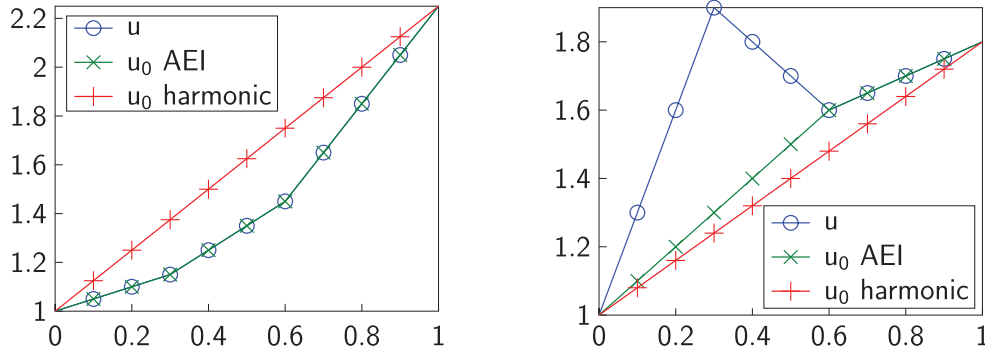


Figure 1. One-dimensional case. Comparison of adaptive and harmonic choices for u_0 . Left: u monotonically increasing; right: u arbitrary.

Proposition 1. Let $u : \bar{\Omega} \rightarrow \mathbb{R}$, $\Omega = (a, b)$, $a < b$, u continuous, piecewise differentiable and with $C \geq u'(x) \geq \varepsilon > 0$ or $-C \leq u'(x) \leq -\varepsilon < 0$, at every differentiability point $x \in \Omega$, $u(a) = c^2(a)$ and $u(b) = c^2(b)$. If u_0 solves (13) then $u_0(x) = u(x)$, $\forall x \in \bar{\Omega}$.

Proof. We first assume that u is strictly increasing and therefore $u'(x) \geq \varepsilon > 0$. Since $\mu(x) = 1/\max\{|u'(x)|, \varepsilon\} = 1/u'(x)$, u itself also satisfies (13) with $u_0(a) = c^2(a)$ and $u_0(b) = c^2(b)$. As μ is strictly positive through Ω , the (weak) solution of (13) is unique and therefore $u_0(x) = u(x)$, $\forall x \in \bar{\Omega}$.

If u is strictly decreasing, the proof is similar for $\mu(x) = -1/u'(x) > 0$. $\square$

Remark 2. From proposition 1 we conclude that our choice for u_0 is in fact *optimal in one space dimension* when u' is strictly positive (or negative) throughout Ω , since it automatically yields u itself. The same conclusion immediately holds in higher dimensions, if the background medium is layered, that is if $u(x_1, x_2, \dots, x_n)$ depends on a single variable. In general, however, u_0 will not equal u , in particular when u is not monotonic.

First, we illustrate the usefulness of our particular choice for u_0 in (13) and (14) by comparing it to a straightforward choice where u_0 is harmonic and satisfies (13) with $\mu(x) \equiv 1$. If u is strictly increasing (or decreasing) over $\Omega = (0, 1)$, u_0 defined by (13) and (14) automatically coincides with u —see proposition 1. Clearly, if u_0 is merely harmonic over Ω , it does not coincide with u , as shown in figure 1. If u both increases and decreases over Ω , neither the harmonic nor the adaptive u_0 will coincide with u . However, the adaptively computed u_0 will in general better approximate u inside Ω , as shown in figure 1; in fact, the adaptive u_0 coincides with u over $[0.6, 1]$.

Next, we approximate the difference $u - u_0$ in $\text{span}\{\phi_1, \phi_2, \dots, \phi_K\}$, where all ϕ_m satisfy (11) in one space dimension, that is

$$\begin{cases} -\frac{d}{dx}\left(\mu(x)\frac{d}{dx}\phi_m(x)\right) = \lambda_m\phi_m(x) & \forall x \in (a, b), \\ \phi_m(a) = 0, & \phi_m(b) = 0, \end{cases} \quad (15)$$

with μ defined in (14).

Clearly, the local behavior of ϕ_m in the neighborhood of any fixed $x_0 \in \Omega$ is essentially determined by the magnitude of $C = |u'(x_0)|$. If $|u'(x)| \simeq C > \varepsilon$ in a neighborhood of x_0 , ϕ_m essentially behaves like

$$\phi_m(x) \simeq A_m \sin(\sqrt{C\lambda_m} x) + B_m \cos(\sqrt{C\lambda_m} x)$$

near x_0 and hence is strongly oscillatory at higher m . However, if u is essentially constant near x_0 , $|u'(x)| \simeq 0$, then $\mu = 1/\varepsilon$ there and ϕ_m essentially behaves like

$$\phi_m(x) \simeq A_m \sin(\sqrt{\varepsilon\lambda_m} x) + B_m \cos(\sqrt{\varepsilon\lambda_m} x).$$

Since ε is very small, ϕ_m will also remain essentially constant there for moderate values of m .

To illustrate this behavior, we now consider the piecewise linear profile $u(x)$ shown in figure 2, where $u'(x) = 3, 0$ and -9 on the interval $(0, 0.3)$, $(0.3, 0.9)$ and $(0.9, 1)$, respectively. Since u' is piecewise constant, we can immediately determine all eigenfunctions as:

$$\phi_m(x) = \begin{cases} A_{m,1} \sin(\sqrt{3\lambda_m} x) + B_{m,1} \cos(\sqrt{3\lambda_m} x), & x \in [0, 0.3), \\ A_{m,2} \sin(\sqrt{\varepsilon\lambda_m} x) + B_{m,2} \cos(\sqrt{\varepsilon\lambda_m} x), & x \in [0.3, 0.9), \\ A_{m,3} \sin(\sqrt{9\lambda_m} x) + B_{m,3} \cos(\sqrt{9\lambda_m} x), & x \in [0.9, 1]. \end{cases} \quad (16)$$

In figure 2, we show u together with some of the eigenfunctions from (15). On every subinterval, ϕ_m has a different frequency determined by the local value of $\sqrt{C\lambda_m}$. In $[0.3, 0.9)$, the frequency $\varepsilon\lambda_m$ is very small, as $\varepsilon = 10^{-4}$, and ϕ_m appears essentially constant. As λ_m further increases, the frequency $\sqrt{\varepsilon\lambda_m}$ increases as well and oscillations appear. Clearly, the smaller ε , the more eigenfunctions ϕ_m essentially behave as constants wherever u is essentially constant. We remark that ϕ_1 nearly coincides with u up to a scaling factor.

Finally, we consider for u the smooth polynomial profile

$$u(x) = 1 - 89.95x^6 + 281.35x^5 - 339.51x^4 + 199.56x^3 - 60.85x^2 + 9.40x.$$

It is shown in figure 3 together with the first eigenfunction ϕ_1 obtained from (15) either with $\mu \equiv 1$ or μ as in (14). Again, we observe that the adaptively determined eigenfunction ϕ_1 captures well the main features of u up to a scaling factor, unlike the first eigenfunction of the Laplacian. To reach a relative error below 1% for the best approximation with respect to the L^2 -norm, only the first six eigenfunctions $\phi_1, \dots, \phi_6$ are needed in the adaptive case. In contrast, the first six eigenfunctions of the Laplacian yield a seven times larger relative L^2 -error, which drops below 1% only once thirteen eigenfunctions are included in the approximation; hence, the AE basis better captures the essential information about u .

