

# Adaptive Spectral Decompositions For Inverse Medium Problems

Daniel H. Baffet<sup>\*†</sup>   Marcus J. Grote<sup>\*‡</sup>   Jet Hoe Tang<sup>\*§</sup>

June 17, 2020

## Abstract

Inverse medium problems involve the reconstruction of a spatially varying unknown medium from available observations by exploring a restricted search space of possible solutions. Standard grid-based representations are very general but all too often computationally prohibitive due to the high dimension of the search space. *Adaptive spectral (AS) decompositions* instead expand the unknown medium in a basis of eigenfunctions of a judicious elliptic operator, which depends itself on the medium. Here the AS decomposition is combined with a standard inexact Newton-type method for the solution of time-harmonic scattering problems governed by the Helmholtz equation. By repeatedly adapting both the eigenfunction basis and its dimension, the resulting adaptive spectral inversion (ASI) method substantially reduces the dimension of the search space during the nonlinear optimization. Rigorous estimates of the AS decomposition are proved for a general piecewise constant medium. Numerical results illustrate the accuracy and efficiency of the ASI method for time-harmonic inverse scattering problems, including a salt dome model from geophysics.

**Keywords:** Adaptive eigenspace inversion, AEI, full waveform inversion, inverse scattering problem, total variation regularization

## 1 Introduction

Inverse medium problems occur in a wide range of applications such as medical imaging, geophysical exploration and non-destructive testing. Given observations of a physical state variable,  $y$ , from the boundary of a bounded region  $\Omega$ , one seeks to reconstruct (unknown) spatially varying medium properties,  $u(x)$ , inside  $\Omega$ . In inverse scattering problems, for instance,  $u$  characterizes the location, shape or physical properties of the scatterer, typically a collection of bounded or penetrable inclusions, while the scattered wave field  $y$  satisfies the governing (time-dependent or time-harmonic) wave equation. In seismic imaging, in particular, full-waveform inversion reconstructs high-resolution subsurface models of the medium parameters  $u$  (e.g. spatially varying sound speed) from reflected seismic waves at the Earth's

---

<sup>\*</sup>Department of Mathematics and Computer Science, University of Basel, Basel, Switzerland

<sup>†</sup>daniel.baffet@unibas.ch

<sup>‡</sup>marcus.grote@unibas.ch

<sup>§</sup>jet-hoe.tang@univ-grenoble-alpes.fr

surface [27]. To determine  $u(x)$  from boundary measurements of  $y$ , the inverse medium problem is usually reformulated as a PDE-constrained optimization problem for a cost functional  $\mathcal{J}[u]$ , which measures the misfit between the simulated and the observed data [30, 19].

Two difficulties one must address when solving an inverse medium problem numerically are the well-posedness of the problem and the size of the set of candidate functions for  $u$ . It is well-known that the problem of minimizing the misfit over  $L^\infty(\Omega)$  is, in general, ill-posed. To obtain a well-posed problem, the misfit functional is typically modified by adding a Tikhonov-type regularization term [31, 11]. Then the resulting formulation may be discretized and solved numerically [25, 17]. Although Tikhonov regularization generally improves the stability of the inversion, it does not address the cost of solving an optimization problem in a subspace of high dimension, still determined by the number of degrees of freedom of the spatial numerical discretization.

When a parametrization of the unknown medium  $u(x)$  is explicitly known a priori, the inversion can easily be limited to a reduced set of unknown parameters. Moreover, if the search space is of sufficiently low dimension, the discretization itself can have a regularizing effect [4, 22]. In general, however, such a low-dimensional representation is not explicitly known a priori.

To resolve the vexing dilemma of remaining sufficiently general while keeping the dimension of the search space sufficiently small, various sparsity promoting strategies were proposed in recent years. In [6], Daubechies, Defrise and De Mol considered linear inverse problems yet replaced the usual quadratic regularizing penalty by weighted  $\ell_p$ -penalties on the coefficients to promote a sparse expansion of  $u$  with respect to an orthonormal basis. Later, Loris et al. [24] successfully applied  $\ell_1$ -norm soft-thresholding to promote a sparse wavelet representation of  $u$  for seismic tomography. Similarly, a truncated wavelet representation of the acoustic velocity and mass density was used for seismic FWI in [23]. In [20], a curvelet-based representation of the wave field  $y$  was used for seismic data recovery from a regularly sampled grid with traces missing. Recently, Kadu, van Leeuwen and Mulder [21] combined a level-set representation, constructed from radial basis functions, with a Gauss-Newton approximation to regularize FWI in the presence of salt bodies.

In [9], de Buhan and Osses proposed to restrict the search space to the span of a small basis of eigenfunctions of a judicious elliptic operator, repeatedly adapted during the nonlinear iteration. Their approach relies on a decomposition

$$w = \varphi_0 + \sum_{k=1}^{\infty} \beta_k \varphi_k, \quad (1.1)$$

for functions  $w \in W^{1,\infty}(\Omega)$ . Here  $\varphi_0$  satisfies the elliptic boundary-value problem

$$L_\varepsilon[w]\varphi_0 = 0 \quad \text{in } \Omega, \quad \varphi_0 = w \quad \text{on } \partial\Omega, \quad (1.2)$$

and for  $k \geq 1$ , each  $\varphi_k$  is an eigenfunction of a  $w$ -dependent, linear, symmetric, and elliptic operator  $L_\varepsilon[w]$ , that is,  $\varphi_k$  satisfies

$$L_\varepsilon[w]\varphi_k = \lambda_k \varphi_k \quad \text{in } \Omega, \quad \varphi_k = 0 \quad \text{on } \partial\Omega, \quad (1.3)$$

for an eigenvalue  $\lambda_k \in \mathbb{R}$ . The eigenvalues  $(\lambda_k)_{k \geq 1}$  form a nondecreasing sequence, with each eigenvalue repeated according to its multiplicity, and the set  $(\varphi_k)_{k \geq 1}$  is an orthonormal

basis of  $L^2(\Omega)$ . Henceforth we shall refer to (1.1) as the *adaptive spectral (AS) expansion or decomposition*.

The AS decomposition has been used in various iterative Newton-like inversion algorithms [8, 16, 18] as follows: Given an approximation of the medium,  $u^{(m-1)}$ , from the previous iteration, the approximation  $u^{(m)}$  at the current iteration is set as the minimizer of the misfit functional  $\mathcal{J}[u]$  in the affine space  $\varphi_0 + \text{span}(\varphi_k)_{k=1}^K$ , where  $\varphi_0$  and  $\varphi_k$ ,  $k = 1, \dots, K$ , satisfy (1.2) and (1.3) with  $w = u^{(m-1)}$ , respectively. As the approximation  $u^{(m)}$  changes from one iteration to the next, so does the affine search space used for the subsequent minimization.

Clearly, the choice of  $L_\varepsilon[w]$  is crucial in obtaining an efficient approximation of the medium with as few basis functions as possible. The particular choice for  $L_\varepsilon[w]$  used in [9], which essentially coincides with the linearization of the gradient of the penalized total variation (TV) functional [16], yields a remarkably efficient approximation for piecewise constant functions. Although  $L_\varepsilon[u]$  is not well-defined for piecewise constant  $u$ , numerical discretizations of (1.2) or (1.3) can be viewed as approximations with  $w$  replaced by a more regular approximation of  $u$ , such as a projection of  $u$  on an  $H^1$ -conforming finite element (FE) space. The AS decomposition also bears a striking similarity to nonlinear eigenproblems for the (subdifferential of the) TV functional and, more generally, for one homogeneous functionals used in image processing – see [3, 5, 1] and the references therein. Depending on a priori available information about the smoothness or spatial anisotropy of the medium, different choices for  $L_\varepsilon[w]$  will yield more or less efficient AS representations [18].

By combining the adaptive inversion process with the TRAC (time reversed absorbing condition) approach, de Buhan and Kray [8] developed an effective solution strategy for time-dependent inverse scattering problems. In [16], Grote, Kray and Nahum proposed the AEI (adaptive eigenspace inversion) algorithm for inverse scattering problems in the frequency domain. It combines the AS decomposition with frequency stepping and truncated inexact Newton-like methods [28, 25], and may also be used without Tikhonov regularization by progressively increasing the dimension  $K$  of the spectral basis with frequency. In [18], the AEI algorithm was extended to multi-parameter inverse medium problems, including the well-known layered Marmousi subsurface model from geosciences. Recently, it was extended to electromagnetic inverse scattering problems at fixed frequency [7] and also to time-dependent inverse scattering problems when the illuminating source is unknown [15].

So far, the remarkable efficiency and accuracy of the AS decomposition for the approximation of piecewise constant functions is only justified via numerical evidence. Although previous inversion algorithms based on the AS decomposition iteratively adapt the basis functions  $\varphi_k$ , they do not provide any criteria for adapting the dimension  $K$  of the search space. Here, we precisely address these two open questions. In Section 2, we present a strategy for adapting the dimension  $K$  of the search space, by solving a small quadratically constrained quadratic minimization problem to filter basis functions while preserving important features. The resulting new ASI (adaptive spectral inversion) algorithm is listed in Section 2.2. In Section 3, we derive rigorous error estimates for  $\varphi_0$  and for the first eigenvalues and eigenfunctions of the elliptic operator associated with a piecewise constant function. In particular, we prove that  $\varphi_0$  and the first eigenfunctions  $\varphi_k$  are “almost” piecewise constant in the sense that their gradients are small outside a neighborhood of internal discontinuities. We also provide a numerical example which illustrates the theory. In Section 4, we apply the ASI Algorithm to two inverse medium problems, the first where the medium is composed of five simple geometric inclusions, and the second, where the medium corresponds to a two-dimensional model of a salt dome

from geophysics. Finally, we conclude with some remarks in Section 5.

## 2 Inverse scattering problem

First, we consider a time-harmonic inverse scattering problem and reformulate it as a PDE-constrained optimization problem. Then we present the adaptive spectral inversion (ASI) method, list the full ASI Algorithm and discuss in further detail the individual steps in adapting both the basis and its dimension during the nonlinear iteration.

### 2.1 Inverse scattering problem

We consider a time-harmonic inverse scattering problem where the scattered wave field,  $y(x)$ , satisfies the Helmholtz equation:

$$-\nabla \cdot (u(x)\nabla y(x)) - \omega^2 y(x) = f(x), \quad x \in \Omega, \quad (2.1a)$$

$$\frac{\partial y}{\partial n}(x) - i\frac{\omega}{\sqrt{u(x)}} y(x) = g(x), \quad x \in \partial\Omega. \quad (2.1b)$$

Here,  $\Omega \subset \mathbb{R}^d$  is a bounded domain with Lipschitz boundary  $\partial\Omega$  and outward unit normal  $n$ ,  $f \in L^2(\Omega)$  and  $g \in L^2(\partial\Omega)$  are known sources, and  $\omega = 2\pi\nu$  is the angular frequency corresponding to the (regular) frequency  $\nu > 0$ . The squared wave speed,  $u(x)$ , of the (unknown) medium satisfies  $u(x) \geq u_{\min} > 0$  throughout  $\Omega$  and is assumed known on the boundary  $\partial\Omega$ .

Given observations  $y_\ell^{\text{obs}}$  on a subset  $\Gamma$  of  $\partial\Omega$  of the scattered fields,  $y_\ell = y_\ell[u]$ , due to known sources  $f = f_\ell$  and  $g = g_\ell$ ,  $\ell = 1, \dots, N_s$ , we wish to recover the medium  $u(x)$  inside  $\Omega$  by minimizing the cost functional  $\mathcal{J} : L^\infty(\Omega) \rightarrow \mathbb{R}$ ,

$$\mathcal{J}[w] = \frac{1}{2} \sum_{\ell=1}^{N_s} \|y_\ell[w] - y_\ell^{\text{obs}}\|_{L^2(\Gamma)}^2. \quad (2.2)$$

Hence, we consider the PDE-constrained optimization problem,

$$u_* = \underset{w \in W}{\operatorname{argmin}} \mathcal{J}[w], \quad (2.3)$$

where  $W$  is a space of candidate media and  $y_\ell$  solves (2.1) with  $f = f_\ell$  and  $g = g_\ell$ . We regard  $W$  as a subspace of  $L^2(\Omega)$  with its standard inner product and norm, denoted by  $\langle \cdot, \cdot \rangle$  and  $\|\cdot\|$ , respectively.

### 2.2 Adaptive spectral inversion

The adaptive spectral inversion (ASI) algorithm minimizes  $\mathcal{J}$  in (2.2) over a finite-dimensional subspace of  $W$  by building and repeatedly adapting a finite basis as follows. Suppose we have, at iteration  $m$ , a function  $\varphi_0^{(m)} \in W$ , which coincides with the true medium  $u$  on the boundary, and a subspace  $\Psi^{(m)} \subset W$  spanned by an orthonormal set

$$\Psi^{(m)} = \operatorname{span}\{\psi_1^{(m)}, \dots, \psi_{J_m}^{(m)}\} \quad (2.4)$$

of functions that vanish on  $\partial\Omega$ . Then, we determine a (local) minimizer  $u^{(m)}$  of  $\mathcal{J}$  in the  $J_m$ -dimensional affine space  $\varphi_0^{(m)} + \Psi^{(m)}$ , i.e.,

$$u^{(m)} = \underset{v \in \varphi_0^{(m)} + \Psi^{(m)}}{\operatorname{argmin}} \mathcal{J}[v], \quad (2.5)$$

using any standard Newton or quasi-Newton method; the smaller  $J_m$ , the cheaper the numerical solution of the nonlinear optimization problem (2.5).

