Eigenvector Model Descriptors for Solving an Inverse Problem of Helmholtz Equation

Hélène Barucq*, Florian Faucher* and Otmar Scherzer†

March 22, 2019

Abstract

We study the inverse problem for the Helmholtz equation using partial data from one-side illumination. In order to reduce the ill-posedness of the problem, the model to be recovered is represented using a limited number of coefficients associated with a basis of eigenvectors, following regularization by discretization approach. The eigenvectors result from a diffusion equation and we compare several choices of weighting coefficient from image processing theory. We first investigate their efficiency for image decomposition (accuracy of the representation with small number of variables, denoising). Depending on the model geometry, we also highlight potential difficulties in the choice of basis and underlying parameters. Then, we implement the method in the context of iterative reconstruction procedure, following a seismic setup. Here, the basis is defined from an initial model where none of the actual structures are known, thus complicating the process. We note that the method is more appropriate in case of salt dome media (which remains a very challenging situation in seismic), where it can compensate for lack of low frequency information. We carry out two and three-dimensional experiments of reconstruction to illustrate the influence of the basis selection, and give some guidelines for applications.

1 Introduction

We consider the inverse problem associated with the propagation of time-harmonic wave which occurs, for example, in seismic application. Following a non-intrusive geophysical setup, it means that we work with partial boundary measurements: the measurements are restricted to be located at the surface of the Earth. Furthermore, they consist in back-scattered (reflection) data, resulting from a one-side illumination only (from the surface). In the last decades, this problem has encountered a growing interest with the increase in numerical capability and the use of supercomputers. However, accurate recovery of deep subsurface structures remain a challenge, due to the nonlinearity and ill-posedness of the problem, the partial reflection data available, and the scale of the domains investigated.

In the context of seismic, the quantitative reconstruction of physical property using an iterative minimization of a cost function originally follows the work of [45, 67, 69] in the time-domain, and [60, 59, 58] for the frequency approach. The method is commonly referred to as Full Waveform Inversion (FWI), we refer to [71] and the references therein for the review of the method. One of the key is that the gradient of the cost function...
is computed using adjoint-state method \cite{47,22,57} to avoid the formation of the (large) Jacobian matrix. Then, Newton-type algorithms represent the traditional framework to perform iterative minimization. We note that, due to the large computational scale of the domain investigated, seismic experiments may have difficulties to incorporate second order (Hessian) information in the algorithm, and one can instead rely on alternative techniques, see, for example, \cite{58,1,24,52,38}. Quantitative (as opposed to qualitative) reconstruction method based upon iterative minimization is naturally not restricted to seismic and we refer, among others, to \cite{7,10} and the references therein for additional applications.

In order to diminish the ill-posedness of inverse problem, it is common to employ regularization. However, it introduces an additional constraint, which may be complicated to select a priori, and problem dependent, e.g., \cite{43,37,42,41} and the references therein. In the regularization by discretization approach, the model representation plays the role of regularizing the functional, by controlling (and limiting) the number of unknowns in the problem, and possibly defining the general shape of the unknown. Constraining the number of unknowns influences the stability and convergence of the procedure, but also the resolution of the outcome. The use of piecewise constant coefficients appears natural for numerical applications, and is also motivated by stability results, cf. \cite{28,13}. But such representations can lead to artificial ‘block’ representation \cite{13,32} which would give a false impression of the resolution; for this reason, piecewise linear model representation is explored in \cite{3,2}, still motivated by stability properties. Wavelets-based model reduction offers a more flexible framework, in particular, wavelets are used for the purpose of regularization in seismic tomography in the work of \cite{50,49}. In \cite{76,77}, FWI is carried out in the time-domain with model represented from a wavelet-based decomposition.

In our work, we will use a model decomposition based upon eigenvectors of a chosen diffusion operator, as introduced in \cite{25,35,36}. Note that this decomposition is shown (with the right choice of operator) to be related with standard Total Variation (TV) or Tikhonov regularization. In our work, we investigate several possibilities for the choice of operator, following image processing techniques, which traditionally also relies on such (diffusion) PDE, see \cite{74}. Therefore, we first investigate the performance of the decomposition depending on the choice of PDE, and then the performance of such model decomposition as parametrization of the reconstruction procedure in seismic FWI.

In geophysics, we specifically have in mind the reconstruction of salt domes (i.e. media with high contrasting objects), which is particularly challenging, because of (in addition to the usual restrictive data) of the change of the kinematics involved, \cite{9}. It has motivated studies based on regularization, \cite{62,40}, or level-sets \cite{39} in seismic. In our experiments, we insist that the presence of the contrasting object is unknown at first, which prevents the use of shape reconstruction techniques. We shall see that the model parametrization from eigenvectors decomposition is well adapted for FWI in this case. We also show limitation of the method, in particular for models of several layered structures, which may fail to be appropriately decomposed onto the basis.

In Section 2 we define the inverse problem associated with the Helmholtz equation and introduce the iterative method for the reconstruction of the wave speed. In Section 3 we review several possibilities for the model decomposition using eigenvectors of diffusion operators. The process of model (image) decomposition is illustrated in Section 4. Then, in Section 5 we carry out the iterative reconstruction using two and three-dimensional FWI experiments. Here, the model decomposition is based upon the initial model, which does not contain a priori information on the target, hence increasing the differences of performance depending on the basis selection. It allows us to identify the best candidate for the recovery
of salt dome and extract some guidelines for reconstruction applications.

2 Inverse time-harmonic wave problem

2.1 Forward problem

We consider a domain $\Omega$ in two or three dimensions, such that $\Omega \in \mathbb{R}^2$ or $\Omega \in \mathbb{R}^3$. The propagation of waves in acoustic medium is given by the scalar pressure field $p$, solution of the Helmholtz equation

$$(-\Delta - \omega^2 c^{-2}(x))p(x) = f(x), \quad \text{in } \Omega,$$

where $c$ is the wave speed, $f$ the source of the phenomenon and $\omega$ the frequency. We still have to specify the boundary conditions to formulate the appropriate problem.

Following a seismic setup, the boundary of $\Omega$, $\Gamma$ is separated into two. The upper (top) boundary, $\Gamma_1$, represents the interface between the Earth and the air, and is subsequently represented via a free surface boundary condition, where the pressure is null. On the other hand, other part of the boundary corresponds to the numerical necessity to restrict the area of interest. Here, conditions must ensure that entering waves are not reflected back to the domain (because the Earth does not end outside our area of interest), see Figure 1. The two most popular formulations to handle such numerical boundary are either the Perfectly Matched Layers (PML, [12]), or Absorbing Boundary Conditions (ABC, [30]). In our case, we use absorbing boundary conditions so that the complete problem writes as

$$\begin{aligned}
(-\Delta - \omega^2 c^{-2}(x))p(x) &= f(x), \quad \text{in } \Omega, \\
p(x) &= 0, \quad \text{on } \Gamma_1, \\
\partial_\nu p(x) - i\omega c^{-1}(x)p(x) &= 0, \quad \text{on } \Gamma_2,
\end{aligned}
$$

where $\partial_\nu$ is the normal derivative.

For the inverse problem, which aim the recovery of the wave speed $c$ in (2), we have access to a discrete set of measurements (i.e. we work with partial data), which correspond to observation of the wave phenomenon. More precisely, our data consist in the measurement of the pressure field solution to (2), at the (discrete) device locations. We refer to $\Sigma$ for this set of positions where the receivers are located, and define the forward map $F$ (which links the model to the data) such that, at frequency $\omega$,

$$F_\omega : m \rightarrow p(x)|_{\Sigma}.$$

We have introduced $m(x) = c^{-2}(x)$, which is also our choice of parameter for the reconstruction, see Remark 2. Note that, in seismic, the data are actually generated from several point sources (excited one by one). All devices (sources and receivers) remain near the surface $(\Gamma_1)$, as illustrated with $\Sigma$ in Figure 1.

2.2 Quantitative reconstruction using iterative minimization

The inverse problem aims the reconstruction of the unknown medium squared slowness $m$ (i.e. the wave speed as $m = c^{-2}$), from data $d_\omega$ given by the forward map $F_\omega(m_\dag)$ obtained from the reference (target) model $m_\dag$:

$$d_\omega = F_\omega(m_\dag) + \mathcal{E}_\omega,$$

where $\mathcal{E}_\omega$ is the error term.
Figure 1: Illustration of the two-dimensional computational domain. The green plus indicates the (point) source and the red dotted line is the location of the discrete measure points. The complete acquisition consists in several sources, all are located at the same depth, only the lateral position varies; the measure points are fixed and do not depend on the source position.

where $\mathcal{E}$ represents the noise in the measures (from devices accuracy, model error, etc).

We employ quantitative method for the reconstruction, usually referred to as Full Waveform Inversion (FWI) in seismic, e.g., [67, 58]. It relies on an iterative minimization algorithm where the cost function is defined as the difference between the observed data and computational simulations from an initial model:

$$
\min_m J(m) = \frac{1}{2} \sum_\omega \|F_\omega(m) - d_\omega\|^2. 
$$

Note that in Problem 5, we have selected the standard least squares minimization but alternatives have also been studied, for example, in the context of seismic, we refer to [65, 61, 11, 61, 18, 15, 51, 75].

