

Sparse vector reflectivity inversion with full-wavefield LSRTM

Kristian Torres ¹ and Mauricio Sacchi

Abstract

This study introduces Sparse Full Waveform Least-Squares Reverse Time Migration (Sparse FWLSRTM), an innovative approach to seismic imaging that incorporates sparse regularization to enhance the quality of seismic data imaging. By comparing Sparse FWLSRTM with traditional Least-Squares Reverse Time Migration (LSRTM) methods, we demonstrate the former's superior capability in delineating geological reflectors with higher precision and reducing imaging artifacts and sidelobes, even when faced with highly decimated data and imprecise velocity models. Moreover, we show how FWLSRTM, with its parameterization based on vectorized reflectivity, can resolve a broader range of seismic events, including refractions, diving waves, and reflections, offering a superior fit for complex data compared to traditional methods. We outline the theoretical framework for Sparse FWLSRTM, including the derivation of misfit gradients and the application of sparse regularization techniques. Through various experiments and comparisons, we illustrate the potential of Sparse FWLSRTM as a robust tool for seismic imaging in complex geological settings, marking a significant advancement in geophysical prospecting.

¹Email: torresba@ualberta.ca

1.1 Introduction

Least-squares reverse time migration (LSRTM) offers distinct advantages in subsurface imaging, such as compensating for wavefield propagation effects and limited resolution typically seen with conventional migration. This makes it a preferred method to update and improve the reflectivity model in complex geological settings. However, it faces significant drawbacks, primarily due to its dependence on the first-order scattering approximation (Born modeling) for generating predicted data during the forward propagation of the linearized inversion. This limitation confines LSRTM to deal with only primary waves with near- to mid-reflection angles (Korsmo et al., 2022), implying the need for the challenging elimination of other wave modes in the observed data, despite their potential to contain significant information. Such a pre-processing step is essential in conventional LSRTM to prevent crosstalk in the imaging results, especially when dealing with complex seismic data that includes multiple reflections (Wong et al., 2015). Furthermore, due to the ill-posed and ill-conditioned nature (i.e., solutions are not unique and/or unstable), LSRTM results often contain migration artifacts and unwanted sidelobes around reflectors. These issues primarily arise due to the limited acquisition aperture, the restricted bandwidth of noisy seismic data, and inaccuracies in the modeling kernel (Zeng et al., 2014). On top of that, many different factors, such as sub-optimal acquisition design or rapidly changing velocity profiles, can play a role in creating shadow zones in migrated images or, broadly speaking, in images produced through least-squares migration. These shadow regions reside in the null space of the forward operator and exacerbate the problem of non-uniqueness in the solutions.

In an effort to alleviate the side effects of conventional LSRTM, we perform a non-linear least-squares inversion combining the full-wavefield vector reflectivity modeling engine (Whitmore et al., 2020) with a sparse constraint via the hyperbolic penalty function (HPF) (Claerbout, 2014). The full-wavefield modeling parameterizes the acoustic wave equation regarding velocity and vector reflectivity, eliminating the need to construct a density model. It generates seismic data with all waveforms, including multiples and transmission events, directly from the reflectivity without assuming they are linearly related. This enables reflectivity inversion with a more comprehensive and accurate match between observed and calculated data while bypassing the first-order Born approximation and the data pre-processing step needed in classic LSRTM. In addition, imposing the sparsity constraint as model-styling regularization helps to stabilize the solution and mitigate artifacts, thus enhancing the vertical resolution of the estimated seismic images.

Other non-linear LSRTM schemes can also refine the velocity model during the iterative migration process (Yao and Jakubowicz, 2012; Korsmo et al., 2022). Nonetheless, similar to Davydenko and

Verschuur (2017), we assume that a background velocity has been previously estimated and is not updated through iterations, reducing the non-linearity of the problem. In other words, we solely focus on inverting the vector reflectivity model. When contrasted with least-squares inversion lacking regularization, a 2D numerical experiment indicates that applying the HPF combined with the full-wavefield vector reflectivity engine promotes sparsity in the images while displaying fewer artifacts and more clearly delineated reflectors.

The outline of this paper is as follows. First, we review the forward operator of full-wavefield vector reflectivity modeling in a re-parameterized variable-density acoustic earth. We also derive its adjoint wave equation and present the corresponding stencils for forward and backward simulations with a damping boundary condition term in the time domain. These pair of forward and adjoint wavefields characterize what we have termed as the full-wavefield LSRTM framework, which is essentially a local optimization problem. Thus, we elaborate on its misfit function and gradient formulation. Then, we add sparsity constraints to the inverse problem of retrieving the horizontal and vertical vector reflectivity components from the recorded seismic data at the surface. Additionally, we detail the process of deriving a relative density model from the obtained vector reflectivity, which is a secondary result of our analysis. Finally, we demonstrate the effectiveness of the inversion with and without sparsity constraints on the Marmousi2 model (Martin et al., 2006).

1.2 Full-wavefield LSRTM

1.2.1 Forward modeling

In its continuous form, the 2D full-wavefield vector reflectivity acoustic wave equation used in this study as the forward modeling engine is given by (Whitmore et al., 2020)

$$\frac{1}{v_p(x, z)^2} \frac{\partial^2 u_s(x, z, t)}{\partial t^2} - \left(\nabla^2 u_s(x, z, t) + \frac{1}{v_p(x, z)} \nabla v_p(x, z) \cdot \nabla u_s(x, z, t) - 2\vec{r}(x, z) \cdot \nabla u_s(x, z, t) \right) = f_s(t), \quad (1.1)$$

where

$$\vec{r} = \begin{bmatrix} r_x \\ r_z \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \frac{\partial}{\partial x} \ln(z) \\ \frac{\partial}{\partial z} \ln(z) \end{bmatrix} \quad (1.2)$$

denotes the vector reflectivity with horizontal and vertical components, r_x and r_z , respectively, $z = \rho v_p$ represents the acoustic impedance, v_p is the velocity model, ρ the density, f_s is the scaled seismic source at the shot coordinates (x_s, z_s) , and $u_s(x, z, t)$ is the pressure wavefield associated to

that source (hence the subscript), dependent on time and space variables. It is assumed that the pressure wavefield has zero initial conditions, i.e., the wavefield does not have any energy before zero time:

$$u_s(x, z, t = 0) = 0, \quad \frac{\partial u_s(x, z, t = 0)}{\partial t} = 0. \quad (1.3)$$

Whitmore et al. (2020) show that equation 1.1 is simply a re-parameterization of the acoustic wave equation with variable density that avoids the explicit dependence on the density model by defining the vector reflectivity as the normalized rate of impedance change in each vector direction. Utilizing this same modeling engine, Fletcher et al. (2023) also suggest a method to diminish reflections from a specific boundary by averaging the wavefields from two independent forward simulations.

