

RICE UNIVERSITY

**Nonlinear Waveform Inversion with Surface-Oriented  
Extended Modeling**

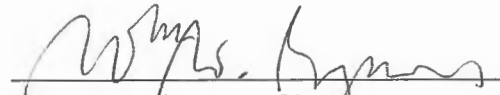
by

**Igor Terentyev**

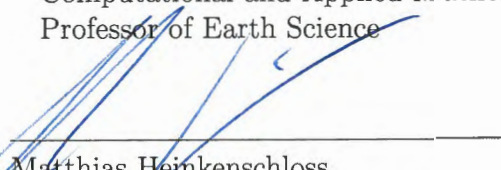
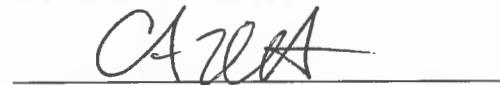
A THESIS SUBMITTED  
IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE

**Doctor of Philosophy**

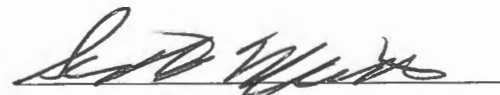
THESIS COMMITTEE:



William W. Symes, Chair,  
Noah G. Harding Professor of  
Computational and Applied Mathematics,  
Professor of Earth Science

  
Matthias Heinkenschloss,  
Noah G. Harding Professor of  
Computational and Applied Mathematics

Colin A. Zelt,  
Professor of Earth Science



Scott A. Morton,  
Ph.D., Hess Corporation

HOUSTON, TEXAS

MARCH, 2017

## **Abstract**

Nonlinear Waveform Inversion with Surface-Oriented Extended Modeling

by

Igor Terentyev

This thesis investigates surface-oriented model extension approach to non-linear full waveform inversion (FWI).

Conventional least-squares (LS) approach is capable of reconstructing highly detailed models of subsurface. Resolution requirements of the realistic problems dictate the use of local descent methods to solve the LS optimization problem. However, in the setting of any characteristic seismic problem, LS objective functional has numerous local extrema, rendering descent methods unsuitable when initial estimate is not kinematically accurate.

The aim of my work is to improve convexity properties of the objective functional. I use the extended modeling approach, and construct an extended optimization functional incorporating differential semblance condition. An important advantage of surface-oriented extensions is that they do not increase the computational complexity of the forward modeling. This approach blends FWI technique with migration velocity analysis (MVA) capability to recover long scale velocity model, producing optimization problems that combine global

convergence properties of the MVA with data fitting approach of FWI. In particular, it takes into account nonlinear physical effects, such as multiple reflections. I employ variable projection approach to solve the extended optimization problem. I validate the method on synthetic models for the constant density acoustics problem.

# Acknowledgements

First and foremost, I would like to express my deepest appreciation to my academic advisor, Prof. William Symes, for providing his guidance and continuous support throughout my research, for his enthusiasm, patience, and immense knowledge. It was a great pleasure working with him.

I am very grateful to my thesis committee, Prof. Matthias Heinkenschloss and Prof. Colin Zelt, for their interest in my work, and for the input and feedback they provided.

My sincere thanks goes to Dr. Scott Morton for useful discussions, encouragement, and for providing access to Hess computational resources.

I would like to thank all the members of the Rice Inversion Group (TRIP) for the stimulating discussions, and sponsors of TRIP for their support.

Last, but never least, my deepest gratitude to my family. My daughter, Polina, deserves special thanks for being such a nice girl. Thanks to my unbelievably supportive and understanding wife, Tania. I cannot express how grateful I am to my parents for their support and encouragement.

# Contents

<b>1</b>	<b>Introduction</b>	<b>9</b>
1.1	Conventional least-squares approach . . . . .	12
1.2	Extended modeling approach . . . . .	18
1.2.1	Full space optimization . . . . .	19
1.2.2	Earth model extensions and differential semblance . . . . .	20
1.2.3	Linearized problem with differential semblance optimization . . . . .	21
1.2.4	Time-lag axis extension . . . . .	22
<b>2</b>	<b>Theory</b>	<b>24</b>
2.1	Extended inverse problem . . . . .	24
2.2	Nested optimization . . . . .	27
2.3	Gradients . . . . .	29
2.4	Variable projection . . . . .	31
<b>3</b>	<b>Model problem</b>	<b>33</b>
3.1	Wave propagation model . . . . .	33
3.2	Plane wave model extension and differential semblance . . . . .	34
<b>4</b>	<b>Numerical experiments</b>	<b>37</b>
4.1	Numerical simulations . . . . .	37

4.2	Source wavelet . . . . .	39
4.3	Optimization . . . . .	43
4.4	Algorithm . . . . .	45
4.5	Dome model . . . . .	47
4.5.1	Conventional inversion experiment . . . . .	48
4.5.2	Absorbing boundary experiment . . . . .	50
4.5.3	Free surface experiment . . . . .	61
4.6	Marmousi model . . . . .	70
4.7	Discussion . . . . .	86
<b>5</b>	<b>Conclusion</b>	<b>87</b>
	<b>Appendices</b>	<b>91</b>
<b>A</b>	<b>Adjoint state method</b>	<b>92</b>
A.1	General derivation . . . . .	92
A.2	Constant density acoustics . . . . .	95
<b>B</b>	<b>Full bandwidth inversion</b>	<b>100</b>
<b>C</b>	<b>Envelope residual operator</b>	<b>103</b>
	<b>Bibliography</b>	<b>104</b>

# List of Figures

4.1	Source: Ricker wavelet with 8 Hz peak frequency passed through a low-cut (4-6 Hz) trapezoidal filter. Left: source wavelet. Right: source spectrum (normalized amplitude). . . . .	40
4.2	Complimentary source: Gaussian wavelet ( $\omega_c = \exp\{-\pi^2 f^2 t^2\}$ ) with $f = 8$ Hz passed through a high-cut (4-6 Hz) trapezoidal filter. Left: source wavelet. Right: source spectrum (normalized amplitude). . .	40
4.3	Full source: linear combination of main source and complimentary source with weight 0.5. Left: source wavelet. Right: source spectrum (normalized amplitude). . . . .	41
4.4	Dome acoustic velocity model. Horizontal layers: 1500, 3000, 3500, 4000, and 2600 m/s. Dome layers: 1800, 2100, and 2600 m/s. . . . .	47
4.5	Dome with absorbing surface example: conventional FWI; initial models (top), and recovered models (bottom). Left: inversion with kinematically correct initial model. Right: inversion with inaccurate long-scale structure of initial model. . . . .	49
4.6	Dome with absorbing surface example: shot records (direct arrivals removed). Left: plane wave with 0 degrees incidence angle. Right: plane wave with 20 degrees incidence angle. . . . .	50
4.7	Dome with absorbing surface example: averaged extended (left) and control (right) models. Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), and 32 (f) outer iterations of the VP method. . . . .	52
4.8	Dome with absorbing surface example: normalized data misfit convergence curve of steepest descent method for outer optimization. . . .	53
4.9	Dome with absorbing surface example: data residual after 32 outer iterations of the VP method. Left: shot record for plane wave with 0 degrees incidence angle. Right: data residual. . . . .	53

4.10	Dome with absorbing surface example: data residual after 32 outer iterations of the VP method. Left: shot record for plane wave with 20 degrees incidence angle. Right: data residual. . . . .	54
4.11	Dome with absorbing surface example: trace comparison of true data (red) and predicted data after 8 iterations (green) and after 32 iterations (blue). 0° incidence angle plane wave shot. Top: leftmost receiver. Bottom: central receiver. . . . .	54
4.12	Dome with absorbing surface example: trace comparison of true data (red) and predicted data after 8 iterations (green) and after 32 iterations (blue). 20° incidence angle plane wave shot. Top: leftmost receiver. Bottom: central receiver. . . . .	55
4.13	Dome acoustic velocity model: gather locations. Location (1) at $x = 2500$ m, (2) at $x = 2854$ m, (3) at $x = 2948$ m, and (4) at $x = 3774$ m. . . . .	56
4.14	Dome with absorbing surface example: model gathers at location (1). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), and 32 (f) outer iterations of the VP method. . . . .	57
4.15	Dome with absorbing surface example: model gathers at location (2). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), and 32 (f) outer iterations of the VP method. . . . .	58
4.16	Dome with absorbing surface example: model gathers at location (3). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), and 32 (f) outer iterations of the VP method. . . . .	59
4.17	Dome with absorbing surface example: model gathers at location (4). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), and 32 (f) outer iterations of the VP method. . . . .	60
4.18	Dome with free surface example: shot records (direct arrivals removed). Left: plane wave with 0 degrees incidence angle. Right: plane wave with 20 degrees incidence angle. . . . .	61
4.19	Dome with free surface example: averaged extended (left) and control (right) models. Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), 32 (f), and 44 (g) outer iterations of the VP method. . . . .	62
4.20	Dome with free surface example: normalized data misfit convergence curve of steepest descent method for outer optimization. . . . .	63



4.21	Dome with free surface example: data residual after 32 outer iterations of the VP method. Left: shot record for plane wave with 0 degrees incidence angle. Right: data residual. . . . .	64
4.22	Dome with free surface example: data residual after 32 outer iterations of the VP method. Left: shot record for plane wave with 20 degrees incidence angle. Right: data residual. . . . .	64
4.23	Dome with free surface example: model gathers at location (1). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), 32 (f), and 44 (g) outer iterations of the VP method. . . . .	65
4.24	Dome with free surface example: model gathers at location (2). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), 32 (f), and 44 (g) outer iterations of the VP method. . . . .	66
4.25	Dome with free surface example: model gathers at location (3). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), 32 (f), and 44 (g) outer iterations of the VP method. . . . .	67
4.26	Dome with free surface example: model gathers at location (4). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), 32 (f), and 44 (g) outer iterations of the VP method. . . . .	68
4.27	Marmousi acoustic velocity model. . . . .	70
4.28	Marmousi with absorbing surface example: shot records (direct arrivals removed). Left: plane wave with 0 degrees incidence angle. Center: plane wave with 30 degrees incidence angle. Right: plane wave with -30 degrees incidence angle. . . . .	71
4.29	Marmousi with absorbing surface example: inversion result. Top: averaged extended model after 70 outer iterations. Bottom: true model resampled to the computational grid. . . . .	73
4.30	Marmousi with absorbing surface example: data residual after 70 outer iterations of the VP method. Left: shot record for plane wave with 0 degrees incidence angle. Right: data residual. . . . .	74
4.31	Marmousi with absorbing surface example: data residual after 70 outer iterations of the VP method. Left: shot record for plane wave with 30 degrees incidence angle. Right: data residual. . . . .	74

4.32	Marmousi with absorbing surface example: trace comparison of true data (red) and predicted data after 70 iterations (blue). $0^\circ$ incidence angle plane wave shot. Top: leftmost receiver. Bottom: central receiver. . . . .	75
4.33	Marmousi with absorbing surface example: trace comparison of true data (red) and predicted data after 70 iterations (blue). $30^\circ$ incidence angle plane wave shot. Top: leftmost receiver. Bottom: central receiver. . . . .	75
4.34	Marmousi with absorbing surface example: inversion of the upper part. Top: top layer of recovered model in full-model inversion (same as upper part of recovered model in Figure 4.29). Middle: inversion result for the top layer and truncated data. Bottom: true model top layer. . . . .	77
4.35	Marmousi with absorbing surface example: gradient QC. Normalized dot product of control model gradient and direction from current model to smoothed averaged extended model. . . . .	78
4.36	Marmousi with absorbing surface example: inversion result. Top: model recovered by layer-stripped EFWI + FWI. Bottom: true model resampled to the computational grid. . . . .	80
4.37	Marmousi with absorbing surface example: final data residual. Left: shot record for plane wave with 0 degrees incidence angle. Right: data residual. . . . .	81
4.38	Marmousi with absorbing surface example: final data residual. Left: shot record for plane wave with 30 degrees incidence angle. Right: data residual. . . . .	81
4.39	Marmousi with absorbing surface example: final data residual. Left: shot record for plane wave with -30 degrees incidence angle. Right: data residual. . . . .	82
4.40	Marmousi with absorbing surface example: trace comparison of true data (red) and final predicted data (blue). $0^\circ$ incidence angle plane wave shot. Top: leftmost receiver. Center: central receiver. Bottom: rightmost receiver. . . . .	83

4.41	Marmousi with absorbing surface example: trace comparison of true data (red) and final predicted data (blue). $30^\circ$ incidence angle plane wave shot. Top: leftmost receiver. Center: central receiver. Bottom: rightmost receiver. . . . .	84
4.42	Marmousi with absorbing surface example: trace comparison of true data (red) and final predicted data (blue). $-30^\circ$ incidence angle plane wave shot. Top: leftmost receiver. Center: central receiver. Bottom: rightmost receiver. . . . .	85
B.1	Dome with absorbing surface example: full bandwidth source. Inversion result for horizontal plane wave source. . . . .	101
B.2	Dome with absorbing surface example: full bandwidth source. Inversion result for $10^\circ$ incidence angle plane wave source. . . . .	101
B.3	Dome with absorbing surface example: full bandwidth source. Inversion result for $20^\circ$ incidence angle plane wave source. . . . .	102

# Chapter 1

## Introduction

The problem of seismic inversion consists in recovering the Earth's physical properties or *models* from geophysical seismic data (seismograms). Typical physical properties of interest are acoustic impedance, compressional and shear velocities, density, etc. Geophysical seismic data is collected during seismic acquisition. An energy source (air-gun array, shot-hole dynamite, seismic vibrator) excites seismic waves propagating through the Earth. Response to the excitation (observed data) is recorded by receivers located near the surface or at the ocean bottom (OBN and OBC surveys) and, in some cases, in boreholes.

Model-based data fitting approach to seismic inversion has been extensively studied by the geoscience community over several decades (see, for example, Bamberg et al., 1979), and has recently coined a name full waveform inversion (FWI). Lines and Treitel, 1984; Tarantola, 2004; Virieux and Operto, 2009 provide excellent overview and history of the subject. Propagation of seismic waves through the Earth is sim-

ulated by forward modeling using partial differential equations (PDE) provided by idealized physical model, such as constant density acoustics, isotropic elasticity, etc. The goal of FWI is to construct a model that provides the best match of the observed and simulated data at various scales. Until last decade, FWI was mostly viewed computationally infeasible by the oil and gas industry due to very large computational cost of solving many forward modeling problems. However, recent advances in computer hardware and high-performance computing made the waveform inversion practical for high resolution industry-scale problems with more and more complex forward modeling physics (isotropic elasticity, visco-elasticity, anisotropy, and so on). In addition to computer advances, various techniques (phase encoding, shot grouping, to name a few) have been developed to substantially increase the computation efficiency of waveform inversion (Romero et al., 2000; Krebs et al., 2009; Ben-Hadj-Ali et al., 2011; Guitton and Diaz, 2012; van Leeuwen and Herrmann, 2013). Some examples of successful application of FWI technology can be found in Plessix et al. (2010); Vigh et al. (2010); Ratcliffe et al. (2011); Sirgue et al. (2011); Warner et al. (2013).

Conventional seismic migration techniques, such as reverse time and Kirchhoff migrations, require accurate earth velocity models to produce images (reflectivity models) with correct positions of the reflectors. A widely used and robust technique to recover earth velocity models is traveltime tomography. However, waveform inversion is preferable to tomography-based model building for multiple reasons. For example, resolution of the traveltime tomography is limited to the bandwidth of the Fresnel zone  $\sim \sqrt{\lambda L}$  (Williamson, 1991), where  $\lambda$  is the wavelength of the propagated wavefield, and  $L$  is the raypath length, while resolution of waveform inversion is on the order of half-wavelength  $\sim \lambda/2$  (Sirgue and Pratt, 2004). There are many circumstances,

such as complex earth structure (with large velocity changes), in which tomography may fail (Vigh et al., 2010). It has been shown (Vigh et al., 2010; Sirgue et al., 2011; Warner et al., 2013) that seismic migration produces improved images (better reflector positioning, resolution, event focusing) when provided with a velocity model updated by FWI.

Important aspects of waveform inversion are lack of information contained in the field seismic data (limited-frequency bandwidth, insufficient coverage), presence of noise in the signal, discrepancy between “true” and forward modeling physics, as well as inaccuracies in seismic source signature, etc. These aspects result in the well-known non-uniqueness of the inverse problem solution, i.e., it is impossible to match the observed data exactly and multiple models may provide best possible data fit. Therefore, the general solution of the inverse problem may involve estimation of an ensemble of models that fit observed data. In probabilistic inversion approach (e.g., Tarantola, 2004), each model is assigned a probability  $p$ , which depends on both the data misfit and an a-priori knowledge of the Earth in the area of interest. Solving the inverse problem consists of finding the probability distribution function  $p$ . Despite the attractiveness of the probabilistic inversion approach, it has no practical application because estimating the probability distribution  $p$  is not currently feasible for any realistic 3D problem.

Conventional deterministic FWI formulation is based on the least squares (LS) data misfit minimization. Due to the very large problem size, carrying out LS minimization is limited to iterative descent methods. Application of local descent methods to the LS minimization in FWI context has one fundamental difficulty. In the setting of any characteristic seismic problem, the objective functional has numerous local

extrema (multimodality property). See, for example, Gauthier et al. (1986); Santosa and Symes (1989); Versteeg and Grau (1991). This phenomenon is often referred to as cycle skipping (phase wrapping in frequency domain) or non-convexity of the LS objective functional.

The oscillatory nature of seismic data is a primary reason for multimodal behavior of the LS objective functional. However, it has been observed (Gauthier et al., 1986; Cao et al., 1990; Bunks et al., 1995; Plessix et al., 1998, to name a few) that if a model is kinematically accurate, the objective functional does not possess this undesirable local extrema property, and FWI process can successfully converge to a desired result. Thus, for successful FWI, the iterative process must start with a sufficiently accurate initial estimate of the velocity model.

To overcome the local minima obstacle and enlarge the basin of attraction of the global minima, one should construct FWI objective functionals that have improved convexity properties. Symes (2008) presented extended modeling concept, which aims to solve the local extrema problem by introducing additional axes to the model space. Various realizations of this concept (Biondi and Almomin, 2014; Gao et al., 2014; Liu et al., 2014) have been able to successfully overcome the cycle-skipping problem.

In this thesis, I employ the inversion strategy based on the extended modeling with surface-oriented extensions, differential semblance (DS) optimization (see, for example, Symes, 1991a; Symes and Carazzone, 1991), and low frequency extension of data source proposed by Sun and Symes (2012). Sun (2012) presents initial investigation of this approach for plane wave model extension and simple layered 2D examples.

## 1.1 Conventional least-squares approach

General formulation of least squares inversion is given in Tarantola and Valette (1982). In the seismic inversion setting, a modeling operator  $F : M \rightarrow D$  relates space  $M$  of all admissible models of the Earth<sup>1</sup> to the data space  $D$  equipped with a Hilbert norm  $\|\cdot\|$ . LS approach seeks to find a model  $m \in M$  that minimizes the data misfit between the observed data  $d_o \in D$  and simulated data  $F[m]$ :

$$\begin{aligned} \min_m J_{LS}[m], \\ J_{LS}[m] = \frac{1}{2} \|F[m] - d_o\|^2. \end{aligned} \tag{1.1}$$

Resolution requirements of the realistic applications lead to large dimension of model space  $M$ . Global search approaches (Scales et al., 1991; Sen and Stoffa, 1991; Stoffa and Sen, 1991) require too many objective functional evaluations to be feasible for multi-dimensional seismic problems. Therefore, use of local descent methods, such as steepest descent, nonlinear conjugate gradient, Gauss- or quasi-Newton (Nocedal and Wright, 1999), becomes mandatory to solve the optimization problem (1.1).

Application of the local descent methods requires computation of the objective functional gradient  $\nabla J_{LS}[m]$ . This computation can be carried out efficiently (without the Fréchet derivatives) via adjoint state method, which was introduced to the seismic community by Bamberger et al. (1982); Lailly (1983); Tarantola (1984). Adjoint state of a hyperbolic PDE was probably first described by Lions (1971) and used in the context of inverse problems by Chavent (1974). Plessix (2006); Fichtner (2010)

---

<sup>1</sup>Typically, admissible models constitute a subset  $M$  of a functional vector space constrained by mathematical bounds, geological and a-priori information, etc.



provide excellent overviews of the method; Symes (2007) gives an example of discrete adjoint derivation.

Some numerical examples (Kolb et al., 1986; Cao et al., 1990; Bunks et al., 1995) demonstrate successful application of this technique to layered medium problems. However, in realistic problems, least-squares objective turns out to contain numerous stationary points even for synthetic noise-free data, thus, rendering local descent methods unsuitable. Early examples of this “misbehavior” were presented by Gauthier et al. (1986); Kolb et al. (1986), and others. In order for local descent methods to converge to a global solution, accurate an low-wavenumber (i.e., kinematically accurate) initial velocity model is required (see, for example, Pica et al., 1990).

This multimodality is a fundamental mathematical property peculiar to the LS objective under characteristic seismic acquisition conditions, in particular, absence of low frequency signal. Santosa and Symes (1989); Symes (1990) give a comprehensive analysis of the behavior of LS objective in the context of horizontally layered medium. They show that it is extremely nonlinear and sensitive with respect to smooth perturbations of the velocity model, while being linear with respect to oscillatory perturbations. Changes in slow varying components of the velocity introduce change of phase and, thus, cycle skipping of high frequency components. As a consequence, objective functional  $J_{LS}[m]$  is extremely non-convex, gradient  $\nabla J_{LS}[m]$  is very unstable, and Jacobian  $DF[m]$  is very ill-conditioned. Jannane et al. (1989) use quasi-sinusoidal perturbations of the layered Earth model to determine model null space for the isotropic elastic problem. They observe travelttime sensitivity to long wavelengths of the velocity, impedance sensitivity to short wavelengths, and sensitivity gap with respect to middle wavelengths in case of small acquisition aperture.

Mora (1989); Santosa and Symes (1989) show that observable wavenumber components of the velocity model are limited from above by high frequency content of the data and from below by acquisition aperture provided that reflector structure is fairly complex, i.e., velocity model has dense (approximately every quarter of the longest data wavelength) and sufficiently strong reflectors. In case of horizontally layered medium, Symes (1991b) proves that sufficiently “rough” velocity can be recovered ( $F$  has continuous left inverse) from the band-limited reflection data. This is an important conclusion, since it implies that out-of-band velocity components can be observable for band-limited source.

Substantial amount of work is devoted to inverse seismic problem with full bandwidth (i.e., impulsive) source. Symes (1986a) considers 1D variable impedance impulse response problem and proves that impedance is continuously invertible. A number of numerical examples (Sacks and Santosa, 1987; Bunks et al., 1995) indicate solvability of the inversion problem with impulsive source in multiple dimensions. Bunks et al. (1995) present a multiscale technique based on frequency continuation that allows them to convexify LS objective functional and converge to the global minimum from kinematically incorrect initial guess.

While originally developed in the time domain, FWI can be also carried out in the frequency domain (see Pratt, 1999; Sirgue and Pratt, 2004, for example). The domain choice is based on particular techniques used in the inversion, capabilities of forward solvers, etc. Frequency domain allows, in particular, inversion at selected frequencies and, thus, is naturally suited for multiscale inversion strategies. Time windowing, on the other hand, requires time representation of the data. If all frequencies are

used simultaneously, frequency domain and time domain approaches are equivalent<sup>2</sup>. Pros and cons of time-domain versus frequency-domain FWI are discussed in Vigh and Starr (2008). Brenders et al. (2012) employ hybrid approach with time-domain forward modeling while carrying out inversion in frequency domain.

Shin and Cha (2008) show that FWI carried out in the Laplace domain is capable of recovering a long-scale velocity model starting from kinematically incorrect initial guess. Their approach is equivalent to matching zero frequency component of damped in time domain seismograms (Shin and Cha, 2009). Laplace domain FWI, however, is less sensitive to mid-scale components of the velocity model and fails to recover short-scale model details. The method is also very sensitive to the values of the damping parameters and maximum data offset (requires transmission data). Smooth velocity model produced by Laplace domain FWI can be used as an initial guess for subsequent frequency domain inversion. This approach coined the name Laplace-Fourier inversion (Shin and Cha, 2008; Ha and Shin, 2012; Kim et al., 2013).

The underlying assumption of the LS inversion is Gaussian distribution of data errors (Tarantola and Valette, 1982). Violation of this assumption, such as large outliers, may negatively affect the results of the LS inversion. A number of publications show that “robust” norms (e.g., Cauchy, Huber, Hybrid, or Student’s) are less sensitive to noise, while allowing same gradient-based optimization machinery as  $L_2$  norm. Crase et al. (1990); Guitton and Symes (2003); Bube and Nemeth (2007); Aravkin et al. (2011) provide examples of improved inversion robustness with such norms. van Leeuwen et al. (2013) note that in order for robust norms to be efficient the outliers have to be localized in some sense, and construct misfit functions in domains that

---

<sup>2</sup>Not considering efficiency and computational requirements of forward solvers.

localize noise.

Seismic inversion can be carried out with different misfit measures. The following several functionals emphasize fitting of phase and can be less sensitive to amplitude mismatch.

Logarithmic misfit (Shin and Min, 2006) is defined in frequency or Laplace domain:

$$E = \ln \frac{d_s(\omega)}{d_o(\omega)},$$

where  $d_s$  and  $d_o$  denote simulated and observed data, respectively, and  $\omega$  denotes complex frequency. It allows for separating amplitude and phase of the data misfit, or combining them in simultaneous phase and amplitude inversion:

$$J = \|\operatorname{Re} E\|^2 + \alpha \|\operatorname{Im} E\|^2,$$

where  $\alpha$  is a weighting parameter.

van Leeuwen and Mulder (2010) propose LS objective functional based on weighted cross-correlation of the observed and simulated data:

$$J = \|T(t)c(t)\|^2,$$

where  $c(t)$  is the cross-correlation:

$$c(t) = \sum_{\omega} \overline{d_s(\omega)} d_o(\omega) e^{2i\omega t},$$

and function  $T(t)$  penalizes energy at nonzero timelags, for example:

$$T(t) = \begin{cases} t & \text{if } |t| \leq t_0 \quad (\text{maximum considered lag}), \\ 0 & \text{otherwise.} \end{cases} \quad (1.2)$$

Warner and Guasch (2014) construct convolutional matching filter  $w(t)$  between observed and simulated data and minimize the difference between  $w(t)$  and zero-lag filter ( $\delta$ -function):

$$J = \|T(t)w(t)\|^2, \quad (1.3)$$

where  $T(t)$  penalizes filter coefficients away from zero-lag, similarly to expression (1.2).

Essentially same approach was proposed by Luo and Sava (2011). They construct deconvolution  $w(t)$  in Fourier domain:

$$w(t) = \sum_{\omega} \frac{\overline{d_o(\omega)}d_s(\omega)e^{-2i\omega t}}{\overline{d_o(\omega)}d_o(\omega) + \varepsilon^2},$$

and minimize analogous to (1.3) objective function.

Modifications of the inversion algorithm described in this section, as well as various other practical techniques (see Virieux and Operto, 2009), aim to increase efficiency and robustness, and alleviate the issues of the LS approach. While capable of enlarging the basin of attraction of local descent methods in certain situations, they cannot resolve the fundamental problem of the multimodality of the least-squares objective and subsequent convergence difficulties of the local descent approach.

In practice, FWI is typically used jointly with other approaches that provide accurate initial models. Commonly used methods to construct starting models for FWI

include reflection tomography, first arrival travelttime tomography (FATT), stereotomography, etc. Brenders and Pratt (2007b,a); Pratt and Brenders (2004) successfully employ joint FATT-FWI approach to both hydrocarbon-exploration scale and lithospheric-scale problems (blind tests).

## 1.2 Extended modeling approach

The overall seismic inversion problem can be cast as a problem of finding (1) state variables  $u$  (pressure, displacement, etc.), and (2) control variables  $m$  (Earth model, source, etc.) satisfying the following two constraints:

$$H[m, u] = 0, \tag{1.4}$$

$$G[u, d_o] = 0, \tag{1.5}$$

in which  $H$  is a wave operator, and  $G$  is a “misfit” operator<sup>3</sup>.

In the conventional FWI approach described in the previous section, forward equation (1.4) is satisfied at all stages (iterations) of the inversion process, and the goal of iterative optimization process is to find a model  $m$  that satisfies data fitting constraint (1.5).

Extended modeling approach recasts the problem by introducing extra degrees of freedom to the model<sup>4</sup>. The goal of model extension is to make data fitting constraint (1.5) easily satisfied at all stages of optimization (“data is always fit”), thus,

---

<sup>3</sup>Typically,  $G[u, d_o] = S[u] - d_o$ , where  $S$  is a sampling operator.

<sup>4</sup>Term model is used here in a broad sense: properties of the Earth, energy source, initial conditions, acquisition parameters, etc.

avoiding cycle skipping problem altogether. Extended models  $\bar{m}$  in general are not physical. An additional constraint is introduced to recover physically plausible models:

$$A[\bar{m}] = 0,$$

where *annihilator*  $A$  is an operator whose null space consists of non-extended models.

A number of methods utilizing extended modeling concept have shown to be able to converge to a global minimum from (kinematically) inaccurate initial models. In this section, I review a number of inversion approaches that employ extended modeling.

### 1.2.1 Full space optimization

van Leeuwen and Herrmann (2013); Peters et al. (2014) consider optimization in full space, i.e., over both control variables  $m$  and state variables  $u$ :

$$\begin{aligned} \min_{m,u} J[m, u], \\ J[m, u] = \|H[m, u]\|^2 + \alpha \|G[u, d_o]\|^2, \end{aligned} \tag{1.6}$$

where  $\alpha$  is a weighting parameter.