**Remark 1** (Multi-frequency algorithm). For the choice of frequency in Problem 5, applications commonly use sequences of increasing frequency during the iterative process, e.g., [19, 60, 66, 17, 10, 32, 9] and the references therein. Namely, one starts with low frequency and minimize for the fixed frequency content. Then, after convergence is obtained, or after a prescribed number of iterations, the frequency is updated (increased) and the iterations continue, cf. Algorithm 2.

**Remark 2** (Parametrization of the unknown). For the reconstruction, we invert the squared slowness $m = c^{-2}$ instead of the velocity. The choice of parameter is first motivated by the Helmholtz equation (2). However, this choice (i.e., velocity, slowness or squared slowness) can lead to important difference in the efficiency of the reconstruction procedure. It is discussed, for example, in [55, 10, 44]; in particular, we motivate our choice from reconstruction comparisons provided in the context of seismic inverse problem in [32, Section 5.4].

The resolution of Problem 5 follows an iterative minimization, traditionally in the Newton method framework. Starting with an initial guess $m^{(0)}$, the model is updated at each iteration $k$ ($k > 0$), using a search direction $s^{(k)}$, such that

$$
m^{(k+1)} = m^{(k)} + \mu s^{(k)}. 
$$

Several possibilities exist for the search direction (e.g. Newton, Gauss-Newton, BFGS, gradient descent, etc.) and we refer to [53] for extensive review of methods. Here, $\mu$ is a scalar coefficient which is approximated using line search algorithm, [53]. In our implementation,
we use gradient-based optimization, because of the numerical burden of computing second order derivative (as mentioned in introduction, geophysical applications can avoid the use of Hessian for this reason). Namely, we use the nonlinear conjugate gradient method and backtracking line search (see [53]), as indicated in Remark [6]. Review of the performance of first order based minimization and influence of line search step selection is investigated in [10], in the context of inverse scattering.

For the computation of the gradient of the misfit functional, we employ the adjoint-state method, which avoids the formation of the dense Jacobian. This method arose in the work of [47] with early application in [22]. It is now relatively standard for the seismic inverse problem, and is reviewed in [57], specificities for complex valued fields can be found in [9].

3 Regularization by discretization: model decomposition

In this section, we introduce a representation of the unknown, i.e., model decomposition, based upon the eigenvectors of the diffusion equation. The objective is to reduce the dimension of the unknown to mitigate the ill-posedness of the inverse problem. We provide several possibilities for the choice of eigenvectors, following image processing literature.

3.1 Regularization and diffusion operator

The resolution of the inverse problem using quantitative method introduces an optimization problem [5] where our misfit functional $\mathcal{J}$ only accounts for a fidelity term: the lone objective is to match the observations. Having to solve an optimization problem is also common in the context of image processing (e.g., denoising, edge enhancement) where the fidelity term corresponds to the matching between the original and processed image. Then, it is relatively common (for both quantitative reconstruction methods and in image processing) to incorporate an additional term in the minimization, for the purpose of regularization. This additional term primary function is to reduce the ill-posedness of the problem, by adding constraint, see, e.g., [43, 37, 21, 63, 64, 72, 48, 42, 62, 41]. The regularized minimization problem writes as

$$\min_m \mathcal{J}_r(m) = \frac{1}{2} \sum_\omega \| \mathcal{F}_\omega(m) - d_\omega \|^2 + \mathcal{I}(m),$$

where $\mathcal{I}$ stands for the regularization term.

In the context of image processing, $\mathcal{I}$ is usually defined to only depend on the gradient of the variable (image), such that

$$\mathcal{I}(|\nabla m|) = \int_\Omega \phi(|\nabla m|) d\Omega.$$

The minimization of the regularization term $\mathcal{I}$ can be recast using the Euler–Lagrange formulation, e.g. [31, Chapter 8]. It implies that the minimizer of $\mathcal{I}$ is in fact the solution to the diffusion equation

$$\nabla \cdot \left( \frac{\phi'(|\nabla m|)}{|\nabla m|} \nabla m \right) = 0 \quad \text{in } \Omega.$$

For the sake of clarity, we introduce the following notation:

$$\mathcal{A}(m, \eta) := -\nabla \cdot (\eta(m) \nabla), \quad \text{with} \quad \eta(m) = \frac{\phi'(|\nabla m|)}{|\nabla m|}.$$
In the following, we present several choices for the diffusion PDE coefficient $\eta$, following image processing theory.

**Remark 3.** The minimization of $J_r$ in Problem (7) can be performed using traditional gradient descent or Newton types algorithms. Another alternative, in particular when rewriting with the Euler–Lagrange formulation in the context of image processing, is to recast the problem as a time dependent evolution, see, e.g., [74, 20, 64, 4].

**Remark 4.** The diffusion equation (9) is obtained using the fact that $\phi = \phi(\nabla m)$ only depends on $\nabla m$. In case of dependency of the function with $m$, or higher order derivatives, the Euler–Lagrange formulation must obviously be adapted.

### 3.2 Diffusion coefficient from image processing

There exist several possibilities for the choice of diffusion coefficient $\eta$ (also referred to as *weighting function*) in (10), inherited from image processing theory and applications. In the following, we investigate the most common formulations, see Table 1 for which we have mainly followed the ones that are reviewed in [14, Table 1] and [63]. Furthermore, we incorporate a scaling coefficient $\beta > 0$ for the diffusion coefficient, which impacts the magnitude.

For consistency in the different models we investigate, the norms that we employed are scaled with the maximal values so that they remain between 0 and 1; we define

$$
\nabla_{1}\!m(x) = \frac{|\nabla m(x)|}{\max(|\nabla m(x)|)}, \quad \nabla_{2}\!m(x) = \frac{|\nabla m(x)|^2}{\max(|\nabla m(x)|^2)},
$$

with $|\nabla m(x)| = \sqrt{\sum_{k=1}^{d} \left(\frac{\partial m}{\partial x_k}\right)^2}$,  \hspace{1cm} (11)  

where $d$ is the space dimension (we restrict ourselves to $d = 2$ or $d = 3$ in our experiments). In order to clarify the expression, we will omit the space dependency in the following definitions.

**Remark 5.** We can make the following comments regarding the nine diffusion coefficients that are introduced in Table 1:

- **The PDE (9)** using the Tikhonov diffusion coefficient $\eta_9$ coincides with the Laplacian equation.

- For the formulation of $\eta_4$ and $\eta_8$, we have to impose a threshold as the coefficient is not defined for the points where the gradient is zero. In the computations, we impose that $\eta_4 = \eta_8 = 1$ for the points $x_i$ where $\nabla m(x_i) < 10^{-12}$.

- The first Perona–Malik formula $\eta_1$, is very similar to the Lorentzian approach, $\eta_6$: only the position of $\beta$ differs. Namely, the Perona–Malik formula would rather use small $\beta$ while the Lorentzian formula would use large $\beta$.

- The second Perona–Malik formula, $\eta_2$, is very similar to the Gaussian criterion $\eta_7$, which only includes an additional dependency on $\beta$.

- The formulation of $\eta_8$ corresponds to the Total Variation (TV) regularization, cf. [72].
| reference (name) | definition | $\beta \to 0$ | $\beta \to \infty$ |
|-----------------|------------|---------------|-----------------|
| [55, 56, 74] (Perona–Malik 1) | $\eta_1(m, \beta) = \frac{\beta}{\beta + \nabla |_2 m}$ | 0 | 1 |
| [55, 56] (Perona–Malik 2) | $\eta_2(m, \beta) = \exp\left(-\frac{\nabla |_2 m}{\beta}\right)$ | 0 | 1 |
| [33] (Geman–Reynolds) | $\eta_3(m, \beta) = \frac{2\beta}{(\beta + \nabla |_2 m)^2}$ | 0 | 0 |
| [34] (Green) | $\eta_4(m, \beta) = \tanh\left(\frac{\nabla |_1 m}{\beta}\right)\left(\frac{1}{\beta \nabla |_1 m}\right)$ | $+\infty$ | 0 |
| [21] (Charbonnier–Blanc-Féraud–Aubert–Barlaud) | $\eta_5(m, \beta) = \frac{1}{\beta}\left(\beta + \nabla |_2 m\right)^{-1/2}$ | $+\infty$ | 0 |
| [36] (Lorentzian) | $\eta_6(m, \beta) = \frac{\beta}{(1 + \beta \nabla |_2 m)^2}$ | 0 | 0 |
| [36] (Gaussian) | $\eta_7(m, \beta) = \left(\beta \exp\left(\frac{\nabla |_2 m}{\beta}\right)\right)^{-1}$ | 0 | 0 |
| [36] (Rudin–Osher–Fatemi) | $\eta_8(m) = \frac{1}{\nabla |_1 m}$ | n/a | n/a |
| (Tikhonov) | $\eta_9 = 1$ | n/a | n/a |

Table 1: List of formula used for the coefficient in the diffusion operator, we refer to [14, 63], where the associated function $\phi$ in (10) can also be found.

### 3.3 Eigenvector model decomposition in FWI

In our work, we employ *regularization by discretization*: instead of adding the regularization term $I$ in the minimization problem, we remain with Problem [5], and use specific representation for the model (unknown). We follow the work of [27, 26, 35, 36] with the “Adaptive Inversion” or “Adaptive Eigenspace Inversion” method. Namely, the unknown is represented via a decomposition into the basis of eigenvectors computed from a diffusion PDE. The purpose is to control the number of unknown in the representation, and consequently reduce the ill-posedness of the inverse problem. The decomposition uses the steps (given in [35]) depicted Algorithm [1].

