EFFICIENT SEISMIC DEPTH IMAGING AND FULL-WAVEFORM INVERSION VIA GENERALIZED MULTISCALE FINITE ELEMENT

A Dissertation

by

YONGCHAE CHO

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Richard L. Gibson, Jr.
Committee Members,	Mark Everett
	Robert Reece
	Yalchin Efendiev
Head of Department,	Mike Pope

May 2019

Major Subject: Geophysics

ABSTRACT

Reverse-time migration (RTM) and full-waveform inversion (FWI) are widely used because they are able to recover complex geological structures. However, these wave-equation based imaging techniques also have a drawback, as they require significant computational cost. In both methods, wave modeling accounts for the largest part of the computing cost for calculating forward- and backward-propagated wavefields before constructing an imaging condition or a model update term. For this reason, I applied a model reduction technique, the generalized multiscale finite element method (GMsFEM), which solves local spectral problems on a fine grid for fast simulation of wave propagation on a coarser grid. This approach can enhance the speed of computation without sacrificing accuracy by utilizing coarser grids for lower frequency waves. In the proposed method, one can control the size of the coarse grid and level of heterogeneity of the wave solutions to tune the trade-off between speedup and accuracy. As I increase the expected level of complexity of the wave solutions, the GMsFEM wave modeling can capture more detailed features of waves by applying a finer coarse grid and a larger number of basis functions. After computing the forward- and backward-wavefield on the coarse grid, the coarse-scale solutions are projected onto the original fine grid. Therefore, although wave solutions are computed on a coarse grid, it still provides the images for RTM and FWI without reducing the image resolution by projecting coarse wave solutions to the fine grid. In the multiscale finite element approach, one can apply flexible wave modeling parameters (i.e., grid size, number of basis functions) according to the target frequency components, which makes the method an attractive tool for the practical applications of the RTM and FWI. I demonstrated the multiscale FWI using the BP and Marmousi-2 synthetic model. In addition, I show FWI examples of the field data obtained in the Gulf of Mexico region. In the field data examples, I demonstrate that applying the proposed multiscale RTM and FWI with a relatively small number of basis functions can quickly construct a macro velocity model using low frequency. I also propose a strategy to maximize the efficiency of the multiscale FWI by utilizing frequency-adaptive multiscale basis functions based on the target frequency group.

DEDICATION

To my family.

ACKNOWLEDGMENTS

It was a quite long journey to arrive at this point for me, and I would not be able to achieve all these research outcomes without my family and advisor's support. I would like to express my deepest appreciation to Dr. Richard Gibson, who supported me financially and guided me academically. I also like thank Dr. Richard Gibson for giving me an opportunity to get involved in a research project of applying multiscale method, which was challenging for me at the early stage of my school life as a Ph.D. student. However, thanks to all of the generous support from my advisors, finally, I could make a small contribution to the science world. Without his instructions and advice from the colleagues in the Department of Mathematics at Texas A&M University, it would not have been possible. Also, Dr. Richard Gibson helped me to get exposed to multiple research topics such as wave modeling in fractured media and Markov-chain Monte Carlo method, which helped to broaden my perspectives in the geophysical study. I would also like to thank Crisman/Berg-Hughes Center for Petroleum and Sedimentary Studies for financially supporting me.

I am grateful to all the committee members: Dr. Mark Everett and Dr. Robert Reece in Department of Geology & Geophysics, and Dr. Yalchin Efendiev in Department of Mathematics in Texas A&M University, for reviewing this dissertation and providing constructive comments to improve the quality of the dissertation. I would also like to thank my former academic advisor Dr. Changsoo Shin at Seoul National University for instructing all the background theory of seismic depth imaging and full-waveform inversion, which played an important role in making a progress in the dissertation researches for the last four years. Lastly, I would like to thank all the staffs in Texas A&M High-Performance Research Computing Center for helping me to enhance the efficiency of the source code and providing valuable comments for implementing parallel computing.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Richard L. Gibson, Jr., Mark Everett, and Robert Reece of the Department of Geology & Geophysics and Professor Yalchin Efendiev of the Department of Mathematics.

The data used for demonstrating the efficiency of the proposed algorithm in Chapter 4 was provided by the ION Geophysical Corporation. All the work conducted for the dissertation was completed by the student independently.

All the results of seismic depth imaging (reverse-time migration) and full-waveform inversion are created by using the self-developed program. For the implementation of the solver of Helmholtz equation, the `Pardiso` solver from the Intel Math Kernel Library (MKL) is used, and `Spare BLAS Level 3` is used for the sparse matrix multiplication. I used `Python` `matplotlib` & `numpy`, `Seismic Unix`, and `OmniGraffle` for plotting the graphs and figures. The dissertation is written by using LaTeX.

Funding Sources

Graduate study was supported by the Crisman & Berg-Hughes Center for Petroleum and Sedimentary Systems studies at Texas A&M University. Also, part of the research was supported by the Energy Institute of Texas A&M University. The study was partially supported by the SEG-Anadarko scholarship.

NOMENCLATURE

$\left\| \quad \right\|$	Absolute value
$\left\| \quad \right\|$	L_2- norm
$\odot$	Element-wise multiplication
∂	The first partial derivative
$*$	Complex conjugate
$\mathbf{A}^{\mathrm{T}}$	Transpose of matrix $\mathbf{A}$
$\mathbf{I}$	Identity matrix
$\mathbf{J}$	Jacobian matrix
$\mathbf{H}$	Full Hessian matrix
$\mathbf{H}_{\alpha}$	Approximated Hessian matrix
$\mathbf{H}_{p}$	Pseudo Hessian matrix
$\mathbf{S}$	Impedance matrix
$\mathbf{M}$, $\mathbf{K}$	Mass and stiffness matrix
$\mathbf{C}$	Damping matrix
$\mathbf{R}$	Projection matrix
$\mathbf{u}$, $\mathbf{d}$	Modeled and observed data (pressure field)
$\mathbf{f}$	Point source vector
$\mathbf{v}$	Virtual source vector
$\mathbf{f_v}$	Virtual source matrix
$\mathbf{r}$	Residual vector
f_{max}	Maximum frequency
h, H	Fine-scale and coarse-scale mesh
K_i, ∂K_i	Coarse neighborhood and its edge at i coarse node
L_i	The number of multiscale basis functions

n_s, n_r	The number of source and receiver
i, j, k	Iteration index
α	Step length for model update
ρ, v	Density and velocity of medium
ψ	Partition of unity
λ, φ	Eigenvalue (or eigen-frequency) and eigenfunctions
ω	Angular frequency
Φ_i	Multisacle basis functions at ith coarse neighborhood
Ω, $\partial\Omega$	The whole computation domain and its boundary
$\mathfrak{R}$	Real part of the complex number
diag	Taking diagonal component of a matrix
FDM	Finite Difference Method
CG FEM	Continuous-Galerkin Finite Element Method
DG FEM	Discontinuous-Galerkin Finite Element Method
GMsFEM	Generalized Multiscale Finite Element Method
RTM	Reverse-Time Migration
FWI	Full-Waveform Inversion
PML	Perfectly Matched Layer
LU	Lower upper decomposition

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION AND LITERATURE REVIEW

1.1 Research motivation

When we build a seismic image in the depth domain, constructing an accurate velocity model is critical to resolve the complicated earth structures and to determine the correct location of each stratum. A variety of research has been conducted to improve the velocity-building schemes such as tomography or full-waveform inversion. Tomography (Chiu et al., 1986; Zhang and Toksöz, 1998; Murphy and Gray, 1999) has been broadly used to infer the subsurface velocity information by using the travel-time of the seismic wave. However, in the tomography method, each key reflection (or refraction) event needs to be interpreted manually. This is a non-trivial task when we handle field seismic data, since multiple seismic events are associated each other and it is not easy to distinguish a specific seismic event due to complicated constructive or destructive interference among multiple seismic events. In addition, if the interpreted travel-time information is not accurate, we might not be able to expect a good velocity model.

As an alternative, wave-equation based seismic imaging and waveform inversion methods have drawn widespread attention by resolving complicated subsurface structure (Baysal et al., 1983; McMechan, 1983; Whitmore, 1983; Biondi and Shan, 2002; Symes, 2007) by overcoming the limits of asymptotic ray theory. The full-waveform inversion (FWI) and reverse-time migration (RTM) have been greatly advanced by the enhancement of the back-propagation theory (Lailly and Bednar, 1983; Tarantola, 1984) for the practical implementation. However, it still requires considerable computational cost when dealing with large data volumes.

In the wave-equation based imaging method, the wave simulation accounts for the largest part of the computing cost for calculating forward- and backward-propagated wavefields before applying an imaging condition or a model update term. Therefore, by accelerating the wave modeling, we could expect the largest reduction of the computational cost. Thus, recent work has developed methods to accelerate the wave simulation. For example, Fomel et al. (2013) employed the

low-rank approximation of a wavenumber matrix for a wavefield extrapolation. Another method is proposed by Nunes and Minkoff (2014). They obtained the wave solution rapidly through subgrid upscaling which makes the coarse grid represent the fine-scale heterogeneity.

In this research, to accelerate a waveform inversion and seismic imaging, I applied another model reduction technique; Generalized Multiscale Finite Elements Methods (GMsFEM). The idea of GMsFEM (Efendiev et al., 2011; Chung et al., 2014; Gibson and Fu, 2015) is somewhat similar to Nunes and Minkoff (2014)'s work in that it utilizes multiscale grids; however, the key difference is that GMsFEM directly computes the basis functions including fine-scale heterogeneity without performing a subgrid upscaling or a model homogenization. In GMsFEM, the basis functions are computed only once for a specific model, and all simulations are computed on the coarse grid. These basis functions incorporate the most dominant wave modes computed from the local spectral problems, so utilizing them for the simulation on a coarse grid can greatly accelerate computation by efficiently reducing the number of unknowns. This advantage manifests itself when we simulate multiple shot and receiver sets such as the RTM or FWI cases.

There exist two different domains for the implementation of the wave-equation based imaging tools: time- and frequency-domain. The time-domain is more widely used than the frequency domain due to its easiness of implementation and relatively light computing memory consumption. However, working the frequency-domain also has advantages in that wavefields can be separated into independent frequency components, which enable us to accelerate the computation by using different wave simulation parameters for different frequencies. Most conventional numerical methods apply the same spatial grid for the entire target frequency band, which might be oversampled for the modeling of low-frequency waves.

Therefore, I propose a method to improve the efficiency of the wave modeling by combining the frequency-domain wave modeling and the GMsFEM with flexible modeling parameters (i.e., different sizes of a spatial grid for different frequency bands). For example, when combining frequency-domain wave simulation with GMsFEM, I can assign a different number of basis functions for different frequencies to tune the trade-off between accuracy and speed. In other words, in

coarse grid simulation, I used a smaller number of basis functions to model low-frequency waves, while I need to use a greater number of basis functions to capture all the detailed change of high-frequency waves. Another advantage of utilizing the frequency-domain is that the dispersion error can be suppressed in the frequency-domain with less effort. For instance, Hustedt et al. (2004) proposed an extension of the mixed-grid method for accurate wave simulation. Chen (2012) introduced an average-derivative optimal scheme to utilize a rectangular mesh with different grid spacing. Another method, known as the compact high-order method (Turkel et al., 2013), is proposed for an accurate and efficient solution of the Helmholtz equation. The compact high-order approach can realize sixth-order accuracy by maintaining the sparsity of the numerical operator (second-order finite difference method). In addition to the easiness of dealing with the dispersion condition, the frequency-domain has an advantage in that I can incorporate high-frequency components without considering a reduced time step.

In the frequency-domain wave equation, a linear system can be formulated through a numerical discretization of the given equation. There are two different methods for solving such a system: a direct solver and an iterative solver. The direct solver is more accurate than the iterative solver and can solve multiple sources (right-hand side, RHS) at one-time matrix inversion. Nevertheless, it requires a significant computational cost, which becomes more critical when solving a 3-dimensional or multi-parameter problems. Nowadays, however, recent advances in direct solvers can mitigate the computational burden by applying low-rank simplifications (Wang et al., 2011; Ghysels et al., 2016) which reduces the cost of LU decomposition. The iterative solver is feasible even in a machine with relatively small memory; however, it is sensitive to the preconditioner (Plessix, 2007). In this research, I will focus on the method to reduce the computational cost of wave modeling for the acceleration of the seismic imaging and waveform inversion using the direct solver (Parallel direct sparse solver, PARDISO) (Lawson et al., 1979; Dongarra et al., 1988; Dodson et al., 1991).

1.2 Literature review

The main efforts of inferring subsurface information from the quantification of the physical properties start from using travel-time information (Oldham, 1906; Rogers, 1914). The initial form

of utilizing seismic wave modeling is to estimate the differential seismograms estimated through the Born approximation, which could provide macro-scale upper-mantle tomography (Gilbert and Dziewonski, 1975; Woodhouse and Dziewonski, 1984). Then, many techniques of exploration seismology for developing the high-resolution seismic images have arisen as a consequence of designing dense and multifold seismic acquisition system. However, recovering the earth structure without having a good prior information is still a challenging problem.

Hence, many scientists used a two-step workflow: first, build a macro property model, and then apply the amplitude projection by using various types of migration techniques (Claerbout and Doherty, 1972; Gazdag, 1978; Stolt, 1978; Biondi and Symes, 2004). This approach might be useful when dealing with relatively simple geological structures; however, as the demand for seismic imaging with complicated earth structure such as salt, shale, volcanic diapirs, thrust belts, karst, or foothills has been increased, Lailly and Bednar (1983) and Tarantola (1984) proposed a local optimization problem, which aimed to minimize the misfit between the modeled and observed seismic data combined with the back-propagation method. The back-propagation algorithm with modern enhancement of the computing resources make both RTM and FWI feasible. They also showed that the perturbation model which is calculated from the first iteration of the FWI is similar to the RTM results. The only difference is that the RTM uses the time-reversed data for the source of back-propagation, while the FWI back-propagates the data misfit.

Unlike the RTM case, FWI requires multiple iterations to obtain the final results, and it is a more challenging problem due to 1) local minima, 2) sensitivity to the initial guess, and 3) expensive computational cost. As a solution of local minimum and sensitivity problems, Shin and Cha (2008) proposed the Laplace-domain FWI, which uses a zero frequency component by fixing the real part of complex frequency to zero. Laplace FWI uses a damped wavefield which dramatically reduces the possibility of converging to the local minimum. Given that the Laplace-domain FWI can provide a robust initial velocity model, Shin and Cha (2009) proposed a FWI which combines the Laplace- and frequency-domain termed the Laplace-Fourier-domain FWI to improve the resolution of the FWI result. Ha et al. (2010) analyzed the difference of Laplace- and

frequency-domain FWI by comparing two different approaches. For more practical application of the Laplace-Fourier FWI, Shin et al. (2010) proposed a strategy to utilize a sequentially ordered frequency to acquire high-resolution velocity model. For practical application of the Laplace-Fourier-domain FWI, Koo et al. (2011) introduced a method for source estimation and direct wave reconstruction, and Cho et al. (2016) showed an application of Laplace-Fourier FWI to the field data which is obtained in the deep-sea area with limited offset.

As both techniques, RTM and FWI, require accurate simulation of the seismic waves for a successful implementation, the techniques of full-waveform (or forward) modeling for solving a partial derivative equation numerically have been actively studied. The methods that are used for wave modeling varies according to the discretization schemes such as finite difference method (Virieux, 1986; Levander, 1988), finite element method (Marfurt, 1984; Min et al., 2003), finite-volume method (Brossier et al., 2008), spectral element method (Komatitsch and Vilotte, 1998; Komatitsch and Tromp, 1999), and pseudo-spectral method (Danecek and Seriani, 2008).

In the frequency domain, the wave equation reduces to a linearized system (Marfurt, 1984). The main task of solving this system is to solve an inverse of the impedance matrix, which has a complex-valued symmetric geometry; however, in most of the practical implementation of the wave modeling, the impedance matrix contains a damping term for the absorbing boundary conditions, which makes the impedance matrix non-symmetric (Hustedt et al., 2004; Operto et al., 2007). One can consider two different approaches to solve the linearized system: iterative and direct solver. The iterative solver, which is implemented with Krylov subspace method (Saad, 2003), consumes little computing memory. Hence, it might be useful when we need to compute 3-dimensional waves. Nevertheless, in the iterative solver, it is a non-trivial task to obtain a good preconditioning parameter for stable modeling (Plessix, 2007). In contrast, the direct solver has limitations for applying to large-scale problems due to its memory complexities of LU factorization (Virieux and Operto, 2009); however, for the 2-dimensional problem, the direct solver provides robust and accurate wave solutions (Jo et al., 1996; Štekl and Pratt, 1998).

There have been many studies to improve the accuracy of wave modeling. For instance, Hustedt

et al. (2004) proposed an extension of the mixed-grid method for accurate wave simulation. Chen (2012) introduced an average-derivative optimal scheme to utilize a rectangular mesh with different grid spacing. Another method, known as the compact high-order method (Turkel et al., 2013), is proposed for an accurate and efficient solution of the Helmholtz equation. The compact high-order approach can realize sixth-order accuracy by maintaining the sparsity of the numerical operator (second-order finite difference method). In addition to the easiness of dealing with the dispersion condition, the frequency-domain has an advantage in that we can incorporate high-frequency components without considering a reduced time step. Also, the wave solutions of multiple sources can be acquired by solving one-time matrix inversion (Wu and Alkhalifah, 2018).

Once we acquire accurate wave solutions, then speed becomes important to obtain the solutions rapidly. Fomel et al. (2013) proposed a method for low-rank wave extrapolation using a two-step time marching approach with a real-valued phase function. A low-rank scheme decomposes the original wave propagation matrix into a small set of spatial locations with corresponding representative wavenumbers. Here the rank of the approximation tunes the tradeoff between accuracy and computational efficiency. Sun et al. (2015) proposed a one-step wave extrapolation scheme which applies a complex-values low-rank decomposition to approximate the space-wavenumber domain waves. Another method which performs subgrid upscaling with seismic imaging was introduced by Nunes and Minkoff (2014). The operator upscaling algorithm is widely used for the flow simulation (Arbogast, 2003; Arbogast et al., 2002). Vdovina et al. (2009) applied the operator upscaling method for 3-dimensional elastic wave propagation, and Nunes and Minkoff (2014) employed the method for 2-dimensional acoustic wave simulation with the RTM examples.

Hou and Wu (1997) proposed the pioneering work, which solves the elliptic partial differential equation. This work could be a first step of the GMsFEM to build a basis function in a coarse-mesh by incorporating highly heterogeneous background properties. Efendiev et al. (2011) demonstrated that using multiple basis functions could enhance the accuracy of the coarse-scale wave modeling. Chen (2012) applied the multiscale method to the pressure-velocity formulation to solve the wave equation, and Gibson et al. (2014) proposed discretization examples of the GMsFEM for solving

the acoustic wave equation to reduce the computational cost in time-domain modeling. Inspired by Gibson et al. (2014)'s work, Gao et al. (2015) proposed an extension for solving the elastic wave equation by using both CG- and DG-approximation of the multiscale method. Chung et al. (2016) and Cho et al. (2017b) presented an application of the elastic wave modeling using GMsFEM in fractured media, which includes fractures explicitly in the model. Considering wave simulations, the GMsFEM (Efendiev et al., 2011; Chung et al., 2014; Gibson and Fu, 2015; Cho et al., 2017b) can be considered as a method similar to that proposed by Nunes and Minkoff (2014) since we also utilize multiscale grids; however, the key difference is that GMsFEM directly computes the basis functions including fine-scale heterogeneity without performing any subgrid upscaling. For the Helmholtz equation using GMsFEM (Fu and Gao, 2017), it becomes quite straightforward to utilize various size of spatial grids.

Gibson and Fu (2015) demonstrated the first example of the RTM using the GMsFEM modeling engine. They used the time-domain wave modeling with zero-lag cross-correlation imaging condition, and Cho et al. (2017a) proposed a frequency-domain multiscale RTM with virtual-source imaging condition. They showed that the multiscale RTM could have the flexibility to alter the modeling parameters such as the number of basis functions according to the target frequency band. Cho and Gibson (2018) expanded the frequency-domain RTM by employing a frequency adaptive spatial grid, which uses multiple coarse-meshes to alter the grid size based on the frequencies. Cho et al. (2018) presented the first application of the GMsFEM to the FWI and demonstrated that the multiscale FWI could reduce the computational burden by accelerating the wave modeling.

1.3 Outline

In this dissertation, I introduce the first application of the multiscale methods to frequency-domain RTM and FWI in isotropic acoustic media. Since the frequency-domain RTM is tantamount to the first iteration of FWI, I will start the discussion with the multiscale RTM examples.

In Chapter 2, I will demonstrate the influence of the multiscale basis functions on the seismic imaging via RTM. I first define the multiscale mesh which consists of fine- and coarse-scale meshes. I also determine the local linear spectral problem for building the multiscale basis func-

tions. Then, I investigate the trade-off between the accuracy and speedup in the wave modeling via GMsFEM. In this research, the accuracy of the wave solution acquired by the GMsFEM is validated by comparing it with the solutions obtained by the CG FEM (reference solution). After showing that the GMsFEM with a few basis functions can calculate the low-frequency components more rapidly than the reference solutions without sacrificing accuracy, I will show the multiscale RTM examples by using the large-scale synthetic model (BP 2004 benchmark model). The multiscale RTM with a various number of multiscale basis functions will be presented to demonstrate the influence of the number of multiscale basis functions to the final image quality.

In Chapter 3, I will further develop the frequency-domain multiscale RTM to improve the computational efficiency. In the examples shown in Chapter 2 use the multiscale mesh which consists of two different sizes of mesh: fine- and coarse-mesh. However, in Chapter 3, I superimpose multiple coarse-scale meshes with corresponding projection matrices. Hence, the key idea of this part is to utilize the flexible coarse-grid sizes with an appropriate number of basis functions to accelerate the frequency-domain acoustic wave modeling without losing accuracy. Then, I will apply the same technique to build the RTM imaging condition. For example, by separating the frequency components, I calculate the low-frequency images in the larger size of coarse-grid, then gradually reduced the size of the coarse-scale grid as the frequency increases. In the RTM with the multiscale spatial grid, although wave solutions are computed on a coarse grid, it still provides the RTM images without degrading the image resolution by projecting coarse wave solutions to the reference grid. I will demonstrate the efficiency of the proposed imaging method using the Marmousi-2 synthetic model.

In Chapter 4, I will show the first examples of multiscale FWI, which can reduce the runtime of FWI through the fast wavefield calculation. In the frequency domain, the RTM algorithm can be altered to the FWI algorithm by replacing the back-propagation source (complex conjugated data vector) with the residual vector. In this chapter, in addition to the synthetic examples (Marmousi-2), I will present the field dataset example using data obtained in the Gulf of Mexico area. Through the field data examples, I will demonstrate the influence of the different number of basis functions

by analyzing the forward and backward wavefields, then I will show that applying the proposed multiscale FWI with a relatively small number of basis functions can quickly construct a macro-velocity model using low frequency. I will also propose a strategy to maximize the efficiency of the multiscale FWI by utilizing frequency-adaptive multiscale basis functions based on the target frequency group.

In Conclusions, I will summarize the multiscale RTM and FWI results, and will also propose some potential applications and possible improvements for the future study.

2. FREQUENCY-DOMAIN REVERSE-TIME MIGRATION WITH ACCELERATED WAVE SIMULATION VIA GENERALIZED MULTISCALE FINITE ELEMENT*

2.1 Introduction

Reverse-time migration (RTM), a wave-equation based imaging method, has been widely used as a powerful tool to recover complex subsurface structure (Baysal et al., 1983; McMechan, 1983; Whitmore, 1983; Biondi and Shan, 2002; Symes, 2007) by overcoming the limits of asymptotic ray theory. However, RTM still requires considerable computational cost when applied to large data volumes such as 3-dimensional or 4-dimensional seismic data. Thus, recent work has developed methods to accelerate the wave propagation simulation. For example, Fomel et al. (2013) employed a low-rank approximation of a wavenumber matrix for wavefield extrapolation. Another method is proposed by Nunes and Minkoff (2014). They obtained the wave solution rapidly through subgrid upscaling which makes the coarse grid represent the fine-scale heterogeneity.

We applied another model reduction technique, the Generalized Multiscale Finite Element Method (GMsFEM), to accelerate RTM imaging. The idea of GMsFEM (Efendiev et al., 2011; Chung et al., 2014; Gibson and Fu, 2015) is somewhat similar to that of Nunes and Minkoff (2014) in that we utilize multiscale grids, though the key difference is that GMsFEM directly computes the basis functions including fine-scale heterogeneity without subgrid upscaling. In GMsFEM, the basis functions are computed only once for a specific model, and all simulations are computed on the coarse grid. These basis functions incorporate the most dominant modes computed from local spectral problems, so utilizing them for the simulation on coarse grid can greatly accelerate computations. This advantage becomes more evident when we simulate multiple shots and receiver sets such as the RTM case, since the basis functions are the same for different shot-receiver pairs in the same model.

Frequency-domain wave simulation can accelerate computations using different wave simula-

*Reprinted with permission from "Frequency-domain reverse-time migration with accelerated wave simulation via generalized multiscale finite element" by Cho et al., 2019. Journal of Applied Geophysics, 160, 103-120, Copyright 2019 by Elsevier.

tion parameters for different frequencies. For example, when combining frequency-domain RTM with GMsFEM, we can assign a different number of basis functions for different frequency components to tune the accuracy. In addition, in frequency-domain RTM, the solution can be obtained by a one-time matrix inversion without incorporating reduced time stepping. At early stages of velocity model delineation, we use only lower frequencies in RTM to construct a lower resolution structural image to facilitate the structural interpretation, which requires fewer multiscale basis functions and therefore shorter computation time. After several iterations, when the velocity model has improved accuracy for imaging, we can both increase the number of basis functions and the range of frequencies to enhance resolution.

There are several alternatives such as Kirchhoff or beam migration to resolve seismic images rapidly during the earlier stage of a model construction. Nevertheless, RTM shows better performance to resolve the images in complicated geological structures, although it requires more computational cost. Therefore, to fully utilize the strength of the RTM, we are aiming to accelerate the RTM imaging engines combined with the GMsFEM wave simulation. From our previous work (Gao et al., 2015; Fu and Gao, 2017), we demonstrated that lower frequency components can be simulated with the smaller number of multiscale basis functions without sacrificing accuracy. Also, Artemyev et al. (2015) proposed a method to combine structured and unstructured mesh to incorporate topography for elastic wave modeling using GMsFEM, and Cho et al. (2017b)'s work showed that the GMsFEM can use triangular mesh to delineate complex fracture networks. As the goal of this paper is demonstrating the application of GMsFEM wave modeling to RTM and analyzing the influence of coarse grid size and the number of basis functions on the final RTM images, we did not include any numerical examples with complicate topography which may require local refinement of the mesh.

In this paper, we will briefly illustrate the application of GMsFEM to solve the Helmholtz equation, then we show a set of multiscale RTM examples to demonstrate how we can tune the trade-off between the speed and accuracy. The application of multiscale frequency-domain RTM helps optimize the velocity modeling and imaging workflow.

2.2 Method

Seismic wave simulation in acoustic media is critical for RTM. Therefore, to begin with, we introduce the acoustic Helmholtz equation as follows:

$$-\frac{\omega^2}{\rho v^2}u = \nabla \cdot (\frac{1}{\rho}\nabla u) + f, \tag{2.1}$$

where ρ is the density and v is velocity, and u and f denote pressure fields and a source term, respectively. The angular frequency is represented by ω. We apply this Helmholtz equation to simulate compressional waves in the frequency-domain.

2.2.1 Generalized multiscale finite element solver

The key idea of the generalized multiscale finite element method (GMsFEM) is to solve the Helmholtz equation by utilizing a multiscale mesh Ω that consists of superposed coarse H and fine h meshes as shown in Figure 2.1, where K_i $(i = 1, \cdots, N)$ refers to coarse neighborhood cells of the ith coarse node. The GMsFEM, an efficient model reduction approach, includes two main procedures, an offline and an online stage (Chung et al., 2014; Gao et al., 2015; Gibson and Fu, 2015). The detailed implementation of each stage will be described in the following subsections. In the summary, the goal of each stage is as follows.

1. *Offline stage:* construction of multiscale basis functions by solving local spectral problems for each coarse cell to represent the effects of fine scale heterogeneity.

2. *Online stage:* simulation of seismic waves (inversion of global impedance matrix) on the coarse grid with multiscale basis functions

2.2.1.1 Offline stage

The offline stage is only computed once for a given velocity and density model and is independent of source position. The basis functions are applied in the online stage to simulate wave propagations for all sources on the coarse grid, reducing the number of unknowns and reducing computation time while maintaining good accuracy.

The offline stage starts with the following local eigenproblem in each coarse neighborhood (Efendiev et al., 2011):

$$-\nabla\cdot(\frac{1}{\rho}\nabla\varphi)=\lambda\kappa\varphi,\qquad \kappa=\int_K\frac{1}{\rho}|\nabla\phi|^2d\mathbf{x}, \tag{2.2}$$

ãĚĄwhere ϕ is the function satisfying the following condition with linear boundary condition on K_i:

$$-\nabla\cdot(\frac{1}{\rho}\nabla\phi)=0. \tag{2.3}$$

The function ϕ satisfies the relation shown in equation 2.3, and at the same time it starts from zero and increases linearly to one, or starts from one and decreases to zero at the boundary of coarse cell ∂K. An example of this condition in a coarse neighborhood is displayed in Figure 2.2.

Figure 2.2(b) shows boundary functions with bilinear form on the edge of each coarse cell which is under the procedure for constructing a partition of unity. Figure 2.3 shows the complete shape of a partition of unity ψ_i at ith coarse neighborhood after applying equation 2.3 to compute the inner part of the coarse neighborhood. After the calculation of the partition of unity, to solve the local spectral problem in equation 2.2, we write the generalized eigenvalue problem in discrete

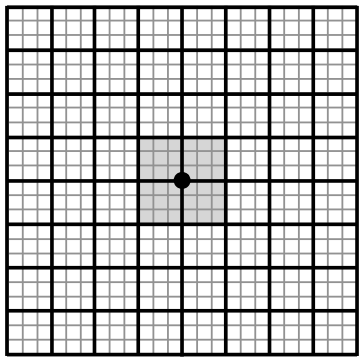
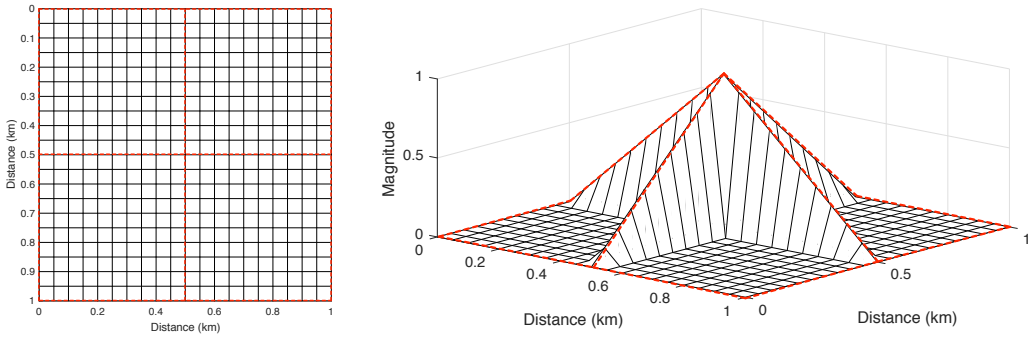


Figure 2.1: Schematic sketch of multiscale domain Ω ($N = 25$); fine grid in gray color and coarse grid in black bold line. Bold points are coarse grid nodes, and grey zone K_i represents the neighborhood of the ith coarse node.



(a) Coarse neighborhood (b) Boundary values at ∂K

Figure 2.2: Mesh definition for conventional partition of unity approach. A coarse neighborhood and corresponding linear boundary condition on ∂K highlighted with red dashed line. The peak point (value=1) is at the center point of K_i coarse neighborhood, which diminishes to the edge of each coarse neighborhood boundary.

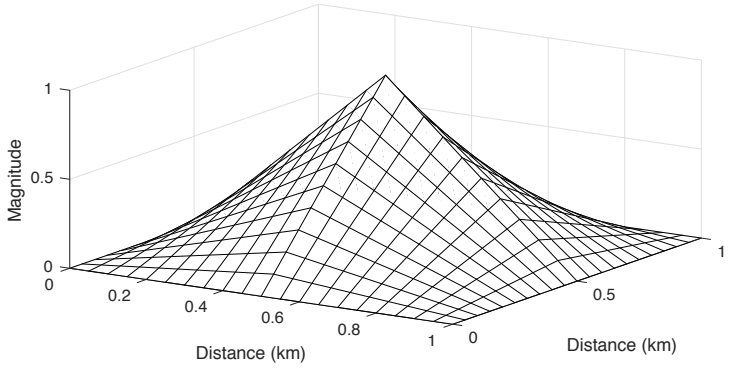


Figure 2.3: Partition of unity ψ_i of a coarse neighborhood K_i.

form as

$$\mathbf{K}\varphi = \lambda \mathbf{M}\varphi, \tag{2.4}$$

where

$$\mathbf{M} = \int_K \kappa\zeta\xi d\mathbf{x}, \qquad \mathbf{K} = \int_K \frac{1}{\rho}\nabla\zeta \cdot \nabla\xi d\mathbf{x}, \tag{2.5}$$

where ζ and ξ are polynomial basis functions that are applied on the fine mesh. $\mathbf{M}$ and $\mathbf{K}$ denote the mass and stiffness matrices for the given coarse cell's neighborhood K_i. When we solve the eigenvalue problem on the fine grid in K, the total number of eigenvalues is identical to the number of the degrees of freedom in a coarse neiborhood K_i. Each matrix is written in the form of products of first order polynomial basis functions, ζ and ξ, on the fine scale grid as presented in equation 2.5. We can calculate a set of eigenvalues with corresponding eigenvectors for each node's coarse neighborhood by solving equation 2.4 (Fu and Gao, 2017). As solutions of local eigenvalue problems for each coarse neighborhood are independent of the location of K_i, the calculation of offline stage can be easily parallelized so it adds little effort to the total computational cost. After solving the local spectral problem (equation 2.4), we order the eigenvalues with corresponding eigenfunctions in an ascending way for each of K_i as:

$$\lambda_i^1 \leq \lambda_i^2 \leq \cdots \leq \lambda_i^{L_i} \leq \cdots \leq \lambda_i^{N_f}, \tag{2.6}$$

where, N_f and L_i are the number of fine nodes within a coarse neighborhood K_i and the number of multiscale basis functions that are required for wave simulations, respectively. Given the spectral problem defined in equation 2.4, each large eigenvalue corresponds to a high wave mode that incorporates higher level of heterogeneity of background models. Therefore, the first L_i eigenfunctions φ_i^j, $(j = 1, \cdots, L_i)$ that we need to select in multiscale finite element solver are controlled by the degree of model heterogeneity on the fine-scale grid. This is analogous to the modes of a vibrating string with fixed ends (constant zero boundary condition on ∂K_i in Figure 2.2) where the fundamental mode is related to the lowest eigen-frequency. Therefore, if we know information about the degree of required accuracy (i.e., appropriate wavelength for the given velocity model), we can

select an appropriate number of basis functions L_i to enhance the efficiency of wave simulations. In other words, we can greatly accelerate the computation without sacrificing the accuracy by choosing the eigenfunctions incorporating the details of heterogeneity which is under our desired accuracy level.