They recast problem (1.6) using variable projection method and numerically validate improved convexity properties of the objective functional. Decoupling variables  $m$  and  $u$  leads to one important advantage of the method: simplicity of the computation of the gradient and Hessian of  $J$ . Computational (especially memory) requirements present the main drawback of this approach currently rendering it inapplicable

to industrial problems.

It should be noted that full space optimization approach fits into the extended modeling framework with extended model  $\bar{m} = \{m, s_v\}$ , where  $s_v$  denotes distributed virtual sources. Forward equation (1.4) then becomes  $H[m, u] = s_v$ . Correspondingly, annihilator  $A[\bar{m}] = s_v$  penalizes non-zero virtual sources.

## 1.2.2 Earth model extensions and differential semblance

Earth model extensions are based on extending Earth model (such as velocity, reflectivity, etc.) by introducing additional dimension  $h$ :  $\bar{m} = \bar{m}(\mathbf{x}, h)$ . Symes (2008) discusses two kinds of Earth model extensions: surface-oriented extensions and subsurface-oriented extensions.

For surface-oriented extensions,  $h$  represents a surface acquisition parameter: shot coordinate, plane wave slowness, surface offset, etc. Thus, surface-oriented extension allows the model to depend on experiment. Data is binned accordingly:  $d = d(\cdot, h)$ . In many cases, inverse problem  $F[\bar{m}(\cdot, h)] = d(\cdot, h)$  is under-determined for any fixed  $h$ , and allows exact (up to a noise level) data fit solution. Obviously, extended model is physically plausible if and only if it does not depend on acquisition parameter  $h$  (“same Earth”). Symes (1986b) introduced *differential semblance* annihilator:

$$A[\bar{m}] = \frac{\partial \bar{m}}{\partial h}. \quad (1.7)$$

This is essentially the only kind of annihilator that leads to smooth objective function (Stolk and Symes, 2003).



In a subsurface-oriented extension approach, Earth model (coefficients of wave equation)  $m(\mathbf{x})$  is generalized by a positive self-adjoint operator. Non-extended models constitute a subspace, such as diagonal operators (see example in Symes, 2008, p. 775) with physical models focused at  $h = 0$ :  $\bar{m}(\mathbf{x}, h) = m(x)\delta(h)$  (Rickett and Sava, 2002).

Symes (2008) compares advantages and disadvantages of two kinds of Earth model extensions. Important advantage of surface extensions is that they do not increase the computational complexity of the forward modeling. In case of subsurface extensions, timestepping-based forward modeling may require substantially higher computational effort than non-extended forward modeling.

### 1.2.3 Linearized problem with differential semblance optimization

DS optimization approach was first applied to a linearized scattering (Born) problem with horizontally layered acoustic model and plane-wave data. Symes (1990, 1991a); Symes and Carazzone (1991) present detailed analysis of the problem. They show that under certain conditions (reflector-rich model, sufficient aperture): (1) DS objective functional is regular and, thus, much better suited for local descent methods than conventional LS functional; (2) solution of the DS optimization problem is stable with respect to noise.

Symes (1993); Symes and Kern (1994) further develop this technique for horizontally heterogeneous acoustic model and shot-gather data. The authors consider variable projection approach and show that it results in a smooth and convex objec-

tive functional, and optimization problem is stable. Recent results on this approach can be found in Huang and Symes (2015); Huang (2016).

Biondi and Almomin (2012); Almomin and Biondi (2012); Liu et al. (2014) use subsurface extensions of the reflectivity and demonstrate on numerical examples that this method is capable of recovering both high- and low-wavenumber components of the velocity model.

A very closely related group of methods, differential semblance velocity analysis (DSVA), is based on minimization of the DS of extended image (image volume). DSVA has been used with a number of extensions, such as surface offset, subsurface offset, common scattering angle, etc. Symes (2008) provides an excellent overview of the subject of migration velocity analysis (MVA), pros and cons of different types of model extensions, connection between MVA and LS inversion, and extensive list of references. For recent results, see Shen and Symes (2008, 2015); Fei and Williamson (2010); Li et al. (2014); Weibull and Arntsen (2014a,b); Maciel et al. (2011).

#### 1.2.4 Time-lag axis extension

Biondi and Almomin (2014) consider a multiple scattering problem:

$$\begin{aligned} \left( \frac{\partial^2}{\partial t^2} - v_0^2 \nabla_{\mathbf{x}}^2 \right) p_0 &= f, \\ \left( \frac{\partial^2}{\partial t^2} - v^2 \nabla_{\mathbf{x}}^2 \right) \delta p &= \delta v^2 \nabla_{\mathbf{x}}^2 p_0, \end{aligned} \tag{1.8}$$

where  $v_0$  and  $\delta v$  denote background velocity and velocity perturbation respectively, and  $v^2 = v_0^2 + \delta v^2$ . They introduce time-lag extension of velocity perturbation  $\delta \bar{v}(\mathbf{x}, t)$

and replace second equation in (1.8) by

$$\left( \frac{\partial^2}{\partial t^2} - \bar{v}^2|_{t=0} \nabla_{\mathbf{x}}^2 \right) \delta p = \delta \bar{v}^2 * \nabla_{\mathbf{x}}^2 p_0,$$

where  $\bar{v}^2 = v_0^2 + \delta \bar{v}^2$ .

The objective function is a compromise of data-fitting term and zero-lag focusing of the extended model:

$$J[\bar{v}] = \|\bar{F}[\bar{v}] - d_o\|^2 + \alpha \|t\delta\bar{v}\|^2. \quad (1.9)$$

Numerical examples show that this approach can converge to a global minimum from kinematically wrong initial models. However, convergence rates are too low to apply this method to industrial 3D problems.

# Chapter 2

## Theory

In this chapter, I provide an abstract formulation of the extended modeling algorithm. Application of this approach to constant density acoustics with plane wave source extension is described in Chapter 3.

### 2.1 Extended inverse problem

In an abstract inversion setting, a modeling operator  $F : M \rightarrow D$  relates a physical model space  $M$  to a data space  $D$ . The inversion problem of finding a model  $m \in M$  that predicts the observed data  $d_o \in D$ ,

$$F[m] \approx d_o, \tag{2.1}$$

is solved by least squares optimization:

$$\min_m J_{LS}[m], \quad (2.2)$$

$$J_{LS}[m] = \frac{1}{2} \|F[m] - d_o\|_D^2.$$

Extended modeling approach introduces extended model space  $\bar{M}$ . Physical model space  $M$  is an embedded linear subspace of  $\bar{M}$  via a linear extension map  $E : M \rightarrow \bar{M}$ . Image  $E[M]$  is sometimes called “physical models”. Spaces  $M$ ,  $\bar{M}$ , and  $D$  are Hilbert spaces with inner products denoted by  $\langle \cdot, \cdot \rangle_M$ ,  $\langle \cdot, \cdot \rangle_{\bar{M}}$ , and  $\langle \cdot, \cdot \rangle_D$  respectively.

Modeling operator  $F$  is extended by  $\bar{F} : \bar{M} \rightarrow D$  so that

$$\bar{F}[E[m]] = F[m]. \quad (2.3)$$

As discussed in Chapter 1.2, physically consistent extended models constitute a kernel of a linear operator (annihilator)  $A : \bar{M} \rightarrow N$ , where  $N$  denotes another Hilbert space:

$$A\bar{m} = 0 \iff \bar{m} = E[m]. \quad (2.4)$$

The extended inversion problem consists of finding an extended model<sup>1</sup>  $\bar{m} \in \bar{M}$  so that

$$\bar{F}[\bar{m}] \approx d_o, \quad (2.5)$$

$$A\bar{m} \approx 0.$$

If a model  $m \in M$  solves original problem (2.1), then from 2.3 and 2.4 follows that

---

<sup>1</sup>Extraction of the physical model  $m$  from extended solution  $\bar{m}$  of (2.5) is assumed to be a trivial operation.

extended model  $\bar{m} = E[m]$  solves problem (2.5). Conversely, for solution  $\bar{m}$  of problem (2.5) condition (2.4) implies that there exists model  $m$ , such that  $E[m] = \bar{m}$ , hence  $F[m] = \bar{F}[E[m]] \approx d_o$ .

Motivated by solvability of the least squares optimization problem for impulsive source<sup>2</sup>, Sun (2012) introduces control model  $m_c$  to problem (2.5):

$$\begin{aligned} \bar{F}[\bar{m}] + \bar{F}_c[\bar{m}] &\approx F_c[m_c] + d_o, \\ A\bar{m} &\approx 0, \end{aligned} \tag{2.6}$$

where  $\bar{F}_c$  denotes modeling operator with complimentary low-frequency source.

Denoting modeling operator with a full bandwidth source (full bandwidth operator) by

$$\bar{F}_f[\bar{m}] = \bar{F}[\bar{m}] + \bar{F}_c[\bar{m}], \tag{2.7}$$

I obtain the following extended optimization problem: find a pair  $\{m_c \in M, \bar{m} \in \bar{M}\}$ , such that

$$\begin{aligned} \bar{F}_f[\bar{m}] &\approx F_c[m_c] + d_o, \\ A\bar{m} &\approx 0. \end{aligned} \tag{2.8}$$

For a low-frequency complimentary source, control data generated by  $F_c[m_c]$  is not sensitive to short-scale perturbations of the control model  $m_c$ , hence solution of problem (2.8) is not unique. When solving (2.8) numerically, regularization needs to be applied to control model  $m_c$ . The goal of the regularization is to enforce smoothness of the control model. A suitable choice of the regularizer is the following operator

$$\Lambda = (I - L)^{-p}, \tag{2.9}$$

---

<sup>2</sup>See Appendix B for numerical validation of this fact for a simple 2D problem.

where  $p > 0$ , and anisotropic Laplace operator

$$L = \omega_z \frac{\partial^2}{\partial z^2} + \omega_x \frac{\partial^2}{\partial x^2} + \omega_y \frac{\partial^2}{\partial y^2}. \quad (2.10)$$

See (Symes and Kern, 1994) for analysis and (Huang, 2016) for implementation details.

## 2.2 Nested optimization

To solve problem (2.8), I introduce a compromise functional

$$J[m_c, \bar{m}; \alpha] = \frac{1 - \alpha}{2} \|\bar{F}_f[\bar{m}] - (F_c[m_c] + d_o)\|_D^2 + \frac{\alpha}{2} \|A\bar{m}\|_M^2, \quad (2.11)$$

where weight  $\alpha \in [0, 1]$  represents a compromise between data fit and physicality of the extended model.

Objective functional  $J$  will suffer the same multimodality problem as the conventional FWI functional  $J_{LS}$  (2.2), if optimized jointly over  $m_c$  and  $\bar{m}$ . For the extended Born waveform inversion, Huang (2016) attempted to optimize compromise functional simultaneously over the control (background) model and extended (reflectivity) model and observed no kinematic correction of the background model, thus keeping reflectors at initial incorrect positions. On the contrary, nested (in particular, variable projection) approach produced useful kinematic updates of the model.

Nested optimization approach introduces a reduced objective functional, which

essentially re-parametrizes extended model  $\bar{m}$  with the control model  $m_c$ :

$$J_r[m_c; \alpha] = \min_{\bar{m}} (J[m_c, \bar{m}; \alpha] + \mathcal{R}(\bar{m})), \quad (2.12)$$

where  $\mathcal{R}$  denotes possible regularization term. Solvability of the impulsive response problem suggests that in case of full bandwidth operator  $\bar{F}$ , reduced (also referred to as inner) problem (2.12) can be efficiently solved (can converge to a global minimizer) with local descent methods. Minimizer of the inner problem is denoted by

$$\bar{\mu} = \operatorname{argmin}_{\bar{m}} J[m_c, \bar{m}; \alpha]. \quad (2.13)$$

The outer optimization is performed over control model  $m_c$ :

$$\begin{aligned} \min_{m_c} \mathcal{J}[m_c], \\ \mathcal{J}[m_c] = J[m_c, \bar{\mu}; \tilde{\alpha}], \end{aligned} \quad (2.14)$$

with a possibly different weight  $\tilde{\alpha}$ .

Various approaches to choosing compromise weights  $\alpha$  and  $\tilde{\alpha}$  can be considered. Sun (2012) uses a compromise functional for the inner problem:  $0 < \alpha < 1$ , while only penalizing non-physicality of the extended model in the outer problem:  $\tilde{\alpha} = 1$ . Another possibility is a “pure data fitting” inner problem:  $\alpha = 0$ . Thus, for each extended model  $\bar{m}(\cdot, h)$ , a decoupled conventional FWI problem is solved. The outer optimization functional penalizes non-physicality of the extended model:  $\tilde{\alpha} = 1$ .

In my work, I consider a variable projection (VP) method (Golub and Pereyra, 2003) described in Section 2.4 of this chapter.



## 2.3 Gradients

Gradient for the inner problem (2.12) is

$$\nabla_{\bar{m}} J[m_c, \bar{m}; \alpha] = (1 - \alpha) D\bar{F}_f[\bar{m}]^* (\bar{F}_f[\bar{m}] - (F_c[m_c] + d_o)) + \alpha A^* A \bar{m}. \quad (2.15)$$

Necessary condition for a stationary point  $\bar{\mu}$  is

$$\nabla_{\bar{m}} J[m_c, \bar{\mu}; \alpha] = 0, \quad (2.16)$$

i.e.,

$$(1 - \alpha) D\bar{F}_f[\bar{\mu}]^* (\bar{F}_f[\bar{\mu}] - (F_c[m_c] + d_o)) + \alpha A^* A \bar{\mu} = 0. \quad (2.17)$$

To derive a gradient for the outer problem (2.14), consider perturbation  $m_c + \delta m_c$  of the control model and corresponding perturbed solution  $\bar{\mu} + \delta \bar{\mu}$  of the inner problem (2.12). Necessary condition (2.16) for perturbed solution gives

$$\begin{aligned} (1 - \alpha) D\bar{F}_f[\bar{\mu} + \delta \bar{\mu}]^* (\bar{F}_f[\bar{\mu} + \delta \bar{\mu}] - (F_c[m_c + \delta m_c] + d_o)) \\ + \alpha A^* A (\bar{\mu} + \delta \bar{\mu}) = 0. \end{aligned} \quad (2.18)$$

Expansion of this equation to the first order yields:

$$\begin{aligned} (1 - \alpha) (D\bar{F}_f[\bar{\mu}] + D^2\bar{F}_f[\bar{\mu}] \delta \bar{\mu})^* \\ (\bar{F}_f[\bar{\mu}] + D\bar{F}_f[\bar{\mu}] \delta \bar{\mu} - (F_c[m_c] + DF_c[m_c] \delta m_c + d_o)) \\ + \alpha A^* A (\bar{\mu} + \delta \bar{\mu}) = 0. \end{aligned} \quad (2.19)$$

After regrouping terms, we have

$$\begin{aligned}
& (1 - \alpha)D\bar{F}_f[\bar{\mu}]^* (\bar{F}_f[\bar{\mu}] - (F_c[m_c] + d_o)) + \alpha A^* A \bar{\mu} \\
& + (1 - \alpha)\{D\bar{F}_f[\bar{\mu}]^* D\bar{F}_f[\bar{\mu}]\delta\bar{\mu} + (D^2\bar{F}_f[\bar{\mu}]\delta\bar{\mu})^* (\bar{F}_f[\bar{\mu}] - (F_c[m_c] + d_o))\} \\
& \quad + \alpha A^* A \delta\bar{\mu} \\
& = (1 - \alpha)\{D\bar{F}_f[\bar{\mu}]^* DF_c[m_c]\delta m_c\} \\
& + (1 - \alpha)\{(D^2\bar{F}_f[\bar{\mu}]\delta\bar{\mu})^* (DF_c[m_c]\delta m_c - D\bar{F}_f[\bar{\mu}]\delta\bar{\mu})\}.
\end{aligned} \tag{2.20}$$

First line of the above equation is zero due to (2.17). Dropping higher-order terms on the right-hand side, we obtain

$$Q[\bar{\mu}, \bar{F}_f[\bar{\mu}] - (F_c[m_c] + d_o)] \delta\bar{\mu} = D\bar{F}_f[\bar{\mu}]^* DF_c[m_c]\delta m_c, \tag{2.21}$$

where self-adjoint linear (of argument  $\delta\bar{m}$ ) operator  $Q$  is defined by

$$Q[\bar{\mu}, d] \delta\bar{\mu} = \{N[\bar{\mu}] + W[\bar{\mu}, d] + \frac{\alpha}{1 - \alpha} A^* A\} \delta\bar{\mu}, \tag{2.22}$$

normal operator  $N$  is given by

$$N[\bar{\mu}] \delta\bar{\mu} = D\bar{F}_f[\bar{\mu}]^* D\bar{F}_f[\bar{\mu}] \delta\bar{\mu}, \tag{2.23}$$

and a so-called tomographic operator  $W$  (Biondi and Sava, 2004; Biondi and Al-momin, 2012) is defined by

$$W[\bar{\mu}, d] \delta\bar{\mu} = (D^2\bar{F}_f[\bar{\mu}] \delta\bar{\mu})^* d. \tag{2.24}$$

From equation (2.21) follows

$$D\bar{\mu} = Q[\bar{\mu}, \bar{F}_f[\bar{\mu}] - (F_c[m_c] + d_o)]^{-1} D\bar{F}_f[\bar{\mu}]^* DF_c[m_c]. \quad (2.25)$$

Since operator  $Q$  is self-adjoint, gradient for the outer problem (2.14) is

$$\begin{aligned} \nabla \mathcal{J}[m_c] = & -(1 - \tilde{\alpha}) DF_c[m_c]^* (\bar{F}_f[\bar{\mu}] - (F_c[m_c] + d_o)) \\ & + DF_c[m_c]^* D\bar{F}_f[\bar{\mu}] Q[\bar{\mu}, \bar{F}_f[\bar{\mu}] - (F_c[m_c] + d_o)]^{-1} \nabla_{\bar{m}} J[m_c, \bar{\mu}; \tilde{\alpha}]. \end{aligned} \quad (2.26)$$

## 2.4 Variable projection

The VP method differs from other nested optimization methods by using same penalty weight  $\tilde{\alpha} = \alpha$  for both the inner (2.12) and outer (2.14) optimization problems. Huang (2016); Xu et al. (2012); van Leeuwen and Mulder (2009); Symes and Kern (1994) utilize VP optimization for linearized (Born) and extended Born waveform inversions.

To reduce notation, I drop  $(1 - \alpha)$  weight from the data-fitting term of the extended objective functional:

$$J[m_c, \bar{m}] = \frac{1}{2} \|\bar{F}_f[\bar{m}] - (F_c[m_c] + d_o)\|_D^2 + \frac{\alpha}{2} \|A\bar{m}\|_M^2. \quad (2.27)$$

Gradient (2.15) for the inner problem then becomes

$$\nabla_{\bar{m}} J[m_c, \bar{m}] = D\bar{F}_f[\bar{m}]^* (\bar{F}_f[\bar{m}] - F_c[m_c] + d_o) + \alpha A^* A \bar{m}. \quad (2.28)$$

One of the advantages of the VP method is significant simplification of the outer

gradient. If the necessary condition (2.17) is satisfied, the second term in the outer gradient expression (2.26) vanishes:

$$\nabla \mathcal{J}[m_c] = -DF_c[m_c]^* (\bar{F}_f[\bar{\mu}] - (F_c[m_c] + d_o)). \quad (2.29)$$

However, in order for this method to work, the inner problem has to be solved very accurately. Since the second term in the outer gradient (2.26) contains an unbounded operator, deviation of the inner gradient (2.28) from zero at the solution  $\bar{\mu}$  of the inner problem may result in large error in the outer gradient. For the extended Born waveform inversion, Symes and Kern (1994) indicate that the accuracy of the solution of the inner problem is critical for the accuracy in the background model gradient computation. Huang (2016) quantifies their suggestion.

# Chapter 3

## Model problem

This chapter presents the wave propagation PDE, model parametrization, and model extension type, considered in this study.

### 3.1 Wave propagation model

I adopt the simplest wave propagation model, *constant density acoustics*, based on the second-order wave equation

$$\begin{aligned} \left( \frac{1}{c^2(\mathbf{x})} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) u(\mathbf{x}, t) &= f(\mathbf{x}, t), \\ u(\mathbf{x}, t) &\equiv 0, \quad t < 0 \quad (\text{casual field}), \end{aligned} \tag{3.1}$$

where position  $\mathbf{x} \in \Omega$  for a domain of interest  $\Omega$ , and time  $t \in [0, T]$  for some simulation time  $T$ . Here

- $c(\mathbf{x})$  is an acoustic velocity,
- $u(\mathbf{x}, t)$  is an acoustic potential,
- $f(\mathbf{x}, t)$  is a causal source representing a body force.

I apply absorbing boundary condition on three sides of the domain  $\Omega$ , and absorbing or zero Dirichlet (free surface) boundary condition on the remaining top side.

Acoustic velocity is restricted to an a-priory known range of velocities  $c \in (c_{min}, c_{max})$ . In order to satisfy this restriction, I introduce a model parameter  $m(\mathbf{x})$  defined by

$$m = \tanh^{-1} \left\{ \left( c^\gamma - \frac{c_{max}^\gamma + c_{min}^\gamma}{2} \right) \frac{2}{c_{max}^\gamma - c_{min}^\gamma} \right\}, \quad (3.2)$$

where  $\gamma$  is some non-zero power. Equation (3.2) monotonically maps range of velocities  $(c_{min}, c_{max})$  to  $(-\infty, \infty)$ , thus “soft-clipping” velocities to a desired range.

Inversion problem is solved in terms of the model  $m$  defined by equation (3.2). To avoid notation clutter, I will keep using  $c(\mathbf{x})$  in the equation (3.1) and its derivatives, implying  $c(\mathbf{x}) = c[m(\mathbf{x})]$ .

## 3.2 Plane wave model extension and differential semblance

In this thesis, I consider surface-oriented extension with respect to plane wave sources. Extended models are parameterized by an acquisition parameter  $h$ , which encodes plane wave source. For two-dimensional problems,  $h$  is a one-dimensional parameter,

such as plane wave incidence angle or slowness.

Data  $d(t, \mathbf{x}_r; h) \in D$  is a collection of traces at discrete receiver positions  $\mathbf{x}_r \in \Sigma_r \subset \Omega$  for a range of the acquisition parameter  $h$ . Modeling operator  $F : M \rightarrow D$  is defined by

$$F[m](\mathbf{x}_r, t; h) = u(\mathbf{x}_r, t; h), \quad (3.3)$$

where  $u(\mathbf{x}, t; h)$  solves equation (3.1) with a plane wave source  $f(\dots; h)$ .

Extended model space  $\bar{M} = \{\bar{m}(\mathbf{x}; h)\}$  is a set of models of position  $\mathbf{x}$  and acquisition parameter  $h$ .

Modeling operator  $F$  trivially extends to  $\bar{F} : \bar{M} \rightarrow D$  by

$$\bar{F}[\bar{m}](\mathbf{x}_r, t; h) = F[\bar{m}(\cdot; h)](\mathbf{x}_r, t; h), \quad (3.4)$$

leading to a wave equation with a separate model for each plane wave source:

$$\left( \frac{1}{c^2(\mathbf{x}; h)} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) u(\mathbf{x}, t; h) = f(\mathbf{x}, t; h), \quad (3.5)$$

Such model extension allows to fit observed data for plane wave shot  $f(\dots; h)$  with data predicted by a separate model  $\bar{m}(\mathbf{x}; h)$ .

As described in Section 1.2.2, differential semblance operator is the only appropriate choice of annihilator for surface-oriented extension considered in this thesis. Annihilator  $A$  from equation (2.11) is given by

$$A\bar{m} = \frac{\partial \bar{m}}{\partial h}. \quad (3.6)$$

Operator  $A^*A$  used in gradient computation (2.15) takes the following form:

$$A^*A\bar{m} = -\frac{\partial^2 \bar{m}}{\partial h^2} + \frac{\partial \bar{m}}{\partial h} \Big|_{h_0}^{h_1}, \quad (3.7)$$

where  $h \in [h_0, h_1]$ .



# Chapter 4

## Numerical experiments

In this chapter, I present numerical examples with two synthetic two-dimensional models: Dome model and Marmousi (Bourgeois et al., 1991) model. For the Dome model, I consider two boundary condition cases: absorbing and free surface.

### 4.1 Numerical simulations

I use conventional regular grid finite difference (FD) scheme to solve equation (3.1) numerically.

Computational rectangular domain  $\Omega$  is discretized with a uniform grid  $\mathbf{x}_{ij} = (i\Delta x, j\Delta x)$ . Spatial sampling  $\Delta x$  is chosen based on effective source bandwidth and velocity range  $(c_{min}, c_{max})$  from equation (3.2) in order to have insignificant numerical dispersion (Cohen, 2001; Gustafsson, 2008). For numerical experiments with

synthetic models, I subsample true models to a computational sub-grid via mass lumping (Symes and Terentyev, 2009).

Time axis discretization is  $t^n = n\Delta t$ . Time sampling  $\Delta t$  is fixed during entire inversion process, and is chosen based on maximum velocity  $c_{max}$  from equation (3.2) in order to satisfy CFL stability conditions (Cohen, 2001).

I use second-order in time and eighth-order in space (2-8) FD scheme given by

$$u^{n+1} = 2u^n - u^{n-1} + \Delta t^2 c^2 [m] L u^n + \Delta t^2 f^n, \quad (4.1)$$

where  $u^n = u(x_{ij}, t^n)$ ,  $L$  is an eighth-order discrete Laplacian, and  $f^n = f(x_{ij}, t^n)$ .

Absorbing boundary conditions are implemented using convolutional perfectly matched layer (PML) approach (Komatitsch and Martin, 2007) adopted to a second order acoustic wave equation. The computational domain is extended in order to include the absorbing layers.

Discretized Born and adjoint approximations for the wave propagation and PML equations are generated using automatic differentiation (AD) software **Tapenade** (Hascoet and Pascual, 2013). To validate and simplify code produced by AD tools, it is useful to have continuous Born operators  $DF$ ,  $D^2F$  and adjoint operators  $DF^*$ ,  $D^2F^*$ , which are derived in Appendix A. Discrete adjoint  $DF^*$  used in the inner gradient computation of the VP method requires the following FD scheme propagated backwards in time:

$$v^{n-1} = 2v^n - v^{n+1} + \Delta t^2 L(c^2 [m] v^n) + r^{n-1}, \quad (4.2)$$

where  $v^n = v(x_{ij}, t^n)$  is back-propagated wavefield,  $r^n = r(x_{ij}, t^n)$  denotes residuals.

Annihilator (3.6) is discretized by a 1-st order one-sided finite difference scheme. Operator  $A^*A$  (3.7) is discretized using a 3-point finite difference method respectively.

Plane wave source with source wavelet  $\omega(t)$  is implemented by placing isotropic point sources with the same wavelet at each grid point at the top of the computational grid and activating them with a delay proportional to source position. For a plane wave with an incidence angle  $\alpha$ , the delay for point source with an index  $j$  is

$$\tau_j = \frac{j\Delta x \sin \alpha}{c}, \quad (4.3)$$

where  $c$  is homogeneous velocity at the top of a model (water velocity).

## 4.2 Source wavelet

The (main) source wavelet  $\omega_m$  employed in all the numerical experiments is 8 Hz peak frequency Ricker (1940) passed through a low-cut (4-6 Hz) trapezoidal filter. Source wavelet and its spectrum are shown in Figure 4.1.

For a complimentary source  $\omega_c$ , I use a Gaussian wavelet ( $\omega_c = \exp\{-\pi^2 f^2 t^2\}$ ) with  $f = 8$  Hz passed through a high-cut (4-6 Hz) trapezoidal filter. Complimentary source wavelet and its spectrum are shown in Figure 4.2.

Full source  $\omega$  is a linear combination of the main source and the complimentary

source with weight  $w_c$ :

$$\omega = \omega_m + w_c \omega_c. \quad (4.4)$$

Full source with  $w_c = 0.5$  and its spectrum are shown in Figure 4.3.

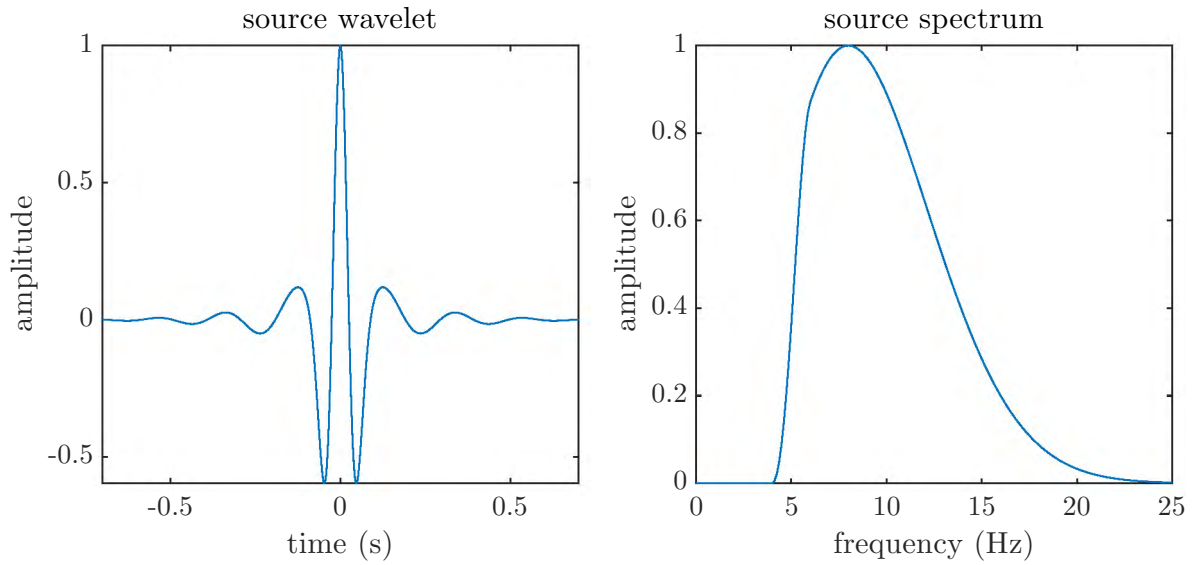


Figure 4.1: Source: Ricker wavelet with 8 Hz peak frequency passed through a low-cut (4-6 Hz) trapezoidal filter. Left: source wavelet. Right: source spectrum (normalized amplitude).