Following Algorithm 1 we introduce the notation,

$$\psi(m, \eta, N) = \{\psi_k\}_{k=1,...,N}$$

the set of $N$ vectors associated with model $m$ and diffusion coefficient $\eta$, computed from (12) and (13): 

and

$$m(m, N, \psi)$$

is the decomposition of model $m$ using $N$ vectors of $\psi$, (14). 

Therefore, the model is represented via $N$ coefficients $\alpha$ of (14) in the basis given by the diffusion operator. The reconstruction procedure follows an iterative minimization of (5).

7
Eigenvector decomposition: given an initial model $m(x)$, a selected integer value $N > 0$, and the choice of diffusion coefficient $\eta$.

1. Compute $m_0$ solution to
   \[
   \begin{align*}
   &\mathcal{A}(m, \eta)m_0 = 0, \quad \text{in } \Omega, \\
   &m_0 = m, \quad \text{on } \Gamma. 
   \end{align*}
   \tag{12}
   \]

2. Compute the subset of $N$ eigenvectors $\{\psi_k\}_{k=1,\ldots,N}$ which are associated to the $N$ smallest eigenvalues $\{\lambda_k\}_{k=1,\ldots,N}$ such that, for all $k$,
   \[
   \begin{align*}
   &\mathcal{A}(m, \eta)\psi_k = \lambda_k \psi_k, \quad \text{in } \Omega, \\
   &\psi_k = 0, \quad \text{on } \Gamma. 
   \end{align*}
   \tag{13}
   \]

3. Compute the model decomposition using $N$ eigenvectors:
   \[
   m = m_0 + \sum_{k=1}^{N} \alpha_k \psi_k, 
   \tag{14}
   \]
   where $\alpha_k$ is a scalar and $\psi_k$ a vector. Here, the set of $\alpha$ is chosen to minimize $\|m - m_0\|^2$; and the $\psi_k$, $k = 1, \ldots, N$ are the eigenvectors associated with the $N$ smallest eigenvalues $\lambda_k$, computed in Step 2.

Algorithm 1: Model decomposition using eigenvectors of the diffusion equation associated to the smallest eigenvalues. We refer to the model decomposition as $m(m, N, \psi)$ where $\psi(m, \eta, N)$ is the set of eigenvectors, see \([15]\) and \([16]\).

and performs successive update of the coefficients $\alpha$. The key is that $N$ is much smaller than the dimension of the original representation of $m$, but allows an accurate resolution, as we illustrate in Sections 4 and 5. Algorithm 2 details the procedure.

Remark 6 (Minimization algorithm). For the minimization procedure depicted in Algorithm 2, we use a non-linear conjugate gradient method for the search direction. This method has the advantage that it only necessitates the computation of the gradient of the cost function, \([53]\). Then, to control the update step $\mu$ in Algorithm 2, line search algorithm is typically employed, e.g., \([29, 53, 23, 10]\). This operation is complex in practice because an accurate estimation would require intensive computational operations (with an additional minimization problem to solve). Here, we employ a simple backtracking algorithm, see \([53]\).

Remark 7 (Gradient computation). The gradient of the cost function is computed using the first order adjoint-state method, \([47, 22]\), which is standard in seismic application, see \([57]\). It avoids the formation of dense Jacobian matrix and instead requires the resolution of an additional PDE, which is the adjoint of the forward PDE, with right-hand sides defined from the difference between measurements and the simulation, see \([7, 22, 10, 9]\) for detailed computation technique.

In our implementation, the gradient is first computed with respect to its original (nodal) representation and we use the chain rule to retrieve the gradient with respect to the decomposition coefficient $\alpha$:

\[
\frac{\partial J}{\partial \alpha} = \frac{\partial J}{\partial m} \frac{\partial m}{\partial \alpha}. 
\tag{20}
\]

It is straightforward, from \([14]\), that the derivation for a chosen coefficient $\alpha_l$ gives $\partial_{\alpha_l} J = \psi_l$. Therefore, it is computationally easy to introduce the formulation with respect to the
**Initialization:** multi-frequency measurements \(d_\omega\); initial model \(m^{(0)}\); number of iterations \(n_{\text{iter}}\); list of frequencies \(\omega_i, i = 1, \ldots, n_\omega\); decomposition dimension associated with frequency: \(N_i, i = 1, \ldots, n_\omega\).

Using Algorithm 1,

1. compute the basis associated with \(m^{(0)}\) and selected \(\eta\), using the highest decomposition dimension \(N_{\text{max}} = \max\{\{N_i\}^{n_\omega}_{i=1}\}\),

\[
\psi_{\text{loc}} = \psi(m, \eta, N_{\text{max}}). \tag{17}
\]

2. Decompose the initial model using the initial decomposition dimension:

\[
m^{(0)} = m(m^{(0)}, N_1, \psi_{\text{loc}}) = m_0 + \sum_{l=1}^{N} \alpha_l^{(0)} \psi_l. \tag{18}
\]

**Algorithm 2:** Iterative minimization algorithm (FWI) using model decomposition to control the number of variables. The model is represented with the eigenvector basis, for which the weights are updated along with the iterations.

**Remark 8.** In Algorithm 2, the basis of eigenvectors remain the same for the complete set of iterations, and is extracted from the initial model. Only the number of basis taken for the representation, \(N_i\), changes. Namely, from \(N_1\) to \(N_2 > N_1\), the decomposition using \(N_2\) still has the same \(N_1\) first eigenvectors in its representation (with different weights \(\alpha\)), and additional \((N_2 - N_1)\) eigenvectors. As an alternative, we investigate the performance of an algorithm where the basis changes at each frequency (i.e. it is recomputed from the current iteration model), see Appendix A.
3.4 Numerical implementation

Our code is developed in Fortran90, it uses both mpi and OpenMP parallelism and run on cluster1 for efficiency. The forward wave operator is discretized using a Finite Differences scheme, e.g. 70, 54, 73. The discretization of the Helmholtz operator generates a large sparse matrix, for which we use the direct solver MUMPS (5, 6) for its factorization and the resolution of linear system. Note that this solver is particularly optimized and designed for this type of linear algebra problems, i.e. large, sparse matrices. Our preference for direct solver instead of iterative ones is mainly motivated for two reasons:

- seismic acquisition is composed of a large amount of sources, i.e. a large amount of right-hand sides (rhs) to be processed for the linear system. Using direct solver, the resolution time is very low once the factorization is performed, hence it is well adapted for the multi-rhs seismic configuration.

- For the minimization algorithm, the gradient is computed via the adjoint-state method (see Remark7). It means that an additional linear system has to be solved, which is actually the adjoint of the forward one. Here, the factors obtained from the factorization of the forward operator can be directly reused, and allow a reduced computational cost, see [10] and the references therein.

The next step is the computation of the eigenvectors associated with the smallest eigenvalues for the diffusion operator. We use the package ARPACK2, which is devoted to solve large sparse eigenvalue problems using iterative methods. More precisely, it uses implicitly restarted Lanczos or Arnoldi methods, respectively for symmetric and non-symmetric matrices, [46]. Several options are available in the package, including the maximum number of iterations allowed, or a tolerance parameter for the accuracy of acceptable solution. In particular, we have observed important time cost reduction when allowing some flexibility in the accuracy with this threshold criterion. However, in the computational experiments presented in this paper, we do not use this option, as the numerical efficiency is not the primary objective.

Remark 9 (Efficient computation of the eigenvectors associated with the lowest eigenvalues). The Lanczos and Arnoldi methods are particularly efficient to compute the largest eigenvalues and associated eigenvectors of the matrix, and only require matrix vector multiplication. However, we are interested in the lowest eigenvalues for our decomposition. The idea is simply to use that the lowest eigenvalues of the discretized diffusion matrix, say A, are simply the largest eigenvalues of the matrix $A^{-1}$. Then, the matrix-vector multiplication, say $Av$ for a vector $v$, becomes a resolution of a linear system $A^{-1}v$. It may appear computationally expensive but it is not thanks to the use of the direct solver MUMPS (see above), which, once the factorization is obtained, is very efficient for the resolution procedure. Hence, the computation follows the steps:

1. The experiments have been performed on the cluster PlaFRIM (Plateforme Fédérative pour la Recherche en Informatique et Mathématiques, https://www.plafrim.fr/fr) with the following node specification: 2 Dodeca-core Haswell Intel Xeon E5–2680 v3 (2.5GHz); 128Go RAM; Infiniband QDR TrueScale: 40Gbs$^{-1}$, Omnipath 100Gbs$^{-1}$.
2. ARPACK uses sequential computation, hence, contrary to the rest of our code, this part does not use parallelism. Future developments include the implementation of the parallel version of the package: PARPACK.
3. ARPACK has the possibility to compute the smallest eigenvalues using matrix-vector multiplication, however, we have observed a drastic increase of the computational time compared to using the inverse matrix and resolution of linear system.
1. compute the (sparse) matrix discretization of the selected diffusion operator: \( A \);
2. compute the factorization of the matrix \( A \) using MUMPS,
3. use the package ARPACK to compute the largest eigenvalues of \( A^{-1} \), by replacing the matrix-vector multiplication step in the iterations by the resolution of a linear system using MUMPS.