In this work, to simulate the forward modeling of the wavefield on the subsurface, we incorporate an extra dampening term, $\eta(x, z) \frac{\partial u(t, x, z)}{\partial t}$, to mimic an infinite domain (Sochacki et al., 1987). The factor η is set to 0 within the boundaries of the physical domain and progressively increases from the interior to the exterior within the damping layer. Depending on the width of the damping layer and the frequency band, η may exhibit either a linear or exponential increase. At the boundary of the artificial domain, we assume the solution fulfills the radiation boundary conditions

$$u_s(x, z, t) \Big|_{(x, z) \rightarrow \infty} \rightarrow 0, \quad \nabla u_s(x, z, t) \Big|_{(x, z) \rightarrow \infty} \rightarrow 0, \quad (1.4)$$

indicating that the wavefield is progressively damped as we move away from the source towards the computational domain's edge. Omitting variable dependence on space dimensions to simplify the notation, the expression for the vector reflectivity modeling with the absorbing boundary term is

$$\frac{1}{v_p^2} \frac{\partial^2 u_s(t)}{\partial t^2} - \left(\nabla^2 u_s(t) + \frac{1}{v_p} (\nabla v_p \cdot \nabla u_s(t)) - 2(\vec{r} \cdot \nabla u_s(t)) \right) + \eta \frac{\partial u_s(t)}{\partial t} = f_s(t). \quad (1.5)$$

Using a second-order approximation in time, we can construct the stencil necessary to compute the full solution of the acoustic wavefield over time:

$$\begin{aligned} u_s(t + dt) = & \frac{\frac{\partial}{\partial x} u_s(t) \frac{\partial}{\partial x} v_p + \frac{\partial}{\partial y} u_s(t) \frac{\partial}{\partial y} v_p}{v_p} - \frac{-\frac{2u_s(t)}{dt^2} + \frac{u_s(t-dt)}{dt^2}}{v_p^2} - 2r_x \frac{\partial}{\partial x} u_s(t) - 2r_y \frac{\partial}{\partial y} u_s(t) + \frac{\partial^2}{\partial x^2} u_s(t) + \frac{\partial^2}{\partial y^2} u_s(t) + \frac{\eta u_s(t)}{dt} \\ & \frac{\frac{\eta}{dt} + \frac{1}{dt^2 v_p^2}}{dt} \\ & + dt v_p^2 f_s(t) \end{aligned} \quad (1.6)$$

where dt represents the time-stepping interval, and $\frac{\partial}{\partial x}$, $\frac{\partial^2}{\partial x^2}$, $\frac{\partial}{\partial z}$, and $\frac{\partial^2}{\partial z^2}$ denote the first and second

order spatial derivatives in horizontal and vertical dimensions, which can be computed with central finite difference approximations.

Figure 1.1 compares the snapshots of the forward-propagated wavefield at time=2.5 seconds between the full-wavefield vector reflectivity modeling and Born modeling on the BP synthetic model. Similarly, Figure 1.2 compares a shot gather obtained via full-wavefield vector reflectivity modeling with a shot gather computed using Born modeling. As can be noticed in both images, the full-wavefield vector reflectivity resolves a broader range of seismic events, including refractions, diving waves, and reflections beyond the Born approximation.

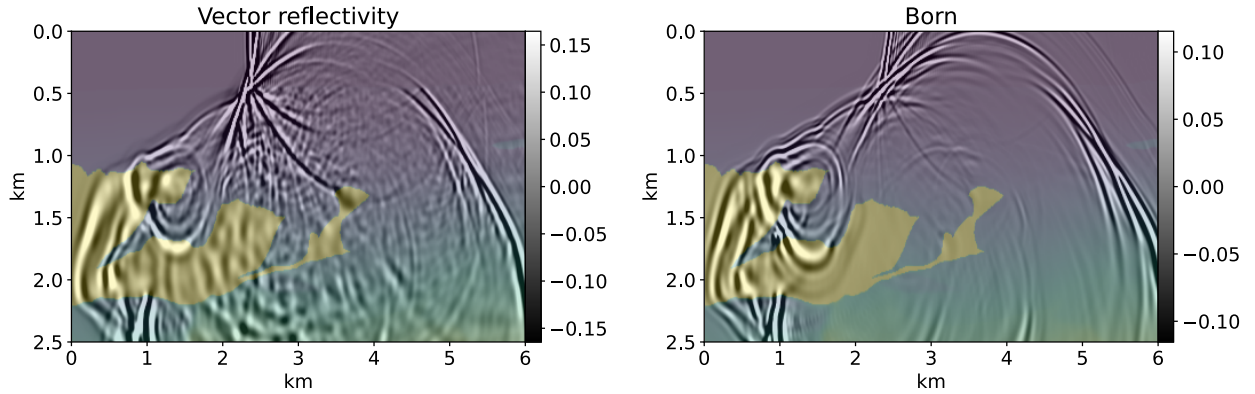


Figure 1.1: Forward propagation experiment on the BP synthetic model. The seismic source is located at $x = 6$ km and $z = 0.02$ km. Left: Acoustic wavefield at $t = 2.5$ seconds modeled with the full-wavefield vector reflectivity engine. Right: Born modeling for the same time snapshot. The velocity model containing a salt body is shown on the background for reference.

1.2.2 Adjoint of the full-wavefield vector reflectivity wave equation

Assuming a fixed background velocity model and known seismic sources, the simplest form of the inverse problem relating the vector reflectivity model, $\vec{r} \in \mathbb{R}^n$, with the observed data recorded at the surface, $d_{\text{obs}} \in \mathbb{R}^m$, can be expressed as the minimization of the objective function

$$\begin{aligned} \min_{\vec{r}} E(\vec{r}) &= \frac{1}{2} \sum_{s,g} \int_0^T dt \|F_{s,g}(\vec{r}, t) - d_{\text{obs},s,g}(t)\|_2^2, \\ &= \frac{1}{2} \sum_{s,g} \int_0^T dt \|d_{\text{res},s,g}(t)\|_2^2, \end{aligned} \quad (1.7)$$

