

Cooperative Gravity and Full Waveform Inversion: Elastic Case



Raul U. Silva-Ávalos, Jonás D. De Basabe, and Mrinal K. Sen

Abstract A variety of geophysical methods can be used to acquire measurements to estimate a subsurface model. However, each method typically yields a different model and it can be challenging to merge with those obtained using other methods. We need to be able to combine the data from different geophysical methods to obtain a more detailed and consistent subsurface model. In this chapter, we present a scheme for cooperative inversion of seismic and gravity measurements. This scheme performs iteratively full-waveform inversion (FWI) and gravimetric inversion to minimize the misfit between the observed and synthetic data. We explain how to use the adjoint-state method to compute the gradient needed for FWI, the constrained conjugate gradient least-squares method to compute the gravimetric inversion, and how to incorporate petrophysical relationships to merge these methods in a cooperative scheme. The final system to be solved is large and sparse, hence the implementation relies on a large sparse matrix storage and high-performance computing. Finally, we show examples using the proposed inversion scheme and compare the results with those of FWI. The cooperative scheme yields more accurate models than those obtained from FWI with negligible additional computational cost.

R. U. Silva-Ávalos

Departamento de Geofísica Aplicada, División de Ciencias de la Tierra, CICESE, Ensenada, Mexico

J. D. De Basabe (✉)

Departamento de Sismología, División de Ciencias de la Tierra, CICESE, Ensenada, Mexico
e-mail: jonas@cicese.mx

M. K. Sen

Institute for Geophysics, John A. and Katherine G. Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, USA
e-mail: mrinal@utexas.edu

1 Introduction

Exploration geophysicists have developed a variety of methods to probe the subsurface using measurements that can be gathered on the ground. The interpretation of the data from these geophysical methods yields an assortment of subsurface models, and the conundrum is to merge these models into a unified model that better reflects the geometry and properties of the area of interest and fits all the available data.

Among the exploration-geophysics methods, the seismic method has been particularly successful. This method was at the heart of the energy transition that took place over one century ago [13] and continues to be widely used with an ever increasing number of applications [47]. Over its long history, this method has evolved together with the technology, becoming the basis for other state-of-the-art methods such as Full Waveform Inversion (FWI) [11].

FWI [38, 39] is a powerful seismic-imaging method used to estimate a seismic-velocity model such that the discrepancies between observed and synthetic seismograms are minimized. This method has become a popular [44] and in the recent years has improved, reducing the computational cost and enhancing the resolution of seismic images.

FWI consists of three main steps performed iteratively. The first step is to do the forward modeling starting from an initial model to compute the synthetic data, and obtain the residual by subtracting the observed data. Several authors have used the Finite Difference Method (FDM) [2, 43] for waveform modeling, however, the Finite Element Method [26], the Spectral Element Method (SEM) [22] or other methods can also be used. The second step is to back-propagate the residual wave field to obtain the adjoint field. This step includes computing a cross-correlation between the forward and the adjoint wavefield and adding over all the data points to obtain a velocity gradient. This is the well-known adjoint method [32], which reduces significantly the computational cost because only two forward modelings are required in each iteration of the inversion process. In the final step, the velocity model is updated by adding to the starting model the scaled velocity gradient using a line-search method to determine the increment. If the observed and synthetic data do not match, these steps are repeated until a stopping criterion is reached. This methodology has provided good results for stratigraphic and predominantly horizontal layered models. Despite the good results both in acoustic and elastic media, density variations have largely been ignored [44].

The gravimetric data is directly linked to the density variations in the subsoil. The observed data can be the gravity or the gravity gradient tensor on the surface [49]. As in FWI, the interpretation of this data relies on solving the forward problem. Several forward modeling methods exist to compute gravity anomalies by solving Poisson's equation for the gravitational potential. Among the best-known methods is the analytical solution for prismatic bodies [4, 29], however, solutions for other geometries are readily available [16, 20, 37, 46]. In this work, we will use the solution for uniform rectangular prisms to be congruent with the grid used in finite differences for waveform modeling.

Gravimetric inversion (GI) for density estimation is a linear problem. This method is well known for estimating structures with horizontal changes of mass distribution. The solution is straightforward using Gauss-Newton minimization [35] to obtain a density model inverting the square matrix on a single step. This method is widely used by geophysicists because of its fast convergence, however, it is computationally expensive and unfeasible for large-scale problems. One alternative to this problem is to use the Conjugate Gradient Least Squares (CGLS) method. This method solves the inverse problem without the need to form and store the square matrix [35].

Nowadays, the exploration of a region of interest for underground resources requires the measurements of several geophysical datasets which need to be interpreted for characterization. Joint inversion allows integrating these different datasets into a consistent Earth-property model. Usually the strategy consists in combining all the methods into one single inverse problem. Vozoff and Jupp [45] were the first to perform joint inversion for different geophysical data sets, namely resistivity and magnetotelluric data. Following this, numerous methodologies and different geophysical data-inversion schemes emerged for the reduction of non-uniqueness and ambiguity in the interpretation of the Earth model. Depending on the constraints in the optimization problem, the joint inversion schemes can be classified into petrophysical, structural, or statistical. Petrophysical joint inversion is based on empirical relationships of the model parameters [23, 27, 48], structural joint inversion seeks to minimize the cross product of the gradient of each model parameter [14, 15] and statistical joint inversion tries to solve the problem attaching to each grid cell of the model a mean point (fuzzy c-mean) depending on the number of c-means [31, 33].

The cooperative inversion of seismic and gravimetric data has attracted significant attention since these methods complement each other and both theories depend on the density. For example, Roy et al. [34] performed first-arrival travel time inversion jointly with gravity data using very-fast simulated annealing. Other groups have done further work using seismic and gravity data [9, 24, 25, 41]. In particular, Blom et al. [6] stressed the importance of density in geological processes and studied the role of density using seismic and gravimetric data, concluding that density estimation requires a strong a priori model to be able to determine it as an independent parameter.

In this work, we present a novel method to obtain a unified inverted model using FWI and GI in a sequential and cooperative scheme. This chapter is divided into five sections, Sect. 1 being the introduction. Section 2 presents the forward modeling on a geophysical framework and is divided in two parts for gravimetric and seismic data modeling. For gravity, we discuss Newton's law of universal gravitation and present the forward modeling based on the response of a rectangular body. For seismic modeling, we give a brief introduction to elastodynamic theory and the forward modeling for elastic and acoustic media. Section 3 discusses the inverse problem and follows the same organization as Chap. 2 for each geophysical method. We first present the general basis on inverse theory. Then we discuss the separate inversion for each method and present the the sequential inversion. The results are presented in

Sect. 4 for two synthetic models using using conventional and cooperative inversion. The conclusions are included in Sect. 5.

2 Forward Modeling of Geophysical Data

This section presents the theoretical framework for the gravimetric and seismic geophysical methods. For the gravity data, we present the solution of Newton's Law of gravitation for a parallelepiped of constant density. For seismic data, we discuss the wave equations for elastic and acoustic media and show how to solve them using finite-difference methods.

2.1 Gravimetric Forward Modeling

Newton's law of gravitation [5] provides the gravitational potential ϕ at an observation point \mathbf{r} due to a body on Earth with density distribution ρ (Fig. 1) as

$$\phi(\mathbf{r}) = \int_{\Omega} \gamma \frac{\rho(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|} dV, \quad (1)$$

where $\gamma = 6.672 \times 10^{-11} \text{ m}^3\text{kg}^{-1}\text{s}^{-2}$ is the universal gravitation constant, \mathbf{r}' is the position for each differential element of density over the volume Ω and $\|\cdot\|$ denotes the vector norm. The gravity acceleration field is given by the gradient of

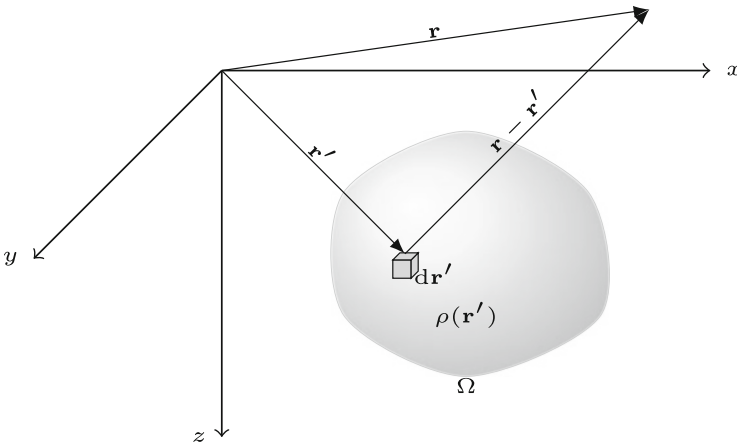


Fig. 1 Observation vector \mathbf{r} and position vector \mathbf{r}' for each differential volume element $d\mathbf{r}'$ for a continuous of density ρ in Cartesian coordinates system

the potential,

$$\mathbf{g}(\mathbf{r}) = \nabla\phi(\mathbf{r}). \quad (2)$$

Consider an arbitrary continuous body of density ρ in Cartesian coordinates (Fig. 1), the components of the gravity acceleration are given by

$$g_x(x, y, z) = \frac{\partial\phi}{\partial x} = \gamma \int_{\Omega} \rho \frac{x - x'}{[(x - x')^2 + (y - y')^2 + (z - z')^2]^{3/2}} dx' dy' dz', \quad (3)$$

$$g_y(x, y, z) = \frac{\partial\phi}{\partial y} = \gamma \int_{\Omega} \rho \frac{y - y'}{[(x - x')^2 + (y - y')^2 + (z - z')^2]^{3/2}} dx' dy' dz', \quad (4)$$

$$g_z(x, y, z) = \frac{\partial\phi}{\partial z} = \gamma \int_{\Omega} \rho \frac{z - z'}{[(x - x')^2 + (y - y')^2 + (z - z')^2]^{3/2}} dx' dy' dz'. \quad (5)$$

In this work, we consider only the vertical component of the gravity acceleration g_z , as usually done in geophysics.

2.2 Gravimetric Forward Modeling

In order to compute the gravimetric response at any observation point on the surface, we need a discretization of the Earth model. Given that Eq. 5 is valid for a continuous body of arbitrary shape and density distribution and taking advantage of the superposition theorem for Newton's law of gravitation, the Earth model can be discretized as a set of rectangular prism of constant density (Fig. 2). For each prism, the analytic solution of Eq. 5 is given by Banerjee and Das Gupta [3]

$$g_z = \left\{ \gamma \left[z \tan^{-1} \left(\frac{xy}{z|\Delta r|} \right) - x \ln(y + |\Delta r|) - y \ln(x + |\Delta r|) \right] \right|_{\Delta x'_1}^{\Delta x'_2} \left|_{\Delta y'_1}^{\Delta y'_2} \left|_{\Delta z'_1}^{\Delta z'_2} \right. \right\} \rho, \quad (6)$$

where the prime coordinates are the corners of the prism, $|\Delta r| = \sqrt{x^2 + y^2 + z^2}$, $\Delta x'_k = x - x'_k$, $\Delta y'_k = y - y'_k$ and $\Delta z'_k = z - z'_k$ $k = 1, 2$. This expression corresponds to the gravity measurement at the point (x, y, z) due to the prism and the part within the braces is the gravity kernel.

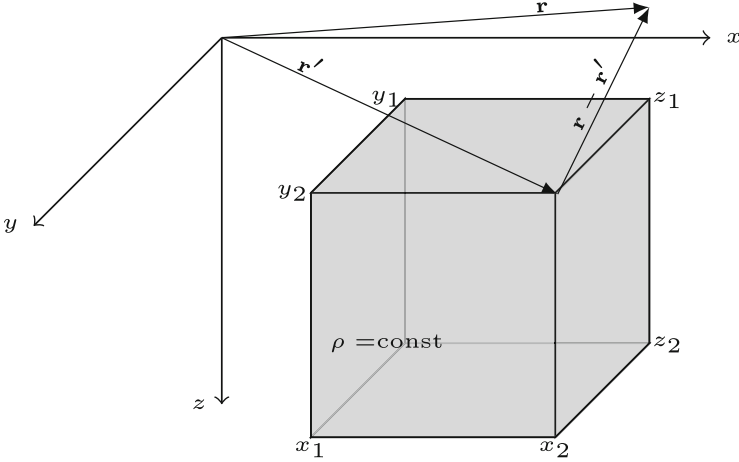


Fig. 2 Rectangular prism of constant density ρ . The coordinates x_i, y_i, z_i are the corners of the prism for $i = 1, 2$

A typical data acquisition is done on the surface for N_s observation points (gravimetric stations), hence $\mathbf{g}_z = [g_{z_1}, g_{z_2}, \dots, g_{z_{N_s}}]^T \in \mathbb{R}^{N_s}$. Considering a model parametrization of $M = n_x \times n_y \times n_z$ prisms where n_x, n_y and n_z are the number of prisms for x, y and z directions respectively, a model vector can be arranged as $\mathbf{m}_\rho = [\rho_1, \rho_2, \dots, \rho_M]^T \in \mathbb{R}^M$. Given this vector notation, the corresponding matrix for the kernel \mathbf{A} in index notation is

$$A_{ij} = \gamma \left[z_i \tan^{-1} \left(\frac{x_i y_i}{z_i |\Delta r_i|} \right) - x_i \ln(y_i + |\Delta r_i|) - y_i \ln(x_i + |\Delta r_i|) \right] \begin{vmatrix} \Delta x'_{2j} & \Delta y'_{2j} & \Delta z'_{2j} \\ \Delta x'_{1j} & \Delta y'_{1j} & \Delta z'_{1j} \end{vmatrix}, \quad (7)$$

where $\mathbf{A} \in \mathbb{R}^{N_s \times M}$, thus the gravity data vector can be represented in a matrix form as

$$\mathbf{g}_z = \mathbf{A} \mathbf{m}_\rho, \quad (8)$$

corresponding to the forward modeling of the gravimetric data. This is a linear problem with respect to density.