Finally, the last step is to retrieve the appropriate coefficients \( \alpha_k \) in (14) for the decomposition. It basically consists in the resolution of a dense linear system (from least squares method). We use LAPACK, \([8]\) (contrary to MUMPS, LAPACK is adapted to dense linear system). Note that, because we usually consider a few hundreds of coefficients for the decomposition, this operation remains relatively cheap compared to the eigenvectors computation. We compare the computational time for the eigenvectors computation and model decomposition in Figure 2 for different values of \( N \) and \( \eta \). We first note that the choice of \( \eta \) does not really modify the computational time. Then, we see that the two operations are mostly linear in \( N \), and that the time to solve the least squares problem with LAPACK is much smaller than the time to compute the eigenvectors with ARPACK, namely, hundred time smaller. For the largest case: \( N = 1000 \) for a squared matrix of size 277,221, it takes about 30min to retrieve the eigenvectors, and 10 s to compute the \( \alpha \).

4 Illustration of model decomposition

First, we illustrate the eigenvector model decomposition with geophysical media in two dimensions. The original model is represented on a structured grid by \( n_x \times n_z \) coefficients, and we have here \( 921 \times 301 = 277,221 \) coefficients. We consider two media of different nature:

- the Marmousi velocity model, which consists in structures and faults, see Figure 3(a);
- a model encompassing salt domes: objects of high contrast velocity, see Figure 3(b).

![Figure 2](image.png)

(a) Computational time to retrieve the eigenvectors associated with the lowest eigenvalues, (13), using sequential ARPACK.

(b) Computational time to obtain the coefficients \( \alpha \) in (14) using least squares method and sequential LAPACK.

Figure 2: Computational time for the eigenvector decomposition for different \( \eta \). The computation of eigenvectors uses discretized matrix of the diffusion operator of size 277,221 × 277,221. It corresponds with the decomposition of model Figure 3(b) which is further illustrated in Section 4.
The two models are of the same size (9.2 × 3 km) and use the same number of coefficients for their representations, in order to produce consistent comparisons.

Figure 3: Seismic velocity models used to illustrate the eigenvectors decomposition. They are both of size 9.2 × 3 km and represented with 921 × 301 nodal coefficients.

In this section, we perform the decomposition of the models with application of Algorithm 1 and steps (12), (13) and (14). We study the main parameters that affect the decomposition:

- the choice of \( \eta \), with possibilities given in Table 1,
- the choice of scaling parameter \( \beta \) in the formulation of \( \eta \) (Table 1),
- the number of eigenvectors \( N \) employed for the decomposition in (14).

The accuracy of the decomposition is estimated using the \( L^2 \) norm of the relative difference between the decomposition and the original representation, we define

\[
E = 100 \frac{\|m - m\|}{\|m\|}, \quad \text{Relative Error} \ (%),
\]

where \( m \) is the original model (Figure 3) and \( m \) the decomposition using the eigenvectors basis from Algorithm 1.

### 4.1 Decomposition of noise-free model

We decompose the salt and Marmousi models using the nine possibilities for \( \eta \), that are given in Table 1. For the choice of scaling coefficient \( \beta \) (which does not affect \( \eta_8 \) and \( \eta_9 \)), we roughly cover an interval from \( 10^{-7} \) up to \( 10^6 \), namely: \( \{10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 5 \times 10^{-2}, 10^{-1}, 5 \times 10^{-1}, 1, 5, 10, 10^2, 10^3, 10^4, 10^5, 10^6\} \). In Tables 2 and 3, we show the best relative error (i.e. minimal value) obtained for the Marmousi model of Figure 3(a) and the salt model of Figure 3(b). We test all choices of \( \eta \) and values of \( N \) between 10 (coarse) and 500 (refined). The corresponding values of the scaling parameter \( \beta \) which gives the best (i.e. the minimal) error are also given in parenthesis.

As expected, we observe that the more eigenvectors are chosen (higher \( N \)), the better will be the decomposition. When using 500 eigenvectors, which represents about 2% of the original number of coefficients (921 × 301 in Figure 3), the error is of a few percent only. This can be explained by the redundancy of information provided by the original fine grid where the model is represented (e.g. the upper part of Figure 3(b) and the three salt bodies are basically constant). Comparing the methods and models, we see that
\[ \beta \]

- The Marmousi model (Table 2) is harder to retrieve than the salt model (Table 3) as it gives higher error. In particular for low \( N \), the salt model can be acutely decomposed (possibly 3% error with \( N = 10 \)).

- For both models, it appears that four methods stand out: \( \eta_1 \) (Perona–Malik), \( \eta_3 \) (Lorentzian), and \( \eta_6 \) (Lorentzian), with a slight advantage towards \( \eta_1 \).

- The scaling coefficient that minimizes the error is consistent with respect to \( N \), with similar amplitude. However, changing the model may require the modification of \( \beta \): between the salt and Marmousi decomposition, \( \eta_1 \) and \( \eta_6 \) have to adapt \( \beta \).

To investigate further the last point, we show the evolution of relative error \( \mathcal{E} \) with respect to the scaling coefficient \( \beta \) for the decomposition of the Marmousi and salt models in Figure 4, where we compare four selected formulations for \( \eta \). We observe some flexibility in the choice of the optimal \( \beta \) for \( \eta_1 \) and \( \eta_5 \). On the other hand, \( \eta_3 \) and \( \eta_6 \) show sharper functions, which mean that the selection of \( \beta \) must be more careful in these cases. In addition, the range of efficient \( \beta \) changes depending on the model decomposed, except for \( \eta_5 \). It demonstrates that the choice of \( \beta \) for optimality is not trivial in general, and is model dependent.
Figure 4: Decomposition of the Marmousi and salt velocity models of Figures 3(a) and 3(b) using \( N = 50 \) eigenvectors and following Algorithm 1. The relative error is computed from (21) for four selected formulation of \( \eta \) (see Table 1) and different scaling parameter \( \beta \).

We can eventually pictures the resulting images obtained after the decomposition of both models. For the sake of clarity, these are in Appendix B, Section B.1, where we show the decomposition of Marmousi using \( N = 50 \) (Figure 22) and the decomposition of the salt using \( N = 20 \) (Figure 23). The pictures illustrate correctly the observations of the tables and the differences between the formulation. For the salt model (Figure 23), we notice that the upper boundary is usually well defined but the bottom of the contrasting object appears smooth for all methods, at the exception of the formulation \( \eta_5 \), and slightly behind, \( \eta_1 \). Those two formulations for the diffusion coefficients clearly stand out. For the decomposition of the Marmousi model, Figure 22, which requires more eigenvectors, it appears that \( \eta_1 \), \( \eta_3 \) and \( \eta_6 \) give the best results and are the only ones able to capture the model structures, they are indeed the ones giving the lowest errors in Table 2.

4.2 Decomposition of noisy model

We incorporate noise in the representation, motivated as we aim to implement the technique for inverse problem, where few information on the target is initially known. Hence, we reproduce the model decomposition, this time incorporating noise in the original pictures. For every nodal velocity (of Figure 3), we recast the values using an uniform distribution that covers \( \pm 20\% \) of the noiseless value. The resulting media are illustrated in Figure 5.

We apply the model decomposition using the different formulations of \( \eta \) and choice of scaling coefficient \( \beta \), following the procedure employed for the noiseless model. In Tables 4 and 5, we show the evolution of best relative error with \( N \), for the noisy Marmousi and salt...
models respectively. Here, the relative error is computed from the difference between the noiseless model, and the decomposition of the noisy one. The objective of the regularization is to preserve the structures while smoothing out the noise effect.

| diffusion coeff. | $N = 10$ | $N = 20$ | $N = 50$ | $N = 100$ | $N = 250$ |
|------------------|---------|---------|---------|---------|---------|
| $\eta_1$        | 16% $(10^{-4})$ | 15% $(10^{-4})$ | 14% $(10^{-7})$ | 13% $(10^{-7})$ | 11% $(10^{-7})$ |
| $\eta_2$        | 16% $(5.10^{-2})$ | 16% $(5.10^{-2})$ | 15% $(5.10^{-2})$ | 14% $(5.10^{-2})$ | 13% $(5.10^{-2})$ |
| $\eta_3$        | 16% $(10^{-2})$ | 14% $(10^{-3})$ | 13% $(10^{-4})$ | 12% $(10^{-4})$ | 10% $(10^{-4})$ |
| $\eta_4$        | 16% $(10^1)$ | 16% $(10^1)$ | 16% $(5.10^1)$ | 15% $(5.10^1)$ | 14% $(10^2)$ |
| $\eta_5$        | 16% $(10^{-6})$ | 16% $(10^{-7})$ | 15% $(10^{-7})$ | 15% $(10^{-7})$ | 14% $(10^{-6})$ |
| $\eta_6$        | 16% $(5.10^2)$ | 14% $(5.10^2)$ | 13% $(10^4)$ | 12% $(10^4)$ | 10% $(10^4)$ |
| $\eta_7$        | 16% $(5.10^{-2})$ | 16% $(5.10^{-2})$ | 15% $(5.10^{-2})$ | 14% $(5.10^{-2})$ | 13% $(5.10^{-2})$ |
| $\eta_8$        | 17% (n/a) | 16% (n/a) | 16% (n/a) | 15% (n/a) | 14% (n/a) |
| $\eta_9$        | 17% (n/a) | 17% (n/a) | 16% (n/a) | 16% (n/a) | 14% (n/a) |

Table 4: Minimal relative error obtained and associated scaling coefficient: $E(\beta)$. The error is computed with respect to the noiseless model Figure 3(a) but the decomposition uses the noisy model of Figure 5(a). The definition of $\eta$ is given Table 1.