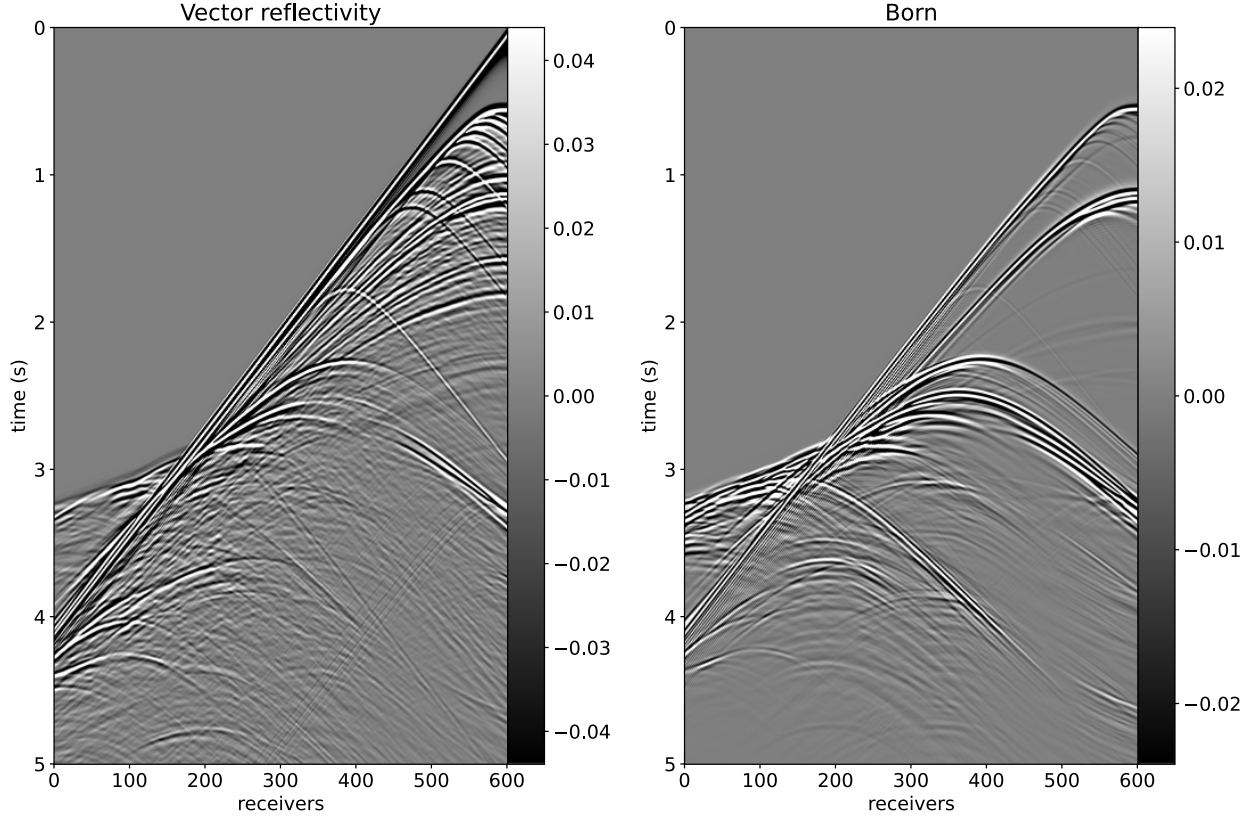


Figure 1.2: Left: Recorded traces for the full-wavefield acoustic vector reflectivity modeling on the BP synthetic model. Right: Born modeling for the same experiment. LSRTM uses Born approximation, focusing on first-order scattering in forward modeling, limiting its use to near-to-mid reflection angles and neglecting other wave modes such as diving and head waves.

where $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ describes the forward operator that propagates the wavefield in Equation 1.1 in the subsurface and registers its solution at each receiver position, s and g are the shot and receiver positions, T is the maximum recording time, and d_{res} represents the data residuals. To find a solution to the problem in equation 2, we use gradient-descent-based algorithms such as the L-BFGS method (Liu and Nocedal, 1989), which require only the value of the misfit function and its gradient to estimate quasi-Newton updates efficiently.

Different from previous work using the heuristic impedance kernel with the inverse scattering imaging condition to compute vector reflectivity updates (e.g., Yang et al., 2021), we directly derive the expression for the objective’s gradient with the Lagrange multiplier method (Plessix, 2006). This also entails calculating the adjoint vector reflectivity wave equation. As a result, we have a reliable

and accurate adjoint model for the solution of the optimization problem.

Based on the adjoint-state method, we neglect the damping term in Equation 1.5, and form the augmented Lagrangian functional as

$$L(\vec{r}) = E(\vec{r}) + \sum_s \int_0^T \int_{x,z \in \Omega} \Psi_s(t) \times \left[\frac{1}{v_p^2} \frac{\partial^2 u_s(t)}{\partial t^2} - \left(\nabla^2 u_s(t) + \frac{1}{v_p} (\nabla v_p \cdot \nabla u_s(t)) - 2(\vec{r} \cdot \nabla u_s(t)) \right) - f_s(t) \right] dx dz dt, \quad (1.8)$$

where the field Ψ is the Lagrange multiplier function, and Ω denotes the computational domain. According to the previously mentioned initial and boundary conditions that Equation 1.1 obeys, we obtain the following expression after integrating Equation 1.8 by parts:

$$L(\vec{r}) = E(\vec{r}) + \sum_s \int_0^T \int_{x,z \in \Omega} \left\{ u_s(t) \times \left[\frac{1}{v_p^2} \frac{\partial^2 \Psi_s(t)}{\partial t^2} - \left(\nabla^2 \Psi_s(t) + \frac{1}{v_p} (\nabla v_p \cdot \nabla \Psi_s(t)) - 2(\vec{r} \cdot \nabla \Psi_s(t)) \right) \right] - \Psi_s(t) f_s(t) \right\} dx dz dt. \quad (1.9)$$

Taking the adjoint-state equation $\partial L(\vec{r}) / \partial u_s = 0$ yields

$$\frac{1}{v_p^2} \frac{\partial^2 \Psi_s(t)}{\partial t^2} - \left(\nabla^2 \Psi_s(t) + \frac{1}{v_p} (\nabla v_p \cdot \nabla \Psi_s(t)) - 2(\vec{r} \cdot \nabla \Psi_s(t)) \right) = d_{\text{res}_{s,g}}(t), \quad (1.10)$$

with final conditions

$$\Psi_s(t = T) = 0, \quad \frac{\partial \Psi_s(t = T)}{\partial t} = 0. \quad (1.11)$$

If we define the adjoint wavefield λ_s in terms of the Lagrange multiplier function as

$$\lambda_s(T - t) = \Psi_s(t), \quad (1.12)$$

then Equation 1.10 becomes

$$\frac{1}{v_p^2} \frac{\partial^2 \lambda_s(T - t)}{\partial t^2} - \left(\nabla^2 \lambda_s(T - t) + \frac{1}{v_p} (\nabla v_p \cdot \nabla \lambda_s(T - t)) - 2(\vec{r} \cdot \nabla \lambda_s(T - t)) \right) = d_{\text{res}_{s,g}}(T - t), \quad (1.13)$$

with initial conditions

$$\lambda_s(t = 0) = 0, \quad \frac{\partial \lambda_s(t = 0)}{\partial t} = 0. \quad (1.14)$$

If we ignore the absorbing boundary conditions, the left-hand side in Equation 1.10 indicates that the wave equation for vector reflectivity is self-adjoint ($F = F^T$), enabling the computation of both the forward and adjoint wavefields using the same numerical solver (Fichtner et al., 2006). Thus, the corresponding expression for the update of the adjoint wavefield follows a similar form as the stencil presented in Equation 1.6, but solving a time-stepping loop backwards in time:

$$\begin{aligned} \lambda_s(t - dt) = & \frac{\frac{\partial}{\partial x} \lambda_s(t) \frac{\partial}{\partial x} v_p + \frac{\partial}{\partial y} \lambda_s(t) \frac{\partial}{\partial y} v_p}{v_p} - \frac{-\frac{2\lambda_s(t)}{dt^2} + \frac{\lambda_s(t+dt)}{dt^2}}{v_p^2} - 2r_x \frac{\partial}{\partial x} \lambda_s(t) - 2r_y \frac{\partial}{\partial y} \lambda_s(t) + \frac{\partial^2}{\partial x^2} \lambda_s(t) + \frac{\partial^2}{\partial y^2} \lambda_s(t) + \frac{\eta \lambda_s(t)}{dt} \\ & \frac{\frac{\eta}{dt} + \frac{1}{dt^2 v_p^2}}{dt} \\ & + dt v_p^2 d_{\text{res}_s}(t). \end{aligned} \quad (1.15)$$

In other words, the discretization of Equation 1.15 represents an upper triangular matrix that is computed by simulating wave propagation starting from the last time step using the data residuals as sources to update the previous adjoint wavefield, $\lambda_s(t - dt)$, rather than the next wavefield.