Once we have determined  $u^{(m)}$ , we must update the search space needed at the next iteration, that is, determine  $\varphi_0^{(m+1)} \in W$  and  $\Psi^{(m+1)} \subset W$ . First, we compute  $\varphi_0^{(m+1)}$ , by solving

$$L_\varepsilon[u^{(m)}]\varphi_0^{(m+1)} = 0 \quad \text{in } \Omega, \quad \varphi_0^{(m+1)} = u \quad \text{on } \partial\Omega, \quad (2.6)$$

where  $L_\varepsilon[w]$  is a  $w$ -dependent, symmetric and elliptic operator; the particular form used here is specified below, but other choices are possible – see Section 3 and [18]. Recall that  $u$  is assumed known on the boundary  $\partial\Omega$ . Then, we compute the first  $K_{m+1}$  (orthonormal) eigenfunctions  $\varphi_1^{(m+1)}, \dots, \varphi_{K_{m+1}}^{(m+1)}$  of  $L_\varepsilon[u^{(m)} - \varphi_0^{(m+1)}]$  (or possibly  $L_\varepsilon[u^{(m)}]$ , cf. [16, 18]) by solving the eigenvalue problem

$$L_\varepsilon[u^{(m)} - \varphi_0^{(m+1)}]\varphi_k^{(m+1)} = \lambda_k \varphi_k^{(m+1)} \quad \text{in } \Omega, \quad \varphi_k^{(m+1)} = 0 \quad \text{on } \partial\Omega, \quad (2.7)$$

for the smallest, not necessarily distinct eigenvalues  $0 < \lambda_1 \leq \dots \leq \lambda_{K_{m+1}}$ . The eigenfunctions  $\varphi_1^{(m+1)}, \dots, \varphi_{K_{m+1}}^{(m+1)}$  should enable an efficient representation of the current iterate  $u^{(m)}$  but also enrich the search space by introducing new potential candidate media for further minimizing  $\mathcal{J}$  at the next iteration.

Next, we merge the current search space  $\Psi^{(m)}$  with the space spanned by the new eigenfunctions as

$$\tilde{\Psi}^{(m+1)} = \Psi^{(m)} + \operatorname{span} \{ \varphi_0^{(m+1)} - \varphi_0^{(m)}, \varphi_1^{(m+1)}, \dots, \varphi_{K_{m+1}}^{(m+1)} \} \quad (2.8)$$

to ensure that the current solution  $u^{(m)} \in \varphi_0^{(m)} + \Psi^{(m)}$  also belongs to  $\varphi_0^{(m+1)} + \tilde{\Psi}^{(m+1)}$ . By a standard Gram-Schmidt procedure, we compute an  $L^2$ -orthonormal basis for the merged space,

$$\tilde{\Psi}^{(m+1)} = \operatorname{span} \{ \psi_1^{(m+1)}, \dots, \psi_{J_{m+1}}^{(m+1)} \}.$$

In the process of adapting the search space, it is crucial to also adapt its new dimension, which not only directly impacts the computational cost but also acts as regularization. Hence, we shall reduce  $\tilde{\Psi}^{(m+1)}$  and thus obtain the new search space  $\Psi^{(m+1)} \subset \tilde{\Psi}^{(m+1)}$  of dimension  $J_{m+1}$  while imposing some regularity on the solution. To decide which basis functions of  $\tilde{\Psi}^{(m+1)}$  to keep and which to discard, we compute an indicator  $\tilde{u}^{(m+1)} \in \tilde{\Psi}^{(m+1)}$  as a filtered approximation of  $u^{(m)} - \varphi_0^{(m+1)}$  in  $\tilde{\Psi}^{(m+1)}$ . Then, we reorder the basis functions  $\psi_k^{(m+1)}$  in decreasing order of its Fourier coefficients

$$\gamma_k = \left\langle \tilde{u}^{(m+1)}, \psi_k^{(m+1)} \right\rangle, \quad (2.9)$$

and discard any  $\psi_k^{(m+1)}$  associated with small  $\gamma_k$ . This yields the new search space  $\Psi^{(m+1)}$  spanned by the remaining  $J_{m+1}$  basis functions,

$$\Psi^{(m+1)} = \operatorname{span} \{ \psi_1^{(m+1)}, \dots, \psi_{J_{m+1}}^{(m+1)} \}. \quad (2.10)$$

**Algorithm 1:** ASI Algorithm

---

**Input:** initial guess  $u^{(0)}$  of  $u$ , affine space  $\varphi_0^{(1)} + \Psi^{(1)}$   
**Output:** reconstruction  $u^{(m)}$  of  $u$

- 1 **for**  $m = 1, 2, \dots$  **do**
- 2     Determine (local) minimizer  $u^{(m)}$  of  $\mathcal{J}[v]$  in the  $J_m$ -dimensional affine space  $\varphi_0^{(m)} + \Psi^{(m)}$  by solving (2.5).
- 3     **if**  $\|u^{(m)} - u^{(m-1)}\| < \varepsilon_{tol}$  **then**
- 4         **return**  $u^{(m)}$
- 5     Compute  $\varphi_0^{(m+1)}$  and the first eigenfunctions  $\varphi_1^{(m+1)}, \dots, \varphi_{K_{m+1}}^{(m+1)}$  as in (2.6), (2.7) for the current medium  $u^{(m)}$ .
- 6     *Merge:* Compute an orthonormal basis  $\{\psi_k^{(m+1)}\}_{k=1}^{\tilde{J}_{m+1}}$  for the merged space  $\tilde{\Psi}^{(m+1)}$  given by (2.8).
- 7     *Filter:* Compute the indicator  $\tilde{u}^{(m+1)}$  as a filtered (regularized) approximation of  $u^{(m)} - \varphi_0^{(m+1)}$  in  $\tilde{\Psi}^{(m+1)}$ .
- 8     *Truncate:* Reduce the dimension of  $\tilde{\Psi}^{(m+1)}$  by discarding those  $\psi_i^{(m+1)}$  with smallest Fourier coefficients (2.9). This yields  $\Psi^{(m+1)}$  of dimension  $J_{m+1}$ .

---

Below we summarize the entire adaptive spectral inversion (ASI) Algorithm:

In the above ASI Algorithm, most of the computational work occurs in Step 2, which involves multiple numerical solutions of the (forward) problem (2.1). For (2.6) and (2.7) in Step 5, we always use  $L_\varepsilon[w]$  of the form

$$L_\varepsilon[w]v = -\nabla \cdot (\mu_\varepsilon[w] \nabla v), \quad (2.11)$$

where the  $w$ -dependent weight function  $\mu_\varepsilon[w]$  is given by

$$\mu_\varepsilon[w] = \frac{1}{\sqrt{|\nabla w|^2 + \varepsilon^2}}, \quad (2.12)$$

with  $\varepsilon > 0$  a small parameter; in our computations, we always set  $\varepsilon = 10^{-8}$ . Note that the theoretical properties of the adaptive spectral decomposition proved in Section 3 also apply to more general  $\mu_\varepsilon[w]$ . We now describe in more detail the individual steps of the above ASI Algorithm.

To initialize the ASI Algorithm, we require an initial guess  $u^{(0)}$ , an approximation of the background  $\varphi_0^{(1)}$  that coincides with the medium on the boundary, and a finite dimensional space  $\Psi^{(1)}$  given as the span of an orthonormal basis

$$\Psi^{(1)} = \text{span} \{\psi_1^{(1)}, \dots, \psi_{J_1}^{(1)}\}.$$

The initial background  $\varphi_0^{(1)}$  and orthonormal basis of  $\Psi^{(1)}$  can be determined by solving (2.6) and (2.7) with  $m = 0$  and setting  $\psi_k^{(1)} = \varphi_k^{(1)}$ . For instance, if  $u^{(0)}$  is constant throughout

$\Omega$  and equal to  $u$  on the boundary, then  $\varphi_0^{(1)} = u^{(0)}$  whereas  $\varphi_1^{(1)}, \varphi_2^{(1)}, \dots$  correspond to the first eigenfunctions of the Laplacian. Alternatively, one may choose  $\varphi_0^{(1)}$  or the basis of  $\Psi^{(1)}$  a priori, independently of  $u^{(0)}$ .

For the minimization of (2.5) in Step 2, we opt for a quasi-Newton method with rank-two updates (BFGS, [26]), but other choices are possible. As initial guess for  $m > 1$ , one can use the  $L^2$ -projection onto  $\varphi_0^{(m)} + \Psi^{(m)}$  either of  $u^{(m-1)}$  or its filtered approximation  $\varphi_0^{(m)} + \tilde{u}^{(m)}$ .

In Step 5, the number of new eigenfunctions  $K_{m+1}$  is somewhat arbitrary, since the dimension of the new basis is anyway adapted subsequently; here, we typically take  $K_{m+1} = J_m$  with  $J_1 = 100$ . To obtain the  $L^2$ -orthonormal basis  $\{\psi_i^{(m+1)}\}_{i=1}^{\tilde{J}_{m+1}}$  for  $\tilde{\Psi}^{(m+1)}$  in Step 6, we apply the standard modified Gram-Schmidt algorithm to the ordered set

$$\left( \varphi_1^{(m+1)}, \dots, \varphi_{K_{m+1}}^{(m+1)}, \varphi_0^{(m+1)} - \varphi_0^{(m)}, \psi_1^{(m)}, \dots, \psi_{J_m}^{(m)} \right)$$

spanning  $\tilde{\Psi}^{(m+1)}$ .

In Step 7, we compute the indicator  $\tilde{u}^{(m+1)}$  used subsequently for truncating the space  $\tilde{\Psi}^{(m+1)}$  as follows. The basis of the truncated new search space  $\Psi^{(m+1)}$  ought to preserve edges in the medium but also suppress noise. Hence, we seek an indicator which is close to  $u^{(m)} - \varphi_0^{(m+1)}$ , yet with minimal TV, by considering the constrained minimization problem

$$\begin{aligned} \min_{v \in \tilde{\Psi}^{(m+1)}} \int_{\Omega} |\nabla v|, \quad \text{subject to} \\ \|v - (u^{(m)} - \varphi_0^{(m+1)})\|^2 \leq \varepsilon_{\Psi}^2 \|u^{(m)} - \varphi_0^{(m+1)}\|^2, \end{aligned} \quad (2.13)$$

for a prescribed tolerance  $\varepsilon_{\Psi} > 0$ . Since the solution of (2.13) is expensive, we replace the TV functional

$$\int_{\Omega} |\nabla v|,$$

by the approximation [16, 18]

$$\int_{\Omega} \mu_{\varepsilon}[u^{(m)} - \varphi_0^{(m+1)}] |\nabla v|^2. \quad (2.14)$$

Hence, we compute the indicator  $\tilde{u}^{(m+1)}$  by solving the quadratically constrained quadratic minimization problem for fixed  $w = u^{(m)} - \varphi_0^{(m+1)}$ ,

$$\begin{aligned} \tilde{u}^{(m+1)} = \operatorname{argmin}_{v \in \tilde{\Psi}^{(m+1)}} \int_{\Omega} \mu_{\varepsilon}[w] |\nabla v|^2, \quad \text{subject to} \\ \|v - w\|^2 \leq \varepsilon_{\Psi}^2 \|w\|^2. \end{aligned} \quad (2.15)$$

which is cheap.

Once  $\tilde{u}^{(m+1)}$  has been computed, we truncate  $\tilde{\Psi}^{(m+1)}$  in Step 8 as follows. Given the Fourier coefficients  $\gamma_{\ell}$  of  $\tilde{u}^{(m+1)}$  arranged in decreasing order, with  $\gamma_{\ell}$  as in (2.9), we remove the maximal number of basis functions with corresponding smallest Fourier coefficients, setting

$$J_{m+1}^0 = \min \left\{ J \in \{1, \dots, \tilde{J}_{m+1}\} : \sum_{\ell=J+1}^{\tilde{J}_{m+1}} \gamma_{\ell}^2 \leq \varepsilon_{\Psi}^2 |\gamma|^2 \right\}, \quad \rho = \frac{J_{m+1}^0}{J_m} \quad (2.16)$$

for a given tolerance  $\varepsilon_\Psi$ . To avoid abrupt changes in the number of basis functions from one iteration to the next, we accept the value  $J_{m+1} = J_{m+1}^0$  without further change if

$$\rho \in [\rho_0, \rho_1], \quad (2.17)$$

for some prescribed  $0 \leq \rho_0 \leq 1 \leq \rho_1$ . Otherwise if  $\rho > \rho_1$ , we still set  $J_{m+1} = J_{m+1}^0$  but also increase  $\varepsilon_\Psi$ , e.g., by doubling it, to avoid such an overly large increase in dimension at the next iteration. On the other hand if  $\rho < \rho_0$ , we set  $J_{m+1} = \lceil \rho_0 J_m \rceil$  to avoid an overly rapid drop in dimension and thus lose important information; moreover, we decrease  $\varepsilon_\Psi$  in (2.16), e.g., by halving it. Typically, we choose  $\rho_0 = 0.9$ ,  $\rho_1 = 1.1$ , and  $\varepsilon_\Psi = 10^{-3}$ .

Finally, we embed the ASI algorithm in a standard frequency stepping approach, where the inverse problem is solved at increasingly higher frequencies  $\nu$ . Whenever two consecutive iterates satisfy

$$\|u^{(m)} - u^{(m-1)}\| < \varepsilon_\nu, \quad (2.18)$$

for a given tolerance  $\varepsilon_\nu$ , we proceed to the observations obtained at a higher frequency.

### 3 Analysis of adaptive spectral decompositions

In this section we derive estimates for  $\varphi_0$  and the eigenvalues  $(\lambda_k)_k$  and eigenfunctions  $(\varphi_k)_k$  of the linear operator  $L_\varepsilon[u_\delta]$  defined in (2.11) with  $u_\delta$  an  $H^1$ -regular approximation of a given piecewise constant function  $u$ . Typically, the medium-dependent weight function  $\mu_\varepsilon[w]$  which characterizes the AS decomposition has either the form

$$\mu_\varepsilon[w](x) = \frac{1}{(|\nabla w(x)|^q + \varepsilon^q)^{1/q}}, \quad (3.1)$$

for some  $q \in [1, \infty)$ , or

$$\mu_\varepsilon[w](x) = \frac{1}{\max(|\nabla w(x)|, \varepsilon)}. \quad (3.2)$$

For the analysis below, we allow  $\mu_\varepsilon[w]$  to have a more general form.