The decomposition of noisy pictures requires more eigenvectors for an accurate representation. Then, the salt model, with high contrast objects, still behaves much better than the many structures of the Marmousi model. For the decomposition of the noisy Marmousi model, none of the formulations really stands out and the error never reach below 10% using at most $N = 250$.

In Figures 6 and 7, we picture the resulting decomposition for the two media. For the decomposition of the Marmousi model, we use $N = 200$; and $N = 50$ for the salt model. It corresponds with higher values compared to the pictures show for the noiseless models (Figures 22 and 23).

The decomposition of the salt model remains acceptable, and we easily distinguish the main contrasting object. The smaller objects also appear, in a smooth representation. The formulations using $\eta_1$, $\eta_3$ and $\eta_6$ provide sharper boundary for the contrasting objects, in particular for the upper interface. Regarding the decomposition of the noisy Marmousi model, it illustrates the limitation of the method, where none of the formulations is really able to reproduce the structures, and most edges are lost. In particular, the central part of the model is mostly missing and the amplitude of the values has been reduced. It seems
Table 5: Minimal relative error obtained and associated scaling coefficient: $E(\beta)$. The error is computed with respect to the noiseless model Figure 3(b) but the decomposition uses the noisy model of Figure 5(b). The definition of $\eta$ is given Table 1.

| diffusion | $N = 10$ | $N = 20$ | $N = 50$ | $N = 100$ | $N = 250$ |
|-----------|---------|---------|---------|----------|----------|
| $\eta_1$  | 14% $(10^{-6})$ | 14% $(10^{-6})$ | 11% $(10^{-5})$ | 9% $(10^{-5})$ | 5% $(10^{-5})$ |
| $\eta_2$  | 17% $(5.10^{-2})$ | 15% $(5.10^{-3})$ | 11% $(5.10^{-3})$ | 8% $(5.10^{-2})$ | 6% $(5.10^{-2})$ |
| $\eta_3$  | 10% $(10^{-4})$ | 10% $(10^{-4})$ | 8% $(10^{-4})$ | 6% $(10^{-4})$ | 5% $(10^{-4})$ |
| $\eta_4$  | 18% $(10^{-3})$ | 15% $(1)$ | 12% $(10^{-1})$ | 10% $(10^{-3})$ | 6% $(10^{-3})$ |
| $\eta_5$  | 17% $(10^{-5})$ | 15% $(5.10^{-3})$ | 12% $(10^{-4})$ | 9% $(10^{-1})$ | 6% $(10^{-5})$ |
| $\eta_6$  | 10% $(10^0)$ | 9% $(10^2)$ | 8% $(10^4)$ | 6% $(5.10^2)$ | 5% $(10^2)$ |
| $\eta_7$  | 17% $(5.10^{-2})$ | 15% $(10^0)$ | 11% $(10^{-2})$ | 8% $(5.10^{-2})$ | 6% $(5.10^{-2})$ |
| $\eta_8$  | 18% (n/a) | 16% (n/a) | 12% (n/a) | 9% (n/a) | 6% (n/a) |
| $\eta_9$  | 21% (n/a) | 15% (n/a) | 13% (n/a) | 11% (n/a) | 8% (n/a) |

that $\eta_1$, $\eta_3$ and $\eta_6$ are slightly more robust and gives (relatively speaking) the best results. To conclude, these three formulations appear less sensitive (for both media) to noise than the other choices.

Figure 6: Decomposition of the Marmousi velocity model of Figure 5(a) using $N = 200$ and the formulations of $\eta$ from Table 1. The selected value of $\beta$ for every method corresponds to the value given in Table 4. The color scale follows the one of Figure 5(a) with values between 1500 and 5500 m s$^{-1}$.

Regarding image decomposition, we can conclude that the eigenvector decomposition

- is very efficient to represent media with contrasting shapes (e.g., salt domes), even if noise is contained in the images.
In this case, the choice of $\eta$ does not really affect the representation of the objects, and all methods behave well, see Figure 7.

The performance of the decomposition depends on the media, and diminishes with structured layered such as the Marmousi model.

In this case and without noise, an appropriate choice of formulation ($\eta_1$, $\eta_3$, $\eta_6$ from Table 2) can provide the accurate representation. However, when incorporating noise, the performance further deteriorates, and, by smoothing the image, the edges contrast are mostly lost by the decomposition.

5 Experiment of reconstruction with FWI

In this section, we show numerical experiments of reconstruction following Algorithm 2. We focus on models encompassing salt domes, as we have shown that the decomposition is more appropriate to this kind of media. We consider initial guess that has no information on the subsurface structures. We remain in the seismic context, where the forward problem is given from (2) and the data are restricted to be acquired near the surface (see Figure 1). In addition to the challenge of working with backscattered, partial data from one side illumination, we also avoid the use of (unrealistically) low frequencies (i.e. below 2 Hz for seismic application).
5.1 Reconstruction of two-dimensional salt model

We consider the two-dimensional salt model that has been introduced above. It is of size $9.2 \times 3$ km, and we generate the data using 91 sources and 183 receivers (i.e. data points) per source. Both devices are located near the surface: sources are positioned on a line at 20 m depth and receivers at 80 m depth.

In order to establish a realistic situation despite having synthetic experiments, the data are generated in the time-domain and we incorporate white noise in the measurements. The level of the signal to noise ratio in the seismic trace is of 20 dB, the noise is generated independently for all receivers record associated with every source. Then we proceed to the discrete Fourier transform to obtain the signals to be used in the reconstruction algorithm. In Figure 8, we show the time-domain data with noise and corresponding frequency data for one source, located at 20 m depth, in the middle of the $x$-axis.

![Figure 8](image)

(a) Time-domain trace with included noise. (b) Real parts of the discrete Fourier transform.

Figure 8: Data associated with a single centered source. The data are first generated in the time-domain, then we incorporate white noise and proceed to the Fourier transform. In this experiment, the complete seismic acquisition is composed of 91 independent sources and 183 receivers for each source.

For the reconstruction of the salt dome model, the starting and true model are given Figure 9. We do not assume any a priori knowledge of the contrasting object in the subsurface, and start with a one-dimensional variation, which has drastically lower amplitude. Therefore, even the velocity background is unknown at start.

5.1.1 Fixed decomposition, single frequency reconstruction

For the reconstruction, we first only use 2 Hz frequency data, and perform 30 minimization iterations. In Figure 10, we show the reconstruction where the decomposition has not been employed, i.e. the model representation follows the original piecewise constant decomposition of the model (one value per node). In Figure 11, we compare the reconstruction using Algorithms for the different formulations of $\eta$ given in Table 1 using fixed $N = 50$. We use $N = 100$ for Figure 12. For the sake of clarity, we focus on the most effective formulations ($\eta_3$ and $\eta_8$) and move the complete pictures with comparison of all formulations of $\eta$ in Appendix B Figure 24 for $N = 100$. 

18
Figure 9: Target model and starting model for FWI. The models are of size $9.2\text{km} \times 3\text{km}$. The initial model corresponds to a one-dimensional variation.

Figure 10: Reconstruction of the salt velocity model from starting medium Figure 9 using $2\text{ Hz}$ frequency data. The reconstruction does not apply eigenvector decomposition. The model is parametrized following the domain discretization, using piecewise constant representation with one value per node on a $921 \times 301$ grid.

Figure 11: Reconstruction of the salt velocity model from starting medium Figure 9 using $2\text{ Hz}$ frequency data. The eigenvector decomposition employs $N = 50$ and the formulations of $\eta$ from Table 1. It uses the same color scale as Figure 9.
We observe that

- the traditional FWI algorithm (without decomposition), see Figure 10, fails to recover any dome. It only shows some thin layers of increasing velocity, with amplitudes much lower than the original ones.

- The decomposition using $N = 50$ is able to discover the largest object with formulation $\eta_3$, $\eta_5$, $\eta_6$ and $\eta_8$, see Figure 11. Also $\eta_5$ and $\eta_8$ show some artifacts in the lower right corner. Any other decomposition seems to fail. We note that, due to the lack of velocity background information, the positions of the domes are slightly above the correct ones.

- In this experiment, the method behaves much better with restrictive number of eigenvectors. With $N = 100$ (Figures 12 and 24), the reconstruction produces mostly artifacts. The restrictive number of eigenvectors provide a regularization of the problem by reducing the number of parameters, which is fundamental for the convergence.

The decomposition in eigenvectors using noisy images, see Figure 5(b), has provided very similar results for all choices of $\eta$, which were all very effective. However, in the context of quantitative reconstruction for inverse problem, there is a drastic difference between the formulations, and the wrong choice easily leads to the failure of the procedure. In addition, the number of eigenvectors for the representation has to be carefully selected. Note also that the only $\eta$ that were performing well for the Marmousi decomposition ($\eta_1$, $\eta_3$ and $\eta_6$ from Table 2) are not the ones that perform well here ($\eta_3$, $\eta_5$, $\eta_6$ and $\eta_8$), at the exception of $\eta_3$.