Finally, to obtain the expression for the misfit gradient, we set the derivative of the augmented Lagrangian functional in Equation 1.9 with respect to the vector reflectivity equal to zero, $\frac{\partial L(\vec{r})}{\partial \vec{r}} = 0$, which will produce

$$\nabla E(\vec{r}) = \begin{bmatrix} \nabla_{r_x} E \\ \nabla_{r_z} E \end{bmatrix} = -2 \sum_s \int_0^T dt \lambda_s(T - t) \nabla u_s(t) \quad (1.16)$$

$$= \sum_s \begin{bmatrix} J_{r_{x_s}}^T \\ J_{r_{z_s}}^T \end{bmatrix} d_{\text{res}_s}. \quad (1.17)$$

This approach is more efficient than calculating the misfit gradient via the Jacobian or Frechet derivative matrix, J , as it requires solving only two partial differential equations (PDEs) for each shot (as denoted in Equation 1.16) instead of solving a set of PDEs equal to the total number of parameters for each shot when employing the Jacobian. The gradient for each shot is therefore given by the product between two wavefields propagating in opposite time directions: the forward wavefield, u_s , simulated on a known background velocity, and the retro-propagated wavefield, λ_s , which is generated by injecting the data residuals from the position of the receivers as the adjoint seismic source, backwards in time. A naive computational approach to estimating the gradient entails storing one of these two wavefields in the entire domain and then accessing its time slices to perform the product with the other wavefield on-the-fly while stacking through the time dimension. However, more advanced wavefield reconstruction schemes can also be implemented to ease the

computational load, especially in 3D, where we deal with 4D wavefields (e.g., Symes, 2007; Nguyen and McMechan, 2015).

Once the misfit gradient has been computed, we can obtain vector reflectivity updates via iterative gradient-based minimization

$$\vec{r}^{k+1} = \vec{r}^k + \gamma^k p^k (\nabla E(\vec{r}^k)), \quad (1.18)$$

where the superscript k indicates the current iteration, γ is the step length computed via a line-search algorithm, and p is the gradient search direction, which we compute with a quasi-Newton algorithm.

1.2.3 Gradient verification

This section focuses on two specific tests: the dot-product test and the gradient test. As mentioned in the previous section, the solution to the inverse problem relies heavily on the accurate estimation of the misfit gradient, which depends on both the forward and the adjoint wavefields. Therefore, it is important first to confirm the exactness of the derivation of Equation 1.10 and its discretization as the true numerical adjoint wave equation of the full-wavefield vector reflectivity modeling engine. To do this, we employ the dot-product test (Claerbout, 2014) given by the inner product

$$\frac{\langle F\vec{a}, b \rangle - \langle \vec{a}, F^T b \rangle}{\langle F\vec{a}, b \rangle} = 0, \quad (1.19)$$

where $\vec{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ are any random vectors spanned in the model and data spaces, respectively. To perform the dot-product test, we use the experiment shown in Figure 1.3, where the long vector \vec{a} in column form is selected as a single flat reflector model, given by the concatenation of the two vector reflectivity x and z components (Figure 1.3c and 1.3d)

$$\vec{a} = \begin{bmatrix} a_x \\ a_z \end{bmatrix}. \quad (1.20)$$

Similarly, Figure 1.4 shows the model b represented as the data residuals between the "observed" and "synthetic" forward responses for a source located at the middle of the model's surface. We use a zero vector reflectivity to compute the latter. The horizontal and vertical components of the gradient using 400 receivers, also deployed at the model's surface, are presented in Figure 1.5. Our numerical results indicate that the discrete adjoint is accurate, passing the dot product test up to (single) numerical precision.

For comparison, we use the exact same configuration to compute the gradient response for the standard LSRTM problem using the adjoint of the Born modeling operator to migrate the data residuals. An important observation is the lack of low-wavenumber energy in both components of the misfit gradient related to the vector reflectivity engine (Figures 1.5a and 1.5b) compared to the image displaying the traditional LSRTM gradient (Figure 1.5c). Furthermore, in the vertical component of the vector reflectivity gradient, the "rabbit-ear" wavepaths associated with the source- and receiver-side wavefields have opposite phases. In seismic acquisitions with a high density of shots and receivers, these opposing wave-paths cancel out, rendering the low-frequency "rabbit-ear" effect negligible on the stacked vertical component of the vector reflectivity. Conversely, the "rabbit-ear" wave-paths have the same phase in the horizontal component of the vector reflectivity gradient, and the migration's "smile" exhibits varying positive and negative values across each side of the shot. Consequently, a multi-shot experiment results in a stacked horizontal component of vector reflectivity with smaller amplitudes and higher low-frequency artifacts. We show this in section 1.5 in a more realistic experiment.

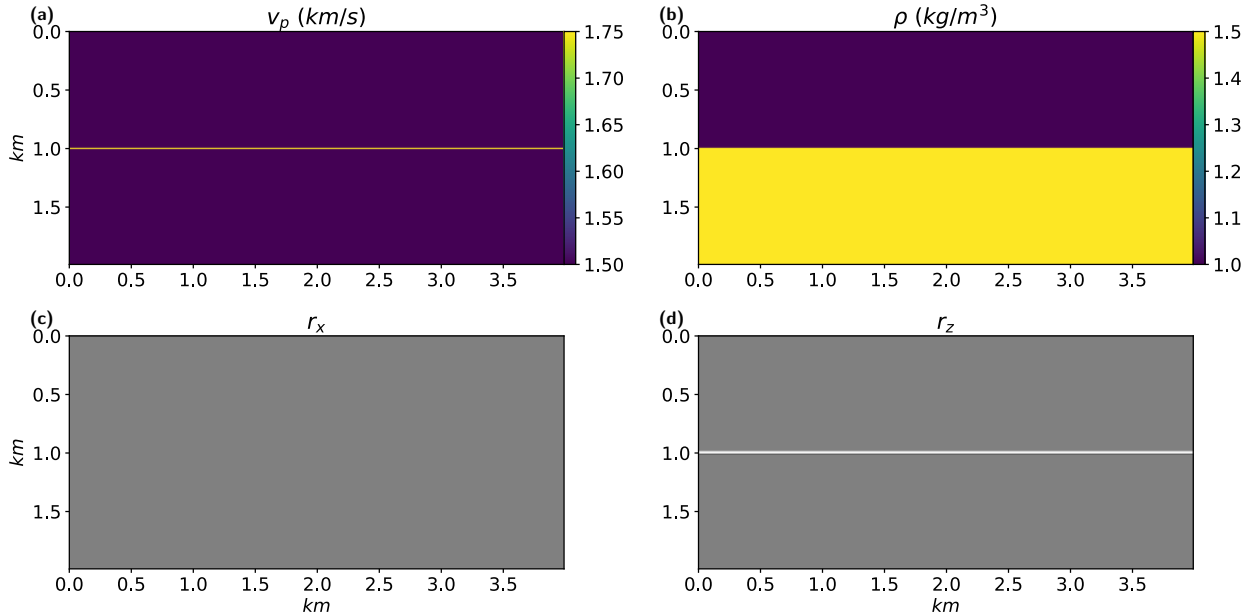


Figure 1.3: Flat reflector experiment to calculate the misfit gradient response of a single shot located at the middle of the model's surface and registered with 400 receivers, also at the surface of the model. (a) velocity model, (b) density model, (c) horizontal and (d) vertical components of the vector reflectivity model.

Secondly, we test for the accuracy of the misfit gradient. To do this straightforwardly, we calculate individual elements of the two gradient components using a finite difference central approximation.