3.2. Two-dimensional case

To illustrate the remarkable approximation properties of the AE basis in two space dimensions, we now consider the profile $u(x)$, $x = (x_1, x_2)$, shown in figure 4. Next, we compute u_0 from (9) with $\varepsilon = 10^{-6}$ and μ as in (10). In figure 5, we observe that u_0 matches with remarkable accuracy the background medium but misses the embedded kite-shaped obstacle. The first eigenfunction ϕ_1 from (11), however, ignores the background and captures precisely the remaining obstacle. Using u_0 and ϕ_1 , we expand u as in (12) with $K = 1$ and compute its best L^2 -approximation. Shown in figure 5, it is hardly distinguishable from the true u with well-defined sharp contours and a relative L^2 -error below 2%.

In contrast, if we repeat the same experiment with $\mu \equiv 1$, we observe in figure 5 how u_0 indeed matches the boundary values of u but fails to capture any additional features inside Ω . Similarly, the first eigenfunction of the Laplacian is independent of $u(x)$ and thus, as expected,

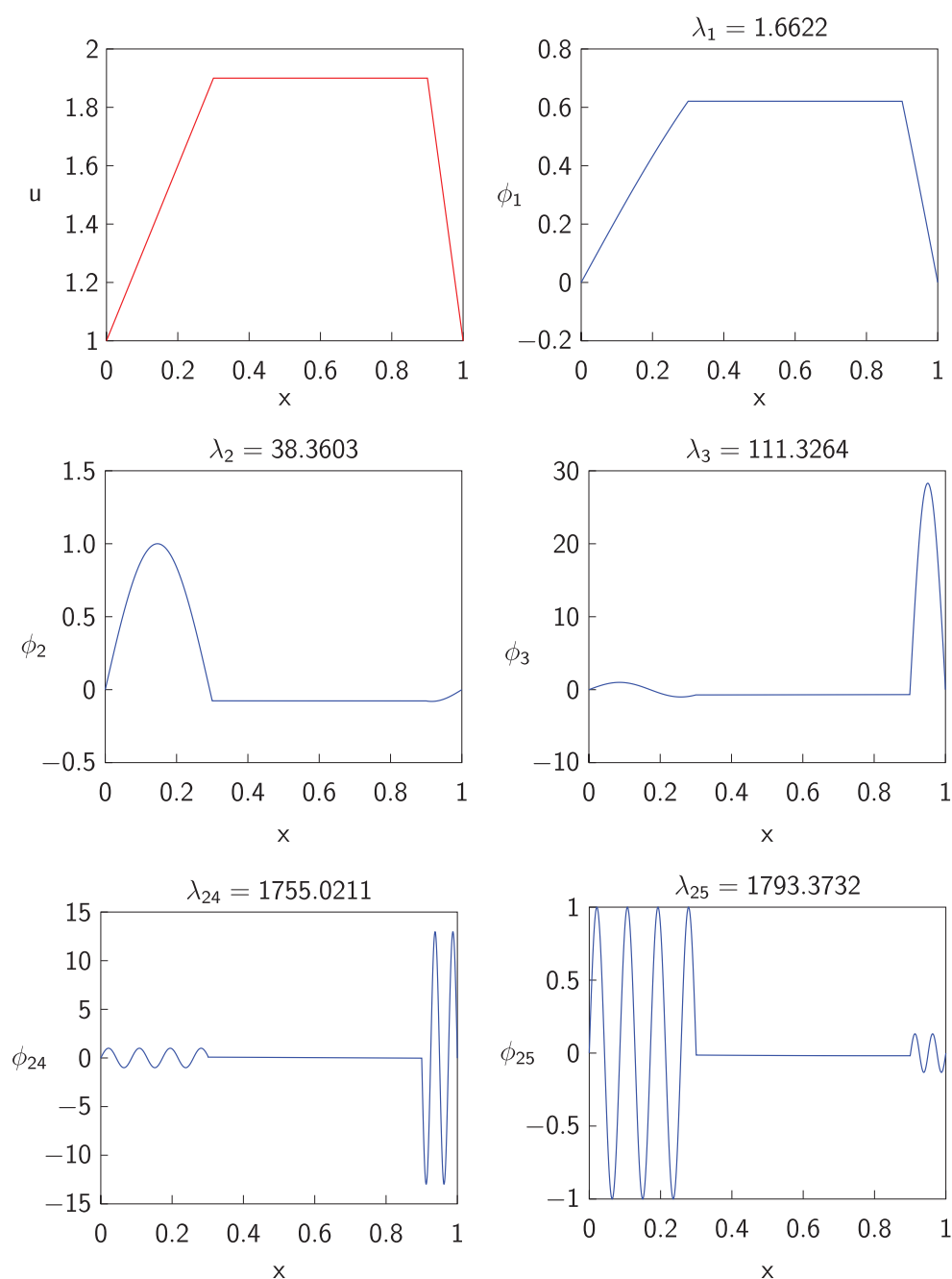


Figure 2. One-dimensional case. The true profile u (top left), together with the eigenfunctions $\phi_1, \phi_2, \phi_3, \phi_{24}$ and ϕ_{25} from (15) with $\varepsilon = 10^{-4}$.

carries no information about it. Again, we expand u as in (12) but now use the first 1000 eigenfunctions of the Laplacian instead. Although the L^2 best approximation with 1000 Laplacian eigenfunctions now yields a reasonable approximation of u with 6% relative L^2 -error, the contours are blurred while small high-frequency ripples appear due to the well-known

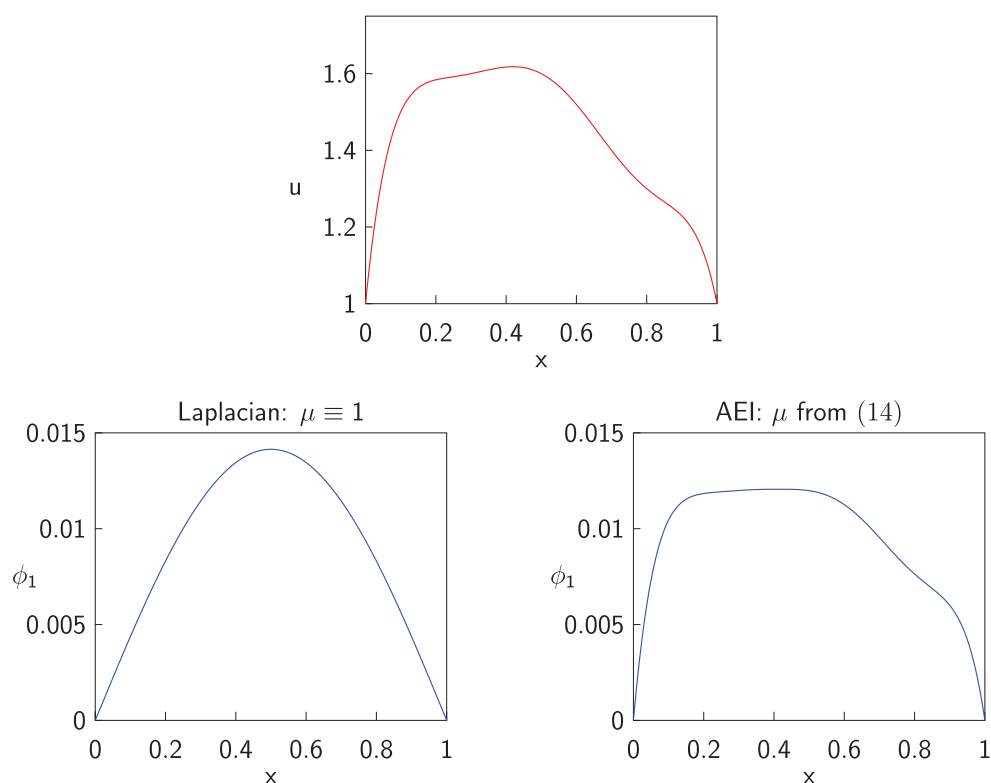


Figure 3. One-dimensional case. The true smooth profile u (top) and the first eigenfunction ϕ_1 from (15): $\mu \equiv 1$ (left), and μ as in (14) (right).