5.1.2 Experiments with evolutive decomposition, multiple frequencies

We investigate the performance of the eigenvectors decomposition for multiple frequency data, and with progressive evolution of the number of eigenvectors in the representation $N$. We have a total of four different experiments, which are summarized in Table 6. The reconstructions, for $\eta_3$ and $\eta_8$, are shown Figure 13. The results for all $\eta$ of Table 1 are pictured in Appendix B, Figures 25, 26 and 27.

From these experiments using multiple $N$ and/or frequency contents, we observe the performance of the method.

- The best option appears to use single 2 Hz frequency with progression of increasing $N$: Experiment 2. Compared to our preliminary test (Figure 11), it discovers larger
Table 6: List of experiments for the reconstruction of the two-dimensional salt dome model (Figure 9). For each combination of frequency and associated number of eigenvectors $N$ in the decomposition, 30 iterations are performed ($n_{\text{iter}}$ in Algorithm 2).

| Experiment | Frequency list | $N$ | Total number of iterations | References |
|------------|----------------|-----|---------------------------|------------|
| 1          | 2 Hz           | 50  | 30                        | Figures 11, 24 |
| 2          | 2 Hz           | {50, 60, 70, 80, 90, 100} | 180         | Figures 13, 25 |
| 3          | 2, 3, 4, 5 Hz  | 50  | 120                       | Figure 13, 26 |
| 4          | 2, 3, 4, 5 Hz  | {50, 60, 70, 80}  | 120         | Figure 13, 27 |

Figure 13: Results of Experiments 2, 3 and 4 of Table 6 for the reconstruction of the salt velocity model Figure 9. The comparison of all formulations of $\eta$ (Table 1) is pictured in Appendix B, Figures 25, 26 and 27.

- Using multiple frequencies surprisingly deteriorates the results, Figures 26 and 27. It may be that the object we are looking for contains only one wavelength (one size and velocity variation) such that incorporating other wavelength (in particular shorter) does not add useful information.
Therefore, we deduce that the most robust way to conduct the iterations is to follow the Experiment 2 process: using a single frequency, with a slow increase of $N$ in parallel.

5.1.3 On the choice of eigenfunction number

We have shown in Figures 11 and 12 that taking $N$ too high to start the algorithm leads to the failure of the procedure. It remains to check if the appropriate $N$ can be selected ‘a priori’, or based upon minimal experiments. In Figure 14(a), we show the evolution of cost function with thirty iterations, for different values of $N$, from 10 to 250. We compare, in Figure 14(b) with the progression of $N$, which follows Experiment 2 of Table 6.

![Graph](image)

Figure 14: Evolution of the misfit functional (scaled with the first iteration value) with iterations depending on the choice of $N$, using 2 Hz frequency.

From Figure 14(a) we see that all choices of $N$ have the same pattern: first the decrease of the functional and then its stagnation. Then, we remark that the good choice for $N$ is not reflected by the misfit function. Indeed, it shows lower error for larger $N$, while they are shown to result in erroneous reconstruction (Figure 12 compared to Figure 11). It is most likely that using larger $N$ leads to local minima and/or deteriorates the stability (see, for the piecewise constant case, [13]). It results in the false impression (from the misfit functional) that it would improve the reconstruction. Using progression in $N$, Figure 14(b) eventually gives the same misfit functional value than the large $N$, but it needs more iterations. This increase of iterations and ‘slow’ convergence is actually required, because it leads to an appropriate reconstruction, see Figure 13.

Therefore, we cannot anticipate a good choice for $N$ a priori (with a few evaluation of the misfit functional). the guideline we propose, as a safest choice, is the progression of increasing $N$, from low to high, it costs more in terms of iterations, but it appears to converge properly.

5.2 Three-dimensional experiment

The method extends for three-dimensional model reconstruction, simply incurring a larger computational cost (as larger matrices are involved for the eigenvectors decomposition and
the forward problem discretization). We proceed with a three-dimensional experiment, where we consider a subsurface medium of size $2.46 \times 1.56 \times 1.2$ km, encompassing several salt domes, illustrated in Figure 15. The seismic acquisition consists in 96 sources, positioned on a two-dimensional plane at 10 m depth; the 1000 receivers are positioned at 100 m depth. Similarly as the previous experiment, the data are first generated in the time-domain and we incorporate noise before we proceed to the Fourier transform. Figure 17 shows the time-domain data associated with a centrally located source, and the corresponding Fourier transform at 5 Hz frequency. For the reconstruction, we start with a one-dimensional variation, in depth only, where none of the objects are intuited, see Figure 16, and the velocity background is also unknown.

**Remark 10.** In [2], a similar test-case is presented in the context of reconstruction with, however, major differences. Firstly Cauchy data used in [2], while we only employ the measurement of the pressure field in the hereby experiment. Secondly, the experiment uses a single frequency: 4 Hz; here, the challenge is to start with higher frequency, 5 Hz data; thirdly, piecewise linear model representation is used in [2], while we use the eigenvector discretization here. Also the discretization method for the wave equation differs (finite differences and continuous Galerkin).

Figure 15: Three-dimensional model incorporating contrasting objects. The domain is of size $2.46 \times 1.56 \times 1.2$ km. We highlight an horizontal section at 550 m depth and vertical section at $y = 670$ m.

In Figure 18, we show the reconstruction without employing the eigenvector decomposition, where the wave speed has a piecewise constant representation on a $124 \times 79 \times 61$ nodal grid. We only use 5 Hz frequency data, and 30 iterations. Then, we employ the eigenvector model representation with Algorithm 2. Following the observations of the two-dimensional reconstruction, we select $\eta_3$, which appear the most robust and try two situations:

- **single frequency** (5 Hz), **fixed** $N$ reconstruction using $\eta_3$, $N = 50$ and 30 iterations, the final reconstruction is Figure 19.

- **single frequency** (5 Hz), **multiple** $N$ reconstruction using $\eta_3$, $N = \{20, 30, 50, 75, 100\}$ and 30 iterations per $N$, i.e. 180 iterations in total: the final reconstruction is Figure 20.
Figure 16: Initial model taken for the reconstruction of the three-dimensional medium with vertical section at $y = 670$ m. It consists in a one-dimensional variation in depth.

Figure 17: Time-domain data and corresponding Fourier transform at 5 Hz frequency. The three-dimensional trace corresponds with the evolution receivers recordings (positioned on a 2D map in the $x$–$y$ plane) with time. There are 1000 data points per time step (i.e. 1000 receivers on the domain) and we highlight sections at fixed time (0.5 s) and for a line of receivers (positioned at $y = 710$ m).

For visualization, we focus on two-dimensional vertical and horizontal sections which illustrate the objects positions and shapes.

This experiment is consistent with our previous results, and we observe the following.

- The classical FWI reconstruction fails to discover the subsurface objects, with only a narrow layer, misplaced, see Figure 18.

- The reconstruction with eigenvector decomposition is able to accurately capture the largest subsurface salt domes, see Figures 19 and 20.

- The best results are obtained when we use the progression of $N$, with a single frequency, see Figure 20. The main salt dome is captured and smaller ones start to appear, including near the boundary. Using a single $N$ is also a good candidate, Figure 19 as
Figure 18: Velocity model reconstruction with vertical section at $y = 670$ m, from starting medium Figure [16] using 5 Hz frequency data. The reconstruction does not apply eigenvector decomposition. The model is parametrized following the domain discretization, using piecewise constant representation with one value per node on a $124 \times 79 \times 61$ grid.

Figure 19: Velocity model reconstruction with vertical section at $y = 670$ m, from starting medium Figure [16] using 5 Hz frequency data. The reconstruction apply eigenvector decomposition with $\eta_3$ and $N = 50$ with 30 iterations, see Algorithm 2.

Figure 20: Velocity model reconstruction with vertical section at $y = 670$ m and horizontal section at 550 m depth, from starting medium Figure [16] and using 5 Hz frequency data. The reconstruction apply eigenvector decomposition with $\eta_3$ and progression of $N = \{20, 30, 50, 75, 100\}$, with 30 iterations per $N$, see Algorithm 2.
it necessitates much less iterations (30 instead of 180) hence less computational time.

6 Conclusion

In this paper, we have investigated the use of eigenvectors model representation for image decomposition and in reconstruction procedure associated with seismic inverse problems, using two and three-dimensional experiments. We have implemented nine diffusion coefficients, and compared their performance, depending on the target medium. Then, we have demonstrated the utility of this approach, in particular in the context of subsurface salt domes reconstruction. For an accurate decomposition, the case of contrasting objects is clearly much more appropriate than when having layered media. Also in this case, all diffusion coefficients provide satisfactory results for salt domes, even in the presence of noise. However for layered patterns, only a few are performing well ($\eta_1$, $\eta_3$, $\eta_6$) and none works very well in the presence of noise. Here, it would be interesting to investigate further the performance of anisotropic diffusion. [74, 36].