2.3 *Waveform Forward Modeling*

An elastic body is governed by the generalized Hooke's law. For small deformations and ignoring attenuation, the stress and strain are directly proportional as

$$\boldsymbol{\tau} = \mathbf{c} : \boldsymbol{\epsilon}, \quad (9)$$

where $\boldsymbol{\tau}$ is the stress tensor, $\boldsymbol{\epsilon}$ the strain tensor, \mathbf{c} represents the fourth-order stiffness tensor containing the constants that characterize the elastic properties of the solid, and $:$ is the double dot product for tensors. In index notation, Eq. 9 can be represented as

$$\tau_{ij} = c_{ijkl}\epsilon_{kl}. \quad (10)$$

for $i, j, k, l = 1, 2, 3$. Taking into consideration that the strain is proportional to the gradient of the displacement [1], $\boldsymbol{\epsilon} = \frac{1}{2} [\nabla u + (\nabla u)^T]$, Eq. 10 can be written as

$$\tau_{ij} = c_{ijkl} \frac{\partial u_k}{\partial x_l}, \quad (11)$$

where $u_k = \{u_x(x, t), u_y(x, t), u_z(x, t)\}$ is the displacement vector. Following [1] and assuming that the elastic body is subject to Newton's second law ($F = ma$) normalized over a volume, an equation relating displacement and stresses can be obtained

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \tau_{ij}}{\partial x_j} + f_i, \quad \text{in } \Omega, \quad (12)$$

where Ω is the spatial domain, f_i represent an external force per unit volume, ρ is the density and the acceleration is written as the second derivative of the displacement u_i . The elastodynamic wave equation is obtained combining Eqs. 11 and 12, to obtain

$$\rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial}{\partial x_j} \left[c_{ijkl} \frac{\partial u_k(x, t)}{\partial x_l} \right] = f_i, \quad (13)$$

valid for heterogeneous, elastic and anisotropic media, ignoring attenuation or viscoelastic effects. In this work, only isotropic media will be considered. In this case, the stiffness tensor is reduced to

$$c_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}), \quad (14)$$

where λ and μ are the Lamè parameters and δ_{ij} is the Kronecker delta function. Substituting Eq. 14 into 13 and reducing indexes, the elastic wave equation for

isotropic media is obtained as follows

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial}{\partial x_i} \left(\lambda \frac{\partial u_j}{\partial x_j} \right) + \frac{\partial}{\partial x_j} \left[\mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right] + f_i. \quad (15)$$

The media parameters of the wave equation were reduced to 3: Lamé's first parameter λ , the shear modulus μ , and the density ρ . There are other ways to write Eq. 15 depending on the choice of the elastic parameters, for example, the bulk modulus $\kappa = \lambda + \frac{2}{3}\mu$ is commonly used instead of λ . In general, these parameters can be expressed in terms of the P-wave velocity, $V_P = \sqrt{\frac{\lambda+2\mu}{\rho}}$, and the S-wave velocity, $V_S = \sqrt{\frac{\mu}{\rho}}$, which will be the parameters estimated on the inverse problem.

The elastodynamic wave equation can be simplified considering the wave propagation through acoustic media (fluids, melted bodies, liquid bodies) where there are no shear forces and therefore $\mu = 0$. Substituting this in Eq. 13 and defining $P = \lambda \nabla \cdot \mathbf{u}$, we obtain

$$\frac{1}{\lambda} \frac{\partial^2 P(\mathbf{x}, t)}{\partial t^2} - \nabla \cdot \left[\frac{1}{\rho} \nabla P(\mathbf{x}, t) \right] = \tilde{f}(\mathbf{x}, t), \quad (16)$$

where the scalar field P is the pressure propagated in the media due to an external force \tilde{f} . For constant density, this expression is simplified to the well-know acoustic wave equation

$$\frac{1}{V_P^2} \frac{\partial^2 P(\mathbf{x}, t)}{\partial t^2} - \nabla^2 P(\mathbf{x}, t) = \tilde{f}(\mathbf{x}, t), \quad (17)$$

where $V_P^2 = \frac{\lambda}{\rho}$ is the P-wave velocity. Let $\partial\Omega$ be the boundary of Ω and \hat{n} be the outward unit normal vector defined in the boundary. The boundary can be decomposed as $\partial\Omega = \Gamma_D \cup \Gamma_N$, $\Gamma_D \cap \Gamma_N = \emptyset$, where Γ_D and Γ_N are the boundaries where Dirichlet and Neumann conditions are defined. The boundary conditions for Eq. 17 are given by

$$P = P_D \quad \text{on} \quad \Gamma_D, \quad (18)$$

$$\nabla P \cdot \hat{n} = P_N \quad \text{on} \quad \Gamma_N. \quad (19)$$

Despite the fact that this equation is valid for acoustic media, it is often used for forward modeling in elastic media, FWI and RTM since it is computationally less expensive than the elastodynamic wave equation and, more importantly, the results are acceptable for many applications.

In order obtain the synthetic seismograms for displacement, velocity or pressure, Eqs. 15 and 17 need to be solved under some initial conditions. Among the most used techniques for wave propagation, we have FDM for acoustic [2] or elastic media [21], SEM for acoustic [8] or elastic media [22] and Finite Difference using

Staggered Grids (SGFD) for elastic media [43]. In this work, the acoustic wave equation will be solved using FDM and the elastic wave equation using SGFD.

Consider the following standard-grid discretization for the space–time domain

$$t_n = n\Delta t, \quad n = 0, 1, 2, \dots, n_t, \quad (20)$$

$$x_i = x_0 + i\Delta x, \quad i = 1, 2, 3, \dots, n_x, \quad (21)$$

$$y_j = y_0 + j\Delta y, \quad j = 1, 2, 3, \dots, n_y, \quad (22)$$

$$z_k = z_0 + k\Delta z, \quad k = 1, 2, 3, \dots, n_z, \quad (23)$$

$$P_{i,j,k}^n = P(x_i, y_j, z_k, t_n), \quad (24)$$

where n_x , n_y and n_z are the total number of grid points in each direction, n_t is the number of time steps, Δx , Δy , Δz and Δt are the spatial and time increments, and x_0 , y_0 and z_0 are the coordinates of the reference point. First, let us consider the acoustic problem. The discrete form for the spatial and time derivatives is given by Alford et al. [2]

$$P_{i,j,k}^{n+1} = 2P_{i,j,k}^n - P_{i,j,k}^{n-1} + V_P^2 \Delta t^2 \left(D_x^2 P_{i,j,k}^n + D_y^2 P_{i,j,k}^n + D_z^2 P_{i,j,k}^n \right), \quad (25)$$

where D_x^2 , D_y^2 and D_z^2 are the discrete operators for the second derivative. For example, the second-order discrete operator for the second derivative centered at x is given by

$$\left(\frac{\partial^2 P}{\partial x^2} \right)_n = \frac{P_{i+1,j,k}^n - 2P_{i,j,k}^n + P_{i-1,j,k}^n}{\Delta x^2} + O(\Delta x^2) \quad (26)$$

with $O(\Delta x^2)$ the truncation error. For this order, only 3 grid points in time are required to compute the second derivative of the pressure. Since the resolution depends on the parametrization of the velocity model in space, it is preferable to use more grid points for x , y and z , as seen in Table 1 for second derivatives for different orders of precision. The visual representation of the reference and neighbouring nodes for the discretization of the acoustic wave equation in 2D is shown in Fig. 3.

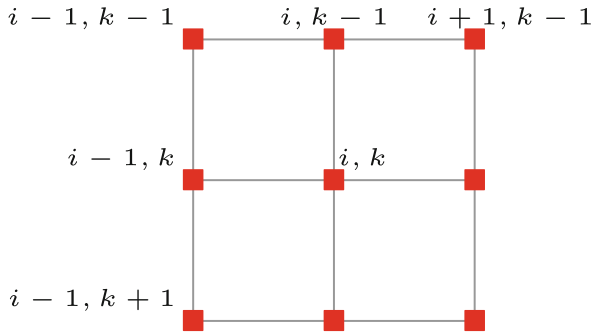
The numerical simulation of Eq. 26 involves the recursive computation of the pressure P over the time steps n_t . However, this recursive computation can present incremental error over time because of the truncation of the approximated solution or because of the machine rounding error. In order to set the discretization parameters such that the errors are bounded, a Von Neumann analysis is required. Based on the work of Alford et al. [2], a stability condition can be obtained by substituting a plane-wave solution into Eq. 26 and performing some standard algebraic simplifications, to obtain

$$\Delta t \leq \frac{\Delta x}{v_{\text{MAX}} \sqrt{n_D}} \left(\sum_{i=1}^M a_m \right)^{-1/2}, \quad (27)$$

Table 1 Central differences coefficients for second order derivative for accuracies of 2, 4, 6 and 8th order with uniform grid spacing

Approximation at $x = 0$ central differences										
Order of accuracy ↓	-4	-3	-2	-1	0	1	2	3	4	
2				1	-2	1				
4			$-\frac{1}{12}$	$\frac{4}{3}$	$-\frac{5}{2}$	$\frac{4}{3}$	$-\frac{1}{12}$			
6		$\frac{1}{90}$	$-\frac{3}{20}$	$\frac{3}{2}$	$-\frac{49}{18}$	$\frac{3}{2}$	$-\frac{3}{20}$	$\frac{1}{90}$		
8	$-\frac{1}{560}$	$\frac{8}{315}$	$-\frac{1}{5}$	$\frac{8}{5}$	$-\frac{205}{72}$	$\frac{8}{5}$	$-\frac{1}{5}$	$\frac{8}{315}$	$-\frac{1}{560}$	

Fig. 3 Visual representation of a standard grid discretization for a 2D acoustic media for the pressure field P



Where $\blacksquare \equiv P$

where v_{MAX} is the maximum value of the velocity model, $\sum_{i=1}^M a_m$ is the sum over the coefficients of Table 1 for each order of precision excluding the central point, and $n_D = 1, 2, 3$ is the dimension (1D, 2D, or 3D). This condition is very important for the inverse problem; given that it depends on the maximum velocity, the velocity model obtained has to be inspected in every iteration for stability.

In order to simulate the wave propagation in time a source has to be applied at any point of the space. In this example and in all the following results for this work, a Ricker wavelet is used, given by

$$w(t) = 2(\pi v)^2 \left\{ 1 - 2 [\pi v (t - t_0)]^2 \right\} e^{-[\pi v (t - t_0)]^2} \tag{28}$$

where v is the peak frequency of the pulse and t_0 is the time shift. The Ricker wavelet is also called the Mexican-hat wavelet because of its distinctive shape (see Fig. 4 for $t_0 = 0.0$ and $v = [2, 5, 10, 15, 25]$ Hz). For low frequencies the wavelet becomes wider and *vice versa* for high frequencies.

Concerning wave propagation in elastic media, FDM with a standard grid has grid-dispersion problems when there are significant contrast of properties [10], therefore, the forward modeling will be performed using Staggered Grid Finite Differences (SGFD) [43]. The isotropic wave equation can be expressed as the

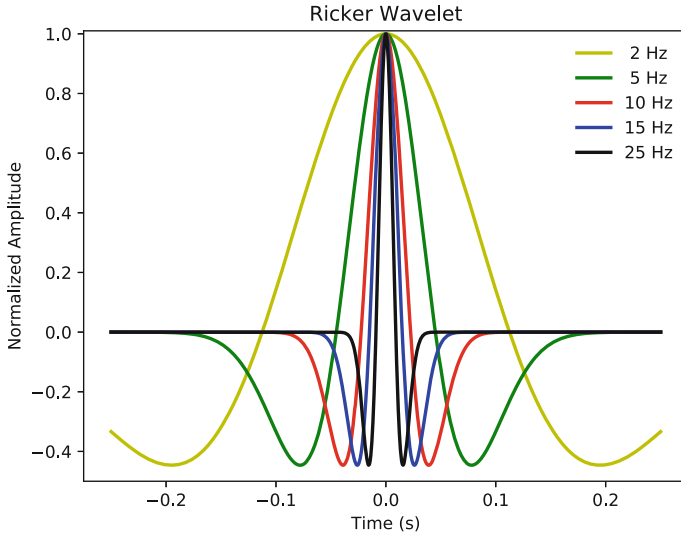


Fig. 4 Ricker wavelet function for peak frequencies 2,5,10,15 and 25 Hz. The function is centered at $t_0 = 0$

following set of equations

$$\rho \partial_{tt} u_x = \partial_x \tau_{xx} + \partial_y \tau_{xy} + \partial_z \tau_{xz} + f_x, \tag{29}$$

$$\rho \partial_{tt} u_y = \partial_x \tau_{xy} + \partial_y \tau_{yy} + \partial_z \tau_{yz} + f_y, \tag{30}$$

$$\rho \partial_{tt} u_z = \partial_x \tau_{xz} + \partial_y \tau_{yz} + \partial_z \tau_{zz} + f_z, \tag{31}$$

$$\tau_{xx} = (\lambda + 2\mu) \partial_x u_x + \lambda(\partial_y u_y + \partial_z u_z), \tag{32}$$

$$\tau_{yy} = (\lambda + 2\mu) \partial_y u_y + \lambda(\partial_x u_x + \partial_z u_z), \tag{33}$$

$$\tau_{zz} = (\lambda + 2\mu) \partial_z u_z + \lambda(\partial_x u_x + \partial_y u_y), \tag{34}$$

$$\tau_{xy} = \mu(\partial_y u_x + \partial_x u_y), \tag{35}$$

$$\tau_{xz} = \mu(\partial_z u_x + \partial_x u_z), \tag{36}$$

$$\tau_{yz} = \mu(\partial_z u_y + \partial_y u_z). \tag{37}$$

The discretization of the elastodynamic wave equation in the displacement-stress formulation is given by Virieux [43]