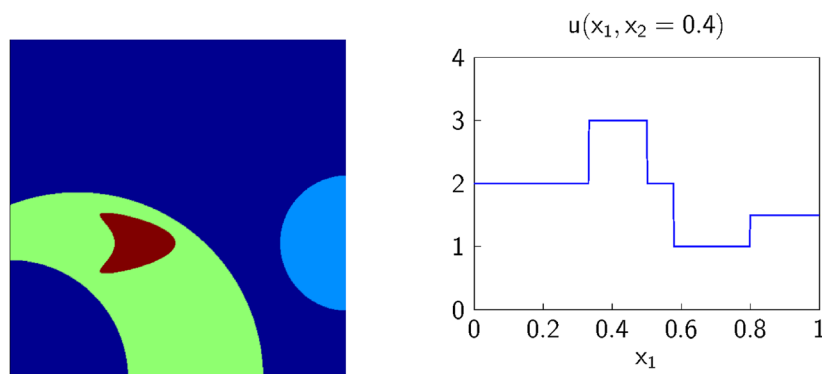


Figure 4. Two-dimensional case. Left: the exact profile u . Right: cross-cut at $x_2 = 0.4$.

Gibbs-phenomenon. These results illustrate the remarkable accuracy even of but a few eigenfunctions of the AE basis.

Remark 3. For a given profile u , the corresponding AE basis $\{u_0, \phi_1, \dots, \phi_K\}$ of relatively small dimension usually yields a remarkably accurate representation of u . At higher eigenvalues λ_m , however, the (mutually orthogonal) AE basis functions ϕ_m become increasingly

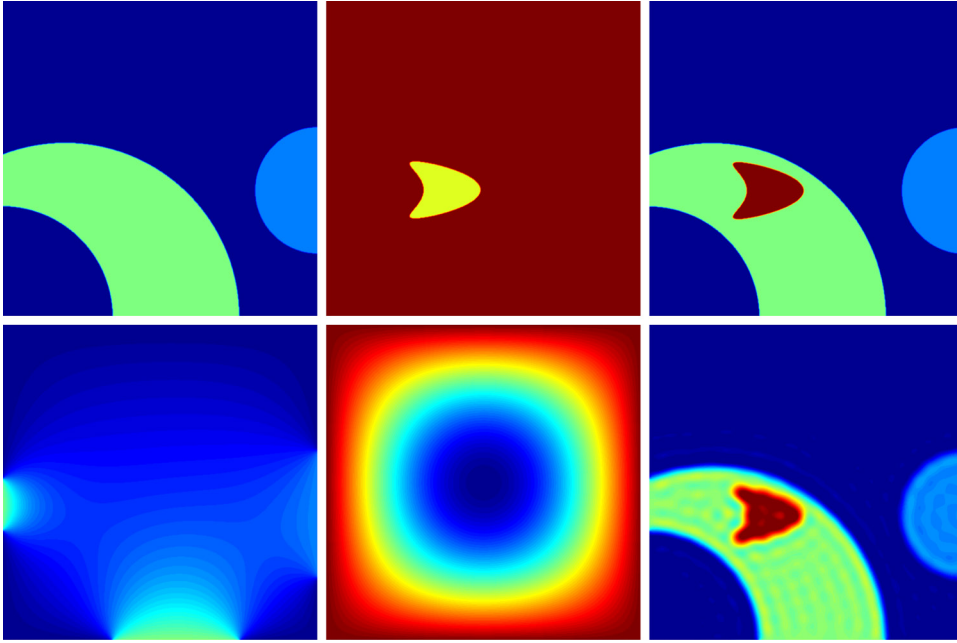


Figure 5. Two-dimensional case. Top, from left to right: the background u_0 from (9), the first AE eigenfunction ϕ_1 from (11), the L^2 best approximation of u using $\{u_0, \phi_1\}$. Bottom, from left to right: the harmonic u_0 , the first Laplace eigenfunction ϕ_1 , the L^2 best approximation of u using the first 1000 Laplace eigenfunctions $\{u_0, \phi_1, \dots, \phi_{1000}\}$.

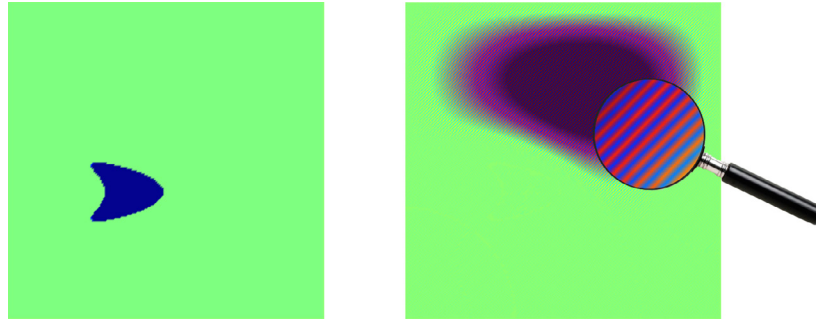


Figure 6. First AEI eigenfunction ϕ_1 (left) and last AEI eigenfunction ϕ_N (right) on a 201×201 grid.

oscillatory and no longer carry useful information about u —see figure 6. Truncating the expansion in (12) at a finite value, K , thus effectively builds regularization into the AEI approach.

4. Numerical results

We shall now illustrate the usefulness and versatility of the AEI method through a series of numerical experiments. Clearly, the squared velocity u of the medium is now unknown and we shall attempt to recover it from boundary measurements by solving (2).

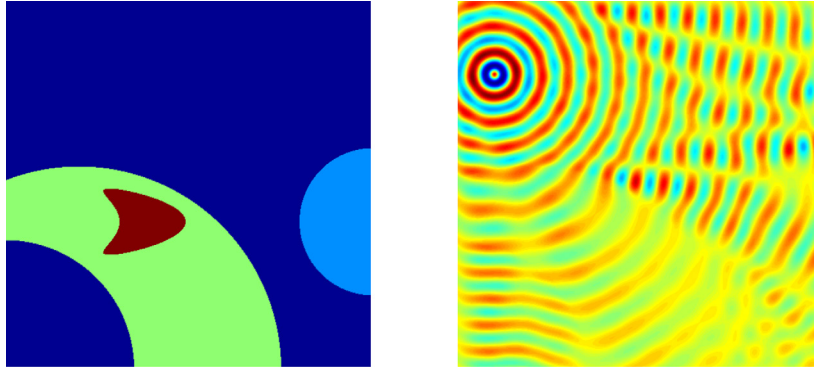


Figure 7. Left: true profile u ; right: real part of y_ℓ with f_ℓ at $(0.1, 0.8)$ and $\omega = 90$.