#### 3.1 Assumptions and definitions

##### 3.1.1 Medium dependent weight function

Let  $\Omega \subset \mathbb{R}^d$  be a bounded and connected Lipschitz domain. We assume that for  $w : \Omega \rightarrow \mathbb{R}$ , with  $\nabla w \in L^\infty(\Omega)$ ,  $\mu_\varepsilon[w]$  is given by

$$\mu_\varepsilon[w](x) = \hat{\mu}_\varepsilon(|\nabla w(x)|), \quad x \in \Omega, \quad (3.3)$$

where  $\hat{\mu}_\varepsilon : [0, \infty) \rightarrow \mathbb{R}$  is a non-increasing function that satisfies

$$\hat{\mu}_\varepsilon(0) = \varepsilon^{-1}, \quad 0 < \hat{\mu}_\varepsilon(t), \quad t\hat{\mu}_\varepsilon(t) \leq 1, \quad t \geq 0. \quad (3.4)$$

In particular, these assumptions yield

$$\mu_\varepsilon[w](x)|\nabla w(x)| \leq 1, \quad \text{a.e. } x \in \Omega, \quad (3.5)$$

and

$$0 < \hat{\mu}_\varepsilon(\|\nabla w\|_{L^\infty(\Omega)}) \leq \mu_\varepsilon[w](x) \quad \text{a.e. } x \in \Omega. \quad (3.6)$$

Although this framework encompasses weight functions as in (3.1) or (3.2), it does not include weight functions, associated with the Lorentzian or Gaussian penalty terms [18], for instance, which do not satisfy the last inequality of (3.4). However, the above framework easily extends to more general weight functions that satisfy an estimate

$$t^r \hat{\mu}_\varepsilon(t) \leq 1, \quad t \geq 0, \quad (3.7)$$

for some  $r \in [1, 2]$ . This extension, indeed, addresses the Lorentzian and Gaussian weight functions which satisfy (3.7) with  $r = 2$ . For the analysis, we assume (3.4), i.e., (3.7) with  $r = 1$ , and comment on the extension to the more general case  $r \in [1, 2]$  in Remark 8 below.

### 3.1.2 Elliptic boundary value problems

To include FE approximations in the analysis, we formulate the boundary value problems (1.2) and (1.3) in closed subspaces  $\mathcal{V}^\delta \subset H^1(\Omega)$  and  $\mathcal{V}_0^\delta = \mathcal{V}^\delta \cap H_0^1(\Omega)$ , respectively. We say that  $\varphi_0 \in \mathcal{V}^\delta$  satisfies

$$L_\varepsilon[u_\delta]\varphi_0 = 0 \quad \text{in } \Omega, \quad \varphi_0 = u_\delta \quad \text{on } \partial\Omega \quad (3.8)$$

in  $\mathcal{V}_0^\delta$ , if

$$B[\varphi_0, v] = 0, \quad \forall v \in \mathcal{V}_0^\delta \quad (3.9a)$$

$$\varphi_0 = u_\delta, \quad \text{on } \partial\Omega, \quad (3.9b)$$

where  $B[\cdot, \cdot]$  is the bilinear form given by

$$B[v, w] = \langle \mu_\varepsilon[u_\delta]\nabla v, \nabla w \rangle, \quad (3.10)$$

with  $\langle \cdot, \cdot \rangle$  denoting the standard  $L^2(\Omega)$  inner product. We say that  $\lambda \in \mathbb{R}$  is an eigenvalue of  $L_\varepsilon[u_\delta]$  in  $\mathcal{V}_0^\delta$ , if there exists  $0 \neq \varphi \in \mathcal{V}_0^\delta$  such that

$$L_\varepsilon[u_\delta]\varphi = \lambda\varphi \quad \text{in } \Omega, \quad \varphi = 0 \quad \text{on } \partial\Omega, \quad (3.11)$$

in  $\mathcal{V}_0^\delta$ , that is, if

$$B[\varphi, v] = \lambda \langle \varphi, v \rangle, \quad \forall v \in \mathcal{V}_0^\delta. \quad (3.12)$$

As shown below, the operator  $L_\varepsilon[u_\delta]$  is symmetric and uniformly elliptic. Thus, we let  $(\lambda_k)$  be the nondecreasing (perhaps finite) sequence of (real positive) eigenvalues of  $L_\varepsilon[u_\delta]$  in  $\mathcal{V}_0^\delta$ , with each eigenvalue repeated according to its multiplicity, and  $(\varphi_k)$  be a basis of  $\mathcal{V}_0^\delta$  of corresponding eigenfunctions, orthonormal with respect to the  $L^2$  inner product. Finally, we denote by  $\|\cdot\|$  the norm of  $L^2(\Omega)$ , and by  $|\cdot|$  the  $\ell^2$ -norm.

### 3.1.3 Piecewise constant medium

Let the medium  $u : \Omega \rightarrow \mathbb{R}$  be piecewise constant admitting a decomposition

$$u(x) = u^0(x) + \tilde{u}(x), \quad x \in \bar{\Omega} \quad (3.13)$$

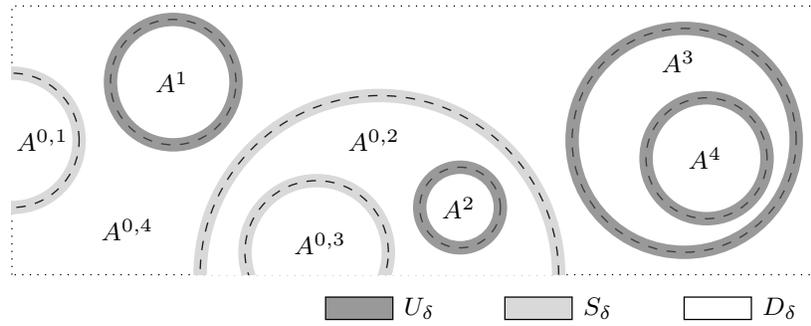


Figure 1: Typical arrangement in two dimensions of the sets  $A^k$ ,  $A^{0,m}$  composing the piecewise constant medium  $u$ , and  $U_\delta$ ,  $S_\delta$ , and  $D_\delta$  given by (3.19), (3.18), and (3.20), respectively.

into a piecewise constant background  $u^0$ , composed of features reaching the boundary  $\partial\Omega$ , and a piecewise constant perturbation  $\tilde{u}$ , composed of a finite number of inclusions separated from the boundary. More precisely, we assume

$$u^0 = \sum_{m=1}^M \hat{u}^{0,m} \chi^{0,m} \quad (3.14)$$

where  $\hat{u}^{0,m} \in \mathbb{R}$ , and  $\chi^{0,1}, \dots, \chi^{0,M}$  are the characteristic functions of mutually disjoint connected open sets  $A^{0,1}, \dots, A^{0,M}$  such that

$$\overline{\Omega} \subset \overline{\bigcup_{m=1}^M A^{0,m}}, \quad \mathcal{H}^{d-1}(\partial A^{0,m} \cap \partial\Omega) > 0, \quad (3.15)$$

where  $\mathcal{H}^{d-1}$  denotes the  $(d-1)$ -dimensional Hausdorff measure. Although the coefficients  $\hat{u}^{0,k}$  must satisfy  $\hat{u}^{0,k} > 0$  for the Helmholtz equation, they may take any real value in the present analysis. We further assume that

$$\tilde{u} = \sum_{k=1}^K \hat{u}^k \chi^k, \quad \hat{u}^k \neq 0, \quad (3.16)$$

where for each  $k = 1, \dots, K$ ,  $\chi^k$  is the characteristic function of a connected open set  $A^k \subset\subset \Omega \setminus S$ , where

$$S = \left( \bigcup_{m=1}^M \partial A^{0,m} \right) \setminus \partial\Omega \quad (3.17)$$

is the set of jump discontinuities, or interfaces, in the background  $u^0$ . Moreover, we assume that the sets  $A^1, \dots, A^K$  have mutually disjoint boundaries; the full inclusion of one set inside another, however, is allowed.

For  $\delta > 0$ , let  $S_\delta$  denote the  $\delta$ -neighborhood of  $S$ ,

$$S_\delta = \{x \in \Omega \mid \text{dist}(x, S) < \delta\}, \quad (3.18)$$

and, similarly, let  $U_\delta^k$  denote the  $\delta$ -neighborhood of  $\partial A^k$ ,

$$U_\delta^k = \left\{ x \in \Omega \mid \text{dist}(x, \partial A^k) < \delta \right\}, \quad \text{and} \quad U_\delta = \bigcup_{k=1}^M U_\delta^k. \quad (3.19)$$

Then we define the open complement  $D_\delta$  as

$$D_\delta = \Omega \setminus (\overline{U_\delta \cup S_\delta}), \quad (3.20)$$

the  $\delta$ -interior  $A_\delta^k$  of  $A^k$  as

$$A_\delta^k = \left( \Omega \setminus \overline{U_\delta^k} \right) \cap A^k, \quad \text{and} \quad A_\delta = \bigcup_{k=1}^K A_\delta^k. \quad (3.21)$$

Figure 1 shows a typical arrangement in two dimensions.

Let  $u_\delta$  be an approximation of  $u$  obtained by a linear method. For example,  $u_\delta$  may be the interpolant of  $u$  in an  $H^1$ -conforming FE space; then, the parameter  $\delta$  corresponds to the mesh size. We assume that for each  $m = 1, \dots, M$ , the approximations  $\chi_\delta^{0,m} \in \mathcal{V}^\delta$  of  $\chi^{0,m}$  satisfy  $\lim_{\delta \rightarrow 0} \chi_\delta^{0,m} = \chi^{0,m}$  in  $L^2(\Omega)$ , and similarly, for each  $k = 1, \dots, K$ , that the approximations  $\chi_\delta^k \in \mathcal{V}_0^\delta$  of  $\chi^k$  satisfy  $\lim_{\delta \rightarrow 0} \chi_\delta^k = \chi^k$  in  $L^2(\Omega)$ . For each  $\delta > 0$ , the  $H^1$ -regular approximation  $u_\delta$  is thus given by

$$u_\delta = u_\delta^0 + \tilde{u}_\delta, \quad (3.22)$$

where

$$u_\delta^0 = \sum_{m=1}^M \hat{u}^{0,m} \chi_\delta^{0,m} \in \mathcal{V}^\delta, \quad \tilde{u}_\delta = \sum_{k=1}^K \hat{u}^k \chi_\delta^k \in \mathcal{V}_0^\delta. \quad (3.23)$$

### 3.2 Statement of main results and discussion

To simplify the presentation, we include in this section only short proofs and proofs of the main results; the remaining proofs are provided in Section 3.3. The main result, given by Theorem 5, provides estimates for the approximation  $\varphi_0$  of the background  $u^0$ , and for the first  $K$  eigenvalues  $\lambda_k$  and eigenfunctions  $\varphi_k$  of  $L_\varepsilon[u_\delta]$  in (2.11). From Theorem 5, we deduce Corollary 6 which provides similar estimates for finite element formulations or for  $u_\delta$  obtained from the convolution of  $u$  with a mollifier.

Theorem 5 relies on Lemmas 2 and 4 which require the approximation  $\{\chi_\delta\}_\delta$  of each characteristic function  $\chi = \chi^k$ , and  $\chi = \chi^{0,m}$ , to satisfy

$$\|\nabla \chi_\delta\|_{L^1(\Omega)} \leq C, \quad (3.24)$$

for every  $\delta > 0$  sufficiently small. Note that since  $\chi_\delta$  converges to a function  $\chi$  with jump discontinuities, the gradients of  $\chi_\delta$  need not be bounded uniformly with respect to  $\delta$ , for  $\delta$  close to zero. Whether (3.24) is satisfied depends on geometric properties of the method by which  $\chi_\delta$  is obtained, as well as on properties of the set  $A = \chi^{-1}(\{1\})$ . The following lemma provides sufficient conditions for (3.24) to hold.

**Lemma 1.** *Let  $A \subset \mathbb{R}^d$  be a bounded Lipschitz domain,*

$$U_\delta = \left\{ x \in \mathbb{R}^d \mid \text{dist}(x, \partial A) < \delta \right\}$$

*with  $\delta \in (0, \eta]$ , for  $\eta > 0$ , and  $\mathcal{L}$  the Lebesgue measure. Then, the following assertions hold:*

1. There exists a constant  $C > 0$  such that  $\delta^{-1}\mathcal{L}(U_\delta) < C$ , for every  $\delta \in (0, \eta]$ .
2. If  $\{g_\delta\}_{\delta \in (0, \eta]} \subset L^p(\mathbb{R}^d)$  such that  $\text{supp}(g_\delta) \subset \overline{U_\delta}$  for all  $\delta \in (0, \eta]$ , and if

$$\delta^{1-1/p} \|g_\delta\|_{L^p(\mathbb{R}^d)} \leq C_1, \quad \forall \delta \in (0, \eta] \quad (3.25)$$

for some  $p \in [1, \infty]$  (with the usual convention  $1/\infty := 0$ ), then there exists a constant  $C > 0$ , such that for every  $\delta \in (0, \eta]$ ,

$$\|g_\delta\|_{L^1(\mathbb{R}^d)} \leq C. \quad (3.26)$$

In particular, for  $p = \infty$ , (3.25) reduces to  $\delta|g_\delta| \leq C_1$  a.e. in  $\mathbb{R}^d$ , for all  $\delta \in (0, \eta]$ . Also note that for  $p = 1$ , the conclusion of the lemma is trivial.