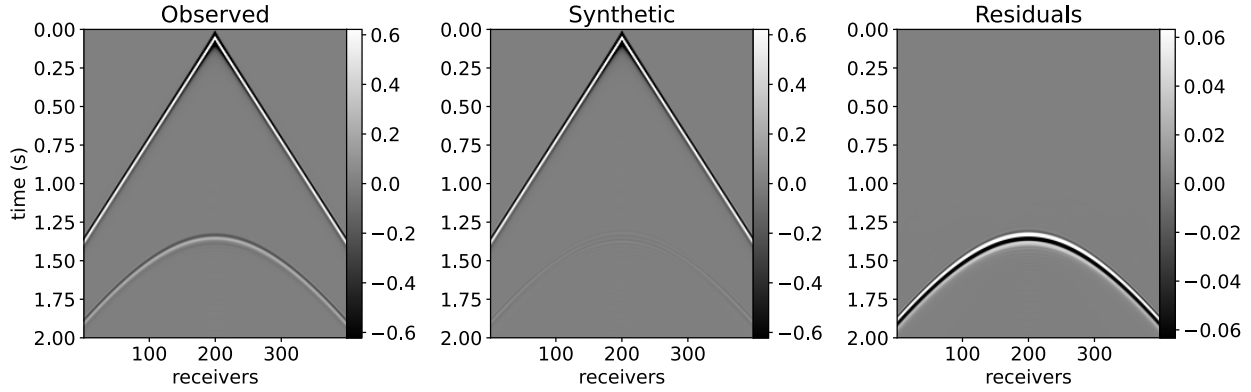


Figure 1.4: Left: "observed" data generated with a variable density acoustic wave equation using the models from Figures 1.3a and 1.3b. Middle: "synthetic" or "calculated" data generated with the full-wavefield vector reflectivity forward modeling using the true velocity model (Figure 1.3a), but a zero vector reflectivity model. Right: the data residuals.

Specifically, we approximate the derivative of the objective function with respect to the horizontal and vertical components of the vector reflectivity by varying the model parameters and observing the change in the objective function. We test this approximation for a few different points in the computational domain following

$$\frac{\partial E}{\partial r_{x_i}} \approx \frac{E(r_{x_i} + \Delta r_{x_i}, r_{z_i}) - E(r_{x_i} - \Delta r_{x_i}, r_{z_i})}{2\Delta r_{x_i}}, \quad (1.21)$$

$$\frac{\partial E}{\partial r_{z_i}} \approx \frac{E(r_{x_i}, r_{z_i} + \Delta r_{z_i}) - E(r_{x_i}, r_{z_i} - \Delta r_{z_i})}{2\Delta r_{z_i}}, \quad (1.22)$$

where x_i and z_i denote discrete points in the computational grid for each gradient component, and Δr_{x_i} and Δr_{z_i} are small perturbations. Varying the value of these model perturbations between 10^{-6} and 10^{-1} , we observe numerical errors that are below 10^{-2} when compared with the adjoint-state-derived gradient.

1.3 Regularization

Regularization strategies are essential for controlling stability and ambiguity in the solution, particularly when faced with insufficient seismic data. Furthermore, even with sufficient data, maintaining a regularization term is crucial to account for the noise inherent in real seismograms and avoid overfitting, which could result in inaccurate results. Therefore, we modify the original problem

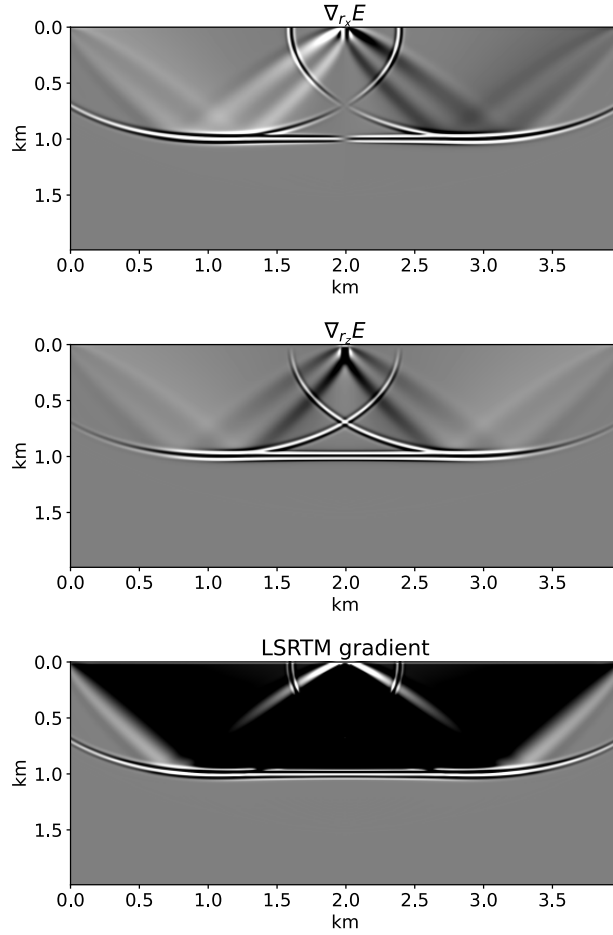


Figure 1.5: Single-shot gradient components for the flat reflector experiment. Top: horizontal component of the vector reflectivity gradient. Middle: vertical component of the vector reflectivity gradient. Bottom: standard LSRTM gradient.

described in Equation 1.7 by incorporating the HPF as the regularization term, which is differentiable and thus facilitates the use of straightforward optimization methods based on gradient descent techniques. This allows us to retrieve a sparse model of the vector reflectivity, characterized by a long-tail prior, which proves beneficial when the model components are expected to exhibit spiky characteristics. The regularized cost-functional for the sparsity promoting full-wavefield LSRTM takes the additive form

$$\min_{\vec{r}} J(\vec{r}) = E(\vec{r}) + \mu R_{\epsilon}(\vec{r}), \quad (1.23)$$

with $\mu > 0$ as the weight balancing the importance of sparseness of the vector reflectivity and data fitting, and

$$R_\epsilon(\vec{r}) = \sqrt{r_x^2 + r_z^2 + \epsilon^2} - \epsilon, \quad (1.24)$$

denotes the HPF term, where ϵ is the l_1/l_2 norm threshold.

The gradient search direction then becomes

$$\nabla J(\vec{r}) = \nabla E(\vec{r}) + \mu \nabla R_\epsilon(\vec{r}), \quad (1.25)$$

with

$$\nabla R_\epsilon(\vec{r}) = \begin{bmatrix} \frac{r_x}{\sqrt{r_x^2 + r_z^2 + \epsilon^2}} \\ \frac{r_z}{\sqrt{r_x^2 + r_z^2 + \epsilon^2}} \end{bmatrix}. \quad (1.26)$$

1.4 Deriving density models from the vector reflectivity

In the oil and gas industry, both density and velocity play an important role in seismic interpretation steps. Therefore, once we obtain the final vector reflectivity, a relative density model can be calculated as a byproduct by solving the normal equations

$$\operatorname{argmin}_m \left(\left\| \frac{1}{2} \begin{bmatrix} D_x \\ D_z \end{bmatrix} m - \begin{bmatrix} r_x \\ r_z \end{bmatrix} \right\|_2^2 + \alpha \|m\|_2^2 \right) \quad (1.27)$$

with a linear solver such as CGLS. In Equation 1.27, D_x and D_z are the first-order derivative operators in the horizontal and vertical dimensions, and α is a damping parameter. The final density model is given by $\rho = \frac{\exp(m)}{v_p}$. This extra step after the inversion extends the applicability of our sparsity-promoting results.

1.5 Application to the Marmousi2 model

We illustrate the performance of our sparse full-wavefield LSRTM workflow on the 2D Marmousi2 model. We generate the observed data from the true velocity and density models (Figures 1.6a and 1.6b) using the variable-density acoustic wave equation to alleviate the inverse crime. This generated data contains multiples, direct arrivals and diving waves. Our survey comprises 25 shots,

modeled with a 20 Hz Ricker wavelet, and 667 receivers evenly distributed at the surface. Figure 1.6c shows a smooth velocity model used as the background migration field for all the tests. Figures 1.7a and 1.7b show the horizontal and vertical components of the true vector reflectivity model. The initial guess for the inversions is set to a zero vector reflectivity, and both ϵ and μ were selected through trial and error and kept fixed throughout iterations.