We displayed the eigenfunctions from different eigenvalues (λ=3, 10, 29, 39). Each corresponding eigenfunction represents different level of heterogeneity of the background properties which are shown in Figure 2.4(a) and 2.4(b). For the construction of multiscale basis functions, we defined a partition of unity since these eigenfunctions are not globally continuous, especially on the boundary of the coarse neighborhoods, as presented in Figure 2.4. Hence, we obtain multiscale basis functions Φ_i^j (Figure 2.5) through the element-wise multiplication of φ_i^j (Figure 2.4) with the partition of unity ψ_i (Figure 2.3) as follows: $\Phi_i^j = \varphi_i^j \psi_i$ (Babuška et al., 1995). Combining the partition of unity and eigenfunctions is one of the key parts to build the multiscale basis functions for the coarse mesh.

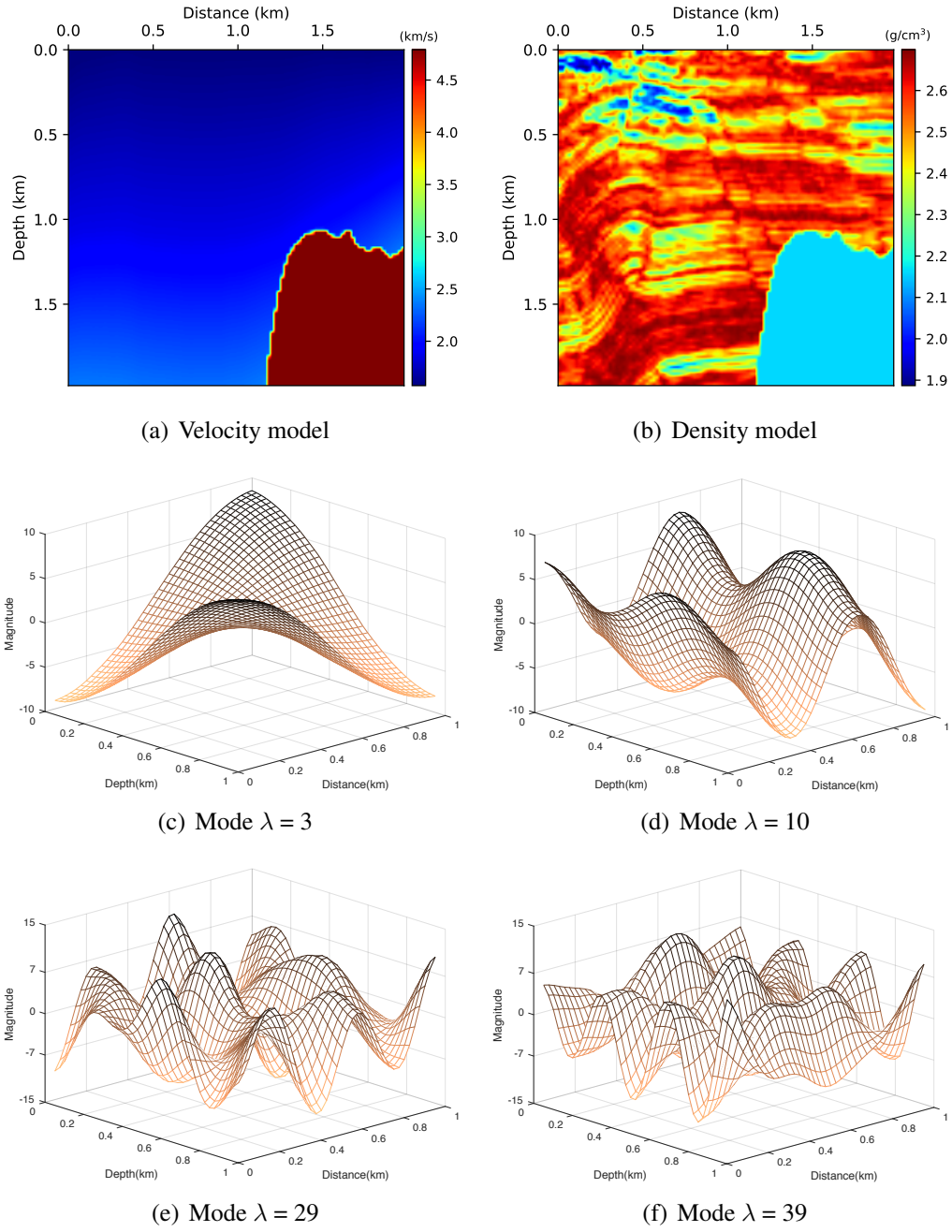
In the GMsFEM, we define the coarse mesh approximation space which consists of multiscale basis functions: $V^H = \text{span}\{\Phi_j^i \mid 1 \leq i \leq N,\ 1 \leq j \leq L_i\}$. A set of multiscale basis functions with different eigenvalues ($\lambda = 3, 10, 29, 39$) is presented in Figure 2.5.

2.2.1.2 Online stage

To solve the Helmholtz equation using GMsFEM on the coarse grid on the online stage, we discretize the system. This discretization of Helmholtz equation follows the approach summarized by Marfurt (1984). We applied the method to implement the wave simulation in coarse grid as

$$\begin{aligned} &\mathbf{S}_H \mathbf{u}_H = \mathbf{f}_H, \\ &\mathbf{S}_H = \mathbf{K}_H + i\omega \mathbf{C}_H + \omega^2 \mathbf{M}_H \end{aligned} \tag{2.7}$$

where $\mathbf{S}_H$ is the global impedance matrix on the coarse grid, and $\mathbf{u}_H$ and $\mathbf{f}_H$ are the frequency-domain pressure field and source term on the coarse grid, respectively. $\mathbf{K}_H$ and $\mathbf{M}_H$ are stiffness and mass matrices. The coarse impedance matrix incorporates the damping term $\mathbf{C}_H$ to suppress



(a) Velocity model

(b) Density model

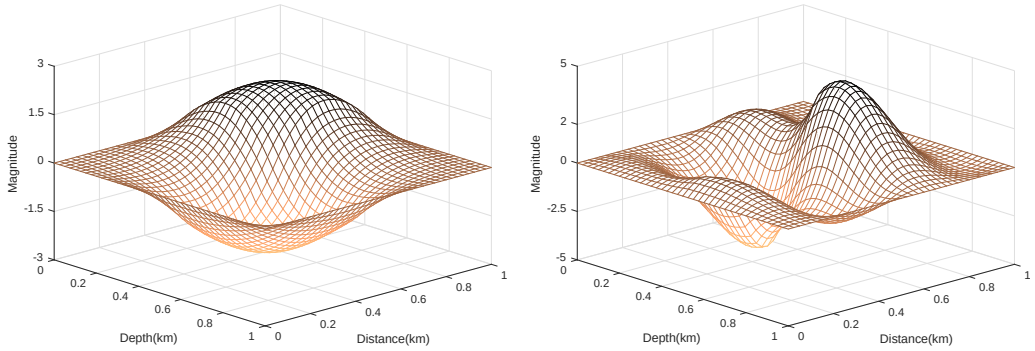
(c) Mode $\lambda = 3$

(d) Mode $\lambda = 10$

(e) Mode $\lambda = 29$

(f) Mode $\lambda = 39$

Figure 2.4: Shape of eigenfunctions which are obtained from equation 2.4. Each eigenfunction represents different wave modes with fine-scale heterogeneity of the background property models shown in (a) and (b).



(a) Mode $\lambda = 3$ (b) Mode $\lambda = 10$

(c) Mode $\lambda = 29$ (d) Mode $\lambda = 39$

Figure 2.5: Multiscale basis functions which are obtained by multiplying the partition of unity (Figure 2.3) and corresponding eigenfunctions in Figure 2.4.

the reflection of outgoing wavefield at the model boundary $\partial\Omega$. In this study, we applied a Perfectly Matched Layer (PML) absorbing boundary condition, and the details of the discrete form to implement $\mathbf{C}_H$ are displayed in Appendix section. Note that we do not construct the coarse-scale stiffness and mass matrices explicitly, but we apply multiscale basis functions to directly compute the coarse-scale impedance matrix from the fine-scale one as detailed below.

The value of the global impedance matrix at the ith coarse node $\mathbf{S}_{H,i}$ can be rewritten as

$$\mathbf{S}_{H,i} = \sum_{k,l} \int_{K_i} \left[-\frac{\omega^2}{\rho v^2} \Phi_i^k \Phi_i^l + \frac{1}{\rho} \nabla \Phi_i^k \cdot \nabla \Phi_i^l \right] d\mathbf{x}, \tag{2.8}$$

Similarly, the local source term $\mathbf{f}_{H,i}$ is as follows:

$$\mathbf{f}_{H,i} = \sum_{k} \int_{K_i} f \Phi_i^k d\mathbf{x}, \tag{2.9}$$

where ω is angular frequency. Φ_i^k and Φ_i^l are multiscale basis functions of the coarse neighborhood K_i. In general, the Helmholtz equation requires a large computational cost even though the matrices are sparse since we need to compute the inversion of the impedance matrix to obtain a wave solution for the given source vector. However, when we construct the impedance matrix on the coarse mesh using multiscale basis functions, the number of unknowns becomes dramatically smaller compared to the fine-scale problem. To assemble the coarse-scale impedance and source matrices, we first need to build the fine-scale matrices. Then we build a projection matrix $\mathbf{R}$, where the column of the $\mathbf{R}$ matrix consists of the discrete multiscale basis functions Φ_i^j from each coarse node neighborhood K_i. The combination of multiscale basis functions in the $\mathbf{R}$ matrix for the selected number of eigen-frequencies can be used to project fine-scale matrices onto the coarse mesh. We can obtain the matrices $\mathbf{S}_H$ and $\mathbf{f}_H$ by multiplying with the projection matrix $\mathbf{R}$ as follows:

$$\begin{aligned} \mathbf{S}_H &= \mathbf{R}^{\mathrm{T}} \mathbf{S}_h \mathbf{R}, \\ \mathbf{f}_H &= \mathbf{R}^{\mathrm{T}} \mathbf{f}_h, \end{aligned} \tag{2.10}$$

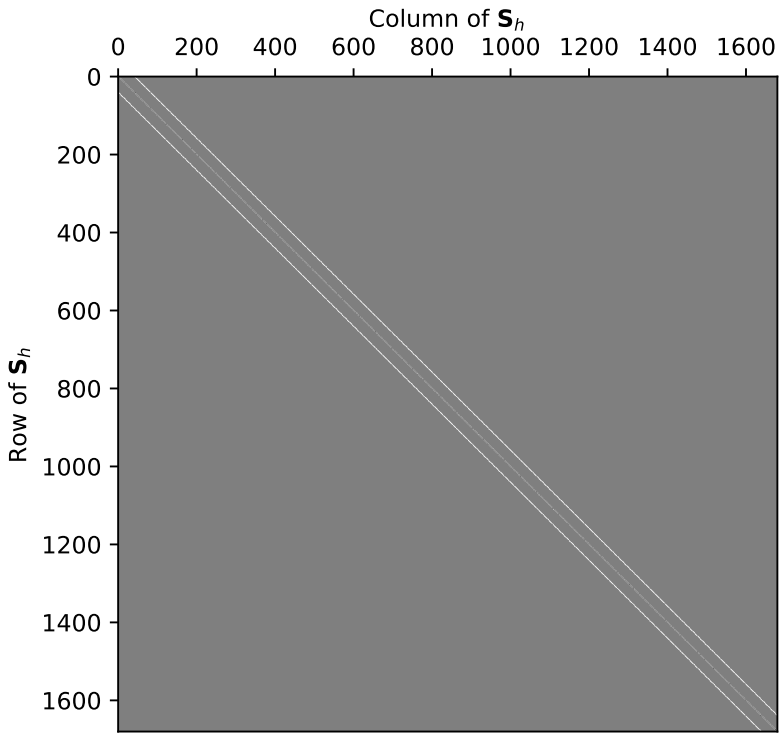


Figure 2.6: Fine-scale impedance matrix $\mathbf{S}_h$ from 40 by 40 model.

where subscript H and h denote coarse and fine mesh, respectively. After obtaining final solutions on coarse mesh, we can simply recover the fine-scale solutions through the multiplication of the projection matrix: $\mathbf{u}_h = \mathbf{R}\mathbf{u}_H$.

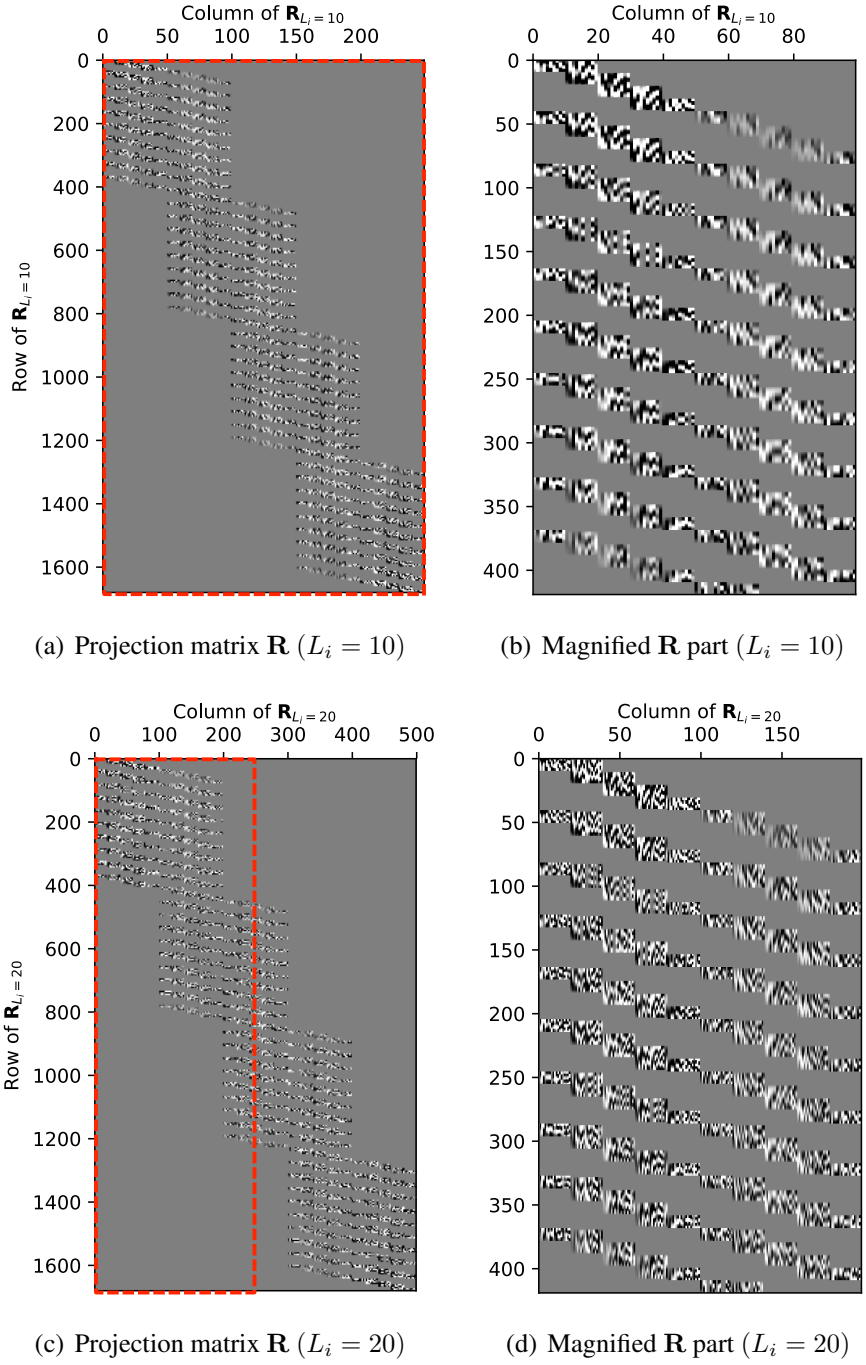
To demonstrate the dramatic reduction of the impedance matrix size graphically, we set the dimension of the model as shown in Figure 2.4(a) and 2.4(b) which consists of 40 by 40 fine-scale cells. Then, the size of impedance matrix for the fine grid can be expressed as the square of fine-scale node points on each axis $(41 \times 41)^2$. In this study, we calculated the fine-scale solutions by utilizing the continuous-Galerkin finite element method (CG FEM), and the fine-scale solution will be considered as a reference to make comparisons between the fine- and coarse-scale wave solutions. In the CG FEM, most of the non-zero values are located around the diagonal components of the impedance matrix (Figure 2.6).

For computing a spectral problem on the coarse grid, we fix the dimension of the coarse mesh to 4 by 4, 16 cells in total. In other words, one coarse cell incorporates 100 fine cells. When we

calculate the coarse-scale impedance matrix by applying the projection matrix $\mathbf{R}$ (Figure 2.7), we can appreciably reduce the size of impedance matrix to accelerate the forward modeling (equation 2.7). The number of rows of the projection matrix is identical to the total number of nodal points of a fine mesh, while the number of columns is related to the number of multiscale basis functions and the nodal points of the coarse mesh. We present the coarse-scale impedance matrices in Figure 2.8 that are obtained by applying equation 2.10. Provided the dimension and the geometry of the fine-scale impedance matrix (Figure 2.6), the geometry of the coarse-scale impedance matrix becomes more complicated as we applied the projection matrix. However, the dimension of the impedance matrix is dramatically decreased in the coarse-scale grid. We calculated two different projection matrices under the same coarse- and fine-mesh by varying the number of multiscale basis functions. Each figure shows a coarse-scale impedance matrix with 10 (Figure 2.8(a)) and 20 (Figure 2.8(b)) basis functions, respectively. A larger number of basis functions increases the size of projection matrix. Therefore, it consumes more computer memory; however, this might be negligible when we consider the amount of reduction in the fine-scale impedance matrix size.

Given that the dimension of impedance matrices for the fine- and coarse-meshes, $\mathbf{S}_h(N_f, N_f)$ and $\mathbf{S}_H(N_cL_i, N_cL_i)$, as well as $\mathbf{R}(N_f, N_cL_i)$, we can roughly estimate how much we can reduce the computational cost. Note that utilizing more multiscale basis functions increases the dimension of coarse-scale matrices. However, we still can attain the impedance matrices which have much smaller dimension than the fine-scale impedance matrix.

When we expand the GMsFEM to 3-D problem, we still solve the same local spectral problem (equation 2.2$\sim$2.4) to compute the multiscale basis functions in the offline stage. Accordingly, we can apply the same algorithm to the 3-dimensional case as far as the computing device has enough memory capacity. However, when the computing resources cannot handle the large impedance and projection matrices duet to memory overflow, we can consider the other multiscale method. As an alternative for reducing the consumption of the massive storage, we may employ high-order multiscale finite element approach (Gao et al., 2018) which does not require any projection matrix construction and explicitly compute the coarse-scale impedance matrices.



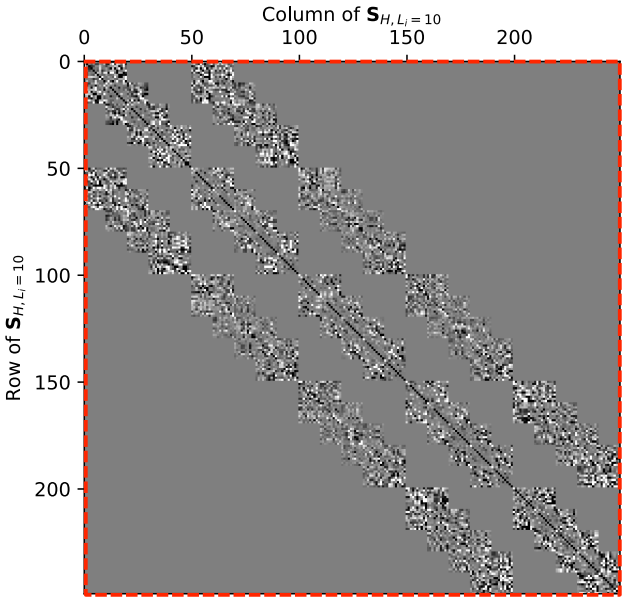
(a) Projection matrix $\mathbf{R}$ ($L_i = 10$)

(b) Magnified $\mathbf{R}$ part ($L_i = 10$)

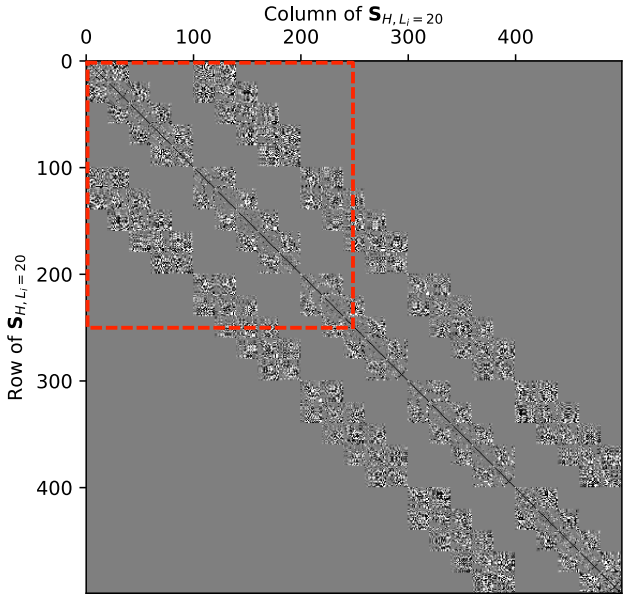
(c) Projection matrix $\mathbf{R}$ ($L_i = 20$)

(d) Magnified $\mathbf{R}$ part ($L_i = 20$)

Figure 2.7: Projection matrix with different number of multiscale basis functions (L_i =10 and 20). Note that the row dimension is same with the dimension of fine-scale mesh. Red dashed line shows the outline of the projection matrix (L_i =10) to guide the comparison of the matrix size.



(a) $\mathbf{S}_H$ ($L_i = 10$)



(b) $\mathbf{S}_H$ ($L_i = 20$)

Figure 2.8: Coarse-scale impedance matrix $\mathbf{S}_H$ with different number of basis functions (L_i =10 and 20), where the red dashed line shows the outline of the impedance matrix (L_i =10) to guide the comparison of the matrix dimension.

2.2.2 RTM imaging conditions using GMsFEM

A conventional approach for imaging in reverse-time migration considers zero-lag crosscorrelation between shot- and receiver-domain wavefield (Baysal et al., 1983), where the receiver wavefield uses time-reversed observed data as a source. This imaging condition can be expressed as

$$\phi_k = \sum_{i=1}^{n_s} \int_0^{T_{\text{max}}} \mathbf{u}_i(t)\mathbf{u}_r(T_{\text{max}} - t)\, dt, \tag{2.11}$$

where, $\mathbf{u}_i$ and $\mathbf{u}_r$ denote the forward- and backward-modeled data, respectively. ϕ_k means the migration image for the kth model parameter, and T_{max} means the data recording time. The index i indicates the shot number. However, in the frequency-domain, this imaging condition needs additional computation due to repetitive forward and inverse Fourier transforms. As an alternative, the migration image can be given by zero-lag crosscorrelation between the partial derivative wavefields (or Born wavefields) with respect to a model parameter (i.e. velocity or density) and the receiver-domain observed data (Chavent and Plessix, 1999):

$$\phi_k = \sum_{i=1}^{n_s} \int_0^{T_{\text{max}}} \left[\frac{\partial \mathbf{u}_i(t)}{\partial m_k}\right]^{\mathrm{T}} \mathbf{d}_i(t)\, dt, \tag{2.12}$$

where $\mathbf{d}_i$ is observed data, and $\frac{\partial \mathbf{u}_i(t)}{\partial m_k}$ is the partial derivative of the wavefield with respect to the model parameter m_k. As mentioned above, using the frequency-domain can be advantageous because it uses only one-time matrix inversion without incorporating reduced time stepping. In addition, when using GMsFEM, different frequency components can easily utilize different input parameters (i.e. number of bases or spatial grid size). Brigham (1988) showed that the migration can be performed using the Fourier transform pairs in the frequency-domain as:

$$\phi_k = \sum_{i=1}^{n_s} \int_0^{\omega_{\text{max}}} \Re \left\{ \left[\frac{\partial \mathbf{u}_i(\omega)}{\partial m_k}\right]^{\mathrm{T}} \mathbf{d}_i^*(\omega) \right\} d\omega, \tag{2.13}$$

where $\mathbf{u}_i$ and $\mathbf{d}_i^*$ indicate the forward modeled wavefield and measured data with complex conjugate in the frequency-domain, respectively. The partial derivative wavefield term shown in equation

2.12 and 2.13 can be acquired from the forward modeling algorithm (Shin et al., 2007). Then, taking the partial derivative of wave equation in matrix form with respect to the model parameter in fine grid gives the following result:

$$\frac{\partial \mathbf{S}_h}{\partial m_k}\mathbf{u}_h + \mathbf{S}_h\frac{\partial \mathbf{u}_h}{\partial m_k} = 0, \tag{2.14}$$

and

$$\frac{\partial \mathbf{u}_h}{\partial m_k} = \mathbf{S}_h^{-1}\mathbf{v}, \tag{2.15}$$

where $\mathbf{v}$ is the virtual source term which consists of Born modeling operator and fine-scale background wavefield as $-\frac{\partial \mathbf{S}_h}{\partial m_k}\mathbf{u}_h$. The virtual source required to generate the Born wavefield and be written as

$$\begin{pmatrix} v_k^1 \\ v_k^2 \\ \vdots \\ v_k^n \end{pmatrix} = -\frac{\partial \mathbf{S}}{\partial m_k}\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}. \tag{2.16}$$

The discrete matrix notation of equation 2.15 at kth arbitrary model parameter can be expressed as

$$\begin{pmatrix} \frac{\partial u_1}{\partial m_k} \\ \frac{\partial u_2}{\partial m_k} \\ \vdots \\ \frac{\partial u_n}{\partial m_k} \end{pmatrix} = \mathbf{S}^{-1}\begin{pmatrix} v_k^1 \\ v_k^2 \\ \vdots \\ v_k^n \end{pmatrix}. \tag{2.17}$$

Taking the transpose of equation 2.17 gives

$$\begin{pmatrix} \frac{\partial u_1}{\partial m_k} & \frac{\partial u_2}{\partial m_k} & \cdots & \frac{\partial u_n}{\partial m_k} \end{pmatrix} = \begin{pmatrix} v_k^1 & v_k^2 & \cdots & v_k^n \end{pmatrix}\left(\mathbf{S}^{-1}\right)^{\mathrm{T}}. \tag{2.18}$$

The virtual source matrix can be acquired by generalizing the kth model parameter m_k, where $k = 1, 2, \cdots, p$. Therefore, we can rewrite above equation as follows:

$$\begin{aligned}
\begin{pmatrix} \frac{\partial u_1}{\partial m_1} & \frac{\partial u_2}{\partial m_1} & \cdots & \frac{\partial u_n}{\partial m_1} \end{pmatrix} &= \begin{pmatrix} v_1^1 & v_1^2 & \cdots & v_1^n \end{pmatrix} \left(\mathbf{S}^{-1}\right)^{\mathrm{T}} \\
\begin{pmatrix} \frac{\partial u_1}{\partial m_2} & \frac{\partial u_2}{\partial m_2} & \cdots & \frac{\partial u_n}{\partial m_2} \end{pmatrix} &= \begin{pmatrix} v_2^1 & v_2^2 & \cdots & v_2^n \end{pmatrix} \left(\mathbf{S}^{-1}\right)^{\mathrm{T}} \\
\begin{pmatrix} \frac{\partial u_1}{\partial m_3} & \frac{\partial u_2}{\partial m_3} & \cdots & \frac{\partial u_n}{\partial m_3} \end{pmatrix} &= \begin{pmatrix} v_3^1 & v_3^2 & \cdots & v_3^n \end{pmatrix} \left(\mathbf{S}^{-1}\right)^{\mathrm{T}} \\
&\vdots \\
\begin{pmatrix} \frac{\partial u_1}{\partial m_p} & \frac{\partial u_2}{\partial m_p} & \cdots & \frac{\partial u_n}{\partial m_p} \end{pmatrix} &= \begin{pmatrix} v_p^1 & v_p^2 & \cdots & v_p^n \end{pmatrix} \left(\mathbf{S}^{-1}\right)^{\mathrm{T}}.
\end{aligned} \tag{2.19}$$

The transpose matrix of the global virtual source matrix can be written as follows:

$$\mathbf{f}_{\mathbf{v}}^{\mathrm{T}} = \begin{pmatrix} v_1^1 & v_1^2 & \cdots & v_1^n \\ v_2^1 & v_2^2 & \cdots & v_2^n \\ \vdots & \vdots & \ddots & \vdots \\ v_p^1 & v_p^2 & \cdots & v_p^n \end{pmatrix} \tag{2.20}$$

The size of the virtual source matrix $\mathbf{f}_{\mathbf{v}}^{\mathrm{T}}$ depends on the number of model parameters that are included in the model domain for imaging. Therefore, it is $n \times n$ matrix provided that all parameters of grid points ($p \rightarrow n$) are included in imaging area.

We can calculate the Born wavefield by taking the inverse of fine scale $\mathbf{S}_h$ impedance matrix and combining it with the virtual source vector (Pratt et al., 1998). Substituting equation 2.15 into equation 2.13 gives the final image condition for the kth model parameter as follows:

$$\phi_k = \sum_{i=1}^{n_s} \int_0^{\omega_{\max}} \Re \left[\mathbf{v}^{\mathrm{T}} \mathbf{R} \mathbf{S}_H^{-1} \mathbf{R}^{\mathrm{T}} \mathbf{d}_i^*(\omega) \right] \, d\omega. \tag{2.21}$$

The combination of impedance matrix and complex conjugate of Fourier transformed measured data, $\mathbf{S}_H^{-1} \mathbf{R}^{\mathrm{T}} \mathbf{d}_i^*$, represents the back propagation of the observed data in coarse grid. By convolving it with the virtual source vector $\mathbf{v}^{\mathrm{T}}$, we acquire the reverse-time migration image at the kth node.

The virtual source vector $\mathbf{v}$ can be extended to matrix form $\mathbf{f_v}$ to consider all image nodes.

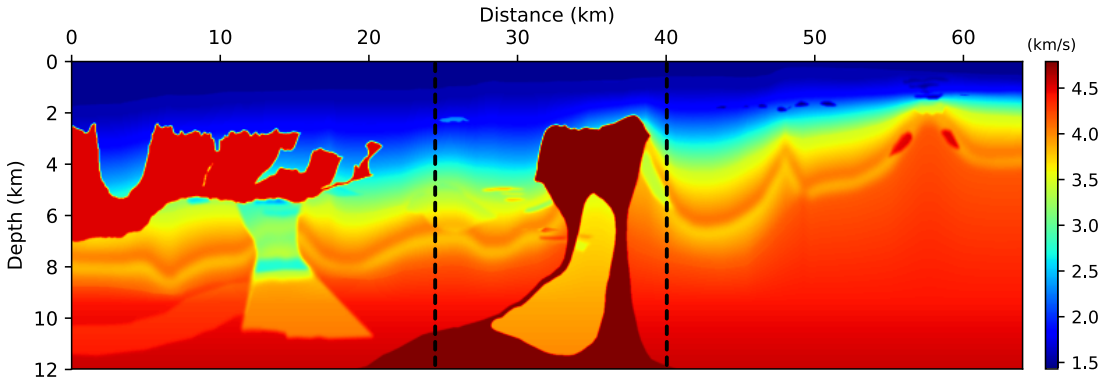
In RTM, we need an additional term to enhance the imaging condition. To put it differently, we are not able to recover the actual amplitude variation using only the condition shown in equation 2.21, especially in high contrast impedance cases. To solve this problem, Claerbout (1971) suggested scaling the cross-correlation image by applying source illumination. Also, Chavent and Plessix (1999) introduced that the migration image can be improved by multiplying the inverse of the Hessian matrix. However, using the full Hessian matrix hinders the practical implementation of migration due to excessive computational cost. Accordingly, we applied the diagonal of the pseudo-Hessian matrix (Shin et al., 2001) as follows:

$$\phi = \frac{\sum_{i=1}^{n_s} \int_0^{\omega_{\max}} \Re \left[\mathbf{f}_{\mathbf{v}}^{\mathrm{T}} \mathbf{R} \mathbf{S}_H^{-1} \mathbf{R}^{\mathrm{T}} \mathbf{d}_i^*(\omega) \right] \, d\omega}{\sum_{i=1}^{n_s} \int_0^{\omega_{\max}} \Re \left[\mathrm{diag}(\mathbf{f}_{\mathbf{v}}^{\mathrm{T}} \mathbf{f}_{\mathbf{v}}) \right] \, d\omega + \lambda}, \tag{2.22}$$

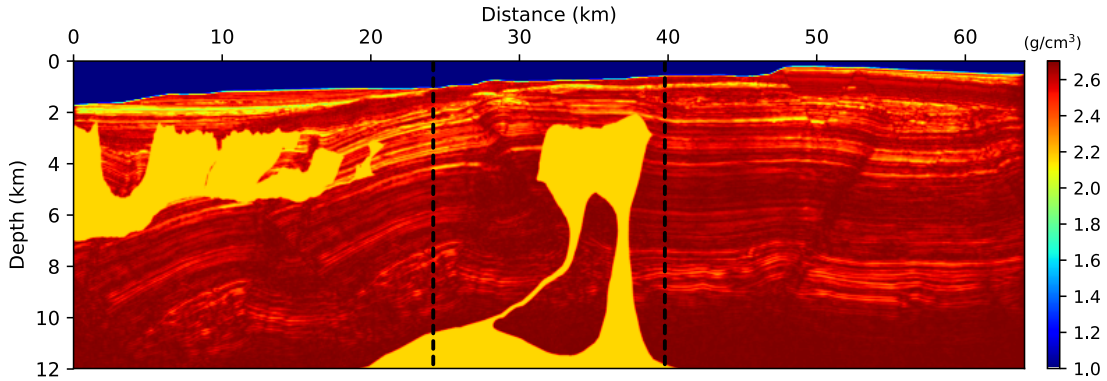
where λ is a damping factor, and $\mathbf{f}_{\mathbf{v}}^{\mathrm{T}} \mathbf{f}_{\mathbf{v}}$ means the pseudo-Hessian matrix. Here we note that the imaging condition ϕ means the RTM image for the entire spatial domain. By utilizing the final imaging condition shown above, we normalize the RTM imaging condition to recover the true amplitude variation with a reasonable computational burden. The advantage of working with the frequency-domain is that we can separate the wavefield into independent frequency components, which enables us to apply the inverse Hessian matrix frequency by frequency for the gradient scaling.