In the following, we assume  $\eta > 0$  sufficiently small such that

$$A_\eta^k \neq \emptyset \quad \forall k, \quad \overline{S_\eta} \cap \overline{U_\eta} = \emptyset, \quad \overline{U_\eta^k} \cap \overline{U_\eta^j} = \emptyset \quad \forall k \neq j, \quad (3.27)$$

and for each connected component  $E_\eta$  of  $D_\eta \setminus A_\eta$ , there holds  $\mathcal{H}^{d-1}(\partial E_\eta \cap \partial\Omega) > 0$ . In other words, we assume for  $\delta \in (0, \eta]$ , that all  $\delta$ -interiors  $A_\delta^k$  of  $A^k$  are non-empty, that all  $\delta$ -neighborhoods of the interfaces of the medium,  $S_\delta$  and  $U_\delta^k$ , do not intersect, and that the only parts of the open complement,  $D_\delta$ , isolated from the boundary  $\partial\Omega$  are the  $\delta$ -interiors  $A_\delta^k$  of the inclusions  $A^k$ . Since the boundaries of  $A^1, \dots, A^K$  are mutually disjoint, and  $A^k \subset\subset \Omega \setminus S$ ,  $k = 1, \dots, K$ , such an  $\eta > 0$  exists. We further suppose that for each  $m = 1, \dots, M$ ,

$$\nabla \chi_\delta^{0,m} \in L^\infty(\Omega), \quad \text{supp}(\nabla \chi_\delta^{0,m}) \subset \overline{S_\delta}, \quad \forall \delta \in (0, \eta], \quad (3.28a)$$

and that for each  $k = 1, \dots, K$ ,

$$\nabla \chi_\delta^k \in L^\infty(\Omega), \quad \text{supp}(\nabla \chi_\delta^k) \subset \overline{U_\delta^k}, \quad \forall \delta \in (0, \eta]. \quad (3.28b)$$

These assumptions imply that

$$\text{supp}(\nabla u_\delta^0) \subset \overline{S_\delta}, \quad \text{supp}(\nabla \tilde{u}_\delta) \subset \overline{U_\delta}, \quad (3.29)$$

and

$$\mu_\varepsilon[u_\delta] = \varepsilon^{-1} \quad \text{a.e. in } D_\delta. \quad (3.30)$$

By (3.28), for each  $\delta \in (0, \eta]$ ,  $\nabla u_\delta \in L^\infty(\Omega)$ , and thus, due to (3.6) with  $w = u_\delta$ , the operator  $L_\varepsilon[u_\delta]$  is uniformly elliptic in  $\Omega$ , for  $\varepsilon > 0$ . A simple but useful conclusion we can draw from (3.30) is

$$\|\nabla \varphi\|_{L^2(D_\delta)}^2 \leq \varepsilon B[\varphi, \varphi] \quad \forall \varphi \in H^1(\Omega). \quad (3.31)$$

For the approximation  $\varphi_0$  of the background  $u^0$ , we have the following estimate.

**Lemma 2.** *For every  $\varepsilon > 0$  and  $\delta \in (0, \eta]$ , there holds*

$$\|\nabla \varphi_0\|_{L^2(D_\delta)}^2 \leq \varepsilon \|\nabla u_\delta^0\|_{L^1(\Omega)}. \quad (3.32)$$

The estimates for the eigenfunctions  $\varphi_j$  are based on the following simple result.

**Proposition 3.** For every  $\varepsilon > 0$  and  $\delta \in (0, \eta]$  there holds

$$\|\nabla\varphi_j\|_{L^2(D_\delta)}^2 \leq \lambda_j(\varepsilon, \delta)\varepsilon \quad \forall j \geq 1. \quad (3.33)$$

*Proof.* Since for each  $j$ ,  $\|\varphi_j\| = 1$ , the proposition follows from (3.31) and (3.12).  $\square$

Thus, to show that an eigenfunction  $\varphi_k$  is “almost” piecewise constant, we need to estimate the corresponding eigenvalue  $\lambda_k$  of  $L_\varepsilon[u_\delta]$  in  $\mathcal{V}_0^\delta$ , for which we rely on the following lemma.

**Lemma 4.** There exists a constant  $C$ , independent of  $\hat{u}^1, \dots, \hat{u}^K$  and  $\hat{u}^{0,1}, \dots, \hat{u}^{0,M}$  such that for every  $\varepsilon > 0$ ,  $\delta \in (0, \eta]$ , and  $k = 1, \dots, K$ , there holds

$$\lambda_k \leq \frac{C|\tau|}{\min_j |\hat{u}^j|}, \quad \tau = (\tau_k) \in \mathbb{R}^K, \quad \tau_k = \tau_k(\delta) = \|\nabla\chi_\delta^k\|_{L^1(\Omega)}. \quad (3.34)$$

Together, the results above yield the following theorem.

**Theorem 5.** Let  $u$  be given by (3.13), the approximation,  $u_\delta$ , of  $u$  be given by (3.22),  $\varphi_0$  be given by (3.9), and  $(\lambda_k, \varphi_k)$  with  $k \geq 1$  satisfy (3.12), where  $(\lambda_k)_k$  is non-decreasing and  $(\varphi_k)_k$  orthonormal in  $L^2(\Omega)$ . Suppose  $A^1, \dots, A^K \subset \Omega$  and  $A^{0,1}, \dots, A^{0,M} \subset \Omega$  have Lipschitz boundaries, and  $\eta > 0$  such that (3.27) is satisfied. If (3.28) hold true, and there exists a constant  $C$  such that for every  $\delta \in (0, \eta]$ , each of the functions  $\chi_\delta = \chi_\delta^{0,m}$ ,  $m = 1, \dots, M$ , or  $\chi_\delta = \chi_\delta^k$ ,  $k = 1, \dots, K$ , satisfies

$$\delta\|\nabla\chi_\delta\|_{L^\infty(\Omega)} \leq C, \quad (3.35)$$

then there exists a constant  $C_1$ , independent of the coefficients  $\hat{u}^{0,1}, \dots, \hat{u}^{0,M}$  and  $\hat{u}^1, \dots, \hat{u}^K$  such that for every  $\varepsilon > 0$  and  $\delta \in (0, \eta]$ , the following estimates hold:

$$\|\nabla\varphi_0\|_{L^2(D_\delta)}^2 \leq C_1 \max_m |\hat{u}^{0,m}| \varepsilon, \quad (3.36)$$

$$\|\nabla\varphi_k\|_{L^2(D_\delta)}^2 \leq \frac{C_1}{\min_j |\hat{u}^j|} \varepsilon, \quad \text{and} \quad \lambda_k \leq \frac{C_1}{\min_j |\hat{u}^j|}, \quad k = 1, \dots, K. \quad (3.37)$$

Moreover, for each Lipschitz domain  $V \subset D_\eta \setminus A_\eta$  with  $\mathcal{H}^{d-1}(\partial V \cap \partial\Omega) > 0$ , there exists a constant  $C_2$ , independent of the coefficients  $\hat{u}^{0,1}, \dots, \hat{u}^{0,M}$  and  $\hat{u}^1, \dots, \hat{u}^K$  such that for every  $\varepsilon > 0$  and  $\delta \in (0, \eta]$ , the following estimates are satisfied

$$\|u^0 - \varphi_0\|_{L^2(V)}^2 \leq C_2 \max_m |\hat{u}^{0,m}| \varepsilon, \quad (3.38)$$

$$\|\varphi_k\|_{L^2(V)}^2 \leq \frac{C_2}{\min_j |\hat{u}^j|} \varepsilon, \quad k = 1, \dots, K. \quad (3.39)$$

*Proof.* Estimates (3.36) and (3.37) easily follow from Lemma 1 with  $p = \infty$ , Lemmas 2 and 4, and Proposition 3. Estimates (3.38) and (3.39) follow from the Poincaré inequality, and estimates (3.36) and (3.37), respectively, since  $\mathcal{H}^{d-1}(\partial V \cap \partial\Omega) > 0$ .  $\square$

Essentially, estimates (3.36) and (3.37) imply that  $\varphi_0$  and the first  $K$  eigenfunctions  $\varphi_k$  of  $L_\varepsilon[u_\delta]$  are “almost” constant in each connected component of  $D_\delta$ . In particular,  $\varphi_0$  is almost constant in each connected component of  $D_\delta \setminus A_\delta$ . Since the Hausdorff measure of  $\partial\Omega \cap \partial A^{0,m}$  is positive for each  $m = 1, \dots, M$ , and  $\varphi_0$  coincides with  $u_\delta^0$  on  $\partial\Omega$ ,  $\varphi_0$  indeed approximates  $u_\delta^0$

well in  $D_\delta \setminus A_\delta$ . Similarly, for  $k = 1, \dots, K$ ,  $\varphi_k$  is almost constant in each connected component of  $D_\delta \setminus A_\delta$  and vanishes on  $\partial\Omega$ ; therefore, every  $\varphi_k$  is small in  $D_\delta \setminus A_\delta$ . Consequently,  $u_\delta$  can be well approximated in  $\varphi_0 + \Phi_K$ , for instance, by its  $L^2$ -best approximant  $\varphi_0 + \Pi_K(u_\delta - \varphi_0)$ , where  $\Pi_K : \mathcal{V}^\delta \rightarrow \Phi_K$  denotes the standard  $L^2$ -orthogonal projection

$$\langle v - \Pi_K v, \varphi \rangle = 0, \quad \forall \varphi \in \Phi_K. \quad (3.40)$$

From Theorem 5, we may deduce estimates (3.36)-(3.39) for specific methods of approximation.

**Corollary 6.** *Suppose  $A^1, \dots, A^K$  and  $A^{0,1}, \dots, A^{0,M}$  have Lipschitz boundaries. Estimates (3.36)-(3.39) hold true in each of the two following cases:*

1. *For each  $\delta \in (0, \eta]$ ,  $u_\delta$  is the convolution of  $u$  with the standard mollifier (e.g., [13]), and  $\mathcal{V}^\delta = H^1(\Omega)$ .*
2. *For each  $\delta \in (0, \eta]$ ,  $u_\delta$  is the Lagrange interpolant of  $u$  in an  $H^1$ -conforming FE space  $V_\delta$  associated with a simplex mesh  $\mathcal{T}_\delta$  with mesh size  $\delta$ , where the family of meshes  $\{\mathcal{T}_\delta\}_{\delta \in (0, \eta]}$  is regular and quasi-uniform (see, e.g., [29]), and either  $\mathcal{V}^\delta = V_\delta$  or  $\mathcal{V}^\delta = H^1(\Omega)$ .*

*Proof.* For the proof, it is sufficient to show that in each case (i), (ii), the hypotheses of Theorem 5 are satisfied, i.e., that the approximation  $\chi_\delta$  of a characteristic function of a Lipschitz domain contained in  $\Omega$  satisfies (3.35). Here we sketch the proof for (i); the result for (ii) follows easily from the definitions of a regular and quasi-uniform family of FE meshes and from basic properties of polynomial Lagrange interpolation.

Let  $\chi$  be the characteristic function of a Lipschitz domain contained in  $\Omega$ . We extend  $\chi$  to  $\mathbb{R}^d$  by setting  $\chi = 0$  outside  $\Omega$ , and set

$$\chi_\delta(x) = \zeta_\delta * \chi = \int_{\mathbb{R}^d} \zeta_\delta(x-y)\chi(y) dy, \quad \zeta_\delta(x) = \delta^{-d}\zeta(x/\delta) \quad (3.41)$$

with  $\zeta$  the standard mollifier. Since

$$\nabla \chi_\delta = \nabla \zeta_\delta * \chi, \quad (3.42)$$

and  $\|\chi\|_{L^\infty(\mathbb{R}^d)} = 1$ , we obtain

$$\delta |\nabla \chi_\delta(x)| \leq \|\chi\|_\infty \int_{|y|<\delta} \delta |\nabla \zeta_\delta(y)| dy = \int_{|z|<1} |\nabla \zeta(z)| dz < \infty \quad (3.43)$$

for a.e.  $x \in \Omega$ , which concludes the proof.  $\square$

**Remark 7.** Let us suppose, for simplicity, that  $K = 1$ ,  $A^1 = A$ ,  $\hat{u}^1 = 1$ ,  $u^0 = 0$ , and the hypotheses of Theorem 5 are satisfied. By following the proof for this simpler case, one finds that for small  $\eta$  and  $\varepsilon$ ,  $\lambda_1$  is bounded from above by a constant arbitrarily close to

$$\frac{\limsup_{\rho \rightarrow 0^+} \|\nabla \chi_\rho\|_{L^1(\Omega)}}{\mathcal{L}(A)}. \quad (3.44)$$

For an appropriate approximation  $\{\chi_\delta\}_\delta$  of the characteristic function  $\chi^1$  of  $A$ , this constant essentially coincides with the eigenvalue

$$\lambda = \frac{\mathcal{H}^{d-1}(\partial A)}{\mathcal{L}(A)} \quad (3.45)$$

of the (subdifferential of the) total variation (TV) functional associated with  $\chi^1$  in two dimensions [1]. However,  $\chi^1$  can be an eigenfunction of the TV functional associated with the eigenvalue (3.45), only if  $A$  is convex and has boundary of class  $C^{1,1}$ . In contrast, the conclusion of Theorem 5 is valid for any Lipschitz domain  $A$ ; in particular,  $A$  need not be convex. Since the TV functional is only lower semi-continuous, we note that the quantities in (3.44) and (3.45) may not be equal.

**Remark 8.** The main results, given by Theorem 5 and Corollary 6, essentially remain true if we replace the last inequality in (3.4) by the more general inequality (3.7) with  $r \in [1, 2]$ . The proofs only require slight modifications, as much of the analysis actually carries over from the case  $r = 1$ . The main adjustments are required in Lemmas 2 and 4, where estimates of  $\mu_\varepsilon[u_\delta^0]|\nabla u_\delta^0|$  and  $\mu_\varepsilon[\tilde{u}_\delta]|\nabla \tilde{u}_\delta|$  are employed. Since the modifications for the two Lemmas are similar, we only address the former here. Instead of the estimate

$$\int_{S_\delta} \mu_\varepsilon[u_\delta^0] |\nabla u_\delta^0|^2 dx \leq \|\nabla u_\delta^0\|_{L^1(\Omega)} \quad (3.46)$$

which relies on (3.4), we obtain

$$\int_{S_\delta} \mu_\varepsilon[u_\delta^0] |\nabla u_\delta^0|^2 dx \leq \mathcal{L}(S_\delta)^{r-1} \|\nabla u_\delta^0\|_{L^1(\Omega)}^{2-r} \quad (3.47)$$

by using (3.7) and Hölder's inequality. As a consequence, for  $r \in (1, 2]$ , the conclusions of Theorem 5 and Corollary 6 essentially remain true, and can even be improved by introducing a small multiplicative term to the right hand sides of the inequalities.