$$[u_x]_{i+1/2,j,k}^{n+1} = 2[u_x]_{i+1/2,j,k}^n - [u_x]_{i+1/2,j,k}^{n-1} + \Delta t^2 [b_x(D_x \tau_{xx} + D_y \tau_{xy} + D_z \tau_{xz} + f_x)]_{i+1/2,j,k}^n, \tag{38}$$

$$[u_y]_{i,j+1/2,k}^{n+1} = 2[u_y]_{i,j+1/2,k}^n - [u_y]_{i,j+1/2,k}^{n-1} + \Delta t^2 [b_y(D_x \tau_{xy} + D_y \tau_{yy} + D_z \tau_{yz} + f_y)]_{i,j+1/2,k}^n, \quad (39)$$

$$[u_z]_{i,j,k+1/2}^{n+1} = 2[u_z]_{i,j,k+1/2}^n - [u_z]_{i,j,k+1/2}^{n-1} + \Delta t^2 [b_z(D_x \tau_{xz} + D_y \tau_{yz} + D_z \tau_{zz} + f_z)]_{i,j,k+1/2}^n, \quad (40)$$

for the displacement calculated on midpoints of the grid. This time D_x , D_y and D_z are the discrete operators for the first derivative in a staggered grid and $b = 1/\rho$. For stresses

$$[\tau_{xx}]_{i,j,k}^n = [(\lambda + 2\mu)D_x u_x + \lambda(D_y u_y + D_z u_z)]_{i,j,k}^n, \quad (41)$$

$$[\tau_{yy}]_{i,j,k}^n = [(\lambda + 2\mu)D_y u_y + \lambda(D_x u_x + D_z u_z)]_{i,j,k}^n, \quad (42)$$

$$[\tau_{zz}]_{i,j,k}^n = [(\lambda + 2\mu)D_z u_z + \lambda(D_x u_x + D_y u_y)]_{i,j,k}^n, \quad (43)$$

$$[\tau_{xy}]_{i+1/2,j+1/2,k}^n = [\mu_{xy}(D_y u_z + D_x u_y)]_{i+1/2,j+1/2,k}^n. \quad (44)$$

$$[\tau_{xz}]_{i+1/2,j,k+1/2}^n = [\mu_{xz}(D_z u_x + D_x u_z)]_{i+1/2,j,k+1/2}^n. \quad (45)$$

$$[\tau_{yz}]_{i,j+1/2,k+1/2}^n = [\mu_{yz}(D_z u_y + D_y u_z)]_{i,j+1/2,k+1/2}^n. \quad (46)$$

The simplification from 3D to 2D media is straightforward ignoring the y -dependent terms. The finite difference coefficients for staggered grid are shown in the Table 2. Figure 5 shows a visual representation of a staggered grid.

Table 2 Central differences coefficients for first order derivatives for accuracies of 2, 4, 6 and 8th order with uniform grid spacing corresponding to staggered grid

	Approximation at $x = 0$							
	x -coordinates at nodes							
Order of accuracy ↓	$-\frac{7}{2}$	$-\frac{5}{2}$	$-\frac{3}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{5}{2}$	$\frac{7}{2}$
2				-1	1			
4			$\frac{1}{24}$	$-\frac{9}{8}$	$\frac{9}{8}$	$-\frac{1}{24}$		
6		$-\frac{3}{640}$	$\frac{25}{384}$	$-\frac{75}{64}$	$\frac{75}{64}$	$-\frac{25}{384}$	$\frac{3}{640}$	
8	$\frac{5}{7168}$	$-\frac{49}{5120}$	$\frac{245}{3072}$	$-\frac{1225}{1024}$	$\frac{1225}{1024}$	$-\frac{245}{3072}$	$\frac{49}{5120}$	$-\frac{5}{7168}$

Fig. 5 Visual representation of a staggered grid discretization for a 2D elastic media in terms of displacements (u_x and u_z) stresses (τ_{xx} , τ_{zz} and τ_{xz})

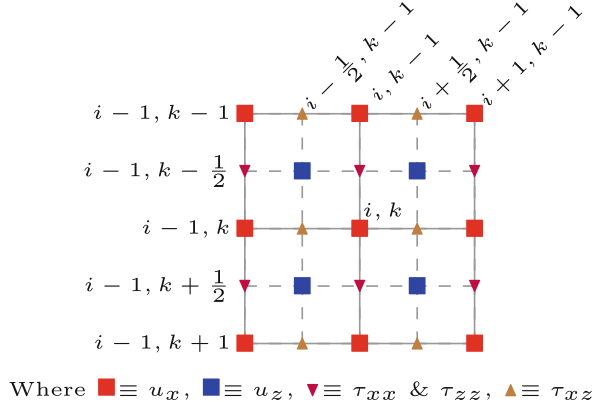
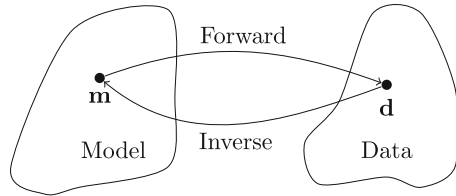


Fig. 6 Illustration of the concept of forward and the inverse problems



3 Inverse Theory for Geophysical Data

This section presents the basic concepts of inverse theory, providing the theoretical framework for GI and FWI for heterogeneous acoustic or elastic media, with an emphasis on the adjoint-state method for FWI. Starting from an Earth model, the forward problem computes theoretical data which will be compared to real data observations. Conversely, the inverse problem starts from the data and aims to compute an Earth model. A simple illustration of this statement is shown in Fig. 6. In general, the inverse problem is computationally more intensive, requires more sophisticated techniques and the interpretation of the results is more involved due to insufficient, inaccurate, noisy or inconsistent data [19]. In order to solve inverse problems, the following elements are essential in its formulation (see Table 3)

- Data,
- Model parameters,
- Forward problem,
- Cost/Objective/Error/Misfit function, and
- Optimization method.

Let us define a general formulation for inverse theory. The function (**F**) that involves such elements needs to be stated. The objective function (also known as cost, error or misfit function) compares the differences between the observed and

Table 3 Elements of inverse theory, where N is the number of data points and M the number of model parameters. In general $N \neq M$

Data vector	$\mathbf{d}^{\text{obs}} = [d_1 \ d_2 \ d_3 \ \dots \ d_N]^T$
Model vector	$\mathbf{m} = [m_1 \ m_2 \ m_3 \ \dots \ m_M]^T$
Forward modeling	$\mathbf{d}^{\text{cal}} = \mathbf{F}(\mathbf{m})$
Objective function	$Q = Q(\mathbf{d}^{\text{obs}}, \mathbf{d}^{\text{cal}} \mathbf{m})$

synthetic data vectors as follows

$$Q(\mathbf{m}) = \|\mathbf{d} - \mathbf{F}(\mathbf{m})\|_p, \quad (47)$$

where $\|\cdot\|_p$ is the L_p norm and Q the objective function. A general form of the L_p norm [28] is defined as

$$\|\mathbf{v}\| := \left[\sum_{i=1}^N |v_i|^p \right]^{1/p} \quad \forall \mathbf{v} \in \mathbb{R}^n, \quad (48)$$

where N is the number of data points and p determines the norm order. Some typical norms are

$$L_1 : \quad \|\mathbf{d} - \mathbf{F}(\mathbf{m})\|_1 := \sum_{i=1}^N |d_i - F_i(\mathbf{m})| \quad (49)$$

$$L_2 : \quad \|\mathbf{d} - \mathbf{F}(\mathbf{m})\|_2 := \left[\sum_{i=1}^N |d_i - F_i(\mathbf{m})|^2 \right]^{1/2} \quad (50)$$

$$L_\infty : \quad \|\mathbf{d} - \mathbf{F}(\mathbf{m})\|_\infty := \max_i |d_i - F_i(\mathbf{m})|. \quad (51)$$

The L_2 norm is used more often in geophysical applications, however, the L_1 norm is also largely studied despite the fact that it has a discontinuity in the derivative. When using the L_2 norm, it is often more practical to work with the square of the objective function, $Q = Q^2$. For illustration purposes, we show in Fig. 7 a comparison of a straight-line fit using the L_1 , L_2 and L_∞ norms. Notice that the norms for higher values of p are more biased towards outliers.

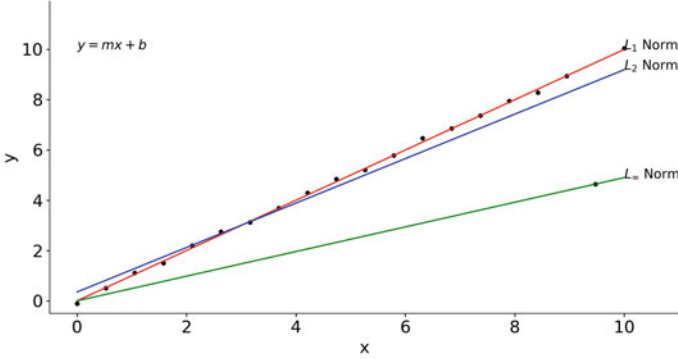


Fig. 7 L_p norm for some values of p corresponding to the fit of a straight line $\mathbf{y} = \mathbf{F}(\mathbf{x}) = \mathbf{a}\mathbf{x} + \mathbf{b}$

3.1 Gravimetric Inversion

The objective function for density estimation due to measurements of the vertical component of the acceleration ($\mathbf{g}_z^{\text{obs}}$) using the L_2 norm is given by

$$Q(\mathbf{m}_\rho) = \sum_{i=1}^{N_x} \left\| \frac{\mathbf{g}_{z_i}^{\text{obs}} - \mathbf{g}_{z_i}^{\text{cal}}}{\sigma_{\mathbf{g}_{z_i}}} \right\|^2 + \alpha_{\text{reg}}^2 \|\mathbf{D}\mathbf{m}_\rho\|^2, \quad (52)$$

where α_{reg} is the regularization parameter, \mathbf{D} is the discrete operator for the gradient and $\sigma_{\mathbf{g}_{z_i}}$ is the standard deviation of the i th data point. Solving the least-squares problem from Eq. 52 using Gauss-Newton method [35] an estimated model \mathbf{m}_ρ can be obtained as

$$\mathbf{m}_\rho = \left[\mathbf{A}^T \mathbf{C}_{dd}^{-1} \mathbf{A} + \alpha_{\text{reg}} \mathbf{D}^T \mathbf{D} \right]^{-1} \mathbf{A}^T \mathbf{C}_{dd}^{-1} \mathbf{g}_z^{\text{obs}}, \quad (53)$$

where \mathbf{C}_{dd}^{-1} is the diagonal covariance matrix and \mathbf{A} is given in equation 7. This least-squares implementation requires to store and invert a square matrix with dimensions depending on the discretization of the model, namely, $M \times M$. We need a fine discretization of the model to achieve a good resolution for the seismic inversion and therefore the joint inversion, nevertheless, we may encounter storage problems in a straight-forward implementation of Eq. 53.

An alternative to solving Eq. 53 is the use of the Conjugate Gradient Least Squares (CGLS) method. This method minimizes the objective function of Eq. 52 without the need to form and store the square matrix from Eq. 53 [35] using a conjugate gradient technique. This method requires as an input \mathbf{G} and \mathbf{d} to find a solution, in this case, the density model (\mathbf{m}_ρ) for $\mathbf{G}\mathbf{m}_\rho = \mathbf{d}$, these matrices are

given by

$$\mathbf{G} = \begin{bmatrix} \mathbf{C}_{dd}^{-1/2} \mathbf{A} \\ \alpha_{\text{reg}} \mathbf{D} \end{bmatrix} \quad (54)$$

$$\mathbf{d} = \begin{bmatrix} \mathbf{C}_{dd}^{-1/2} \mathbf{g}_z^{\text{obs}} \\ 0 \end{bmatrix}, \quad (55)$$

in this case, the matrix \mathbf{G} will be large and sparse due to the discrete operations for the Tikhonov regularization, the model vector \mathbf{m}_ρ is not modified.

3.2 Acoustic Full Waveform Inversion

We now proceed to describe the methodology of Acoustic Full Waveform Inversion (AFWI). The least-squares functional for minimizing the misfit between the observed and the synthetic pressure due to a single shot is given by the L_2 norm of the residual

$$Q = \frac{1}{2} \sum_r \int_0^T \left[P_r^{\text{obs}} - P_r^{\text{cal}} \right]^2 dt, \quad (56)$$

where P_r^{obs} is the observed pressure and P_r^{cal} is the synthetic pressure computed using Eq. 17. T is the total recording time and r denotes the index for the receiver. Implicitly the P_r^{cal} depends on the model parameter \mathbf{m} as $P_r^{\text{cal}} = P_r^{\text{cal}}(\mathbf{m})$. This model needs to be found in such a way Eq. 56 is minimized. Taking the derivative with respect to a model perturbation

$$\partial Q = - \sum_r \int_0^T \int \left[P_r^{\text{obs}} - P_r^{\text{cal}} \right] \delta P(\mathbf{x}, t) dt, \quad (57)$$

where δP is a perturbation of P aiding to compute the Frèchet derivative, which represents the sensibility for each data point and for each model parameter. This derivative is computed by making a small perturbation in each model parameter and performing a forward modeling for each data point, therefore $M \times N$ forward modelings are needed to obtain the derivative, which is impractical to implement even with the advances in computational resources, therefore, alternative methods for minimizing the problem are required.