2.3 Numerical examples

For simulation of waves and imaging, we used the BP benchmark model (Billette and Brandsberg-Dahl, 2004) as shown in Figure 2.9. The model includes a large salt diapir with a complex boundary, and the flank of salt is challenging to resolve due to its steep slope. Through multiple tests of frequency-domain wave simulation and RTM on this model, we demonstrate how effectively the coarse-grid simulation can resolve the image of detailed features with high impedance contrast. For references, RTM examples of a classical approach of the RTM using time-domain modeling can be found from other scientist's work as follows: Liu et al. (2012); Foltinek et al. (2009); Pestana



(a) Velocity model



(b) Density model

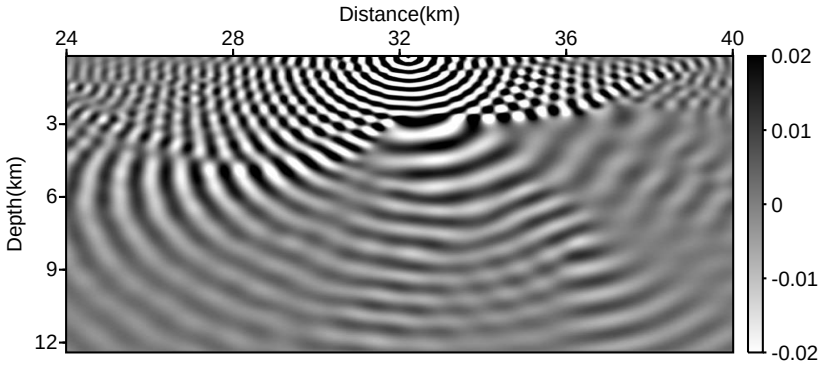
Figure 2.9: BP 2004 Benchmark model adapted from Billette and Brandsberg-Dahl (2004).

and Stoffa (2010).

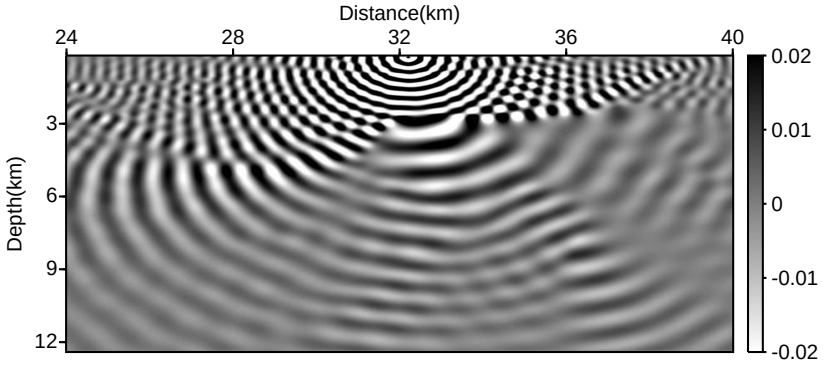
First, we demonstrate the simulation of wave propagation on the coarse grid. For the wave simulation, we applied a single source which is located at the center of the free surface. The corresponding modeling results are displayed from Figure 2.10 to Figure 2.12, and computation time is displayed in Table 2.1. For each frequency, the reference solution is calculated by using the CG FEM. We calculated the wave solution for the entire domain Ω, and displayed only the area marked with dashed black line in Figure 2.9 to highlight the waves around the source location. For the source wavelet, we used Ricker wavelet with three different frequencies (5 Hz, 10 Hz, and 15 Hz). The sizes of fine and coarse cell are 20 m and 100 m, respectively.

Each example shows waves from a different frequency but the same number of multiscale basis functions: 5 and 10. We can observe that the smaller number of basis functions ($L_i = 5$) can model the low frequency waves accurately within a shorter amount of computational time as shown in Figure 2.10(a), while the error increases as the frequency becomes higher (Figure 2.11(a) and 2.13(a)). Nevertheless, we can tune the accuracy by applying more basis functions (Figure 2.13(b)), and multiscale forward modeling still gives the faster solution than the reference solution. The set of results shows that the accuracy of GMsFEM solutions is highly correlated with the number of basis functions L_i and the corresponding frequency. For example, a smaller number of basis functions can simulate low-frequency waves accurately, while more basis functions are required to obtain accurate wave solutions with high-frequencies. In other words, a greater number of basis functions incorporates more detailed wave modes, so we can capture the details of short wavelengths.

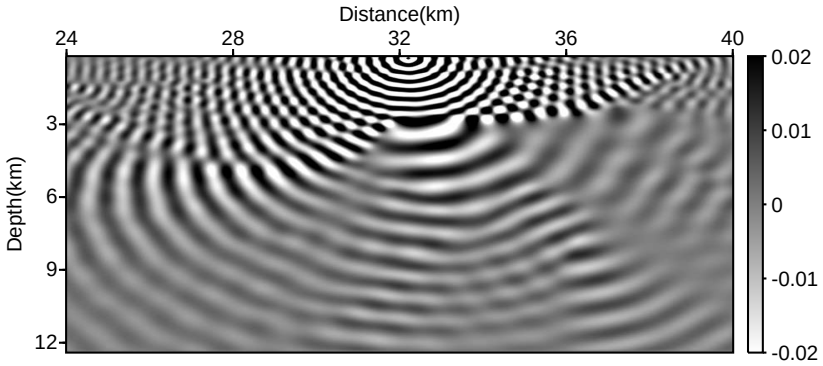
In addition to the number of basis functions, the coarse cell dimension also has an influence on the performance of multiscale wave modeling. To demonstrate this point, we performed the wave simulation in two different sizes of coarse cell (100 m and 200 m) with the same range of frequencies, and measured the l_2-error for the different number of basis functions. As we solve the eigenvalue problem (equation 2.4), the number of eigenvalues is equivalent to the degrees of freedom in a coarse neighborhood K_i. Therefore, we assigned more basis functions for a larger



(a) GMsFEM ($L_i = 5$), $\|e_{5\text{Hz}}\|_{l_2} = 0.539$

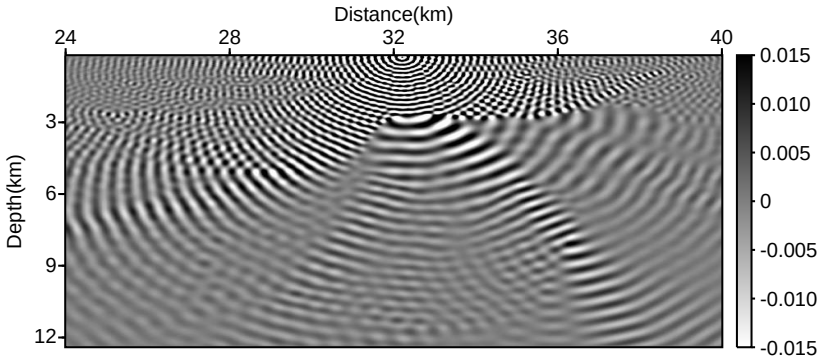


(b) GMsFEM ($L_i = 10$), $\|e_{5\text{Hz}}\|_{l_2} = 0.058$

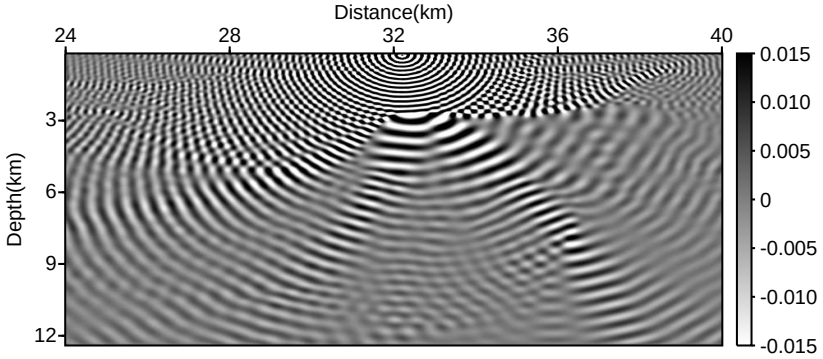


(c) Reference solution (5 Hz)

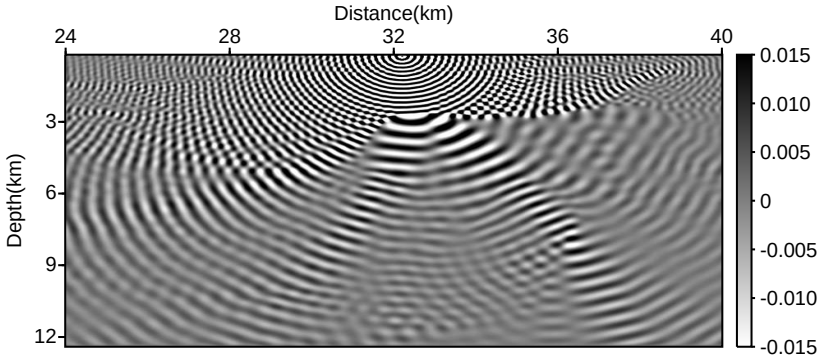
Figure 2.10: GMsFEM solution vs. reference solution from 5 Hz source frequency.



(a) GMsFEM ($L_i = 5$), $\|e_{10\text{Hz}}\|_{l_2} = 0.928$

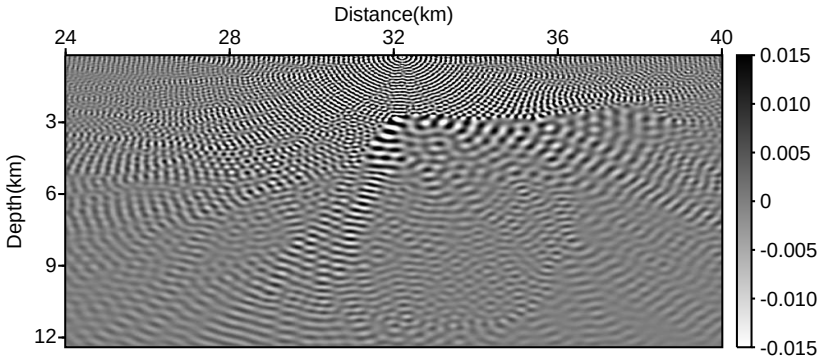


(b) GMsFEM ($L_i = 10$), $\|e_{10\text{Hz}}\|_{l_2} = 0.085$

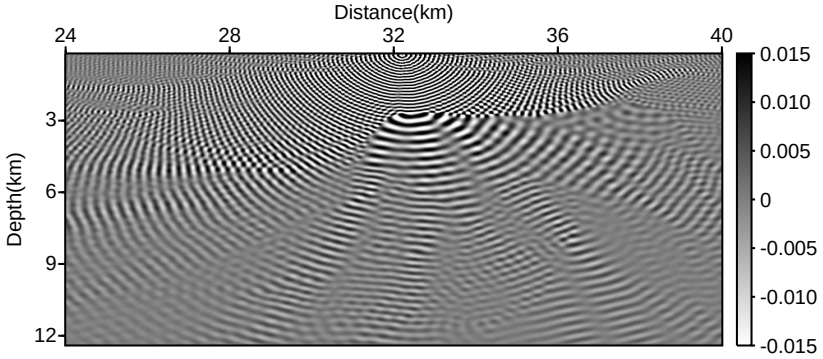


(c) Reference solution (10 Hz)

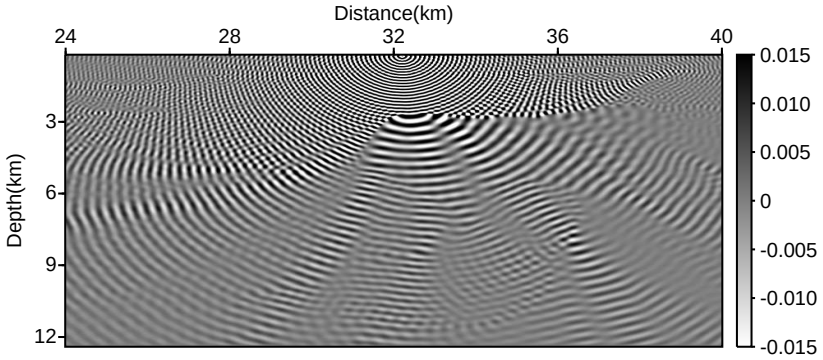
Figure 2.11: GMsFEM solution vs. reference solution from 10 Hz source frequency.



(a) GMsFEM ($L_i = 5$), $\|e_{15\text{Hz}}\|_{l_2} = 1.282$

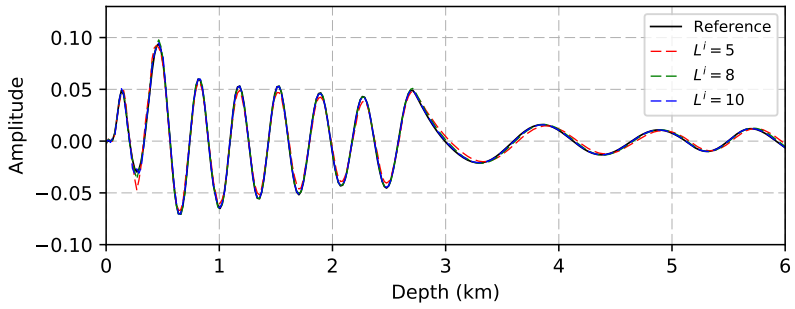


(b) GMsFEM ($L_i = 10$), $\|e_{15\text{Hz}}\|_{l_2} = 0.105$

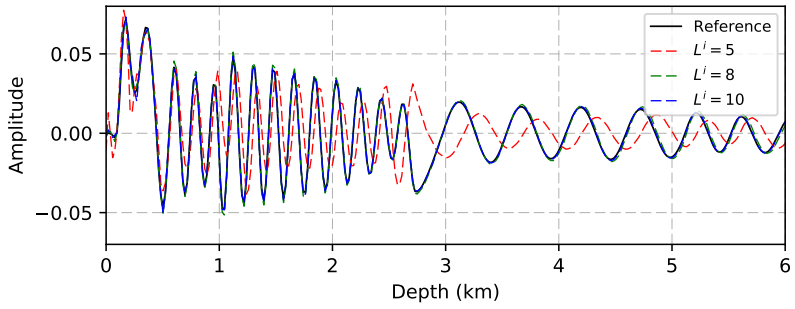


(c) Reference solution (15 Hz)

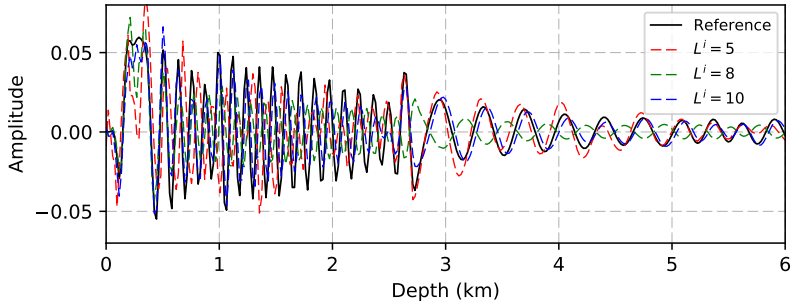
Figure 2.12: GMsFEM solution vs. reference solution from 15 Hz source frequency.



(a)



(b)



(c)

Figure 2.13: Comparison of waves from GMsFEM solution and reference (CG FEM) solution from different source frequency: (a) 5 Hz, (b) 10 Hz, and (c) 15 Hz.

coarse grid as shown in Table 2.1. Considering the 5 Hz case, a larger coarse grid with more basis functions could simulate the waves more rapidly with higher accuracy. In contrast, the simulation in the smaller coarse grid (100 m) performs better at higher frequency (15 Hz case) by employing smaller number of basis functions.

The fine-scale mesh of the GMsFEM depends on the dimension of the background property. In other words, in the GMsFEM, the fine grid cell size is fixed, and only coarse cell size is altered to tune the computational speed and accuracy. However, utilizing too large grid size than the given frequency band may cause dispersion error in wave simulations. The GMsFEM with a large coarse cell (i.e., 100 m in 15 Hz) can have a better suppression of the dispersion error compared to the other classical methods (i.e., finite-difference or finite-element method) with 100 m grid size by applying multiple basis functions. Here the key difference of the GMsFEM is that it uses basis functions which incorporate fine-scale heterogeneity information to enhance the accuracy of the coarse-scale modeling.

In the GMsFEM, defining the dispersion condition and performing a stability analysis is still challenging task since the multiscale approach is an approximation of a numerical method (CG FEM in this case). Hence, it is difficult to build a direct relationship between numerical solutions of the GMsFEM and analytic solutions. For this reason, the modeling parameters of the multiscale method such as the size of coarse cell and required number of basis functions are somewhat empirical parameters. Gibson et al. (2014) showed examples of the numerical analysis of dispersion error for different GMsFEM parameters in the acoustic wave problem. In this study, we measured the l_2-relative errors to measure the accuracy of the multiscale wave solutions. The l_2-error and observed runtime with corresponding speedup are displayed in Table 2.1. We measured the runtime using single computing processor unit.

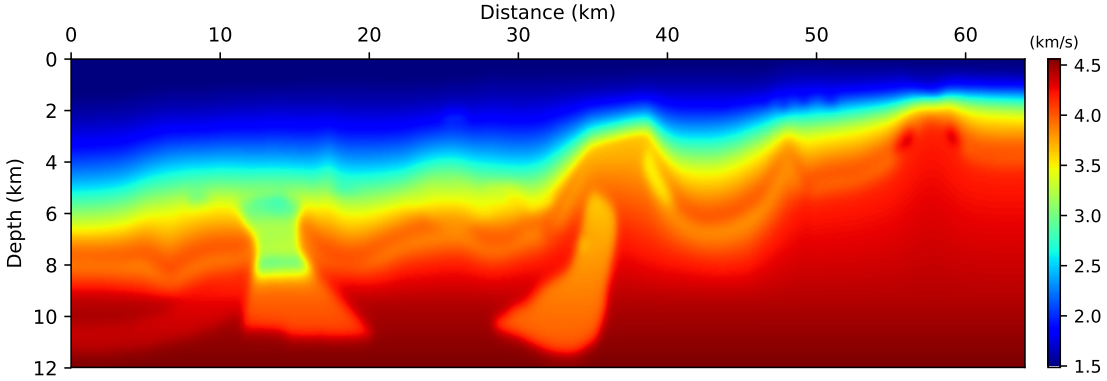
We can take two examples of utilizing the multiscale RTM in velocity model building cycle. 1) First, when we combine this tool with salt plotting or horizon picking to delineate complex salt structures, the multiscale RTM can provides the result rapidly. As the multiscale forward algorithm performs the offline stage only once for the entire model and applies the same bases for

Table 2.1: Runtime and l_2-relative errors for the different number of basis functions. Runtime for the fine-scale reference solution is 4.21 s. The number in parenthesis shows the speedup compared to the reference (CG FEM) case.

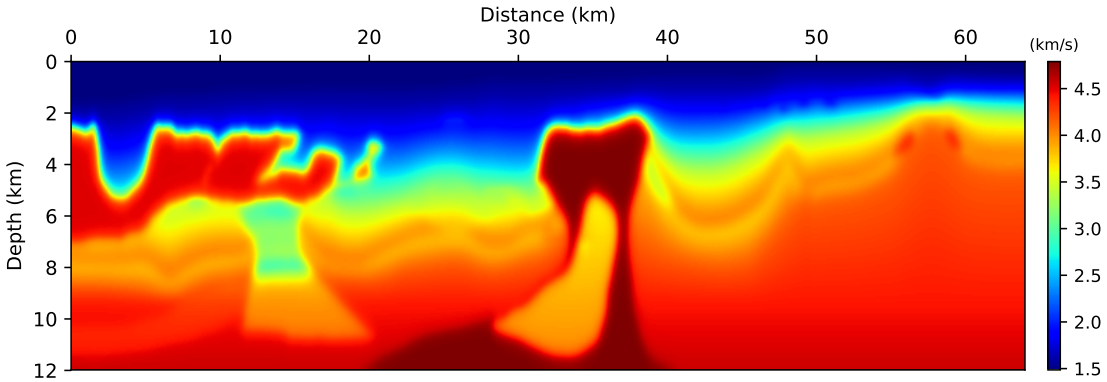
H=100 m	Time (s)	$\|e_{5Hz}\|$	$\|e_{10Hz}\|$	$\|e_{15Hz}\|$
$L_i = 5$	0.83 (4.96)	0.539	0.928	1.282
$L_i = 6$	1.01 (4.08)	0.234	0.423	0.707
$L_i = 7$	1.31 (3.13)	0.218	0.338	0.694
$L_i = 8$	1.56 (2.64)	0.221	0.299	0.565
$L_i = 9$	1.88 (2.19)	0.205	0.283	0.382
$L_i = 10$	2.12 (1.94)	0.058	0.085	0.105
H=200 m	Time (s)	$\|e_{5Hz}\|$	$\|e_{10Hz}\|$	$\|e_{15Hz}\|$
$L_i = 10$	0.72 (5.72)	0.092	1.408	1.156
$L_i = 12$	0.82 (5.02)	0.062	1.151	1.382
$L_i = 14$	1.08 (3.81)	0.057	1.065	1.163
$L_i = 16$	1.27 (3.24)	0.047	0.645	0.875
$L_i = 18$	1.53 (2.69)	0.041	0.113	0.423
$L_i = 20$	1.83 (2.25)	0.038	0.061	0.227

multiple shots with different locations, the cost savings increase when we simulate a large number of shots. In addition, given that applying GMsFEM with a smaller number of bases can capture macro-scale features within a short period of time, applying this model reduction technique to RTM will be beneficial by manipulating the level of accuracy adaptively according to the stage of velocity updating. For instance, when a velocity model is actively updated and therefore less accurate, we use a smaller number of basis to test the large-scale structural image. Once detailed velocity structure is determined, we assign more basis functions to resolve the subtle variation of stratigraphic features. 2) Another advantage of taking GMsFEM for a wave modeling is that we can accelerate the wave calculation which directly contributes to acceleration of the wave-equation based imaging technique such as FWI or least-squares migration. The amount of speedup in FWI or least-square RTM can be greater than in RTM case since they requires multiple iterations of a model or an image update.

To demonstrate the influence of multiscale wave modeling on the final migration image, we applied the RTM with a varying number of basis functions to demonstrate these advantages of multiscale RTM to optimize the velocity modeling workflow. We evaluated images using three



(a)



(b)

Figure 2.14: Velocity models at different stage of model update: (a) velocity model without salt diapir after intermediate model updating, and (b) accurate velocity model after finalizing the model updating.

hypothetical velocity models representing different stages of model design. We started with a velocity model (Figure 2.14(a)) with accurate strata. We then performed the RTM with the smoothed velocity model (Figure 2.14(b)), which includes salt diapirs, to acquire the final image. When we update a velocity model combined with the GMsFEM RTM, the multiscale basis functions need to be updated when a certain velocity model update create a large velocity contrast. Therefore, we need to regenerate the basis functions when the velocity model is altered from Figure 2.14(a) to Figure 2.14(b). However, the offline stage adds little effort compared to the total computation of RTM.

We generated synthetic data for the fine grid through a forward modeling with a Ricker wavelet

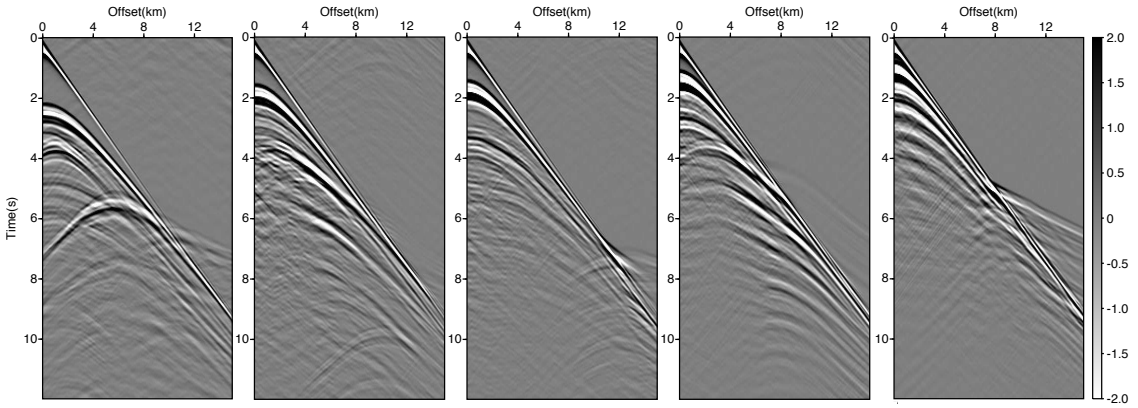
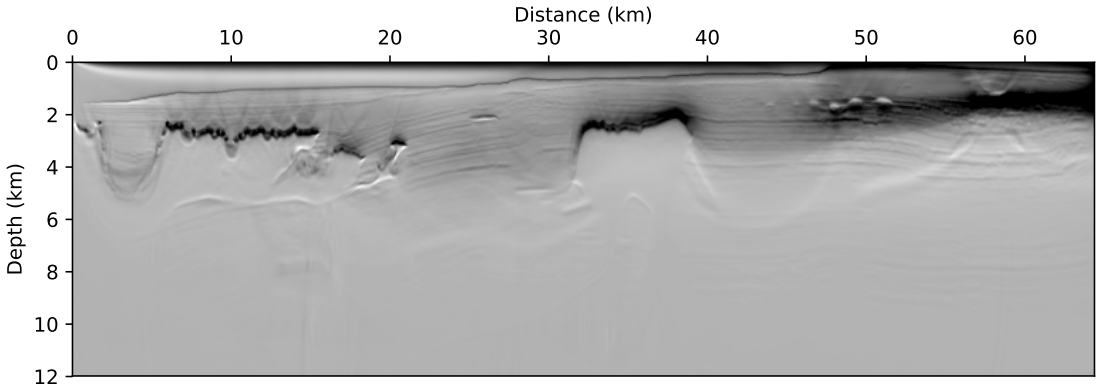
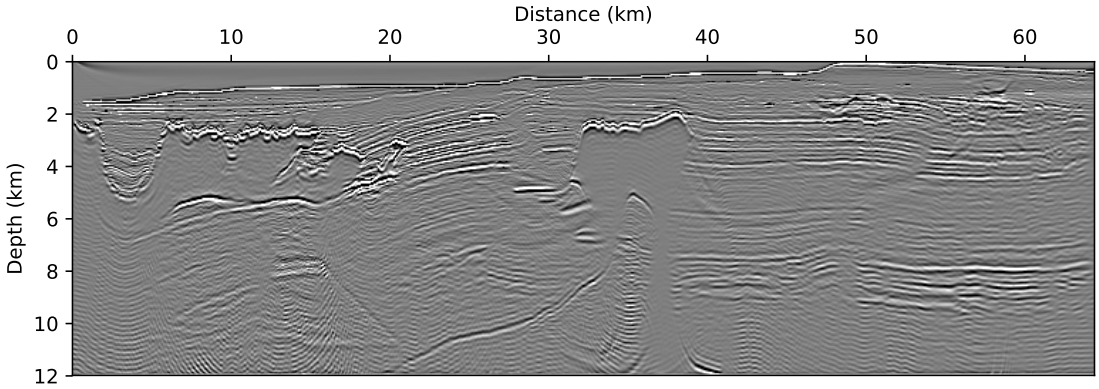


Figure 2.15: An example of the synthetic shot gathers.



(a) Gradient of fine-scale RTM image.



(b) Reference (fine-scale) RTM image, Runtime=213.61 min.

Figure 2.16: Reference RTM images ($f_{\max}$=15Hz) that are performed using fine-scale Continuous-Galerkin finite element method: (a) the gradient image and (b) final RTM image after applying the Laplacian filter. The average memory usage per processor unit is 861.52 MB.

(central frequency $= 15$ Hz). The size of grid is 20 m. Receiver and shot interval are 20 m and 100 m, respectively. The length of the streamer is 15 km and 639 shots are applied at the fixed 20 m depth, and the total recording time is 12 s. An example of a shot gather with the corresponding spectrum is displayed in Figure 2.15. For the RTM on the fine-scale grid, we employed the CG FEM and consider the result as a reference (Figure 2.16); however, we used GMsFEM for the simulation of forward and backward wavefields in multiscale RTM. The imaging condition shown in equation 2.22 yields a gradient image as presented in Figure 2.16(a). Accordingly, we applied a Laplacian filter and made a correction for the geometrical spreading to obtain the final RTM image (Figure 2.16(b)). The results from the multiscale RTM at different stages of the model update are displayed in Figure 2.17 and 2.18. We also displayed the gradient image to observe any existence of noisy signals since a clean gradient is important for earlier stages of velocity model building for the model update. The gradient images show high contrast around the salt boundary, and it does not include any artifact. Therefore, we can conclude that the artifacts becomes visible after applying the Laplacian filter.

The runtime for the coarse-scale results measures the online stage, while values for the fine-scale results are for the impedance matrix inversion ($\mathbf{S}_h^{-1}$). The online time includes the time taken for the coarse impedance matrix inversion ($\mathbf{S}_H^{-1}$) and the large sparse matrices multiplication ($\mathbf{R}^{\mathrm{T}}\mathbf{S}_h^{-1}\mathbf{R}$ and $\mathbf{R}^{\mathrm{T}}\mathbf{f}_h$). In multiscale RTM, we need to perform additional matrix operations; however, by reducing the size of impedance matrix significantly from (N_f,N_f) to (N_c,N_c), the multiscale modeling approach rapidly calculates the matrix inversion.

We performed RTM on a coarse mesh which has the same grid size as the minimum wavelength ($\approx$ 100 m). For all the RTM examples with different background velocities, we applied the same source frequency (15 Hz) and varied the number of basis functions (L_i =5, 8, 10).

According to the table 2.2, when we applied 5 multiscale bases, we could obtain images with CPU time reduced by 76% from the reference case; however, as commonly observed from the different RTM examples with few multiscale basis functions ($L_i = 5$), numerical artifacts that are caused by the lack of basis functions contaminates the RTM image, especially the deeper part

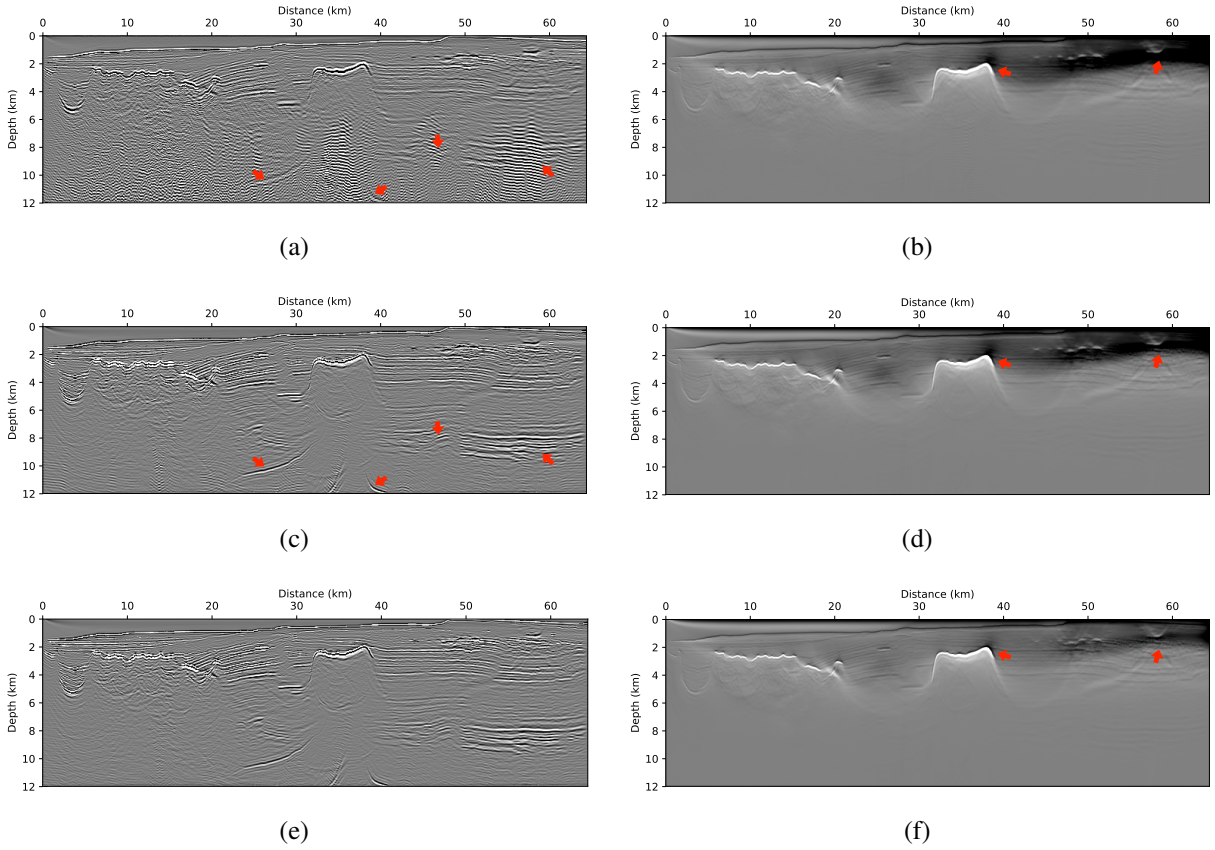


Figure 2.17: Multiscale RTM images obtained from the correct strata velocity (Figure 2.14(a)) with different number of basis functions: (a, b) $L_i = 5$, (c, d) $L_i = 7$, and (e, f) $L_i = 9$. Right panels (b, d, f) show the corresponding gradient image of the RTM images shown in left panels (a, c, e). Relative l_2-errors are as follows: (a) $\|e_{L_i=5}\|$=0.869, (b) $\|e_{L_i=7}\|$=0.211, and (c) $\|e_{L_i=9}\|$=0.087.

of the model. Although deeper part of the strata are not obvious, multiscale RTM with 5 basis functions still can capture the large scale structures. As the goal of imaging in earlier stage of velocity modeling is not defining all the detailed strata but locating macro geological structures quickly, the multiscale RTM image with 5 basis functions is good enough for verifying the near surface structures. Then, we apply more basis functions to resolve deeper part of the subsurface. As presented in Figure 2.17, when we use the correct strata velocity model without salt diapir, the multiscale RTM could successfully resolve the salt top even in the case with fewer basis functions (Figure 2.17(a)). After defining the salt top and the shallower part of the model through several iteration of velocity modeling using the multiscale RTM with small number of basis functions, we

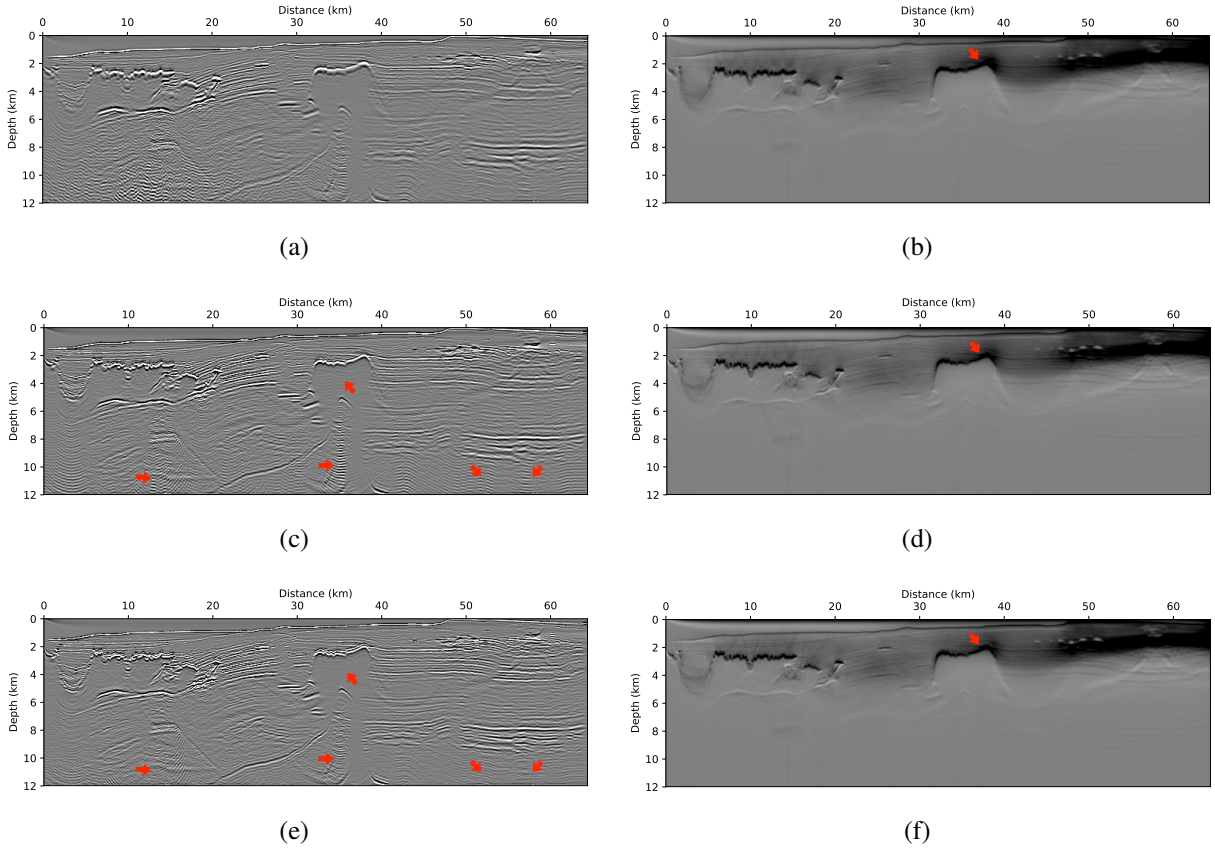


Figure 2.18: Multiscale RTM images obtained from the smoothed true velocity (FIgure 2.14(b)) with different number of basis functions: (a, b) $L_i = 5$, (c, d) $L_i = 7$, and (e, f) $L_i = 9$. Right panels (b, d, f) show the corresponding gradient image of the RTM images shown in left panels (a, c, e). Relative l_2-errors are as follows: (a) $\|e_{L_i=5}\|$=0.859, (b) $\|e_{L_i=7}\|$=0.236, and (c) $\|e_{L_i=9}\|$=0.075.

then can apply more number of basis functions to delineate salt bottom and to resolve deeper part of the models as pointed with red arrows in Figure 2.17(a) and 2.17(c). Comparing the cases with 7 and 9 basis functions, we can observe that the RTM images with an accurate background velocity exhibit improvements on suppressing dispersion in the image with 9 basis functions, as indicated with red arrows (Figure 2.18). Considering the multiscale RTM results shown in Figure 2.18(c), 9 multiscale basis functions could resolve the image in a shorter amount of time without sacrificing the accuracy. The multiscale RTM with 9 basis functions could accelerate the imaging by reducing computation time; 46% faster than the RTM with continuous CG FEM wave simulation engine.