### 3.3 Proofs

*Proof of Lemma 1.* 1. We show that  $\psi : [0, \eta] \rightarrow \mathbb{R}$  given by

$$\psi(\delta) = \begin{cases} \delta^{-1} \mathcal{L}(U_\delta) & \delta \neq 0 \\ 2\mathcal{H}^{d-1}(\partial A) & \delta = 0 \end{cases} \quad (3.48)$$

is bounded by proving that  $\psi$  is continuous in  $[0, \eta]$ . Since  $A$  is a bounded Lipschitz domain, its boundary  $\partial A$  is  $(d-1)$ -rectifiable, i.e., there exists a Lipschitz function from a bounded subset of  $\mathbb{R}^{d-1}$  onto  $\partial A$ . Hence the  $(d-1)$ -dimensional Minkowski content and the  $(d-1)$ -dimensional Hausdorff measure of  $\partial A$  coincide [14, Theorem 3.2.39], that is

$$\lim_{\delta \rightarrow 0^+} \delta^{-1} \mathcal{L}(U_\delta) = 2\mathcal{H}^{d-1}(\partial A). \quad (3.49)$$

Therefore  $\lim_{\delta \rightarrow 0^+} \psi(\delta) = \psi(0)$  and  $\psi$  is continuous at  $\delta = 0$ . Since  $\rho(x) = \text{dist}(x, \partial A)$  is Lipschitz and satisfies  $|\nabla \rho(x)| = 1$ , a.e.  $x \in U_\delta$  [10, §6], we have

$$\mathcal{L}(U_\delta) = \int_0^\delta \mathcal{H}^{d-1}(\rho^{-1}(\{t\})) dt, \quad (3.50)$$

by the co-area formula [13, §3.4]. Combining this with (3.49), we conclude that the function  $\delta \mapsto \mathcal{L}(U_\delta)$  is continuous in  $[0, \eta]$ . It follows that  $\psi$  is continuous in  $(0, \eta]$ . Since  $\psi$  is also continuous at  $\delta = 0$ , we have that it is continuous in the entire closed interval  $[0, \eta]$ , and is, therefore, bounded.

2. By Hölder's inequality we have

$$\|g_\delta\|_{L^1(\mathbb{R}^d)} = \int_{U_\delta} |g_\delta(x)| dx \leq \left(\mathcal{L}(U_\delta)\right)^{1-1/p} \|g_\delta\|_{L^p(\mathbb{R}^d)}, \quad (3.51)$$

which, together with (3.25), yields

$$\|g_\delta\|_{L^1(\mathbb{R}^d)} \leq C_1 \left(\delta^{-1} \mathcal{L}(U_\delta)\right)^{1-1/p}. \quad (3.52)$$

Thus the conclusion follows from assertion 1.  $\square$

*Proof of Lemma 2.* By Dirichlet's principle,  $\varphi_0$  is the unique minimizer of the functional  $I(v) = B[v, v]$  in  $u_\delta + \mathcal{V}_0^\delta$ , i.e.,

$$B[\varphi_0, \varphi_0] \leq B[u_\delta + v, u_\delta + v], \quad \forall v \in \mathcal{V}_0^\delta,$$

or, since  $u_\delta^0 - u_\delta = -\tilde{u}_\delta \in \mathcal{V}_0^\delta$ , equivalently

$$B[\varphi_0, \varphi_0] \leq B[u_\delta^0 + v, u_\delta^0 + v], \quad \forall v \in \mathcal{V}_0^\delta. \quad (3.53)$$

By (3.53) with  $v = 0$ , and using (3.29)

$$B[\varphi_0, \varphi_0] \leq B[u_\delta^0, u_\delta^0] = \int_{\Omega} \mu_\varepsilon[u_\delta] |\nabla u_\delta^0|^2 = \int_{S_\delta} \mu_\varepsilon[u_\delta] |\nabla u_\delta^0|^2. \quad (3.54)$$

Then, using (3.3),  $\nabla u_\delta = \nabla u_\delta^0$ , which holds in  $S_\delta$ , and (3.5) we obtain

$$B[\varphi_0, \varphi_0] \leq \int_{S_\delta} \mu_\varepsilon[u_\delta] |\nabla u_\delta^0|^2 = \int_{S_\delta} \mu_\varepsilon[u_\delta^0] |\nabla u_\delta^0|^2 \leq \|\nabla u_\delta^0\|_{L^1(S_\delta)}. \quad (3.55)$$

Finally, the conclusion follows from (3.31).  $\square$

The proof of Lemma 4 requires the following result.

**Lemma 9.** *If the sets  $A_\eta^k$ ,  $k = 1, \dots, K$ , are nonempty and (3.27) is satisfied, for some  $\eta > 0$ , then the following assertions hold true.*

1. *There exists  $1 \leq j \leq K$ , such that*

$$\mathcal{B}_\eta^j \setminus \bigcup_{\substack{k=1 \\ k \neq j}}^K \overline{\mathcal{B}_\eta^k} \neq \emptyset, \quad \mathcal{B}_\eta^\ell = A_\eta^\ell \cap D_\eta. \quad (3.56)$$

2. *The restrictions of the functions  $\chi^k$ ,  $k = 1, \dots, K$ , to  $D_\eta$  are linearly independent.*

3. If for every  $k = 1, \dots, K$ , (3.28b) is satisfied, then there exists a constant  $C > 0$ , such that for every  $\delta \in [0, \eta]$  and  $\widehat{\psi} = (\psi^k) \in \mathbb{R}^K$ , the following estimates are satisfied:

$$\|\psi_\delta\|^2 \geq \|\psi\|_{L^2(D_\eta)}^2 \geq C|\widehat{\psi}|^2, \quad (3.57)$$

where  $\psi = \sum_k \psi^k \chi^k$ , and  $\psi_\delta = \sum_k \psi^k \chi_\delta^k$ .

*Proof.* 1. Since one set  $A_\eta^k$  may be included inside another, the index  $j$  must be such that the boundary of  $A_\eta^j$  is contained in the boundary of the union  $A_\eta = \cup_k A_\eta^k$ . We show that this requirement is sufficient: Fix some  $x \in \partial A_\eta$ . Then, for each neighborhood  $W$  of  $x$ , there holds

$$W \cap A_\eta \neq \emptyset, \quad W \cap (\Omega \setminus A_\eta) \neq \emptyset. \quad (3.58)$$

Since

$$W \cap A_\eta = \bigcup_{k=1}^K (W \cap A_\eta^k), \quad (3.59)$$

and

$$W \cap (\Omega \setminus A_\eta) = \bigcap_{k=1}^K (W \cap (\Omega \setminus A_\eta^k)), \quad (3.60)$$

we have  $x \in \partial A_\eta^j$ , for some  $1 \leq j \leq K$ . Since  $\partial A_\eta^j \subset \partial U_\eta^j$  and, by (3.27), the sets  $\overline{U_\eta^k}$  are mutually disjoint,  $x \notin \partial A_\eta^k$ , for all  $k \neq j$ . Similarly, we have  $x \notin \partial S_\eta$ . Hence, there exists an open neighborhood  $W$  of  $x$ ,  $W \cap (\overline{A_\eta^j} \cup \overline{U_\eta^j}) = \emptyset$ , for all  $k \neq j$ , and  $W \cap \overline{S_\eta} = \emptyset$ . Moreover, since  $x \in \partial A_\eta^j$ , the intersection  $V = W \cap A_\eta^j$  is nonempty, while  $V \cap \overline{U_\eta^k} = \emptyset$ , for all  $k = 1, \dots, K$ , and  $V \cap \overline{S_\eta} = \emptyset$ . Thus,  $V$  is nonempty, open and satisfies  $V \subset A_\eta^j \cap D_\eta = \mathcal{B}_\eta^j$  and  $V \cap \overline{\mathcal{B}_\eta^k} \subset V \cap \overline{A_\eta^k} = \emptyset$ , for  $k \neq j$ , by (3.21) and (3.27). This yields (3.56).

2. Suppose

$$\sum_{k=1}^K \psi^k \chi^k = 0 \quad \text{a.e. in } D_\eta \quad (3.61)$$

for some  $\psi^k \in \mathbb{R}$ ,  $k = 1, \dots, K$ . According to 1, there is an index  $j$  for which (3.56) holds. Note that the set on the left-hand side of (3.56) is open and therefore is of positive measure. Without loss of generality, assume that  $j = K$  satisfies (3.56). However, by (3.61) this can be, only if  $\psi^K = 0$ . The above argument may be repeated by induction. Since there is a finite number of functions  $\chi^k$ , the procedure stops only when a single function remains, say  $\chi^1$ . Then (3.61) reduces to  $\psi^1 \chi^1 = 0$ . Again, since  $A_\eta^1 \cap D_\eta$  is open and nonempty, this implies  $\psi^1 = 0$  and the restrictions of  $\chi^k$  to  $D_\eta$  are thus linearly independent.

3. Let  $\widehat{\psi} = (\psi^k) \in \mathbb{R}^K$ ,  $\delta \in [0, \eta]$ ,  $\psi = \sum_k \psi^k \chi^k$ , and  $\psi_\delta = \sum_k \psi^k \chi_\delta^k$ , where  $\chi_0^k = \chi^k$ . As each  $\chi_\delta^k$  coincides with  $\chi^k$  in  $D_\delta$ , we have

$$\|\psi_\delta\|^2 \geq \int_{D_\delta} \left| \sum_{k=1}^K \psi^k \chi_\delta^k \right|^2 = \int_{D_\delta} \left| \sum_{k=1}^K \psi^k \chi^k \right|^2 = \|\psi\|_{L^2(D_\delta)}^2. \quad (3.62)$$

Since  $D_\eta \subset D_\delta$ ,  $\delta \leq \eta$ , we obtain the lower bound

$$\|\psi_\delta\|^2 \geq \|\psi\|_{L^2(D_\delta)}^2 \geq \|\psi\|_{L^2(D_\eta)}^2. \quad (3.63)$$

Expanding the term on the right-hand side yields

$$\|\psi\|_{L^2(D_\eta)}^2 = \sum_{k,j=1}^K \psi^k \psi^j \int_{D_\eta} \chi^k \chi^j = \sum_{k,j=1}^K M_{kj} \psi^k \psi^j \quad (3.64)$$

where

$$M_{kj} = \int_{D_\eta} \chi^k \chi^j = \mathcal{L}(A^k \cap A^j \cap D_\eta) \quad k, j = 1, \dots, K. \quad (3.65)$$

Hence  $\|\psi\|_{L^2(D_\eta)}^2$  is a quadratic form in the vector  $\widehat{\psi}$ , represented by the symmetric  $K \times K$  matrix  $M = (M_{kj})$ , and it is positive for any  $\widehat{\psi} \neq 0$  by assertion 2 of this lemma. Therefore,  $M$  is symmetric positive definite, which implies (3.57) and hence the conclusion.  $\square$

*Proof of Lemma 4.* From the spectral theory for symmetric elliptic operators [12, §6.5 – 6.6], it follows that for each  $n \geq 1$ ,

$$\lambda_n = \min_{0 \neq v \in \Phi_{n-1}^\perp} \frac{B[v, v]}{\|v\|^2}, \quad (3.66)$$

where  $\Phi_{n-1} = \text{span}\{\varphi_j\}_{j=1}^{n-1}$  for  $n \geq 2$ , or  $\Phi_0 = \{0\}$  for  $n = 1$ , and  $W^\perp$  denotes the orthogonal complement of  $W$  in  $\mathcal{V}_0^\delta$  with respect to the  $L^2(\Omega)$  inner product. For any

$$\psi_\delta = \sum_{k=1}^K \psi^k \chi_\delta^k, \quad \psi^k \in \mathbb{R}, \quad (3.67)$$

there holds

$$\langle \varphi_j, \psi_\delta \rangle = \sum_{k=1}^K \psi^k \langle \varphi_j, \chi_\delta^k \rangle. \quad (3.68)$$

Now, let  $1 \leq n \leq K$  and  $\widehat{\psi} = (\psi^k) \in \mathbb{R}^K$ , with  $|\widehat{\psi}| = 1$ , such that  $\psi_\delta$  given by (3.67) satisfies

$$\langle \varphi_j, \psi_\delta \rangle = 0, \quad j = 1, \dots, n-1, \quad (3.69)$$

for  $n \geq 2$ , or no condition at all for  $n = 1$ . One can always find such a vector of coefficients  $\widehat{\psi}$ , since  $n \leq K$  and hence the homogeneous linear system (3.69) has more unknowns than equations. By Lemma 9 there exists a constant  $C_1 > 0$ , independent of  $|\widehat{\psi}| = 1$ ,  $\delta \in (0, \eta]$  or  $\varepsilon$ , such that  $\|\psi_\delta\|^2 \geq C_1$ . Therefore,  $\psi_\delta \in \Phi_{n-1}^\perp$ , is not identically zero, and we obtain from (3.66)

$$\lambda_n \leq \frac{B[\psi_\delta, \psi_\delta]}{\|\psi_\delta\|^2} \leq \frac{B[\psi_\delta, \psi_\delta]}{C_1}. \quad (3.70)$$

Thus, it is left to estimate  $B[\psi_\delta, \psi_\delta]$ . By (3.28b), we have

$$B[\psi_\delta, \psi_\delta] = \int_\Omega \mu_\varepsilon[u_\delta] |\nabla \psi_\delta|^2 dx = \int_{U_\delta} \mu_\varepsilon[u_\delta] |\nabla \psi_\delta|^2 dx. \quad (3.71)$$

Furthermore, by employing (3.3) and the equality  $\nabla u_\delta = \nabla \widetilde{u}_\delta$  which holds a.e. in  $U_\delta$  because  $S_\delta \cap U_\delta = \emptyset$ , we get

$$B[\psi_\delta, \psi_\delta] = \int_{U_\delta} \mu_\varepsilon[\widetilde{u}_\delta] |\nabla \psi_\delta|^2 dx. \quad (3.72)$$

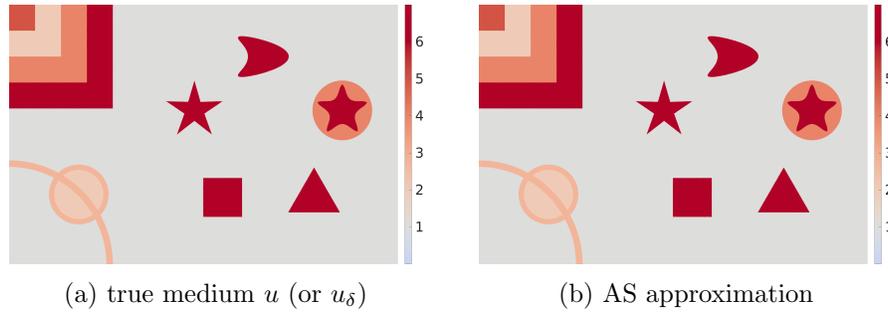


Figure 2: Two-dimensional piecewise constant medium: (a) true medium  $u$  (or  $u_\delta$ ), (b) AS approximation  $\varphi_0 + \Pi_K[u_\delta - \varphi_0]$ . Note that  $u$  and its FE-interpolant  $u_\delta$  cannot be distinguished here.