A more efficient way to minimize Eq. 56 relies on the use of the adjoint state method for the acoustic waveform. Let us minimize the augmented misfit function subject to the wave equation multiplied by an arbitrary, well behaved and derivable

Lagrange multiplier $\Lambda := \Lambda(\mathbf{x}, t)$ remaining to be defined [32] then

$$\begin{aligned} Q &= \frac{1}{2} \sum_r \int_0^T \int_{\Omega} \left[P_r^{\text{obs}} - P_r^{\text{cal}} \right]^2 \delta(\mathbf{x} - \mathbf{x}_r) d^3x dt \\ &\quad - \int_0^T \int_{\Omega} \Lambda \left[\frac{1}{V_P^2} \frac{\partial^2 P}{\partial t^2} - \nabla^2 P - f \right]^2 d^3x dt, \end{aligned} \quad (58)$$

notice that the last term of Eq. 58 is zero, corresponding to the wave equation acting as constrictor, therefore the problem is consistent. Taking the total derivative

$$\begin{aligned} \delta Q &= - \sum_r \int_0^T \int_{\Omega} \left[P_r^{\text{obs}} - P_r^{\text{cal}} \right] \delta(\mathbf{x} - \mathbf{x}_r) \delta P d^3x dt \\ &\quad - \int_0^T \int_{\Omega} \Lambda \left[-\frac{2}{V_P^3} \frac{\partial^2 P}{\partial t^2} \right] d^3x dt \\ &\quad - \int_0^T \int_{\Omega} \Lambda \left[\frac{1}{V_P^2} \frac{\partial^2 \delta P}{\partial t^2} - \nabla^2 \delta P \right] d^3x dt, \end{aligned} \quad (59)$$

where the source is considered as independent of the model parameter perturbation. Notice that the perturbation δP appears on the first and last term. In the last term, the linear operator of the wave equation ($\mathcal{L} = \frac{1}{V_P^2} \frac{\partial^2}{\partial t^2} - \nabla^2$) is acting over δP which is a computation that we are looking to avoid. For this, let us first integrate by parts two times for t as

$$\int_0^T \Lambda \left[\frac{\partial^2 \delta P}{\partial t^2} \right] dt = \left(\Lambda \frac{\partial \delta P}{\partial t} \right) \Big|_0^T - \left(\frac{\partial \Lambda}{\partial t} \delta P \right) \Big|_0^T + \int_0^T \left[\frac{\partial^2 \Lambda}{\partial t^2} \right] \delta P dt. \quad (60)$$

Setting $\Lambda(x, t = T) = \frac{\partial \Lambda}{\partial t}(x, t = T) = 0$, yields

$$\int_0^T \Lambda \left[\frac{\partial^2 \delta P}{\partial t^2} \right] dt = \int_0^T \left[\frac{\partial^2 \Lambda}{\partial t^2} \right] \delta P dt, \quad (61)$$

this means that the second derivative is a self-adjoint operator ($\mathcal{L} = \mathcal{L}^*$). For the Laplacian operator ∇ the same procedure can be done, setting the correct boundary conditions in space. Consider the last term of Eq. 59,

$$\int_{\Omega} \Lambda \left[\nabla^2 \delta P \right] d^3x. \quad (62)$$

Taking into consideration the identity $\psi \nabla^2 \phi - \phi \nabla^2 \psi = \nabla \cdot (\psi \nabla \phi - \phi \nabla \psi)$, then

$$\int_{\Omega} \Lambda [\nabla^2 \delta P] d^3x = \int_{\Omega} \delta P [\nabla^2 \Lambda] d^3x + \int_{\Omega} \nabla \cdot [\Lambda \nabla \delta P - \delta P \nabla \Lambda] d^3x. \quad (63)$$

Applying Gauss theorem on the last term of the equation

$$\int_{\Omega} \nabla \cdot [\Lambda \nabla \delta P - \delta P \nabla \Lambda] d^3x = \int_{\partial\Omega} [\Lambda \nabla \delta P - \delta P \nabla \Lambda] \cdot \hat{n} dS, \quad (64)$$

where the integral was changed from volumetric to surface. In order to cancel the boundary integral in the above equation, we set the following boundary conditions for δP and Λ [12]

$$\delta P = 0, \quad \text{on } \Gamma_D, \quad \nabla \delta P \cdot \hat{n} = 0, \quad \text{on } \Gamma_N, \quad (65)$$

and

$$\Lambda = 0, \quad \text{on } \Gamma_D, \quad \nabla \Lambda \cdot \hat{n} = 0, \quad \text{on } \Gamma_N. \quad (66)$$

Therefore,

$$\int_{\partial\Omega} [\Lambda \nabla \delta P - \delta P \nabla \Lambda] \cdot \hat{n} dS = 0. \quad (67)$$

In this way, Eq. 59 can be rewritten as

$$\begin{aligned} \delta Q &= - \sum_r \int_0^T \int_{\Omega} [P_r^{\text{obs}} - P_r^{\text{cal}}] \delta(\mathbf{x} - \mathbf{x}_r) \delta P d^3x dt \\ &\quad - \int_0^T \int_{\Omega} \Lambda \left[-\frac{2}{V_P^3} \frac{\partial^2 P}{\partial t^2} \right] d^3x dt \\ &\quad - \int_0^T \int_{\Omega} \left[\frac{1}{V_P^2} \frac{\partial^2 \Lambda}{\partial t^2} - \nabla^2 \Lambda \right] \delta P d^3x dt, \\ &= - \int_0^T \int_{\Omega} \left\{ \frac{1}{V_P^2} \frac{\partial^2 \Lambda}{\partial t^2} - \nabla^2 \Lambda + \sum_r [P_s^{\text{obs}} - P_s^{\text{cal}}] \delta(x - x_r) \right\} \delta P d^3x dt \\ &\quad - \int_0^T \int_{\Omega} \Lambda \left[-\frac{2}{V_P^3} \frac{\partial^2 P}{\partial t^2} \right] d^3x dt. \end{aligned} \quad (69)$$

Let us define the Lagrange multiplier Λ in such a way that the first term of Eq. 69 is canceled. Then

$$\frac{1}{V_P^2} \frac{\partial^2 \Lambda}{\partial t^2} - \nabla^2 \Lambda = - \sum_r \left[P_r^{\text{obs}} - P_r^{\text{cal}} \right] \delta(\mathbf{x} - \mathbf{x}_r), \quad (70)$$

which corresponds to another wave equation using the residuals at the seismogram locations as a source. The importance of this result relies on the computation of the gradient without the need to compute the perturbation of P and therefore Fréchet derivatives, instead, a single additional forward modeling needs to be performed using the same wave propagation method but with the residuals as a source. Finally, to give more meaning to the Lagrange multiplier let us define $\Lambda(x, t) \equiv P^\dagger(x, T - t)$, thus the gradient

$$\partial V = \frac{2}{V_P^3} \int_0^T P^\dagger(\mathbf{x}, T - t) \frac{\partial^2 P(\mathbf{x}, t)}{\partial t^2} dt, \quad (71)$$

which is a convolution of the pressure and adjoint wave fields. Using multiple seismic sources requires a summation as follows

$$\partial V = \frac{2}{V_P^3} \sum_s^{n_s} \int_0^T P^\dagger(\mathbf{x}, T - t) \frac{\partial^2 P(\mathbf{x}, t)}{\partial t^2} dt. \quad (72)$$

where n_s is the total number of shots. Notice that the pressure and adjoint wavefields are computed in opposite directions for the time stepping: $P(\mathbf{x}, t)$ is going forward in time and $P^\dagger(\mathbf{x}, T - t)$ is going backward in time.

3.3 Gradient Based Optimization

With the velocity gradient, we can update the velocity model minimizing the cost of 56, but first let us illustrate how such gradient is constructed. Consider the modified Marmousi model and a starting 2D velocity model of Fig. 8. This model involves slightly folded layers similar to a bookshelf sliding fault system and a discordance event at the bottom. The velocity range was shortened to 1500–3500 km/s covering a depth of 1000 m and a horizontal distance of 2000 m on a grid of $n_x = 200$ and $n_z = 100$ grid nodes. The starting model is a smoothed version (Gaussian smoothing) of the true velocity model and the water layer is considered to be known in both models. Table 4 summarizes the parameters used for the forward modeling and the construction of the gradient for this example. The parameters satisfy the stability condition of Eq. 27 for a 10th order FDM in 2D media. The receivers and sources locations are equally spaced along the surface, 10 m spacing between seismograms and 20 m spacing for sources (shots). For this example, the seismic traces are shown in Fig. 9 for some shots. This data acquisition correspond

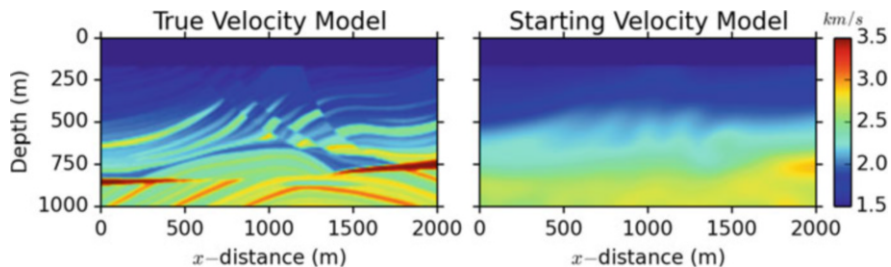


Fig. 8 Modified Marmousi velocity model (left) and starting velocity model (right). The velocity was shortened to 3500 m/s

Table 4 Parameters used for the construction of the gradient for the Marmousi model

Parameter	Value	Unit	Description
x_{MIN}	0	m	Starting horizontal distance
x_{MAX}	2000	m	Final horizontal distance
z_{MIN}	0	m	Starting depth
z_{MAX}	1000	m	Final depth
n_x	200		Grid nodes for x
n_z	100		Grid nodes for z
T	2	s	Recording time
n_t	1500		Grid nodes for t
v_{MAX}	3500	m/s	Maximum velocity
f	15	Hz	Ricket wavelet's peak frequency
n_r	200		Number of receivers
n_s	100		Number of sources

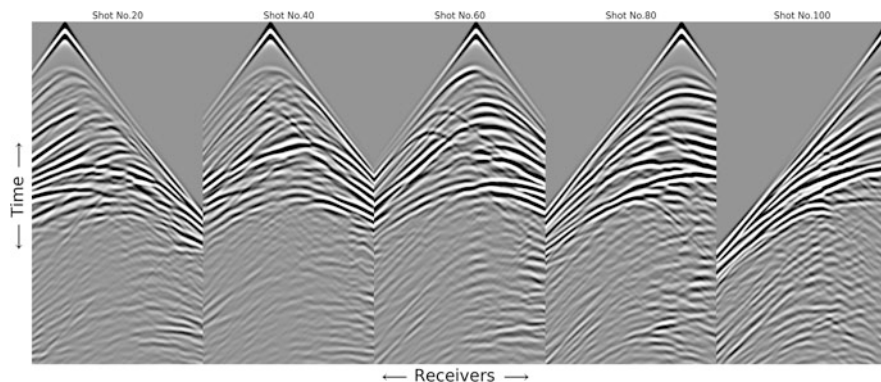


Fig. 9 Synthetic seismic data acquisition for the Marmousi model example at shots number 20, 40, 60, 80 and 100 corresponding to 200 receivers equally spaced along the surface

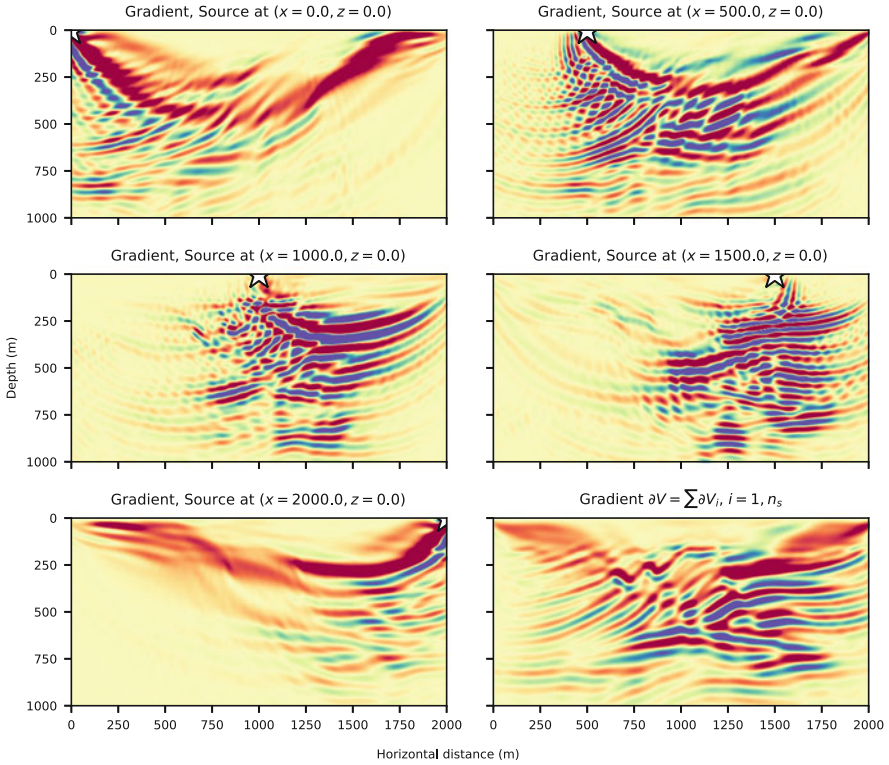


Fig. 10 Velocity gradient for several source locations $x_s = 0, 500, 1000, 1500$ and 2000 using the whole stream of receivers (200). The white star represents the different source positions. The bottom right gradient consist on the addition of all gradients

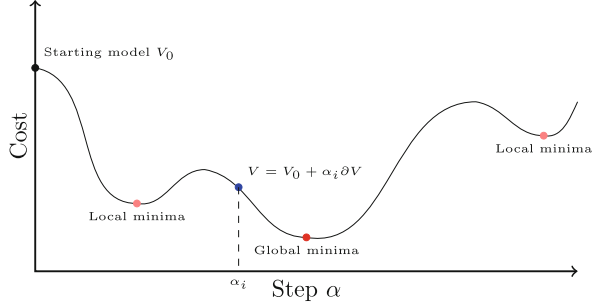
to the observed data vector $\mathbf{P}^{\text{obs}} \in \mathbb{R}^N$, $N = n_i n_r n_s$, which in this case is a vector of $1500 \times 200 \times 100 = 30$ Million data elements.