Then, we have considered the quantitative reconstruction procedure in seismic, where only restrictive partial, backscattered data are available. Here, the choice of coefficients makes a difference, and should be carefully selected. We recommend $\eta_5$, from [33], which was the most robust in our applications, even with fixed $N$ and a few iterations. However, we have shown that the choice of $N$ is not trivial, and one cannot rely on the misfit functional evaluation. Therefore, we have proposed a progression of increasing $N$, which appear to stabilize the reconstruction. The FWI algorithm based upon eigenvectors model representation has shown successful reconstruction of subsurface salt dome media, in particular avoiding the low frequency requirement, which is a crucial aspect for realistic seismic applications. Yet, it is not appropriate for layered media.

Following these analysis, some difficulties remain regarding the optimal choice of parameters. For instance, it is possible that $\eta$ has to be selected differently depending on the model (as illustrated with the Marmousi decomposition). Similarly, the scaling coefficient $\beta$ would affect the performance but the acceptable range appears, hopefully, quite large.

Acknowledgments

The authors would like to thank Prof. Marcus Grote for thoughtful comments and discussions. OS is supported by the Austrian Science Fund (FWF), with SFB F68, project F6807-N36 (Tomography with Uncertainties). OS also acknowledges support from the Austrian Science Fund (FWF) within the national research network Geometry and Simulation, project S11704 (Variational Methods for Imaging on Manifolds) and I3661-N27 (Novel Error Measures and Source Conditions of Regularization Methods for Inverse Problems). The research of FF is supported by the Inria–TOTAL strategic action DIP.

A Two-dimensional reconstruction with evolution of basis

In this appendix, we experiment the re-computation of the eigenvectors basis along with the iterations. We consider the same test-case as Experiment 2 of Table 6, with one major difference: the set of eigenvectors is recomputed from the current iteration model every time we update $N$. Compared to Algorithm 2 instead of using a fixed $\psi_{loc}$ from the initial
model, it is recomputed from the current $m^{(k)}$ every time we change the value of $N$. The reconstructions using the update of basis are shown Figure 21 for all choices of $\eta$ in Table 1.

We observe that changing the basis along with the iterations does not provide any benefits for the accuracy of the reconstruction. On the contrary, it seems to deteriorate the recovery compared to keeping the same basis from the starting model (see Figures 13 and 25).

Figure 21: Reconstruction of the salt velocity model from starting medium Figure 9 (with the same color scale) using 2 Hz frequency data. The eigenvector decomposition employs a progression of eigenvectors with $N = \{50, 60, 70, 80, 90, 100\}$ and 30 iterations per $N$. Here, the set of eigenvectors is changed when $N$ changes (i.e. every 30 iterations), and is computed from the current model reconstruction. The formulations of $\eta$ are from Table 1.

B Extended set of figures

B.1 Noise-free model decomposition from Subsection 4.1
B.2 Two-dimensional salt model FWI from Subsection 5.1
Figure 22: Decomposition of the noise-free Marmousi velocity model of Figure 3(a) using $N = 50$, the formulation for $\eta$ are from Table 1 and the respective $\mu$ values extracted from Table 2. The color scale varies between 1500 and 5500 m s$^{-1}$.

Figure 23: Decomposition of the noise-free salt velocity model of Figure 3(b) using $N = 20$, the formulation for $\eta$ are from Table 1 and the respective $\mu$ values extracted from Table 3. The color scale varies between 1500 and 4500 m s$^{-1}$.

References

[1] V. Akcelik, G. Biros, and O. Ghattas, Parallel multiscale gauss-newton-krylov methods for inverse wave propagation, in Supercomputing, ACM/IEEE 2002 Conference, Nov 2002.
Figure 24: Reconstruction of the salt velocity model from starting medium Figure 9 (with the same color scale) using 2 Hz frequency data. The eigenvector decomposition employs $N = 100$ and the formulations of $\eta$ from Table 1.

Figure 25: Experiment 2 of Table 6 (multi $N$, single frequency data) for the reconstruction of the salt velocity model from starting medium Figure 9 (with the same color scale). The formulations of $\eta$ follow Table 1.

[2] G. Alessandrini, M. V. de Hoop, F. Faucher, R. Gaburro, and E. Sincich, Inverse problem for
Figure 26: Experiment 3 of Table 6 (single $N$, multiple frequency data) for the reconstruction of the salt velocity model from starting medium Figure 9 (with the same color scale), the formulations of $\eta$ follow Table 1.

Figure 27: Experiment 4 of Table 6 (multiple $N$, multiple frequency data) for the reconstruction of the salt velocity model from starting medium Figure 9 (with the same color scale), the formulations of $\eta$ follow Table 1.

*the helmholtz equation with cauchy data: reconstruction with conditional well-posedness driven iterative
regularization, arXiv preprint arXiv:1712.00398, (2017).

[3] G. Alessandrini, M. V. de Hoop, R. Gaburro, and E. Sincich, Lipschitz stability for a piecewise linear Schrödinger potential from local Cauchy data. Asymptotic Analysis, 108 (2018), pp. 115–149.

[4] L. Alvarez, P.-L. Lions, and J.-M. Morel, Image selective smoothing and edge detection by nonlinear diffusion. ii, SIAM Journal on numerical analysis, 29 (1992), pp. 845–866.

[5] P. R. Amestoy, I. S. Duff, J.-Y. L’Excellent, and J. Koster, A fully asynchronous multifrontal solver using distributed dynamic scheduling, SIAM Journal on Matrix Analysis and Applications, 23 (2001), pp. 15–41.

[6] P. R. Amestoy, A. Guermouche, J.-Y. L’Excellent, and S. Pralet, Hybrid scheduling for the parallel solution of linear systems, Parallel computing, 32 (2006), pp. 136–156.

[7] H. Ammari, J. K. Seo, and L. Zhou, Viscoelastic modulus reconstruction using time harmonic vibrations, Mathematical Modelling and Analysis, 108 (2018), pp. 115–149.

[8] L. Alvarez, P.-L. Lions, and J.-M. Morel, Image selective smoothing and edge detection by nonlinear diffusion. ii, SIAM Journal on numerical analysis, 29 (1992), pp. 182–193.

[9] H. Barucq, H. Calandra, G. Chavent, and F. Faucher, A priori estimates of attraction basins for velocity model reconstruction by time-harmonic Full Waveform Inversion and Data Space Reflectivity formulation, Research Report RR-9253, Magique 3D ; Inria Bordeaux Sud-Ouest ; Université de Pau et des Pays de l’Adour, Feb. 2019.

[10] H. Barucq, F. Faucher, and H. Pham, Localization of small obstacles from back-scattered data at limited incident angles with full-waveform inversion, Journal of Computational Physics, 370 (2018), pp. 1–24.

[11] J. B. Bednar, C. Shin, and S. Pyun, Comparison of waveform inversion, part 2: phase approach, Geophysical Prospecting, 55 (2007), pp. 465–475.

[12] J.-P. Bérenger, A perfectly matched layer for the absorption of electromagnetic waves, Journal of Computational Physics, 114 (1994), pp. 185 – 200.

[13] E. Beretta, M. V. de Hoop, F. Faucher, and O. Scherzer, Inverse boundary value problem for the helmholtz equation: quantitative conditional lipschitz stability estimates, SIAM Journal on Mathematical Analysis, 48 (2016), pp. 3962–3983.

[14] L. Blanc-Féraud, P. Charbonnier, G. Aubert, and M. Barlaud, Nonlinear image processing: modeling and fast algorithm for regularization with edge detection, in Image Processing, 1995. Proceedings., International Conference on, vol. 1, IEEE, 1995, pp. 474–477.

[15] E. Bozdağ, J. Trampert, and J. Tromp, Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, Geophysical Journal International, 185 (2011), pp. 845–870.

[16] R. Brossier, Two-dimensional frequency-domain visco-elastic full waveform inversion: Parallel algorithms, optimization and performance, Computers & Geosciences, 37 (2011), pp. 444–455.

[17] R. Brossier, S. Operto, and J. Virieux, Seismic imaging of complex onshore structures by 2d elastic frequency-domain full-waveform inversion, Geophysics, 74 (2009), pp. WCC105–WCC118.

[18] R. Brossier, S. Operto, and J. Virieux, Which data residual norm for robust elastic frequency-domain full waveform inversion?, Geophysics, 75 (2010), pp. R37–R46.

[19] C. Bunks, F. M. Saleck, S. Zaleski, and G. Chavent, Multiscale seismic waveform inversion, Geophysics, 60 (1995), pp. 1457–1473.

[20] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll, Image selective smoothing and edge detection by nonlinear diffusion, SIAM Journal on Numerical Analysis, 29 (1992), pp. 182–193.

[21] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, Two deterministic half-quadratic regularization algorithms for computed imaging, in Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference, vol. 2, IEEE, 1994, pp. 168–172.

[22] G. Chavent, Identification of functional parameters in partial differential equations, in Identification of Parameters in Distributed Systems, R. E. Goodson and M. Polis, eds., ASME, New York, 1974, pp. 31–48.

[23] G. Chavent, K. Gadyushin, and V. Tcheverda, Reflection fwi in mbti formulation, in 77th EAGE Conference and Exhibition 2015, 2015.
[24] Y. Choi, D.-J. Min, and C. Shin, Frequency-domain elastic full waveform inversion using the new pseudo-hessian matrix: Experience of elastic marmousi-2 synthetic data, Bulletin of the Seismological Society of America, 98 (2008), pp. 2402–2415.