First, we compare the adaptive eigenspace representation to a standard grid-based nodal representation of the control u to demonstrate the resulting significant reduction in degrees of freedom. Next we add yet another level of adaptivity by solving the auxiliary elliptic eigenvalue problem with an adaptive finite element method. We also present numerical experiments that underpin our choices for the background profile u_0 and the varying dimension of V_K with increasing frequency. Then, we show that the AE basis functions ϕ_m are highly localized, and hence easily sparsified; thus, the memory requirement of the adaptive eigenspace basis is kept to a minimum. Finally, we demonstrate the robustness of the AEI approach in the presence of noise or missing data and combine it with the sample average approximation (SAA) approach [14].

In the entire section, we consider the profile u displayed in figure 7, which mimics a layered material with regions of different wave speed. Unless specified otherwise, the typical parameter settings in the numerical experiments are the following: nine equispaced Gaussian sources are located along the upper boundary at $(0.1, 0.8), \dots, (0.9, 0.8)$, whereas the receivers are located on the four lateral boundaries of $\Omega = (0, 1) \times (0, 1)$. We use second-order staggered finite differences on a 500×500 Cartesian mesh for the discretization of (2), (9)–(11). To avoid any inverse crime, the reference solution is computed on a separate finer mesh, which does not contain the coarser computational mesh.

In the AEI algorithm described in section 2.3, we always set the initial guess to $u(x) \equiv 1$. Starting at the lowest frequency $\omega = 8$, we progressively increase the frequency $\omega = 10, 12, 14, \dots, 90$. The number of eigenfunctions K starts at $K = 16$ or 32 and increases linearly with the frequency ω . For the optimization, we use a standard truncated Gauss–Newton method [13, 25] without extra regularization term. The search direction is by a truncated Conjugate-Gradient iteration with the Eisenstat–Walker stopping criterion [10, 11, 22, 24] and the step-size by a standard Armijo rule. In the definition of μ in (10), we always set $\varepsilon = 10^{-6}$.

In the following numerical experiments, we either use finite differences for the *discretize-then-optimize* approach, or finite elements for the *optimize-then-discretize* approach. We shall not emphasize any particular choice for the discretization as it did not affect the results.

4.1. Adaptive eigenspace versus nodal basis

The use of an adaptive eigenspace (AE) basis for the control variable u instead of a standard grid-based nodal basis is the distinguishing feature of the AEI method. Thus, we now compare the AE versus a standard nodal representation for the reconstruction of u . In both cases, we omit extra Tikhonov-type regularization.

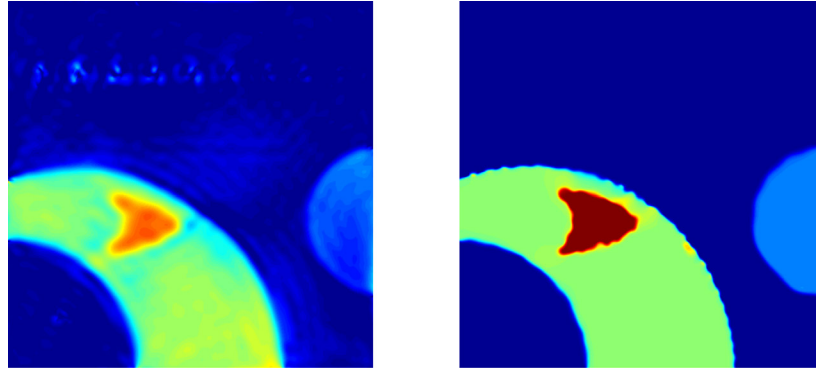


Figure 8. Adaptive eigenspace versus nodal basis. Full boundary data: reconstruction with a nodal basis (left) or an AE basis (right).

Table 1. Adaptive eigenspace versus nodal basis. Relative L^2 -error and number of degrees of freedom for u .

Basis	Full data		Partial data	
	L^2 -error (%)	N_u	L^2 -error (%)	N_u
Nodal	15.91	501 000	30.24	501 000
AEI	4.65	≤ 360	4.80	≤ 360

First, we include an additional tenth source located at $(0.15, 0.15)$. The two corresponding reconstructed profiles are shown in figure 8. Although both methods recover the essential features of the medium, the AEI method clearly yields much crisper boundaries but also higher accuracy inside the various subregions. Moreover, the AEI method achieves the higher accuracy with fewer than $N_u = 360$ degrees of freedom versus $N_u = 501\,000$ for the standard nodal representation—see table 1. Clearly, adding regularization would certainly remove some of the artifacts in the grid-based approach and thus yield a smoother, but not necessarily more accurate, reconstruction.

Next, we repeat the previous experiment but now omit all the receivers at the lower boundary of Ω together with the tenth source located at $(0.15, 0.15)$. Hence, much less information about the lower part of the medium is available in the data. Nevertheless, as shown in figure 9, the AEI method is still able to recover u everywhere inside Ω , unlike the standard grid-based approach. Indeed, as shown in table 1, the relative L^2 -error for the nodal approach has now almost doubled whereas the error for the AEI method has hardly changed. Although the AEI method uses much fewer control variables than the grid-based approach, i.e. $N_u = K$, the reconstructions appear remarkably accurate and tolerant to missing data.

4.2. Adaptive finite element discretization

The AEI method uses two separate computational meshes, the first for the forward and adjoint problems (1) and (5) and the second for the auxiliary elliptic eigenvalue problems (9) and (11). So far both meshes were spatially uniform. Here we include yet another level of adaptivity by adapting the finite element (FE) mesh in the solution of (9) and (11) to better capture small-scale features of u .

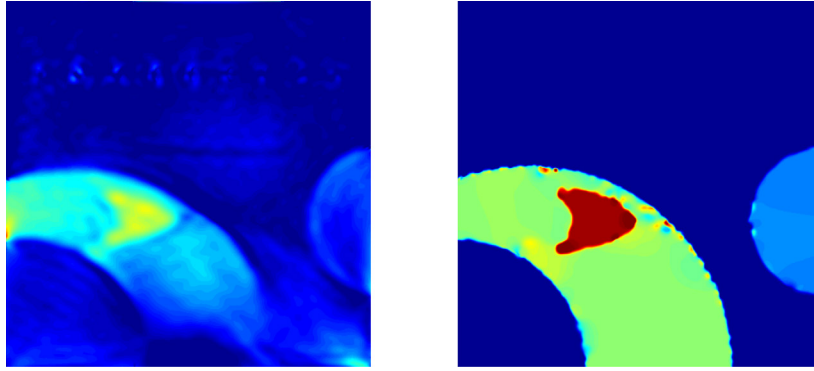


Figure 9. Adaptive eigenspace versus nodal basis. Missing boundary data: reconstruction with a nodal basis (left) or an AE basis (right).