In short, multiscale RTM can rapidly resolve the image of the salt boundary and capture the

Table 2.2: Runtime of multiscale RTM, memory usage, and speedup for the different number of basis functions.

Method	# Basis functions (L_i)	RTM runtime	Speedup	Memory usage
CGFEM RTM	-	213.61 min	-	861.52 Mb
GMsFEM RTM	5	51.58 min	4.14	1157.96 Mb
GMsFEM RTM	7	101.52 min	2.11	1494.63 Mb
GMsFEM RTM	9	146.27 min	1.46	1743.07 Mb

large-scale geological structures by applying a small number of basis functions. Although multiscale RTM with few basis functions cannot capture the detailed variation of geological structures, it can be useful when the velocity model is actively updated since it resolves macro structures rapidly. Utilizing more basis functions ($L_i = 7$), we could obtain more clear subsurface strata and the salt boundary by suppressing numerical artifacts contaminating deeper part of the model. By allocating additional basis functions ($L_i = 9$), the signal to noise ratio in RTM image is enhanced, and all the strata in deeper part are clearly interpretable. We may consider a flexible approach to utilize the GMsFEM modeling engine for improving the efficiency of the wave modeling. For instance, it is adequate to build only a low-resolution image using lower frequencies with a few basis functions at early stage of model building. We then gradually increase the maximum frequency and the number of basis functions as the velocity model becomes closer to the true velocity. Such strategy can also save the computational time by applying the flexible frequency band and the number of basis functions.

Note that multiscale RTM can be a rapid imaging solution, but it consumes more memory to store the projection matrix (Table 2.2). Again, given that the early stage of velocity modeling does not require high frequency to capture all the detailed stratigraphy due to active model updating, we can apply a smaller number of basis with lower source frequency to build the macro structural image. The images and corresponding runtime demonstrates how the frequency-domain multiscale RTM can improve the efficiency of velocity modeling and geological structure imaging.

2.4 Conclusions

The GMsFEM algorithm efficiently simulates waves in the frequency domain by reducing the number of unknowns. We demonstrated the reduction of the impedance matrix size using a smaller number of multiscale basis functions. The projection matrix which consists of multiscale basis functions plays an important role to tune the trade-off point between the accuracy and the computational speed. By allocating appropriate basis functions, the simulation on a coarse-grid can still capture the influence of fine-scale heterogeneity without sacrificing accuracy. We introduced the first results of applying frequency-domain wave modeling using the GMsFEM to RTM, and the migration images show the influence of the number of basis functions to manipulate the calculation speed. The proposed RTM with a small number of basis functions helps to accelerate updating of velocity model by computing large-scale structural image within short amount of time. In addition, using more basis functions directly contributes to enhance the quality of the final RTM images. In this study, we considered the RTM using continuous-Galerkin FEM modeling engine (fine-scale solution) as a reference case. The proposed multiscale RTM algorithm still obtains images more rapidly than when performing the computations on the original fine-scale grid. As a future study, robust performance comparisons between the different imaging algorithms such as finite-difference implementation or ray-based method (i.e., beam-migration) and the GMsFEM RTM need to be made for rigorous demonstration of the efficiency of the multiscale appoach.

3. REVERSE-TIME MIGRATION VIA FREQUENCY-ADAPTIVE MULTISCALE SPATIAL GRIDS*

3.1 Introduction

Reverse-time migration (RTM) is a powerful tool to resolve images in complex subsurface structure (Baysal et al., 1983; McMechan, 1983; Whitmore, 1983; Biondi and Shan, 2002; Symes, 2007) such as salt diapirs, complex faults, or karst structures by overcoming the limits of asymptotic ray theory. However, the RTM, a wave equation based imaging tool, requires considerable computational cost to model the wavefields. In the RTM, the wave simulation for computing both forward and backward propagated wavefields takes the largest portion in computational cost. For accurate imaging and stable wave solutions, the size of the spatial grid needs to satisfy the dispersion condition at the highest frequency component. Accordingly, there has been many research to suppress the numerical dispersion error in the time-domain wave simulation. For instance, Fomel et al. (2013) proposed a method for low-rank wave extrapolation using a two-step time marching approach with a real valued phase function. A low-rank scheme decomposes the original wave propagation matrix into a small set of spatial locations with corresponding representative wave numbers. Here the rank of the approximation tunes the tradeoff between accuracy and computational efficiency. Sun et al. (2015) proposed a one-step wave extrapolation scheme which applies a complex-values low-rank decomposition to approximate the space-wavenumber domain waves. Another method which performs subgrid upscaling with seismic imaging was introduced by Nunes and Minkoff (2014). The operator upscaling algorithm is widely used for the flow simulation (Arbogast, 2003; Arbogast et al., 2002). Vdovina et al. (2009) applied the operator upscaling method for 3-dimensional elastic wave propagation, and Nunes and Minkoff (2014) employed the method for 2-dimensional acoustic wave simulation with the RTM examples.

Although the time-domain wave modeling is widely used, we can suppress the dispersion er-

*Reprinted with permission from "Reverse time migration via frequency-adaptive multiscale spatial grids" by Cho and Gibson, 2019. Geophysics, 84(2), Copyright 2019 by Society of Exploration Geophysicists.

ror in the frequency-domain with less effort. For instance, Hustedt et al. (2004) proposed an extension of the mixed-grid method for accurate wave simulation. Chen (2012) introduced an average-derivative optimal scheme to utilize a rectangular mesh with different grid spacing. Another method, so called compact high-order method (Turkel et al., 2013), is proposed for an accurate and efficient solution of the Helmholtz equation. The compact high-order approach can realize sixth-order accuracy by maintaining the sparsity of the numerical operator (second-order finite difference method). In addition to the easiness of dealing with the dispersion condition, the frequency-domain has advantage in that we can incorporate high-frequency components without considering a reduced time step. Also, the wave solutions of multiple sources can be acquired by solving one-time matrix inversion.

In the frequency-domain wave equation, we can formulate a linear system through a numerical discretization of the given equation. We can take two different methods for solving such a system: a direct solver and an iterative solver. The direct solver is more accurate than the iterative solver, and it can solve multiple sources (right-hand side) at one time. Nevertheless, it requires significant computational cost, which becomes more critical when we solve a 3-dimensional or multi-parameter problems. These days, however, recent advances in direct solvers could mitigate the computational burden by applying low-rank simplifications (Wang et al., 2011; Ghysels et al., 2016) which reduces the cost of LU decomposition. The iterative solver is feasible even in a machine with relatively small memory; however, it is sensitive to the preconditioner (Plessix, 2007). In this research, we will focus on the method to reduce the computational cost of wave modeling and imaging using the direct solver.

Most conventional numerical methods apply the same spatial grid for the entire target frequency band, which might be oversampled for the modeling of low frequency waves. Therefore, in this paper, we applied a model reduction technique, so called Generalized Multiscale Finite Elements Method (GMsFEM), to apply different sizes of spatial grid for different frequency bands. For example, we apply larger spatial grid size to simulate wavefields with long wave length, while applying a finer grid to simulate higher frequency wavefields. The key idea of the proposed method

is that we superpose multiple coarse grids for simulation onto a fixed fine imaging mesh to maintain the dimension of the final RTM image. The GMsFEM carries out simulations on the coarse grid using a small number of finite element basis functions computed on the fine grid for a given earth model.

Considering wave simulations, the GMsFEM (Efendiev et al., 2011; Chung et al., 2014; Gibson and Fu, 2015; Cho et al., 2017b) can be considered as a method similar to that proposed by Nunes and Minkoff (2014) since we also utilize multiscale grids; however, the key difference is that GMsFEM directly computes the basis functions including fine-scale heterogeneity without performing any subgrid upscaling. We solve the wave problem in the frequency-domain to make each frequency component independent. For the Helmholtz equation using GMsFEM, it becomes quite straightforward to utilize various size of spatial grids. In addition, the advantage of working with the frequency-domain is that we can apply different number of eigenfrequencies to capture the detailed feature of wavefields at different frequencies. For example, when we solve a local spectral problem, we consider a smaller number of eigenfrequencies for large wavelengths, and examine a larger number of eigenfrequencies for shorter wavelength features. In the frequency-domain, we can also compute the wave solution through the one-time matrix inversion without considering a reduced time step.

Our previous work (Gao et al., 2015), the time-domain elastic wave modeling using the GMsFEM, shows that lower frequency components can be simulated with a smaller number of multiscale basis functions without sacrificing accuracy. Also, Cho et al. (2017a) showed that GMsFEM approach can enhance the efficiency of seismic imaging by accelerating the wave modeling. Thus, a key goal of this study is to apply GMsFEM wave modeling using frequency-adaptive multiscale spatial grids to compute the RTM image condition. In this paper, we will briefly illustrate an application of GMsFEM to solve the Helmholtz equation by varying the eigenvalue and the coarse grid size, then we show a set of multiscale RTM examples using the Marmousi-2 synthetic model to show how the frequency-adaptive spatial grids can contribute to enhance the efficiency of the RTM.

3.2 Theory and method

To begin, we introduce the acoustic Helmholtz equation as follows:

$$-\frac{\omega^2}{\rho v^2}u = \nabla \cdot (\frac{1}{\rho}\nabla u) + f, \tag{3.1}$$

where ρ and v are the density and velocity, respectively. u is pressure, and f denotes the source term. The angular frequency is represented by ω. The Helmholtz (equation 3.1) models compressional waves in the frequency domain. The key idea of the proposed method is that we vary the sizes of grid cells based on frequency to enhance the efficiency of wave modeling. In the multiscale method, a coarse grid is superposed onto the fine-scale grid occupying the same domain, and material properties are specified on the fine-scale grid.

3.2.1 Generalized multiscale finite element method

We applied the GMsFEM to handle the multiscale mesh Ω that consists of a fine grid h and coarser grids H_i $(i = 1, \cdots, M)$ as shown in Figure 3.2, where M is the number of coarse grids. In GMsFEM, first, we need to make a link between the fine grid and the coarse grids by calculating basis functions in each coarse cell solving local spectral eigenvalue problems to represent fine scale heterogeneity of the property models. The calculation of the basis functions, which is called the offline stage in GMsFEM, need to be performed for each coarse grid that is superposed on the fine grid. The basis functions are determined by the coarse mesh and the material properties, and they are independent of the source and receiver locations. Therefore, we can keep applying the same basis functions once the velocity and density model are determined. To put it differently, the offline stage is performed only once for the given model, and it accounts for a small portion of the total computational effort. Also, the multiscale basis functions are built in each coarse neighborhood K_i $(i = 1, \cdots, N)$ at ith coarse node, a shaded area of the mesh, in Figure 3.2. Then, for the actual wave simulation, each coarse grid utilizes the multiscale basis functions to calculate forward and backward wavefields during the online stage. In summary, the GMsFEM, an efficient model reduction approach, includes two main procedures: offline and online stage. In the offline stage, we

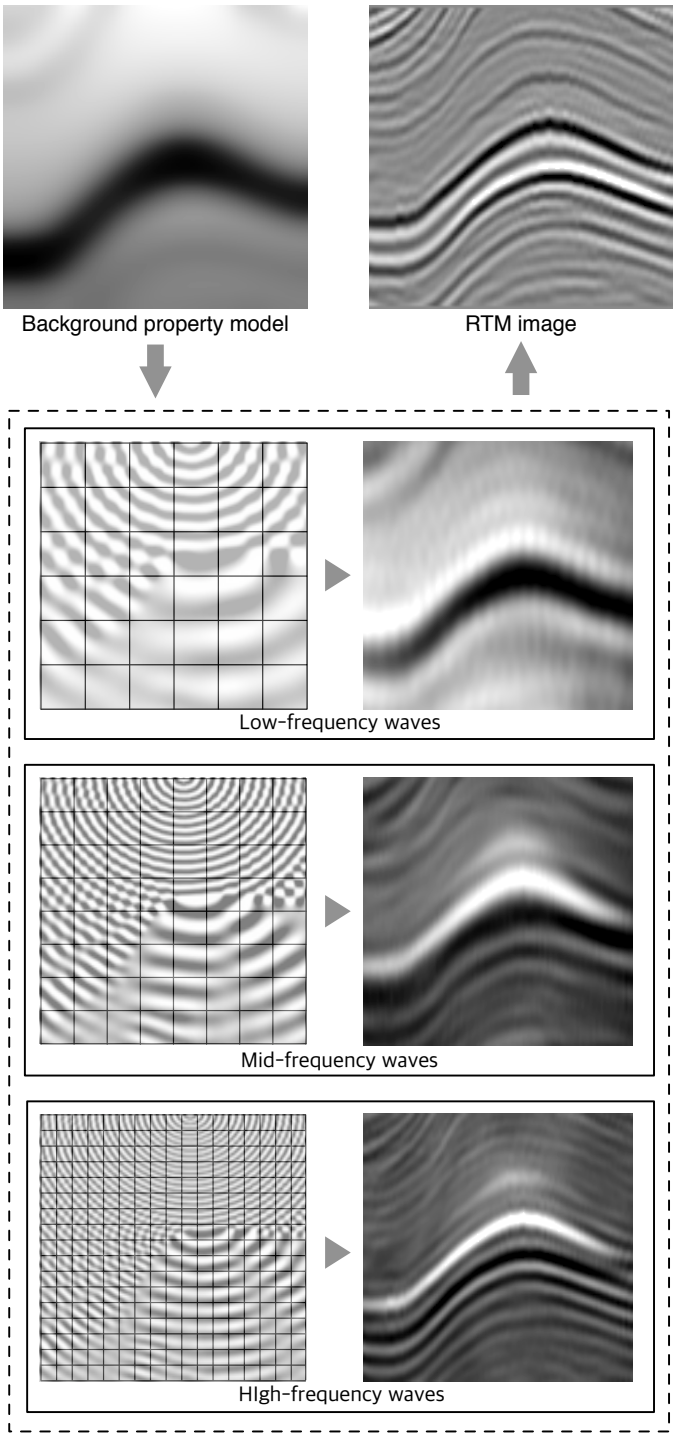
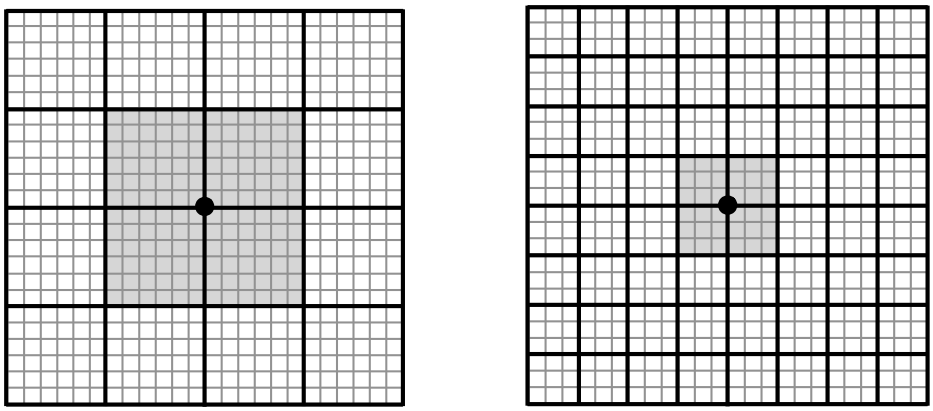


Figure 3.1: Conceptual sketch of the multiscale RTM using the frequency-adaptive spatial grids. Note that the gradient image construction in the box with dashed line is performed independently depends on the frequencies, so we applied different size of coarse mesh for the different frequency bands. The dimension of the background property model, the RTM image, and the gradient images at different frequencies are all identical with the dimension of the fine mesh.



(a) (b)

Figure 3.2: The grids in GMsFEM for different frequency band: (a) A grid for low frequency band, and (b) a grid for high frequency band.

calculate the multiscale basis functions incorporating the heterogeneity of the background property model. These basis functions incorporate the most dominant modes computed from local spectral problems, so utilizing them for the simulation on coarse grid can significantly enhance the computational efficiency. After the construction of multiscale basis functions, in the frequency domain, we implement an impedance matrix inversion to compute the wave solution during the online stage. The amount of reduction in computational cost becomes greater as the frequency decreases since we apply a coarser spatial grid for longer wavelengths. In addition, wave modeling using GMsFEM is more beneficial when we deal with multiple shot-receiver pairs as in RTM case as the basis functions are independent of the source and receiver locations.

3.2.1.1 Construction of multiscale basis functions: Offline stage

The offline stage starts by solving the local spectral problem in each coarse neighborhood K_i (Efendiev et al., 2011) as follows:

$$-\nabla \cdot (\frac{1}{\rho}\nabla\varphi) = \lambda\kappa\varphi, \qquad \kappa = \int_{K_i} \frac{1}{\rho}|\nabla\phi|^2 d\mathbf{x}, \tag{3.2}$$

where λ and φ are the eigenvalue and corresponding eigenfunctions, respectively. When we solve the eigenvalue problem on the fine grid in a coarse neighborhood K_i, the total number of eigenvalues is identical to the number of the degrees of freedom in K_i. ϕ is the function satisfying the condition $-\nabla \cdot (\frac{1}{\rho}\nabla\phi) = 0$ inside of each coarse cell H_i, while the function ϕ is a linear function which starts from 0 and increases linearly to 1, or starts from 1 and decreases to 0 at the boundary of coarse cell ∂H_i. To enhance the accuracy of wave solutions, Fu and Gao (2017) proposed an alternate form of the eigenvalue problem that includes an additional term ψ^2:

$$-\nabla \cdot (\frac{\psi_i^2}{\rho}\nabla\varphi) = \lambda\kappa'\varphi, \qquad \kappa' = \int_{K_i} \frac{\psi_i^2}{\rho}|\nabla\phi|^2 d\mathbf{x}, \tag{3.3}$$

where ψ indicates the element-wise partition of unity (Figure 3.3(a)). Given the solutions for the function ϕ, from either equation 3.2 and 3.3 we can rewrite the generalized eigenvalue problem (equation 3.3) in discrete form as

$$\mathbf{K}\varphi = \lambda\mathbf{M}\varphi, \tag{3.4}$$

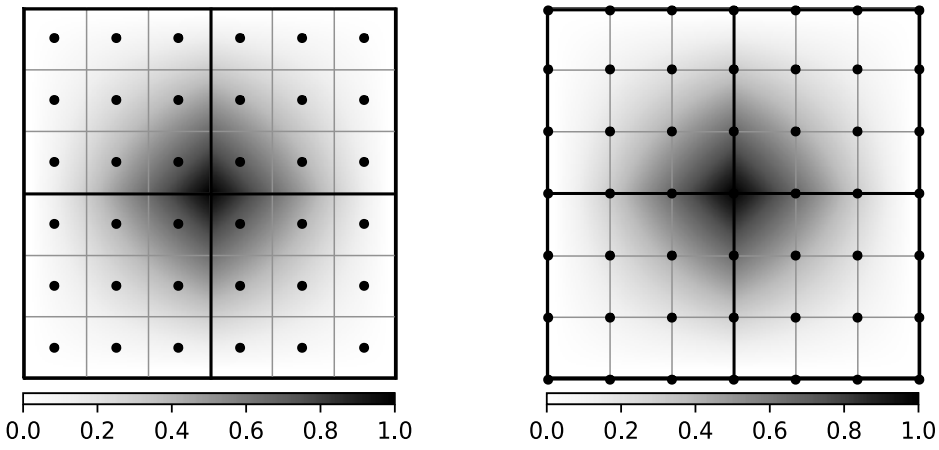
where

$$\mathbf{M} = \int_{K_i} \kappa\zeta\xi d\mathbf{x}, \qquad \mathbf{K} = \int_{K_i} \frac{\psi_i^2}{\rho}\nabla\zeta \cdot \nabla\xi d\mathbf{x}, \tag{3.5}$$

where $\mathbf{M}$ and $\mathbf{K}$ are the mass and stiffness matrices for the given coarse neighborhood K_i. ζ and ξ are the first order polynomial basis functions that are applied on the fine mesh. Below we will compare results for both approaches for computing basis functions. We can calculate a set of eigenvalues with corresponding eigenvectors for each node's coarse neighborhood K_i by solving equation 3.4. After solving the local spectral problem, to define the most dominant wave modes with corresponding eigenfunctions, we sort the eigenvalues in an ascending way for each of K_i as follows:

$$\lambda_i^1 \leq \lambda_i^2 \leq \cdots \leq \lambda_i^{L_i} \leq \cdots \leq \lambda_i^{N_h}, \tag{3.6}$$

where, N_h and L_i are the number of fine nodes within a coarse neighborhood K_i and the number of multiscale basis functions that are required for wave simulations, respectively. The key idea of



(a) (b)

Figure 3.3: Partition of unity being discretized in different method: (a) Element-wise partition of unity, and (b) nodal point-wise partition of unity.

GMsFEM is that we only consider a small number of modes to capture the details of wavefields more efficiently without loosing accuracy. Therefore, when we simulate waves in a low frequency band, we can reduce the computational cost by applying L_i basis functions that is enough to capture the macro variation of the waves with long wavelength. In high frequency bands, however, we can tune the accuracy by increasing the number of basis functions L_i since larger eigenvalues correspond to higher modes that incorporate more detailed variation of the wave solutions. In short, the first L_i eigenfunctions φ_i^j, $(j = 1, \cdots, L_i)$ that are required in the GMsFEM solver are controlled by the level of heterogeneity on the fine-scale grid. As a final step of building basis functions at K_i, we defined a nodal point based partition of unity ψ_i' (Figure 3.3(b)) since the eigenfunctions are not globally continuous. We obtain multiscale basis functions Φ_i through the multiplication of φ_i with the partition of unity ψ_i' as follows: $\Phi_i = \varphi_i \odot \psi_i'$ (Babuška et al., 1995), where $\odot$ denote the element-wise multiplication. The above equations illustrate the method to calculate basis functions in a fixed coarse grid. We apply the same relations for computing the set of multiscale basis functions for the different coarse grids as well. To do so, we define a multiscale

basis space in the coarse grid as

$$V_k^H = \text{span}\left\{\Phi_j^i | 1 \leq i \leq N, 1 \leq j \leq L_i\right\}, \tag{3.7}$$

where V_k^H is the coarse mesh approximation space in the kth coarse grid. In a fixed fine-scale dimension of Ω, the total number of basis functions N becomes smaller as the size of coarse cell H_i becomes larger. Also, we can apply smaller number of multiscale basis functions L_i in a smaller size of coarse cell.

3.2.1.2 Wave modeling in the coarse grids: Online stage

In the online stage, we perform the actual wave modeling, and the multiscale basis functions being built in the offline stage are used for the wave simulation. We follow the approach summarized by Marfurt (1984) for the discretization of the Helmholtz equation,

$$\mathbf{S}_k^H \mathbf{u}_k^H = \mathbf{f}_k^H, \tag{3.8}$$

where the impedance matrix $\mathbf{S}_k^H$ at the kth coarse grid can be expanded as follows: $\mathbf{S}_k^H = \mathbf{K}_k^H + i\omega\mathbf{C}_k^H + \omega^2\mathbf{M}_k^H$. $\mathbf{K}_k^H$ and $\mathbf{M}_k^H$ are stiffness and mass matrices. $\mathbf{u}_k^H$ and $\mathbf{f}_k^H$ are the pressure field and source term, respectively. The coarse impedance matrix also includes the damping term $\mathbf{C}_k^H$ to suppress the reflection of outgoing wavefield at the model boundary $\partial\Omega$. In this paper, we employed a Perfectly Matched Layer (PML) absorbing boundary condition, and the detailed implementation of the damping term in the GMsFEM is introduced by Fu and Gao (2017).

We can conceptually express the coarse-scale impedance matrix and the source term at the ith coarse node in an arbitrary coarse grid by using the multiscale basis functions Φ_i as

$$\begin{aligned}
\mathbf{S}_i^H &= \sum_{p,q} \int_{K_i} \left[-\frac{\omega^2}{\rho v^2} \Phi_i^p \Phi_i^q + \frac{1}{\rho} \nabla \Phi_i^p \cdot \nabla \Phi_i^q \right] d\mathbf{x}, \\
\mathbf{f}_i^H &= \sum_{p} \int_{K_i} f \Phi_i^p d\mathbf{x},
\end{aligned} \tag{3.9}$$

where ω is angular frequency. However, instead of building the coarse-scale mass and stiffness matrices explicitly, we define a projection matrix to $\mathbf{R}$ by using the multiscale basis functions and coarse grid information. It makes the coarse grid simulation represent the fine-scale heterogeneity, and the global projection matrix $\mathbf{R}$ can be expressed as

$$\begin{aligned} \mathbf{R} &= \left(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_N\right)^{\mathrm{T}}, \\ \mathbf{R}_i &= \left(\Phi_1, \Phi_2, \ldots, \Phi_{L_i}\right), \end{aligned} \tag{3.10}$$

where the column of the $\mathbf{R}_i$ matrix are the discrete multiscale basis functions Φ_j from each K_i. The row dimension of the global $\mathbf{R}$ matrix represents the total number of nodal points in the fine grid, while column dimension is identical to the product of the number of multiscale basis functions and the nodal points in the coarse mesh. Note that the projection matrix $\mathbf{R}$ is determined by the coarse space V_k^H and the number of basis functions L_i. Thus, we need to obtain the projection matrix for each kth coarse grid, then directly calculate the coarse-scale mass and stiffness matrices from the fine-scale matrices. Therefore, we can rewrite the ith multiscale basis function Φ_i in the space domain with coarse grid V_k^H as $\Phi_i = \sum_j r_{ij}\phi_j$, where r_{ij} is the coefficient and ϕ_j is a piece-wise linear basis functions in the fine-scale domain V^h. The coarse impedance matrix then can be written as

$$\begin{aligned} \mathbf{S}_{i,j}^H &= \int_\Omega \left[-\frac{\omega^2}{\rho v^2}\Phi_i\Phi_j + \frac{1}{\rho}\nabla\Phi_i \cdot \nabla\Phi_j \right] d\mathbf{x}, \\ &= \sum_{m,n} r_{mi} \left[\int_\Omega -\frac{\omega^2}{\rho v^2}\phi_m\phi_n d\mathbf{x} + \int_\Omega \frac{1}{\rho}\nabla\phi_m \cdot \nabla\phi_n d\mathbf{x} \right] r_{jn}, \end{aligned} \tag{3.11}$$

and the source vector can be set as

$$\mathbf{F}_j^H = \int_\Omega f\Phi_j d\mathbf{x} = \sum_l r_{jl} \int_\Omega f\phi_l d\mathbf{x}, \tag{3.12}$$

where, r_{ij} is the component of the projection matrix $\mathbf{R}$. We can express the fine-scale impedance

matrix as $\mathbf{S}^h = s^h_{mn}$ and $s_{mn} = s(\phi_m, \phi_n)$, we then have

$$\begin{aligned} \mathbf{S}^H &= \mathbf{R}^{\mathrm{T}} \mathbf{S}^h \mathbf{R}, \\ \mathbf{f}^H &= \mathbf{R}^{\mathrm{T}} \mathbf{f}^h. \end{aligned} \tag{3.13}$$

After obtaining final solutions in the coarse mesh, we can simply recover the fine-scale solutions through the multiplication of the projection matrix: $\mathbf{u}^h = \mathbf{R}\mathbf{u}^H$. In general, wave modeling in the frequency domain requires a large computational cost due to saving a large impedance matrix in a memory space and calculation of the matrix inversion. As an alternative, when we construct the impedance matrix on the coarse grid using multiscale basis functions, we can dramatically reduce the number of unknowns in matrix operations compared to the fine-scale problem. When we deal with 3-dimensional problems in the GMsFEM, we still solve the same local eigenvalue problems to obtain the multiscale basis functions. However, the memory consumption might increase significantly when we store the multiple projection matrices $\mathbf{R}_i$. Hence, in this case, we may consider an alternative proposed by Gao et al. (2018), which is a high-order multiscale finite element approach. This method helps reduce the memory consumption since it does not require matrix projection since it computes the coarse-scale matrices explicitly.

When we build the mass and stiffness matrices, the size of the grid satisfying the dispersion condition is controlled by the maximum frequency, as it requires a larger number of sampling points per wavelength. Therefore, when we use a conventional approach with a fixed size of spatial grid for the entire frequency band, the simulation of low frequency waves is performed in the mesh that is too fine compared to the wavelength. To overcome this problem, we designed a multiscale mesh (Figure 3.2) by varying the size of the spatial grid depending on the frequency bands. A central task in utilizing the frequency-adaptive spatial grid is that we construct a different projection matrix $\mathbf{R}$ for each pair of the coarse- and fine-scale meshes. We consider a fine grid h with 101 by 101 dimension (10 m interval) to demonstrate the reduction of the impedance matrix size obtained by this adaptive approach. We also compare two coarser grids H_1 and H_2 for different frequency bands on the fine grid h, where H_1 (100 m) and H_2 (50 m) are for simulating

waves in low- and mid-frequency band, respectively. In this case, the dimension of the fine-scale impedance matrix $\mathbf{S}^h$ is $10,201^2$. Given the size of each coarse grid, the dimension of the projection matrix for each grid can be as $\mathbf{R}_1[10,201 \times 121L_i]$ and $\mathbf{R}_2[10,201 \times 2,601L_i]$, respectively. After applying these projection matrices for each coarse grid (equation 3.13), we then acquire the coarse-scale impedance matrices. The dimension of $\mathbf{S}_1^H$ and $\mathbf{S}_2^H$ become $(121L_i)^2$ and $(2,601L_i)^2$, respectively. Setting an appropriate number of multiscale basis functions L_i plays an important role to effectively reduce the numerical error in a fixed size of coarse grid. Considering the amount of reduction in matrix size, we can roughly estimate how much we can enhance the efficiency of the computation. Note that utilizing more multiscale basis functions L_i increases the dimension of coarse-scale matrices, but we still can attain the impedance matrices which have much smaller dimension than the fine-scale impedance matrix (Cho et al., 2017a). To obtain accurate solutions from the GMsFEM, we need to apply a sufficient number of basis functions to incorporate detailed feature of wavefields into the solutions; however, this number decreases as the cell size decreases. For example, if 10 basis functions are required to obtain accurate solution in H_1, we can calculate accurate wave solutions in H_2 only by using 6 or 7 basis functions.

3.2.2 Multiscale RTM imaging condition

A popular imaging condition for RTM is to calculate zero-lag cross correlation between shot- and time reversed receiver-domain wavefields (Baysal et al., 1983). However, this imaging condition results in additional computational cost when we work in the frequency-domain because of the repeated Fourier transform. Thus, we used an imaging condition using zero-lag cross-correlation between the Born wavefields and the observed seismic data (Chavent and Plessix, 1999). Brigham (1988) showed that the RTM imaging condition can be obtained throughout the Fourier transform pairs in the frequency domain as follows:

$$\phi_k = \sum_{i=1}^{N_s} \int_0^{\omega_{\max}} \Re \left\{ \left[\frac{\partial \mathbf{u}_i(\omega)}{\partial m_k} \right]^{\mathrm{T}} \mathbf{d}_i^*(\omega) \right\} d\omega, \tag{3.14}$$

where ϕ_k denotes the image value at the kth model parameter. $\mathbf{u}_i$ and $\mathbf{d}_i^*$ indicate the forward modeled wavefield and measured data, and superscript * means the complex conjugate in the frequency domain.