Let us consider several source positions at the surface, $x_s = 0, 500, 1000, 1500$ and 2000 m for a depth $z = 0$. The gradient for each source as well as the gradient stacked for all sources ($\partial V = \sum \partial V_i$) is shown at Fig. 10. Each gradient exhibits more sensibility beneath its position at the surface, even though the surface is fully covered with receivers. While the image is not clear for each one, the addition of all gradients into a single one produces a velocity gradient with fine resolution.

Computing the gradient is readily parallelizable. We implemented this part of our problem using Message Passing Interface (MPI) in Fortran 90 and compute the gradient for each source in parallel in a computer cluster.¹ Taking advantage

¹ We used the cluster Lamb of the supercomputing lab at the Specialized Labs System of the Earth Sciences Division of CICESE. Each node is equipped with 20 cores and we used for all examples 5 nodes, for a total of $5 \times 20 = 100$ MPI cores, equal to the number of sources.

Fig. 11 Illustration of a cost function as a function of the steps α_i . An ideal step size would be the one leading closer to the global minimum



of the fact that each source-gradient computation is independent of the others, we compute each gradient simultaneously following Eq. 71, then Eq. 72 is obtained by combining the values from all processes into a main MPI core using an *MPI_ALLREDUCE* operation.

The total gradient (Fig. 10 bottom right) resembles the footprint of the layers for the Marmousi model and it is similar to a typical seismic migration. This velocity model is added to the starting model using a scalar factor that needs to be carefully chosen. The gradient-based optimization minimizes Eq. 56 by updating the velocity model iteratively as follows

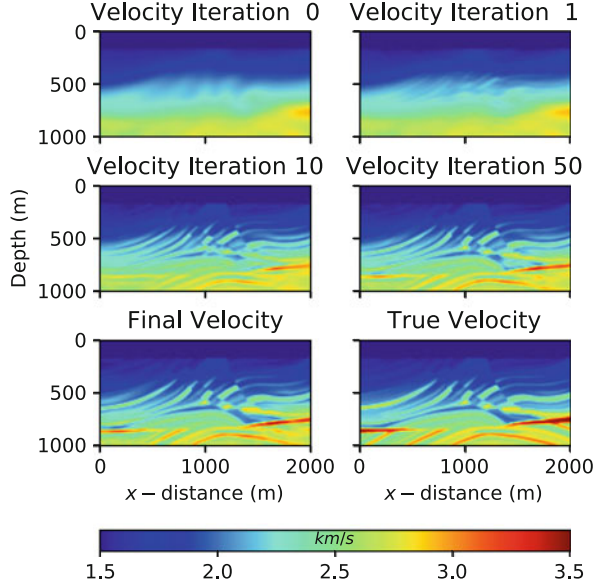
$$V_{n+1} = V_n + \alpha_n \partial V, \quad (73)$$

where the scalar α_n is the step length which represents how much the current model V_n moves along the direction ∂V at the n -th iteration. The efficiency of the minimization depends on the choice of the step α_n which can lead to local or global minima as illustrated in Fig. 11.

There are several algorithms for the search of the optimal step length α_n [30]. For this work we used a step line search method using interval reduction. Consider the range of values of steps $\alpha_1 < \alpha_2 < \alpha_3 < \dots < \alpha_k$ with k the number of test points with their respective costs $\text{cost}_1, \text{cost}_2, \text{cost}_3, \dots, \text{cost}_k$. In this method, we select the value of α which corresponds to the minimum cost. If the minimum cost corresponds to the first test value then for the next iteration a *zoom in* is performed for the test points $[\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_k] \times \text{zoom}$ with $\text{zoom} < 1$, on the other hand if the optimal step correspond to the final point test a *zoom out* is performed as $[\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_k]/\text{zoom}$. A typical value of zoom is $\frac{\sqrt{5}-1}{2}$ corresponding to the reciprocal of the Golden ratio. The evaluation of the cost function for k different step sizes is computationally expensive, however, it is compensated by the effectiveness due to the optimal choice of α_n . The AFWI iterative scheme combining all the components is shown in algorithm 1.

Continuing with the example from Fig. 8 and Table 4, we show in Fig. 12 the inversion results at some of the iterations. For this example we used 10 test points in the step line search method ($k = 10$). In the first 10 iterations the stratigraphic features are recovered. Thereafter, the velocity values at each point of the model are

Fig. 12 Velocity Marmousi model after some iterations of FWI. The true velocity model is at the bottom-right



steadily recovered, with more resolution on the central part of the survey. The final velocity model after 228 iterations (Fig. 12 bottom left) closely resembles the true model.

To further illustrate how AFWI works, we show in Fig. 13 the seismogram from the station at (1000, 0) m with 5% of Gaussian noise and the seismograms computed using the starting and final models. Notice that the seismogram from the final model closely follows the observed seismogram.

A more accurate indicator for the quality of the FWI iterative process is the analysis of the objective function for each iteration (Fig. 14). The objective function for this example is reduced faster at early iterations and becomes slower for later iterations, because the stratigraphic information has been recovered first and, at the end of the process, only the velocity value is getting recovered slowly.

Algorithm 1: Typical AFWI process

Data: Observed pressure $P^{\text{obs}}[n_t, n_r, n_s]$, starting velocity model V

Input : FWI iterations n_{iter} .

for $i_{\text{iter}} = 1 : n_{\text{iter}}$ **do**

- ┌ Computes forward modeling and adjoint field;
- ├ Performs step-line search for α ;
- └ Update velocity model $V_{i_{\text{iter}}+1} = V_{i_{\text{iter}}} + \alpha \partial V$

Output: Velocity model V , data residual $R[n_t, n_r, n_s]$

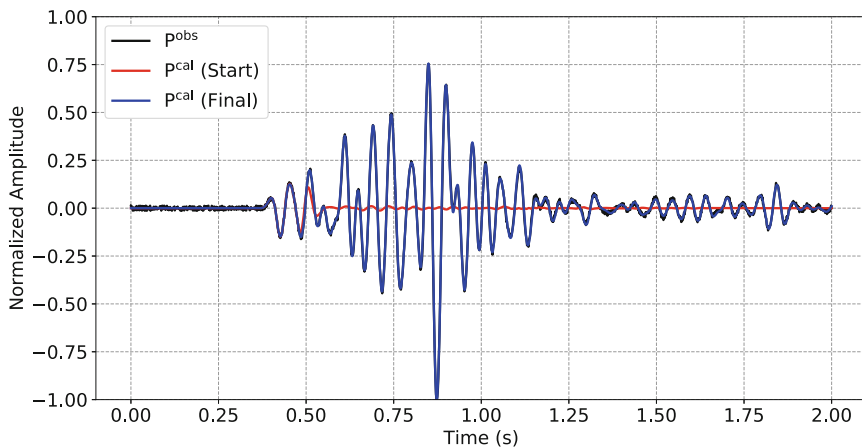


Fig. 13 Seismogram comparison for starting (red) and final (blue) synthetic data with respect to the observed data corresponding to a single source and a single receiver for the Marmousi model FWI example

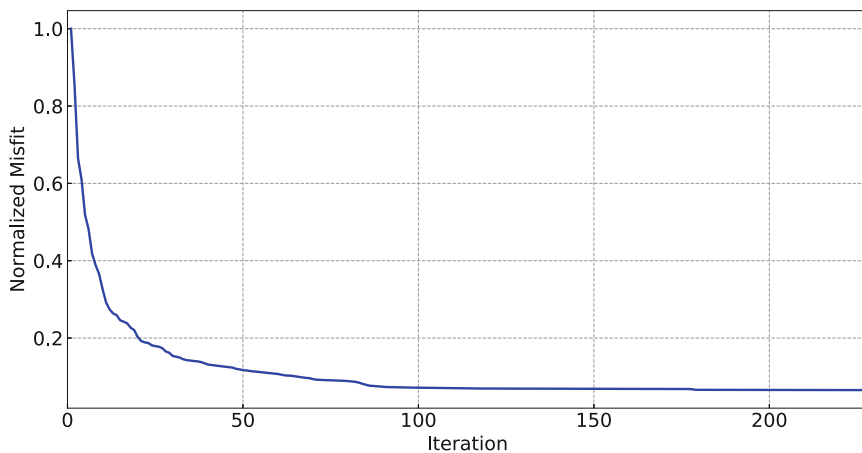


Fig. 14 Objective function (cost, misfit) reduction for 228 iterations of FWI for the Marmousi model example

3.4 Elastic Full Waveform Inversion

The tools and algorithm applied to the acoustic case can be used for Elastic FWI (EFWI) by replacing the forward modeling. The objective function for EFWI is given by

$$Q(m) = \frac{1}{2} \sum_s \sum_r \int_0^T \|\mathbf{u}_{r,s}^{\text{obs}} - \mathbf{u}_{r,s}^{\text{cal}}\|^2 dt, \quad (74)$$

where $\mathbf{u}_{r,s}^{\text{obs}}$ is the observed displacement and $\mathbf{u}_{r,s}^{\text{cal}}$ is the synthetic displacement computed using the elastodynamic wave equation. T is the total time of recording, r is the receiver index and s is the source index. The displacements can be u_x , u_y and/or u_z (or velocities v_x , v_y , v_z) for a model m which depends on the Lamé parameters and density (or velocities V_P and V_S).

As in the acoustic case, the direct minimization of Eq. 74 involves the computation of the perturbations, which increase even more the computational cost for elastic media because the displacement (or velocity) fields are vectors. The same procedure as AFWI can be pursued using the adjoint-state method. The mathematical deduction of the gradients will not be detailed, however, notice that the second-order derivatives are self-adjoint operators. See [42] for further details of the adjoint method for elastic media.

For an isotropic media we require the gradients for density (δ_ρ), shear modulus (δ_μ) and bulks modulus (δ_κ), given by Tromp et al. [42]

$$\delta_\rho(\mathbf{x}) = - \sum_r \int_0^T \rho(\mathbf{x}) \mathbf{u}^\dagger(\mathbf{x}, T-t) \cdot \partial_t^2 \mathbf{u}(\mathbf{x}, t) dt \quad (75)$$

$$\delta_\mu(\mathbf{x}) = - \sum_r \int_0^T 2\mu(\mathbf{x}) \mathbf{D}^\dagger(\mathbf{x}, T-t) : \mathbf{D}(\mathbf{x}, t) dt \quad (76)$$

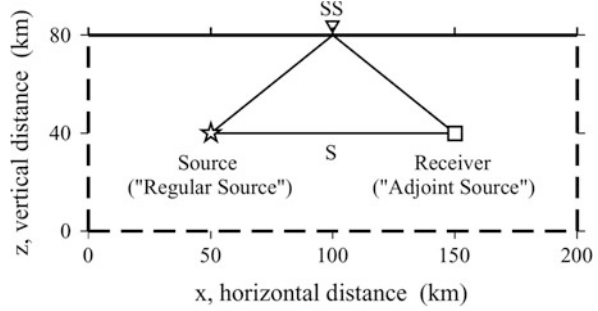
$$\delta_\kappa(\mathbf{x}) = - \sum_r \int_0^T \kappa(\mathbf{x}) [\nabla \cdot \mathbf{u}^\dagger(\mathbf{x}, T-t)] [\nabla \cdot \mathbf{u}(\mathbf{x}, t)] dt \quad (77)$$

where $:$ is a double dot product operator between tensors, and \mathbf{D} denotes the deviatoric strain, defined as

$$\mathbf{D} = \frac{1}{2} [\nabla \mathbf{u} + (\nabla \mathbf{u})^T] - \frac{1}{3} (\nabla \cdot \mathbf{u}) \mathbf{I}. \quad (78)$$

Notice that these computations involve more complex operations than in the acoustic case. The adjoint deviatoric strain \mathbf{D}^\dagger is computed using the equation for \mathbf{D} but using \mathbf{u}^\dagger . The elastic gradient can be expressed in terms of the shear-wave

Fig. 15 Source—receiver geometry for the computation of the elastic kernels. Taken from [42]



velocity

$$\delta_{V_S}(\mathbf{x}) = 2 \left(\delta_{\mu} - \frac{4}{3} \frac{\mu}{\kappa} \delta_{\kappa} \right), \quad (79)$$

and the compressional-wave velocity

$$\delta_{V_P}(\mathbf{x}) = 2 \left(1 + \frac{4}{3} \frac{\mu}{\kappa} \delta_{\kappa} \right), \quad (80)$$

then, a step line search can be used to obtain the model parameters iteratively. Following the work of Tromp et al. [42], the source-receiver geometry for an isotropic elastic media with homogeneous properties (Fig. 15) is used.

Following the same procedure as in the previous section for acoustic media, the wave propagation for the horizontal displacement and the back-propagation for the adjoint horizontal displacement is shown in Fig. 16 for 52 seconds of recording time. For illustration purposes, the P-wave velocity kernel is shown in the third column. The gradient shows the so-called banana-doughnut shape, which is related to the ray path [42].