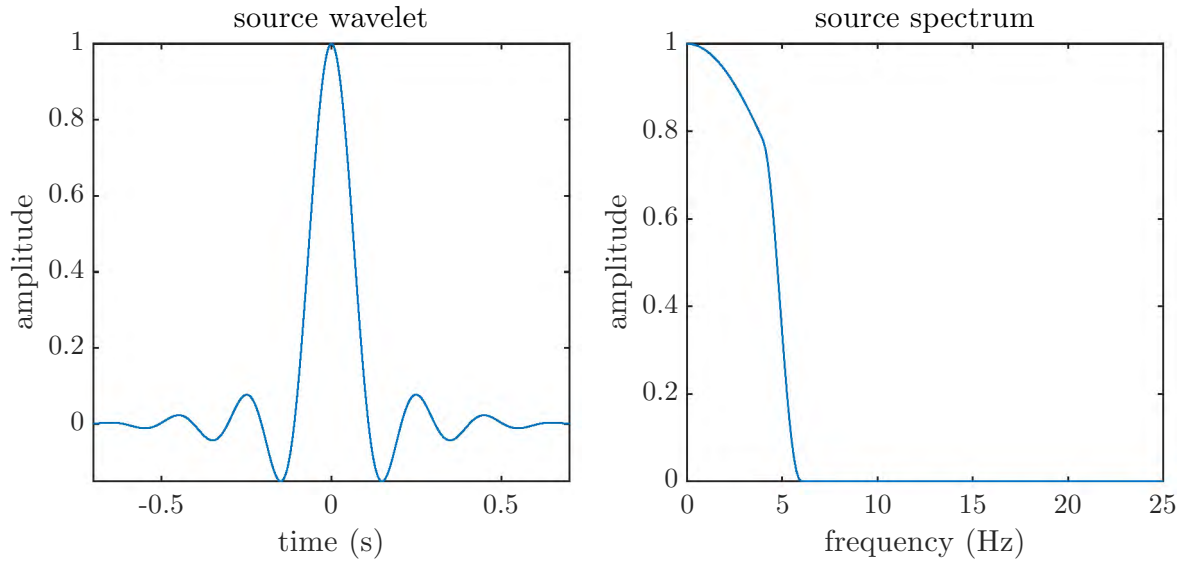


Figure 4.2: Complimentary source: Gaussian wavelet ( $\omega_c = \exp\{-\pi^2 f^2 t^2\}$ ) with  $f = 8$  Hz passed through a high-cut (4-6 Hz) trapezoidal filter. Left: source wavelet. Right: source spectrum (normalized amplitude).

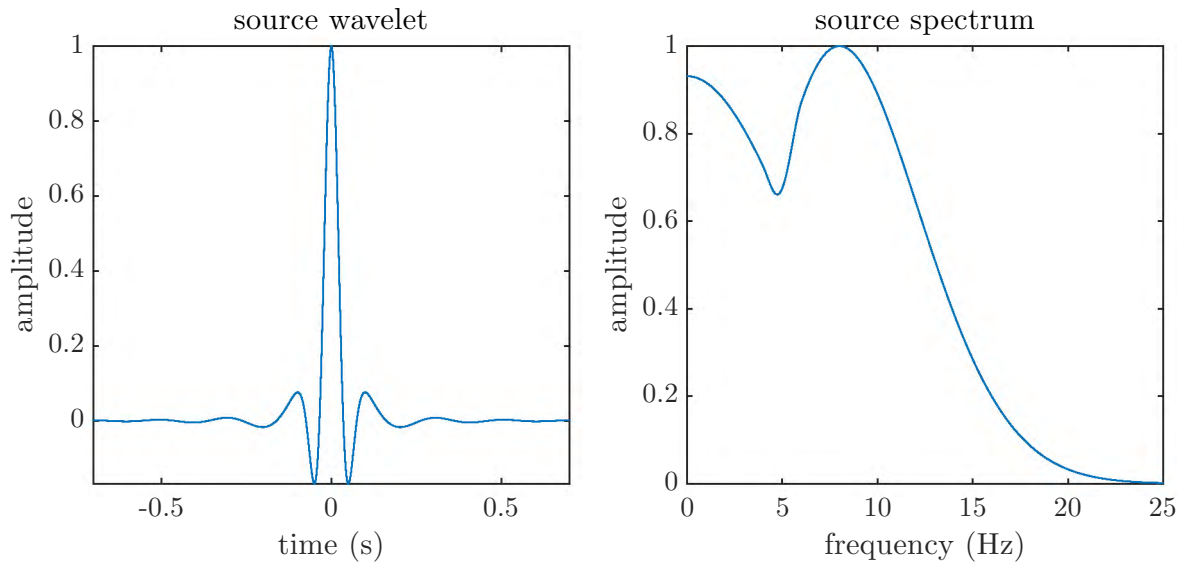


Figure 4.3: Full source: linear combination of main source and complimentary source with weight 0.5. Left: source wavelet. Right: source spectrum (normalized amplitude).

Numerical experiments showed that choice of complimentary source weight  $w_c$  is crucial for extended FWI algorithm. If complimentary source weight  $w_c$  is too large,

the inner problem is dominated by low frequency control data  $F_c[m_c]$ , which is fixed during each inner optimization. This allows to solve inner problem faster and more robustly, in particular, it eliminates the need in frequency continuation technique or allows for less accurate initial guess for the inner problem. However, the rest of the data (observed data  $d_o$  and predicted data  $F[\bar{m}]$ ) can be considered “noise” compared to the data generated by the control source. Thus, control model  $m_c$  behaves as an attractor for extended models, which are updated during the inner optimization. Since the initial guess for the first iteration of the outer problem is a constant “water model” for both  $m_c$  and  $\bar{m}$ , the algorithm is unable to improve these models. In absence of numerical and data noise and if the inner problem is solved accurately enough to produce reliable control model gradient, increasing control source weight  $w_c$  should result in slower convergence of the algorithm. However, in my numerical experiments choosing large  $w_c$  resulted in algorithm failure.

Border case  $w_c = 0$  obviously produces zero control data  $\bar{F}_c[\bar{m}]$  and  $F_c[m_c]$  regardless of models  $\bar{m}, m_c$  and results in zero gradient of the control model gradient. Similarly, if  $w_c$  is too small the control model gradient is not computed accurately and the inversion fails.

In my numerical experiments, weight of the control source was chosen on the basis of trial and error for each experiment.

## 4.3 Optimization

I solve the inner optimization problem with Polak-Ribière variation of the nonlinear CG method (Nocedal and Wright, 1999) with quadratic fit line search. The main stopping criteria is small reduction of the objective functional (with respect to a previous iteration) during several consecutive CG iterations. In all numerical examples, inner optimization process stopped when reduction of the objective functional was less than 1% in 8 consecutive CG iterations.

At  $k$ -th iteration of the outer optimization loop, inner optimization problems are solved for trial control models  $m_c^k = m_c^{k-1} + \alpha s_c^{k-1}$ , where  $\alpha$  is a step size and  $s_c^{k-1}$  is a search direction. To solve inner optimization problem, the extended model initial guess  $\bar{\mu}_0^k$  must be chosen. Choice of extended model initial guess played an important role in successful application of the extended modeling algorithm. I considered the following options:

- (1) trial control model as initial guess:  $\bar{\mu}_0^k = m_c^k$ ,
- (2) previous control model as initial guess:  $\bar{\mu}_0^k = m_c^{k-1}$ ,
- (3) solution of the inner problem for previous control model:  $\bar{\mu}_0^k = \bar{\mu}^{k-1}$ ,
- (4) solution of the inner problem for previous control model averaged over extension parameter  $h$ :  $\bar{\mu}_0^k = H[\bar{\mu}^{k-1}]$ , where  $H$  is an averaging operator over the acquisition parameter  $h$ .

Options (1) and (2) resulted in algorithm failure even in numerical tests with simplest models. Option (4) resulted in successful application of the algorithm in more

complicated numerical examples, such as Dome and Marmousi tests presented in this section. In my tests, I averaged squared slowness:

$$s_{ave}^2 = \frac{\sum_{i=1}^N s_i^2}{N}, \quad (4.5)$$

where  $N$  denotes number of extended models,  $s_i = 1/c_i[\bar{m}_i]$  denotes  $i$ -th slowness computed using inverse of equation (3.2). Averaged model is then computed by  $H[\bar{m}] = m[1/s_{ave}]$  via equation (3.2).

To accelerate the convergence of the outer problem, the following ad hoc modification of the option (4) can be used:

$$\bar{\mu}_0^k = (1 - \gamma)H[\bar{\mu}^{k-1}] + \gamma \max_h \bar{\mu}^{k-1}(\cdot; h), \quad (4.6)$$

where  $\gamma \in (0, 1)$  is a weight parameter. Using small weight ( $\gamma \approx 0.05 \sim 0.1$ ) significantly improved convergence during first iterations of the outer optimization loop. In my numerical tests, if  $\gamma$  was not reduced or set to zero at later iterations, the algorithm failed as the velocity models became too fast. Obviously, this modification works only if I start from slow-velocity initial guess (water) and the models get “faster” during optimization.

To solve the outer optimization problem, I considered Polak-Ribière CG and steepest descent methods. CG method did not show any advantage over the steepest descent, which might be due to inaccuracy of the solution of the inner problem and outer gradient calculation. A typical line search involves a sequence of step lengths, accepting the first that satisfies certain conditions. Numerical tests showed that choosing a very large step length can still lead to a sufficient decrease of the VP functional, thus,



eventually producing models with very high velocities and failing the inversion process. Typically (when starting optimization process from slow-velocity initial guess, such as water), extended models tend to be faster than the control model. A practical way to limit the step length  $\alpha$  in the outer optimization problem is to bound it by a distance to the average extended model:

$$\alpha_{max} \|\Lambda \nabla g^k\| < \|m_c^k - \Lambda H(\bar{\mu}^k)\|, \quad (4.7)$$

where  $m_c^k$  is the control model at  $k$ -th iteration,  $\bar{\mu}^k$  denotes a corresponding solution of the inner problem, and  $g_k$  denotes the outer gradient at  $\{m_c^k, \bar{\mu}^k\}$ .

In numerical tests presented in this thesis, I observed that the control model gradient tends to point in the direction of averaged extended model at first iterations of the VP method. Figure 4.35 shows the normalized dot product (angle cosine) of the negative outer gradient ( $-\Lambda \nabla \mathcal{J}[m_c]$ ) and the difference of the averaged extended model and the control model ( $\Lambda H(\bar{\mu}) - m_c$ ). With more iterations this quantity decreases and becomes less stable, which usually corresponds to stagnation of the optimization process.

## 4.4 Algorithm

The algorithm for extended FWI integrated with variable projection optimization approach used in this thesis is listed below. The algorithm is implemented in C++ with OpenMP-based parallelization (OpenMP, 2017). I use FFTW software package (Frigo and Johnson, 2005) to compute Fourier transforms used in filtering procedures.

---

**Algorithm 1** Extended FWI with VP optimization

---

- 1: Setup. Major parameters listed below:
  - 2:   Choose starting model  $m_{start}$
  - 3:   Choose regularization term and its weight
  - 4:   Choose annihilator weight
  - 5: Solve inner problem at 0-th iteration:
  - 6:    $m_c^0 \leftarrow m_{start}$
  - 7:   Prepare extended initial guess:  $\bar{m}(h) \leftarrow m_{start}$
  - 8:   CG with initial guess  $\bar{m}(h)$ :  $J^0 \leftarrow J_r[m_c^0]$  with minimizer  $\bar{\mu}^0$
  - 9: **for**  $k := 0, 1, \dots$  **do**
  - 10:   Compute outer gradient:  $g^k \leftarrow \nabla \mathcal{J}$  at  $\{m_c^k, \bar{\mu}^k\}$  using equation (2.29)
  - 11:   Smooth outer gradient:  $s^k \leftarrow \Lambda g_k$  using equation (2.9)
  - 12:   Prepare extended initial guess:  $\bar{m}(h) \leftarrow H[\bar{\mu}^k]$
  - 13:   Estimate initial step  $\alpha$  according to (4.7)
  - 14:   **loop** backtracking line search
  - 15:      $m_c^{k+1} \leftarrow m_c^k + \alpha s^k$
  - 16:     Solve inner problem:
  - 17:       CG with initial guess  $\bar{m}(h)$ :  $J^{k+1} \leftarrow J_r[m_c^{k+1}]$  with minimizer  $\bar{\mu}^{k+1}$
  - 18:     **if**  $J^{k+1} < J^k$  **then**
  - 19:       Accept  $m_c^{k+1}$ ,  $\bar{\mu}^{k+1}$ , and  $J^{k+1}$
  - 20:       **break**
  - 21:     **else**
  - 22:       reduce step  $\alpha$
  - 23:     **end if**
  - 24:     **if**  $\alpha$  too small **then**
  - 25:       Resume inner solver & solve inner problem more accurately:
  - 26:       Prepare extended initial guess:  $\bar{m}(h) \leftarrow \bar{\mu}^k(h)$
  - 27:       CG with initial guess  $\bar{m}(h)$ :  $J^k \leftarrow J_r[m_c^k]$  with updated minimizer  $\bar{\mu}^k$
  - 28:       **goto** 10
  - 29:     **end if**
  - 30:   **end loop**
  - 31: **end for**
-

## 4.5 Dome model

This section shows two numerical experiments (absorbing top boundary and free surface top boundary) with Dome model. Both experiments have identical setup (acquisition parameters, optimization parameters and strategy, etc.) described below.

The Dome model shown in Figure 4.4 is resampled to a  $23.35 \text{ m} \times 23.35 \text{ m}$  simulation grid ( $49 \times 213$  grid points) and padded with an extra 100 m water layer on top. The model velocity range is 1500–4000 m/s.

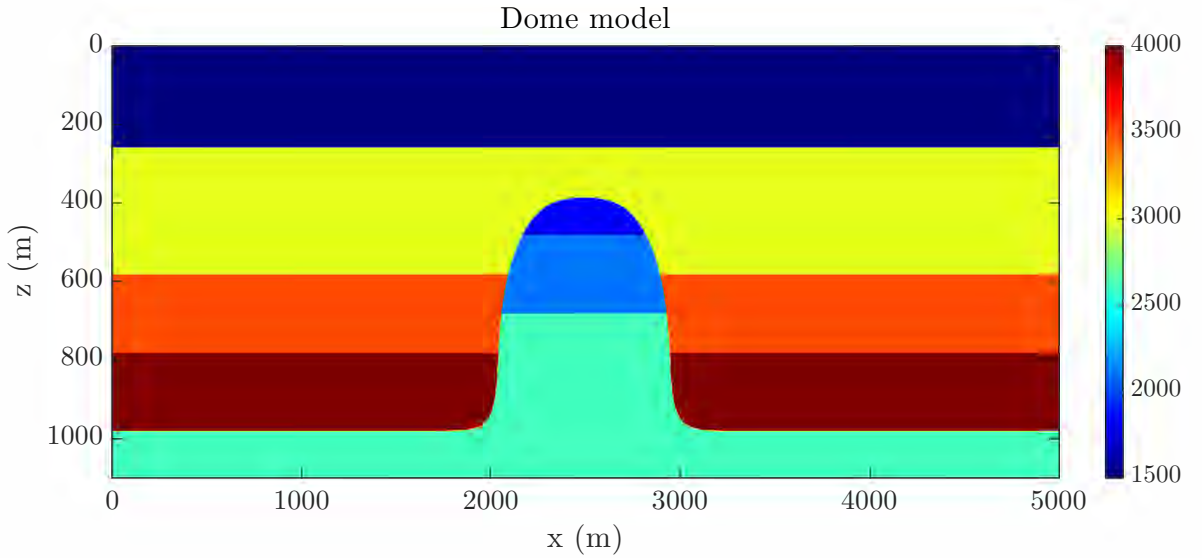


Figure 4.4: Dome acoustic velocity model. Horizontal layers: 1500, 3000, 3500, 4000, and 2600 m/s. Dome layers: 1800, 2100, and 2600 m/s.

I use 27 plane wave shots with plane wave incidence angle  $\alpha$  ranging from  $-25^\circ$  to  $25^\circ$ . Acquisition parameter  $h = \text{sgn} \alpha \sin^2 \alpha$  represents squared slowness, and is uniformly gridded between its end values. In all experiments, complimentary weight  $w_c = 0.2$  (see Section 4.2 and equation (4.4) for its description). This value of  $w_c$  is chosen on the basis of trial and error.

213 receivers are placed at grid points at the top of the domain. All traces for all shots are 1.5 s long. I use reduced-time shot records: each trace (for each shot) is recorded starting from its first arrival time. Therefore, each shot has different simulation time, increasing with shot incidence angle.

For initial guess (initial control and extended models in the first outer iteration of VP method), I use a homogeneous “water model”  $c = 1500$  m/s. Top 200 meter layer of control and extended models is fixed to water velocity at all inner and outer iterations (all gradients are muted).

I add total variation (TV) regularization term to the objective functional of the inner problem. TV regularization promotes sparsity of the model spatial gradient (Aster et al., 2012) and works well for “blocky” models such as Dome model. Weight of the TV regularization term is chosen on the basis of trial and error for each numerical experiment and is same for all iterations of the outer optimization problem.

### 4.5.1 Conventional inversion experiment

Here, I present results of two conventional inversion experiments: (1) — with kinematically correct initial guess, and (2) — with initial guess that does not provide accurate enough long-scale velocity structure.

Conventional optimization problem is solved with Polak-Ribière CG method, similarly to the inner problem of the VP method for the extended modeling approach.

Top row of Figure 4.5 shows initial velocity models for conventional experiments: initial model (1) is produced by smoothing the true slowness with a moving 500 m

box average. Initial model (2) is a vertical velocity gradient model.

Bottom row of Figure 4.5 shows conventional FWI results. Inversion (1) is capable of recovering the model very precisely. However, in case of inaccurate initial guess, which is still better than constant initial “water model” used in extended approach investigated in this thesis, the conventional inversion (2) fails.

Relative objective functional values are 0.2% and 41% for experiments (1) and (2) respectively. Recovered model for experiment (1) has a noticeable model error at the bottom under the dome (bottom-left model in Figure 4.5). Nevertheless, data is fitted very accurately (0.2% relative error), which indicates that the model error “belongs” to the null space of the objective functional.

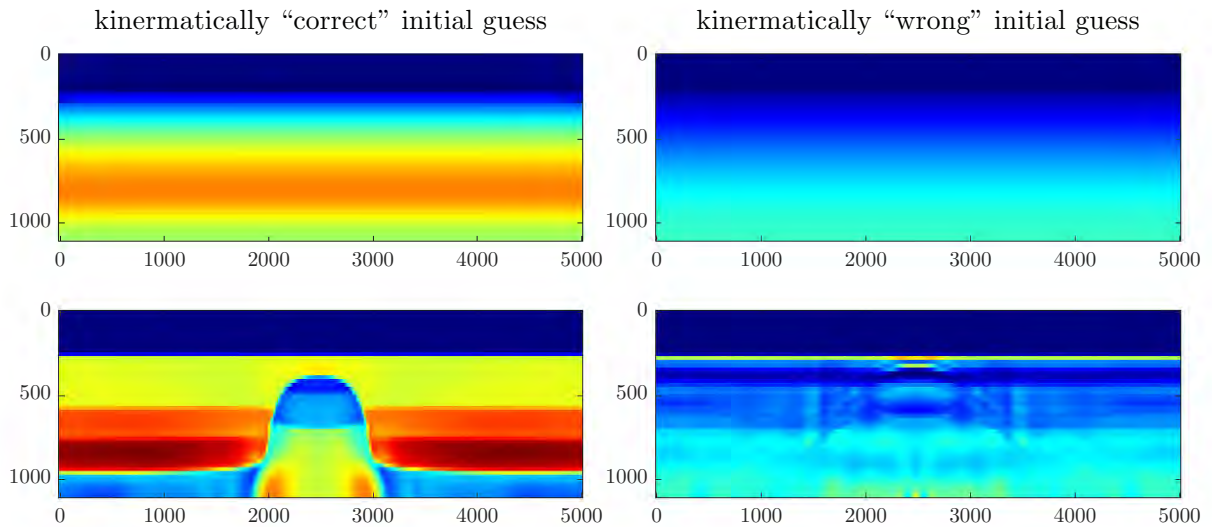


Figure 4.5: Dome with absorbing surface example: conventional FWI; initial models (top), and recovered models (bottom). Left: inversion with kinematically correct initial model. Right: inversion with inaccurate long-scale structure of initial model.

## 4.5.2 Absorbing boundary experiment

In this section, I present numerical test results for the Dome model with absorbing boundary conditions on all four sides of the model.

Figure 4.6 shows observed data (with direct arrival removed) for two shots:  $0^\circ$  incidence angle (horizontal plane-wave) and  $20^\circ$  incidence angle.

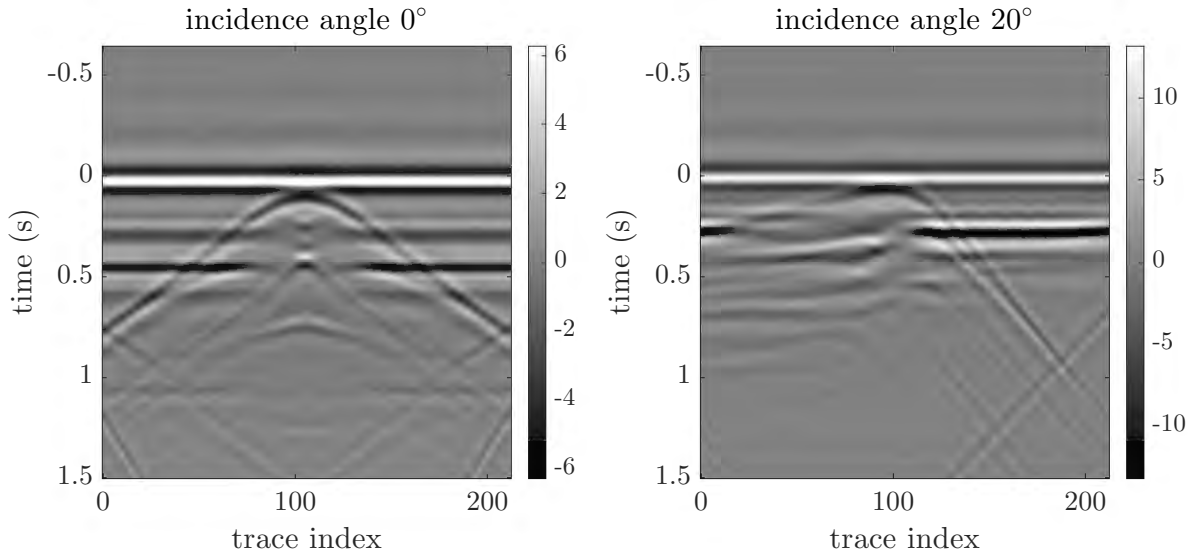


Figure 4.6: Dome with absorbing surface example: shot records (direct arrivals removed). Left: plane wave with 0 degrees incidence angle. Right: plane wave with 20 degrees incidence angle.

Figure 4.7 shows intermediate results of the inversion process at outer iterations 1, 2, 4, 8, 16, and 32. Left column presents averaged (over extension axis) extended models. Right column shows the control models. Data misfit history, shown in Figure 4.8, indicates that there is little improvement in the error after 25 iterations. As shown, after 32 iterations, the model is recovered quite accurately (less than 1% data fitting error). After 20 or so iterations, the model has sufficiently accurate

long-scale information to switch to a conventional FWI, presented in Section 4.5.1. As expected, control model does not contain short-scale information and resembles smoothed version of the extended model.

Figures 4.9 and 4.10 show shot records (direct arrivals removed) of observed data (left column) and data residuals (right column) after 32 iterations for shots with  $0^\circ$  and  $20^\circ$  incidence angles respectively.

Figure 4.11 compares leftmost (receiver at the left model edge) and central (receiver over dome top) traces of predicted and true data for  $0^\circ$  incidence angle plane wave shot. Figure 4.12 gives similar comparison for  $20^\circ$  incidence angle plane wave shot. These two figures show that data is fitted very well for both offsets.

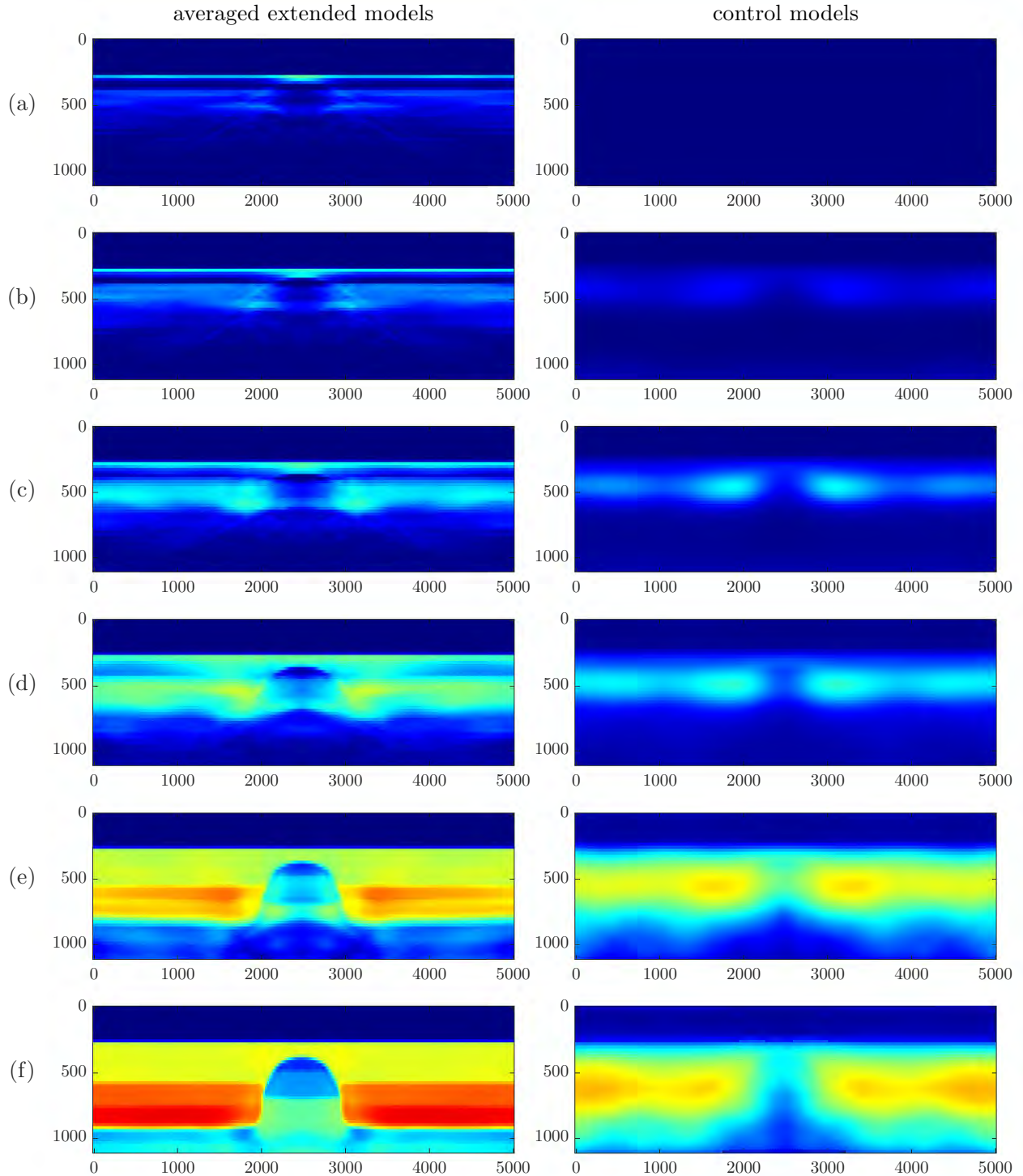


Figure 4.7: Dome with absorbing surface example: averaged extended (left) and control (right) models. Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), and 32 (f) outer iterations of the VP method.