[25] 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 (2013), p. 085009.

[26] ______, A new approach to solve the inverse scattering problem for waves: combining the trac and the adaptive inversion methods, Inverse Problems, 29 (2013), p. 085009.

[27] M. De Buhan and A. Osses, Logarithmic stability in determination of a 3d viscoelastic coefficient and a numerical example, Inverse Problems, 26 (2010), p. 095006.

[28] M. V. de Hoop, L. Qiu, and O. Scherzer, A convergence analysis of a multi-level projected steepest descent iteration for nonlinear inverse problems in banach spaces subject to stability constraints, arXiv preprint arXiv:1206.3706, (2012).

[29] S. C. Eisenstat and H. F. Walker, Globally convergent inexact newton methods, SIAM Journal on Optimization, 4 (1994), pp. 393–422.

[30] B. Engquist and A. Majda, Absorbing boundary conditions for numerical simulation of waves, Proceedings of the National Academy of Sciences, 74 (1977), pp. 1765–1766.

[31] L. C. Evans, Partial differential equations, (2010).

[32] F. Faucher, Contributions to seismic full waveform inversion for time harmonic wave equations: stability estimates, convergence analysis, numerical experiments involving large scale optimization algorithms, PhD thesis, Université de Pau et Pays de l’Ardour, 2017.

[33] D. Geman and G. Reynolds, Constrained restoration and the recovery of discontinuities, IEEE Transactions on Pattern Analysis & Machine Intelligence, (1992), pp. 367–383.

[34] P. J. Green, Bayesian reconstructions from emission tomography data using a modified em algorithm, IEEE transactions on medical imaging, 9 (1990), pp. 84–93.

[35] M. J. Grote, M. Kray, and U. Nahum, Adaptive eigenspace method for inverse scattering problems in the frequency domain, Inverse Problems, 33 (2017), p. 025006.

[36] M. J. Grote and U. Nahum, Adaptive eigenspace for multi-parameter inverse scattering problems, Computers & Mathematics with Applications, (2019).

[37] V. Isakov, Inverse problems for partial differential equations, vol. 127, Springer, 2006.

[38] H. Jun, E. Park, and C. Shin, Weighted pseudo-hessian for frequency-domain elastic full waveform inversion, Journal of Applied Geophysics, 123 (2015), pp. 1–17.

[39] A. Kadu, T. Van Leeuwen, and W. Mulder, Parametric level-set full-waveform inversion in the presence of salt bodies, in SEG Technical Program Expanded Abstracts 2017, Society of Exploration Geophysicists, 2017, pp. 1518–1522.

[40] M. Kalita, V. Kazei, Y. Choi, and T. Alkhalifah, Regularized full-waveform inversion for salt bodies, in SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 2018, pp. 1043–1047.

[41] B. Kaltenbacher, Minimization based formulations of inverse problems and their regularization, SIAM Journal on Optimization, 28 (2018), pp. 620–645.

[42] M. Kern, Numerical Methods for Inverse Problems, John Wiley & Sons, 2016.

[43] A. Kirsch, An introduction to the mathematical theory of inverse problems, Applied mathematical sciences, Springer, New York, 1996.

[44] D. Köhn, D. De Nil, A. Kurzmann, A. Przebindowska, and T. Bohlen, On the influence of model parametrization in elastic full waveform tomography, Geophysical Journal International, 191 (2012), pp. 325–345.

[45] P. Lailly, The seismic inverse problem as a sequence of before stack migrations, in Conference on Inverse Scattering: Theory and Application, J. B. Bednar, ed., Society for Industrial and Applied Mathematics, 1983, pp. 206–220.

[46] R. B. Lehoucq and D. C. Sorensen, Deflation techniques for an implicitly restarted arnoldi iteration, SIAM Journal on Matrix Analysis and Applications, 17 (1996), pp. 789–821.

[47] J. L. Lions and S. K. Mitter, Optimal control of systems governed by partial differential equations, vol. 1200, Springer Berlin, 1971.
[48] P. Lobel, L. Blanc-Féraud, C. Pichot, and M. Barlaud, A new regularization scheme for inverse scattering, Inverse Problems, 13 (1997), p. 403.

[49] I. Loris, H. Douma, G. Nolet, I. Daubechies, and C. Regone, Nonlinear regularization techniques for seismic tomography, Journal of Computational Physics, 229 (2010), pp. 890–905.

[50] I. Loris, G. Nolet, I. Daubechies, and F. Dahlen, Tomographic inversion using $l_1$-norm regularization of wavelet coefficients, Geophysical Journal International, 170 (2007), pp. 359–370.

[51] L. Métivier, R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux, Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion, Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society, 205 (2016), pp. 345–377.

[52] L. Métivier, R. Brossier, J. Virieux, and S. Operto, Full waveform inversion and the truncated newton method, SIAM Journal on Scientific Computing, 35 (2013), pp. B401–B437.

[53] J. Nocedal and S. J. Wright, Numerical Optimization, Springer Series in Operations Research, 2 ed., 2006.

[54] S. Operto, J. Virieux, A. Ribodetti, and J. Anderson, Finite-difference frequency-domain modeling of viscoacoustic wave propagation in 2d tilted transversely isotropic (tti) media, Geophysics, 74 (2009), pp. 75–95.

[55] J. Nocedal and J. Malik, A network for multiscale image segmentation, in Circuits and Systems, 1988., IEEE International Symposium on, IEEE, 1988, pp. 2565–2568.

[56] ———, Scale-space and edge detection using anisotropic diffusion, IEEE Transactions on pattern analysis and machine intelligence, 12 (1990), pp. 629–639.

[57] R.-E. Plessix, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, Geophysical Journal International, 167 (2006), pp. 495–503.

[58] R. G. Pratt, C. Shin, and G. J. Hick, Gauss–newton and full newton methods in frequency–space seismic waveform inversion, Geophysical Journal International, 133 (1998), pp. 341–362.

[59] R. G. Pratt, Z.-M. Song, P. Williamson, and M. Warner, Two-dimensional velocity models from wide-angle seismic data by waveform inversion, Geophysical Journal International, 124 (1996), pp. 323–340.

[60] R. G. Pratt and M. H. Worthington, Inverse theory applied to multi-source cross-hole tomography., Geophysical Prospecting, 38 (1990), pp. 287–310.

[61] S. Pyun, C. Shin, and J. B. Bednar, Comparison of waveform inversion, part 3: amplitude approach, Geophysical Prospecting, 55 (2007), pp. 477–485.

[62] L. Qi, N. Chemingui, Z. Zou, and A. Valenciano, Full-waveform inversion with steerable variation regularization, in SEG Technical Program Expanded Abstracts 2016, Society of Exploration Geophysicists, 2016, pp. 1174–1178.

[63] L. Robert and R. Deriche, Dense depth map reconstruction: A minimization and regularization approach which preserves discontinuities, in European Conference on Computer Vision, Springer, 1996, pp. 439–451.

[64] L. I. Rudin, S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D: nonlinear phenomena, 60 (1992), pp. 259–268.

[65] C. Shin, S. Pyun, and J. B. Bednar, Comparison of waveform inversion, part 1: conventional wavefield vs logarithmic wavefield, Geophysical Prospecting, 55 (2007), pp. 449–464.

[66] L. Sirgue and R. G. Pratt, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies, Geophysics, 69 (2004), pp. 231–248.

[67] A. Tarantola, Inversion of seismic reflection data in the acoustic approximation, Geophysics, 49 (1984), pp. 1259–1266.

[68] A. Tarantola, A strategy for nonlinear elastic inversion of seismic reflection data, Geophysics, 51 (1986), pp. 1893–1903.

[69] A. Tarantola, Inversion of travel times and seismic waveforms, in Seismic tomography, Springer, 1987, pp. 135–157.

[70] J. Virieux, Sh-wave propagation in heterogeneous media: velocity-stress finite-difference method, Geophysics, 49 (1984), pp. 1933–1942.
[71] J. Virieux and S. Operto, *An overview of full-waveform inversion in exploration geophysics*, Geophysics, 74 (2009), pp. WCC1–WCC26.

[72] C. R. Vogel and M. E. Oman, *Iterative methods for total variation denoising*, SIAM Journal on Scientific Computing, 17 (1996), pp. 227–238.

[73] S. Wang, M. V. de Hoop, and J. Xia, *On 3d modeling of seismic wave propagation via a structured parallel multifrontal direct helmholtz solver*, Geophysical Prospecting, 59 (2011), pp. 857–873.

[74] J. Weickert, *Anisotropic diffusion in image processing*, vol. 1, Teubner Stuttgart, 1998.

[75] Y. Yang, B. Engquist, J. Sun, and B. F. Hamfeldt, *Application of optimal transport and the quadratic wasserstein metric to full-waveform inversion*, Geophysics, 83 (2018), pp. R43–R62.

[76] Y. O. Yuan and F. J. Simons, *Multiscale adjoint waveform-difference tomography using wavelets*, Geophysics, 79 (2014), pp. WA79–WA95.

[77] Y. O. Yuan, F. J. Simons, and E. Bozdağ, *Multiscale adjoint waveform tomography for surface and body waves*, Geophysics, 80 (2015), pp. R281–R302.