3.5 Cooperative Inversion

In a joint-inversion scheme, different geophysical forward problems are solved to obtain a consistent Earth-property model that matches the respective data sets measured at the surface. Usually, the strategy consists of combining all the parameters into one objective function, leading to a large system of often disparate parameters [34]. There are mainly three types of joint inversion techniques, depending on the construction of the cost function:

- Petrophysical joint inversion, where the models are constrained by an empirical relationship [6, 25, 34, 36],

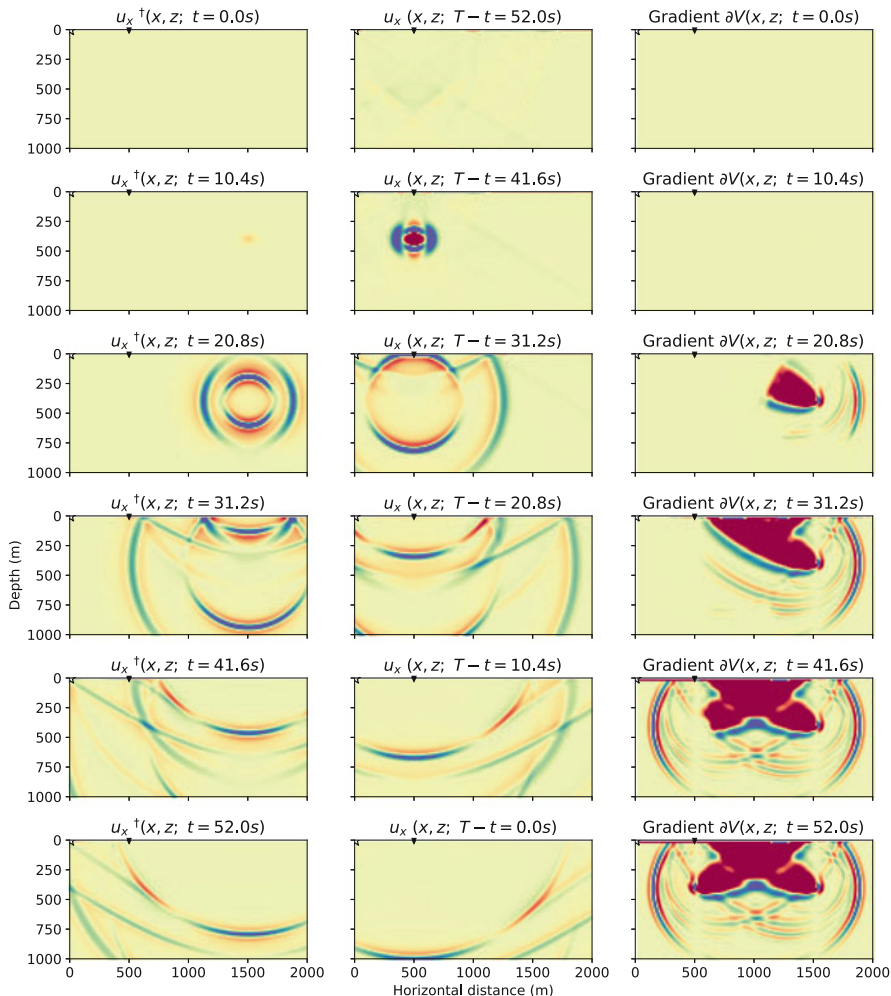


Fig. 16 Regular displacement u_x and adjoint displacement u_x^\dagger wave propagation for 52 seconds of recording time for the construction of the P-wave velocity kernel

- Structural joint inversion [14, 15], where the functional is used to match the structure for both models through the cross gradient, and
- Statistical joint inversion, e.g. using the fuzzy c -means technique [31, 33].

We will focus on the petrophysical joint inversion to combine FWI and GI. We propose a cooperative and sequential approach in which we solve at different stages for the densities and velocities. The resulting system is, therefore, more manageable and there is more control over the parameters at each stage. We call this a cooperative strategy to distinguish it from the joint strategies that solve all the geophysical parameters together at every iteration. Unlike conventional

joint inversions, where the problem is to minimize a two-part objective function (e.g. seismic and gravity errors), this cooperative inversion is based on alternately minimizing the errors in seismic and gravity data iteratively [36]. The main reasons to perform these sequentially are to increase robustness, reduce the computational cost, and keep always a strong control in the GI, avoiding the natural behavior of this potential method to yield shallower models. Furthermore, in the proposed scheme we do not need to impose depth-dependent weights or constraints to the GI to avoid shallower models, this is achieved instead by using the velocity model from FWI as the a priori gravimetric model. Another advantage of this approach is that, regardless of the model obtained from fitting a gravity anomaly, the total mass is uniquely recovered as implied by Gauss' theorem [18]. This means that, although gravity is a low-resolution geophysical tool, it does provide unique information linked to the velocity model. We seek to minimize the gravimetric data constrained with the velocity model obtained after an FWI process using the following objective function

$$Q(\mathbf{m}_\rho) = \sum_{i=1}^{N_s} \left\| \frac{\mathbf{g}_{z_i}^{\text{obs}} - \mathbf{A}\mathbf{m}_\rho}{\sigma_{\mathbf{g}_{z_i}}} \right\|^2 + \alpha_{\text{reg}}^2 \|\nabla \mathbf{m}_\rho\|^2 + \beta^2 \|\mathbf{m}_\rho - \mathbf{m}_{\rho(V)}\|^2, \quad (81)$$

where \mathbf{m}_ρ is the density model obtained using a petrophysical relationship as a function of the velocity model obtained from AFWI or EFWI. β is the parameter that weights the role on the inversion of seismic versus GI. Higher values of β yield results closer to the seismic model and vice versa. Our results will focus more on the velocity model from FWI to avoid shallower models due to a weakly-restricted GI. Then the density model will give feedback to the velocity model using an empirical relationship for the next FWI iteration. We use the following relationship from Gardner et al. [17] as petrophysical constraint,

$$\mathbf{m}_{\rho(V)} = \rho = \rho_0 V_P^{k_0}, \quad (82)$$

with $\rho_0 = 0.31 \text{ g/cm}^3$ and $k_0 = 0.25$. Other density-velocity petrophysical relationships are readily available in the literature and can easily be incorporated into our proposed scheme. For example, Brocher [7] computed the following polynomial fits for density as a function of velocity

$$\rho(\text{g/cm}^3) = 1.6612V_P(\text{km/s}) - 0.4721V_P^2 + 0.0671V_P^3 - 0.0043V_P^4 + 0.000106V_P^5, \quad (83)$$

and velocity as a function of density

$$V_P(\rho) = 39.128\rho - 63.064\rho^2 + 37.083\rho^3 - 9.1819\rho^4 + 0.8228\rho^5. \quad (84)$$

These are valid for densities between $2.0 < \rho < 3.5 \text{ g/cm}^3$ and velocities in the range $1.5 < V_p < 8.5 \text{ km/s}$ respectively. However, since both of Brocher's equations are based on polynomial fits, they are not inversely related. An iterative procedure using Eqs. 83 and 84 will not lead to the same velocity-density values. For example, starting from a velocity of 3500 m/s, a density of 2.318 g/cm^3 is computed using Eq. 83, then, using Eq. 84 to get the corresponding velocity, we obtain a value of 3692.34 m/s, a change of 192.34 m/s (5.49%). Therefore, since we require that the two functions be inverse of each other, we would have to do some adjustments to incorporate these petrophysical relations into our scheme.

The CGLS method is implemented in a straightforward way modifying \mathbf{G} and \mathbf{d} from Eqs. 54 and 55 as follows

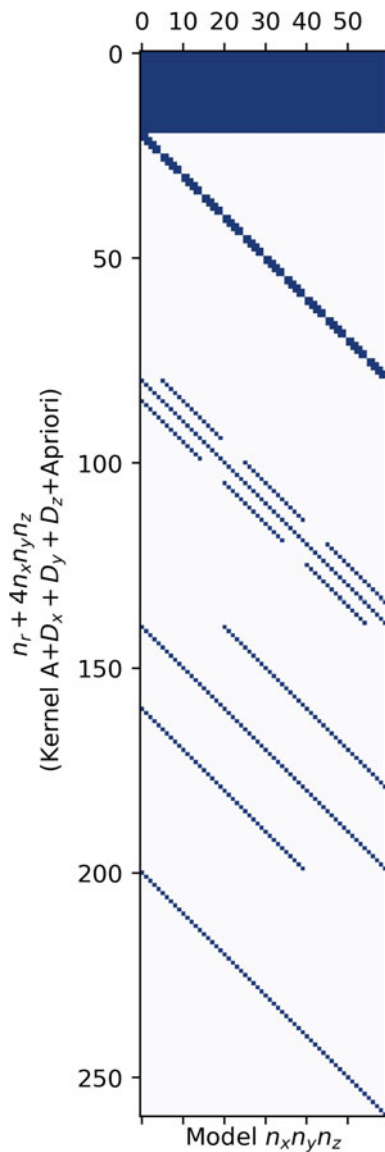
$$\mathbf{G} = \begin{bmatrix} \mathbf{C}_{dd}^{-1/2} \mathbf{A} \\ \alpha_{\text{reg}} \mathbf{D} \\ \beta \mathbf{I} \end{bmatrix}, \quad (85)$$

$$\mathbf{d}^{\text{CG}} = \begin{bmatrix} \mathbf{C}_{dd}^{-1/2} \mathbf{g}_z^{\text{obs}} \\ 0 \\ \beta \mathbf{m}_{\rho(V_p)} \end{bmatrix}, \quad (86)$$

where \mathbf{I} is the identity matrix. Once again, \mathbf{G} is a large and sparse matrix. For example, for a discretization with $n_x = 5$, $n_y = 4$ and $n_z = 3$, we would have a matrix of 15,600 element, of which only 1706 are non zero elements (a sparsity of 10.9%, see Fig. 17), whereas the square matrix of a Gauss-Newton implementation would have to store 3600 elements. The procedure to solve the system $\mathbf{Gm} = \mathbf{d}^{\text{CG}}$ is shown in algorithm 2 [35]. An efficient implementation of this algorithm requires that all the matrices be stored in a sparse representation, we use Coordinate Format (COO) sparse matrices for this.

In summary, this cooperative inversion scheme for gravity and seismic data consists of the following iterative steps: From a starting velocity model, we perform FWI to update the velocity model, then, using Gardner's density-velocity relationship, we perform constrained GI to update the density model, finally, using Gardner's velocity-density relationship, a velocity model is obtained that will be the starting model to solve FWI.

Fig. 17 Large sparse structure of the matrix \mathbf{G} for the CGLS method. The blue spots represent non-zero elements and the white spaces zero elements



Algorithm 2: CGLS algorithm to iteratively solve the problem $Gm = d^{CG}$

Data:
 $m_0, d_0 = d^{CG}, r_0 = G^T d^{CG}$
 $p_0 = r_0, t_0 = Ap_0, n$ -iterations
Result: model m_k
while $k < n$ **and** $\|r_k\| < 10^{-15}$ **do**
 $\alpha_k = \|r_{k-1}\|^2 / \|t_{k-1}\|^2$
 $m_k = m_{k-1} + \alpha_k p_{k-1}$
 $d_k = d_{k-1} - \alpha_k t_{k-1}$
 $r_k = G^T d_k$
 $\beta_k = \|r_k\|^2 / \|r_{k-1}\|^2$
 $p_k = r_k + \beta_k p_{k-1}$
 $t_k = G p_k$
 $k \leftarrow k + 1$

4 Results

In order to test the proposed cooperative inversion algorithm and demonstrate its advantages, we apply this method on two synthetic examples for 2D elastic media.

4.1 EFWI: Marmousi Model

Let us consider again the Marmousi model. The geometry and parameters are the same as those of the example in Sect. 3.3 (see Fig. 8 and Table 4). The S -wave velocity is computed using $V_S = V_P / \sqrt{3}$ and the density is obtained using Gardner petrophysical relationship; the models for V_S and ρ are not shown. Only the vertical component of the displacement is considered and we use 10 test points for the step line search. We did not add Gaussian noise to the data in this example.

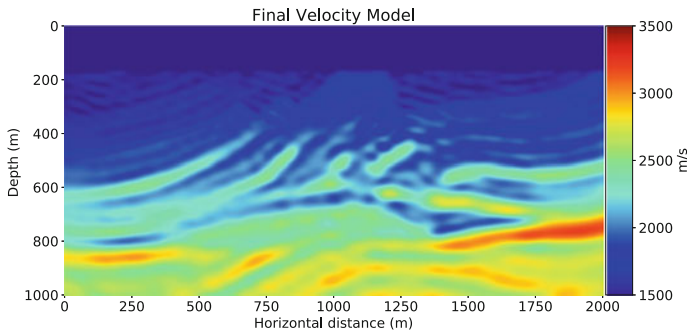


Fig. 18 The final velocity model after 100 iterations of EFWI for the Marmousi example

The velocity model obtained after 100 iterations, shown in Fig. 18, resembles the stratigraphic information of the Marmousi model (Fig. 8). This result shows spurious artifacts in the final model. This problem is attributed to the problem of the limited bandwidth of the observed data [40]. The artifacts are also related to the S-waves since they are usually not present in AFWI. These artifacts yield small errors between the observed and computed seismograms but contaminates the iterative process and affect the convergence.

The convergence of the EFWI iterative process can be analysed from the behaviour of the objective function. We observe from Fig. 19 that the objective function for this example converges fast at early iterations and becomes stagnant for later iterations. This is mostly due to the presence of the spurious artifacts mentioned before.