Before deriving the imaging condition using a coarse grid wave simulation, we first take the partial derivative of the Helmholtz equation with respect to the kth model parameter m_k in the fine-scale grid as

$$\frac{\partial \mathbf{S}^h}{\partial m_k}\mathbf{u}^h + \mathbf{S}^h\frac{\partial \mathbf{u}^h}{\partial m_k} = 0, \tag{3.15}$$

then we can rearrange above equation as

$$\frac{\partial \mathbf{u}^h}{\partial m_k} = [\mathbf{S}^h]^{-1}\left[-\frac{\partial \mathbf{S}^h}{\partial m_k}\mathbf{u}^h\right], \tag{3.16}$$

where we can rewrite the term consisting of the Born modeling operator and the background wavefields as $\mathbf{v}$ for conciseness, which is referred to as a virtual source vector. Therefore, the Born wavefield can be also expressed as $[\mathbf{S}^h]^{-1}\mathbf{v}$. The Born wavefield $\frac{\partial \mathbf{u}^h}{\partial m_k}$ can be calculated by using the forward modeling algorithm, which is an impedance matrix inversion (Pratt et al., 1998; Shin et al., 2007) in the frequency domain. By applying the projection matrix $\mathbf{R}$, we define the Born wavefield in the coarse grid as

$$\frac{\partial \mathbf{u}^H}{\partial m_k} = [\mathbf{R}^{\mathrm{T}}\mathbf{S}^h\mathbf{R}]^{-1}\mathbf{v}. \tag{3.17}$$

We can derive the final multiscale imaging condition by plugging in equation 3.17 into equation 3.14 as

$$\phi_k = \sum_{i=1}^{N_s}\int_0^{\omega_{\max}} \Re\left[\mathbf{v}^{\mathrm{T}}\mathbf{R}(\mathbf{S}^H)^{-1}\mathbf{R}^{\mathrm{T}}\mathbf{d}_i^*(\omega)\right] d\omega, \tag{3.18}$$

where N_s is the total number of shots. We can calculate the image value for the kth model parameter by applying equation 3.18. Note that even though we performed the wave simulation in the coarse grid, the dimension of the RTM image is the same as that of the background property model (or fine mesh). After calculating the wave solutions, we project the coarse grid wavefields to the fine grid by using $\mathbf{R}$ before applying the virtual source vector. Therefore, the final multiscale RTM

result would have the same resolution with the reference RTM image. Based on equation 3.18, we can develop the imaging condition which includes multiple coarse grids and corresponding projection matrices as

$$\phi_k = \sum_{i=1}^{N_s} \sum_{j=1}^{N_b} \int_{\omega_{j-1}}^{\omega_j} \Re \left[\mathbf{v}^{\mathrm{T}} \mathbf{R}_j (\mathbf{S}^H)^{-1} \mathbf{R}_j^{\mathrm{T}} \mathbf{d}_i^*(\omega) \right] d\omega, \tag{3.19}$$

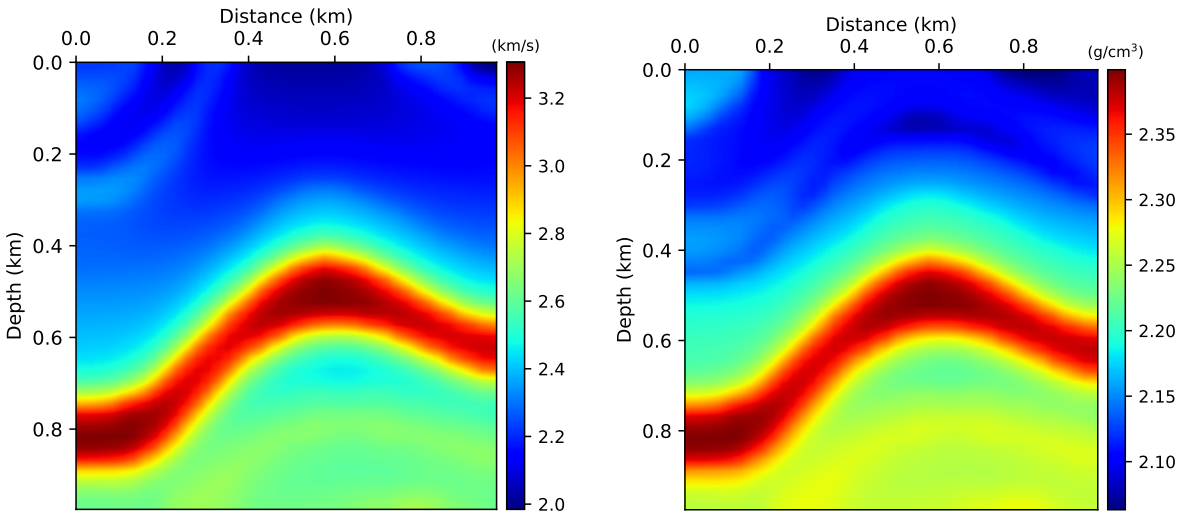
where ϕ_k denotes the image value at kth model parameter. N_s is the total number of shots. N_b is the selected number of frequency bands – for example, when we simply divide the target frequency band into low, middle, and high frequencies, then N_b is three. Also, each frequency band will be allocated to a different coarse grid. The imaging condition (equation 3.19) does not take into account incident wave energy, and this problem becomes more critical when we consider cases with large impedance contrast such as salt diapir or karst structures. As a solution, Claerbout (1971) proposed a source illumination term to recover clearer images at larger depth by scaling the RTM image. In addition, inspired by Chavent and Plessix (1999)'s work showing the migration image can be improved by dividing with the Hessian matrix, we applied a diagonal of the pseudo-Hessian matrix (Shin et al., 2001) to reduce the computational cost as follows:

$$\phi = \frac{\sum_{i=1}^{N_s} \sum_{j=1}^{N_b} \int_{\omega_{j-1}}^{\omega_j} \Re \left[\mathbf{f}_{\mathbf{v}}^{\mathrm{T}} \mathbf{R}_j (\mathbf{S}^H)^{-1} \mathbf{R}_j^{\mathrm{T}} \mathbf{d}_i^*(\omega) \right] d\omega}{\sum_{i=1}^{N_s} \sum_{j=1}^{N_b} \int_{\omega_{j-1}}^{\omega_j} \Re \left[\mathrm{diag}(\mathbf{f}_{\mathbf{v}}^{\mathrm{T}} \mathbf{f}_{\mathbf{v}}) \right] d\omega + \lambda}, \tag{3.20}$$

where $\mathbf{f}_{\mathbf{v}}$ is a matrix form of the virtual source vectors $\mathbf{v}$ to represent the entire target image domain. λ is a damping factor to stabilize the numerical operation, and $\mathbf{f}_{\mathbf{v}}^{\mathrm{T}} \mathbf{f}_{\mathbf{v}}$ is the approximated-Hessian matrix. By taking the diagonal component of the approximated-Hessian, we implement the pseudo-Hessian matrix. By applying equation 3.20, we calculate the RTM imaging condition more efficiently combined with a frequency-adaptive multiscale spatial grids.

3.3 Numerical examples

In this section, we first demonstrate how the GMsFEM performs wave modeling by varying the number of basis functions and the size of the coarse mesh. Then we show the RTM examples using



(a) (b)

Figure 3.4: A small size (40×40) property model; (a) velocity and (b) density, to demonstrate the difference of local spectral problems in accuracy.

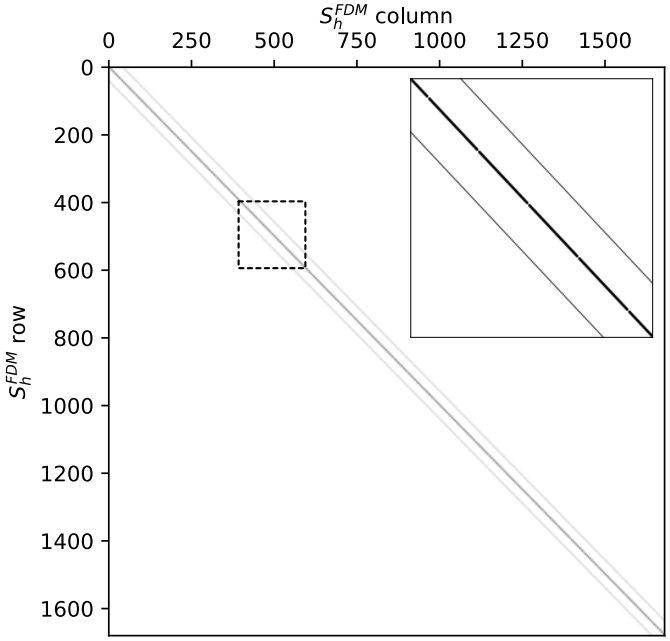
GMsFEM for the forward- and backward-modeling with various combinations of coarse and fine meshes. For generating a synthetic dataset to test the RTM, we applied the continuous-Galerkin finite element method (CG FEM), while GMsFEM is employed to calculate the imaging condition of the multiscale RTM.

3.3.1 Comparison of the local spectral problems

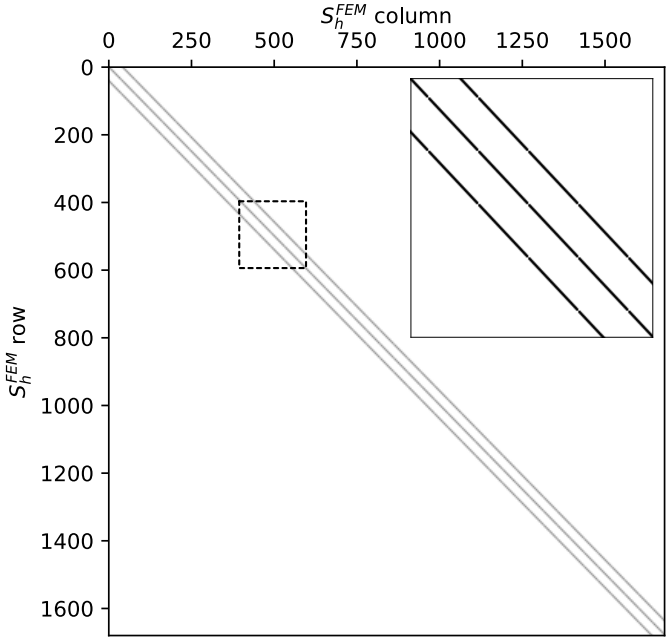
We created a heterogeneous property model ($n_x = n_y = 40$) that includes an anticline structure (Figure 3.4). When the model consists of 40×40 elements, the dimension of the fine-scale impedance matrix $\mathbf{S}^h$ becomes 1681×1681 as shown in Figure 3.5. We compared the impedance matrices that are generated from the finite difference method (FDM) and the finite element method (FEM). We used the second-order FDM and FEM in each case. The FDM operator utilizes four grid points (above, below, left and right of grid point of interest) to solve the PDEs. In contrast, the continuous-Galerkin FEM also utilizes the grid points located on diagonals passing through the grid point of interest, so that values on a total of eight locations are used. Therefore, the impedance matrix of FEM has more off-diagonal components (Figure 3.5(b)) than that of FDM case (Figure

3.5(a)).

As the model size becomes larger, the size of the fine-scale impedance matrix exponentially increases. Therefore, in this research, we apply GMsFEM to reduce the number of unknowns in the impedance matrix inversion. We superposed a 4×4 coarse mesh onto the fine mesh, where a single coarse cell has 100 fine cells. We then need to build a projection matrix which consists of the multiscale basis function that are obtained by solving the local eigenvalue problems (equation 3.3). Figure 3.6 exhibits the projection matrices (equation 3.10) from different local spectral problems. The multiscale basis functions that are calculated by equation 3.2 yield the projection matrix shown in Figure 3.6(a), while equation 3.3 provides the projection matrix which is displayed in Figure 3.6(b). The projection matrix that is calculated from the local eigenvalue problem incorporating the squared partition of unity has a more evenly distributed variation of values than the other one. The influence of this on the numerical performance will be shown below. The row dimension of the projection matrix follows the total number of fine-scale nodes, while the column dimension is governed by the coarse mesh nodes ($H_N = 25$). In this case, we used 10 basis functions ($L_i = 10$). As we apply more basis functions to capture more detailed wavefields, the column dimension ($H_N L_i$) of the projection matrix increases. In Figure 3.7, the inset detailed image of the coarse-scale impedance matrix is displayed for each case. In addition to the evenly distributed magnitudes, when we applied the projection matrix from the spectral problem including the partition of unity, the coarse impedance matrix has more weight on the diagonal components (Figure 3.7(b)). In contrast, the diagonal components of the coarse-scale impedance matrix shown in Figure 3.7(a) are hidden by the surrounding values. We utilize the Marmousi-2 model (Figure 3.8) for calculating frequency-domain wavefields to show how the two approaches to solve the local spectral problems affect the wave solution. For a mono-frequency source at 15 Hz, we calculated a reference wavefield (Figure 3.9(a)) using fine-scale impedance matrices, which took 1.273 s in run-time with a single processor unit. The wave solutions from the GMsFEM are displayed in Figure 3.9(b) and 3.9(d), which took 0.194 s for computing time. Through the GMsFEM, we could compute the solution 6.5 times faster than the reference (CG-FEM). The speedup does not

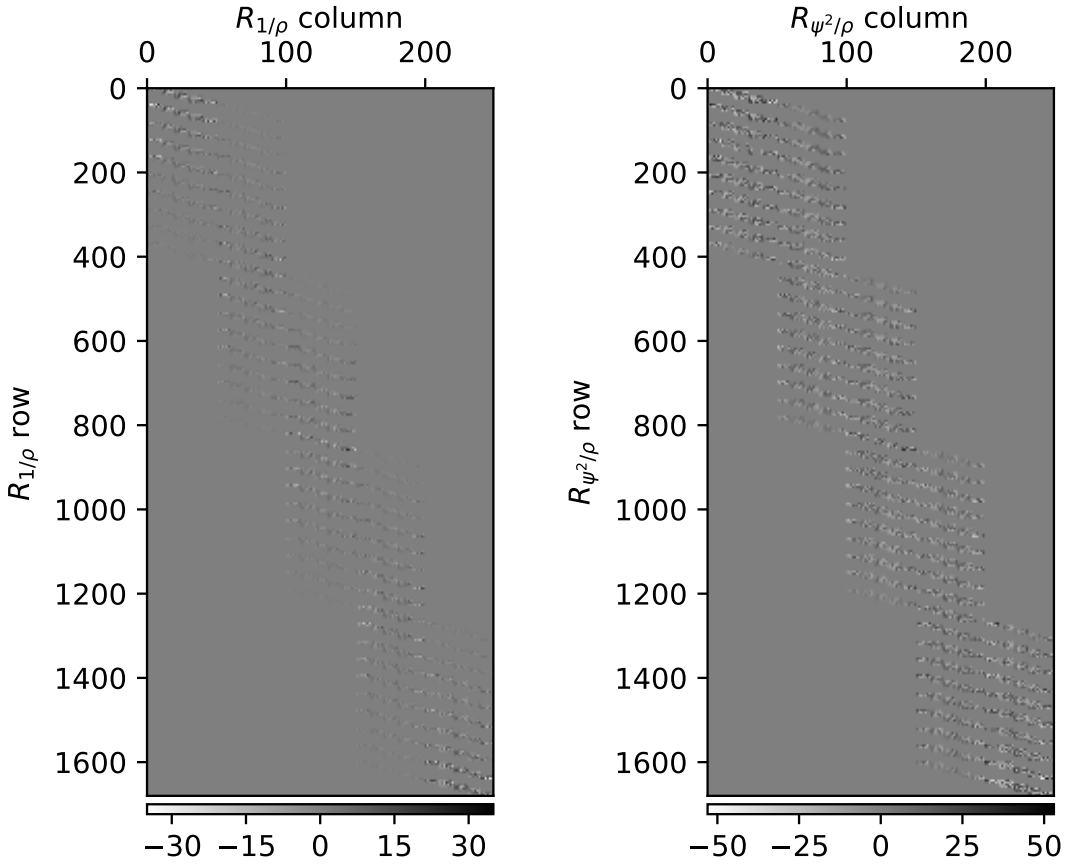


(a)



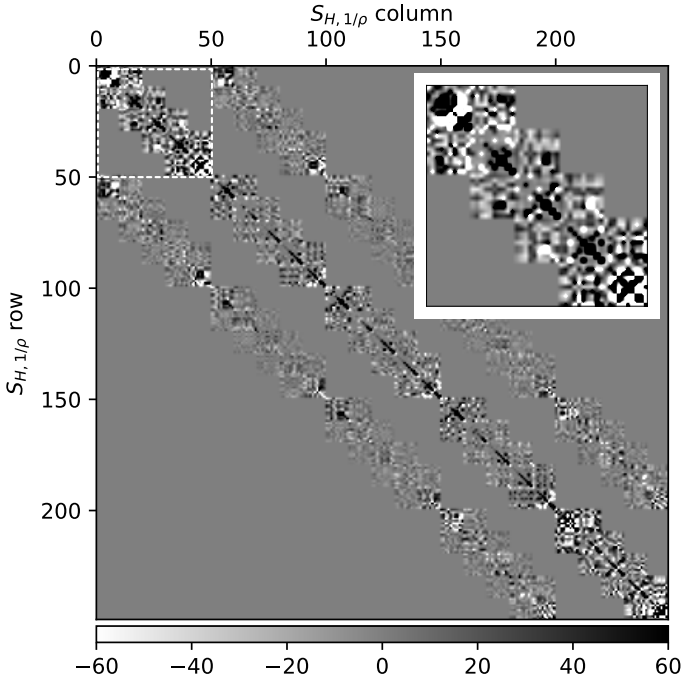
(b)

Figure 3.5: The fine-scale impedance matrices $\mathbf{S}^h$ from the property models shown in Figure 3.4 using two different discretization method: (a) FDM and (b) FEM. The impedance matrix built via FEM contains a greater number of off-diagonals.

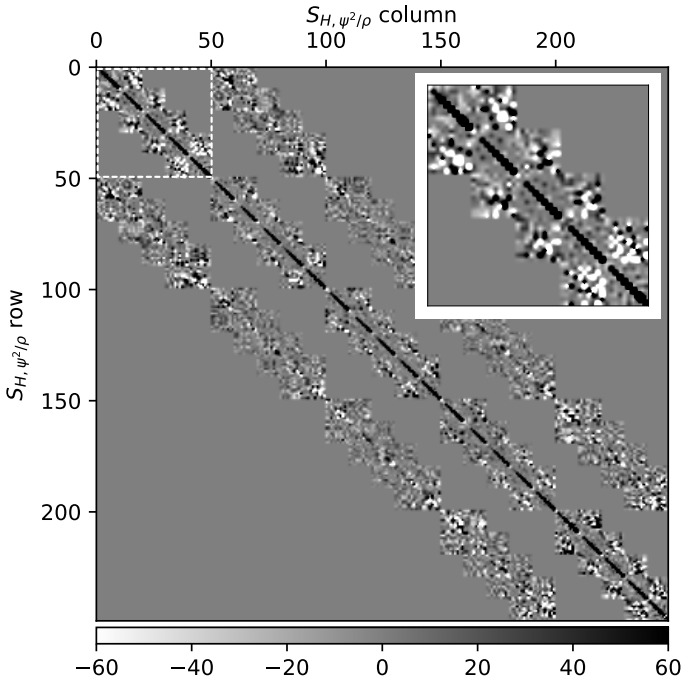


(a) (b)

Figure 3.6: The projection matrices that are calculated from different eigenvalue problems: (a) $-\nabla \cdot (\frac{1}{\rho}\nabla\varphi) = \lambda\kappa\varphi$ and (b) $-\nabla \cdot (\frac{\psi_i^2}{\rho}\nabla\varphi) = \lambda\kappa'\varphi$.

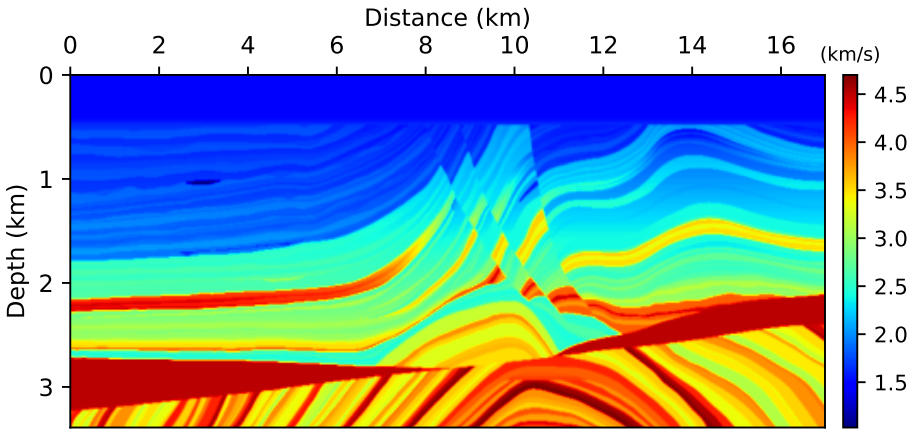


(a)

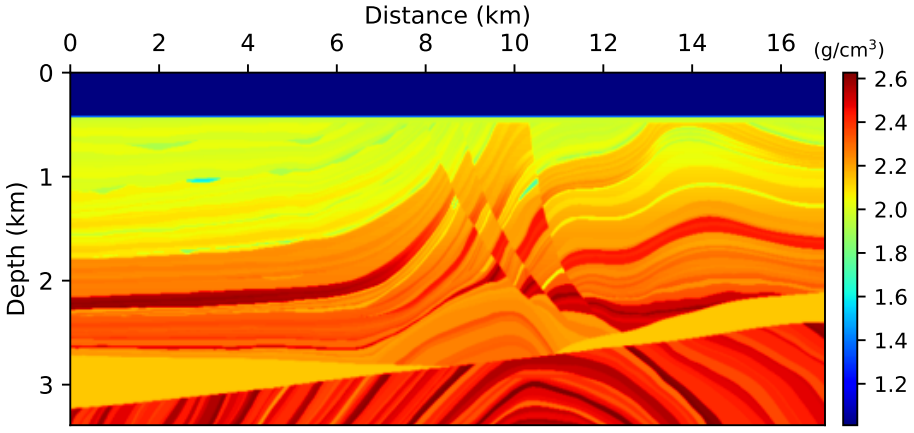


(b)

Figure 3.7: The coarse-scale impedance matrices from the different projection matrices incorporating distinct local spectral problems: (a) $\mathbf{S}^H = \mathbf{R}_{1/\rho}^{\mathrm{T}} \mathbf{S}^h \mathbf{R}_{1/\rho}$ and (b) $\mathbf{S}^H = \mathbf{R}_{\psi^2/\rho}^{\mathrm{T}} \mathbf{S}^h \mathbf{R}_{\psi^2/\rho}$.



(a)



(b)

Figure 3.8: Marmousi-2 model: (a) Velocity, and (b) density models.

proportionally increase as the dimension of the coarse impedance matrix decreases, because we use a sparse direct solver which takes only non-zero values and coordinates of the components into consideration. One might argue that the coarse impedance matrix requires longer computational time to factorize the matrix, as it has more a complicated geometry than the fine-scale one which is almost tridiagonal. Nevertheless, as the matrix factorization is required only once when we construct the coarse-scale matrices, and we can compute results for multiple source locations (or RHS) with a single matrix inversion. Therefore, this task contributes little effort to the total computation.

When we compare the run time between Figure 3.9(b) and 3.9(d), there is no difference since the size of the matrix is the same; however, we demonstrate that the spectral problem incorporating the element-wise partition of unity yields a more accurate solution by subtracting the GMsFEM solutions from the reference solution as shown in Figure 3.9(c) and 3.9(e). Note that the magnitudes of the difference are displayed in the 5% of the color scale of the snapshots in Figure 3.9. Most of the error is located around the source location, and the oscillatory numerical artifacts diminish as the horizontal distance becomes further from the source point. This numerical error which is smaller than 1% of the original amplitude of the waves is negligible when we apply the GMsFEM to the RTM case due to destructive interference among the multiple sources. Given that including the partition of unity in the spectral problem puts more weight on the diagonal component of the coarse impedance matrix, we also tried to apply the weight (lumped matrix) by force; however, weighting the diagonal component of the coarse-scale matrix did not result in any meaningful enhancement of the accuracy. Cho et al. (2017a) introduced the multiscale RTM examples by using the GMsFEM with the basis functions from equation 3.2 (used in Figure 3.9(b)); however, in this research, we applied the local eigenvalue problem including the partition of unity (Figure 3.9(d)) term as it helps to improve the accuracy of the GMsFEM solutions.

Before applying the frequency-adaptive spatial grid to the frequency-domain wave modeling, we first made comparisons between the performance of different numerical methods: FDM, FEM, and GMsFEM, for the waves from a mono-frequency single shot. For each approach, we located the wave source at the center of the model, and calculated wavefields by applying two different

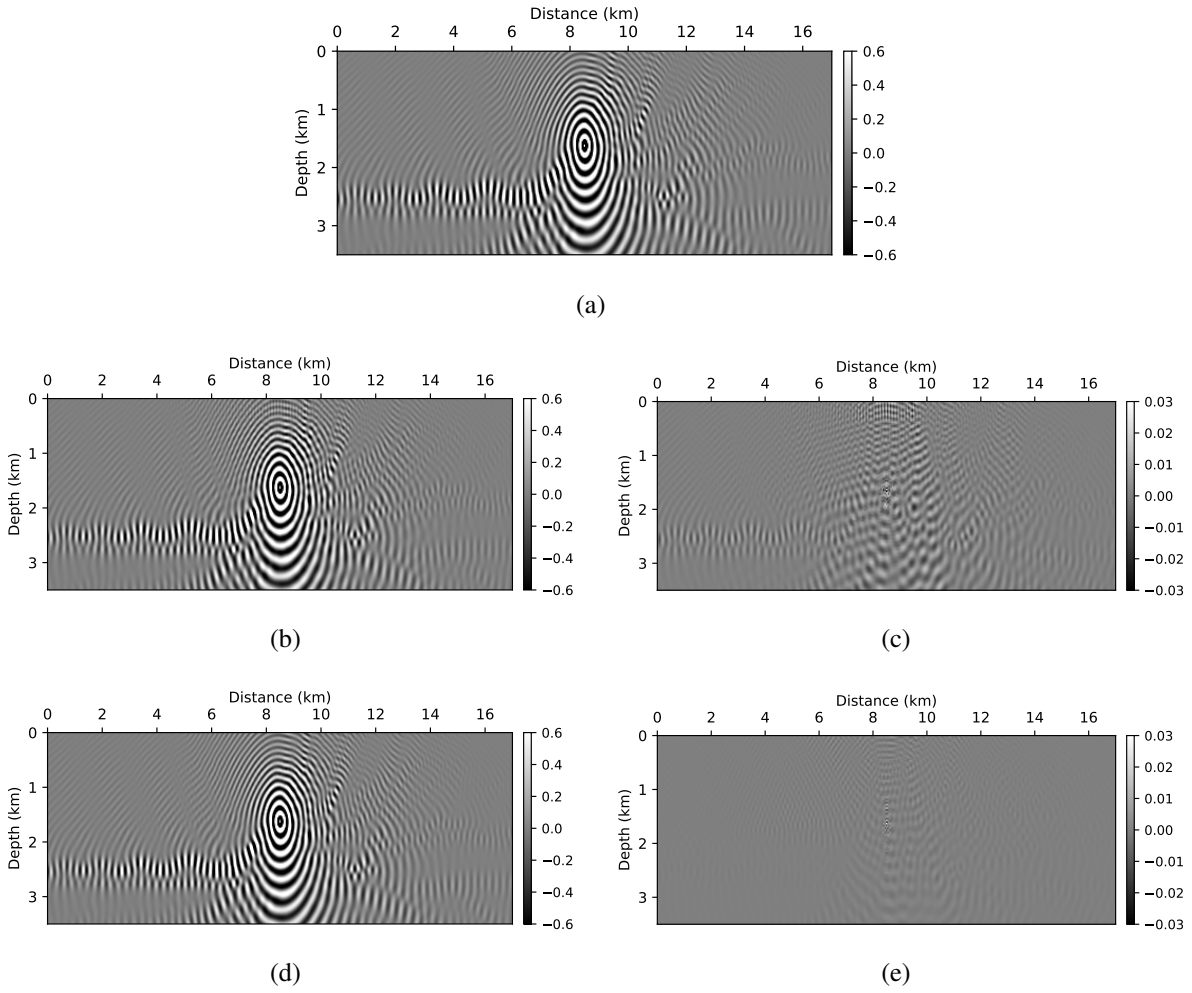


Figure 3.9: The wave solutions of a mono-frequency source (15 Hz) that are calculated from (a) fine-scale impedance matrix and coarse-scale impedance matrix shown in (b) Figure 3.7(a) and in (d) Figure 3.7(b). Figure (c) and (e) exhibit the difference between the reference and Figure (b) and (d), respectively. The eigenvalue problem including the partition of unity term (equation 3.3) provides a better projection matrix for calculating more accurate wave solutions.

source frequencies (5 Hz and 15 Hz) as displayed in Figure 3.10. In the 5 Hz examples (Figure 3.10(a), 3.10(c), 3.10(e), and 3.10(g)), all the wavefield snapshots are almost identical ($\varepsilon_{L^2} < 0.03$). In contrast, for the 15 Hz case, we obtain dispersive artifacts in the snapshot from the GMsFEM with seven basis functions (Figure 3.10(f)), which means that seven basis functions are not enough to capture all the details of the waves with 15 Hz frequency. Nevertheless, we can suppress the artifacts by applying more basis functions, and the GMsFEM ($L_i = 9$) still provides the solutions faster than the reference case. The key strategy of the GMsFEM is to utilize the basis functions and grid size flexibly according to the frequencies of interest. The corresponding L^2-error is displayed in Table **??** to demonstrate the accuracy of the solutions. Here, we used the fine-scale FEM solution as a reference. To quantify the computational efficiency, we presented the run-time of each numerical method in Table **??**. The run-time is divided into three parts: matrix factorization, matrix projection, and matrix inversion. Note that the process of the matrix projection (equation 3.13) is only performed in the GMsFEM case. As demonstrated in Figure 3.5, as the FEM implies more off-diagonals than the FDM, the FEM shows slower computing speed than the FDM. Given the total computing time, the FDM could simulate a single shot 30 % faster than the FEM. The GMsFEM, however, provides the fastest solutions even though it requires additional matrix operations. Comparing only the run-time of the matrix inversion, the GMsFEM ($L_i = 9$) can achieve significant speedup (21 and 11 times faster than FEM and FDM, respectively) due to the substantially reduced dimension of the impedance matrices.

3.3.2 Wave modeling using frequency-adaptive grids

Here we demonstrate how the numerical error of the GMsFEM behaves differently depending on the given parameters: the size of the cell in the coarse mesh, frequency bands, and the number of multiscale basis functions. First, we generated a shot gather by using the CG FEM ($h = 10$ m) as shown in (Figure 3.11). By taking this shot profile as a reference, we compared differences between the reference spectrum and the frequency spectra of the GMsFEM solutions for different numbers of basis functions (L_i =6, 8, 10 and $H = 100$ m) as displayed in Figure 3.12. Note that the size of coarse cell is fixed to 100 m in this example. We measured the run-time using a single computing

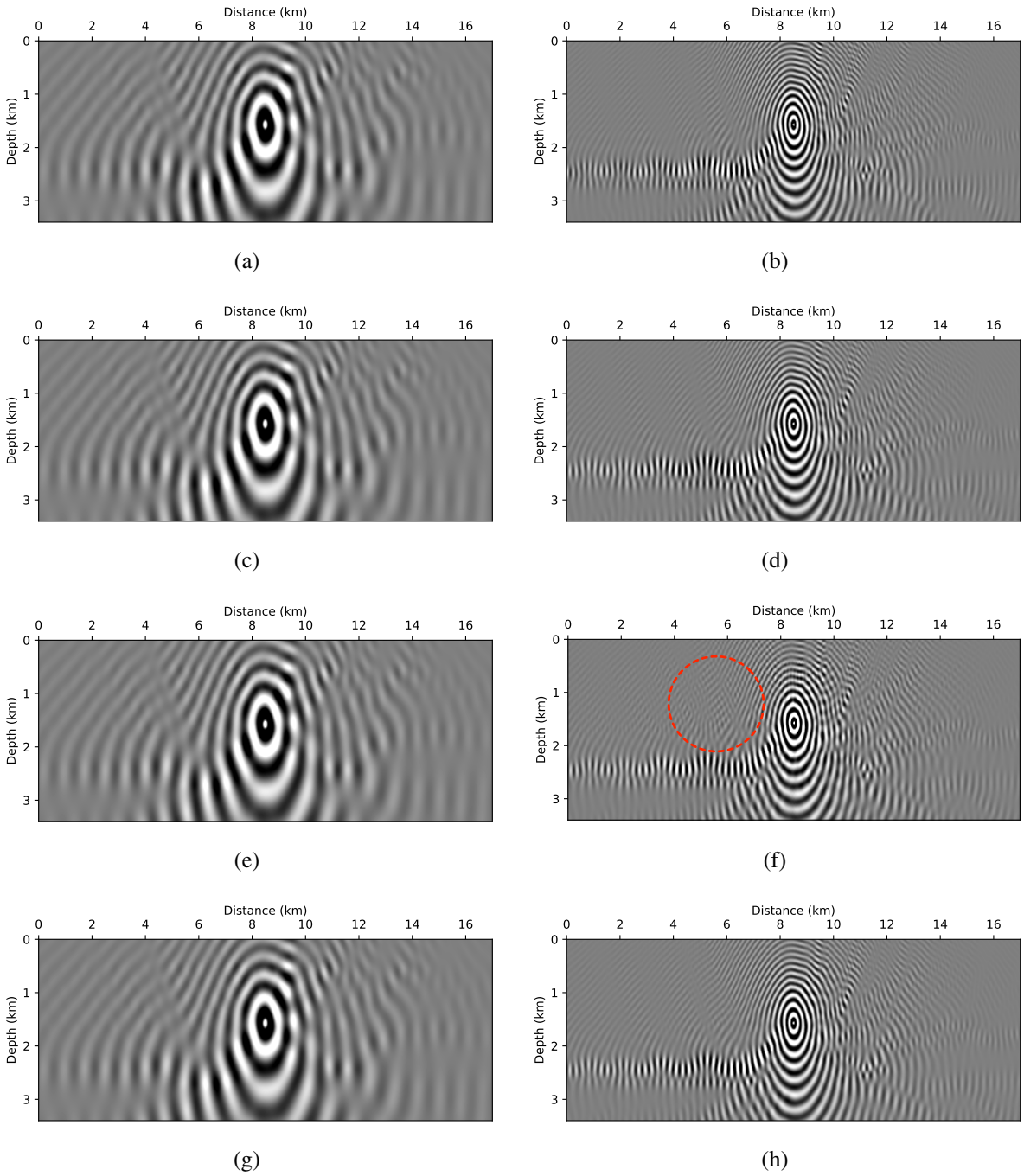


Figure 3.10: Comparisons of the wavefields from varied numerical methods: FDM (a, b), FEM (c, d), GMsFEM with $L_i = 7$ (e, f), and GMsFEM with $L_i = 9$ (g, h). The waves in left and right panel of the figure shows 5 Hz and 15 Hz waves, respectively. The dashed red circle in Figure (f) points the area with dispersive artifacts caused by the insufficient multiscale basis functions.