In the AEI Algorithm, mesh adaptation is performed after the while loop in step 3b and before the update in step 3c. We use a standard *a priori* FE adaptive mesh strategy based on the Hessian of the current u [30], which is available in the open source software FreeFem++ [16]. Again, the mesh to solve the forward problem (1) and the adjoint problem (5) remains fixed and uniform (200×200 Cartesian mesh) throughout the entire computation. For the optimization, we now use a truncated Quasi-Newton (BFGS) method with Wolfe–Powell step-size control.

In figure 10, we present the numerical results for the AEI method with and without mesh adaptation. The reconstruction on a uniform triangular mesh yields 4.86% relative L^2 -error, whereas the reconstruction with adaptive FE strategy yields 4.17% relative L^2 -error, yet with a ten times smaller number of vertices. Hence, we have not only reduced the error in the reconstruction with even crisper and smoother edges, but also greatly reduced the number of degrees of freedom in the control thereby dividing the overall execution time by 2.5. As shown in figure 10, the adapted mesh automatically concentrates the degrees of freedom along variations of u and is refined only where it is needed.

4.3. Adaptive versus harmonic background u_0

As shown in section 3.2, the background state u_0 defined through (9) allows the AEI method to accommodate varying boundary data. Here we compare the accuracy in the reconstruction for two different choices for u_0 :

- harmonic: u_0 is computed only once at the beginning of the AEI Algorithm by solving (9) with $\mu \equiv 1$
- adaptive: u_0 is recomputed at each frequency step by solving (9) with μ as in (10)

In figure 11, we compare the recovered u for the above two different background states u_0 , both either with or without the adaptive mesh strategy from section 4.2. Clearly, the adaptive background u_0 improves the accuracy of the reconstruction as the relative L^2 -errors are approximately halved—see table 2. Moreover, the artifacts near the exterior boundary, visible in the harmonic background approach, are now completely absent.

Hence the more accurate background state u_0 enables the AEI method to better approximate the remainder $u - u_0$ with the AE basis. Still, the adaptive strategy is also slightly more expensive since it adaptively recomputes u_0 for every frequency. These conclusions also appear to hold if an adaptive FE strategy is included in the solution of (9) and (11), which in fact seems to have little impact on the overall accuracy here.

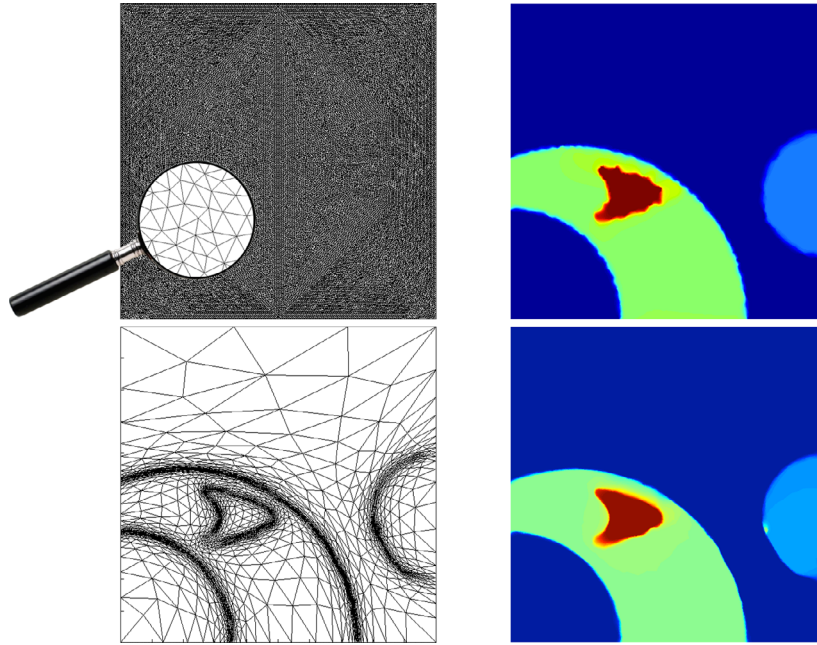


Figure 10. Adaptive FE discretization. Top: uniform triangular mesh with 30534 vertices (left) and recovered u (right). Bottom: with mesh adaptation, final mesh for $\omega = 90$ with 2783 vertices (left) and recovered u (right).

4.4. Constant versus adaptive dimension of V_K

Our AEI method does not require extra regularization term, such as standard Tikhonov regularization. Instead it restricts the optimization to the AE subspace $V_K = \text{span}\{u_0, \phi_1, \dots, \phi_K\}$, which effectively acts as inherent regularization of the inverse problem. Here we focus on the choice of K to understand the regularization effect of the AEI method. To prevent any intrinsic regularization from a very fine mesh or over-abundant data, we omit the receivers at the lower boundary of Ω (missing data) and use a coarser 200×200 Cartesian mesh.

First, we keep the number of eigenfunctions constant through the entire frequency stepping process. In figure 12, we present the numerical results for constant $K = 100$ using either a truncated full-Newton or Gauss–Newton method. For the full-Newton method, the regularization is not sufficient to reduce the perturbation at the top of the computational domain, although the reconstruction is more accurate than with the nodal basis. In contrast, for the Gauss–Newton method, the regularization is too strong and the method has difficulty reconstructing the kite. Similar conclusions hold for other constant values of K , which underlines the need for K to vary with the frequency ω .

Next, we let K vary linearly with ω , starting at a small number of eigenfunctions ($K = 32$) to reduce the risk of landing in a false local minimum. As ω increases during frequency continuation, we slowly increase K to capture smaller details of the scatterer. In figure 12, both reconstructions are now quite accurate with an L^2 -error of 6.50% for the Newton and 6.72% for the Gauss–Newton method, respectively.

Remark 4. The number of eigenfunctions K controls the regularization, similarly to the parameter α in a standard Tikhonov regularization term $\alpha R[u]$. For Tikhonov regularization, α is initially large but then gradually decreases to zero as the nonlinear iteration approaches

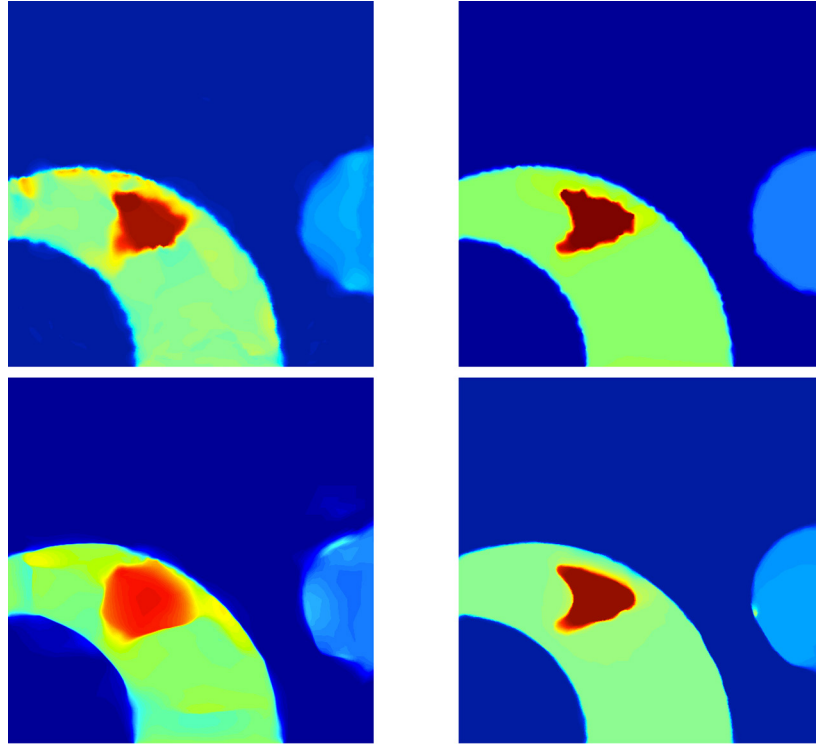


Figure 11. Adaptive versus harmonic background. Top: without mesh adaptation, harmonic u_0 (left) and adaptive u_0 (right). Bottom: with mesh adaptation, harmonic u_0 (left) and adaptive u_0 (right).