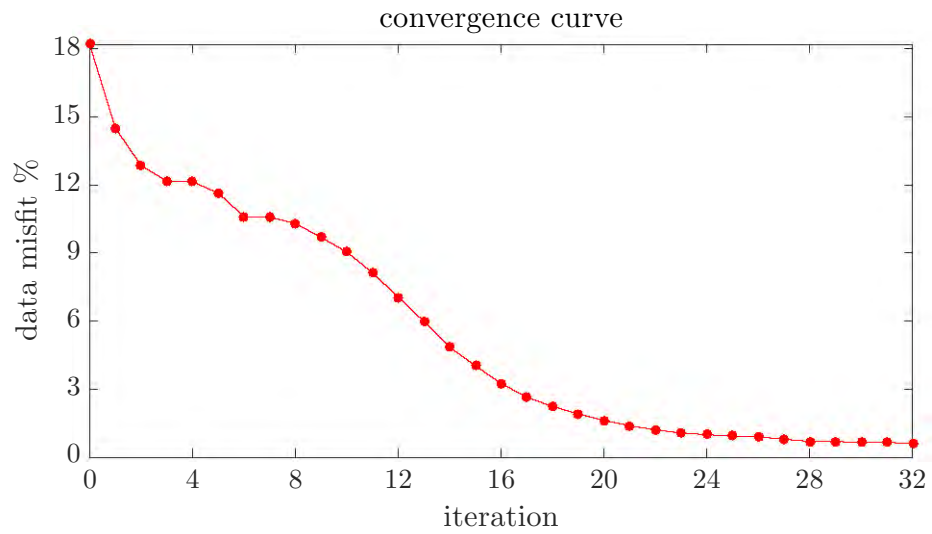


Figure 4.8: Dome with absorbing surface example: normalized data misfit convergence curve of steepest descent method for outer optimization.

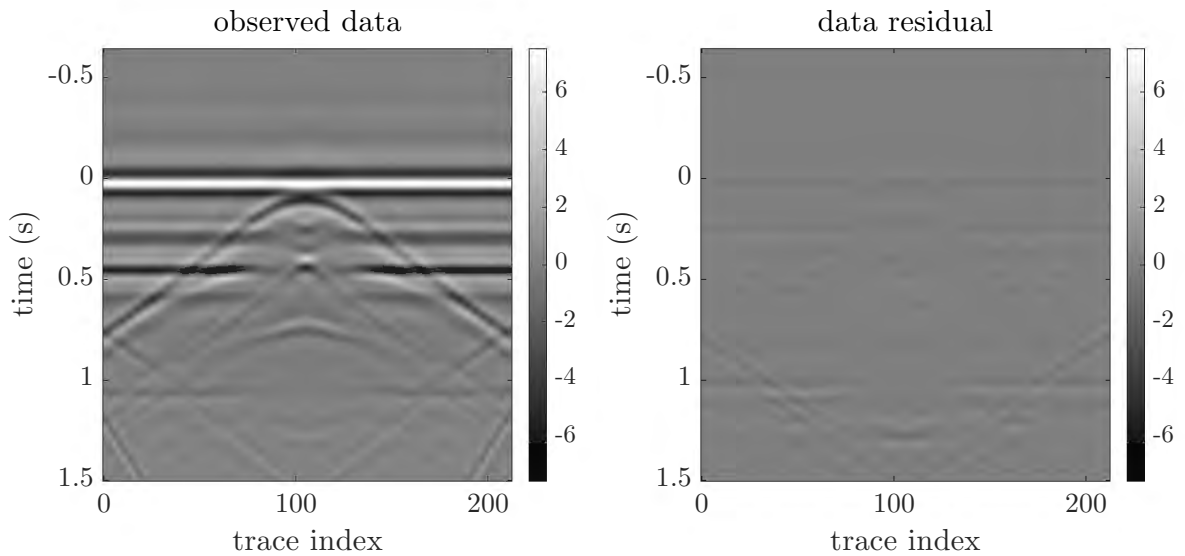


Figure 4.9: Dome with absorbing surface example: data residual after 32 outer iterations of the VP method. Left: shot record for plane wave with 0 degrees incidence angle. Right: data residual.

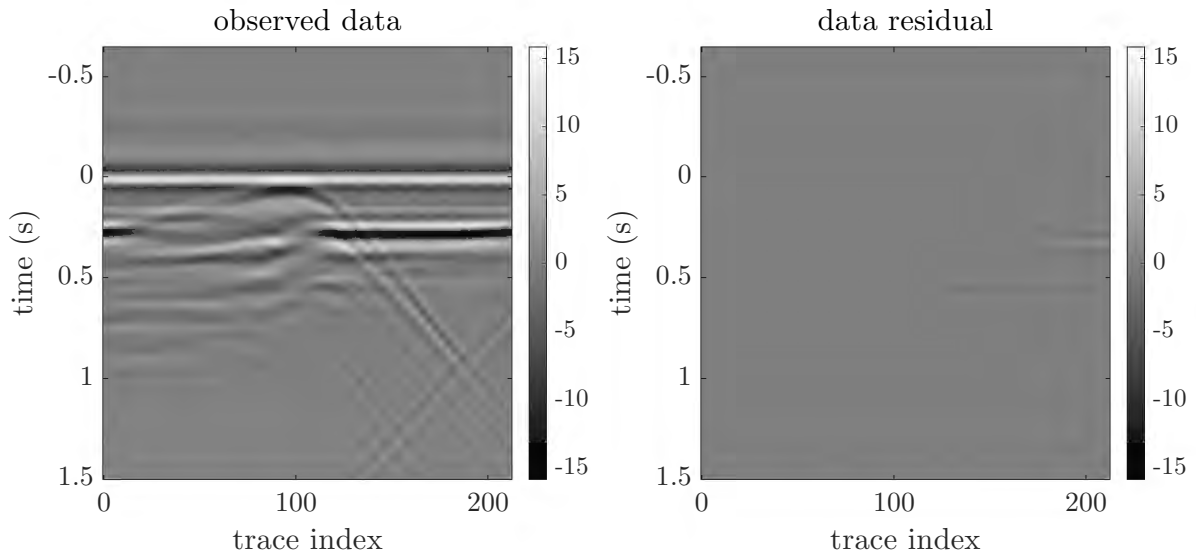


Figure 4.10: Dome with absorbing surface example: data residual after 32 outer iterations of the VP method. Left: shot record for plane wave with 20 degrees incidence angle. Right: data residual.

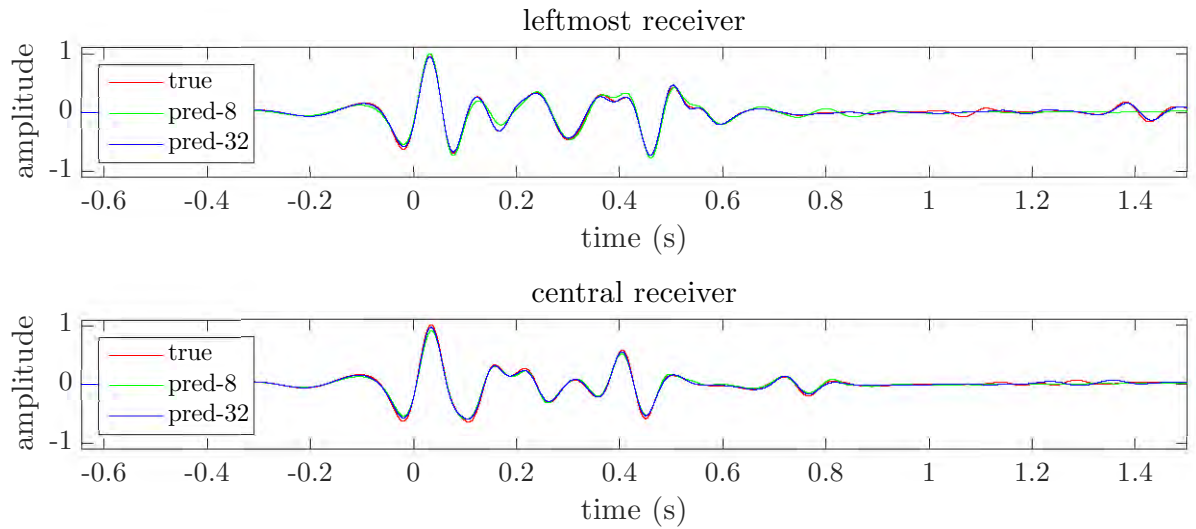


Figure 4.11: Dome with absorbing surface example: trace comparison of true data (red) and predicted data after 8 iterations (green) and after 32 iterations (blue).  $0^\circ$  incidence angle plane wave shot. Top: leftmost receiver. Bottom: central receiver.

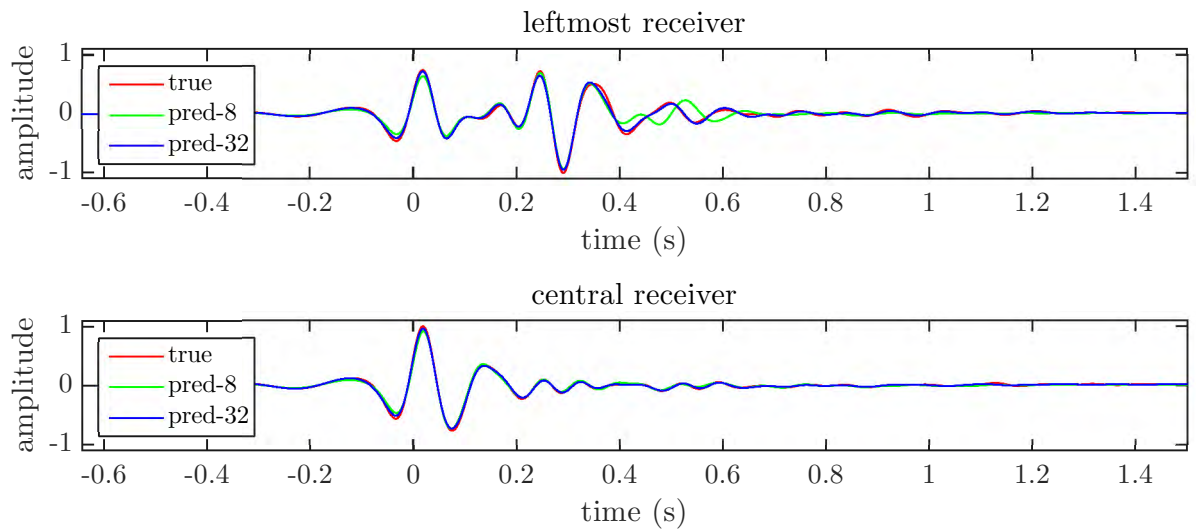


Figure 4.12: Dome with absorbing surface example: trace comparison of true data (red) and predicted data after 8 iterations (green) and after 32 iterations (blue).  $20^\circ$  incidence angle plane wave shot. Top: leftmost receiver. Bottom: central receiver.

Figures 4.14–4.17 present model gathers at gather locations shown on Figure 4.13. Model gathers are flattened after 32 iterations.

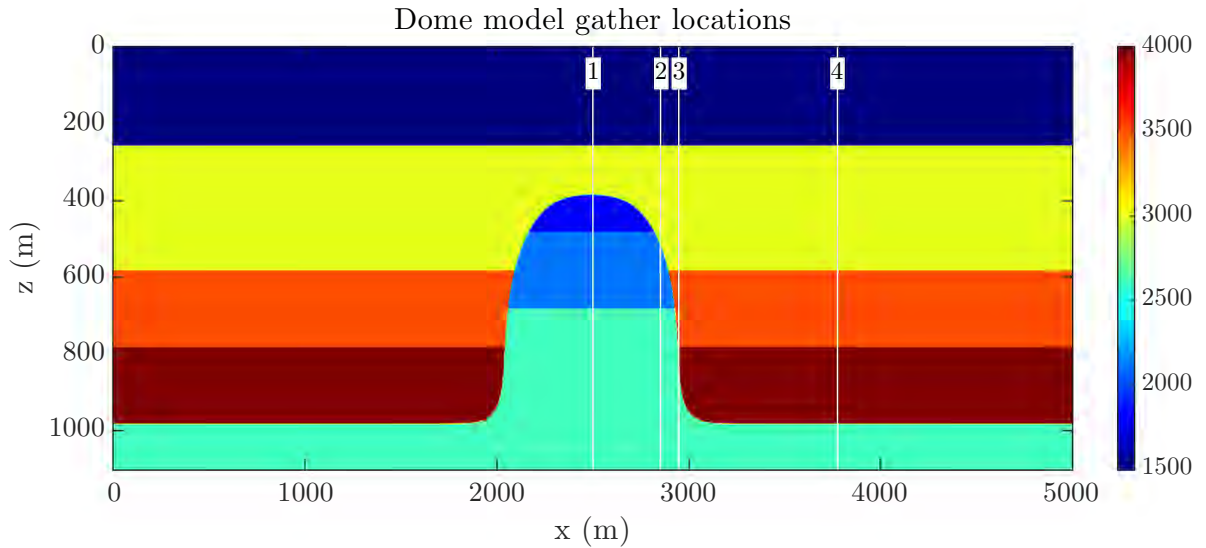


Figure 4.13: Dome acoustic velocity model: gather locations. Location (1) at  $x = 2500$  m, (2) at  $x = 2854$  m, (3) at  $x = 2948$  m, and (4) at  $x = 3774$  m.

For this experiment, the surface-oriented extended FWI approach is capable of recovering an accurate velocity model (both long- and short-scale structure) and achieving a very small data fitting error, while starting from a kinematically wrong initial model. Such results could not be achieved with a conventional FWI approach.

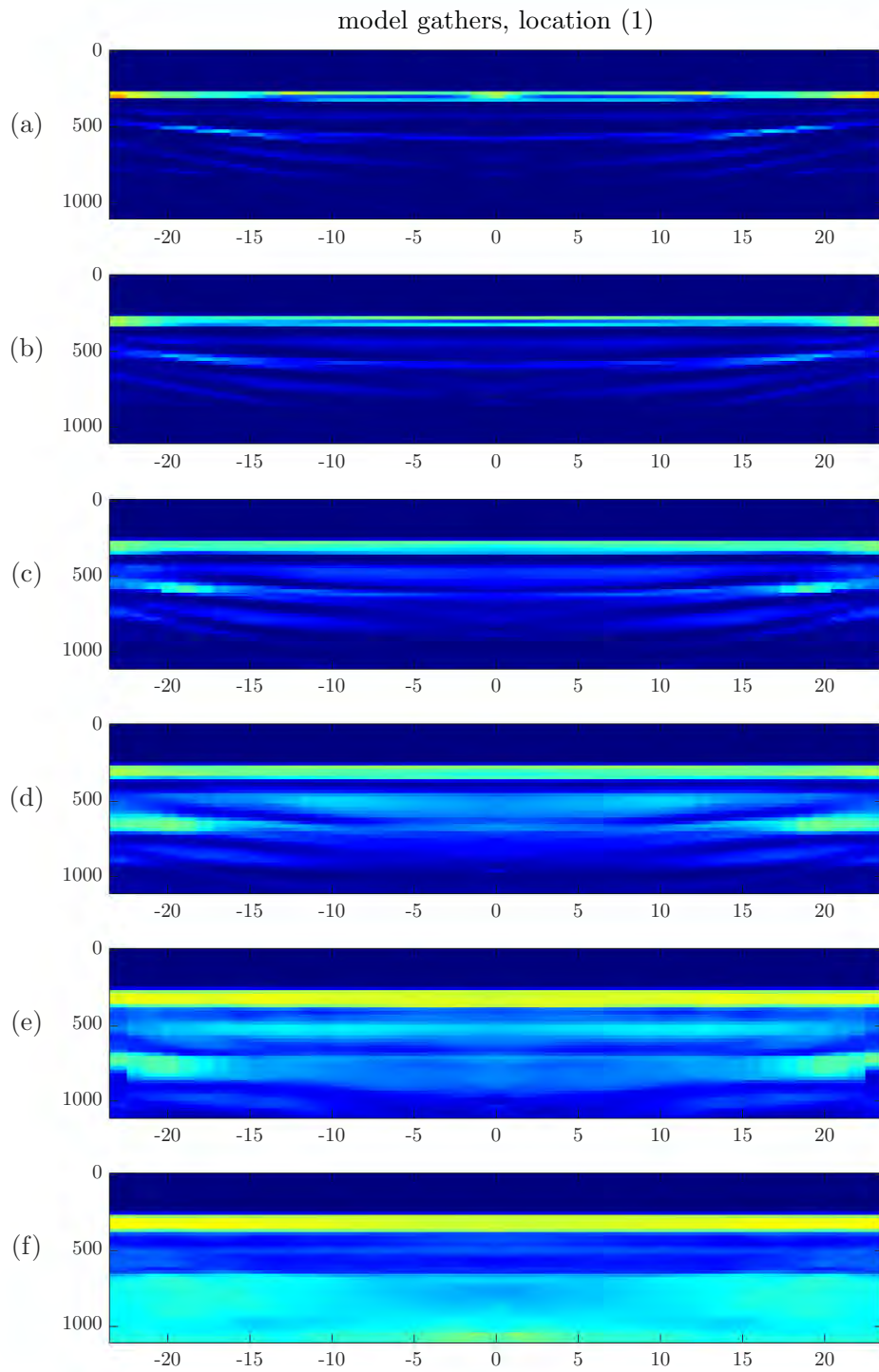


Figure 4.14: Dome with absorbing surface example: model gathers at location (1). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), and 32 (f) outer iterations of the VP method.

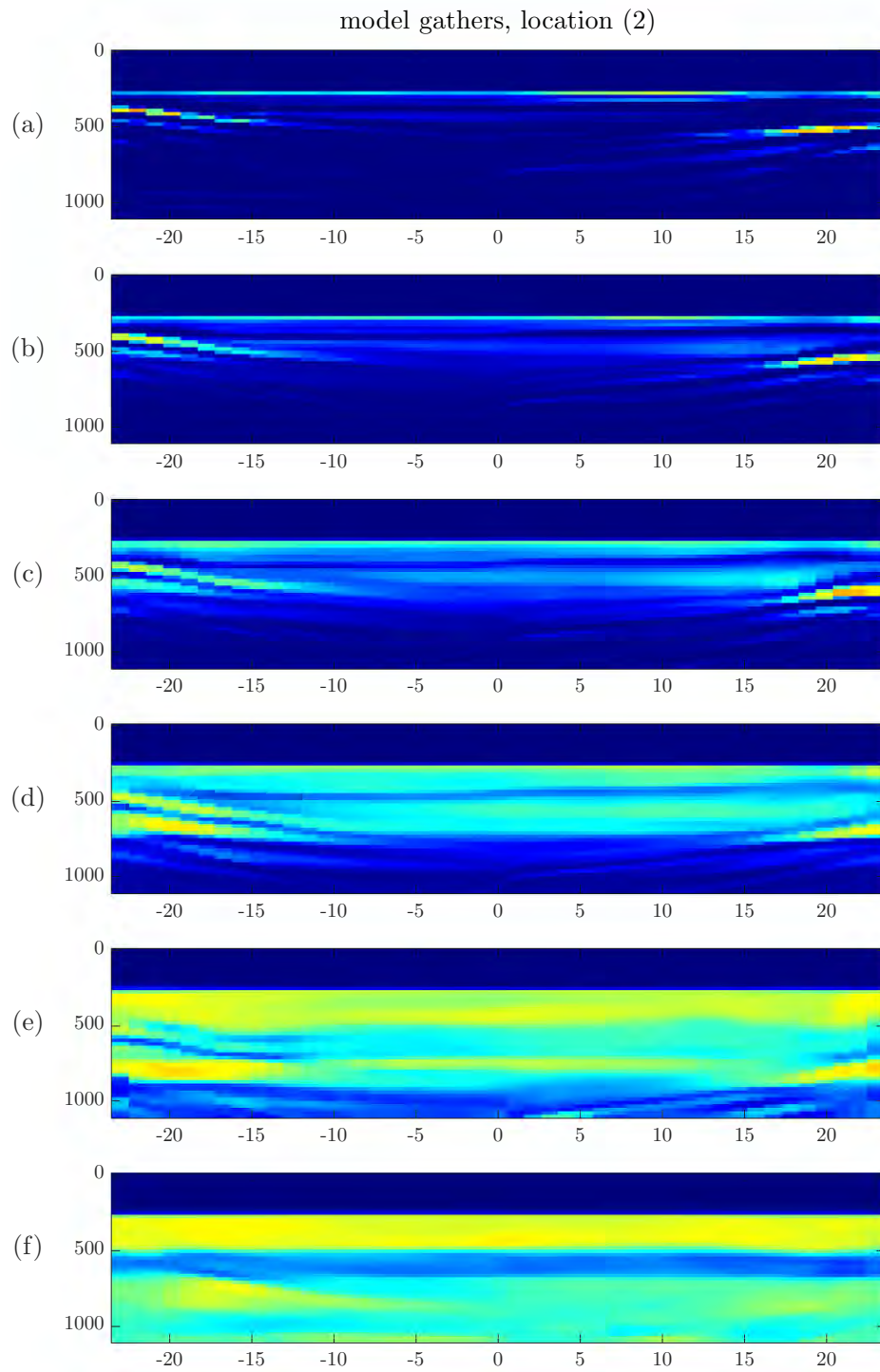


Figure 4.15: Dome with absorbing surface example: model gathers at location (2). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), and 32 (f) outer iterations of the VP method.

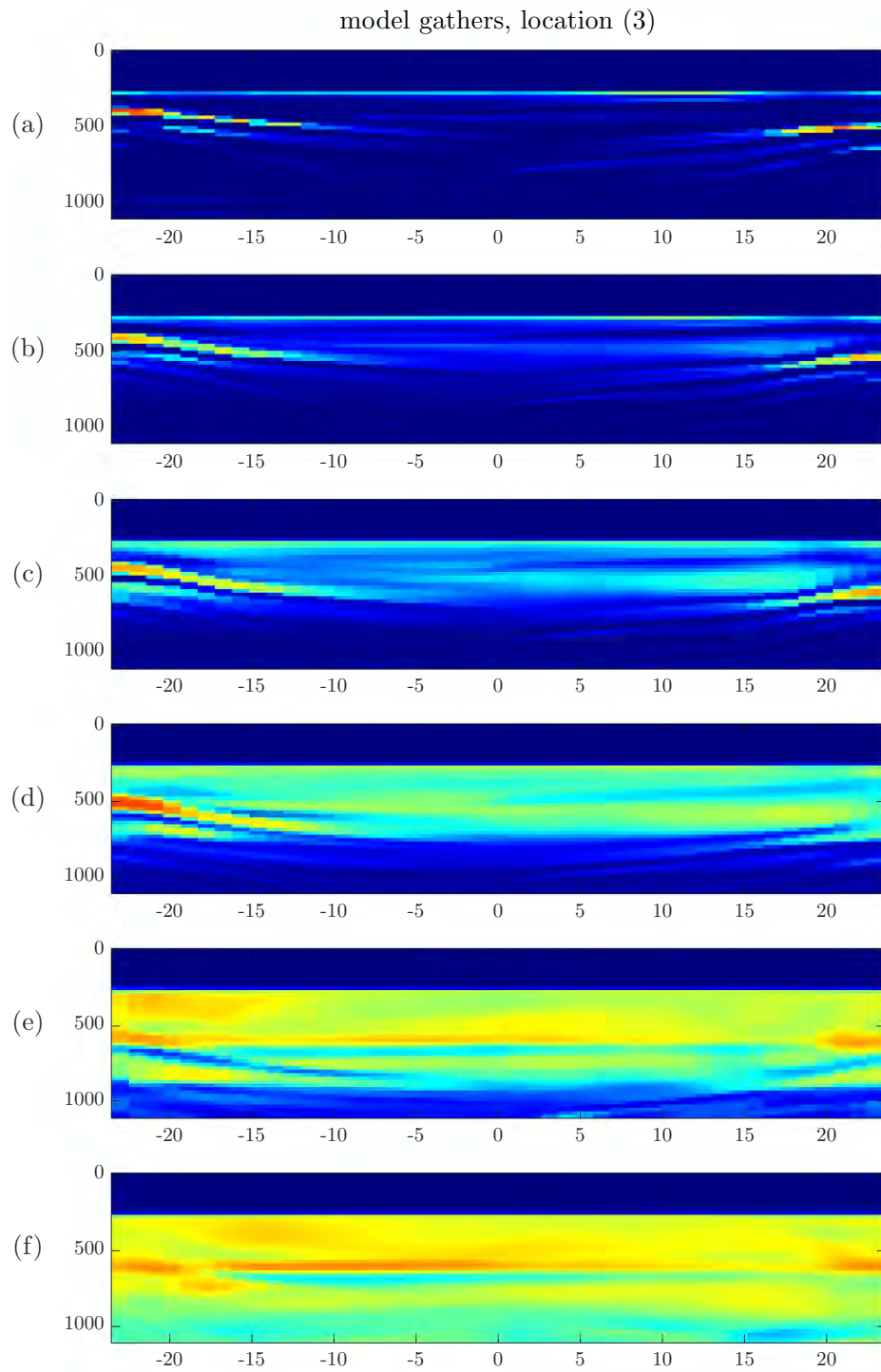


Figure 4.16: Dome with absorbing surface example: model gathers at location (3). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), and 32 (f) outer iterations of the VP method.

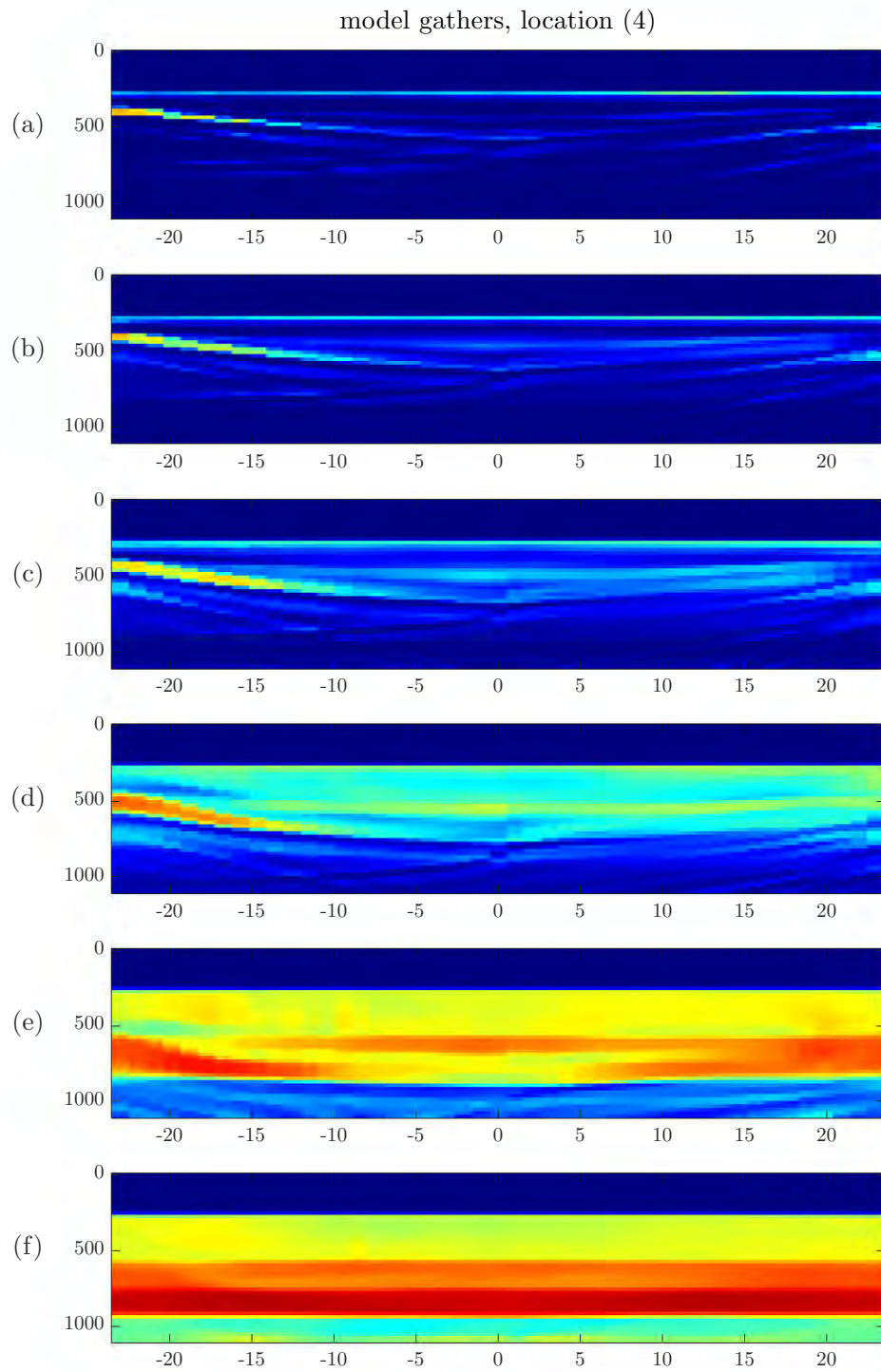


Figure 4.17: Dome with absorbing surface example: model gathers at location (4). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), and 32 (f) outer iterations of the VP method.



### 4.5.3 Free surface experiment

In this section, I present numerical test results for the Dome model with free surface boundary condition on top of the model (absorbing boundary conditions on three other sides).

Figure 4.18 shows observed data (with direct arrival removed) for two shots:  $0^\circ$  incidence angle (horizontal plane-wave) and  $20^\circ$  incidence angle. Significant multiple reflections (multiples) due to the free surface are present.