We obtain the final vector reflectivity components for the seismic data inversion without (Figures 1.7c and 1.7d) and with (Figures 1.7e and 1.7f) HPF regularization after 30 iterations of L-BFGS. Inversion results with this dataset show that the iterative migrations with HPF regularization are sparse and present fewer migration artifacts, less contaminating sidelobes, and better signal content than their non-regularized counterpart. Moreover, all the least-squares results have correct amplitudes and behave as an approximate inversion since forward modelling reproduces the data approximately, as shown in Figure 1.8 for a shot positioned at the surface at $x=3.33$ km. No frequency domain transformations or spatial filtering techniques such as high-pass filters, Laplacian filters, or any other form of image enhancement were employed to produce these final results. Figure 1.9 compares the inverted results, with and without sparse regularization, and the true vertical component of the vector reflectivity for a trace in depth extracted at the middle of the model. We observe that in both cases, the inverted results match well with the corresponding defined reflectivity model. The inverted result with HPF regularization has more high-wavenumber components and better agrees with the true reflectivity.

For comparison, we also conducted 30 iterations of traditional LSRTM (as shown in Figure 1.6d), in which we applied a mute to the data to remove direct, diving and far-offset events. Despite this, the final image still exhibits substantial cross-talk attributable to parameter leakage. This is because standard LSRTM only inverts for velocity perturbations using an imprecise modeling engine. On the other hand, with the full-wavefield LSRTM, reflectivity changes caused by density variations are not erroneously mapped as velocity updates. We recognize that other studies have suggested multi-parameter LSRTM to deal with the cross-talk between model parameters, including elastic (Chen and Sacchi, 2017) and variable-density (Farshad and Chauris, 2021) reformulations. However, they still rely on and are restricted by the assumptions inherent to Born modeling, and require substantially greater computational resources.

Contrasting the results of full-wavefield LSRTM with traditional LSRTM, which heavily relies on an accurate velocity model due to the Born approximation, also reveals insights into how the vector reflectivity parameterization acts as a model-extension approach. It inherently makes the problem less over-determined by increasing the number of unknowns in the system. This comparison suggests

that the full-wavefield LSRTM method is likely more resilient to the challenges of inaccurate velocity models. This observation becomes increasingly evident in the deeper sections of our results. Full-wavefield LSRTM also significantly reduces backscattering noise compared to traditional LSRTM, enhancing seismic imaging.

Figure 1.10 shows the imaging results for the same experiment but changing the migration velocity model to a highly smoothed velocity, as shown in Figure 1.10a. The results indicate that full-wavefield LSRTM still exhibits fewer artifacts and provides more accurate images than traditional LSRTM. The sparse full-wavefield LSRTM further reduces the high-frequency artifacts.

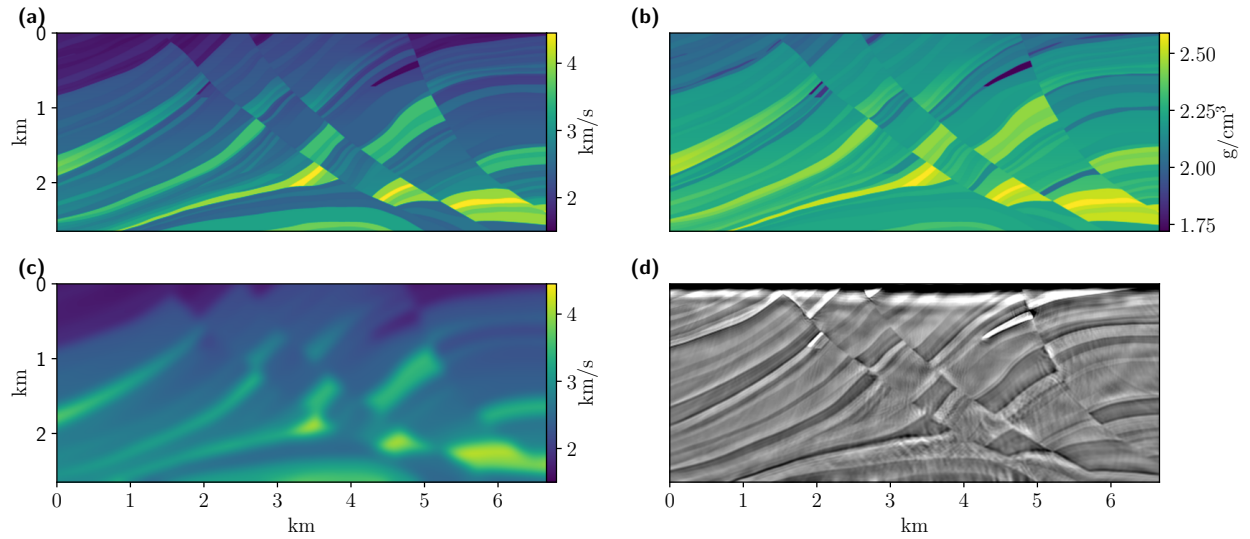


Figure 1.6: The 2D Marmousi2 model: (a) true velocity model, (b) true density model, (c) background migration velocity field, (d) traditional LSRTM inversion.

1.6 Conclusions

Classic LSRTM is solved in a linear inversion framework, which can be restrictive in handling data from complex geological settings where non-linearity might be significant. We successfully obtained enhanced vector reflectivity models by re-formulating LSRTM as a non-linear full-wavefield least-squares inversion in combination with sparse regularization. In our numerical experiment, applying this method to the 2D Marmousi2 model demonstrated its clear superiority. The parameter μ , which controls the regularization strength, requires some tuning. Our results indicate that sparse full-wavefield LSRTM reduced artifacts and sidelobes more effectively and delineated reflectors more

precisely than traditional methods, even in the presence of highly decimated data. Moreover, the ability of our approach to adapt to imprecise velocity models highlights its potential as a robust tool in seismic imaging, especially in complex geological settings.