Table 2. Adaptive versus harmonic background. Relative L^2 -error in u .

Background state u_0	Harmonic (%)	Adaptive (%)
Without mesh adaptation	10.73	4.86
With mesh adaptation	10.90	4.17

the desired minimum. In contrast for our AEI approach, K is initially small but then gradually increases during optimization so that the AE basis includes a higher number of eigenfunctions for better accuracy.

4.5. Sparse AEI method

The AEI method greatly reduces the number of parameter values by restricting the optimization to the subspace $V_K = \text{span}\{u_0, \phi_1, \dots, \phi_K\}$ of much smaller dimension. Since the discretized version of the eigenvalue problem (11) leads to a sparse, symmetric and positive definite matrix, the first K eigenfunctions can be efficiently computed via a standard Lanczos iteration [21]. Still, the storage of the first K eigenfunctions, which are global functions in Ω , may at first appear quite large, especially in three space dimensions.

In fact, much of the information contained in those eigenfunctions is highly localized in space and essentially negligible in most of Ω . Again we consider the example described at the

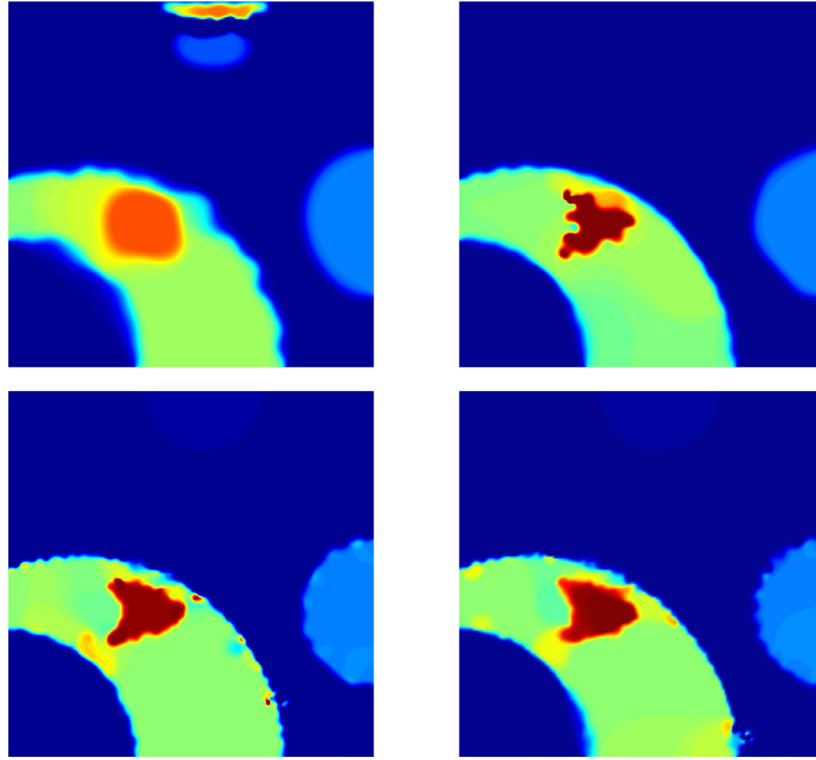


Figure 12. Constant versus adaptive dimension of V_K . Top: for constant $K = 100$, truncated full-Newton method with relative L^2 -error = 14.79% (left) and truncated Gauss-Newton method with relative L^2 -error = 7.95% (right). Bottom: for linearly varying K , truncated full-Newton method with relative L^2 -error = 6.50% (left) and truncated Gauss-Newton method with relative L^2 -error = 6.72% (right).

beginning of section 4. Next, for each eigenfunctions ϕ_m , we monitor all entries smaller than $\eta \cdot \|\phi_m\|_\infty$ in magnitude. In figure 13, we display the percentage of small entries in ϕ_m for each frequency averaged over all $m \leq K$; recall that K increases linearly with ω . As ω increases, the percentage of small entries also increases and quickly saturates above 90% at higher frequencies; hence, more than 90% of all entries are in fact negligible. To save memory space, we can therefore set to zero all those small values and simply replace the eigenfunctions ϕ_m by their sparse approximations.

In figure 14, we show the resulting reconstruction with the sparse AEI approach, where all small entries in the eigenfunctions below $\eta \|\phi_m\|_\infty$ are set to zero for $\eta = 0.1$, 0.05, or 0.01. Remarkably, the reconstruction hardly changes and, in fact, is even slightly better (L^2 -errors below 4%), while saving more than 90% of memory.

4.6. Noisy data

To illustrate the robustness of the AEI method with respect to noise, we now add multiplicative noise to the observations for each observation point x_i and frequency ω_j :

$$y^{\text{obs},\delta}(x_i, \omega_j) = y^{\text{obs}}(x_i, \omega_j)(1 + \delta \cdot \xi_{i,j}), \quad (17)$$

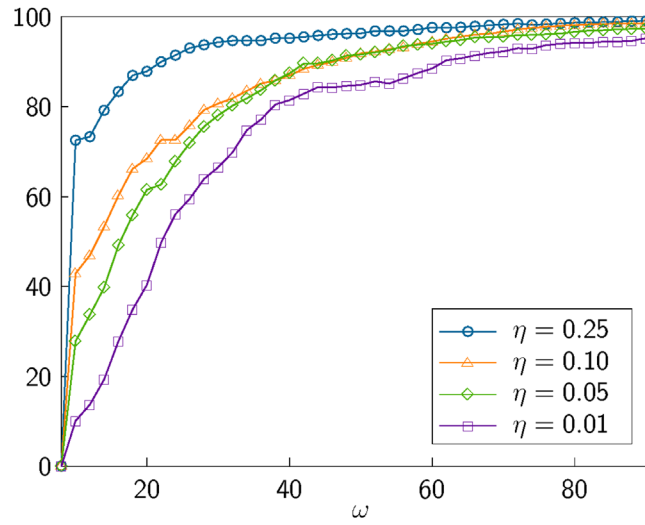


Figure 13. Sparse AEI method. Percentage of entries smaller than $\eta \cdot \|\phi_m\|_\infty$ in magnitude averaged over all $m \leq K$ at each frequency.