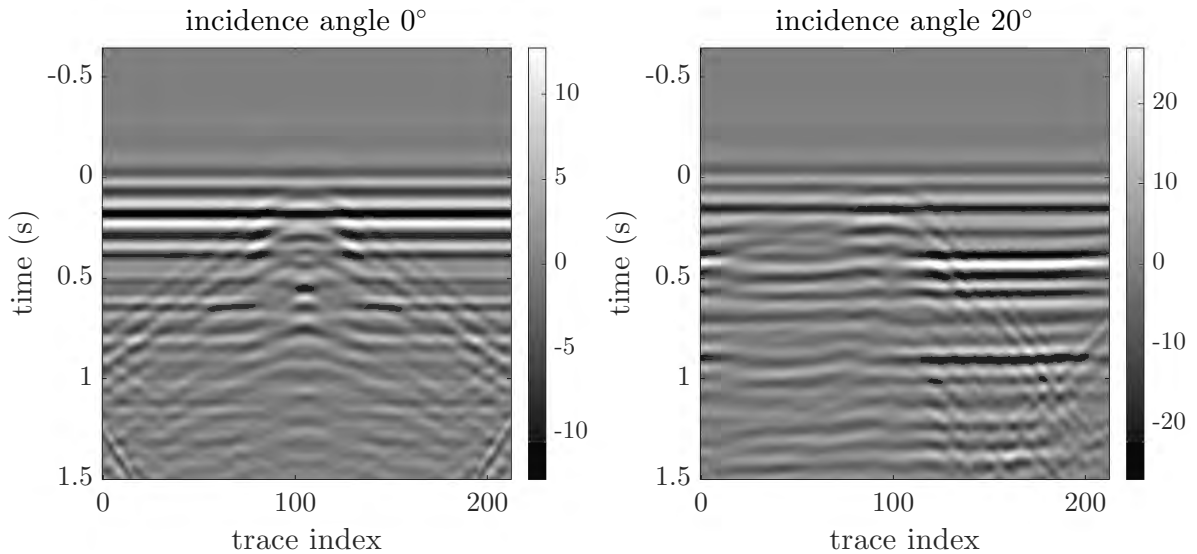


Figure 4.18: Dome with free surface example: shot records (direct arrivals removed). Left: plane wave with 0 degrees incidence angle. Right: plane wave with 20 degrees incidence angle.

Figure 4.19 shows intermediate results of the inversion process at outer iterations 1, 2, 4, 8, 16, 32, and 44. Left column presents averaged (over extension axis) extended models. Right column shows the control models.

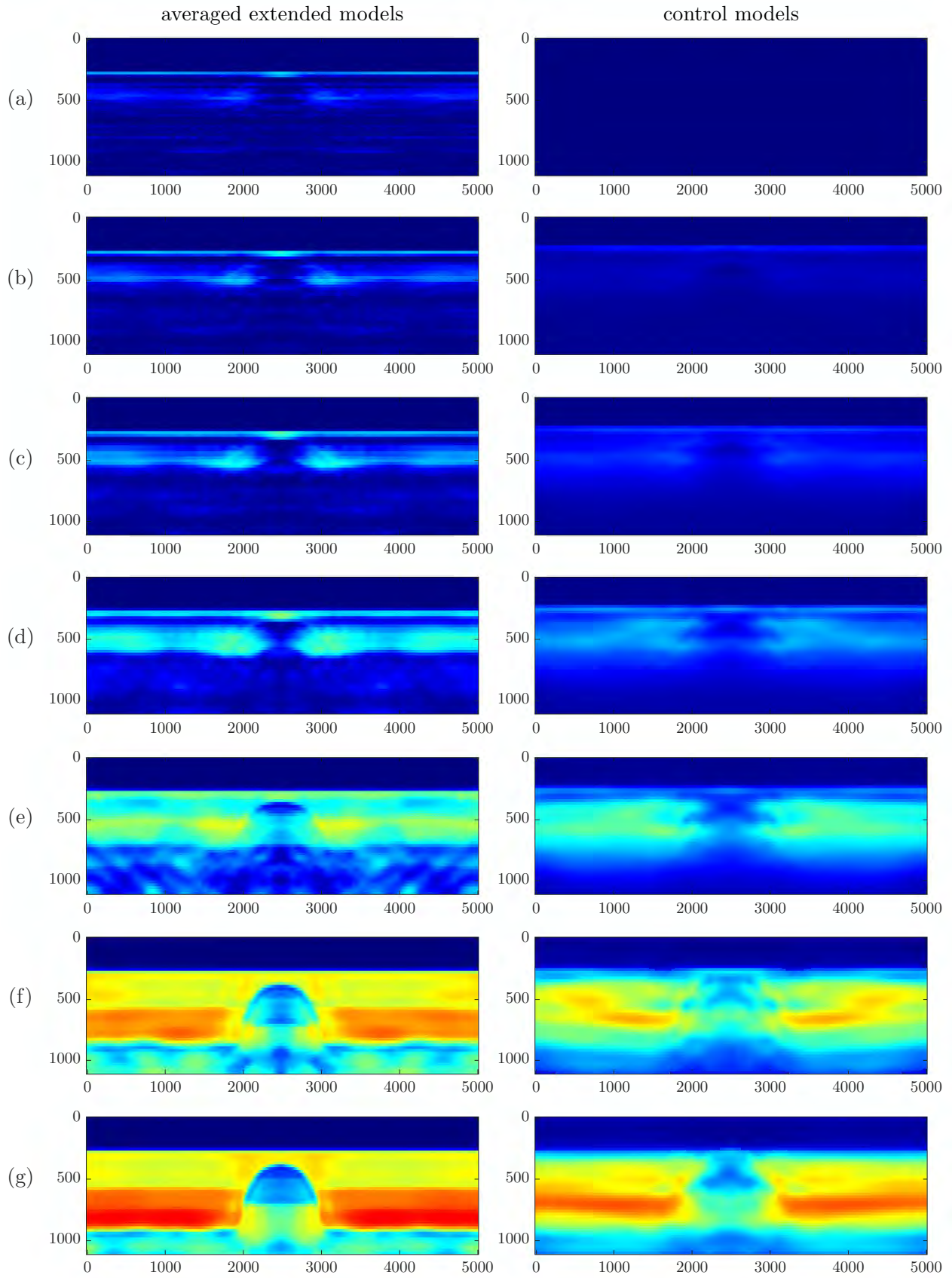


Figure 4.19: Dome with free surface example: averaged extended (left) and control (right) models. Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), 32 (f), and 44 (g) outer iterations of the VP method.

Data misfit history, shown in Figure 4.20, indicates that there is little improvement in the error after 30 iterations. As shown, after 44 iterations, the model is recovered quite accurately (1% data fitting error). After 30 or so iterations, the model has sufficiently accurate long-scale information to switch to a conventional FWI, presented in Section 4.5.1.

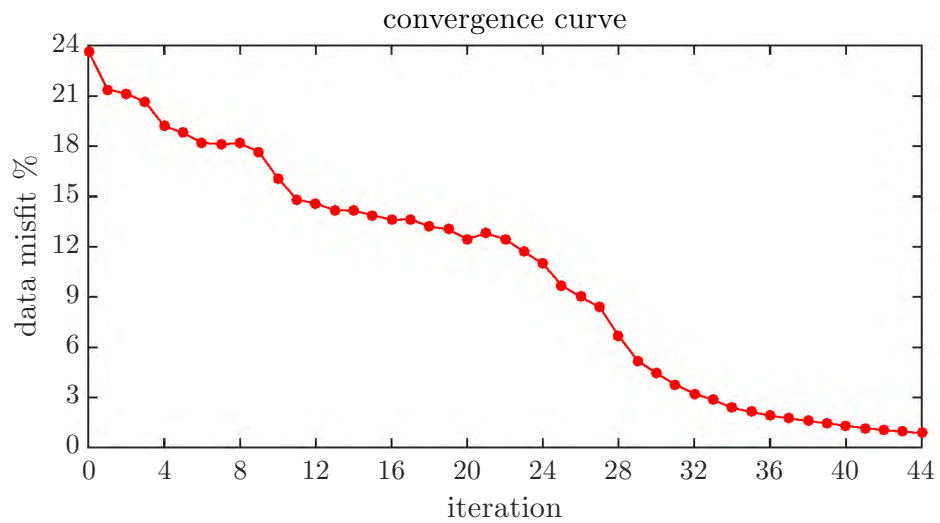


Figure 4.20: Dome with free surface example: normalized data misfit convergence curve of steepest descent method for outer optimization.

Figures 4.21 and 4.22 show shot records (direct arrivals removed) of observed data (left column) and data residuals (right column) after 32 iterations for shots with  $0^\circ$  and  $20^\circ$  incidence angles respectively.

Figures 4.23–4.26 present model gathers at gather locations shown on Figure 4.13. Model gathers are flattened after 44 iterations.

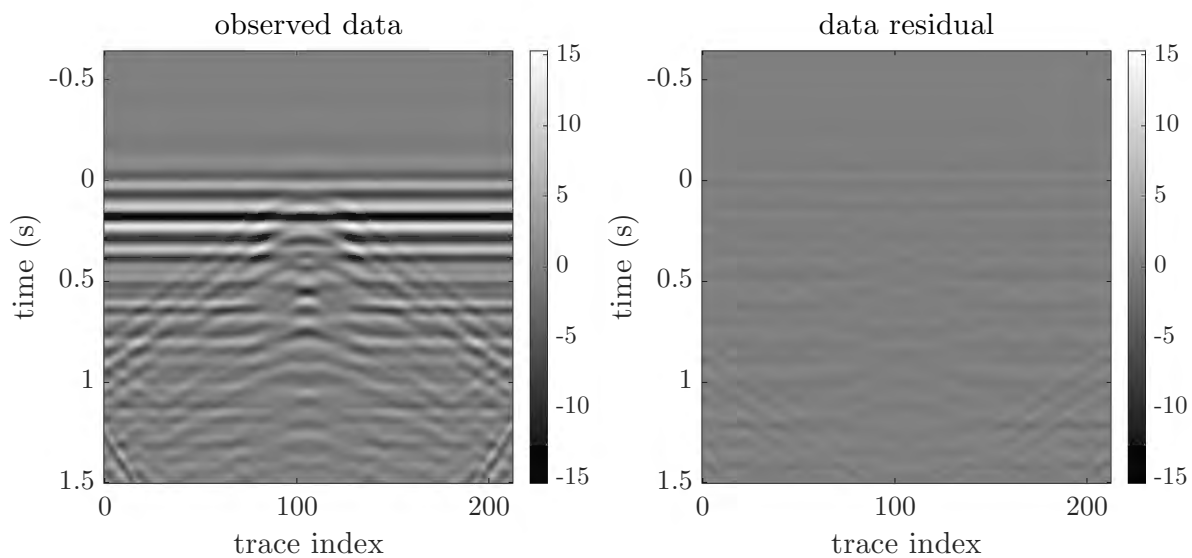


Figure 4.21: Dome with free surface example: data residual after 32 outer iterations of the VP method. Left: shot record for plane wave with 0 degrees incidence angle. Right: data residual.

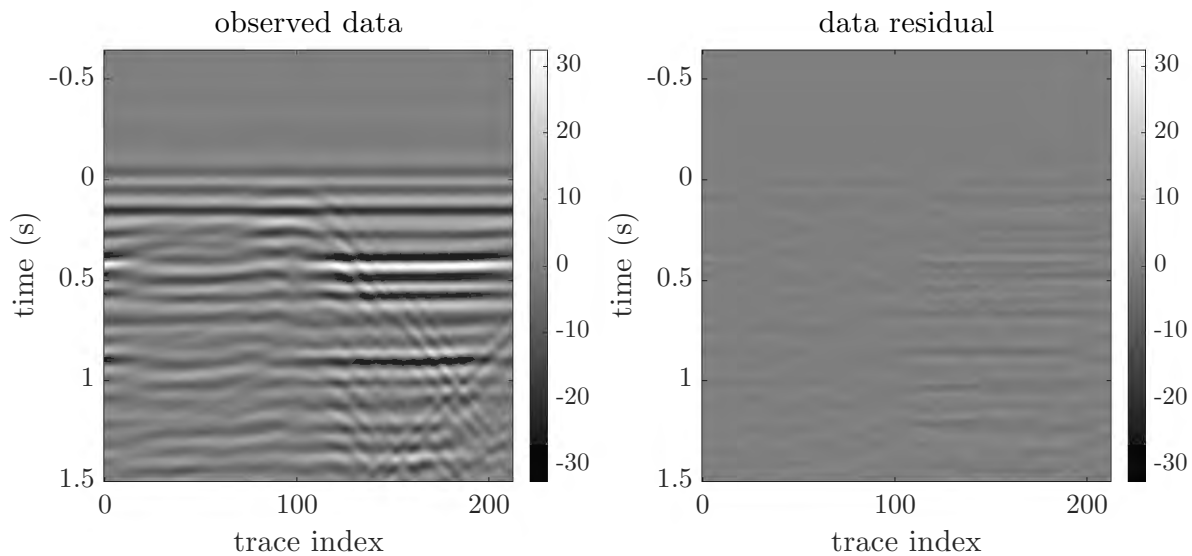


Figure 4.22: Dome with free surface example: data residual after 32 outer iterations of the VP method. Left: shot record for plane wave with 20 degrees incidence angle. Right: data residual.

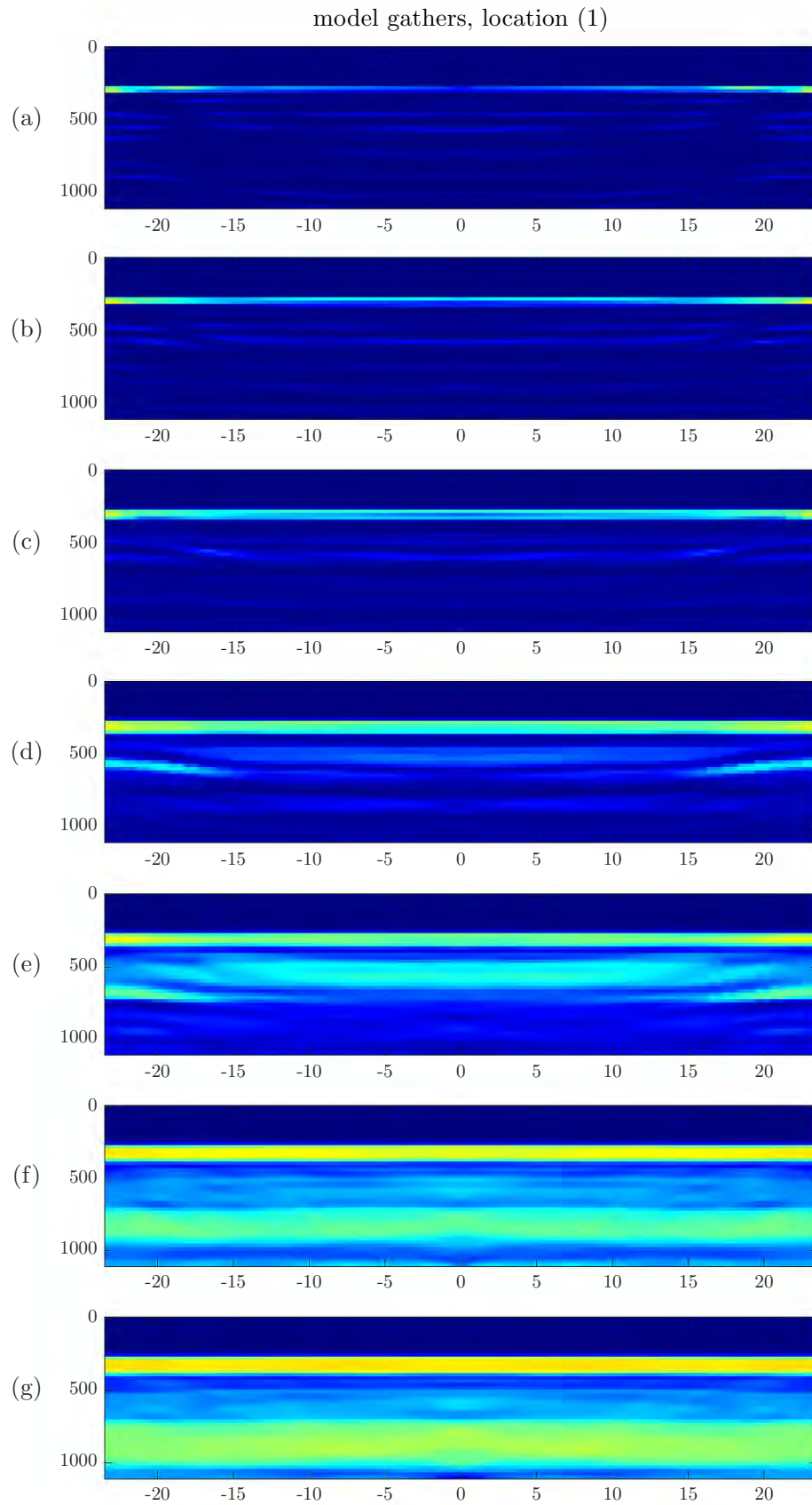


Figure 4.23: Dome with free surface example: model gathers at location (1). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), 32 (f), and 44 (g) outer iterations of the VP method.

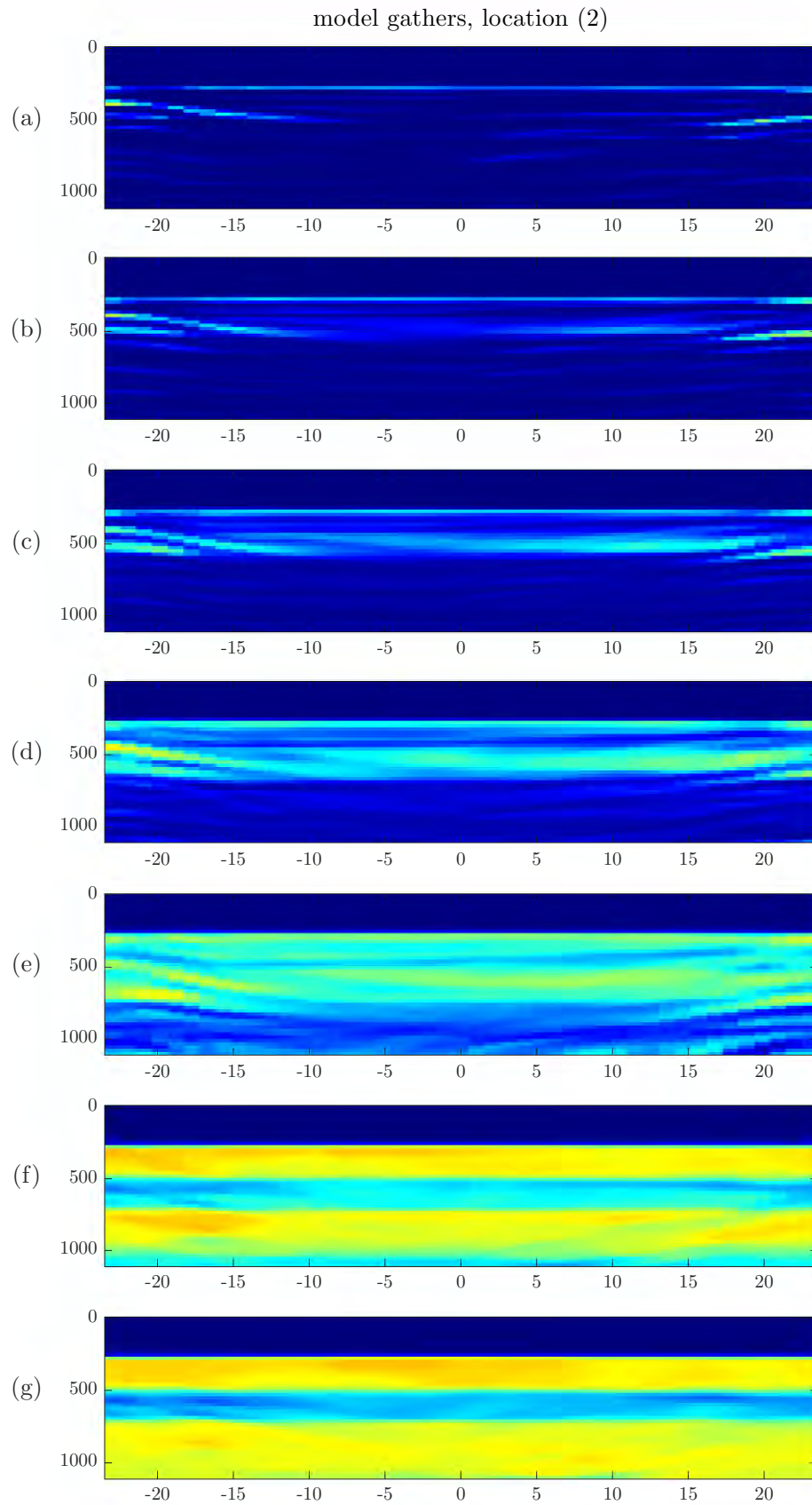


Figure 4.24: Dome with free surface example: model gathers at location (2). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), 32 (f), and 44 (g) outer iterations of the VP method.

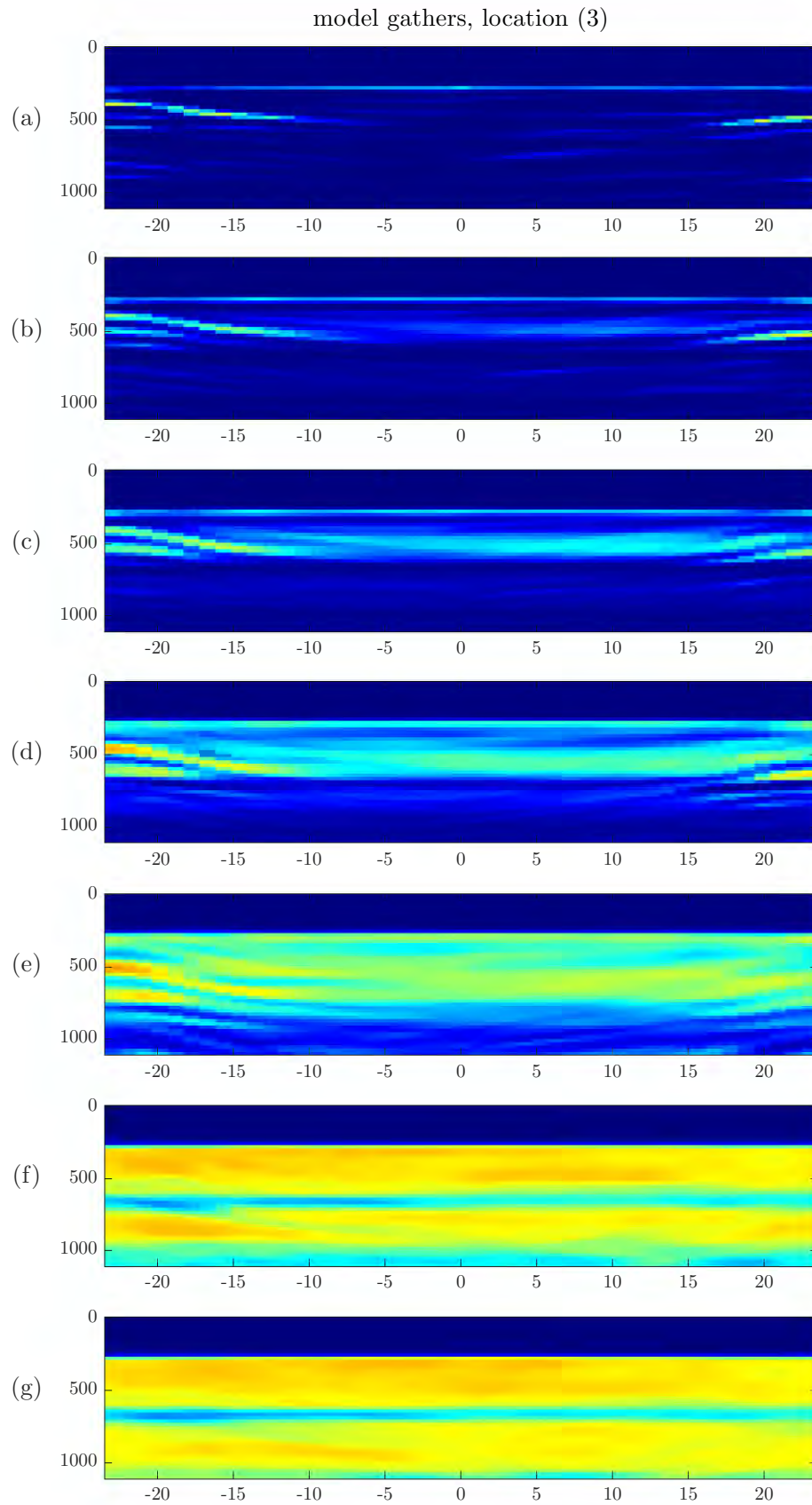


Figure 4.25: Dome with free surface example: model gathers at location (3). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), 32 (f), and 44 (g) outer iterations of the VP method.

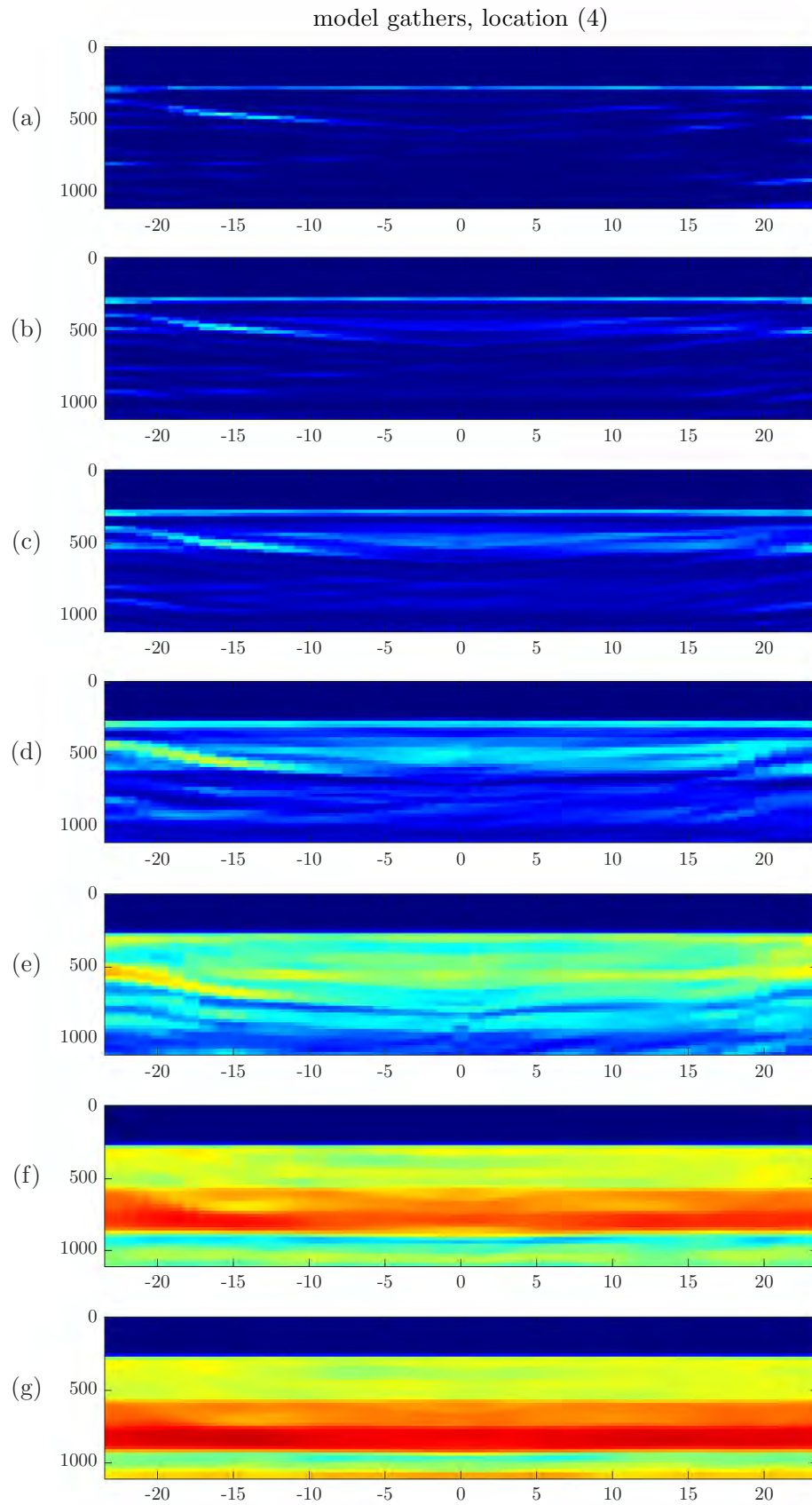


Figure 4.26: Dome with free surface example: model gathers at location (4). Inverted models after 1 (a), 2 (b), 4 (c), 8 (d), 16 (e), 32 (f), and 44 (g) outer iterations of the VP method.



In current state of FWI practice, multiples present significant obstacle to a successful inversion process. Typically, multiples are treated as noise and are removed from the data during the pre-processing stage of the inversion. Thus, presence of strong free surface related multiples would make this experiment much more difficult for conventional FWI. Nevertheless, extended FWI is capable of recovering an accurate velocity model resulting in a small data misfit ( $\sim 1\%$ ). Results of this numerical tests are comparable to the results of the absorbing surface test. However, slower convergence of the method is observed in presence of the multiples.

## 4.6 Marmousi model

This section shows numerical experiment with the Marmousi model (Bourgeois et al., 1991).

The Marmousi model shown in Figure 4.27 is resampled to a  $23.41 \text{ m} \times 23.41 \text{ m}$  simulation grid ( $94 \times 411$  grid points) and padded with an extra 100 m water layer on top. The model velocity range is 1500–4450 m/s.