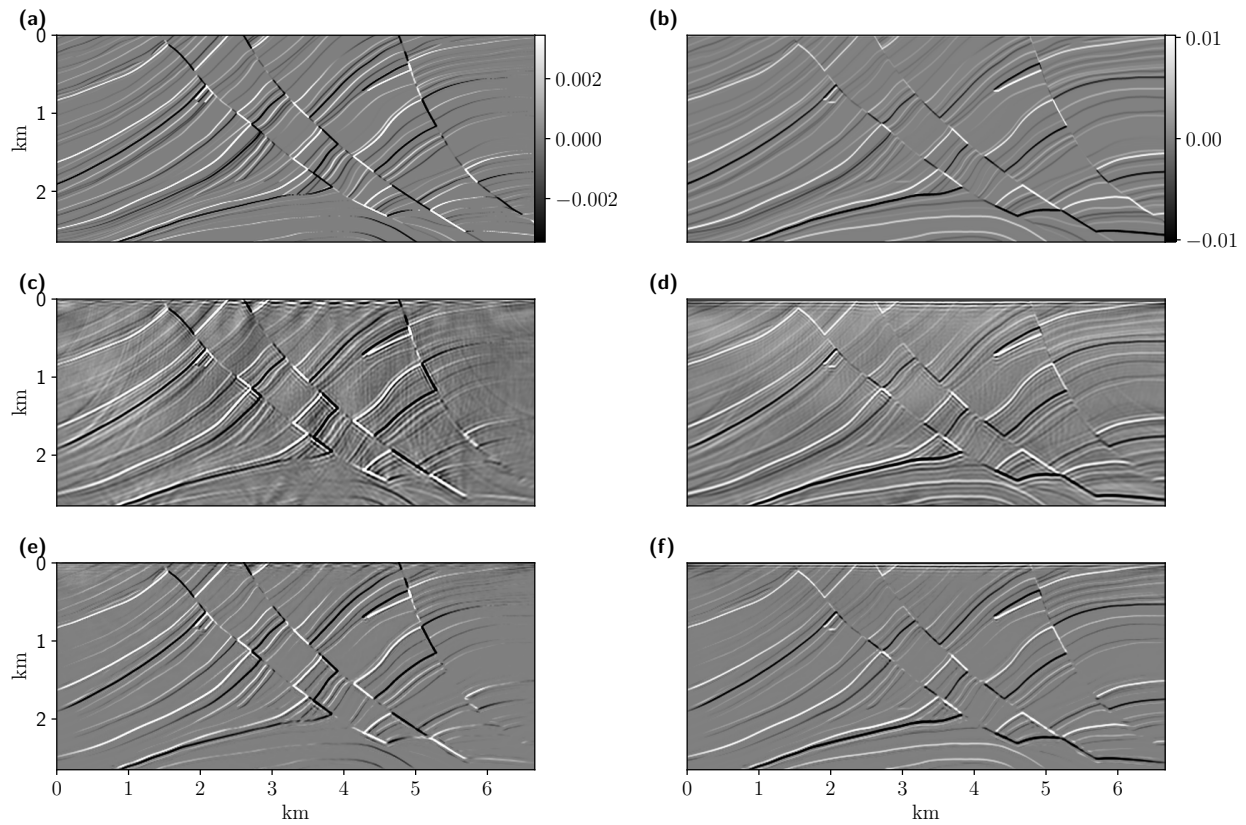


Figure 1.7: Full-wavefield LSRTM inversion of the 2D Marmousi2 model: (a), (c) and (e) show the horizontal components for the true vector reflectivity and the inversions without and with HPF regularization, respectively. Similarly, (b), (d) and (f) show the true and inverted vertical components. All the results are displayed using consistent amplitude scales, as demonstrated in Figures (a) and (b), thereby accurately representing true amplitude reflectivity imaging.

1.7 Acknowledgements

The authors thank the Signal Analysis and Imaging Group (SAIG) sponsors at the University of Alberta, for their continued support.

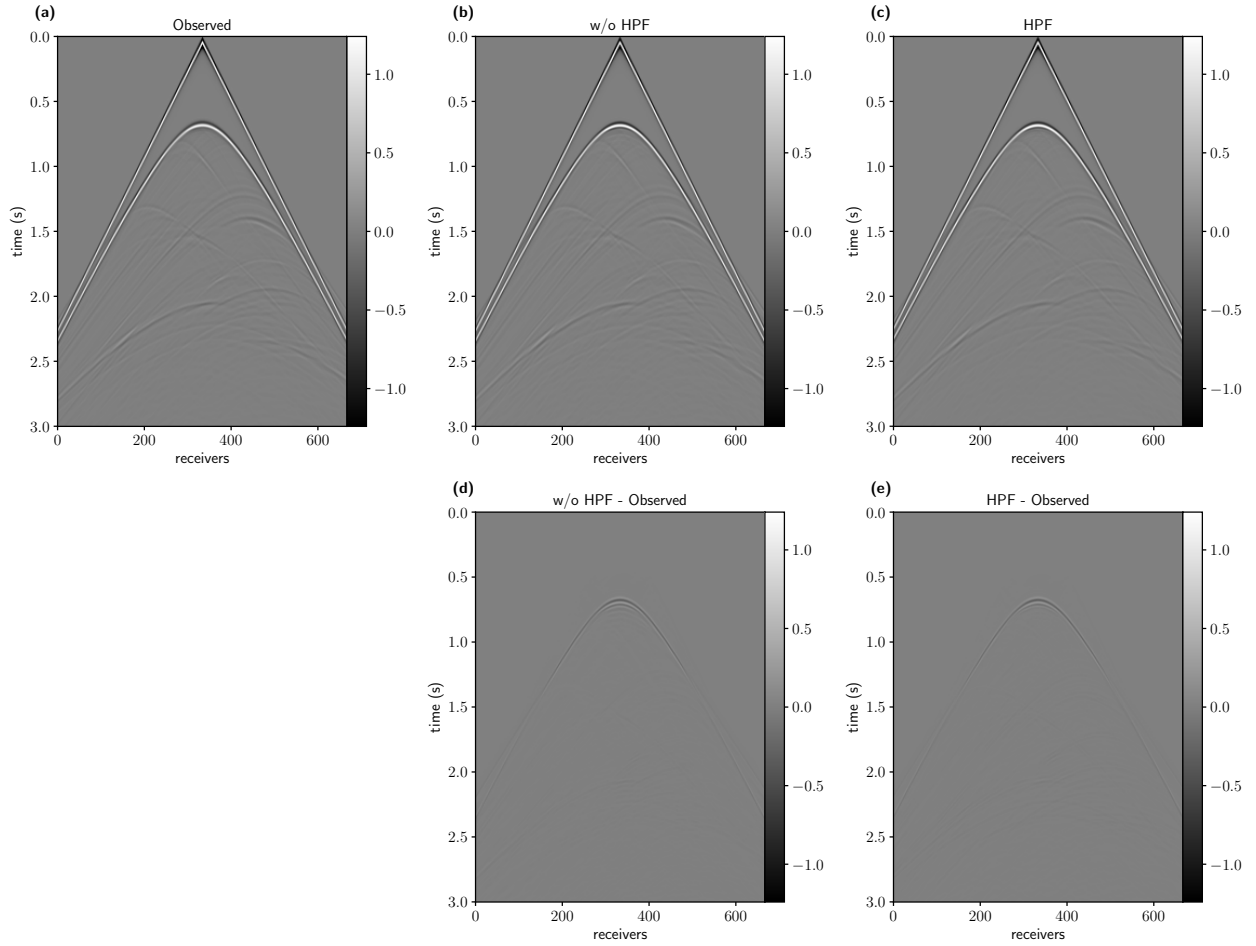


Figure 1.8: Comparison of traces in the data domain for the Marmousi2 inversion experiment for a shot positioned at the surface at $x=3.33$ km. (a) Observed data using the true velocity and density models. (b) and (c) show the forward modelling shot gather obtained with the vector reflectivity engine using the inverted results without and with sparse regularization, respectively. (d) and (e) show the data residuals for each result ((a) - (b) and (a) - (c)).