Next, we shall show that

$$|\nabla\psi_\delta(x)| < \frac{|\nabla\tilde{u}_\delta(x)|}{m}, \quad \text{a.e. } x \in U_\delta, \quad (3.73)$$

with  $m = \min_j |u^j| > 0$ . Since the sets  $U_\delta^1, \dots, U_\delta^K$  are mutually disjoint, each  $x \in U_\delta$  lies in precisely one  $U_\delta^k$ , with  $1 \leq k \leq K$ , and hence at a.e.  $x \in U_\delta$ , at most one of the gradients,  $\nabla\chi_\delta^k(x)$ , is nonzero. Therefore, a.e. in  $U_\delta$  we have

$$\begin{aligned} |\nabla\psi_\delta| &= \left| \sum_{k=1}^K \psi^k \nabla\chi_\delta^k \right| = \sum_{k=1}^K |\psi^k| |\nabla\chi_\delta^k| \\ &\leq \frac{1}{m} \sum_{k=1}^K |u^k| |\nabla\chi_\delta^k| = \frac{1}{m} \left| \sum_{k=1}^K u^k \nabla\chi_\delta^k \right| \end{aligned} \quad (3.74)$$

which completes the proof of (3.73).

Finally, from (3.72), (3.73) and (3.5) with  $w = u_\delta$ , we infer that

$$B[\psi_\delta, \psi_\delta] \leq \frac{1}{m} \int_{U_\delta} |\nabla\psi_\delta| dx = \frac{1}{m} \|\nabla\psi_\delta\|_{L^1(\Omega)}. \quad (3.75)$$

By using the definition of  $\tau_k$  in (3.34) and applying the Cauchy-Schwarz inequality in  $\mathbb{R}^K$  we obtain

$$B[\psi_\delta, \psi_\delta] \leq \frac{1}{\min_j |u^j|} \sum_{k=1}^K |\psi^k| \tau_k \leq \frac{|\tau|}{\min_j |u^j|}, \quad (3.76)$$

which, together with (3.70), yields the conclusion.  $\square$

### 3.4 Numerical example

To illustrate the remarkable approximation properties of the adaptive spectral (AS) representation in Theorem 5 and Corollary 6, we now consider in  $\Omega = (0, 1.5) \times (0, 1)$  the piecewise constant medium  $u = u^0 + \tilde{u}$ , shown in Fig. 2a. It consists of a background,  $u^0$ , and an interior

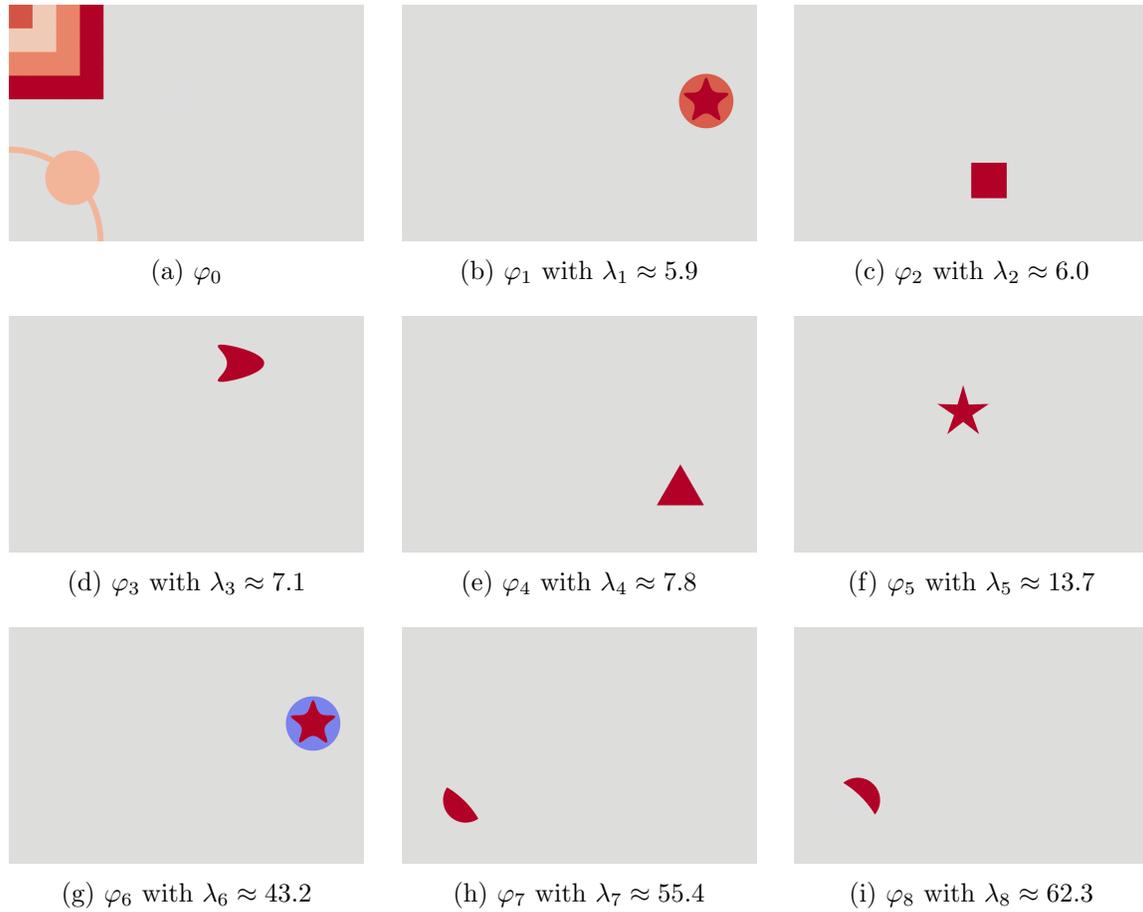


Figure 3: Two-dimensional piecewise constant medium: (a) background  $\varphi_0$  in (3.8) with  $w = u_\delta$  and  $\varepsilon = 10^{-8}$ ; (b)-(i) eigenfunctions  $\varphi_1, \dots, \varphi_8$  of (3.11) corresponding to the first eight eigenvalues  $\lambda_1, \dots, \lambda_8$ .

part,  $\tilde{u}$ , which vanishes on the boundary  $\partial\Omega$ . Both  $u^0$  and  $\tilde{u}$  are linear combinations,

$$u^0(x) = \sum_{m=1}^M \hat{u}^{0,m} \chi^{0,m}(x), \quad \tilde{u}(x) = \sum_{k=1}^K \hat{u}^k \chi^k(x), \quad \hat{u}^{0,m}, \hat{u}^k \in \mathbb{R}$$

of characteristic functions  $\chi^{0,m}$  and  $\chi^k$  associated with given subsets  $A^{0,m}$  and  $A^k$  of  $\Omega$ , respectively, with  $M = 7$  and  $K = 8$ . While the subsets  $A^{0,m}$  composing the background reach the boundary  $\partial\Omega$  in the sense of  $\mathcal{H}^{d-1}(\partial A^{0,m} \cap \partial\Omega) > 0$ , where  $\mathcal{H}^{d-1}$  denotes the  $(d-1)$ -dimensional Hausdorff measure, the inclusions  $A^k$  lie strictly inside  $\Omega$ .

First, we approximate  $u$  by its  $H^1$ -conforming  $\mathcal{P}^1$ -FE interpolant,  $u_\delta \in W^{1,\infty}(\Omega)$ , on a regular triangular mesh, locally refined along discontinuities, with 242'790 elements varying in size between  $4.4 \times 10^{-4}$  and  $3.8 \times 10^{-1}$ . Due to the fine adaptively constructed mesh, the FE-approximation errors are essentially negligible. In fact, since  $u$  and  $u_\delta$  differ by a relative  $L^2$ -error as little as 2.5%, they are hardly distinguishable in Fig. 2a.

Now, we compute the approximation  $\varphi_0$  of the background by solving (3.8) with  $L_\varepsilon[u_\delta]$  and  $\mu_\varepsilon[u_\delta]$  as in (2.11), (2.12) and  $\varepsilon = 10^{-8}$ . As shown in Fig. 3a,  $\varphi_0$  appears essentially piecewise

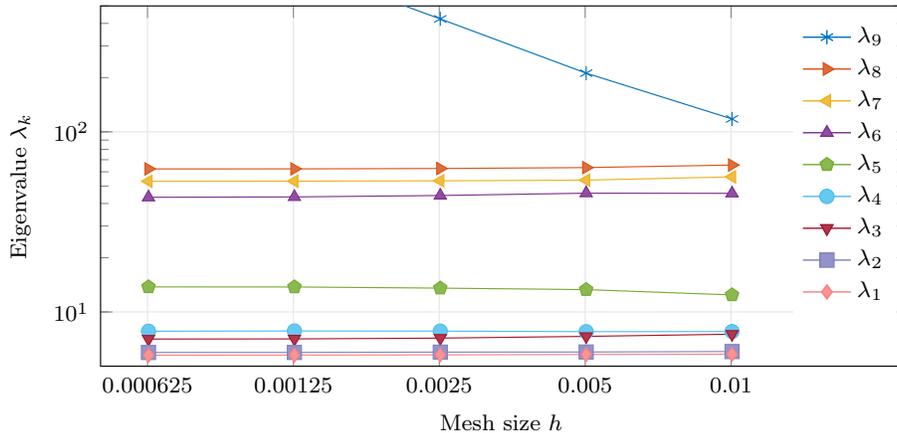


Figure 4: Two-dimensional piecewise constant medium: eigenvalues  $\lambda_1, \dots, \lambda_9$  of  $L_\varepsilon[u_\delta]$  for mesh sizes  $\delta = h = 0.01/2^m$ ,  $m = 0, \dots, 4$ , and fixed  $\varepsilon = 10^{-8}$ .

constant throughout  $\Omega$  while correctly representing the various components of the background  $u^0$ . Similarly, the first eight eigenfunctions  $\varphi_1, \dots, \varphi_8$  of  $L_\varepsilon[u_\delta]$ , defined by (3.11), correctly identify in Figs. 3b – 3i all remaining interior sets (or inclusions). For all isolated inclusions, each eigenfunction  $\varphi_k$  accurately matches a single characteristic function  $\chi^k$ . However, for the two overlapping star- and disk-shaped sets,  $A^1$  and  $A^6$ , each of the remaining two eigenfunctions  $\varphi_1$  and  $\varphi_6$  essentially correspond to a linear combination of  $\chi^1$  and  $\chi^6$ ; hence,  $\varphi_1$  and  $\varphi_6$  span the same two-dimensional subspace as  $\chi^1$  and  $\chi^6$ . Note that neither  $\varphi_0$  nor  $\varphi_1, \dots, \varphi_8$  are truly piecewise constant but, in fact, lie in  $H^1(\Omega)$ . Although their gradients nearly vanish in  $D_\delta$  (i.e., outside the  $\delta$  neighborhoods of the interfaces), due to the upper bounds (3.36) and (3.37) of Theorem 5, the eigenfunctions do vary (slightly) throughout  $\Omega$ .

Next, we subtract the background  $\varphi_0$  from  $u_\delta$  and compute the  $L^2$ -projection  $\Pi_K[u_\delta - \varphi_0]$  into  $\Phi_K = \text{span}\{\varphi_i\}_{i=1}^K$ , defined in (3.40). Since  $\varphi_0$  matches well  $u^0$  and  $\varphi_1, \dots, \varphi_8$  essentially span the same eight-dimensional subspace as  $\chi_1, \dots, \chi_8$ ,  $\Pi_K[u_\delta - \varphi_0]$  approximates  $\tilde{u}$  remarkably well. Hence, the combined AS representation  $\varphi_0 + \Pi_K[u_\delta - \varphi_0]$ , shown in Figure 2b, also approximates remarkably well the entire medium  $u$ , or  $u_\delta$ , with a relative  $L^2$ -error as little as 2.5%, or 0.05%, respectively. In fact,  $u$  is hardly distinguishable from the AS approximation  $\varphi_0 + \Pi_K[u_\delta - \varphi_0]$  in Figure 2.

Finally, we monitor in Fig. 4 the behavior of the first nine eigenvalues of  $L_\varepsilon[u_\delta]$  for a sequence of increasingly finer quasi-uniform FE meshes and fixed  $\varepsilon = 10^{-8}$ . While the first eight eigenvalues remain bounded, thereby validating the upper bound (3.37), the ninth eigenvalue apparently diverges as  $h$  tends to zero.

## 4 Numerical experiments

Here we present two numerical examples which illustrate the accuracy and the usefulness of the ASI method for the solution of inverse medium problems. In the first, the unknown medium is composed of five simple geometric inclusions, and in the second, the medium is a two-dimensional model of a salt dome from geophysics.