We show in Fig. 20 the observed and computed seismograms for the station located at (1000, 0) m corresponding to the 20th source located at (404.04, 0) m. Overall, the seismogram for the final model closely approximates the observed data. The phase of all the events is matched very well, however, there are discrepancies

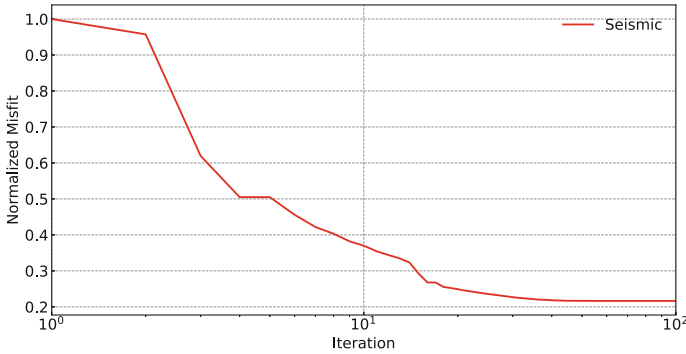


Fig. 19 Normalized objective function for the Marmousi model using EFWI

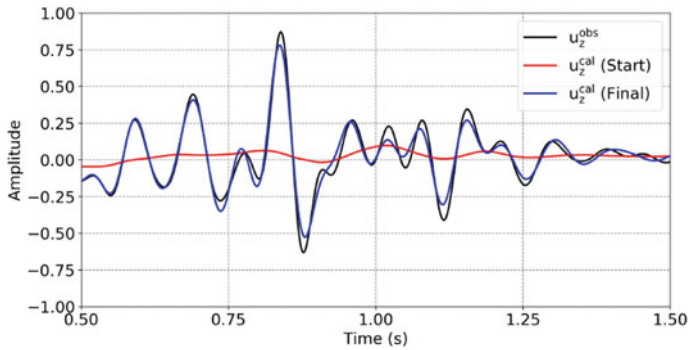


Fig. 20 Seismogram comparison of starting (red) and final (blue) synthetic data for EFWI with the observed data (black) corresponding to a single source and a receiver located at (1000, 0) m

in the amplitudes. Unfortunately, these results can not be improved using more iterations, we would have to rely on the implementation of additional techniques to get a better approximation.

4.2 *EFWI: Texas-Shaped Model*

For this example, we created a laterally heterogeneous layered model covering a horizontal distance of 2000m and a depth of 1000m for both seismic and gravimetric data. Beneath the low-velocity layers, we place a structure with the shape of Texas, as shown in Fig. 21a. The shallow layers have lower velocities (between 1500 and 2000 m/s) with respect to the deepest layer (~ 3500 m/s). The S -wave velocity is computed using $V_S = V_P/\sqrt{3}$ and we used Gardner's density-velocity relationship to obtain the density model. Notice that the maximum velocity used for this example is 3500 m/s, hence Gardner's equation applies to this example. We use the same discretization parameters from the previous example (Table 4).

The initial model and the result after 48 iterations are shown in Fig. 21b, c. The top layers show many spurious artifacts similar to those in the previous example. These artifacts persist if we continue iterating the method. The objective function, shown in Fig. 22, exhibits a convergent behaviour until 40th iterations where stagnation is reached.

4.3 *Cooperative Inversion: Marmousi Model*

Let us apply the cooperative scheme on the Marmousi model (Fig. 8). We used the same model parameters of Table 4, with the addition of 200 equispaced gravimetric stations in the surface. The true velocity model and the starting model are the same as in the AFWI example of Sect. 3.3, and the S -wave velocity and density models are the same as in Sect. 4.1.

In order to compare the result, we performed 100 iterations of the cooperative inversion and show the results, together with those of EFWI, in Fig. 23. Incorporating GI helps to eliminate the spurious artifacts and smooths the model (Fig. 23b). This is because GI acts as a filter in the cooperative inversion. Each iteration has a computational cost of 75 minutes for EFWI and 76 minutes for the cooperative scheme.

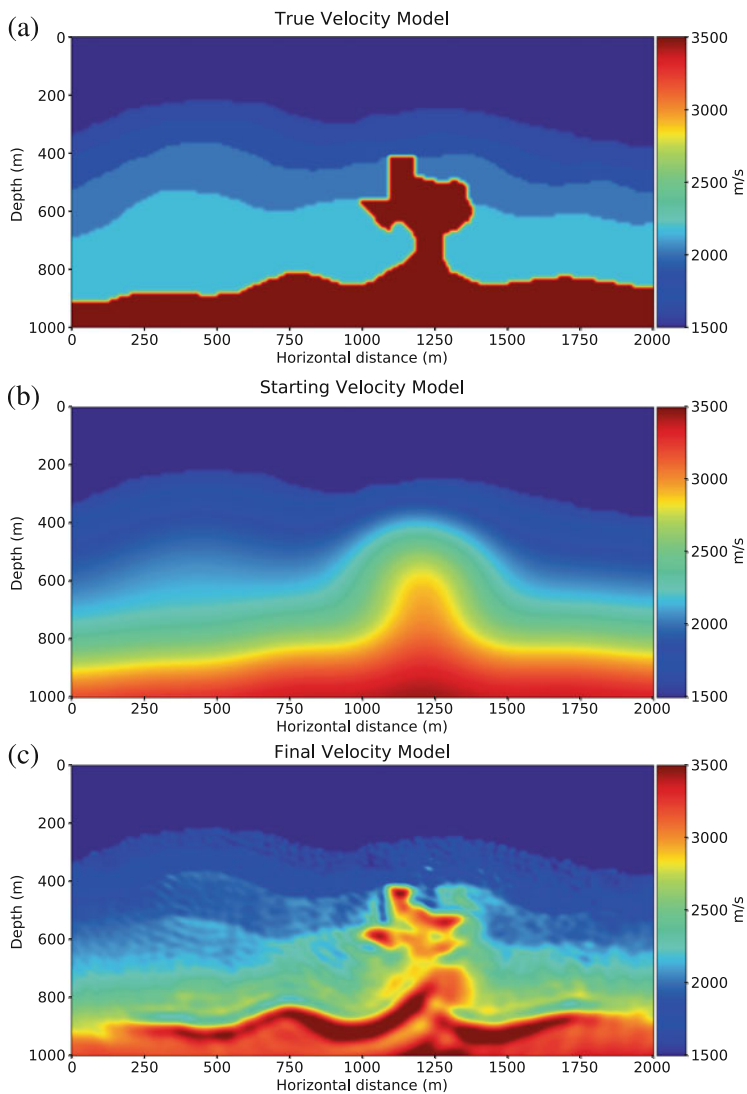


Fig. 21 (a) Texas-shape true velocity model, and (b) and its smoothing set as a starting model. (c) Final velocity model after EFWI

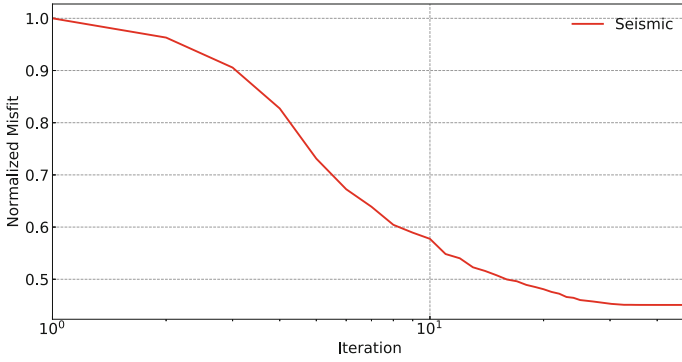


Fig. 22 Normalized misfit for seismic data for the Texas-shape model after EFWI

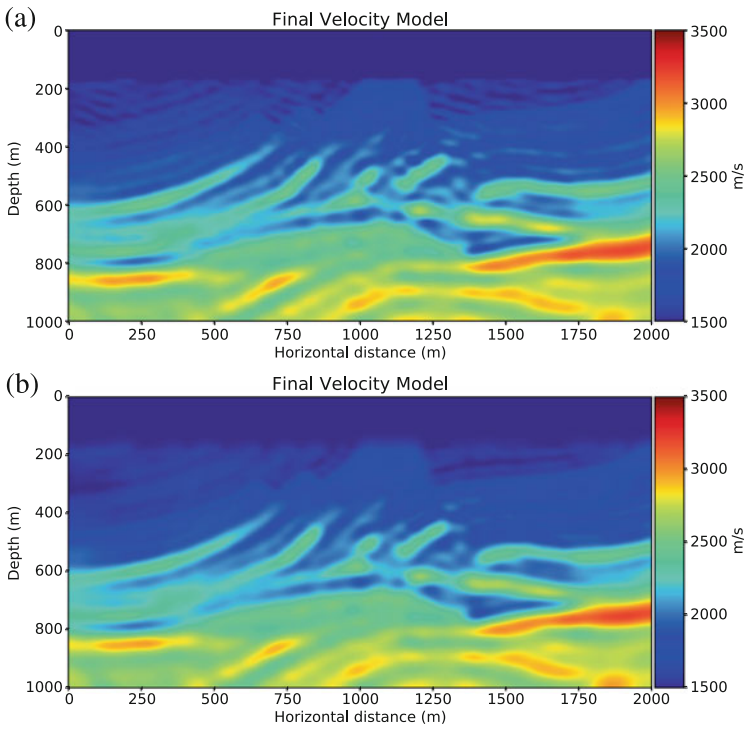


Fig. 23 Marmousi final velocity model obtained using (a) conventional EFWI, and (b) the cooperative inversion (EFWI and GI).

4.4 Cooperative Inversion: Texas-Shaped Model

As a final example, let us apply cooperative inversion to the Texas-shaped model of Fig. 21a. The true velocity model, the starting model and other parameters are the same as in Sect. 4.2. The final velocity model after 48 iterations is shown in Fig. 24 together with the results of EFWI to facilitate the comparison. Comparing Fig. 24a, b, we observe that cooperative inversion reduces the artifacts that pollute the EFWI results. The density models obtained from conventional GI and cooperative inversion are shown in Fig. 25. Figure 25a shows the typical behavior of conventional GI of giving preference to shallower models, whereas the model obtained from cooperative inversion (Fig. 25b) yields significantly better model. In order to illustrate the data fit, we show in Fig. 26 the seismograms for a station at (1000, 0) m from a source at (404.04, 0) m, and the gravimetric anomaly. The seismogram for the final model show a good agreement with the observed seismogram, with small discrepancies in phase and amplitude (Fig. 26a). The computed gravimetric anomaly has small discrepancies with the observed anomaly (Fig. 26b). We emphasize that

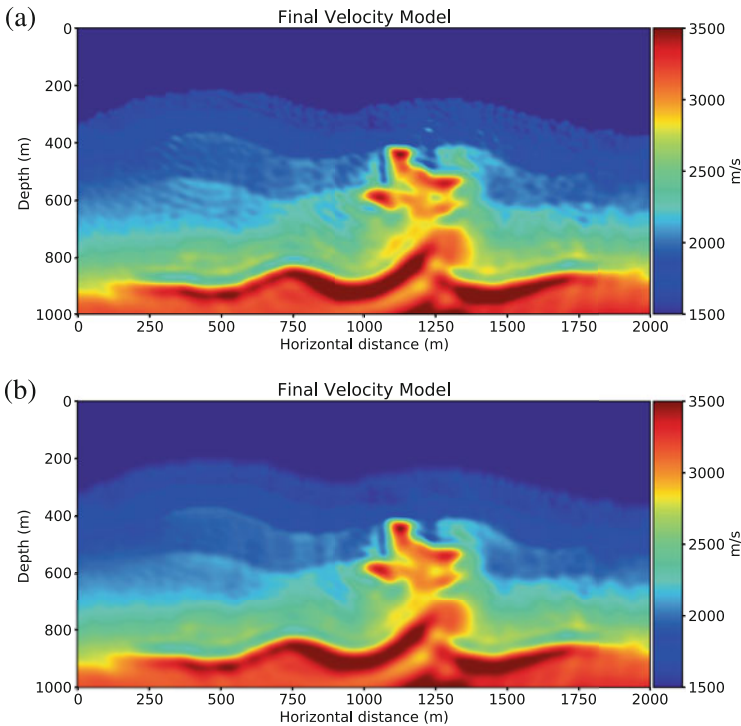


Fig. 24 Texas-shape final velocity model after 48 iterations, comparison between (a) conventional EFWI, and (b) the cooperative inversion

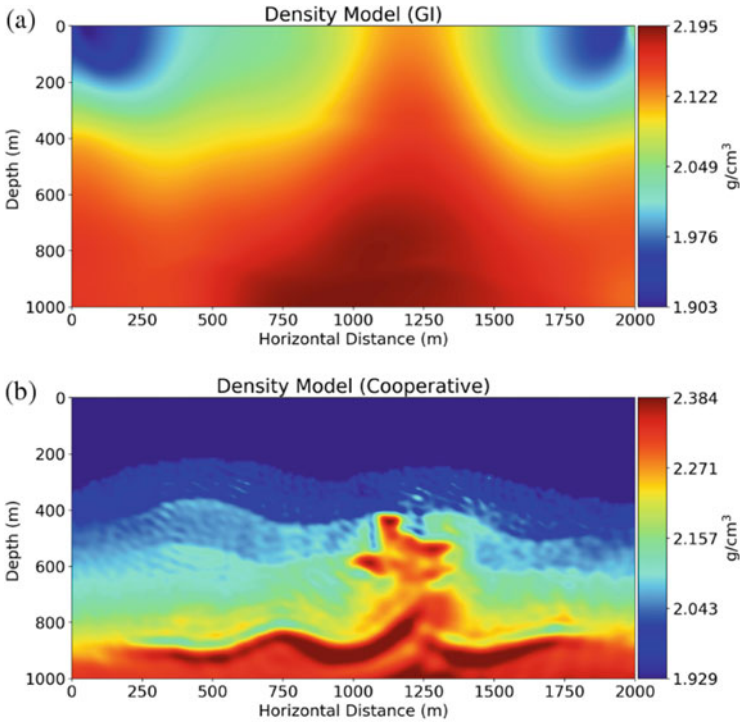


Fig. 25 Texas-shape final density model after 48 iterations, comparison between (a) conventional gravimetric inversion (GI), and (b) the cooperative inversion

the cooperative inversion does not aim to exactly fit all the data but to obtain a realistic model.