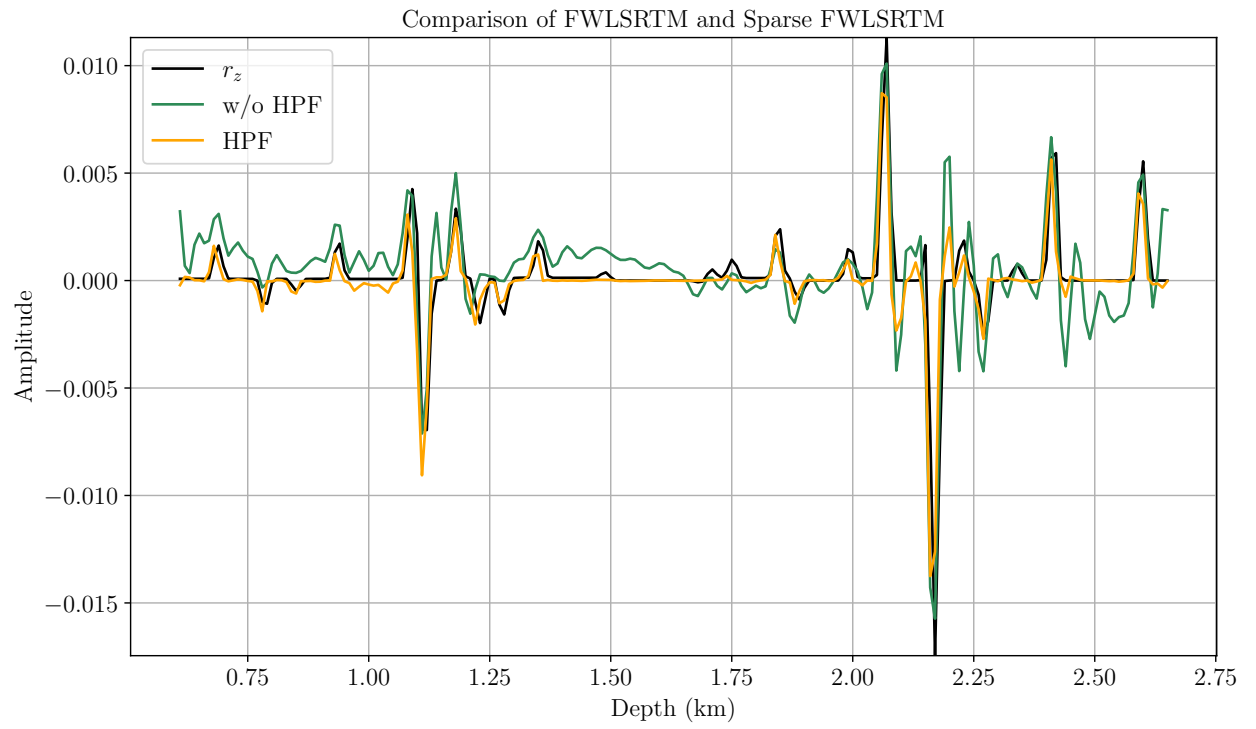


Figure 1.9: Comparison of traces from the vertical component of the inverted results, with (orange) and without (green) sparse regularization, and the true reflectivity model (black) at location $x = 3.33$ km.

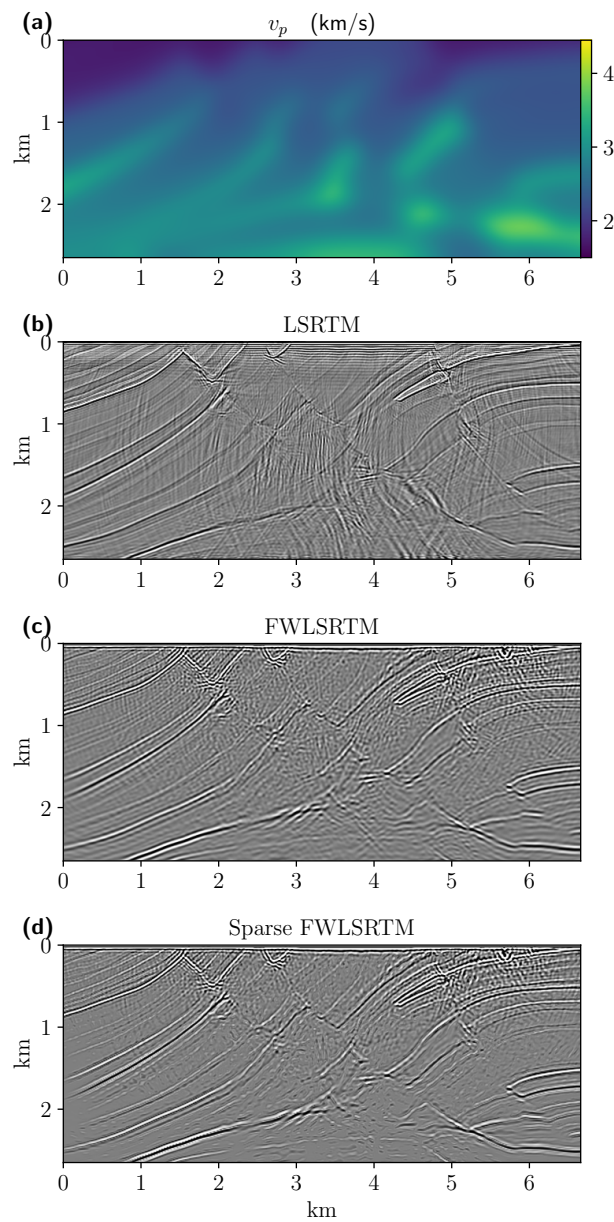


Figure 1.10: Full-wavefield LSRTM inversion of the 2D Marmousi2 model using a highly smoothed velocity model: (a) velocity model, (b) standard LSRTM result, (c) vertical component of full-wavefield LSRTM, (d) vertical component of full-wavefield LSRTM with sparse regularization. A Laplacian filter was applied to the images to remove low-frequency artifacts.

1.8 Appendix: misfit gradient with respect to the velocity model

This work only dealt with adjoint-state-derived vector reflectivity updates assuming a fixed background velocity model, which resembles classic LSRTM. However, following a procedure similar to the one presented in section 1.2.2, we can obtain the gradient with respect to the velocity model under a least-squares inversion framework based on the squared l_2 norm misfit using the Lagrange multiplier method. The expression for the misfit gradient with respect to the velocity based on the full-wavefield vector reflectivity wave equation is

$$\nabla E(v_p) = \sum_s \int_0^T dt \left[\frac{2}{v_p^3} \lambda_s(T-t) \frac{\partial^2 u_s(t)}{\partial t^2} - \frac{1}{v_p^2} \lambda_s(T-t) \nabla \cdot (v_p \nabla u_s(t)) \right]. \quad (1.28)$$

Combining the two adjoint-state-derived gradients allows us to invert velocity and vector reflectivity models simultaneously. This approach will be explored in future research.

References

- Chen, K., and M. D. Sacchi, 2017, Elastic least-squares reverse time migration via linearized elastic full-waveform inversion with pseudo-hessian preconditioning: *Geophysics*, **82**, S341–S358.
- Claerbout, J., 2014, Geophysical image estimation by example: Lulu. com.
- Davydenko, M., and D. Verschuur, 2017, Full-wavefield migration: using surface and internal multiples in imaging: *Geophysical Prospecting*, **65**, 7–21.
- Farshad, M., and H. Chauris, 2021, Accelerating the multi-parameter least-squares reverse time migration using an appropriate preconditioner: *Computational Geosciences*, **25**, 2071–2092.
- Fichtner, A., H.-P. Bunge, and H. Igel, 2006, The adjoint method in seismology: I. theory: *Physics of the Earth and Planetary Interiors*, **157**, 86–104.
- Fletcher, R., J. Hobro, and J. Rickett, 2023, Suppressing reflections with vector reflectivity acoustic modelling: 84th EAGE Annual Conference & Exhibition, European Association of Geoscientists & Engineers, 1–5.
- Korsmo, Ø., Y. Yang, N. Chemingui, A. Pankov, and A. Castiello, 2022, Nonlinear ls-rtm based on the seismic scale separation and full-wavefield multiparameter fwi: Second International Meeting for Applied Geoscience & Energy, Society of Exploration Geophysicists and American Association of Petroleum . . . , 977–981.
- Liu, D. C., and J. Nocedal, 1989, On the limited memory bfgs method for large scale optimization: *Mathematical programming*, **45**, 503–528.