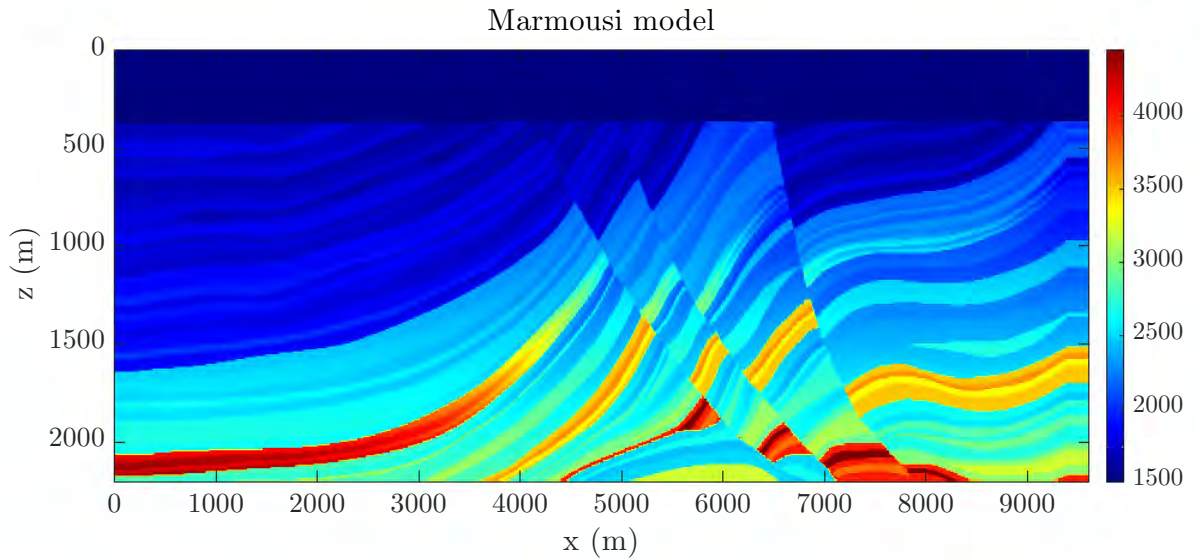


Figure 4.27: Marmousi acoustic velocity model.

I use 29 plane wave shots with plane wave incidence angle  $\alpha$  ranging from  $-30^\circ$  to  $30^\circ$ . Acquisition parameter  $h = \text{sgn } \alpha \sin^2 \alpha$  represents squared slowness, and is uniformly gridded between its end values. In all experiments, complimentary weight  $w_c = 0.1$  (see Section 4.2 and equation (4.4) for its description). This value of  $w_c$  is chosen on the basis of trial and error.

213 receivers are placed at grid points at the top of the domain. All traces for all

shots are 2 s long. As with the Dome experiment, I use reduced-time shot records.

For initial guess (initial control and extended models in the first outer iteration of VP method), I use a homogeneous “water model”  $c = 1500$  m/s. Top 360 meter layer of control and extended models is fixed to water velocity at all inner and outer iterations (all gradients are muted).

I add Tikhonov regularization term to the objective functional of the inner problem. Weight of the Tikhonov regularization term is chosen on the basis of trial and error for each numerical experiment and is same for all iterations of the outer optimization problem.

Figure 4.28 shows observed data (with direct arrival removed) for three shots:  $0^\circ$  incidence angle (horizontal plane-wave),  $30^\circ$  incidence angle, and  $-30^\circ$  incidence angle.

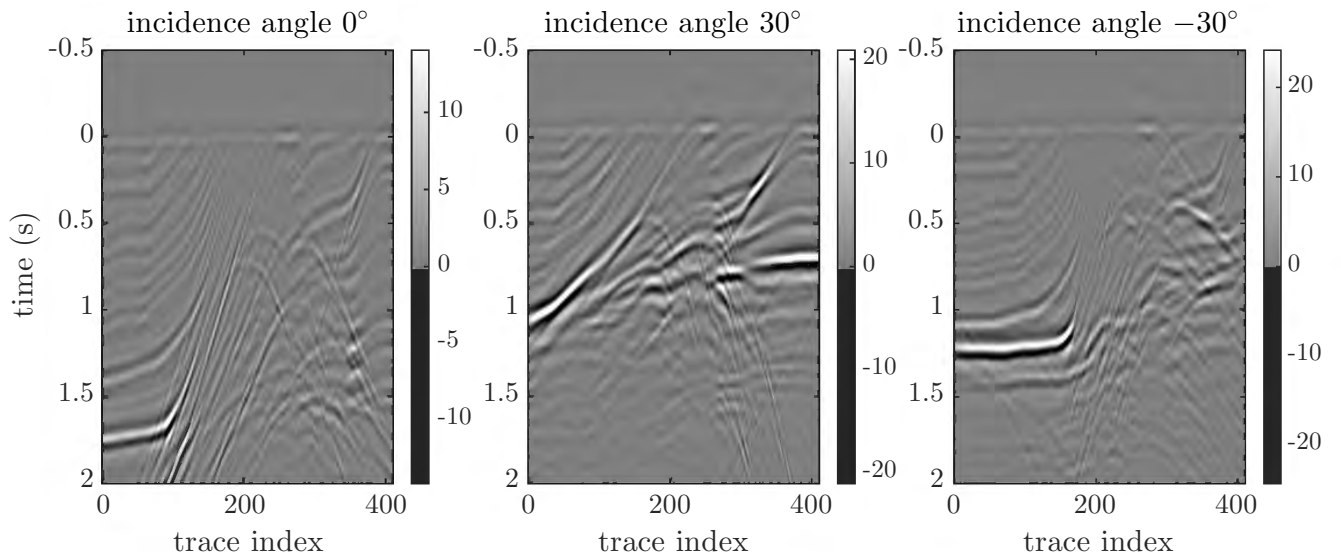


Figure 4.28: Marmousi with absorbing surface example: shot records (direct arrivals removed). Left: plane wave with 0 degrees incidence angle. Center: plane wave with 30 degrees incidence angle. Right: plane wave with -30 degrees incidence angle.

Figure 4.29 shows inversion result after 70 outer iterations of the VP method (top) and true resampled model (bottom). Overall, the Marmousi model is recovered sufficiently accurately. Significant errors in positions of interface (2) and high-velocity layer (3), in the left part of the model are due to low-reflective overburden area (1), whose recovered velocity is slower than exact velocity. Right part of the model with stronger reflectors is recovered much better; for example, high-velocity areas (4) and (5) are positioned correctly.

Figures 4.30 and 4.31 show shot records (direct arrivals removed) of observed data (left column) and data residuals (right column) after 70 iterations for shots with  $0^\circ$  and  $30^\circ$  incidence angles respectively.

Figure 4.32 compares leftmost and central traces of predicted and true data for  $0^\circ$  incidence angle plane wave shot. Figure 4.33 gives similar comparison for  $30^\circ$  incidence angle plane wave shot. These two figures show that data is fitted very well for both offsets.

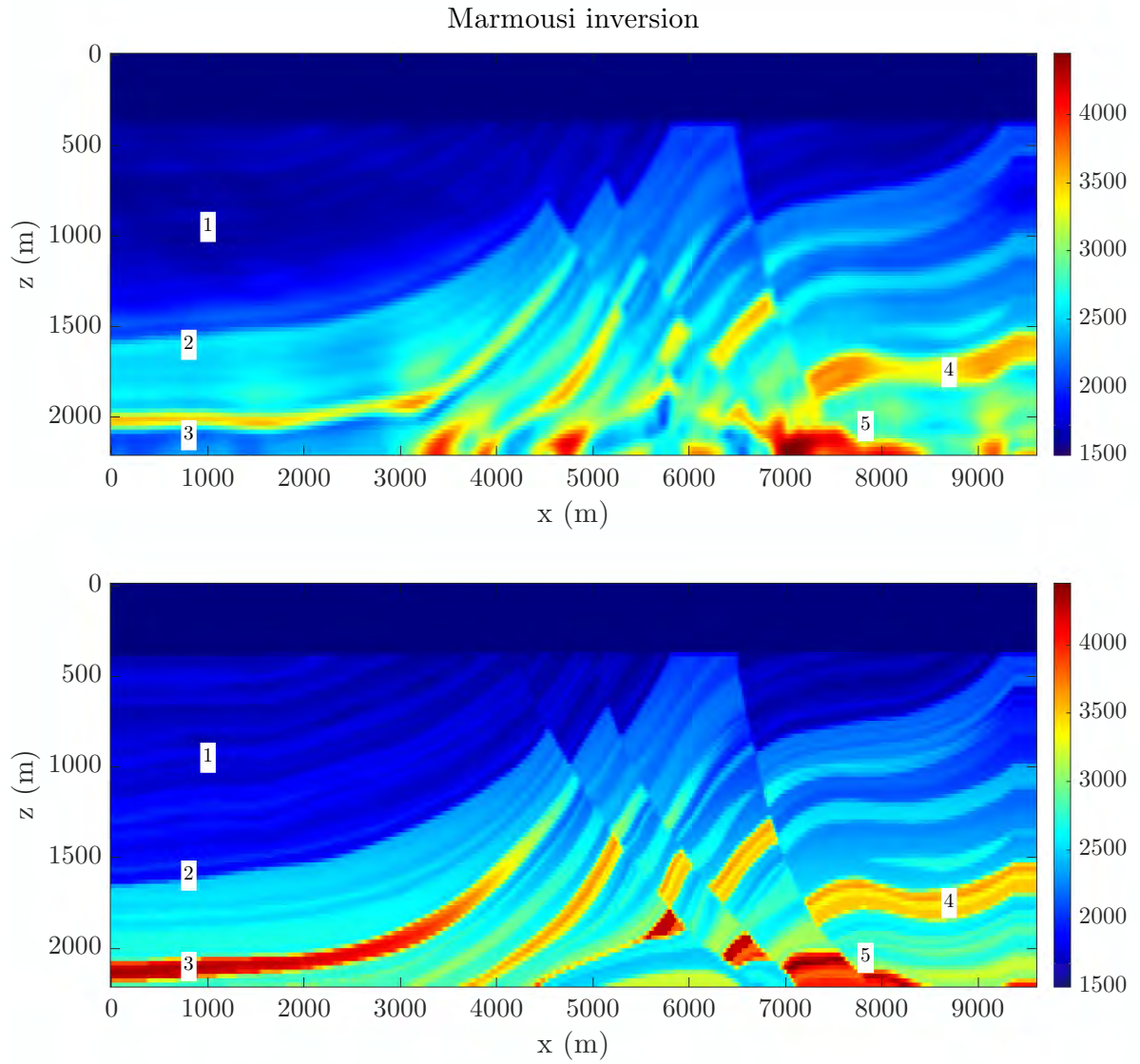


Figure 4.29: Marmousi with absorbing surface example: inversion result. Top: averaged extended model after 70 outer iterations. Bottom: true model resampled to the computational grid.

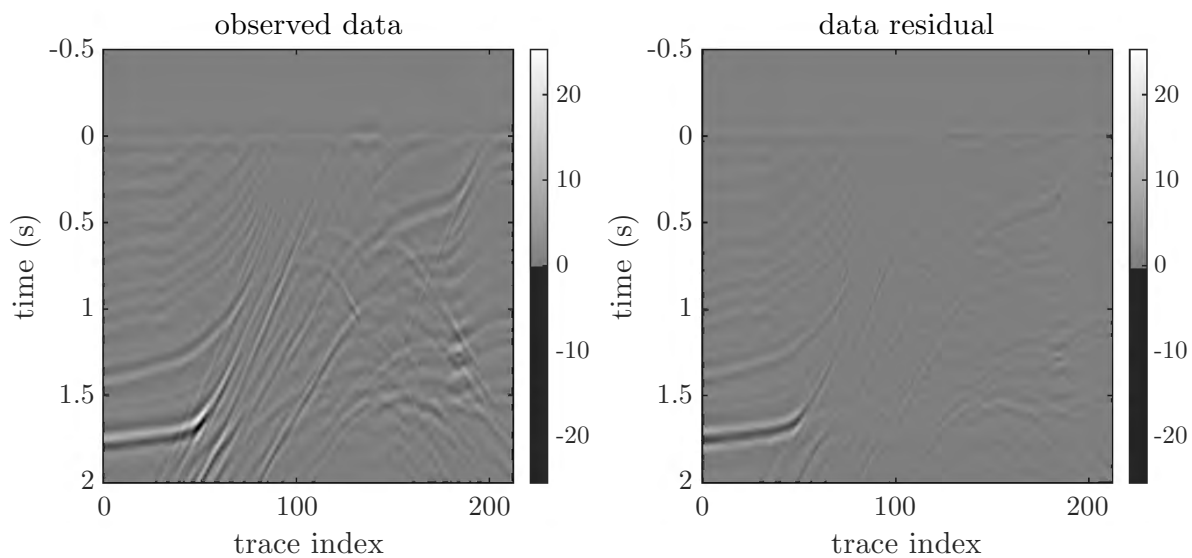


Figure 4.30: Marmoussi with absorbing surface example: data residual after 70 outer iterations of the VP method. Left: shot record for plane wave with 0 degrees incidence angle. Right: data residual.

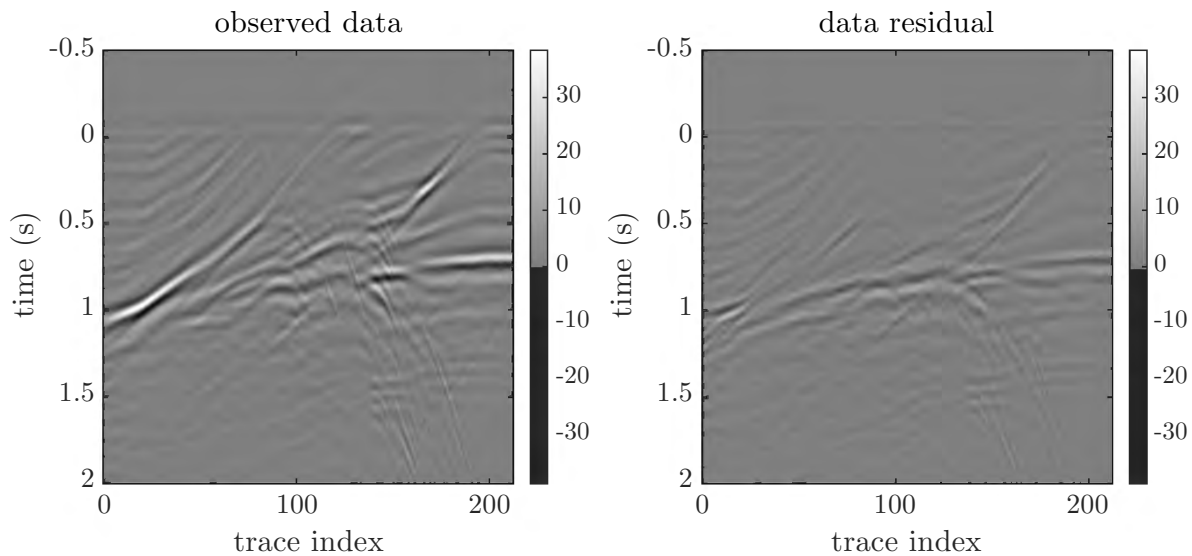


Figure 4.31: Marmoussi with absorbing surface example: data residual after 70 outer iterations of the VP method. Left: shot record for plane wave with 30 degrees incidence angle. Right: data residual.

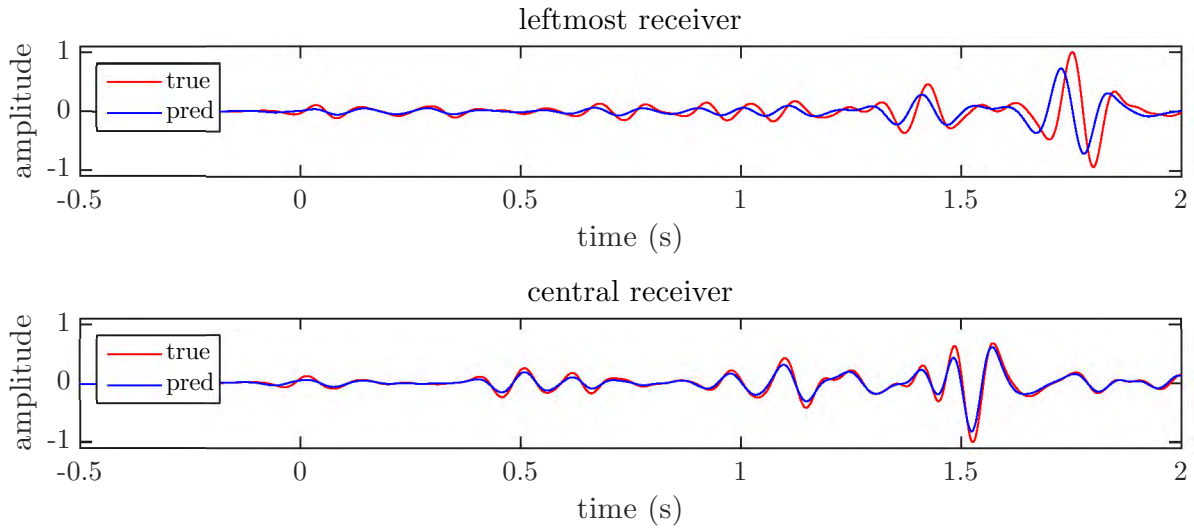


Figure 4.32: Marmousi with absorbing surface example: trace comparison of true data (red) and predicted data after 70 iterations (blue).  $0^\circ$  incidence angle plane wave shot. Top: leftmost receiver. Bottom: central receiver.

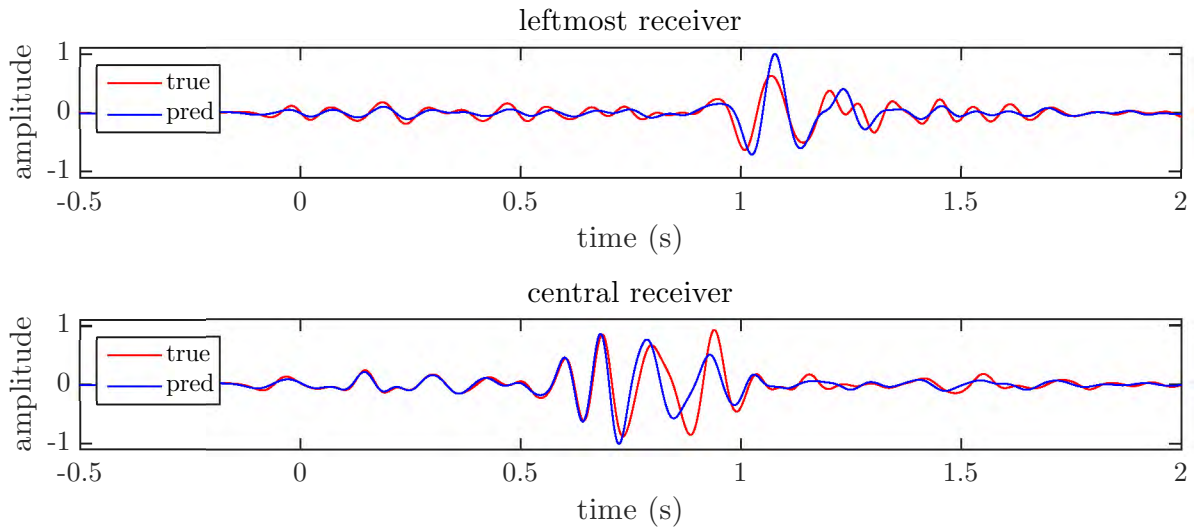


Figure 4.33: Marmousi with absorbing surface example: trace comparison of true data (red) and predicted data after 70 iterations (blue).  $30^\circ$  incidence angle plane wave shot. Top: leftmost receiver. Bottom: central receiver.

In order to improve Marmousi results, layer stripping approach can be used. This approach consists of updating the model in a top-to bottom manner by using, for instance, truncated or damped data. Figure 4.34 shows inversion result for the upper third of Marmousi model and truncated data (1 s) in comparison to the same area of the inverted model from full-model inversion. Significant improvement of the overthrust area in the left of the model is observed.

Figure 4.35 shows the normalized dot product of the negative gradient ( $-\Lambda \nabla \mathcal{J}[m_c]$ ) and the difference of the averaged extended model and the control model ( $\Lambda H(\bar{\mu}) - m_c$ ). See step length limit formula eq:maxstep in Section 4.3. The control model gradient tends to point in the direction of averaged extended model. As the iteration process proceeds, the normalized dot product values decrease, and eventually “randomly oscillate” around zero. This behavior corresponds to stagnation and eventual failure of the descent process.



Top layer Marmousi inversion

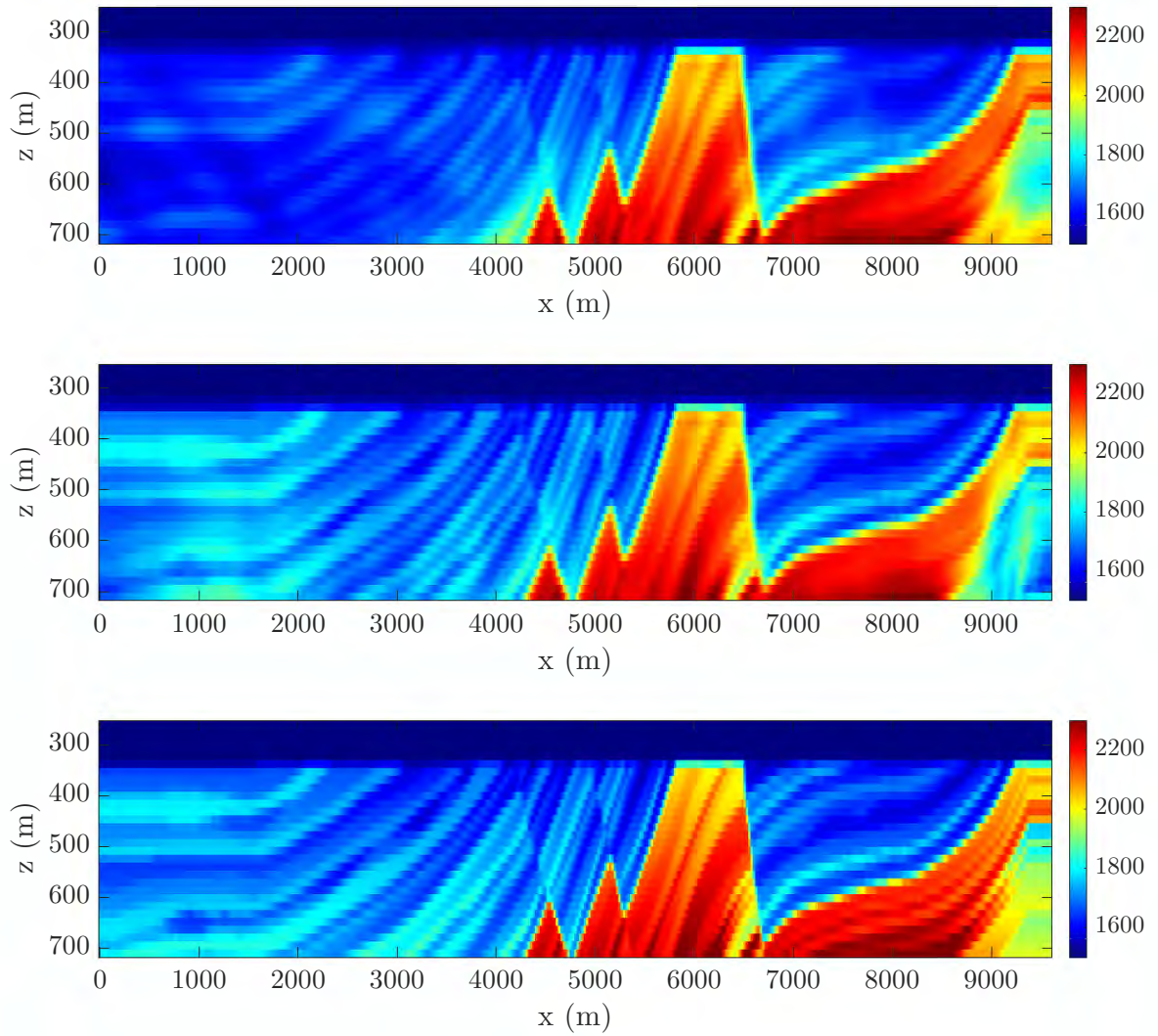


Figure 4.34: Marmousi with absorbing surface example: inversion of the upper part. Top: top layer of recovered model in full-model inversion (same as upper part of recovered model in Figure 4.29). Middle: inversion result for the top layer and truncated data. Bottom: true model top layer.

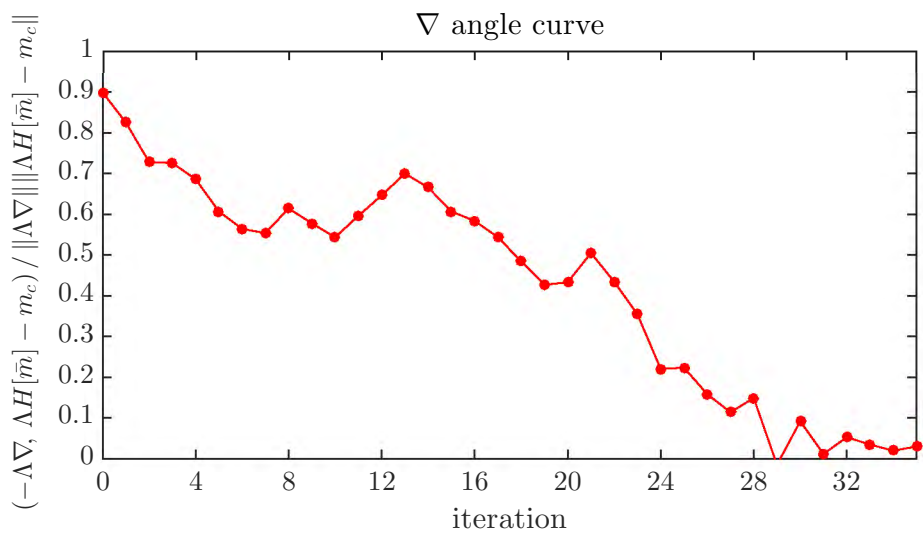


Figure 4.35: Marmousi with absorbing surface example: gradient QC. Normalized dot product of control model gradient and direction from current model to smoothed averaged extended model.

## Marmousi: layer-stripped EFWI + FWI

In this section, I show final inversion result of layer-stripped extended FWI followed by conventional FWI. Extended FWI consisted of three layer-stripping stages: top 1/3 of the domain, traces truncated to 1 second; top 1/2 of the domain, traces truncated to 1.5 seconds; and full domain. Inversion result of each stage was extended downward and used as initial guess for the next stage.

Figure 4.36 shows final recovered model (top) and true model (bottom). Velocities in overburden area (1) and positions of interface (2) and high-velocity layer (3) improved significantly.

Figures 4.37, 4.38, and 4.38 present shot record of observed (left) data and data residuals (right) for  $0^\circ$ ,  $30^\circ$  and  $-30^\circ$  plane wave incidence angle shots respectively.

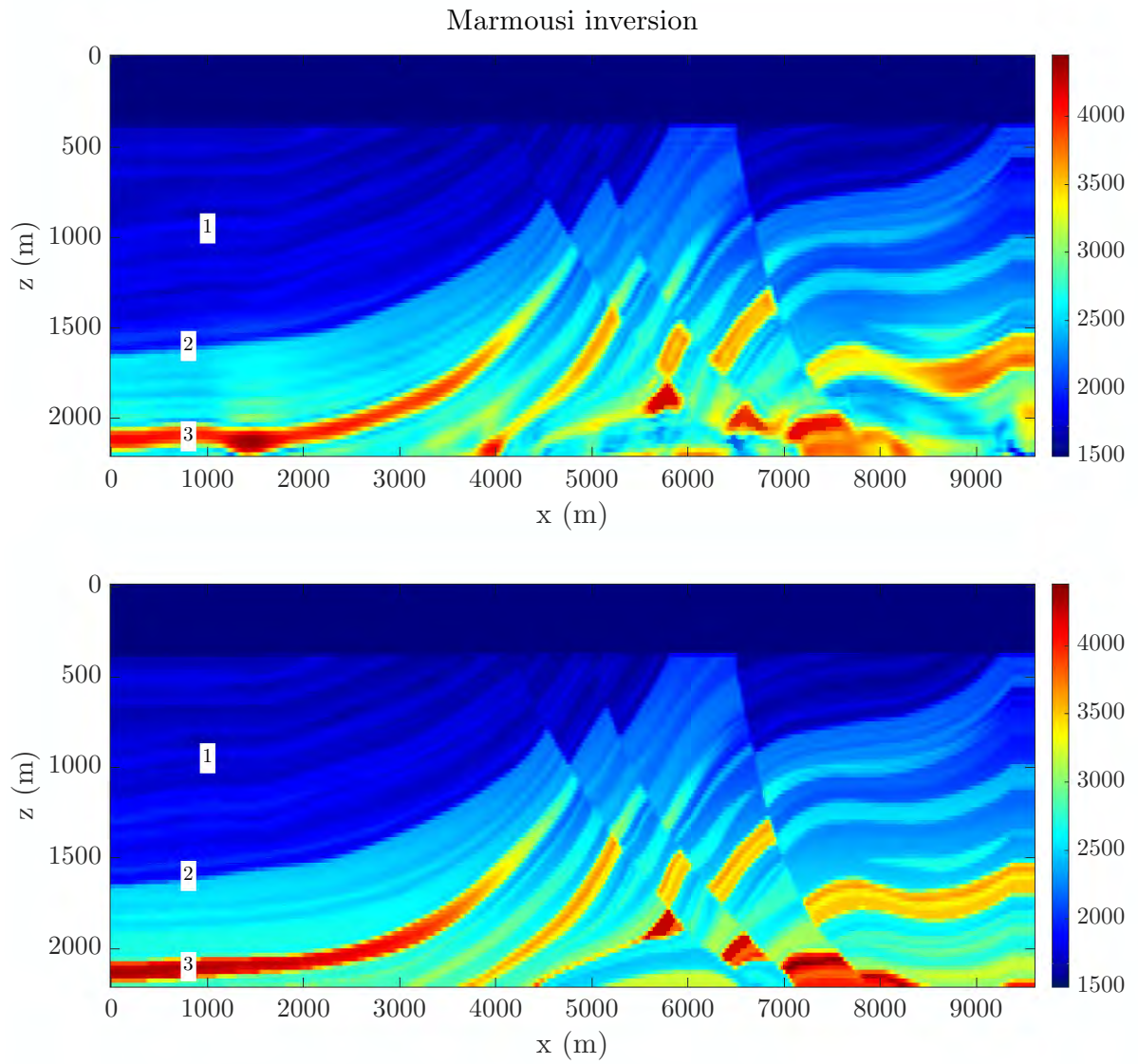


Figure 4.36: Marmousi with absorbing surface example: inversion result. Top: model recovered by layer-stripped EFWI + FWI. Bottom: true model resampled to the computational grid.