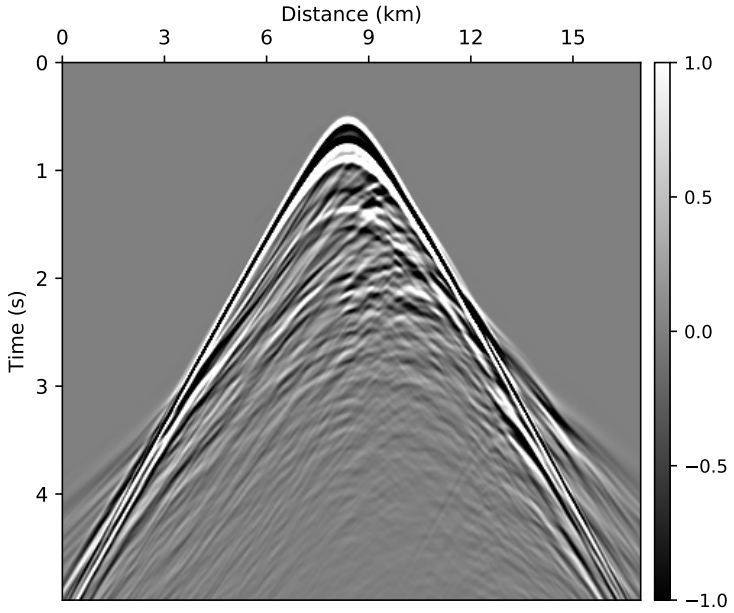
Method	t_F	t_P	t_I	t_{total}	ε_{5Hz}	ε_{15Hz}
FEM	7.6818	-	2.3246	10.006	-	-
FDM	5.9371	-	1.3408	7.2779	-	-
GMsFEM ($L_i = 7$)	1.2578	1.8697	0.0075	3.2025	0.029	0.178
GMsFEM ($L_i = 9$)	2.3756	2.6576	0.1129	5.1461	0.014	0.033

Table 3.1: Comparison of run-time (sec) and L^2-errors for simulating a mono-frequency single shot using different numerical schemes: FDM, FEM, and GMsFEM ($L_i = 7$ and 9), where t_F and t_P denote the run-time took for the matrix factorization and matrix projection, respectively. t_I is the time for the actual matrix inversion.

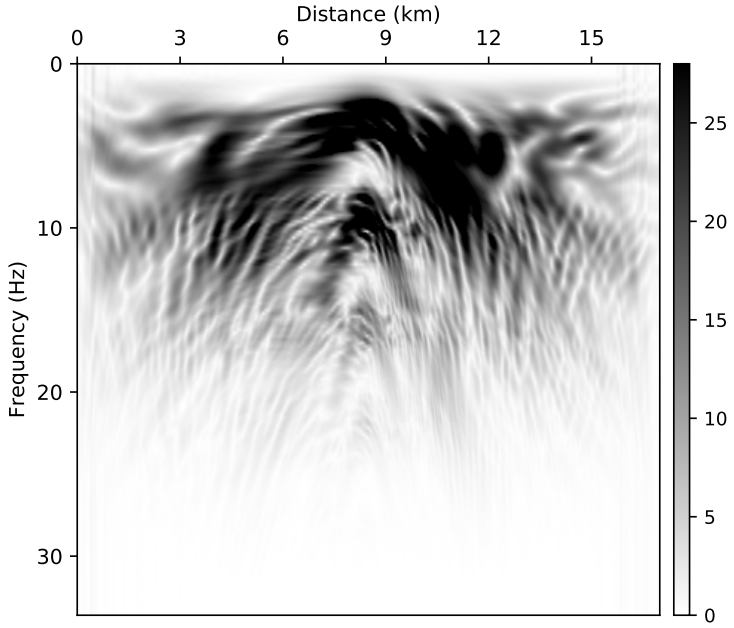
processor (Figure 3.12). Given the run-time being displayed in Figure 3.12, the corresponding speedup for each choice of the number of basis functions is 6.07 ($L_i = 6$), 4.15 ($L_i = 8$), and 2.82 ($L_i = 10$), respectively. Fewer basis functions can simulate the low frequency wave (10 Hz), and we can also calculate the high frequency waves by applying more basis functions. For example, in 30 Hz case, we can compute the solutions using 10 basis function (Figure 3.12(c)), while 6 basis functions are enough to get accurate solutions in 10 Hz wave (Figure 3.12(a)).

We presented the frequency spectra calculated from FEM and GMsFEM in Figure 3.13. Though both of the frequency spectra seem to be identical, the numerical error from GMsFEM is concentrated in the high frequency part. To make a clear comparison, we subtracted the frequency spectra of the GMsFEM from the reference one as presented in Figure 3.14. While the fine-scale cell is fixed to 10 m, we performed wave simulations in different sizes of coarse grids: 100 m and 50 m, to observe the relationship between the size of coarse cell and the number of basis functions in GMsFEM.

We measured the difference of the frequency spectra from the reference result for the mesh with 100 m coarse cell with 6, 7, and 8 basis functions (Figure 3.14(a), 3.14(c), and 3.14(e)). We can observe that the wavefields below 20 Hz are accurate in spite of the fewer basis functions. However, at the higher frequency band (over 20 Hz), the wavefields are sensitive to the number of basis functions. We found that 8 basis functions are enough to suppress the numerical errors over the entire frequency bands as presented in Figure 3.14(e). Similarly, in the spectra examples from

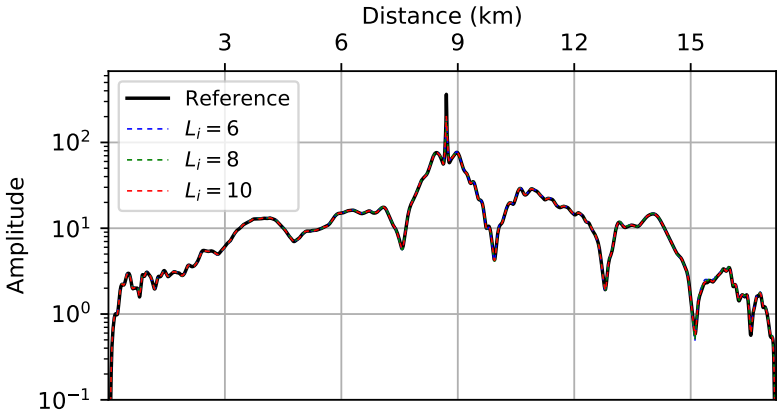


(a)

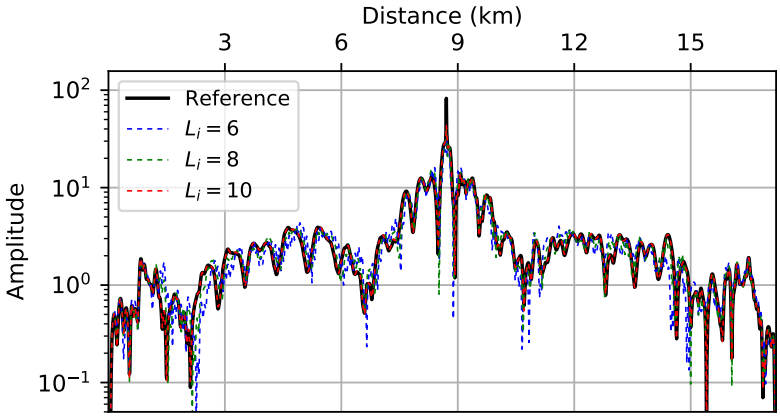


(b)

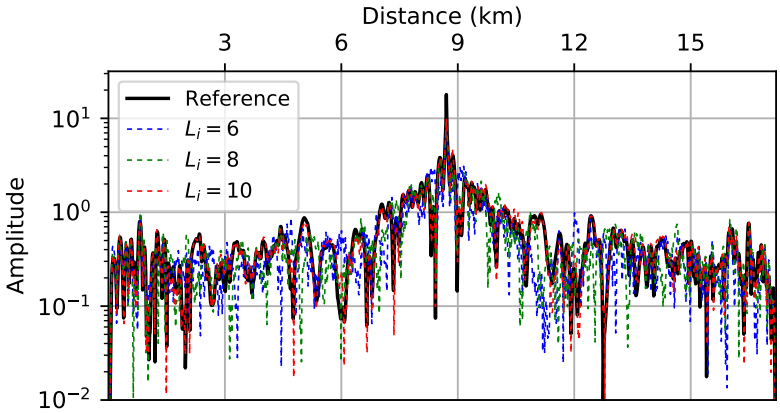
Figure 3.11: Synthetic seismic data (15 Hz central frequency) that are generated by using the continuous-Galerkin finite element method: (a) A seismogram from the source at the center point of the surface with (b) corresponding frequency spectrum of the shot gather.



(a)

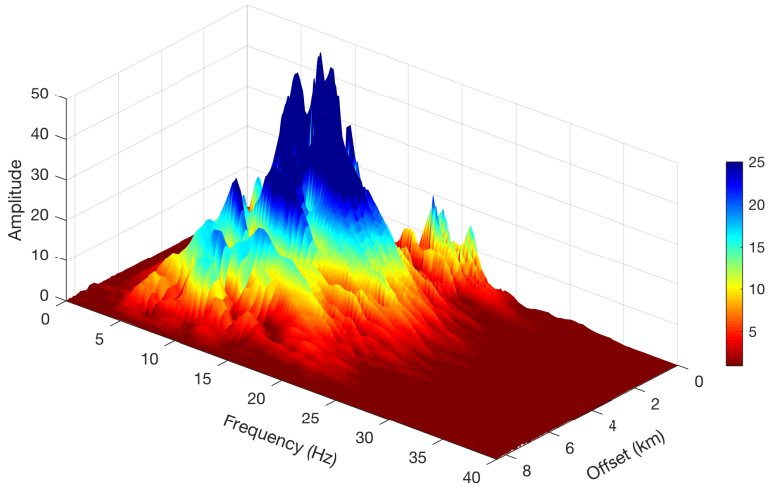


(b)

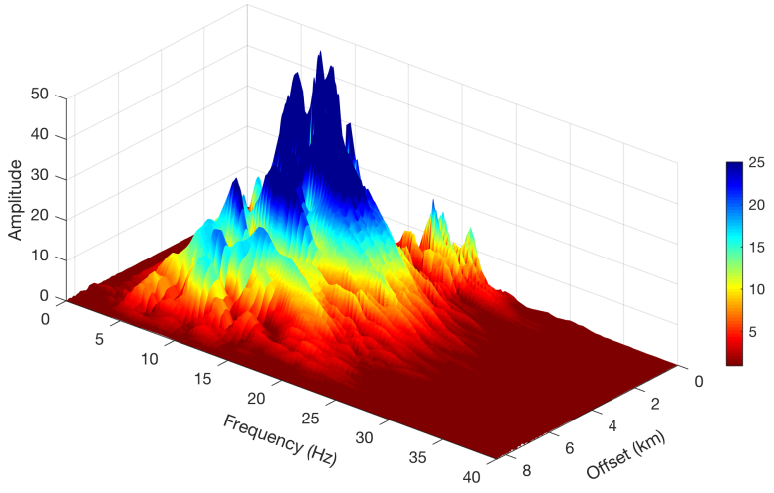


(c)

Figure 3.12: Cross-section of the spectra at frequencies of (a) 10 Hz, (b) 20 Hz, and (c) 30 Hz, where the runtime for each case is 26.1 s, 38.3 s, and 56.9 s. It took 158.4 s to simulate the waves in fine-scale mesh.



(a)



(b)

Figure 3.13: Comparison of the frequency spectra: (a) Reference frequency spectrum that is calculated only by using fine-scale grid, and (b) frequency spectra from GMsFEM ($L_i = 6$) with 100 m coarse grid.

the 50 m coarse cell (Figure 3.14(b), 3.14(d), 3.14(f)), the error decreases as we applied more basis functions; however, in this case, we could achieve an acceptable level of accuracy only by using 5 basis functions. In short, the GMsFEM with smaller coarse mesh can calculate the solutions using fewer basis functions.

In GMsFEM, as there exists a trade-off between the speedup and accuracy, we can define an optimized set of parameters: number of basis functions L_i and coarse cell size H, according to the target frequency band. For example, at lower frequency, we utilize a coarser mesh with an appropriate number of basis functions $L_{i,f_{\text{low}}}$, while using a finer mesh at higher frequencies with $L_{i,f_{\text{high}}}$ bases. Note that we allocate more basis functions in the coarser mesh at lower frequency ($L_{i,f_{\text{low}}} > L_{i,f_{\text{high}}}$). We will harness this feature of GMsFEM in the RTM by varying the size of coarse mesh at different frequency bands.

3.3.3 Multiscale RTM examples

The Marmousi-2 model is used to demonstrate the RTM using multiscale grids. To generate synthetic seismic data, we applied a source with 15 Hz central frequency, and the total recording time is 5 s. The receivers are located on the entire surface of the model every 10 m. The source interval is 50 m, and 338 sources are applied at 10 m depth. We divided the target frequency band into three parts: $0 \sim 10$ Hz, $10 \sim 20$ Hz, and $20 \sim 25$ Hz. Then different sizes of coarse meshes are applied for each frequency band with an appropriate number of basis functions: 200 m ($L_i = 12$), 100 m ($L_i = 8$), and 50 m ($L_i = 6$), while the fine mesh is fixed to 10 m. The RTM gradient images for each frequency band are presented in Figure 3.15. The gradient image from the lowest frequency band builds macro geological structures, and the image from the higher frequency bands can capture the detailed strata. Therefore, by using GMsFEM combined with multiscale spatial grids, we could successfully construct the gradient images without any contamination of numerical artifact. The key part of this method is to utilize appropriate number of basis functions for each frequency band to maximize the efficiency of the computation. We sum all the gradient images in Figure 3.15 to obtain the final RTM gradient images. Then, we applied the Laplacian filter to get the final RTM images.

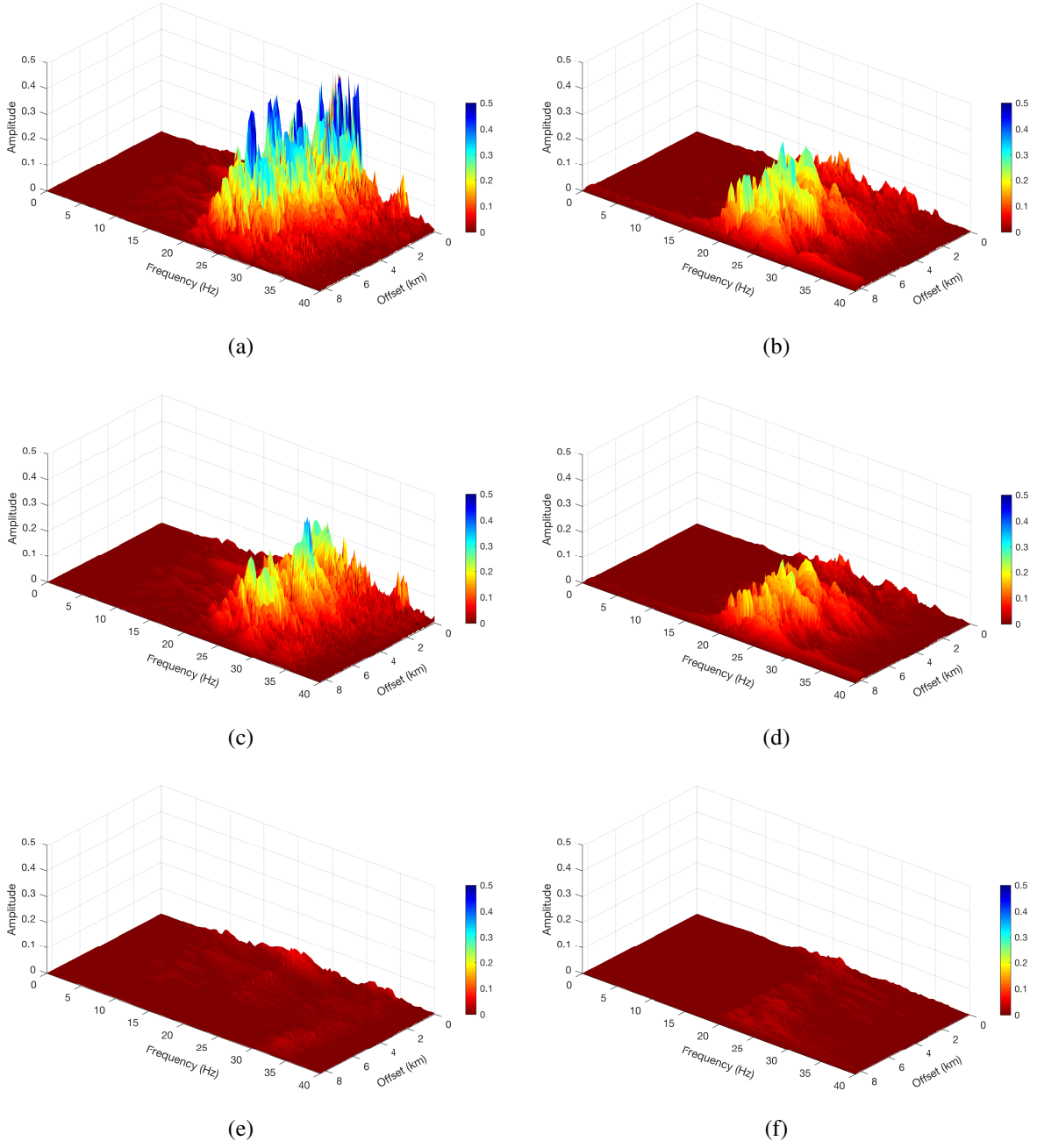


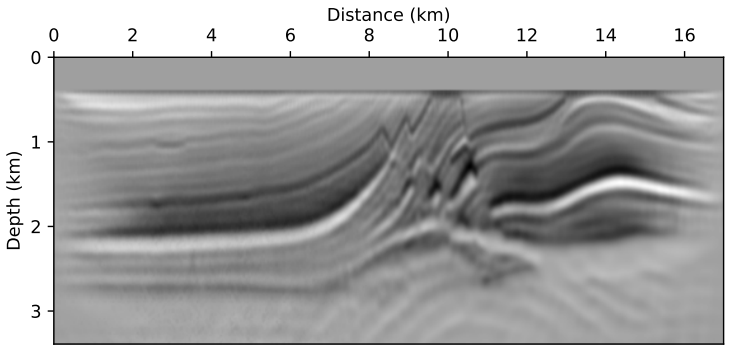
Figure 3.14: Difference between the reference frequency spectrum and the spectra from GMsFEM with different coarse grid: $H = 100\ m$ (a, c, e) and $H = 50\ m$ (b, d, f), varying the number of basis functions: (a) $L_i = 6$, (c) $L_i = 7$, (e) $L_i = 8$ and (b) $L_i = 3$, (d) $L_i = 4$, and (f) $L_i = 5$.

Method	Grid size	Run-time in sec. (speedup)			
		0∼10 Hz	10∼20 Hz	20∼25 Hz	0∼25 Hz
FEM	10 m	373.71 (-)	380.42 (-)	203.11 (-)	957.24 (-)
FDM	10 m	250.39 (1.48)	266.11 (1.43)	138.26 (1.47)	654.76 (1.46)
GMsFEM	100 m	104.63 (3.57)	107.17 (3.55)	57.21 (3.55)	269.01 (3.56)
GMsFEM	Adaptive	146.23 (2.56)	162.74 (2.34)	117.83 (1.72)	426.80 (2.24)

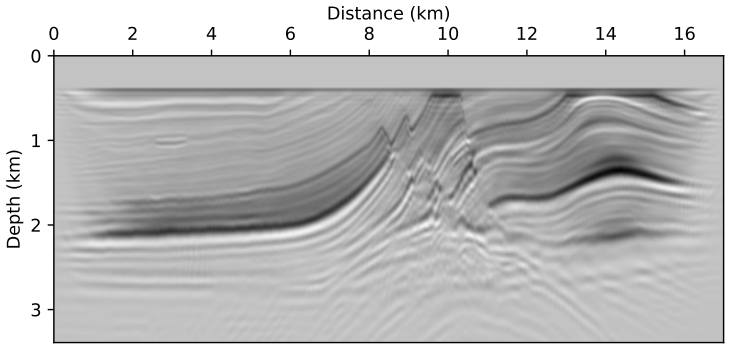
Table 3.2: Run-time for calculating RTM gradient image at each frequency band. In the GMsFEM with 100 m coarse grid size, eight basis functions are applied. We used 200 m (0∼10 Hz), 100 m (10∼20 Hz), and 50 m (20∼25 Hz) coarse grid for calculating the waves at each corresponding frequency band. Through the frequency-adaptive spatial grid and GMsFEM, we can achieve the RTM results 2.24 times faster than the reference case. For the measurement of run-times, we used 100 processor units.

We present the run-time of the RTM in Table 3.2 to compare the computational speed among different numerical schemes. We used 100 computing cores to perform the imaging. Comparing the run-times, the FDM computes the imaging conditions 46 % faster than continuous-Galerkin FEM. Also, when we applied the GMsFEM with constant coarse meshes (100 m, $L_i = 8$), the RTM result could be obtained more rapidly as a factor of 3.56. Note that the FEM is the reference to compute the speedup of the FDM and the GMsFEM RTM examples. In the GMsFEM with frequency-adaptive approach, however, the speedup of GMsFEM varies upon the size of coarse grid size. For example, the speedup becomes larger at lower frequency due to the reduction of the total number of unknowns in the online stage. Observing the run-time in Table 3.2, the RTM method using frequency-adaptive GMsFEM could compute the imaging condition faster (speedup≈1.53) than the finite-difference approach. However, note that the magnitude of speedup varies with the order of numerical method and the performance of the direct solver. We used the PARDISO (Lawson et al., 1979; Dongarra et al., 1988; Dodson et al., 1991) solver in this study. We made comparison between the RTM results from different wave modeling engines. The results of RTM with CG FEM modeling (Figure 3.16(c) and 3.16(d)) delineate more detailed strata than the images with FDM modeling engine (Figure 3.16(a) and 3.16(b)). This is attributed to the number of points that are considered to compute a wavefield at a certain grid points.

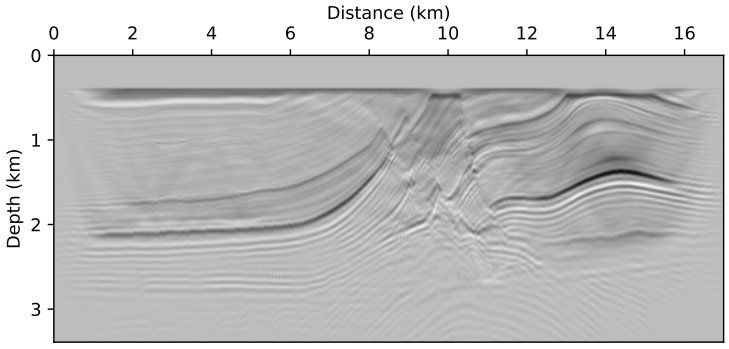
To put it differently, the FEM uses a greater number of points (8) than the FDM (4), so it



(a)



(b)



(c)

Figure 3.15: Multiscale RTM gradient ($h = 10$ m) from different size of coarse mesh for corresponding frequency bands: (a) low-frequency band; 0∼10 Hz ($H = 200$ m, $L_i = 12$), (b) mid-frequency band; 10∼20 Hz ($H = 100$ m, $L_i = 8$), and (c) high-frequency band; 20∼25 Hz ($H = 50$ m, $L_i = 6$).

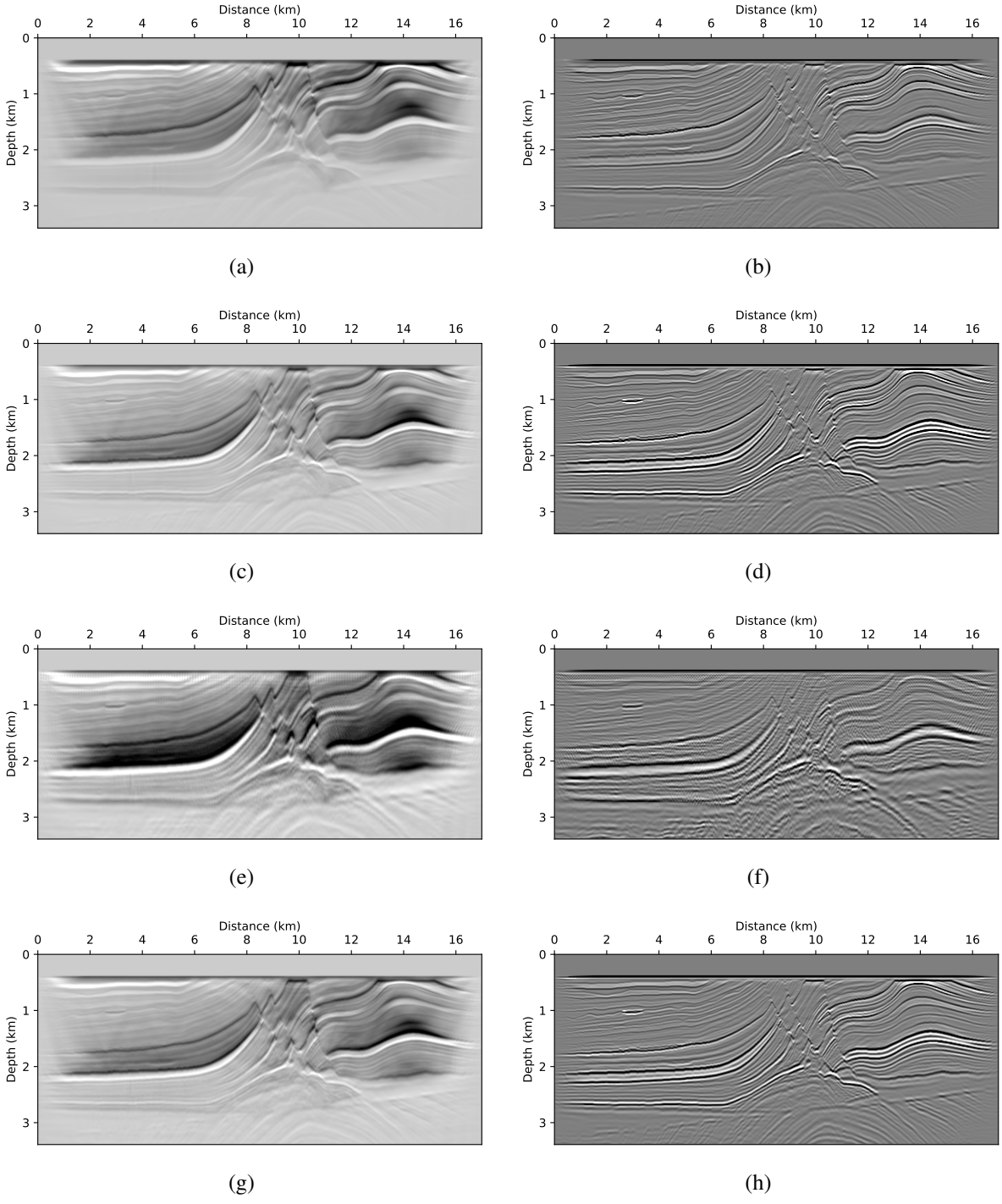


Figure 3.16: The gradient images (left panels) and final RTM images (right panels) from various numerical method for wave modeling: FDM (a, b), FEM (c, d), GMsFEM with $L_i = 8$ (e, f), and GMsFEM with frequency-adaptive grids (g, h). The run-time of each RTM images are presented in Table 3.2.

calculates the solution more accurately compared to the same order of FDM. The results with fixed and adaptive meshes are compared in Figure 3.16(f) and 3.16(h). Despite the fact that the multiscale RTM with a fixed mesh can capture the large-scale variations of the geological structures more rapidly than the reference image, it is hard to resolve the fine strata that are located between the major layers. In addition, after we applied the Laplacian filter to the gradient image, dispersive numerical artifacts are appeared (Figure 3.16(f)). We can apply different solutions to suppress the dispersive artifacts. Reducing the target frequency might be helpful to decrease the contamination of the image; however, eliminating high-frequency components also hinders to resolve the small-scale geological features. Therefore, we applied a flexible spatial grid according to the target frequencies. Figure 3.16(g) and 3.16(h) exhibit the gradient and the final RTM images, respectively, that are calculated by using the frequency -adaptive multiscale spatial grids, and it can resolve all the detailed geological features of the Marmousi-2 model. It also computes the imaging conditions 2.24 times faster than the CG FEM case.

3.4 Conclusions

The GMsFEM algorithm combined with flexible multiscale grids can efficiently simulate waves in the frequency domain. We demonstrate that the GMsFEM using the newer local spectral problem including the partition of unity term provides the multiscale basis functions which can enhance the accuracy of the solutions. We then show the relationship between the size of coarse cell and the number of basis functions in GMsFEM to maximize the speedup. The results show that the coarser mesh requires more basis functions to capture the influence of fine-scale heterogeneity, but we may use a smaller number of basis functions as the coarse mesh becomes finer. We also applied a different size s of coarse mesh and corresponding number of multiscale basis to enhance the computational efficiency of the wave modeling in the RTM. When we compared the multiscale RTM results with a fixed mesh and a flexible mesh, the RTM result with the frequency-adaptive meshes produces better images than the other one.

4. ACCELERATING FULL-WAVEFORM INVERSION VIA GENERALIZED MULTISCALE FINITE ELEMENT METHOD

4.1 Introduction

Full-waveform inversion (FWI) has drawn wide attention for constructing an earth model with high resolution. As a wave-equation based imaging technique, FWI also helps to resolve complex geological structure with corresponding earth properties such as velocity or acoustic impedance. Since Lailly and Bednar (1983) and Tarantola (1984) proposed a back propagation algorithm to perform waveform inversion more efficiently, scientists have been developing FWI algorithms with enhanced robustness (Virieux and Operto, 2009). However, FWI still requires significant computational cost, and many studies also consider efficient implementations of FWI.

There exist two different approaches to implement FWI in terms of the domain of wave modeling: time- and frequency-domain. Time-domain is widely used due to its advantage in computing memory consumption. Also, we can parallelize the time-domain wave modeling by considering sub-domain decomposition. However, working in the frequency domain has its own advantages as well (Pratt, 1999; Wu and Alkhalifah, 2018). For example, in terms of wave modeling, we can obtain the wave solutions by applying a one-time impedance matrix inversion without reducing time stepping. Therefore, if we can reduce the size of impedance matrix, the speed of wave simulation could be greatly enhanced. Also, we do not need to consider a reduced time-step to simulate high frequency waves. Moreover, in the frequency domain FWI, we can improve gradient scaling which helps accelerate convergence of the solution by applying the inverse Hessian at each frequency independently.

To enhance the computational efficiency of FWI, we can consider several different approaches: 1) selection of a suitable sequence of frequency groups, 2) applying a technique for fast model convergence, or 3) accelerating wave modeling. Selecting frequency components which have a large contribution to the final inversion results can reduce the computational burden by removing

unnecessary forward- and backward-modeling. Sirgue and Pratt (2004) presented a FWI strategy with selective frequency components. In their approach, the total number of frequency could be reduced by eliminating redundant information in the wavenumber coverage of the target offset. Sirgue and Pratt (2004) proposed a new strategy of discretization which is determined by the maximum offset presented in the seismic survey. Kwon et al. (2017) applied the same strategy to the Laplace domain FWI by selecting the Laplace constants to minimize the redundancy of the imaginary wavenumber while reserving the continuity of the imaginary wavenumber of the local wave path. However, utilizing a large number of frequency components can help to enhance the signal to noise ratio, especially when we deal with a field data.

Xue et al. (2017) introduced a method to accelerate the convergence rate of the FWI process by replacing the attenuated gradient with a Q-compensated gradient. In this approach, they employed a visco-acoustic wave equation to formulate a Q-compensated FWI. Pan et al. (2017) applied a new preconditioning schemes for the conjugate-gradient method to compute the Hessian-free Gauss-Newton method rapidly. In addition, for more effective wave modeling, Fomel et al. (2013) employed a low rank approximation of a wavenumber matrix for wavefield extrapolation. The low rank method decomposes the initial wave propagation matrix into a small sets with representative wave numbers. Another method for acceleration of wave modeling is proposed by Nunes and Minkoff (2014), which uses a subgrid upscaling technique.

In this paper, the proposed method concentrates on the acceleration of FWI via rapid computation of the wavefield. We apply the Generalized Multiscale Finite Element Method (GMsFEM) (Efendiev et al., 2011; Chung et al., 2014; Gibson and Fu, 2015), which is similar to Nunes and Minkoff (2014)'s approach. However, the key point of the GMsFEM is to reduce the dimension of impedance matrix without performing any upscaling or model homogenization. From our previous work (Gao et al., 2015; Fu and Gao, 2017; Cho et al., 2017b), we demonstrated that the GMsFEM with fewer basis functions can accelerate the simulation of low frequency waves without sacrificing accuracy. Also, Cho et al. (2017a) introduced the first application of GMsFEM forward modeling to frequency domain RTM, showing that the GMsFEM wave modeling engine could resolve

images of complex salt structure rapidly. As FWI requires multiple iterations of model updates, applying the GMsFEM to FWI for resolving subsurface geological structures has the potential of even greater benefit from the accelerations provided by the GMsFEM.

In this paper, we will briefly describe the theory of GMsFEM to solve the Helmholtz equation in acoustic media, then we illustrate the model update term in FWI using multiscale wave modeling. A sensitivity kernel analysis is also presented with FWI results from the Marmousi-2 model. The examples of the sensitivity kernel demonstrate that the correct model update is still maintained even if we apply the coarse scale wave simulation. Also, we will demonstrate the multiscale FWI approach by using field seismic data acquired in the Gulf of Mexico offshore area.

4.2 Method

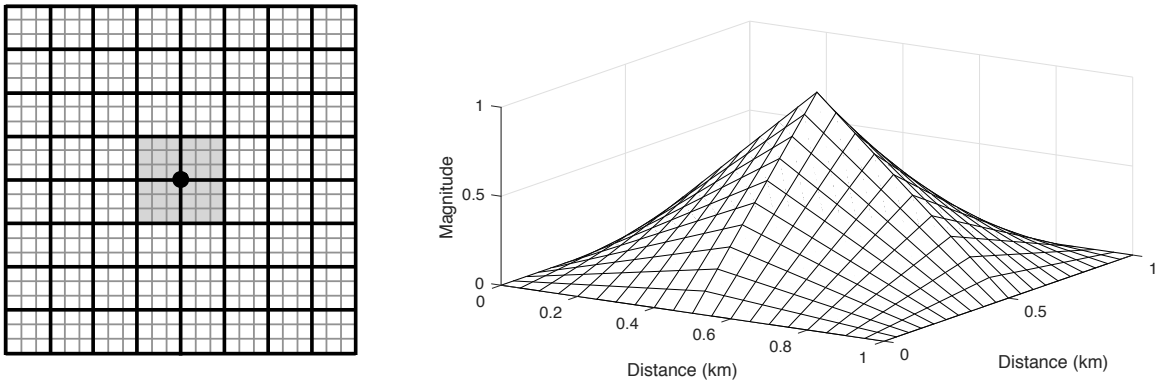
We start the application of the GMsFEM wave modeling to full-waveform inversion with the following acoustic wave equation:

$$-\frac{\omega^2}{\rho v^2}u = \nabla \cdot (\frac{1}{\rho}\nabla u) + f, \tag{4.1}$$

where v and ρ are the velocity and density, respectively. u is the pressure, and f is the source term. In GMsFEM, we use a mesh which includes two different sizes of grids (Figure 4.1(a)). The wave simulation via GMsFEM consists of two main steps: 1) construction of basis functions on the fine grid, and 2) wave modeling in the coarse grid.