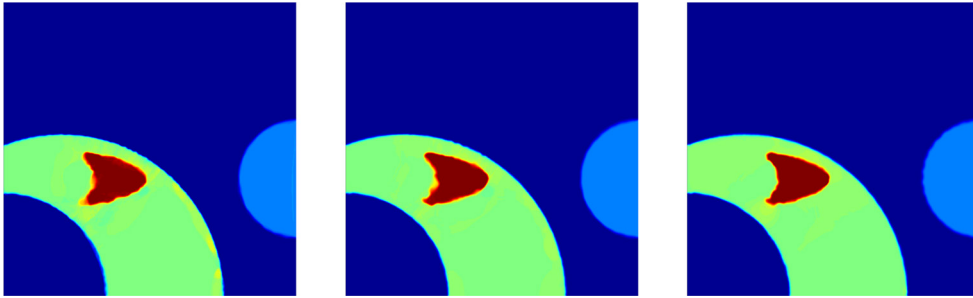


Figure 14. Sparse AEI method. Reconstruction of u with sparsified eigenfunctions. Left: with $\eta = 0.1$ (error = 3.83%). Center: $\eta = 0.05$ (error = 3.81%). Right: $\eta = 0.01$ (error = 3.90%).

where $\xi_{i,j}$ are i.i.d. Gaussian random variables with mean zero and variance equal to one. The level of noise is denoted by δ .

Again, we consider the parameter settings described at the beginning of section 4, but use the adaptive finite element strategy from section 4.2. In figure 15 we present the reconstruction results for two different levels of noise: for $\delta = 10\%$ (left) we obtain a relative L^2 -error of 4.01% and for $\delta = 20\%$ (right) we obtain a relative L^2 -error of 5.22%. As shown in figure 15, the AEI method is still able to reconstruct the profile without any added regularization and without artifacts due to noise. In fact, at the smaller noise level $\delta = 10\%$, the relative L^2 -error is even slightly better than that without noise, 4.01% versus 4.17%; each case, however, leads to (slightly) different meshes and eigenfunctions due to the adaptive finite element strategy.

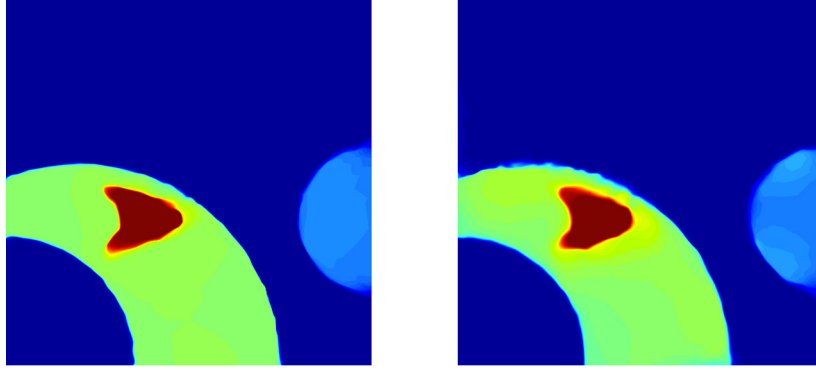


Figure 15. Noisy data for varying noise level δ . Left: $\delta = 10\%$ (L^2 -error = 4.01%). Right: $\delta = 20\%$ (L^2 -error = 5.22%).

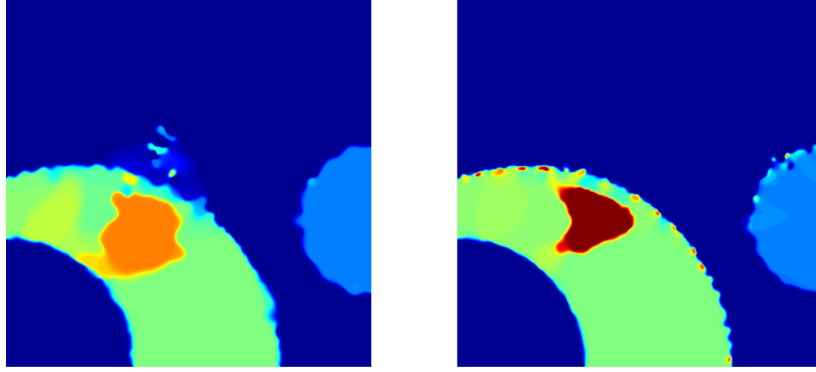


Figure 16. SAA approach. Reconstruction of u with a single source. Left: without SAA and with $N_s = 1$ (10.05% relative error). Right: with a single SAA ‘super-shot’, i.e. $N_r = 1$, from $N_s = 201$ sources (5.79% relative error). Note that the computational effort is identical.

4.7 Sample average approximation

As the number of sources N_s in (2) or (6) increases, the cost of computing N_s forward and adjoint solutions in (1) and (5) may become prohibitive. To limit the computational cost without ignoring any of the available data, we consider the sample average approximation (SAA) approach from [14], which replaces the sources f_ℓ by N_r ‘super-shots’, $N_r \ll N_s$,

$$F_j = \sum_{\ell=1}^{N_s} \xi_\ell^j f_\ell, \quad j = 1, \dots, N_r, \quad (18)$$

where the ξ_ℓ^j are i.i.d. random variables with zero mean and unit variance and corresponding observations

$$Y_j^{\text{obs}} = \sum_{\ell=1}^{N_s} \xi_\ell^j y_\ell^{\text{obs}}, \quad j = 1, \dots, N_r. \quad (19)$$

During frequency stepping, we choose for each ω a different realization of ξ_ℓ^j —here $\xi_\ell^j = \pm 1$ with probability 0.5.

Again, we consider the parameter settings described at the beginning of section 4, but now with $N_s = 201$ Gaussian sources located at $(0.1, 0.8)$, $(0.11, 0.8)$, $\dots$, $(0.89, 0.9)$, $(0.9, 0.8)$. The SAA approach [14] with only a single ‘super-shot’, $N_r = 1$, yields the reconstruction shown in figure 16. For comparison, we also display the reconstruction without SAA with a single source located at $(0.5, 0.8)$. Although the computational cost of both approaches is identical, the SAA approach yields better accuracy, remarkably so, given that only a single (‘super-shot’) source is used for the approximation.

5. Concluding remarks

We have presented a nonlinear optimization method for the solution of inverse scattering problems in the frequency domain, when the scattered field is governed by the Helmholtz equation. Instead of a standard (FD or FE) grid-based representation, the unknown (squared) sound speed u is projected to the finite-dimensional subspace $V_K = \text{span}\{u_0, \phi_1, \dots, \phi_K\}$ of much smaller dimension. The ‘background’ u_0 is determined by solving (9) whereas the remaining orthonormal basis functions ϕ_m are determined by computing the first K eigenfunctions in (11). The time-harmonic inverse medium problem is formulated as a PDE-constrained optimization problem and solved by an inexact truncated Newton or quasi-Newton iteration. During the optimization process, which includes frequency continuation, both the basis and the dimension of V_K are repeatedly adapted to the current iterate. The full adaptive eigenspace inversion (AEI) Algorithm is given in section 2.3.

For monotonic one-dimensional or layered media, we have proved that our choice for u_0 is in fact optimal. For arbitrary media, our numerical results suggest that it is clearly superior to a straightforward harmonic extension from the known boundary values. Together with but a few eigenfunctions, the adaptive eigenspace basis yields a remarkably accurate representation of u . At higher eigenvalues, the eigenfunctions become increasingly oscillatory while no longer carrying useful information about u . Hence adapting the dimension of the eigenspace basis in (12), effectively builds regularization into the inversion, so that no additional Tikhonov regularization is needed—see remark 4. As the frequency ω increases, smaller scale features of the scatterer become visible and the dimension of V_K thus ought to increase accordingly.