We apply the ASI Algorithm from Section 2.2 with frequency stepping for the minimization

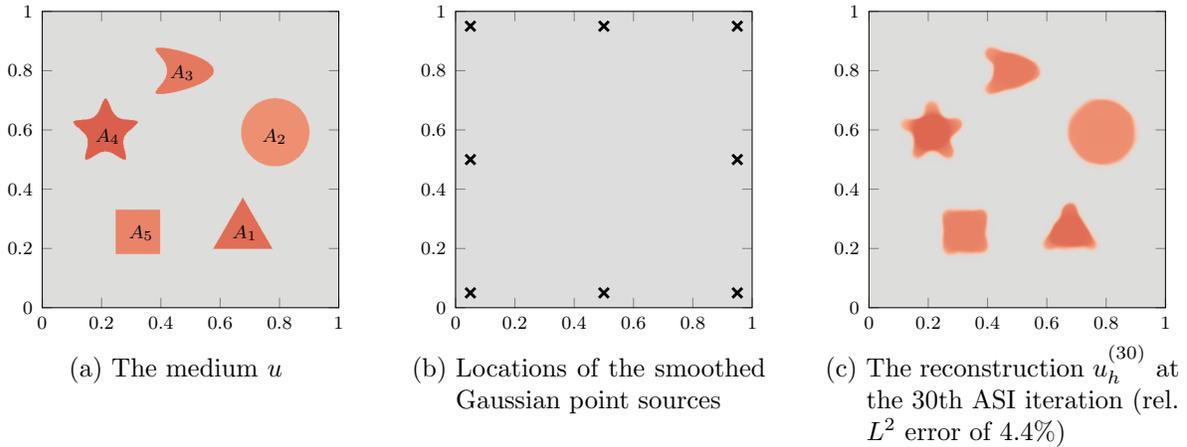


Figure 5: Five simple geometric inclusions: ASI reconstruction  $u_h$  of the unknown medium  $u$  from boundary observations with 20% noise

of the misfit  $\mathcal{J}$ , given by (2.2), where the forward solution operator of the boundary value problem is replaced by a  $\mathcal{P}^3$ -FE Galerkin approximation. For  $w$  a continuous, piecewise linear FE function, the operator  $L_\varepsilon[w]$ , needed in Step 5 of the ASI Algorithm, is given by (2.11), where,  $\mu_\varepsilon = \mu_\varepsilon[w]$  is given by (2.12) and  $\varepsilon = 10^{-8}$ . To find a minimizer  $u_h^{(m)}$  of  $\mathcal{J}$  in  $\varphi_0^{(m)} + \Psi^{(m)}$  in Step 2 of the ASI Algorithm at the  $m$ -th ASI iteration, we use the BFGS quasi-Newton method [26]. We stop the iterations of the BFGS method once the relative norm of the gradient of the function

$$\beta \mapsto \mathcal{J} \left[ \varphi_0^{(m)} + \sum_{j=1}^{J_m} \beta_j \varphi_j^{(m)} \right]$$

is smaller than  $10^{-6}$ . The criterion for incrementing the frequency is given by (2.18) with a tolerance  $\varepsilon_\nu = 0.005$ . In Step 7 of the ASI Algorithm, we use (2.16) with  $\varepsilon_\Psi = \varepsilon_\nu = 0.005$ .

#### 4.1 Five simple geometric inclusions

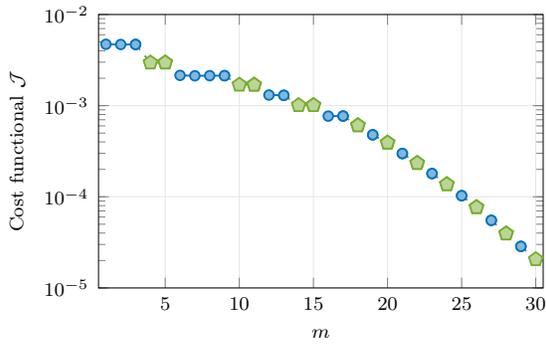
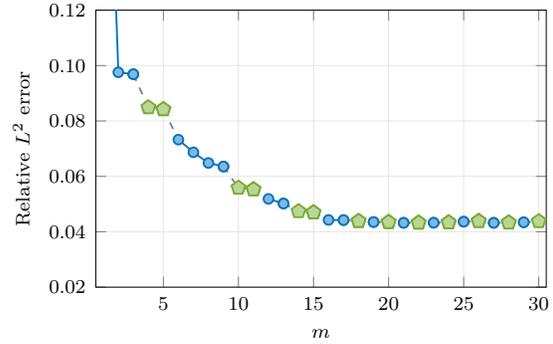
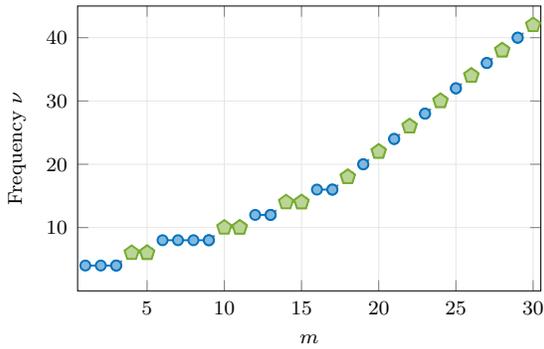
We consider inside  $\Omega = (0, 1) \times (0, 1)$  the unknown piecewise constant medium

$$u(x) = 2 + 1.4\chi_{A_1}(x) + 1.1\chi_{A_2}(x) + 1.3\chi_{A_3}(x) + 1.5\chi_{A_4}(x) + 1.2\chi_{A_5}(x),$$

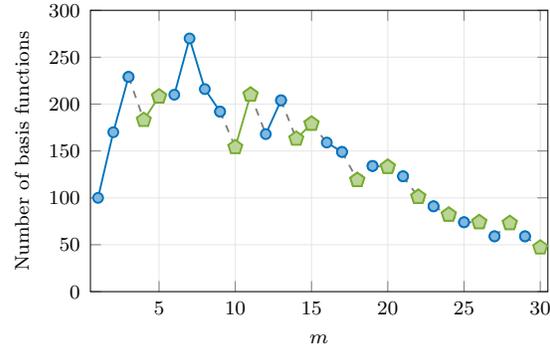
where  $\chi_{A_i}$  denotes the characteristic function of the set  $A_i$ ,  $i = 1, \dots, 5$ , shown in Figure 5a. Given noisy observations  $y_\ell^{\text{obs}}$ ,  $\ell = 1, \dots, N_S$ , on the boundary  $\Gamma = \partial\Omega$  of  $\Omega$ , we seek to reconstruct  $u$  inside  $\Omega$ .

To generate the perturbed observations  $y_\ell^{\text{obs}}$ ,  $\ell = 1, \dots, N_S$ , we add 20% white noise to the numerical solutions of (2.1) with  $g = 0$  for eight separate smoothed Gaussian point sources  $f$ , each centered at a location shown in Figure 5b. Both the observations  $y_\ell^{\text{obs}}$  and the scattered wave fields  $y_\ell = y_\ell[w]$ , needed for the inversion procedure, are discretized with a  $\mathcal{P}^3$ -FEM using about 10 grid points per wavelength. To avoid any inverse crime, the observations  $y_\ell^{\text{obs}}$  are computed on a different, about 30% finer mesh.

We seek our approximation  $u_h$  of the exact medium  $u$  in the  $\mathcal{P}^1$ -FE space for a regular, triangular mesh with 320'801 vertices and 640'000 elements. Since  $u$  is assumed known on the

(a) misfit  $\mathcal{J}$  vs. ASI iteration(b) relative  $L^2$  error vs. ASI iteration

(c) frequency vs. ASI iteration



(d) number of basis functions vs. ASI iteration

Figure 6: Five simple geometric inclusions. A change in marker indicates a change in frequency.

boundary  $\partial\Omega$  of  $\Omega$ , where it is constant, we choose as initial guess  $u_h^{(0)}$  constant throughout  $\Omega$ ; thus, on  $\partial\Omega$ ,  $u_h^{(0)}$  coincides with  $u$ . To initialize the algorithm we must also choose  $\varphi_0^{(1)}$  and  $\Psi^{(1)} = \text{span}\{\psi_1^{(1)}, \dots, \psi_{J_1}^{(1)}\}$ . To do so, we set  $J_1 = 100$  and solve (2.6) and (2.7). Since  $u_h^{(0)}$  is constant,  $L_\varepsilon[u_h^{(0)}]$  simply reduces to the negative Laplacian times  $\varepsilon^{-1}$ . For Step 8 of the ASI Algorithm, we use (2.17) with  $\rho_0 = 0.8$  and  $\rho_1 = 1.2$ .

The reconstruction  $u_h = u_h^{(m)}$  of the medium at the final ASI iteration, with  $m = 30$ , is shown in Figure 5c. Although corners appear smoothed out in the reconstruction,  $u_h$  captures well the locations and shapes of the inclusions composing  $u$  with a relative  $L^2$ -error  $\|u - u_h\|/\|u\|$  of 4.4%.

In Figures 6a and 6b, we show the misfit  $\mathcal{J}$  of the reconstructed medium  $u_h^{(m)}$  and the relative  $L^2$  error at each ASI iteration. Different markers indicate different frequencies, shown in Figure 6c. Both the misfit and the relative  $L^2$  error decrease until the relative error eventually levels around 0.04. In Figure 6d, we monitor the dimension  $J_m$  of the adaptive space  $\Psi^{(m)}$  at each ASI iteration. Starting at  $J_1 = 100$ , first the dimension  $J_m$  increases until it peaks around  $J_m = 270$  at iteration  $m = 7$ , and eventually decreases to about 50.

The first five basis functions  $\psi_1^{(m)}, \dots, \psi_5^{(m)}$  of the adaptive space  $\Psi^{(m)}$  at the final step,  $m = 30$ , shown in Figure 7, illustrate how the adaptive basis captures the span of the first five lowest eigenfunctions in the AS decomposition.

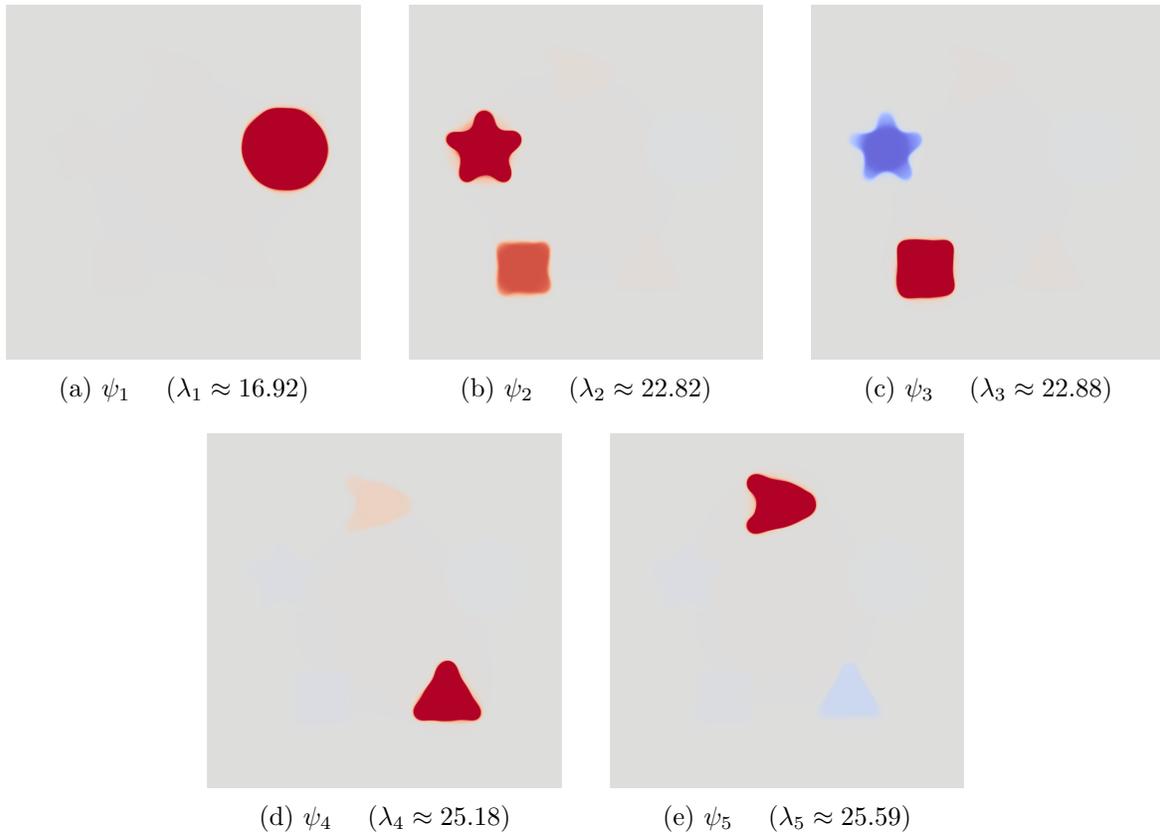


Figure 7: Five simple geometric inclusions: the first five basis functions  $\psi_1, \dots, \psi_5$  and their corresponding Rayleigh quotients  $\lambda_1, \dots, \lambda_5$ , at the final iteration of the ASI method.

## 4.2 Salt-dome model

We consider a two-dimensional (Pluto 1.5) salt dome model from geosciences, generated by the “subsalt multiples attenuation and reduction technology” (SMAART) [2]. Hence, we consider (2.1a) in  $\Omega = (0, 24.4) \times (-9, 0)$  [km] with the squared velocity profile  $u(x) = c^2(x)$  [km/s]<sup>2</sup> shown in Figure 8 (top left). We impose a first order absorbing boundary condition (2.1b) on the two lateral and the lower artificial boundaries, and a homogeneous Neumann condition at the top (physical) boundary  $\Gamma = \{y = 0\} \subset \partial\Omega$ . Synthetic observations at the surface  $\Gamma$  are obtained from the wave fields induced by  $N_s = 100$  source terms located 50 meters beneath the top surface, each about 245 meters apart. Again, we add 20% white noise to the observations.

Now, we apply the ASI Algorithm to reconstruct  $u$  from the surface observations. Here, we seek an approximation  $u_h$  of the unknown medium  $u$  in a  $\mathcal{P}^1$ -FE space with 176’349 vertices and 351’360 elements. The initial guess  $u^{(0)}$ , shown in Figure 8, is generated by extending along the  $x$  direction the known Eastern boundary (borehole) data  $u$  to the entire computational domain. Then, starting with  $\nu = 0.5$  [Hz], the initial guess and the space  $\varphi^{(1)} + \Psi^{(1)}$  spanned by the spectral basis of the negative Laplacian operator with  $J_1 = \dim(\Psi^{(1)}) = 150$ , we solve the inverse problem (2.3) by using the ASI method while progressively increasing the frequency  $\nu$  up to 4.0 [Hz]. In Step 8 of the ASI Algorithm, we set  $\rho_0 = 0.9$  and  $\rho_1 = 1.1$  in (2.17).