4.2.1 Computation of multiscale basis functions

The process for the construction of the basis functions on the coarse grid is termed the “offline stage". The offline stage is performed only once, since the multiscale basis functions depend only on the given velocity and density model. In other words, these basis functions are independent of the acquisition geometry such as source and receiver locations. Therefore, the offline stage adds little effort to the total computation. We present the mesh which is used for the GMsFEM implementation in Figure 4.1(a). In the GMsFEM, the domain Ω consists of three factors: 1) fine-scale mesh, 2) coarse-scale mesh, and 3) a coarse node’s neighborhood. The fine mesh (gray



(a) A spatial mesh for GMsFEM.

(b) Partition of unity at a coarse neighborhood.

Figure 4.1: (a) Fine meshes incorporate the properties (v and ρ), and the actual wave simulation is performed in the coarse mesh. Gray area shows a coarse neighborhood with the corresponding (b) partition of unity.

solid line) stores all the background properties and has the same dimension as the properties (i.e., velocity and density). The coarse mesh is highlighted with bold solid line in Figure 4.1(a), which contains all the nodal points to compute the wave solutions. When we calculate wave solutions at ith coarse node, we need to use the multiscale basis functions, which are built in the ith coarse neighborhood (surrounding coarse cells of ith coarse node) as colored with gray (Figure 4.1(a)).

The offline stage starts with a local spectral problem for each coarse neighborhood. Efendiev et al. (2011) proposed a method of solving the local eigenvalue problem (equation 4.2) for building multiscale basis functions to incorporate the fine-scale heterogeneity of the background properties, which can be expressed as follows:

$$-\nabla\cdot\left(\frac{1}{\rho}\nabla\varphi\right)=\lambda\frac{1}{\rho}\varphi,\qquad\frac{1}{\rho}\frac{\partial\varphi}{\partial\mathbf{n}}=0. \tag{4.2}$$

Fu and Gao (2017) then introduced the following form of eigenvalue problem by including the

partition of unity (Babuška et al., 1995) (Figure 4.1(b)) for more accurate wave solutions:

$$-\nabla\cdot\left(\frac{\psi_i^2}{\rho}\nabla\varphi\right)=\lambda\frac{\psi_i^2}{\rho}\varphi,\qquad\frac{\psi_i^2}{\rho}\frac{\partial\varphi}{\partial\mathbf{n}}=0, \tag{4.3}$$

where ψ is the partition of unity, and $\mathbf{n}$ is the outward-pointing normal vector of K. Cho and Gibson (2018) demonstrated that the local spectral problem including the partition of unity enhances the accuracy of the wave solutions due to the well-balanced distribution of the impedance matrix components and properly weighted diagonal components. To solve the local spectral problems in each coarse neighborhood, we discretize equation 4.3 as

$$\mathbf{K}\varphi=\lambda\mathbf{M}\varphi, \tag{4.4}$$

where

$$\mathbf{M}=\int_{K_i}\frac{1}{\rho}\zeta\xi d\mathbf{x},\qquad\mathbf{K}=\int_{K_i}\frac{1}{\rho}\nabla\zeta\cdot\nabla\xi d\mathbf{x}, \tag{4.5}$$

where ζ and ξ are the first order polynomial basis functions that are applied on the fine mesh. $\mathbf{M}$ and $\mathbf{K}$ are the mass and stiffness matrices for the given coarse neighborhood K_i, and the coarse neighborhood is highlighted with gray color in Figure 4.1(a). Solving the local eigenvalue problems for each coarse neighborhoods yields a set of eigenvalues with corresponding eigenfunctions. We then sort the eigenvalues in ascending order for each of K_i as:

$$\lambda_i^1\leq\lambda_i^2\leq\cdots\leq\lambda_i^{L_i}\leq\cdots\leq\lambda_i^{N_f}, \tag{4.6}$$

where N_f is the degrees of freedom for the given coarse neighborhood, and L_i denotes the number of multiscale basis functions that is selected for actual wave modeling. By considering higher order wave modes, the basis functions used in the GMsFEM capture the influence of more detailed heterogeneity of the wave solutions. The first L_i eigenfunctions φ_i^j, $(j=1,\cdots,L_i)$ that we chose in the GMsFEM solver are controlled by the degree of heterogeneity of the fine-scale background properties and the level of desired accuracy. We obtain the final multiscale basis functions by

multiplying the eigenfunctions and partition of unity ($\Phi_i^j = \varphi_i^j \psi_i$), since the defined eigenfunctions are not globally continuous (Babuška et al., 1995). A detailed procedure for building multiscale basis functions using a continuous-Galerkin formulation is illustrated by Gao et al. (2015); Fu and Gao (2017).

4.2.2 Wave modeling in coarse mesh

The Helmholtz equation incurs a large computational cost to compute the inverse of the global impedance matrix at every frequency component, which can be expressed as

$$\mathbf{Su} = \mathbf{f}, \qquad (\mathbf{S} = \omega^2 \mathbf{M} + i\omega \mathbf{C} + \mathbf{K}). \tag{4.7}$$

In the finite element method, this impedance matrix $\mathbf{S}$ is a combination of the mass $\mathbf{M}$, stiffness $\mathbf{K}$, and damping $\mathbf{C}$ matrices (Marfurt, 1984). For the implementation of the absorbing boundary conditions at the model edge, we applied a Perfectly Matched Layer (PML) boundary condition. The detailed description for the implementation of the damping matrix in coarse-scale modeling is illustrated by Fu and Gao (2017). By taking the inverse of the impedance matrix, we acquire the wave solutions $\mathbf{u}$ at the given frequency ω. For the matrix inversion, we can take two different methods to solve the system (equation 4.7): direct solver or iterative solver. Although the iterative solver might be useful in term of computing memory consumption, accuracy of the solver is sensitive to choice of preconditioner (Plessix, 2007), so it hinders computing accurate solutions especially in high-frequency components. In contrast, a direct solver yields an accurate solution and it is more robust than an iterative solver in terms of stability; however, it requires significant computational cost. In this case, since accurate wave modeling is critical in FWI, we need to use a direct solver (`Pardiso` from MKL library) (Lawson et al., 1979; Dongarra et al., 1988; Dodson et al., 1991).

Therefore, a key goal of this study is to reduce the computational burden working with a direct solver. To achieve this goal, we aim to reduce the dimension of the fine-scale impedance matrix $\mathbf{S}_h$. In our proposed method, we do not construct the mass and stiffness matrices in the coarse mesh

explicitly. Instead, we compute the coarse-scale impedance matrix from the fine-scale impedance matrix by using a projection matrix $\mathbf{R}$, where the column i of the $\mathbf{R}$ matrix is the discrete multiscale basis functions Φ_i^j from each coarse node neighborhood K_i. The projection matrix is expressed as

$$\begin{aligned} \mathbf{R} &= \left(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_N\right)^{\mathrm{T}}, \\ \mathbf{R}_i &= \left(\Phi_1, \Phi_2, \ldots, \Phi_{L_i}\right). \end{aligned} \tag{4.8}$$

The row dimension of the projection matrix is identical to the total number of fine-scale nodes, and the column dimension is equal to the product of the number of multiscale basis functions L_i and the total number of coarse-scale nodes. A primary focus of our approach is that we accelerate the wave simulation (online stage) in the frequency domain by dramatically reducing the size of impedance matrix via this projection. To compute the coarse-scale impedance matrix $\mathbf{S}_H$ and source vector $\mathbf{f}_H$, we apply the projection matrix $\mathbf{R}$ to the existing fine-scale impedance matrix $\mathbf{S}_h$ and source vector $\mathbf{f}_h$, which can be written as follows:

$$\mathbf{S}_H = \mathbf{R}^{\mathrm{T}}\mathbf{S}_h\mathbf{R}, \quad \mathbf{f}_H = \mathbf{R}^{\mathrm{T}}\mathbf{f}_h. \tag{4.9}$$

After applying the projection matrix, we can express the system shown in equation 4.7 in a coarse-scale formulation as

$$\sum_{m,n} r_{mi}\left[\int_\Omega -\frac{\omega^2}{\rho v^2}\phi_m\phi_n d\mathbf{x} + \int_\Omega \frac{1}{\rho}\nabla\phi_m\cdot\nabla\phi_n d\mathbf{x}\right] r_{jn}\sum_k r_{jk}\int_\Omega u\phi_k d\mathbf{x} = \sum_l r_{jl}\int_\Omega f\phi_l d\mathbf{x}, \tag{4.10}$$

where r_{ij} is the component of projection matrix and ϕ_j is a piece-wise linear basis function in the fine-scale finite element formulation. After obtaining final solutions on the coarse mesh, we reapply the projection matrix ($\mathbf{u}_h = \mathbf{R}\mathbf{u}_H$) to recover the original dimension of the solution in fine-scale mesh. Note that even though the wave modeling is performed in the coarse mesh, we still acquire the wave solution which has the same dimension as the fine-scale (reference) mesh.

4.2.3 FWI using GMsFEM

The goal of FWI is to update the property model in the direction which can minimize the misfit of observed and modeled data. Therefore, we applied a steepest descent method to minimize the l_2 objective function, which can be written as

$$E(\mathbf{m}) = \frac{1}{2}\sum_{i=1}^{n_s}\sum_{j=1}^{n_r} \| \left(u_{ij} - d_{ij}\right)^2 \|, \tag{4.11}$$

where $\mathbf{m}$ is the model parameter. $\mathbf{u}_{ij}$ and $\mathbf{d}_{ij}$ are the modeled and observed wavefield at ith source and jth receiver pairs, respectively. By taking the partial derivative of equation 4.11, the steepest descent direction can be determined as

$$\frac{\partial E(\mathbf{m})}{\partial m_k} = \sum_{i=1}^{n_s}\sum_{j=1}^{n_r} \Re\left[\frac{\partial u_{ij}}{\partial m_k}\left(u_{ij} - d_{ij}\right)^*\right], \tag{4.12}$$

where the asterisk denotes a complex conjugate. To compute the Born (or partial derivative) wavefield $\partial_{m_k} u_{ij}$, we used the differentiating form of the Helmholtz equation (equation 4.7). Taking the partial differential of the system with respect to a kth model parameter yields the following:

$$\frac{\partial \mathbf{S}_h}{\partial m_k}\mathbf{u}_h + \mathbf{S}_h\frac{\partial \mathbf{u}_h}{\partial m_k} = 0, \tag{4.13}$$

or

$$\frac{\partial \mathbf{u}_h}{\partial m_k} = \mathbf{S}_h^{-1}\mathbf{v}, \tag{4.14}$$

where $\mathbf{v} = -\left[\partial_{m_k}\mathbf{S}_h\right]\mathbf{u}_h$ is the virtual source term which is required to perturb the model parameter. By substituting equation 4.14 into equation 4.12, we obtain the gradient of the objective function as

$$\frac{\partial E(\mathbf{m})}{\partial m_k} = \int_0^{\omega_{\max}} \sum_{i=1}^{n_s} \Re\left[\mathbf{v}^{\mathrm{T}}\mathbf{S}_h^{-1}\mathbf{r}_i^*(\omega)\right] d\omega, \tag{4.15}$$

where $\mathbf{r}_i^*$ denotes the complex conjugate of the residual. The fine-scale impedance matrix $\mathbf{S}_h$ has a symmetric form and satisfies the reciprocity condition. $\Re$ indicates the real part of the complex

number. As calculating the Born wavefields explicitly requires considerable computational cost, we employed the adjoint-state method (Tarantola, 1984; Plessix, 2006). Equation 4.15 expresses the gradient term using fine-scale wave modeling results. By applying the projection matrix $\mathbf{R}$, we calculate the multiscale gradient using the coarse-scale wave modeling as

$$\frac{\partial E(\mathbf{m})}{\partial m_k} = \int_0^{\omega_{\max}} \sum_{i=1}^{n_s} \Re \left[\mathbf{v}^{\mathrm{T}} \mathbf{R} \mathbf{S}_H^{-1} \mathbf{R}^{\mathrm{T}} \mathbf{r}_i^*(\omega) \right] d\omega. \tag{4.16}$$

In FWI, applying an appropriate gradient scaling is important to resolve the deeper part of the subsurface. In the steepest descent method, we consider the Hessian matrix for the gradient scaling, which can be written as

$$H_{ij} = \frac{\partial^2 E(\mathbf{m})}{\partial m_i \partial m_j}, \quad (i, j = 1, 2, \cdots n), \tag{4.17}$$

where H_{ij} denote the component of the Hessian matrix, and n is the total number of model parameters. Using the complete form of the Hessian (equation 4.17) requires significant computational cost. The Hessian matrix for the iterative solution termed the full-Newton method, which can be expressed as

$$\begin{aligned} H_{ij} &= \frac{\partial}{\partial m_i} \frac{\partial E(\mathbf{m})}{\partial m_j} \\ &= \frac{\partial}{\partial m_i} \Re \left[\left(\frac{\partial \mathbf{u}}{\partial m_i} \right)^{\mathrm{T}} \mathbf{r}^* \right] \\ &= \Re \left[\left(\frac{\partial \mathbf{u}}{\partial m_i} \right)^{\mathrm{T}} \left(\frac{\partial \mathbf{u}}{\partial m_i} \right)^* + \frac{\partial^2 \mathbf{u}}{\partial m_i \partial m_j} \mathbf{r}^* \right]. \end{aligned} \tag{4.18}$$

The equation above can be rewritten in matrix from as

$$H_{ij} = \Re \left\{ \begin{bmatrix} \frac{\partial u_1}{\partial m_i} & \frac{\partial u_2}{\partial m_i} & \cdots & \frac{\partial u_n}{\partial m_i} \end{bmatrix} \begin{bmatrix} \frac{\partial u_1^*}{\partial m_j} \\ \frac{\partial u_2^*}{\partial m_j} \\ \vdots \\ \frac{\partial u_n^*}{\partial m_j} \end{bmatrix} + \begin{bmatrix} \frac{\partial^2 u_1}{\partial m_i \partial m_j} & \frac{\partial^2 u_2}{\partial m_i \partial m_j} & \cdots & \frac{\partial^2 u_n}{\partial m_i \partial m_j} \end{bmatrix} \begin{bmatrix} r_1^* \\ r_2^* \\ \vdots \\ r_n^* \end{bmatrix} \right\}. \tag{4.19}$$

Equation 4.19 can be simplified as

$$\begin{aligned} H_{i.j} &= \Re\left\{\mathbf{J}^{\mathrm{T}}\mathbf{J}^{*}\right\} + \Re\left\{\left[\begin{array}{cccc} \frac{\partial \mathbf{J}^{\mathrm{T}}}{\partial m_1}\mathbf{r}^{*} & \frac{\partial \mathbf{J}^{\mathrm{T}}}{\partial m_w}\mathbf{r}^{*} & \cdots & \frac{\partial \mathbf{J}^{\mathrm{T}}}{\partial m_n}\mathbf{r}^{*} \end{array}\right]\right\} \\ &= \mathbf{H}_a + \mathbf{R}, \end{aligned} \tag{4.20}$$

where the $\mathbf{R}$ term can be ignored for the simplification, and we call the left term $\mathbf{H}_a$ as the approximated Hessian. The method which only considers the term $\Re\left\{\mathbf{J}^{\mathrm{T}}\mathbf{J}^{*}\right\}$ is named the Quasi-Newton method; however, this method also consumes a great deal of computing resources. Therefore, for practical implementation of the gradient scaling, Shin et al. (2001) proposed an efficient scheme called a pseudo-Hessian matrix. The pseudo-Hessian $\mathbf{H}_p$ only considers the diagonal components of the approximated Hessian $\mathbf{H}_a$ to reduce the memory consumption. The derivation of the pseudo-Hessian begins with an approximated Hessian as

$$\begin{aligned} H_{a(i,j)} &= \Re\left\{\mathbf{J}^{\mathrm{T}}\mathbf{J}^{*}\right\} \\ &= \Re\left[\begin{array}{cccc} v_1^{\mathrm{T}} & v_2^{\mathrm{T}} & \cdots & v_n^{\mathrm{T}} \end{array}\right]\left(\mathbf{S^{-1}}\right)^{\mathrm{T}}\mathbf{S^{-1}}\left[\begin{array}{cccc} v_1^{*} & v_2^{*} & \cdots & v_n^{*} \end{array}\right], \end{aligned} \tag{4.21}$$

where $\mathbf{v}$ means the virtual source, and the product of the Green's function $\left(\mathbf{S^{-1}}\right)^{\mathrm{T}}\mathbf{S^{-1}}$ can be approximated by the identity matrix $\mathbf{I}$, then the matrix form of equation 4.21 can be expressed as

$$H_{a(i,j)} = \Re\left[\begin{array}{cccc} v_1^{\mathrm{T}}v_1^{*} & v_2^{\mathrm{T}}v_1^{*} & \cdots & v_n^{\mathrm{T}}v_1^{*} \\ v_1^{\mathrm{T}}v_2^{*} & v_2^{\mathrm{T}}v_2^{*} & \cdots & v_n^{\mathrm{T}}v_2^{*} \\ \vdots & \vdots & \ddots & \vdots \\ v_1^{\mathrm{T}}v_n^{*} & v_2^{\mathrm{T}}v_n^{*} & \cdots & v_n^{\mathrm{T}}v_n^{*} \end{array}\right]. \tag{4.22}$$

By selecting the diagonal components from the approximated Hessian, we can construct the pseudo-Hessian as follows:

$$\mathbf{H}_p = \mathrm{diag}\mathbf{H}_a = \left[\begin{array}{cccc} v_1^{\mathrm{T}}v_1^{*} & v_2^{\mathrm{T}}v_2^{*} & \cdots & v_n^{\mathrm{T}}v_n^{*} \end{array}\right]. \tag{4.23}$$

Hence, the final term for the update of the kth model parameter results from applying a pseudo-Hessian matrix, which can be written as

$$\Delta \mathbf{m} = -\int_0^{\omega_{\max}} \frac{\sum_{i=1}^{n_s} \Re\left[\mathbf{v}^{\mathrm{T}}\mathbf{R}\mathbf{S}_H^{-1}\mathbf{R}^{\mathrm{T}}\mathbf{r}_i^*(\omega)\right]}{\sum_{i=1}^{n_s} \Re\left[\mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{v})\right] + \lambda} d\omega, \tag{4.24}$$

where λ is a damping factor which stabilizes the singularity of the pseudo-Hessian matrix. In the proposed method, we performed the gradient scaling independently at each frequency, then summed the normalized model updates over the target frequency band. In this study, we used a fixed step length ($\alpha = 0.02$). Thus, at kth model parameter, the final form of the model update can be expressed as

$$\mathbf{m}_{k+1} = \mathbf{m}_k - \alpha \int_0^{\omega_{\max}} \frac{\sum_{i=1}^{n_s} \Re\left[\mathbf{v}^{\mathrm{T}}\mathbf{R}\mathbf{S}_H^{-1}\mathbf{R}^{\mathrm{T}}\mathbf{r}_i^*(\omega)\right]}{\sum_{i=1}^{n_s} \Re\left[\mathrm{diag}(\mathbf{v}^{\mathrm{T}}\mathbf{v})\right] + \lambda} d\omega. \tag{4.25}$$

4.2.4 Source estimation

For practical implementation of the FWI, estimating accurate source wavelets is critical for successful FWI results. In this research, I estimate the source wavelet using a full-Newton method with a L_2-norm objective functions (Lines and Treitel, 1984; Pratt, 1999; Shin et al., 2007). The source estimation can be started by expressing the source wavelet in complex form as follows: $g = e + if$, where e and f denote its real and imaginary part, respectively. Similarly, the complex wavefield at jth receiver location can be expressed as

$$\begin{aligned} u_j &= (c_j + id_j)(e + if), \\ &= (c_j e - d_j f) + i(c_j f + d_j e), \qquad (j = 1, 2, \cdots, n_r), \end{aligned} \tag{4.26}$$

where $c_j + id_j)$ is the Green's function computed at jth receiver location. Supposing that the observed data as $s_j = a_j + ib_j$, the residual at each receiver location can be written as

$$\begin{aligned} \delta r_j &= u_j - s_j \\ &= c_j e - d_j f - a_j + i(c_j f + d_j e - b_j), \qquad (j = 1, 2, \cdots, n_r). \end{aligned} \tag{4.27}$$

Given the full-Newton method begins with the following equation:

$$\delta \mathbf{p} = -\mathbf{H}^{-1} \nabla E_{\text{src}}, \tag{4.28}$$

where $\delta \mathbf{p}$ means the update term of the source wavelet, and ∇E_{src} is the steepest-descent direction. As I use L_2-norm in this derivation, the objective function E_{src} can be expressed as

$$\begin{aligned} E_{\text{src}} &= \frac{1}{2} \sum_j \delta r_j \delta r_j^*, \\ &= \frac{1}{2} \sum_j \left\{ (c_j e - d_j f - a_j)^2 + (c_j f + d_j e - b_j)^2 \right\}. \end{aligned} \tag{4.29}$$

The full Hessian matrix is given by

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 E_{\text{src}}}{\partial e^2} & \frac{\partial^2 E_{\text{src}}}{\partial e \partial f} \\ \frac{\partial^2 E_{\text{src}}}{\partial f \partial e} & \frac{\partial^2 E_{\text{src}}}{\partial f^2} \end{pmatrix}, \tag{4.30}$$

Substituting equation 4.30 into equation 4.28 gives

$$\delta \mathbf{p} = - \begin{pmatrix} \frac{\partial^2 E_{\text{src}}}{\partial e^2} & \frac{\partial^2 E_{\text{src}}}{\partial e \partial f} \\ \frac{\partial^2 E_{\text{src}}}{\partial f \partial e} & \frac{\partial^2 E_{\text{src}}}{\partial f^2} \end{pmatrix}^{-1} \begin{pmatrix} \nabla_e E_{\text{src}} \\ \nabla_f E_{\text{src}} \end{pmatrix} \tag{4.31}$$

By taking differentiation of the objective functions with respect to the real part e and imaginary part f of the source wavelet yields

$$\begin{aligned} \nabla_e E_{\text{src}} &= \sum_j \left\{ e(c_j^2 + d_j^2) - a_j c_j - b_j d_j \right\}, \\ \nabla_f E_{\text{src}} &= \sum_j \left\{ f(c_j^2 + d_j^2) + a_j d_j - b_j c_j \right\}. \end{aligned} \tag{4.32}$$

The element of the full Hessian matrix $\mathbf{H}$ can be calculated by taking the second derivative of the objective functions as follows:

$$\begin{pmatrix} \frac{\partial^2 E_{\text{src}}}{\partial e^2} & \frac{\partial^2 E_{\text{src}}}{\partial e \partial f} \\ \frac{\partial^2 E_{\text{src}}}{\partial f \partial e} & \frac{\partial^2 E_{\text{src}}}{\partial f^2} \end{pmatrix} = \begin{pmatrix} \sum_j (c_j^2 + d_j^2) & 0 \\ 0 & \sum_j (c_j^2 + d_j^2) \end{pmatrix}. \tag{4.33}$$

Thus, by performing the matrix operation in equation 4.31, the final update term of the source wavelet can be expressed as

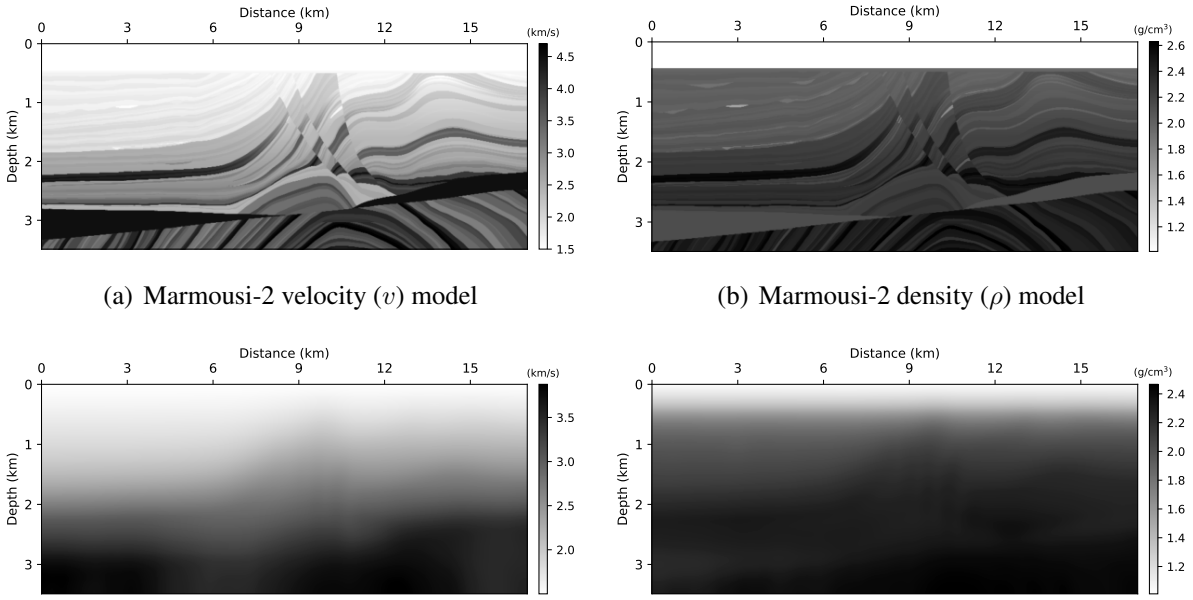
$$\delta \mathbf{p} = \begin{pmatrix} \delta e \\ \delta f \end{pmatrix} = \begin{pmatrix} -e + \frac{\sum_j (a_j c_j + b_j d_j)}{\sum_j (c_j^2 + d_j^2)} \\ -f - \frac{\sum_j (a_j d_j - b_j c_j}{\sum_j (c_j^2 + d_j^2)} \end{pmatrix}. \tag{4.34}$$

4.3 Numerical examples: synthetic data

4.3.1 Performance analysis with multiple sources

We used the Marmousi-2 model (Figure 4.2) (Martin et al., 2006) to demonstrate the proposed FWI with the GMsFEM wave modeling engine. We applied smoothing by using 1.5 km by 1.5 km window to construct an initial model (Figure 4.2(c) and 4.2(d)). To generate a synthetic dataset, we located 188 shots with 100 m interval, and the receiver interval is fixed to 20 m. The total recording time and the central frequency are 5 s and 20 Hz, respectively.

In the GMsFEM, we can accelerate the wave modeling by tuning the number of multiscale basis functions and coarse-scale grid size. (Gibson et al., 2014; Gao et al., 2015; Fu and Gao, 2017; Cho et al., 2017b). However, as we solve the frequency-domain wave equation by using the direct solver, the number of RHS values (the number of sources to solve at the same time via one-time matrix inversion) also has an influence on the computational speed. Thus, we varied the number of RHS in the direct solver to measure the amount of speedup in the coarse-scale modeling. Figure 4.3 presents the measured average runtime per iteration (mono-frequency) with various number of RHS and multiscale basis functions. In this example, we tested with 100 (maximum number of RHS) shots under a single iteration. According to the graph, in the GMsFEM, we can compute



(a) Marmousi-2 velocity (v) model

(b) Marmousi-2 density (ρ) model

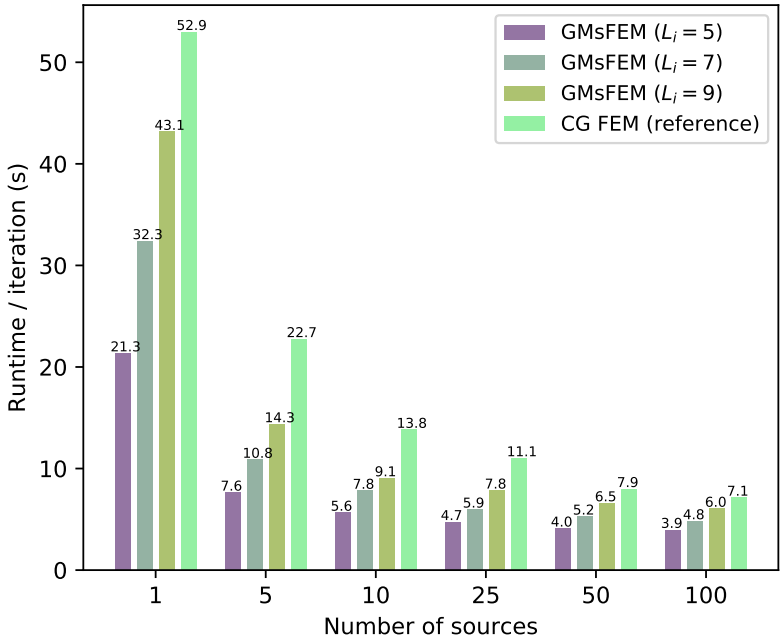
(c) Initial guess (v)

(d) Initial guess (ρ)

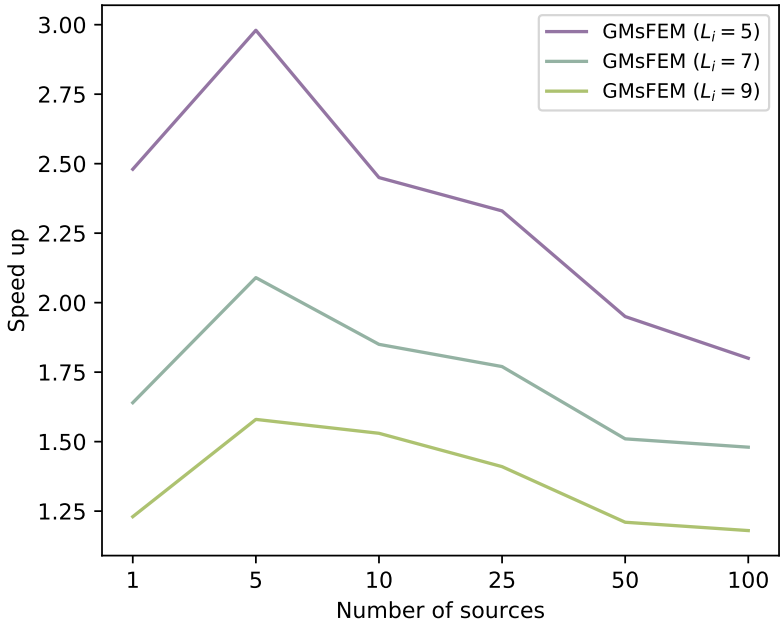
Figure 4.2: True velocity model and initial estimate.

the wave solutions more rapidly than the reference case (CG FEM); however, as we increase the number of RHS, the amount of speedup also decreases. In the case with 5 RHS, the GMsFEM wave modeling with 5 basis functions could compute the solution 2.98 times faster than the reference case. As we apply more basis functions ($L_i = 9$), the GMsFEM provides the results 1.58 times faster. In contrast, when we increase the number of RHS to 100, in both cases ($L_i = 5$ and $L_i = 9$), the speedup reduces to 1.8 and 1.18, respectively. Although the level of speedup decreases, we can still accelerate the wave modeling by using the GMsFEM wave modeling engine.

Since there exists a tradeoff between the accuracy and speedup in GMsFEM (Chung et al., 2011, 2014; Gibson et al., 2014; Gao et al., 2015), we also investigated the accuracy of the wave modeling as presented in Figure 4.4 and 4.5. We first generated the frequency-domain data matrix (Figure 4.4) using 7 Hz mono-frequency source. Observing the differences (Figure 4.4(b) and 4.4(d)) between the GMsFEM and the reference (CG FEM) case, we can conclude that both cases of the GMsFEM ($L_i = 5$ and 9) could perform the forward modeling without sacrificing accuracy. However, considering the difference of the data matrices from 13 Hz mono-frequency source



(a)



(b)

Figure 4.3: Performance analysis of the GMsFEM wave modeling as a function of number of sources included in the right hand side of the system of equation 4.7. The runtime and corresponding speedup presented in bar and line format, respectively.

(Figure 4.5(b)), we can infer that the five basis functions are not enough to capture all the detailed variation of the given frequency. Because of this insufficient basis functions with large coarse-scale grid size (200 m), a dispersive artifact is observed around the source locations (nearby the diagonal components of the data matrix) from Figure 4.4(a). Nevertheless, the noise that are originated from the coarse-scale modeling can be suppressed by applying a larger number of basis functions (Figure 4.5(d)). In short, the GMsFEM with fewer basis functions can accelerate the simulation of waves with low frequency. Also, we can model the high frequency waves accurately by increasing the number of basis functions.

4.3.2 Sensitivity kernel analysis

To demonstrate the influence of the changes in resolution and accuracy of the coarse scale modeling on the calculation of partial-derivative wavefields, we compared the sensitivity kernels that are calculated from the GMsFEM and conventional continuous-Galerkin (CG) FEM (the reference solution). In this test, we computed monochromatic sensitivity kernels for a homogeneous velocity (2.0 km/s). Those sensitivity kernels are generated from a single source (1 km, 2 km) with a single receiver (3 km, 2 km) at a single frequency. Figure 4.6 and 4.7 present the sensitivity kernels that are acquired for 5 Hz and 15 Hz, respectively. As shown in the previous examples with data matrices, we need to incorporate additional basis functions at higher frequencies to correctly simulate the details of high frequency waves. Similar patterns that are observed in the forward modeling using the GMsFEM can be found from the sensitivity kernel. Examining the sensitivity pattern with 5 Hz frequency (Figure 4.6), the GMsFEM with five basis functions ($L_i = 5$) can generate a result almost identical to that generated by the reference solution. Therefore, we can conclude that five basis functions are enough to capture the details of 5 Hz waves and can compute the partial-derivative wavefields rapidly without sacrificing accuracy. In contrast, for the 15 Hz source frequency, the sensitivity kernel of the GMsFEM with five basis functions still maintains the right general pattern; however, the sensitivity kernel is contaminated by numerical artifacts (Figure 4.7(b)). These artifacts can be suppressed by increasing the number of basis functions to over seven ($L_i > 7$), as shown in Figure 4.7(c) and 4.7(d).