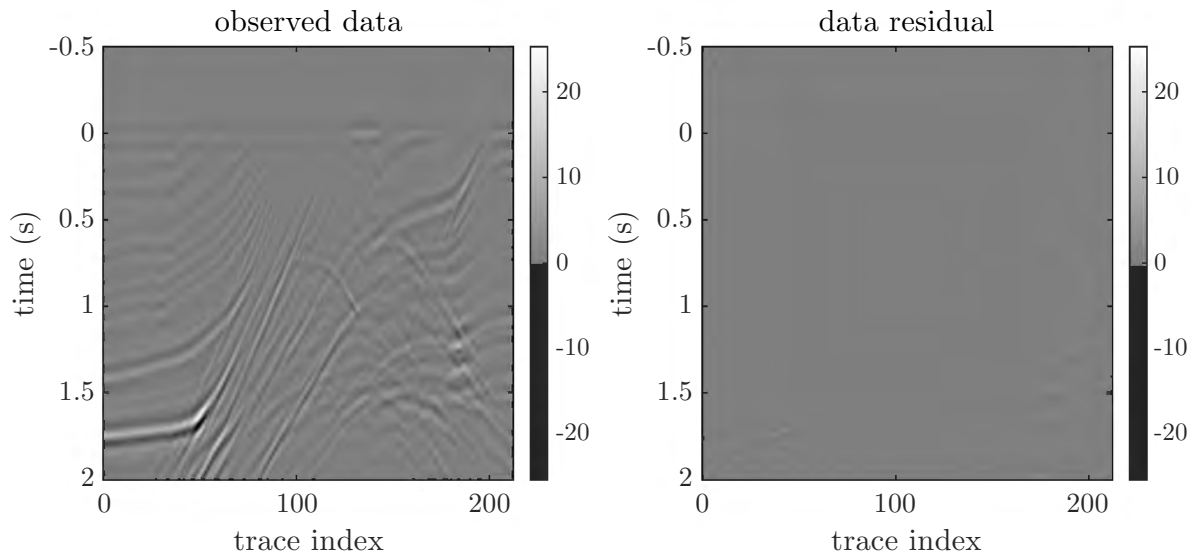


Figure 4.37: Marmousi with absorbing surface example: final data residual. Left: shot record for plane wave with 0 degrees incidence angle. Right: data residual.

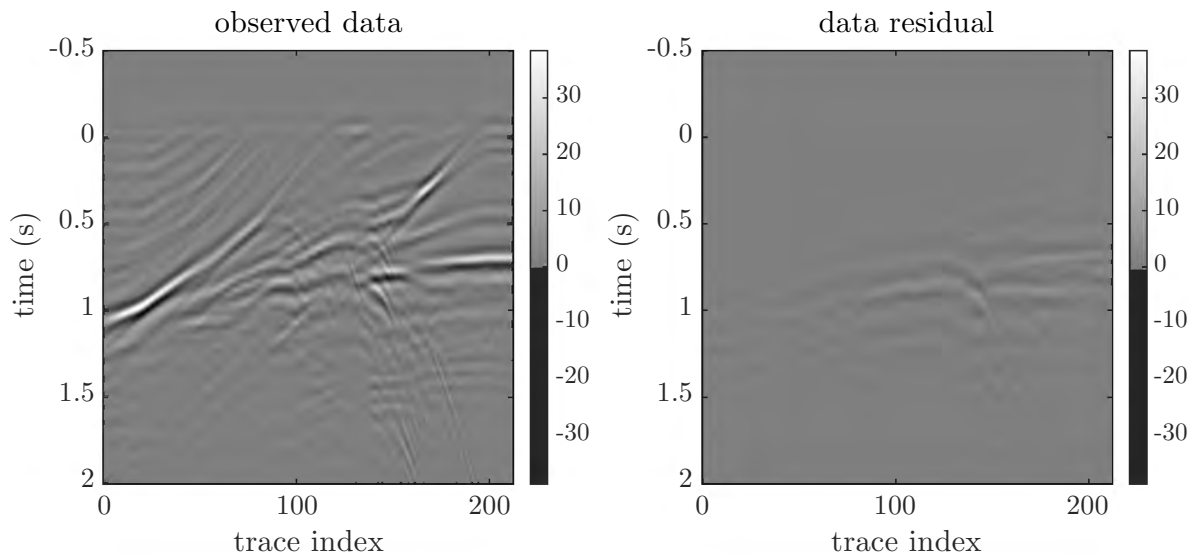


Figure 4.38: Marmousi with absorbing surface example: final data residual. Left: shot record for plane wave with 30 degrees incidence angle. Right: data residual.

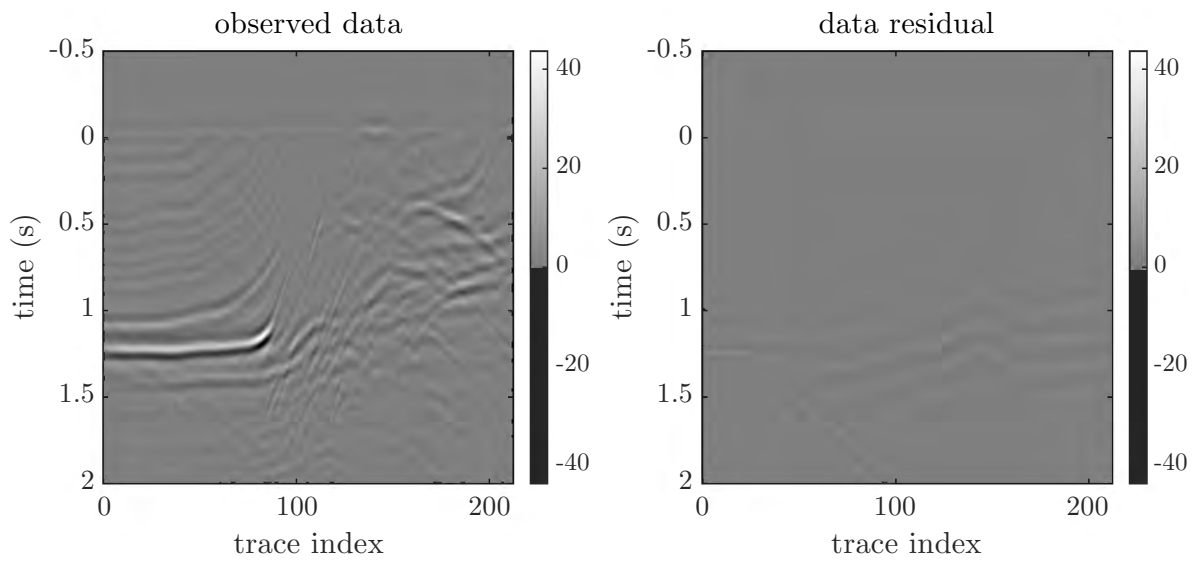


Figure 4.39: Marmousi with absorbing surface example: final data residual. Left: shot record for plane wave with  $-30$  degrees incidence angle. Right: data residual.

Figures 4.40, 4.42 and 4.41 compare leftmost, central and rightmost traces of predicted and true data for  $0^\circ$ ,  $30^\circ$  and  $-30^\circ$  plane wave incidence angle shots respectively.

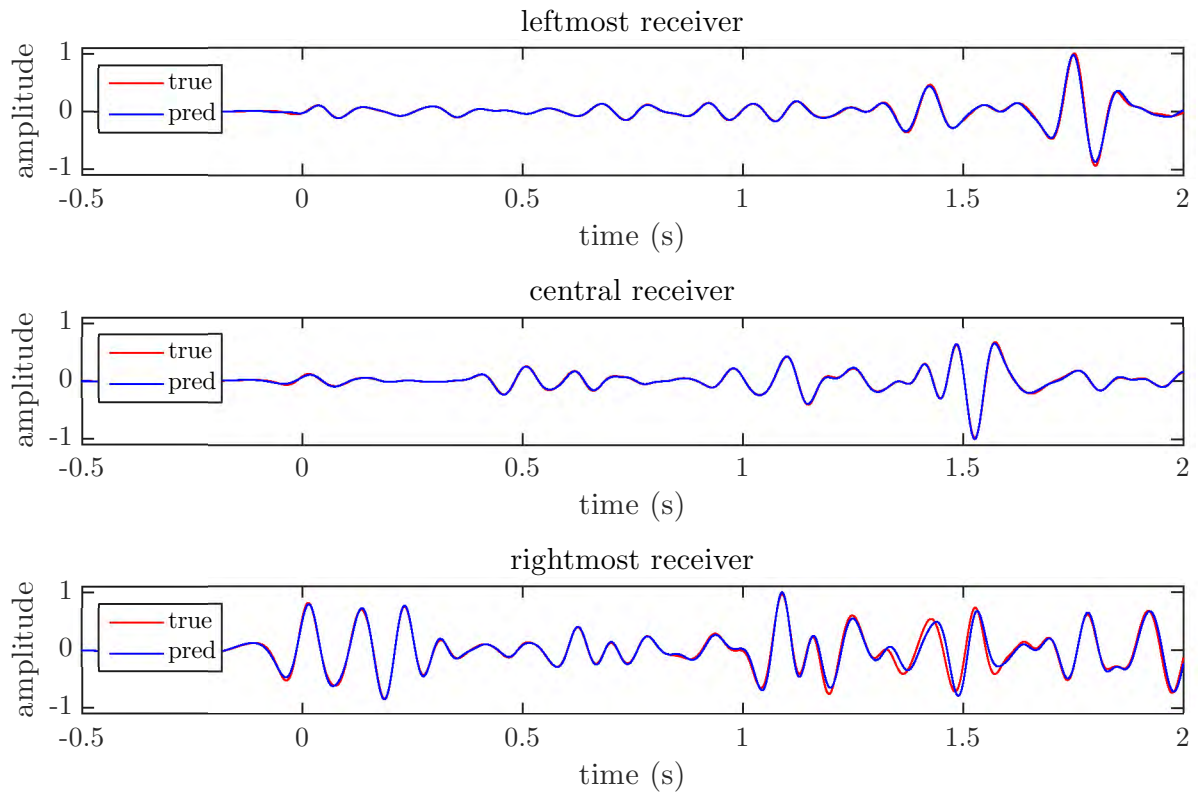


Figure 4.40: Marmousi with absorbing surface example: trace comparison of true data (red) and final predicted data (blue).  $0^\circ$  incidence angle plane wave shot. Top: leftmost receiver. Center: central receiver. Bottom: rightmost receiver.

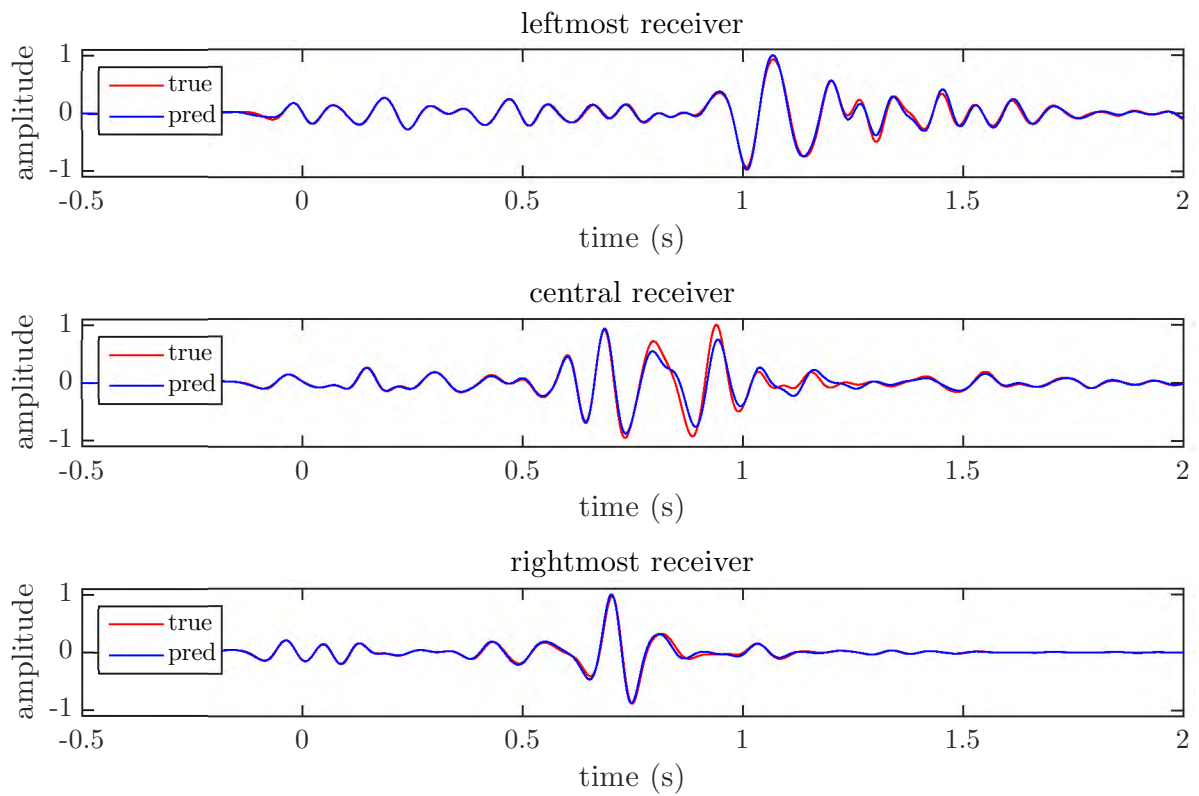


Figure 4.41: Marmousi with absorbing surface example: trace comparison of true data (red) and final predicted data (blue).  $30^\circ$  incidence angle plane wave shot. Top: leftmost receiver. Center: central receiver. Bottom: rightmost receiver.



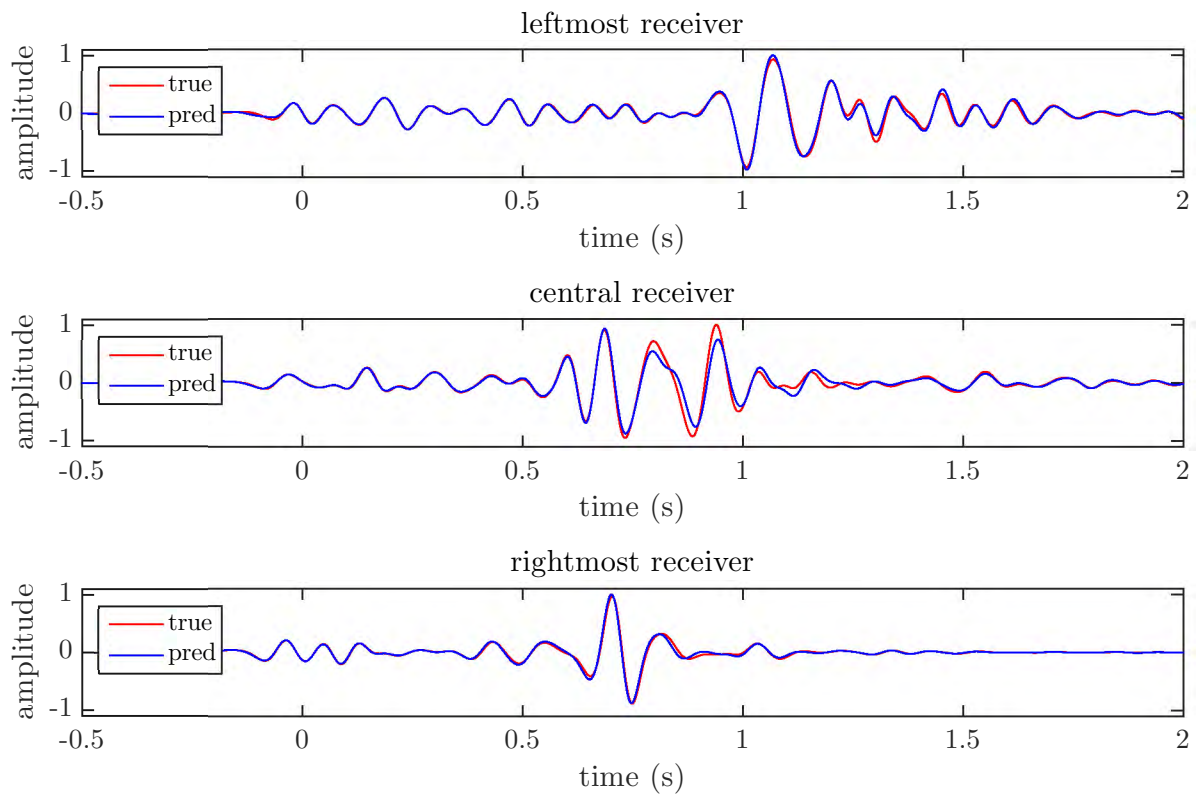


Figure 4.42: Marmousi with absorbing surface example: trace comparison of true data (red) and final predicted data (blue).  $-30^\circ$  incidence angle plane wave shot. Top: leftmost receiver. Center: central receiver. Bottom: rightmost receiver.

## 4.7 Discussion

The examples provided in this chapter possess some of the most important characteristics of a realistic waveform inversion problem, such as absence of low-frequency signal in the data, strong surface multiples (Dome example), complex earth structure (Marmousi example). Conventional FWI, based on local descent methods, cannot recover long- and medium-scale model structure due to numerous local extrema of the objective function.

Numerical tests presented above show that extended waveform inversion coupled with variable projection optimization approach is capable of recovering accurate velocity models and achieving small data misfit while starting from kinematically incorrect initial models. The main drawback of this approach is its high computational cost, due to 1) model extension, 2) necessity to solve inner optimization problem very precisely in order to obtain reliable gradient of the VP method. Certain ad hoc or heuristic strategies (choice of initial guess for the inner problem, choice of step length) used for successful application of the method to the considered examples might not work as well in other problem settings.

# Chapter 5

## Conclusion

This thesis examines a modification of nonlinear full waveform inversion (FWI) method that allows to recover long- and medium-scale earth structure from reflection seismic data. Conventional FWI is based on least squares data misfit optimization approach. Resolution requirements of realistic applications lead to huge model space dimension, hence making local gradient-based methods mandatory for solving the problem numerically. Main characteristic of the reflection seismic data is absence of low-frequency content as well as transmission energy typically present in wide- or full-azimuth acquisition surveys, which results in multiple local extrema of the least squares objective function. This multimodality property makes local descent methods unusable, unless the initial guess is kinematically accurate.

Migration velocity analysis (MVA) uses image-gather flattening concept. An image is an approximation of a reflectivity model for a given (smooth) background and is produced by applying an adjoint of the Born operator to the data. Typically, seis-

mic data is redundant with respect to acquisition parameters, such as shot records, and can be divided into bins. MVA produces separate image for each bin, resulting in a so-called image volume. It then measures discrepancy between images, which is minimized to update the background model. Differential semblance (DS) MVA objective measures image discrepancy using differential operators. MVA approach is widely used in the industry as is not prone to local extrema issue and is capable of recovering long-scale structure of background velocity model. Based on the Born (single-scattering) modeling, MVA does not account for non-linear effects (multiples), and requires scale separation and simple kinematics.

Surface-oriented extended full waveform inversion (EFWI) approach investigated in this thesis combines MVA global convergence properties with FWI data fitting capabilities that take into account non-linear effects, such as multiples. Similarly to MVA, EFWI divides data into subsets. For each subset, EFWI produces a separate model based on non-linear forward modeling, resulting in a model volume. DS is used to measure model volume discrepancy. Sun (2012) introduces low frequency source and control model to parametrize set of extended models satisfying data fitting constraint. Solvability of the least squares optimization problem for impulsive source suggests that this parametrization can be computed robustly by solving conventional FWI problems with full bandwidth source. This control model plays similar role to a background model in the extended Born waveform inversion (Huang, 2016).

I combine EFWI with the variable projection (VP) method: VP functional is a compromise of the conventional data fitting term and DS term, which penalizes extended model non-physicality. The inner optimization problem consists of optimizing over the extended model given a control model; the outer optimization problem con-

sists of optimizing over the control model. In addition, the DS term in the inner problem serves as a regularizer, making model gathers less noisy, which allows to compute the outer gradient more reliably.

In Chapter 2, I present an abstract formulation of the extended modeling algorithm and re-parametrization by a control model. I describe a nested optimization approach (Section 2.2) and provide gradient derivation for inner and outer optimization problems (Section 2.3). For the numerical experiments, I adopt the simplest wave propagation model, constant density acoustics, with plane wave surface extension. Chapter 3 describes concrete formulation of the EFWI for this model.

The examples in Chapter 4 present some important features of a realistic seismic inversion problem: strong surface multiples (Dome with free surface example), complex earth structure (Marmousi example). Numerical tests show that EFWI coupled with variable projection optimization approach is capable of recovering long-scale velocity model and achieving small data misfit starting from kinematically incorrect initial models.

At the same time, model extension approach significantly increases the computational cost of the inversion. For surface-oriented extensions, one inner problem solve is equivalent to conventional FWI in terms of the computational cost. Numerical tests in this work showed that the inner problem has to be solved very precisely in order to produce reasonable control model gradient. Various approaches may be used to improve convergence of the inner problem and reduce the computation cost: second-order methods (l-BFGS, Gauss-Newton, Newton), preconditioners, various practical techniques: layer stripping, frequency descent, different misfit functions (e.g., convolution filter based, envelop residual functionals). However, theoretically, the error

of the outer gradient is not controlled by the inner problem solution accuracy. Mar-mousi experiment, for instance, required much more accurate solution of the inner problem, with many more iterations of the CG method. One of the most important questions left for the further research is: how accurately does the inner problem need to be solved in order to produce reliable outer gradient, and to control outer gradient accuracy? In my work, I used a simple ad hoc “brute force” approach: if the outer gradient does not result in desired/sufficient descent during outer iteration, I resume inner problem optimization process in order to solve the inner problem more accurately.

It is known (Stolk and Symes, 2004; Symes, 2008) that for MVA, image gathers based on surface-oriented data binning may not be flat in case of complex kinematics (substantial multi-pathing), even for correct migration velocity. This is a fundamental limitation of the surface-oriented extensions. In such cases, subsurface extensions may be more appropriate. However, they are much more computationally expensive than surface extensions.

# Appendices

# Appendix A

## Adjoint state method

In this chapter, I provide derivations for continuous first- and second-order Born approximations and adjoints.

### A.1 General derivation

To derive Born and adjoint equations, denote wave operator by  $H : M \times U \rightarrow V$ .

Forward map  $\mathcal{F} : M \rightarrow U$  satisfies

$$H(m, \mathcal{F}[m]) = f.$$

Application of the chain rule to the identity  $DH(m, \mathcal{F}[m]) = 0$  leads to

$$\partial_m H(m, \mathcal{F}[m])\delta m + \partial_u H(m, \mathcal{F}[m])D\mathcal{F}[m]\delta m = 0,$$



where  $\delta m$  is a model perturbation. Linearity of  $H$  with respect to  $u$  simplifies previous equation:

$$\partial_m H(m, \mathcal{F}[m])\delta m + H(m, D\mathcal{F}[m]\delta m) = 0. \quad (\text{A.1})$$

Differentiating equation (A.1) yields

$$\begin{aligned} & \partial_{mm}^2 H(m, \mathcal{F}[m])(m_1, \delta m) + \partial_m H(m, D\mathcal{F}[m]m_1)\delta m \\ & + \partial_m H(m, D\mathcal{F}[m]\delta m)m_1 + H(m, D^2\mathcal{F}[m](m_1, \delta m)) = 0, \end{aligned} \quad (\text{A.2})$$

where  $m_1$  is another model perturbation.

Equation for the first-order Born approximation  $\delta u = D\mathcal{F}[m]\delta m$  follows from equation (A.1):

$$H(m, \delta u) = -\partial_m H(m, \mathcal{F}[m])\delta m. \quad (\text{A.3})$$

Equation for the second-order Born approximation  $\delta^2 u = D^2\mathcal{F}[m](m_1, \delta m)$  follows from equation (A.2):

$$\begin{aligned} H(m, \delta^2 u) = & -\partial_{mm}^2 H(m, \mathcal{F}[m])(m_1, \delta m) \\ & - \partial_m H(m, u_1)\delta m - \partial_m H(m, \delta u)m_1, \end{aligned} \quad (\text{A.4})$$

where another Born approximation  $u_1 = D\mathcal{F}[m]m_1$ .

Taking adjoint of equation (A.1) gives adjoint of  $D\mathcal{F}[m]$ :

$$\begin{aligned} D\mathcal{F}[m]^* \delta u &= -\partial_m H(m, \mathcal{F}[m])^* v, \\ H(m, \cdot)^* v &= \delta u. \end{aligned} \quad (\text{A.5})$$

Taking adjoint of equation (A.2) gives adjoint of  $D^2\mathcal{F}[m]m_1$ :

$$\begin{aligned} (D^2\mathcal{F}[m]m_1)^*\delta u &= -(\partial_{mm}^2 H(m, \mathcal{F}[m])m_1)^*v \\ &\quad - \partial_m H(m, u_1)^*v - D\mathcal{F}[m]^*(\partial_m H(m, \cdot)m_1)^*v, \\ H(m, \cdot)^*v &= \delta u, \end{aligned}$$

which I rewrite by introducing secondary adjoint variable  $v_1$ :

$$\begin{aligned} (D^2\mathcal{F}[m]m_1)^*\delta u &= -(\partial_{mm}^2 H(m, \mathcal{F}[m])m_1)^*v \\ &\quad - \partial_m H(m, u_1)^*v + \partial_m H(m, \mathcal{F}[m])^*v_1, \\ H(m, \cdot)^*v &= \delta u, \\ H(m, \cdot)^*v_1 &= (\partial_m H(m, \cdot)m_1)^*v. \end{aligned} \tag{A.6}$$

Modeling operator  $F : M \rightarrow D$  is usually defined as  $F[m] = S\mathcal{F}[m]$ , where linear operator  $S$  is a sampling operator. In this case, first-order Born approximation  $\delta d$  and second-order Born approximation  $\delta^2 d$  are given by

$$\begin{aligned} \delta d &= S\delta u, \\ \delta^2 d &= S\delta^2 u, \end{aligned}$$

where  $\delta u$  and  $\delta^2 u$  solve (A.3) and (A.4) respectively.

Adjoint of  $DF[m]$  follows directly from (A.5):

$$\begin{aligned} DF[m]^*\delta d &= -\partial_m H(m, \mathcal{F}[m])^*v, \\ H(m, \cdot)^*v &= S^*\delta d. \end{aligned}$$

Adjoint of  $D^2F[m]m_1$  follows directly from (A.6):

$$\begin{aligned}
(D^2F[m]m_1)^*\delta d &= -(\partial_{mm}^2 H(m, \mathcal{F}[m])m_1)^*v \\
&\quad - \partial_m H(m, u_1)^*v + \partial_m H(m, \mathcal{F}[m])^*v_1, \\
H(m, \cdot)^*v &= S^*\delta d, \\
H(m, \cdot)^*v_1 &= (\partial_m H(m, \cdot)m_1)^*v.
\end{aligned}$$

In case when  $(Su)(\mathbf{x}_r, t) = \delta(\mathbf{x} - \mathbf{x}_r) * u(\mathbf{x}, t)$  at discrete receiver positions  $\mathbf{x}_r \in \Sigma_r$ , adjoint of  $S$  is

$$(S^*\delta d)(\mathbf{x}, t) = \sum_{\mathbf{x}_r \in \Sigma_r} \delta d(\mathbf{x}_r, t) \delta(\mathbf{x} - \mathbf{x}_r).$$

## A.2 Constant density acoustics

In case of constant density acoustics, equation (3.1) defines operator  $H$  as

$$H(m, u) = \partial_{tt}^2 u - p(m) \nabla_{\mathbf{x}}^2 u, \quad (\text{A.7})$$

where  $p(m)$  is a function of  $m$ .

For simplicity, I assume  $\Omega = \mathbf{R}^n$ . Space  $U$  is a space of appropriate class of functions with zero initial conditions:  $U = \{u : u(\mathbf{x}, 0) = \partial_t u(\mathbf{x}, 0) = 0\}$ . Spaces  $U$ ,  $V$ , and  $M$  are equipped with  $L^2$  inner products.