After  $m = 27$  ASI iterations, the reconstruction  $u_h = u_h^{(m)}$ , shown in Fig. 8, captures

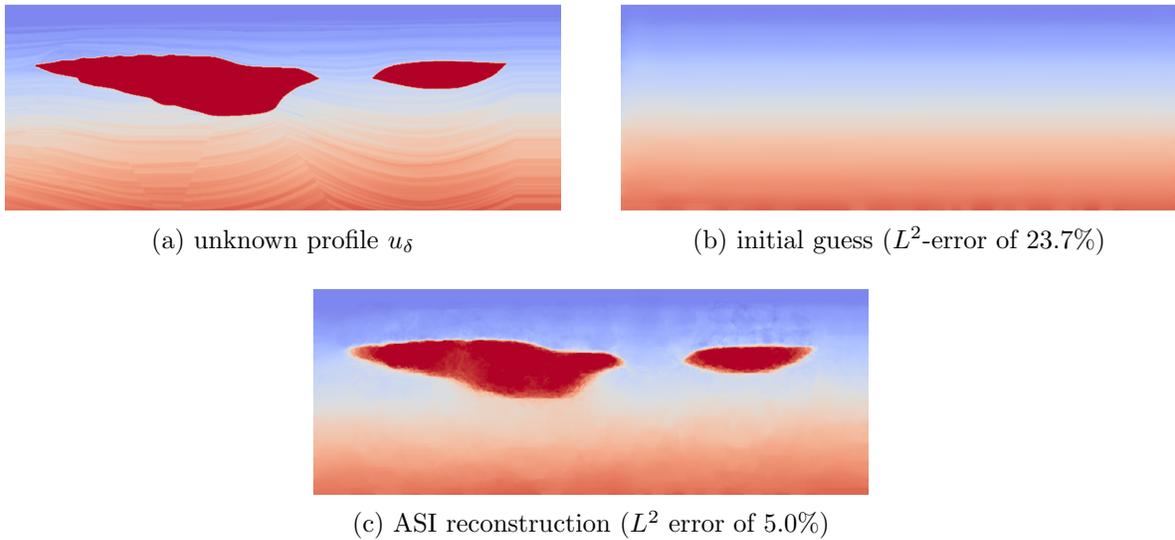


Figure 8: Salt dome model: top left: FE approximation  $u_\delta$  of  $u$ ; top right: initial guess  $u_h^{(0)}$ ; bottom: ASI reconstruction ( $\nu = 4.0[Hz]$ ).

remarkably well the location, size, and inner velocities of the two salt bodies, though not originally present in the initial guess. In Fig. 9, we monitor the misfit  $\mathcal{J}(u_h^{(m)})$ , the dimension  $J_m$  of the search space  $\Psi^{(m)}$ , and the relative  $L^2$ -error at each ASI iteration. Both the relative  $L^2$ -error and the misfit  $\mathcal{J}$  monotonically decrease until the error levels off at about 5%. In Fig. 9d, we observe that the number of basis functions of the AS space  $\Psi^{(m)}$  varies between 150 and 800 basis functions.

## 5 Concluding remarks

Starting from the adaptive spectral (AS) decomposition (1.1)–(1.3), we have proposed a nonlinear optimization method for the solution of inverse medium problems. Instead of a grid-based discrete representation, the unknown medium  $u(x)$  is projected at the  $m$ -th iteration to a finite-dimensional affine subspace  $\varphi_0 + \Psi^{(m)}$ , which is updated repeatedly. The search space  $\Psi^{(m)}$  is constructed by combining the search space from the previous iteration with the "background"  $\varphi_0$ , which satisfies (1.2), and the first  $K$  eigenfunctions  $\varphi_j$  of a judicious linear elliptic operator  $L_\varepsilon$ . Since  $L_\varepsilon$  depends itself on the current iterate,  $\varphi_0$  and the orthonormal basis of eigenfunctions  $\varphi_1, \dots, \varphi_K$  are updated repeatedly. Moreover, the resulting ASI (adaptive spectral inversion) algorithm, listed in Section 2.2, also adapts "on the fly" the dimension  $J_m$  of the search space by solving a small, quadratically constrained, quadratic minimization problem to filter basis functions while preserving important features. Hence the AS decomposition not only substantially reduces the dimension of the search space, but also removes the need for added Tikhonov-type regularization. Our numerical results for the ASI method, when applied to time-harmonic inverse scattering problems governed by the Helmholtz equation, illustrate its accuracy and efficiency even in the presence of noisy or partial boundary data. In particular, the ASI method is able to invert a two-dimensional (Pluto 1.5) salt dome model [2] from noisy surface observations with only a few hundred control variables.

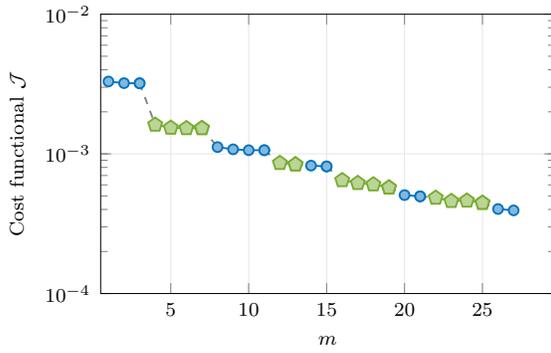
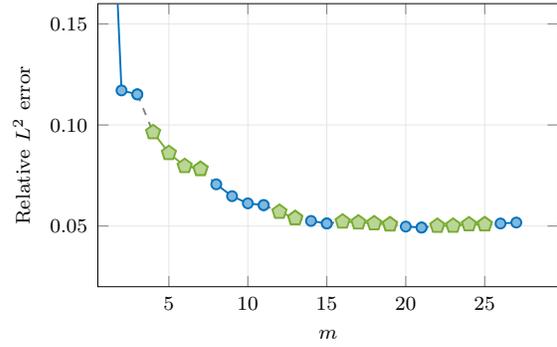
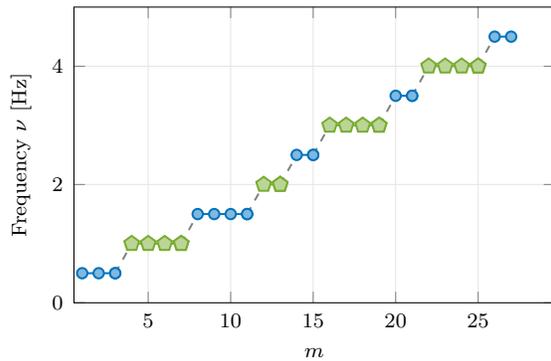
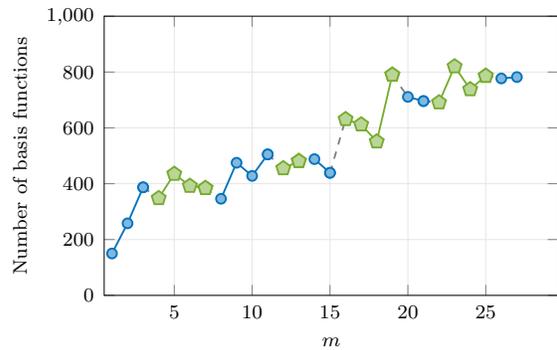
(a) misfit  $\mathcal{J}(u_h^{(m)})$  vs. ASI iteration(b) relative  $L^2$  error in  $u_h^{(m)}$  vs. ASI iteration(c) frequency  $\nu$  vs. ASI iteration(d) number of basis functions  $J_m$  vs. ASI iteration

Figure 9: Salt dome model: misfit (a), relative  $L^2$  error (b), frequency  $\nu$  (c), and number of basis functions (d) in each ASI iteration. A change in marker indicates a change in frequency  $\nu$ .

Our analysis in Section 3 underpins the remarkable accuracy of the AS decomposition observed in practice. In particular, Theorem 5 provides rigorous estimates for the approximation  $\varphi_0$  of the background and for the eigenfunctions  $\varphi_1, \dots, \varphi_K$  of  $L_\varepsilon[u_\delta]$ , when the medium  $u$  consists of  $K$  piecewise constant distinct characteristic functions. Our estimates imply that  $\varphi_0$  and the first  $K$  eigenfunctions  $\varphi_j$  are “almost” constant in each connected component away from interfaces. Hence for small  $\delta$  and  $\varepsilon$ , the background is well approximated by  $\varphi_0$ , whereas the deviation from the background,  $u - \varphi_0$  (or  $u_\delta - \varphi_0$ ), may be well approximated in the span of  $\varphi_1, \dots, \varphi_K$ . The analysis is valid for a wide class of medium-dependent (nonlinear) weight functions  $\mu_\varepsilon$  and also for different types of  $H^1$ -regular approximations  $u_\delta$  of  $u$ . In particular, they also hold for standard  $H^1$ -conforming FE approximations  $u_\delta = u_h$ , where  $\delta$  corresponds to the underlying mesh size  $h$  – see Corollary 6.

The operator  $L_\varepsilon[u_\delta]$  is related to the linearization of the total variation functional [16]. In Remark 7, we discuss similarities between Theorem 5 and the nonlinear spectral theory for the TV functional [5, 3, 1]. The comparison also raises some interesting questions we have not addressed in this work. Since our estimates are uniform in the parameters  $\varepsilon$  and  $\delta$ , they suggest that a limit argument could be used for estimating the asymptotic behavior of the

eigenvalues for  $\varepsilon, \delta \rightarrow 0$ . Interestingly, however, the present estimates are also valid for cases that are specifically excluded by the spectral theory for the TV functional [1]. More precisely, according to [1], if a set  $A$  is not convex, or if its boundary is Lipschitz but not  $C^{1,1}$ , then its characteristic function  $\chi_A$  cannot be an eigenfunction for the eigenvalue (3.45) of the TV functional. Our estimates, however, then still hold true, as supported by our numerical tests in Section 3.4.

Although we have only considered scalar problems here, the AS decomposition can also be used for multiple parameters [18]. As the AS decomposition is independent of the underlying governing PDE, or any particular choice of misfit functional  $\mathcal{J}[u]$ , it is probably also useful for other inverse problems with time-dependent or elliptic governing field equations, or possibly for different applications from image analysis.

## References

- [1] G. Bellettini, V. Caselles, and M. Novaga. The total variation flow in  $\mathbb{R}^N$ . *Journal of Differential Equations*, 184(2):475 – 525, 2002.
- [2] P. Bulant. Sobolev scalar products in the construction of velocity models: Application to model Hess and to SEG/EAGE salt model. *Pure and Applied Geophysics*, 159(7):1487–1506, Jul 2002.
- [3] M. Burger, G. Gilboa, and M. Moeller. Nonlinear spectral analysis via one-homogeneous functionals: Overview and future prospects. *Journal of Mathematical Imaging and Vision*, 56(2):300–319, 2016.
- [4] G. Chavent. *Nonlinear Least Squares for Inverse Problems*. Springer, 2009.
- [5] D. Cremers, G. Gilboa, L. Eckardt, M. Burger, and M. Möller. Spectral decompositions using one-homogeneous functionals. *CoRR*, abs/1601.02912, 2016.
- [6] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11):1413–1457, 2004.
- [7] M. de Buhan and M. Darbas. Numerical resolution of an electromagnetic inverse medium problem at fixed frequency. *Computers and Mathematics with Applications*, 74:3111 – 3128, 2017.
- [8] M. de Buhan and M. Kray. A new approach to solve the inverse scattering problem for waves: combining the TRAC and the adaptive inversion methods. *Inverse Problems*, 29(8):085009, 2013.
- [9] M. de Buhan and A. Osses. Logarithmic stability in determination of a 3D viscoelastic coefficient and a numerical example. *Inverse Problems*, 26(9):095006, 2010.
- [10] M. C. Delfour and J.-P. Zolésio. *Shapes and Geometries*. Vieweg, 2 edition, 2011.

- [11] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Springer-Verlag, 2000.
- [12] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 2010.
- [13] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, 1992.
- [14] H. Federer. *Geometric Measure Theory*. Springer-Verlag, 1969.
- [15] M. Graff, M. J. Grote, F. Nataf, and F. Assous. How to solve inverse scattering problems without knowing the source term: a three-step strategy. *Inverse Problems*, 35:1041001, 2019.
- [16] M. Grote, M. Graff-Kray, and U. Nahum. Adaptive eigenspace method for inverse scattering problems in the frequency domain. *Inverse Problems*, 33:025006, 02 2017.
- [17] M. J. Grote, J. Huber, D. Kourounis, and O. Schenk. Inexact interior-point method for PDE-constrained nonlinear optimization. *SIAM J. Sci. Comp.*, 36(3):A1251–A1276, 2014.
- [18] M. J. Grote and U. Nahum. Adaptive eigenspace for multi-parameter inverse scattering problems. *Computers and Mathematics with Applications*, 2019.
- [19] E. Haber, U. M. Ascher, and D. Oldenburg. On optimization techniques for solving nonlinear inverse problems. *Inverse Problems*, 16:1263, 2000.
- [20] F. J. Herrmann and G. Hennenfent. Non-parametric seismic data recovery with curvelet frames. *Geophys. J. Int.*, (173):233–248, 2008.
- [21] A. Kadu, T. van Leeuwen, and W. A. Mulder. Salt reconstruction in full-waveform inversion with a parametric level-set method. *IEEE Transactions on Computational Imaging*, 3(2):305–315, June 2017.
- [22] B. Kaltenbacher and J. Offtermatt. A convergence analysis of regularization by discretization in preimage space. *Math. Comp.*, 81(280):2049–2069, October 2012.
- [23] Y. Lin, A. Abubakar, and T. M. Habashy. Seismic full-waveform inversion using truncated wavelet representations. pages 1–6, 2012. SEG Annual meeting 2012, Las Vegas.
- [24] I. Loris, H. Douma, G. Nolet, I. Daubechies, and C. Regone. Nonlinear regularization techniques for seismic tomography. *Journal of Computational Physics*, (229):890–905, 2010.
- [25] L. Métivier, R. Brossier, J. Virieux, and S. Operto. Full waveform inversion and the truncated Newton method. *SIAM J. Sci. Comput.*, 35(2):B401–B437, 2013.
- [26] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, 2 edition, 2006.
- [27] S. Operto and J. Virieux. An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74:WCC1–WCC26, 11 2009.

- [28] R. G. Pratt, C. Shin, and G. J. Hicks. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophys. J. Int.*, 133(2):341–362, 1998.
- [29] A. Quarteroni. *Numerical Models for Differential Problems*. Springer, 4 edition, 2008.
- [30] A. Tarantola. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266, 1984.
- [31] A. N. Tikhonov. On the stability of inverse problems. *Dokl. Akad. Nauk SSSR*, 39(5):195–198, 1943.