A comparison of the seismic misfit for EFWI and the cooperative scheme is presented in Fig. 27. The misfits exhibit a similar reduction at earlier iterations, however, later the cooperative inversion adjusts better the seismic traces given the elimination of the artifacts.

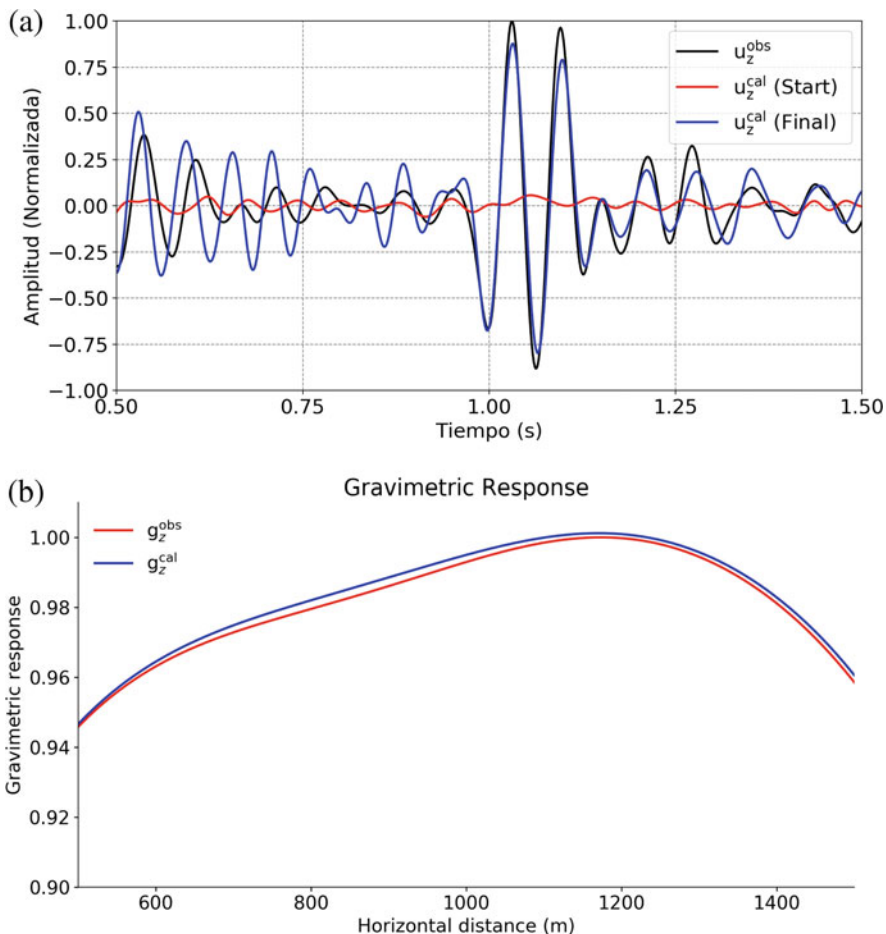


Fig. 26 Data fit for the Texas-shape model after 48 iterations of the cooperative scheme. (a) Vertical-displacement seismograms for a station at (1000, 0) m from a source at (404.04, 0) m, and (b) Observed and computed gravity anomaly

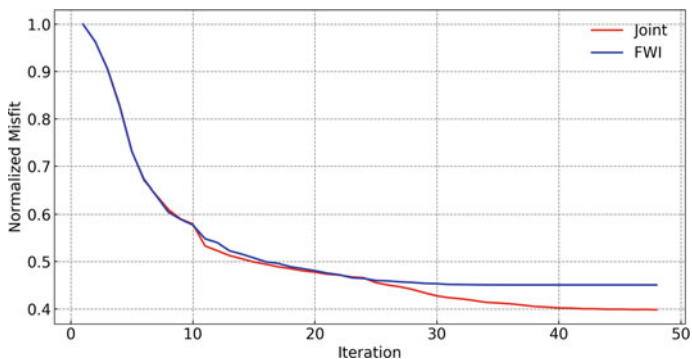


Fig. 27 Normalized misfit reduction for seismic data for cooperative inversion (red) and separated inversion (blue)

5 Conclusions

We have developed a cooperative scheme that combines GI and FWI. The methods are combined in an iterative scheme based on petrophysical relationships that can be used to characterize typical geological environments found in real field data such as irregular high velocity bodies embedded in complex horizontal layers, thrust and dominoes fault systems, for both acoustic and elastic media.

The synthetic examples for elastic media show that both methods converge as long as the starting model is acceptable and FWI has more weight in the cooperative inversion algorithm. The models recover the stratigraphic part, the fault dip, the discordances and the top and shape of the high velocity and density bodies. The weights in the cost function play a critical role in the trade off between the convergence rate and the accuracy of the resulting models. Further analysis is required to determine optimal weights.

Comparing the results obtained by separate and cooperative inversion, we observe that the cooperative scheme helps improve the density models of GI by constraining them to the FWI models. FWI also benefits from the cooperative inversion, avoiding the saturation in parts of the model and reducing the presence of the spurious reflectors. Finally, the sequential implementation of the cooperative scheme has a negligible additional computational cost compared to the classical FWI.

Acknowledgments We want to thank CONACYT and CeMIEGeo for their financial support. The numerical computations were performed using the cluster Lamb of the supercomputing lab at the Specialized Labs System of the Earth Sciences Division of CICESE.

References

1. K. Aki, P.G. Richards, *Quantitative Seismology*, vol. 1 (University Science Books, Sausalito, 2002)
2. R. Alford, K. Kelly, D.M. Boore, Accuracy of finite-difference modeling of the acoustic wave equation. *Geophysics* **39**(6), 834–842 (1974)
3. B. Banerjee, S. Das Gupta, Gravitational attraction of a rectangular parallelepiped. *Geophysics* **42**(5), 1053–1055 (1977)
4. B. Bhattacharyya, L.-K. Leu, Spectral analysis of gravity and magnetic anomalies due to rectangular prismatic bodies. *Geophysics* **42**(1), 41–50 (1977)
5. R.J. Blakely, *Potential Theory in Gravity and Magnetic Applications* (Cambridge University Press, Cambridge, 1996)
6. N. Blom, C. Boehm, A. Fichtner, Synthetic inversions for density using seismic and gravity data. *Geophys. J. Int.* **209**(2), 1204–1220 (2017)
7. T.M. Brocher, Empirical relations between elastic wavespeeds and density in the earth's crust. *Bull. Seismol. Soc. Am.* **95**(6), 2081–2092 (2005)
8. G. Cohen, *Higher-Order Numerical Methods for Transient Wave Equations* (Springer, Berlin, 2012)
9. D. Colombo, D. Rovetta, Coupling strategies in multiparameter geophysical joint inversion. *Geophys. J. Int.* **215**(2), 1171–1184 (2018)

10. J.D. De Basabe, M.K. Sen, A comparison of finite-difference and spectral-element methods for elastic wave propagation in media with a fluid-solid interface. *Geophys. J. Int.* **200**(1), 278–298 (2015)
11. A. Fichtner, *Full Seismic Waveform Modelling and Inversion* (Springer, Berlin, 2010)
12. A. Fichtner, H.-P. Bunge, H. Igel, The adjoint method in seismology: I. Theory. *Phys. Earth Planet. Interiors* **157**(1–2), 86–104 (2006)
13. B. Frehner, Monumental geophysics: J. clarence karcher and the reflection method. *Lead. Edge* **40**(6), 404–407 (2021). <https://doi.org/10.1190/le40060404.1>
14. L.A. Gallardo, M.A. Meju, Characterization of heterogeneous near-surface materials by joint 2D inversion of dc resistivity and seismic data. *Geophys. Res. Lett.* **30**(13), 1658–1661 (2003)
15. L.A. Gallardo, M.A. Meju, Joint two-dimensional dc resistivity and seismic travel time inversion with cross-gradients constraints. *J. Geophys. Res. Solid Earth* **109**(B3), B03311 (2004)
16. J. García-Abdeslem, The gravitational attraction of a right rectangular prism with density varying with depth following a cubic polynomial. *Geophysics* **70**(6), J39–J42 (2005)
17. G. Gardner, L. Gardner, A. Gregory, Formation velocity and density—the diagnostic basics for stratigraphic traps. *Geophysics* **39**(6), 770–780 (1974)
18. F.S. Grant, G.F. West, *Interpretation theory in applied geophysics* (New York: McGraw-Hill, 1965)
19. D.D. Jackson, Interpretation of inaccurate, insufficient and inconsistent data. *Geophys. J. Int.* **28**(2), 97–109 (1972)
20. L.R. Johnson, J.J. Litehiser, A method for computing the gravitational attraction of three-dimensional bodies in a spherical or ellipsoidal earth. *J. Geophys. Res.* **77**(35), 6999–7009 (1972)
21. K. Kelly, R. Ward, S. Treitel, R. Alford, Synthetic seismograms; a finite-difference approach. *Geophysics* **41**(1), 2–27 (1976)
22. D. Komatitsch, J. Tromp, Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophys. J. Int.* **139**(3), 806–822 (1999)
23. J.M. Lees, J. VanDecar, Seismic tomography constrained by bouguer gravity anomalies: applications in Western Washington. *Pure Appl. Geophys.* **135**(1), 31–52 (1991)
24. F.-C. Lin, B. Schmandt, V.C. Tsai, Joint inversion of rayleigh wave phase velocity and ellipticity using usarray: constraining velocity and density structure in the upper crust. *Geophys. Res. Lett.* **39**(12), L12303 (2012)
25. W. Lin, M. Zhdanov, Joint inversion of seismic and gravity gradiometry data using gramian constraints, in *SEG Technical Program Expanded Abstracts 2017* (Society of Exploration Geophysicists, Houston, 2017)
26. K.J. Marfurt, Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations. *Geophysics* **49**(5), 533–549 (1984)
27. V. Menichetti, A. Guillen, Simultaneous interactive magnetic and gravity inversion. *Geophys. Prospect.* **31**(6), 929–944 (1983)
28. W. Menke, *Geophysical Data Analysis: Discrete Inverse Theory* (Academic Press, Cambridge, 2018)
29. D. Nagy, The gravitational attraction of a right rectangular prism. *Geophysics* **31**(2), 362–371 (1966)
30. J. Nocedal, S. Wright, *Numerical Optimization* (Springer, Berlin, 2006)
31. H. Paasche, J. Tronicke, Cooperative inversion of 2D geophysical data sets: a zonal approach based on fuzzy c-means cluster analysis. *Geophysics* **72**(3), A35–A39 (2007)
32. R.-E. Plessix, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophys. J. Int.* **167**(2), 495–503 (2006)
33. A. Romero, L.A. Gallardo, Borehole constrained inversion of geophysical data based on fuzzy clustering, in *SEG Technical Program Expanded Abstracts 2015* (Society of Exploration Geophysicists, Houston, 2015), pp. 1792–1796
34. L. Roy, M.K. Sen, K. McIntosh, P.L. Stoffa, Y. Nakamura, Joint inversion of first arrival seismic travel-time and gravity data. *J. Geophys. Eng.* **2**(3), 277 (2005)

35. M.K. Sen, P.L. Stoffa, *Global Optimization Methods in Geophysical Inversion* (Cambridge University Press, Cambridge, 2013)
36. R.U. Silva, J.D. De Basabe, M.K. Sen, M. Gonzalez-Escobar, E. Gomez-Trevino, S. Solorza-Calderon, Cooperative full waveform and gravimetric inversion. *J. Seismic Exploration* **29**(6), 549–573 (2020)
37. M. Talwani, Computation with the help of a digital computer of magnetic anomalies caused by bodies of arbitrary shape. *Geophysics* **30**(5), 797–817 (1965)
38. A. Tarantola, Inversion of seismic reflection data in the acoustic approximation. *Geophysics* **49**(8), 1259–1266 (1984)
39. A. Tarantola, A strategy for non linear inversion of seismic reflection data. *Geophysics* **51**(10), 1893–1903 (1986)
40. F. ten Kroode, S. Bergler, C. Corsten, J.W. de Maag, F. Strijbos, H. Tijhof, Broadband seismic data – the importance of low frequencies. *Geophysics* **78**(2), WA3–WA14 (2013). <https://doi.org/10.1190/geo2012-0294.1>
41. R. Tondi, U. Achauer, M. Landes, R. Daví, L. Besutiu, Unveiling seismic and density structure beneath the vrancea seismogenic zone, Romania. *J. Geophys. Res. Solid Earth* **114**(B11), B11307 (2009)
42. J. Tromp, C. Tape, Q. Liu, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophys. J. Int.* **160**(1), 195–216 (2005)
43. J. Virieux, P-SV wave propagation in heterogeneous media: velocity-stress finite-difference method. *Geophysics* **51**(4), 889–901 (1986)
44. J. Virieux, S. Operto, An overview of full-waveform inversion in exploration geophysics. *Geophysics* **74**(6), WCC1–WCC26 (2009)
45. K. Vozoff, D. Jupp, Joint inversion of geophysical data. *Geophys. J. Int.* **42**(3), 977–991 (1975)
46. R.A. Werner, The gravitational potential of a homogeneous polyhedron or don't cut corners. *Celestial Mech. Dyn. Astron.* **59**(3), 253–278 (1994)
47. O. Yilmaz, *Seismic Data Analysis* (Society of Exploration Geophysicists, Houston, 2001). ISBN 978-1-56080-098-9. <https://doi.org/10.1190/1.9781560801580>
48. H. Zeyen, J. Pous, 3-D joint inversion of magnetic and gravimetric data with a priori information. *Geophys. J. Int.* **112**(2), 244–256 (1993)
49. M.S. Zhdanov, R. Ellis, S. Mukherjee, Three-dimensional regularized focusing inversion of gravity gradient tensor component data. *Geophysics* **69**(4), 925–937 (2004)