For the rest of this section, I denote by  $u$  the solution of the wave equation

$H(m, u) = f$ , i.e.,

$$\begin{aligned}\partial_{tt}^2 u - p(m) \nabla_{\mathbf{x}}^2 u &= f, \\ u(\mathbf{x}, 0) = \partial_t u(\mathbf{x}, 0) &= 0.\end{aligned}$$

Substituting (A.7) into equation (A.3) yields first-order Born approximation:

$$\begin{aligned}\partial_{tt}^2 \delta u - p(m) \nabla_{\mathbf{x}}^2 \delta u &= p'(m) \delta m \nabla_{\mathbf{x}}^2 u, \\ \delta u(\mathbf{x}, 0) = \partial_t \delta u(\mathbf{x}, 0) &= 0.\end{aligned}\tag{A.8}$$

Substituting (A.7) into equation (A.4) yields second-order Born approximation:

$$\begin{aligned}\partial_{tt}^2 \delta^2 u - p(m) \nabla_{\mathbf{x}}^2 \delta^2 u &= p''(m) m_1 \delta m \nabla_{\mathbf{x}}^2 u \\ &+ p'(m) \delta m \nabla_{\mathbf{x}}^2 u_1 + p'(m) m_1 \nabla_{\mathbf{x}}^2 \delta u, \\ \delta^2 u(\mathbf{x}, 0) = \partial_t \delta^2 u(\mathbf{x}, 0) &= 0.\end{aligned}\tag{A.9}$$

Adjoint of  $\partial_m H(m, u)$  follows from the following equality:

$$\begin{aligned}\langle \partial_m H(m, u) \delta m, v \rangle &= \int_0^T dt \int dx (-p'(m) \delta m \nabla_{\mathbf{x}}^2 u) v \\ &= \int dx \delta m (-p'(m) \int_0^T dt (\nabla_{\mathbf{x}}^2 u) v) = \left\langle \delta m, -p'(m) \int_0^T dt (\nabla_{\mathbf{x}}^2 u) v \right\rangle.\end{aligned}$$

Therefore, adjoint of  $\partial_m H(m, u)$  is given by

$$\partial_m H(m, u)^* v = -p'(m) \int_0^T dt (\nabla_{\mathbf{x}}^2 u) v.\tag{A.10}$$

Similarly, adjoint of  $\partial_{mm}^2 H(m, u)m_1$  is given by

$$(\partial_{mm}^2 H(m, u)m_1)^* v = -p''(m)m_1 \int_0^T dt (\nabla_{\mathbf{x}}^2 u)v. \quad (\text{A.11})$$

To derive adjoint of  $H(m, \cdot)$  consider the following equality:

$$\begin{aligned} \langle H(m, \cdot)\delta u, v \rangle &= \int_0^T dt \int dx (\partial_{tt}^2 \delta u - p(m)\nabla_{\mathbf{x}}^2 \delta u) v \\ &= \int_0^T dt \int dx (\partial_{tt}^2 v - \nabla_{\mathbf{x}}^2 (p(m)v)) \delta u + \int dx (\partial_t \delta u v - \delta u \partial_t v)|_{t=T} \\ &= \langle \delta u, \tilde{H}(m, v) \rangle + \int dx p(m) (\partial_t \delta u v - \delta u \partial_t v)|_{t=T}, \end{aligned}$$

Where

$$\tilde{H}(m, v) = \partial_{tt}^2 v - \nabla_{\mathbf{x}}^2 (p(m)v).$$

Therefore, if  $v(\mathbf{x}, T) = \partial_t v(\mathbf{x}, T) = 0$ , then

$$H(m, \cdot)^* v = \tilde{H}(m, v). \quad (\text{A.12})$$

Similarly, adjoint of  $\partial_m H(m, \cdot)m_1$  is given by

$$(\partial_m H(m, \cdot)m_1)^* v = -\nabla_{\mathbf{x}}^2 (p'(m)m_1 v). \quad (\text{A.13})$$

Substituting equations (A.10) and (A.12) into equation (A.5) yields adjoint of

$D\mathcal{F}[m]$ :

$$D\mathcal{F}[m]^* \delta u = p'(m) \int_0^T dt (\nabla_{\mathbf{x}}^2 u) v,$$

where  $v$  solves reverse-time equation:

$$\partial_{tt}^2 v - \nabla_{\mathbf{x}}^2 (p(m)v) = \delta u,$$

$$v(\mathbf{x}, T) = \partial_t v(\mathbf{x}, T) = 0.$$

(A.14)

Substituting equations (A.10)–(A.13) into equation (A.6) yields adjoint of  $D^2\mathcal{F}[m]m_1$ :

$$\begin{aligned} (D^2\mathcal{F}[m]m_1)^* \delta u &= p''(m)m_1 \int_0^T dt (\nabla_{\mathbf{x}}^2 u) v \\ &+ p'(m) \int_0^T dt (\nabla_{\mathbf{x}}^2 u_1) v - p'(m) \int_0^T dt (\nabla_{\mathbf{x}}^2 u) v_1, \end{aligned}$$

where  $u_1$  solves Born equation:

$$\partial_{tt}^2 u_1 - p(m)\nabla_{\mathbf{x}}^2 u_1 = p'(m)m_1 \nabla_{\mathbf{x}}^2 u,$$

$$u_1(\mathbf{x}, 0) = \partial_t u_1(\mathbf{x}, 0) = 0,$$

$v$  solves primary reverse-time equation:

$$\partial_{tt}^2 v - \nabla_{\mathbf{x}}^2 (p(m)v) = \delta u,$$

$$v(\mathbf{x}, T) = \partial_t v(\mathbf{x}, T) = 0,$$

and  $v_1$  solves secondary reverse-time equation:

$$\partial_{tt}^2 v_1 - \nabla_{\mathbf{x}}^2 (p(m)v_1) = -\nabla_{\mathbf{x}}^2 (p'(m)m_1 v),$$

$$v_1(\mathbf{x}, T) = \partial_t v_1(\mathbf{x}, T) = 0.$$

(A.15)

For a model defined by equation 3.2, I have

$$p(m) = \left( \frac{c_{max}^\gamma + c_{min}^\gamma}{2} + \frac{c_{max}^\gamma - c_{min}^\gamma}{2} \tanh m(\mathbf{x}) \right)^{2/\gamma},$$
$$p'(m) = \left( \frac{c_{max}^\gamma + c_{min}^\gamma}{2} + \frac{c_{max}^\gamma - c_{min}^\gamma}{2} \tanh m(\mathbf{x}) \right)^{(2/\gamma)-1} \frac{c_{max}^\gamma - c_{min}^\gamma}{\gamma} (1 - \tanh^2 m(\mathbf{x})).$$

# Appendix B

## Full bandwidth inversion

This appendix shows an inversion exercise on a Dome model (Figure 4.4) and full bandwidth plane wave source. The experiment setup is similar to the Dome with absorbing boundary setup described in Section 4.5. For a full bandwidth wavelet, I use Ormsby (Ryan, 1994) with 0-0-15-22 Hz frequencies.

The exercise consists of a number of independent conventional LS data fitting inversions with different plane wave incidence angles. All inversions start with homogeneous “water model”  $c = 1500$  m/s.

I employ frequency continuation technique (Bunks et al., 1995). The method consists of running multiple inversions with increasing frequency range of the source wavelet and the observed data. Resulting model of each previous inversion is used as an initial guess for the next one.

Optimization problems are solved with Polak-Ribière conjugate gradient (CG)



method. No preprocessing (muting, tapering, etc.) or preconditioning is used in this test example. Number of CG iterations at each frequency continuation stage is limited by 80.

Figures B.1–B.3 show inversion results for plane waves with  $0^\circ$ ,  $10^\circ$ , and  $20^\circ$  incidence angles. For a full bandwidth source, the model is recovered accurately starting from initial guess that does not contain long-scale information.

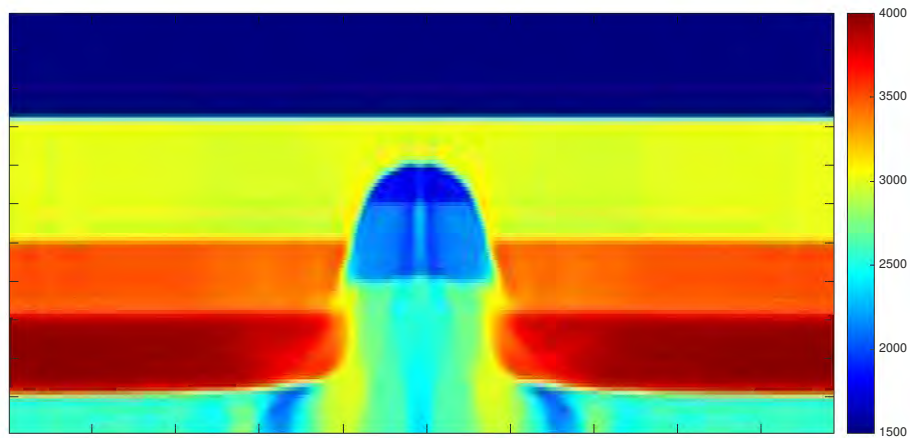


Figure B.1: Dome with absorbing surface example: full bandwidth source. Inversion result for horizontal plane wave source.

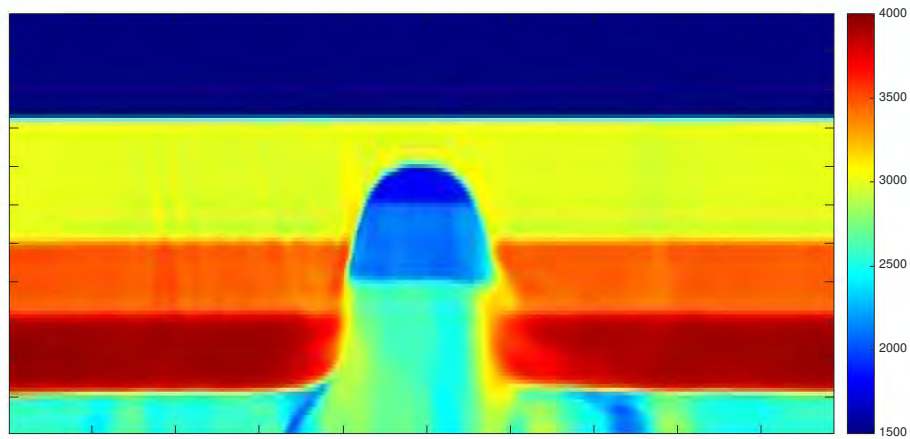


Figure B.2: Dome with absorbing surface example: full bandwidth source. Inversion result for  $10^\circ$  incidence angle plane wave source.

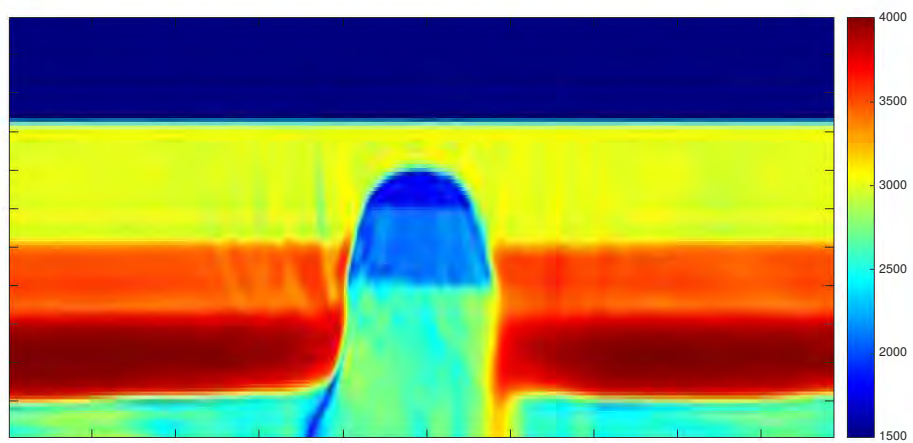


Figure B.3: Dome with absorbing surface example: full bandwidth source. Inversion result for  $20^\circ$  incidence angle plane wave source.

# Appendix C

## Envelope residual operator

Envelope residual operator  $\mathcal{E} : D \rightarrow D$  is defined by

$$\mathcal{E}[d] = d^2 + (Hd)^2, \quad (\text{C.1})$$

where  $H$  denotes Hilbert transform.

Given observed data  $d_o$  and denoting simulated data by  $d_s = F[m]$ , envelope inversion is defined as

$$\min_m J_{\mathcal{E}}[m],$$
$$J_{\mathcal{E}}[m] = \frac{1}{2} \|\mathcal{E}[d_s] - \mathcal{E}[d_o]\|^2.$$

## Gradient computation

Gradient of the objective function  $J_{\mathcal{E}}$  is computed following the adjoint equations (A.14) with new residual

$$\delta u = D\mathcal{E}[d_s]^*(\mathcal{E}[d_s] - \mathcal{E}[d_o]). \quad (\text{C.2})$$

Differentiating equation (C.1) yields

$$D\mathcal{E}[d] \delta d = 2 d \delta d + 2 (Hd)(H\delta d).$$

First term in the above equation is self-adjoint. Adjoint of the second term follows from the skew-adjointness property of the Hilbert transform:

$$\langle (Hd)(H\delta d), v \rangle = \langle \delta d, -H(v Hd) \rangle.$$

Therefore, equation (C.2) becomes

$$\delta u = 2\{d_s(\mathcal{E}[d_s] - \mathcal{E}[d_o]) - H[H[d_s](\mathcal{E}[d_s] - \mathcal{E}[d_o])]\}.$$

# Bibliography

- Almomin, A. and B. Biondi, 2012, Tomographic full waveform inversion: practical and computationally feasible approach: Presented at the 82nd Annual International Meeting, Society of Exploration Geophysicists.
- Aravkin, A., T. van Leeuwen, and F. Herrmann, 2011, Robust full-waveform inversion using the Student's t-distribution: 81st Annual International Meeting, Expanded Abstracts, 2669–2673, Society of Exploration Geophysicists.
- Aster, R. C., B. Borchers, and C. H. Thurber, 2012, Parameter estimation and inverse problems: Elsevier.
- Bamberger, A., G. Chavent, C. Hemon, and P. Lailly, 1982, Inversion of normal incidence seismograms: *Geophysics*, **47**, 757–770.
- Bamberger, A., G. Chavent, and P. Lailly, 1979, About the stability of the inverse problem in 1-D wave equation — application to the interpretation of seismic profiles: *Appl. Math. Opt.*, **5**, 10–47.
- Ben-Hadj-Ali, H., S. Operto, and J. Virieux, 2011, An efficient frequency-domain full waveform inversion method using simultaneous encoded sources: *Geophysics*, **76**, R109–R124.
- Biondi, B. and A. Almomin, 2012, Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis

- sis: Presented at the 82nd Annual International Meeting, Society of Exploration Geophysicists.
- , 2014, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion: *Geophysics*, **79**, WA129–WA140.
- Biondi, B. and P. Sava, 2004, Wave-equation migration velocity analysis - I: Theory, and II: Subsalt imaging examples: *Geophysics*, **52**, 593–623.
- Bourgeois, A., P. Lailly, and R. Versteeg, 1991, The Marmousi model, *in* Versteeg, R. and G. Grau, eds., *The Marmousi Experience*. IFP/Technip.
- Brenders, A. and G. Pratt, 2007a, Efficient waveform tomography for lithospheric imaging: implications for realistic, 2-D acquisition geometries and low frequency data: *Geophysical Journal International*, **168**, 152–170.
- , 2007b, Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model: *Geophysical Journal International*, **168**, 133–151.
- Brenders, A. J., U. Albertin, and J. Mika, 2012, Comparison of 3D time- and frequency-domain waveform inversion: Benefits and insights of a broadband, discrete-frequency strategy: 82nd Annual International Meeting, Expanded Abstracts, 1–5, Society of Exploration Geophysicists.
- Bube, K. P. and T. Nemeth, 2007, Fast line searches for the robust solution of linear systems in the hybrid  $l^1/l^2$  and Huber norms: *Geophysics*, **72**, A13–A17.
- Bunks, C., F. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: *Geophysics*, **60**, 1457–1473.
- Cao, D., W. B. Beydoun, S. C. Singh, and A. Tarantola, 1990, A simultaneous inversion for background velocity and impedance maps: *Geophysics*, **55**, 458–469.
- Chavent, G., 1974, Identification of functional parameters in partial differential equations, *in* Goodson, R. E. and M. Polis, eds., *Identification of parameters in dis-*

- tributed systems, 31–48. ASME.
- Cohen, G., 2001, Higher order numerical methods for transient wave equations: Springer.
- Cruse, E., A. Pica, M. Noble, J. McDonald, and A. Tarantola, 1990, Robust elastic nonlinear waveform inversion: Application to real data: *Geophysics*, **55**, 527–538.
- Fei, W. and P. Williamson, 2010, On the gradient artifacts in migration velocity analysis based on differential semblance optimization: 80th Annual International Meeting, Expanded Abstracts, 4071–4076, Society of Exploration Geophysicists.
- Fichtner, A., 2010, Full seismic waveform modelling and inversion: Springer Berlin Heidelberg.
- Frigo, M. and S. G. Johnson, 2005, The design and implementation of FFTW3: Proceedings of the IEEE, **93**, 216–231. Special issue on “Program Generation, Optimization, and Platform Adaptation”.
- Gao, F., P. Williamson, and R. G. Pratt, 2014, A new objective function for full waveform inversion: Differential semblance in the data domain: 84th Annual International Meeting, Expanded Abstracts, 1178–1183, Society of Exploration Geophysicists.
- Gauthier, O., J. Virieux, and A. Tarantola, 1986, Two-dimensional nonlinear inversion of seismic waveforms: Numerical results: *Geophysics*, **51**, 1387–1403.
- Golub, G. and V. Pereyra, 2003, Separable nonlinear least squares: the variable projection method and its applications: *Inverse problems*, **19**, R1–R26.
- Guittou, A. and E. Diaz, 2012, Attenuating crosstalk noise with simultaneous source full waveform inversion: *Geophysical Prospecting*, **60**, 759–768.
- Guittou, A. and W. W. Symes, 2003, Robust inversion of seismic data using the Huber norm: *Geophysics*, **68**, 1310–1319.
- Gustafsson, B., 2008, High order difference methods for time dependent PDE:

- Springer.
- Ha, W. and C. Shin, 2012, Laplace-domain full-waveform inversion of seismic data lacking low-frequency information: *Geophysics*, **77**, R199–R206.
- Hascoet, L. and V. Pascual, 2013, The Tapenade automatic differentiation tool: principles, model, and specification: *ACM Trans. Math. Softw.*, **39**, 20:1–20:43.
- Huang, Y., 2016, Born waveform inversion in shot coordinate domain: PhD thesis, Rice University, Houston, Texas, USA.
- Huang, Y. and W. W. Symes, 2015, Born waveform inversion via variable projection and shot record model extension: Presented at the 85th Annual International Meeting, Society of Exploration Geophysicists.
- Jannane, M., W. Beydoun, E. Crase, D. Cao, Z. Koren, E. Landa, M. Mendes, A. Pica, M. Noble, G. Roeth, S. Singh, R. Snieder, A. Tarantola, D. Trezeguet, and M. Xie, 1989, Wavelengths of earth structures that can be resolved from seismic reflection data: *Geophysics*, **54**, 906–910.
- Kim, Y., C. Shin, H. Calandra, and D.-J. Min, 2013, An algorithm for 3D acoustic time-Laplace-Fourier-domain hybrid full waveform inversion: *Geophysics*, **78**, R151–R166.
- Kolb, P., F. Collino, and P. Lailly, 1986, Prestack inversion of a 1D medium: *Proceedings of the IEEE*, 498–506.
- Komatitsch, D. and R. Martin, 2007, An unsplit convolutional Perfectly Matched Layer improved at grazing incidence for the seismic wave equation: *Geophysics*, **72**, SM155–SM167.
- Krebs, J. R., J. E. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M.-D. Lacasse, 2009, Fast full-waveform seismic inversion using encoded sources: *Geophysics*, **74**, WCC177–WCC188.
- Lailly, P., 1983, The seismic inverse problem as a sequence of before-stack migrations,



- in* Bednar, J. B., ed., Conference on inverse scattering: theory and applications, 206–220. Society for Industrial and Applied Mathematics.
- Li, Y., B. Biondi, R. Clapp, and D. Nichols, 2014, Wave-equation migration velocity analysis for VTI models: *Geophysics*, **79**, WA59–WA68.
- Lines, L. and S. Treitel, 1984, A review of least-squares inversion and its application to geophysical problems: *Geophys. Prospect.*, **32**, 159–186.
- Lions, J.-L., 1971, Optimal control of systems governed by partial differential equations: Springer Verlag.
- Liu, Y., W. W. Symes, and Z. Li, 2014, Extended reflection waveform inversion via differential semblance optimization: 84th Annual International Meeting, Expanded Abstracts, 1232–1237, Society of Exploration Geophysicists.
- Luo, S. and P. Sava, 2011, A deconvolution-based objective function for wave-equation inversion: 81st Annual International Meeting, Expanded Abstracts, 2788–2792, Society of Exploration Geophysicists.
- Maciel, J. S., J. C. Costa, and J. Schleicher, 2011, Coherence measures in automatic time migration velocity analysis: 81st Annual International Meeting, Expanded Abstracts, 3871–3876, Society of Exploration Geophysicists.
- Mora, P., 1989, Inversion = migration + tomography: *Geophysics*, **54**, 1575–1586.
- Nocedal, J. and S. Wright, 1999, Numerical optimization: Springer Verlag.
- OpenMP, 2017, The OpenMP API specification for parallel programming.
- Peters, B., F. J. Herrmann, and T. van Leeuwen, 2014, Wave-equation based inversion with the penalty method: adjoint-state versus wavefield-reconstruction inversion: Presented at the EAGE Annual Conference Proceedings.
- Pica, A., J. P. Diet, and A. Tarantola, 1990, Nonlinear inversion of seismic reflection data in a laterally invariant medium: *Geophysics*, **55**, 284–292.
- Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient

- of a functional with geophysical applications: *Geophys. J. Int.*, **167**, 495–503.
- Plessix, R.-E., G. Baeten, J. W. de Maag, M. Klaassen, Z. Rujie, and T. Zhifei, 2010, Application of acoustic full waveform inversion to a low-frequency large-offset land data set: 80th Annual International Meeting, Expanded Abstracts, 930–934, Society of Exploration Geophysicists.
- Plessix, R.-E., Y.-H. de Roeck, and G. Chavent, 1998, Waveform inversion of reflection seismic data for kinematic parameters by local optimization: *SIAM J. Sci. Comput.*, **20**, 1033–1052.
- Pratt, G. and A. Brenders, 2004, Frequency-domain waveform tomography using refracted arrivals: Presented at the 66th Annual meeting, European Association of Exploration Geophysicists.
- Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, Part 1: Theory, and verification in a physical scale model: *Geophysics*, **64**, 888–901.
- Ratcliffe, A., C. Win, V. Vinje, G. Conroy, M. Warner, A. Umpleby, I. Stekl, T. Nangoo, and A. Bertrand, 2011, Full waveform inversion: A North Sea OBC case study: 81st Annual International Meeting, Expanded Abstracts, 2384–2388, Society of Exploration Geophysicists.
- Ricker, N., 1940, The form and nature of seismic waves and the structure of seismograms: *Geophysics*, **5**, 348–366.
- Rickett, J. and P. Sava, 2002, Offset and angle-domain common image-point gathers for shot profile migration: *Geophysics*, **67**, 883–889.
- Romero, A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: *Geophysics*, **65**, 426–436.
- Ryan, H., 1994, Ormsby, Klauder, Butterworth — a choice of wavelets: *CSEG Recorder*, **19**, no. 7, 8–9.
- Sacks, P. E. and F. Santosa, 1987, A simple computational scheme for determining the

- sound speed of an acoustic medium from the surface values of its impulse response: SIAM J. Sci. Stat. Comp., **8**, 501–520.
- Santosa, F. and W. W. Symes, 1989, An analysis of least-squares velocity inversion: SEG.
- Scales, J. A., M. L. Smith, and T. L. Fischer, 1991, Global optimization methods for highly nonlinear inverse problems, *in* Cohen, G., L. Halpern, and P. Joly, eds., Mathematical and numerical aspects of wave propagation phenomena, 434–444. Society for Industrial and Applied Mathematics.
- Sen, M. K. and P. L. Stoffa, 1991, Nonlinear one-dimensional seismic waveform inversion using simulated annealing: Geophysics, **56**, 1624–1638.
- Shen, P. and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, VE49–VE59.
- , 2015, Horizontal contraction in image domain for velocity inversion: Geophysics, **80**, R95–R110.
- Shin, C. and Y. H. Cha, 2008, Waveform inversion in the Laplace domain: Geophys. J. Int., **173**, 922–931.
- , 2009, Waveform inversion in the Laplace-Fourier domain: Geophysical Journal International, **177**, 1067–1079.
- Shin, C. and D.-J. Min, 2006, Waveform inversion using a logarithmic wavefield: Geophysics, **71**, R31–R42.
- Sirgue, L., B. Denel, and F. Gao, 2011, Integrating 3D full waveform inversion into depth imaging projects: 81st Annual International Meeting, Expanded Abstracts, 2354–2358, Society of Exploration Geophysicists.
- Sirgue, L. and R.-G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248.
- Stoffa, P. L. and M. K. Sen, 1991, Nonlinear multiparameter optimization using

- genetic algorithms: Inversion of plane-wave seismograms: *Geophysics*, **56**, 1794–1810.
- Stolk, C. C. and W. W. Symes, 2003, Smooth objective functionals for seismic velocity inversion: *Inverse Problems*, **19**, 73–89.
- , 2004, Kinematic artifacts in prestack depth migration: *Geophysics*, **69**, 562–575.
- Sun, D., 2012, A nonlinear differential semblance algorithm for waveform inversion: PhD thesis, Rice University, Houston, Texas, USA.
- Sun, D. and W. W. Symes, 2012, Waveform inversion via nonlinear differential semblance optimization: 82nd Annual International Meeting, Expanded Abstracts, 1–7, Society of Exploration Geophysicists.
- Symes, W. W., 1986a, On the relation between coefficient and boundary values for solutions of Webster’s horn equation: *SIAM J. Math. Anal.*, **17**, 1400–1420.
- , 1986b, Stability and instability results for inverse problems in several-dimensional wave propagation, *in* Glowinski, R. and J.-L. Lions, eds., *Proc. 7th International Conference on Computing Methods in Applied Science and Engineering*. North-Holland.
- , 1990, Velocity inversion: A case study in infinite-dimensional optimization: *Math. Prog.*, **48**, 71–102.
- , 1991a, A differential semblance algorithm for the inverse problem of reflection seismology: *Comput Math Appl*, **22**, 147–178.
- , 1991b, Layered velocity inversion: a model problem from reflection seismology: *SIAM J. Math. Anal.*, **22**, 680–716.
- , 1993, A differential semblance criterion for inversion of multioffset seismic reflection data: *J. Geoph. Res.*, **98**, 2061–2073.
- , 2007, Reverse time migration with optimal checkpointing: *Geophysics*, **72**,

- SM213–SM221.
- , 2008, Migration velocity analysis and waveform inversion: *Geophys. Prospect.*, **56**, 765–790.
- Symes, W. W. and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: *Geophysics*, **56**, 654–663.
- Symes, W. W. and M. Kern, 1994, Inversion of reflection seismograms by differential semblance analysis: algorithm structure and synthetic examples: *Geophys. Prospect.*, **42**, 565–614.
- Symes, W. W. and I. Terentyev, 2009, Subgrid modeling via mass lumping in constant density acoustics: 79th Annual International Meeting, Expanded Abstracts, 2572–2576, Society of Exploration Geophysicists.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266.
- , 2004, *Inverse problem theory and methods for model parameter estimation*: Society for Industrial and Applied Mathematics.
- Tarantola, A. and B. Valette, 1982, Generalized nonlinear inverse problems solved using the least squares criterion: *Rev. Geophys.*, **20**, 219–232.
- van Leeuwen, T., A. Aravkin, H. Calandra, and F. J. Herrmann, 2013, In which domain should we measure the misfit for robust full waveform inversion?: Presented at the EAGE Annual Conference Proceedings.
- van Leeuwen, T. and F. J. Herrmann, 2013, Mitigating local minima in full-waveform inversion by expanding the search space: *Geophys. J. Int.*, **195**, 661–667.
- van Leeuwen, T. and W. A. Mulder, 2009, A variable projection method for waveform inversion: Presented at the EAGE Annual Conference Proceedings. Expanded Abstract.
- , 2010, A correlation-based misfit criterion for wave-equation traveltime tomog-

- raphy: *Geophys. J. Int.*, **182**, 1383–1394.
- Versteeg, R. and G. Grau, eds. 1991, *The Marmousi experience: proceedings of the EAEG workshop on practical aspects of inversion: IFP/Technip.*
- Vigh, D. and E. W. Starr, 2008, Comparisons for waveform inversion, time domain or frequency domain?: 78th Annual International Meeting, Expanded Abstracts, 1890–1894, Society of Exploration Geophysicists.
- Vigh, D., W. Starr, J. Kapoor, and H. Li, 2010, 3D full waveform inversion on a Gulf of Mexico WAZ data set: 81st Annual International Meeting, Expanded Abstracts, 957–961, Society of Exploration Geophysicists.
- Virieux, J. and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, **74**, WCC1–WCC26.
- Warner, M. and L. Guasch, 2014, Adaptive waveform inversion: Theory: 84th Annual International Meeting, Expanded Abstracts, 1089–1093, Society of Exploration Geophysicists.
- Warner, M., A. Ratcliffe, T. Nangoo, J. Morgan, A. Umpleby, N. Shah, V. Vinje, I. Stekl, L. Guasch, C. Win, G. Conroy, and A. Bertrand, 2013, Anisotropic 3D full-waveform inversion: *Geophysics*, **78**, R59–R80.
- Weibull, W. W. and B. Arntsen, 2014a, Anisotropic migration velocity analysis using reverse-time migration: *Geophysics*, **79**, R13–R25.
- , 2014b, Reverse-time demigration using the extended-imaging condition: *Geophysics*, **79**, WA97–WA105.
- Williamson, P. R., 1991, A guide to the limits of resolution imposed by scattering in ray tomography: *Geophysics*, **56**, 202–207.
- Xu, S., D. Wang, F. Chen, Y. Zhang, and G. Lambare, 2012, Full waveform inversion for reflected seismic data: Presented at the EAGE Annual Conference Proceedings.