In contrast to a standard grid-based nodal representation, the AEI method uses much fewer control variables for u . Still the reconstructions are remarkably accurate, display less artifacts and prove more tolerant to partial or missing data. Moreover, our AEI method leads to a significant reduction in execution time and proves robust with respect to added noise. By combining it with the sample average approximation (SAA) approach from [14], it also efficiently handles large numbers of sources.

Since the discrete version of the eigenvalue problem (11) leads to a sparse, symmetric and positive definite matrix, the first K eigenfunctions can be efficiently computed via a standard Lanczos iteration. If finite element mesh adaptation is used for the numerical solution of (11), small-scale features and interfaces are captured with even greater accuracy in the reconstruction, without increasing the computational effort. Although the eigenfunctions are global, their information content is highly localized in space so that most entries are in fact negligible.

Our current work involves the extension of the AEI method to multi-parameter inverse scattering problems [28, 29, 33]. We shall report on those results elsewhere in the near future.

References

- [1] Bao G and Lin J 2011 Imaging of local surface displacement on an infinite ground plane: the multiple frequency case *SIAM J. Appl. Math.* **71** 1733–52
- [2] Beilina L and Klibanov M V 2012 *Approximate Global Convergence, Adaptivity for Coefficient Inverse Problems* (New York: Springer)
- [3] Cakoni F and Colton D L 2006 *Qualitative Methods in Inverse Scattering theory: an Introduction (Interaction of Mechanics and Mathematics Series)* (New York: Springer)
- [4] Cassier M and Hazard C 2013 Multiple scattering of acoustic waves by small sound-soft obstacles in two dimensions: mathematical justification of the Foldy–Lax model *Wave Motion* **50** 18–28
- [5] Chen Y 1997 Inverse scattering via Heisenberg’s uncertainty principle *Inverse Problems* **13** 253
- [6] Cheney M 2001 The linear sampling method and the MUSIC algorithm *Inverse Problems* **17** 591–5
- [7] Colton D and Kirsch A 1996 A simple method for solving inverse scattering problems in the resonance region *Inverse Problems* **12** 383–93
- [8] de Buhan M and Kray M 2013 A new approach to solve the inverse scattering problem for waves: combining the TRAC and the adaptive inversion methods *Inverse Problems* **29** 085009
- [9] de Buhan M and Osses A 2010 Logarithmic stability in determination of a 3D viscoelastic coefficient and a numerical example *Inverse Problems* **26** 95006
- [10] Dembo R S and Steihaug T 1983 Truncated-Newton algorithms for large-scale unconstrained optimization *Math. Program.* **26** 190–212
- [11] Eisenstat S C and Walker H F 1996 Choosing the forcing terms in an inexact Newton method *SIAM J. Sci. Comput.* **17** 16–32
- [12] Grote M J, Huber J, Kourounis D and Schenk O 2014 Inexact interior-point method for PDE-constrained nonlinear optimization *SIAM J. Sci. Comput.* **36** A1251–76
- [13] Haber E, Ascher U M and Oldenburg D 2000 On optimization techniques for solving nonlinear inverse problems *Inverse Problems* **16** 1263
- [14] Haber E, Chung M and Herrmann F 2012 An effective method for parameter estimation with PDE constraints with multiple right-hand sides *SIAM J. Optim.* **22** 739–57
- [15] Haddar H, Lechleiter A and Marmorat S 2013 An improved time domain linear sampling method for Robin and Neumann obstacles *Appl. Anal.* 1–22
- [16] Hecht F 2012 New development in FreeFem++ *J. Numer. Math.* **20** 251–65
- [17] Herzog R and Kunisch K 2010 Algorithms for PDE-constrained optimization *GAMM-Mitt.* **33** 163–76
- [18] Hinze M, Pinnau R, Ulbrich M and Ulbrich S 2009 Optimization with PDE constraints *Mathematical Modelling: Theory and Applications* vol 23, 1st edn (New York: Springer)
- [19] Kirsch A 2002 The MUSIC-algorithm and the factorization method in inverse scattering theory for inhomogeneous media *Inverse Problems* **18** 1025–40
- [20] Lechleiter A 2009 The factorization method is independent of transmission eigenvalues *Inverse Problems Imaging* **3** 123–38
- [21] Lehoucq R B and Sorensen D C 1996 Deflation techniques for an implicitly re-started arnoldi iteration *SIAM J. Matrix Anal. Appl.* **17** 789–821
- [22] Métivier L, Brossier R, Virieux J and Operto S 2013 Full waveform inversion and the truncated Newton method *SIAM J. Sci. Comput.* **35** B401–37
- [23] Mojabi P and LoVetri J 2009 Microwave biomedical imaging using the multiplicative regularized Gauss–Newton inversion *IEEE Antennas Wirel. Propag. Lett.* **8** 645–8
- [24] Nash S G 2000 A survey of truncated-Newton methods *J. Comput. Appl. Math.* **124** 45–59
- [25] Nocedal J and Wright S J 2006 *Numerical Optimization* 2nd edn (New York: Springer)
- [26] Potthast R 2006 A survey on sampling and probe methods for inverse problems *Inverse Problems* **22** R1–R47
- [27] Prada C, Manneville S, Spoliarsky D and Fink M 1996 Decomposition of the time reversal operator: application to detection and selective focusing on two scatterers *J. Acoust. Soc. Am.* **99** 2067–76
- [28] Prieux V, Brossier R, Operto S and Virieux J 2013 Multiparameter full waveform inversion of multicomponent Ocean-bottom-cable data from Valhall field. Part 1: imaging compressional wave speed, density and attenuation *Geophys. J. Int.* **194** 1640–64
- [29] Prieux V, Brossier R, Operto S and Virieux J 2013 Multiparameter full waveform inversion of multicomponent Ocean-bottom-cable data from Valhall field. Part 2: imaging compressional and shear-wave velocities *Geophys. J. Int.* **194** 1665–81

- [30] Quarteroni A 2014 *Numerical Models for Differential Problems* (MS&A vol 8) 2nd edn (New York: Springer)
- [31] Rieder A 2005 Inexact Newton regularization using conjugate gradients as inner iteration *SIAM J. Numer. Anal.* **43** 604–22
- [32] Rudin L I, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Physica D* **60** 259–68
- [33] Tarantola A 1984 Inversion of seismic reflection data in the acoustic approximation *Geophysics* **49** 1259–66
- [34] Tarantola A 2005 *Inverse Problem Theory and Methods For Model Parameter Estimation* (Philadelphia, PA: Society for Industrial and Applied Mathematics)
- [35] Van Doren J F M, Van den Hof P M J, Jansen J D and Bosgra O H 2011 Parameter identification in large-scale models for oil and gas production *Proc. 18th IFAC World Congress (Milano, Italy)* pp 10857–62
- [36] van Leeuwen T and Herrmann F J 2014 3D frequency-domain seismic inversion with controlled sloppiness *SIAM J. Sci. Comput.* **36** S192–217
- [37] Vogel C 2002 *Computational Methods for Inverse Problems* (Philadelphia, PA: Society for Industrial and Applied Mathematics)
- [38] Vogel C R and Oman E 1996 Iterative methods for total variation denoising *SIAM J. Sci. Comput.* **17** 227–38