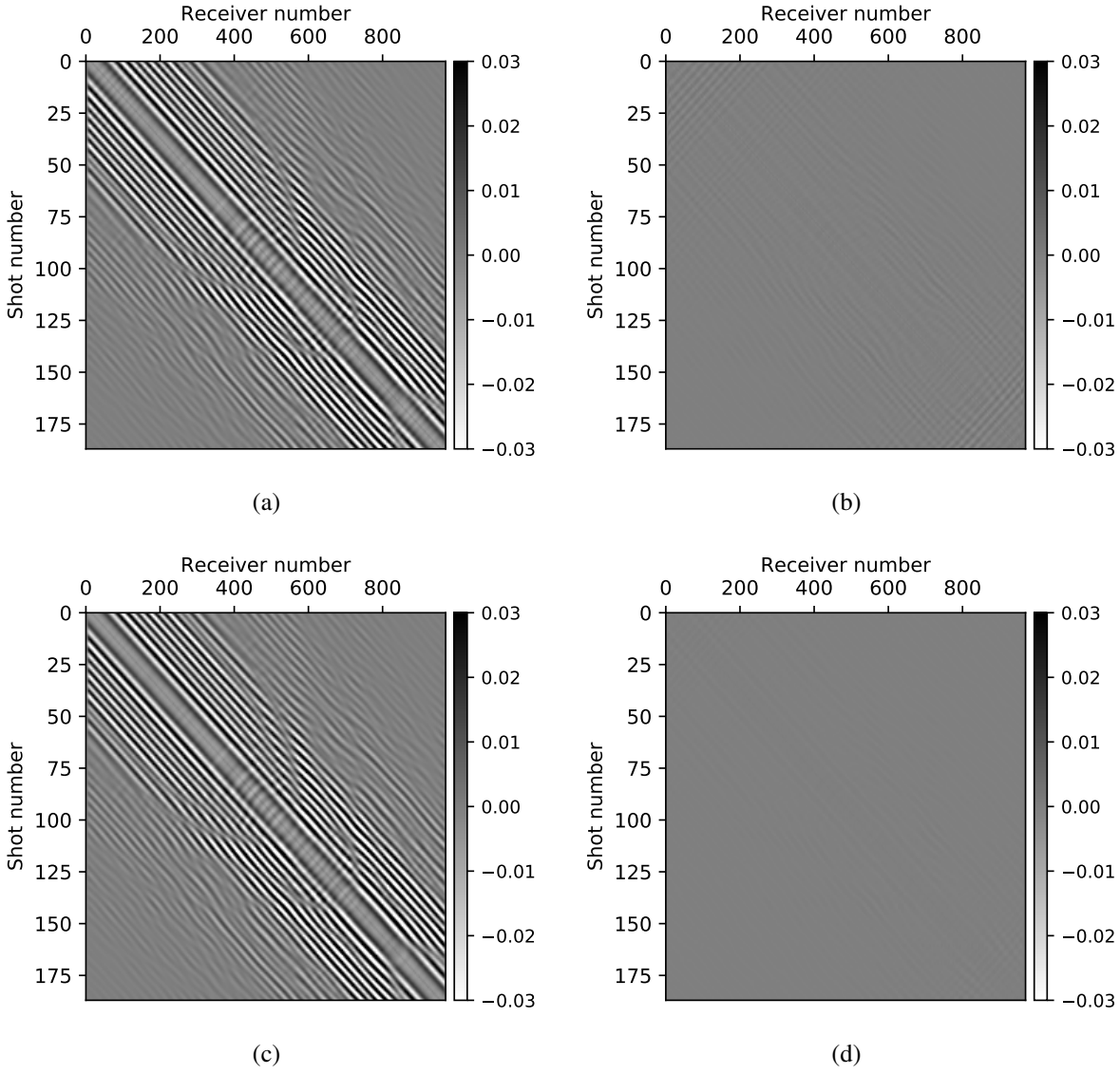


Figure 4.4: Mono frequency (7 Hz) data calculated by using the GMsFEM: (a, b) $L_i = 5$ and (c, d) $L_i = 9$, and the corresponding L_2-error is as follows: $\varepsilon_{L_i=5} \approx 0.08$ and $\varepsilon_{L_i=9} \approx 0.02$. Left panel (a, c) represents the real part of the Fourier components of the modeled data, and right panel (b, d) shows the difference between the reference (CG FEM) and modeled wave from the GMsFEM.

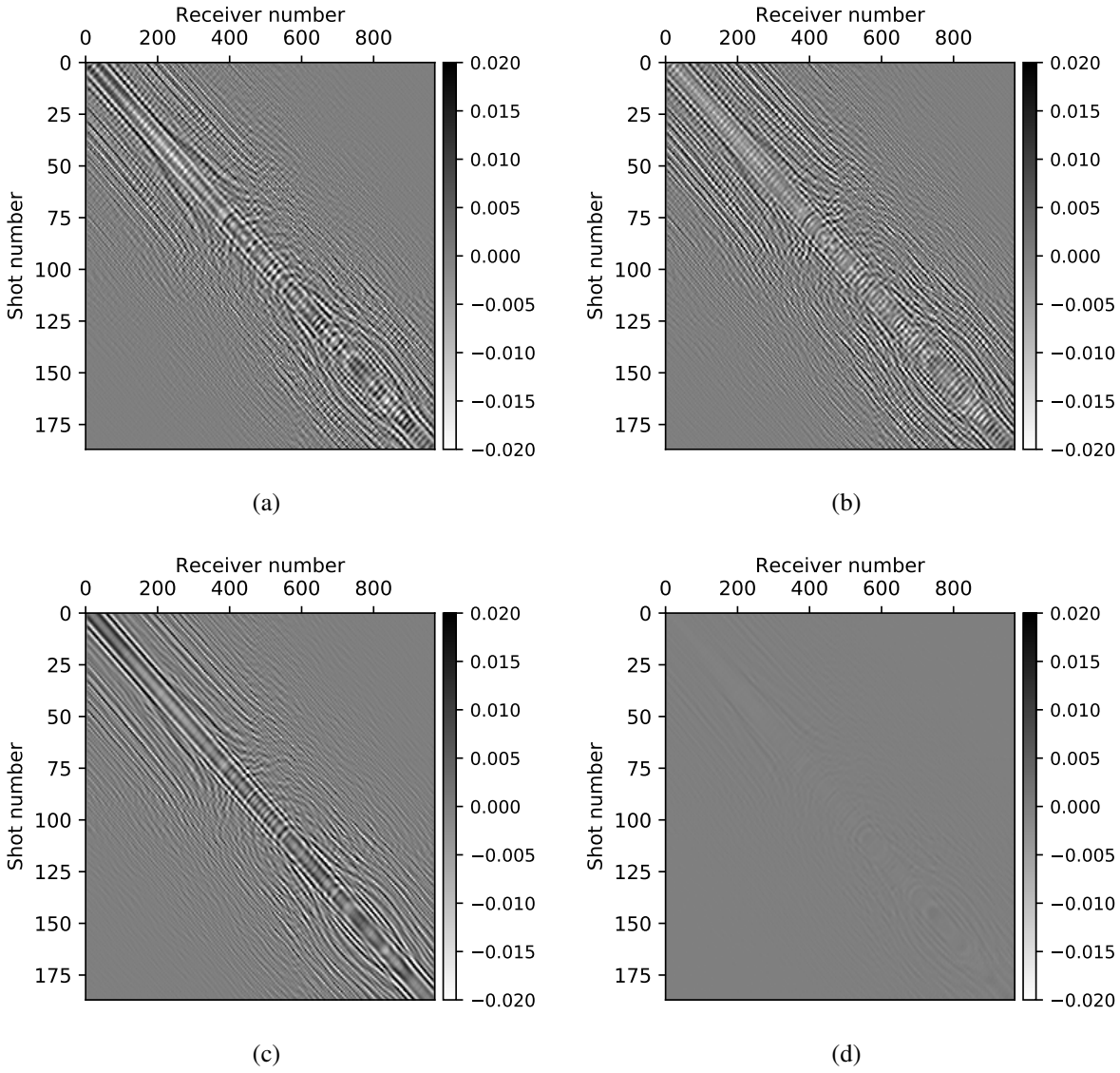
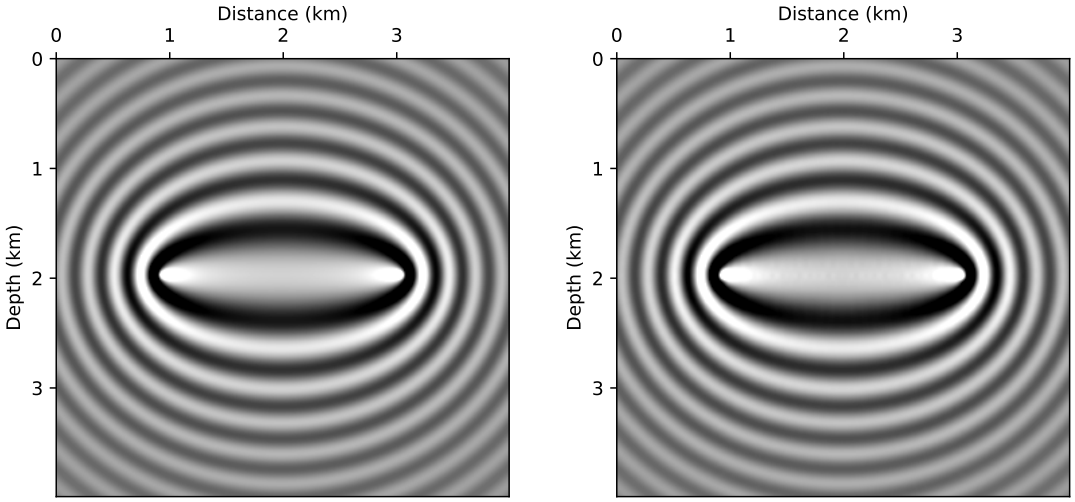


Figure 4.5: Mono frequency (13 Hz) data calculated by using the GMsFEM: (a, b) $L_i = 5$ and (c, d) $L_i = 9$ and the corresponding L_2-error is as follows: $\varepsilon_{L_i=5} \approx 0.87$ and $\varepsilon_{L_i=9} \approx 0.06$. Left panel (a, c) represents the real part of the Fourier components of the modeled data, and right panel (b, d) shows the difference between the reference (CG FEM) and modeled wave from the GMsFEM.



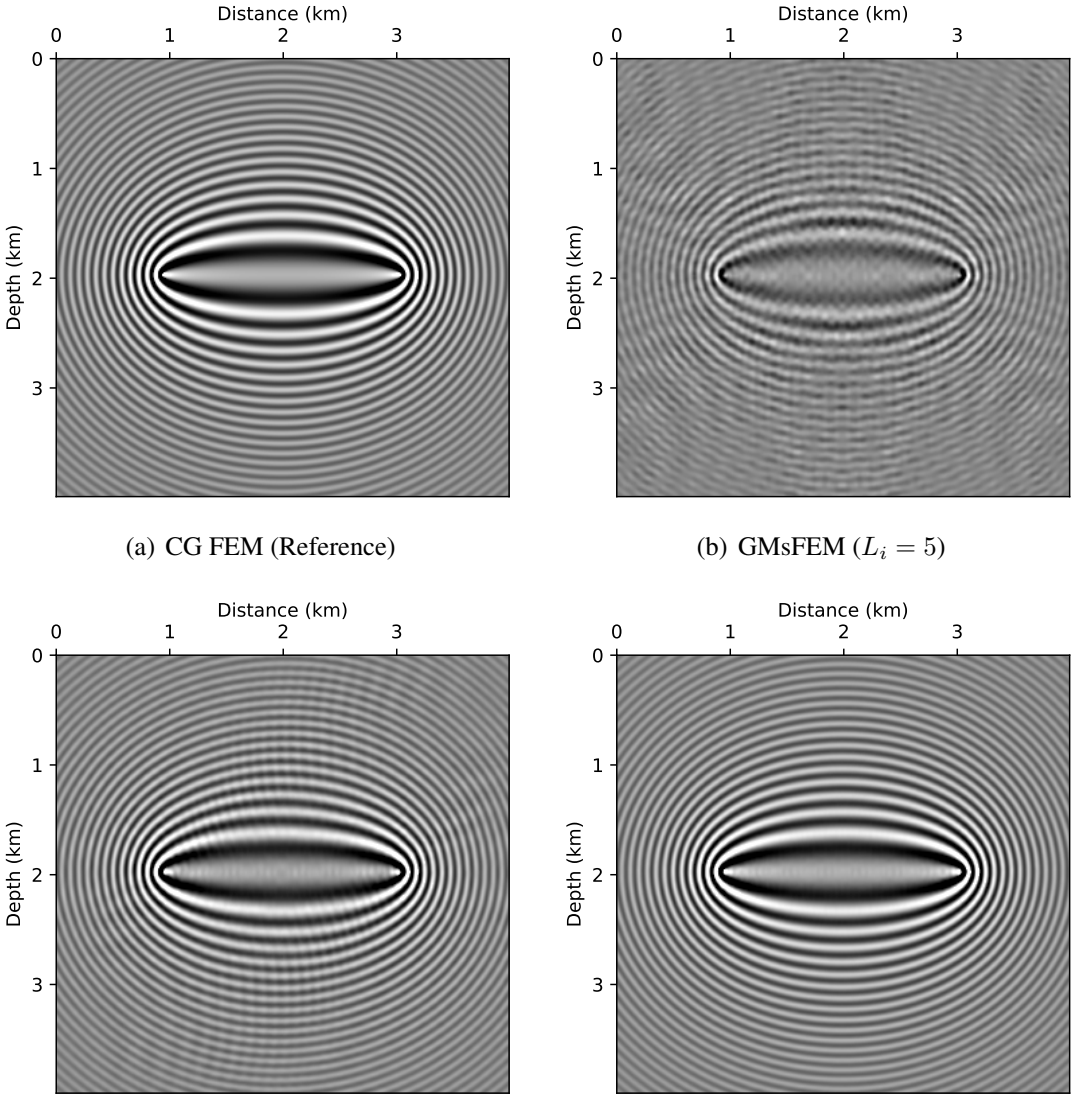
(a) CG FEM (Reference)

(b) GMsFEM ($L_i = 5$)

(c) GMsFEM ($L_i = 7$)

(d) GMsFEM ($L_i = 9$)

Figure 4.6: Comparison of sensitivity kernels (5 Hz).



(a) CG FEM (Reference)

(b) GMsFEM ($L_i = 5$)

(c) GMsFEM ($L_i = 7$)

(d) GMsFEM ($L_i = 9$)

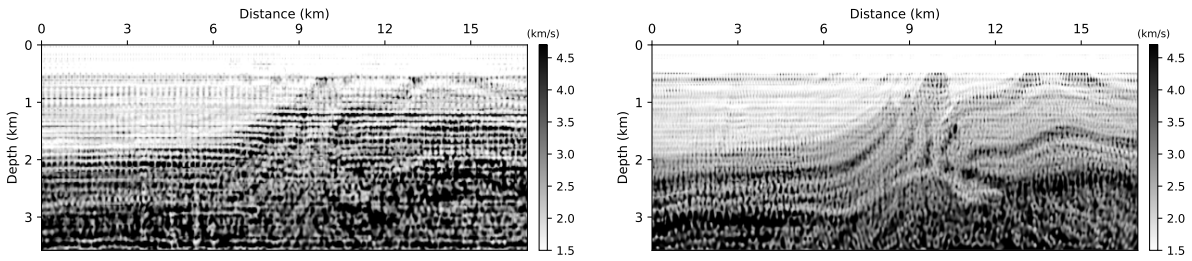
Figure 4.7: Comparison of sensitivity kernels (15 Hz).

4.3.3 Marmousi-2 FWI results

The results of the frequency-domain multiscale FWI are presented in Figure 4.8. The FWI is started from the initial guess which is displayed in Figure 4.2. We applied 500 iterations to acquire the final inversion results. The runtime is measured with 400 processor units. In the GMsFEM, in addition to the number of multiscale basis functions, the coarse-scale grid size can also be altered to enhance the computational efficiency. Figure 4.8(a), 4.8(c), and 4.8(e) present the results of the multiscale FWI with 200 m coarse-scale grid. Although the FWI with five basis functions (Figure 4.8(a)) is contaminated with the trace of coarse mesh, all these artifacts can be attenuated after considering a greater amount of wave modes (Figure 4.8(e)). In this example, we applied the same amount of basis functions to demonstrate the influence of coarse grid size. Given the results shown in Figure 4.8(b), 4.8(d), and 4.8(f), as we expected, a smaller size of coarse grid (100 m) could resolve better subsurface structures with the same basis functions as the larger coarse mesh. In Figure 4.8(b), although the FWI result with five basis functions ($L_i = 5$) is contaminated by numerical artifacts, it can still delineate the macro velocity structure. By utilizing the multiscale FWI approach with nine basis functions (Figure 4.8(f)), we could get the final solutions more rapidly (14,650 s) than the reference case (16,050 s).

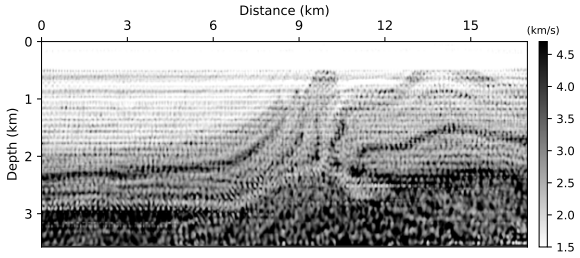
The corresponding error curve as a function of runtime is displayed in Figure 4.9. The error curves with 100 m coarse mesh (Figure 4.9(b), 4.9(d), and 4.9(f)) converge faster than the reference error curve which uses the CG FEM as a modeling engine, though the multiscale FWI with 200 m coarse mesh (Figure 4.9(a), 4.9(c), and 4.9(e)) ended up with a larger error than the reference curve. However, this tool for rapid macro velocity model construction may be useful when we need to quickly construct a background model for the reverse-time migration as it requires smoothed background model. It can also be utilized as an initial guess for the waveform inversion. As Lailly and Bednar (1983) showed that RTM is tantamount to the first iteration of FWI, we can consider the model update in FWI as multiple realizations of gradient computation in RTM. Thus, with this approach we can expect higher speedup in FWI than in the RTM case.

When we need more accurate and detailed earth models, we apply more basis functions as

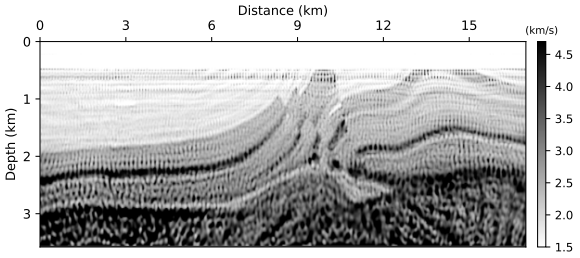


(a) $L_i = 5$ (Total runtime = 3,238 s)

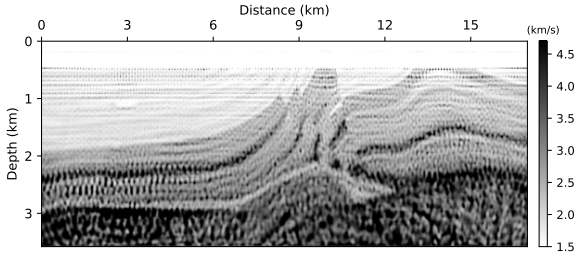
(b) $L_i = 5$ (Total runtime = 6,685 s)



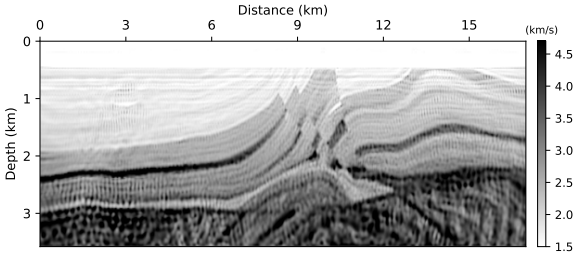
(c) $L_i = 7$ (Total runtime = 8,614 s)



(d) $L_i = 7$ (Total runtime = 10,105 s)



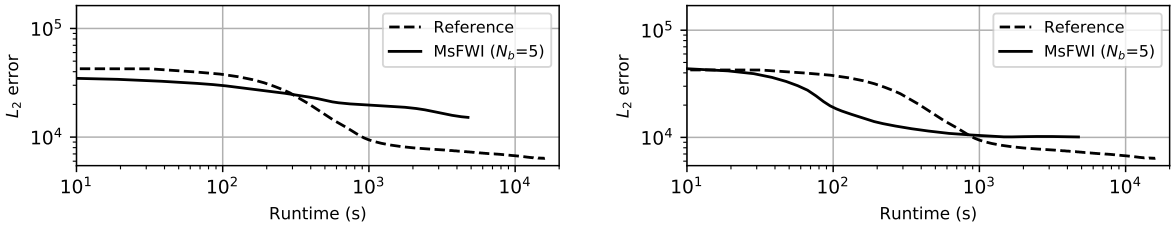
(e) $L_i = 9$ (Total runtime = 10,210 s)



(f) $L_i = 9$ (Total runtime = 14,650 s)

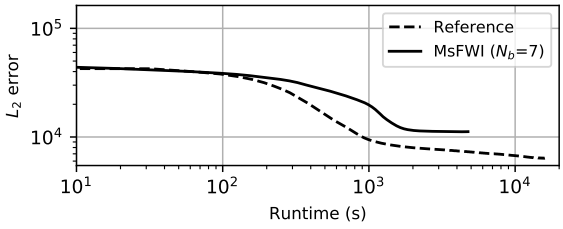
Figure 4.8: GMsFEM FWI results with different numbers of basis functions (reference runtime = 16,050 s).

displayed in Figure 4.8(f). Utilizing nine basis functions ($L_i = 9$) successfully enhances the resolution of the velocity model and helps to capture more detailed geological features. Given the error curve shown in Figure 4.9(d) and 4.9(f), as we achieve the higher accuracy, there is less speedup with GMsFEM. Thus, in GMsFEM, we can enhance the efficiency by tuning the tradeoff between the accuracy and speedup according to the purpose of the FWI.

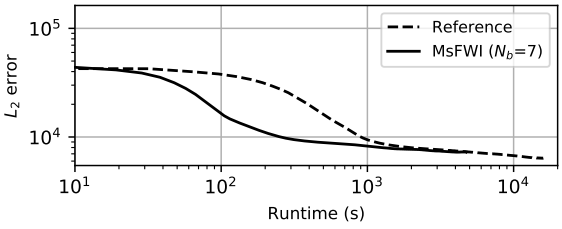


(a) $L_i = 5$ (runtime per iteration = 6.476 s)

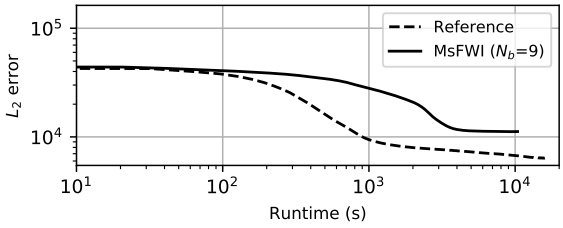
(b) $L_i = 5$ (runtime per iteration = 13.37 s)



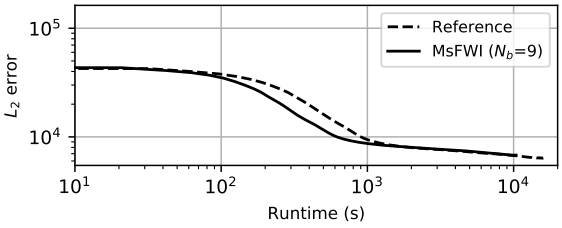
(c) $L_i = 7$ (runtime per iteration = 17.23 s)



(d) $L_i = 7$ (runtime per iteration = 20.21 s)



(e) $L_i = 9$ (runtime per iteration = 20.42 s)



(f) $L_i = 9$ (runtime per iteration = 29.3 s)

Figure 4.9: l_2 error versus runtime. Captions also note average runtime per iteration (reference time = 32.1 s). Note that both axes are displayed in log scale, so the actual convergence rate is faster than the apparent slope of the error curve.

4.4 Numerical examples: field data

We demonstrated the multiscale FWI by using the Marmousi-2 synthetic data. However, even though a certain inversion scheme could recover the subsurface properties successfully, applying the same algorithm to the field dataset is still a challenging problem. Moreover, when the acquisition region includes complicated structure such as karst, salt or volcanic diapirs, constructing a velocity model via FWI becomes more difficult. In this section, we applied the multiscale FWI method to the field data collected on the the Gulf of Mexico. The seismic survey incorporates a salt diapir located the middle of area. The field dataset consists of 399 shots with 408 channels. The shot and receiver intervals are 50 m and 25 m, respectively. The maximum offset is 10,292 m. The total recording time is 12 s with 4 ms of sampling interval. As a part of pre-processing, we applied high-pass frequency filter by removing the frequency components under 4 Hz. We also muted the noise arriving before the first break. We displayed the near-offset gather in Figure 4.10. Provided that the salt is intruded up to the near surface, the main goal of the multiscale FWI would be resolving the flank of the salt dome more rapidly than the reference case without degrading the accuracy or image resolution.

A robust and ideal FWI algorithm might be able to make the solution converge to the global minimum from a simple starting model such as homogeneous or linearly increasing 1-D model. However, in real world problems with FWI, it is hard to recover the velocity model without having a good prior information due to geophysical non-uniqueness problem and issue on the signal to noise ratio. In addition, since we eliminated the low frequency part ($\sim$4 Hz) which might be critical to infer the macro velocity structure, the lack of low frequency information hinders the solution converging to the global minimum. Figure 4.11(b) shows the results of frequency-domain FWI starting from the 1-D velocity model (Figure 4.11(a)). Although the model could roughly delineate the location of salt top and the shape of strata, it could not update the velocity of the salt dome and suffering from the local minimum. There are alternatives to build a good priori model such as RMS velocity picking, travel time tomography, or salt plotting. In this research, before performing the actual frequency-domain FWI, we applied the Laplace-domain FWI (Shin and Cha,

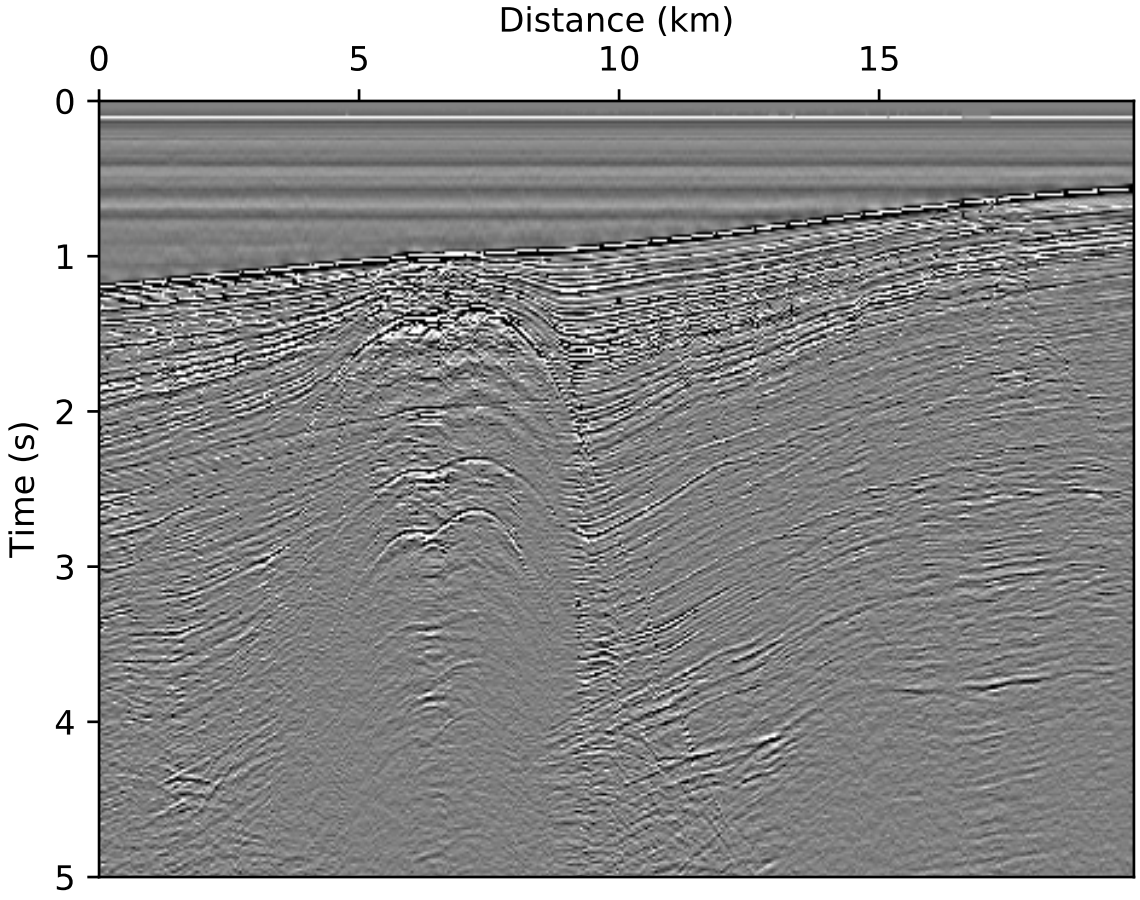


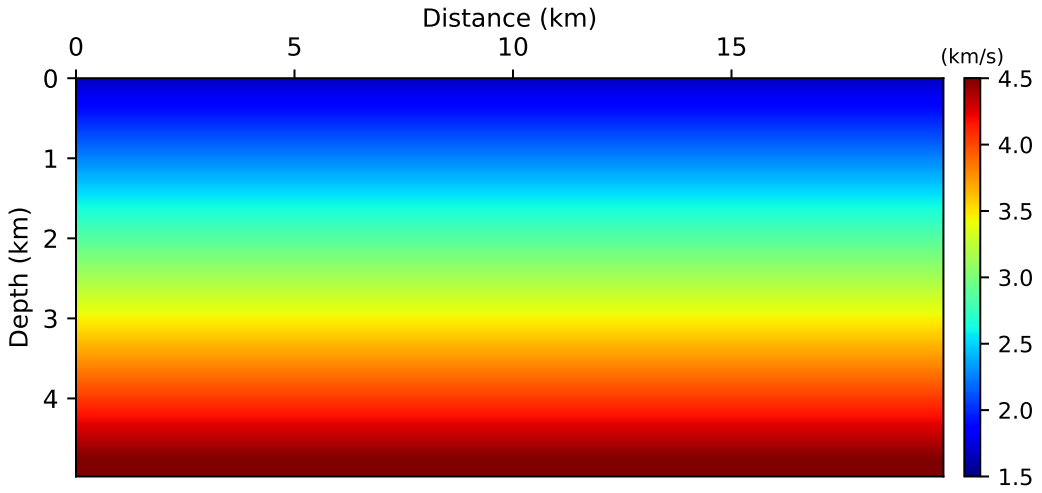
Figure 4.10: Near-offset gather of the field data acquired from Gulf of Mexico region.

2008) to determine an initial velocity model. One notable feature of the Laplace-domain FWI is that it utilizes damped wavefields to enhance the converging rate by generating a smoothed forward and backward modeling results. Also, as the Laplace-domain FWI uses coarser mesh due to a generous dispersion condition, it requires less computational costs compared to the conventional frequency domain FWI. We applied 6, 8, and 10 $(1/s)$ Laplace damping constants combined with low temporal frequencies (0.1$\sim$5 Hz), and acquired smoothed initial velocity model as presented in Figure 4.12(a).

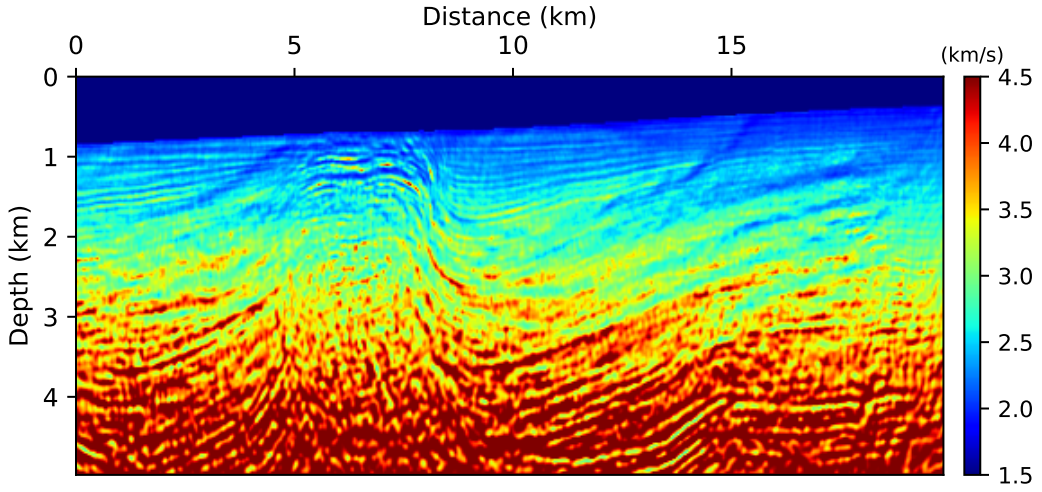
In the GMsFEM approach, we also need to incorporate the density information for computing the multiscale basis functions (Chung et al., 2011; Gibson et al., 2014; Fu and Gao, 2017). Therefore, after obtaining the velocity information through the Laplace-domain FWI, we applied the Gardner's relation (Gardner et al., 1974) to define the density of the clastic sediments. We then assigned a constant density value to the salt body based on the velocity information as presented in Figure 4.12(b). We build the multiscale basis functions by solving the local spectral problems (equation 4.3) based on the given initial model.

4.4.1 Forward & Backward modeling analysis

Many scientists have demonstrated the accuracy and speedup of the GMsFEM scheme using synthetic models (Gao et al., 2015; Cho et al., 2017a; Cho and Gibson, 2018); however, in this paper, we are to evaluate the GMsFEM through comparisons with field data in forward and backward wave modeling. We first investigated the forward modeled wavefields (Figure 4.13). In the previous examples with Marmousi-2 data, we showed data matrices to measure the overall error of the forward modeling; however, in these examples, we present two different mono-frequency solutions to observe the variation of the wavefield as the velocity model updated. In Figure 4.13, left (a, c, e) and right (b, d, f) panels show the amplitude of wavefields from 6 Hz and 12 Hz, respectively. Figure 4.13(a) and 4.13(b) make comparisons between the observed wavefield (gray line) and the modeled wavefield from the initial property models which is shown in Figure 4.12. The offset of amplitude is attributed to the Green's function at the first iteration, which is computed from the point source with amplitude of 1. Since estimation of a source signature to match the am-

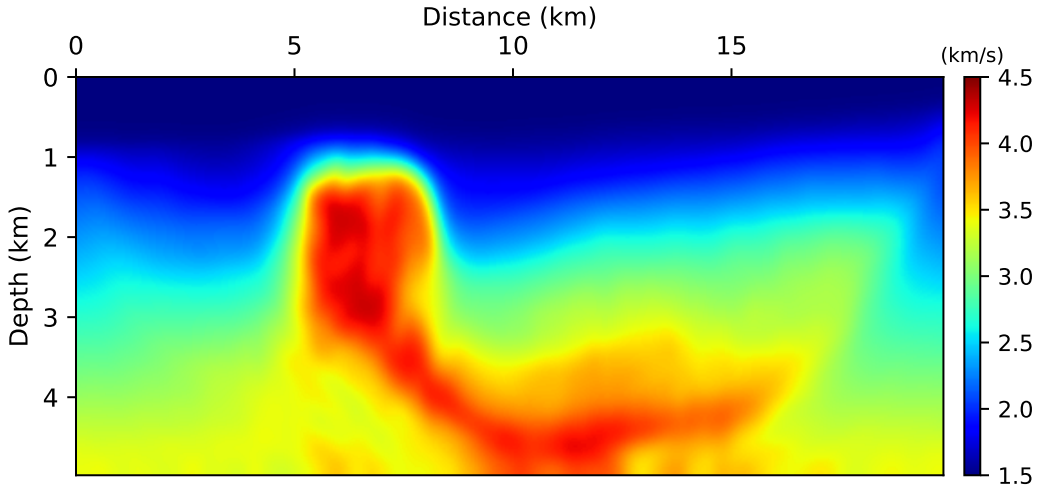


(a)

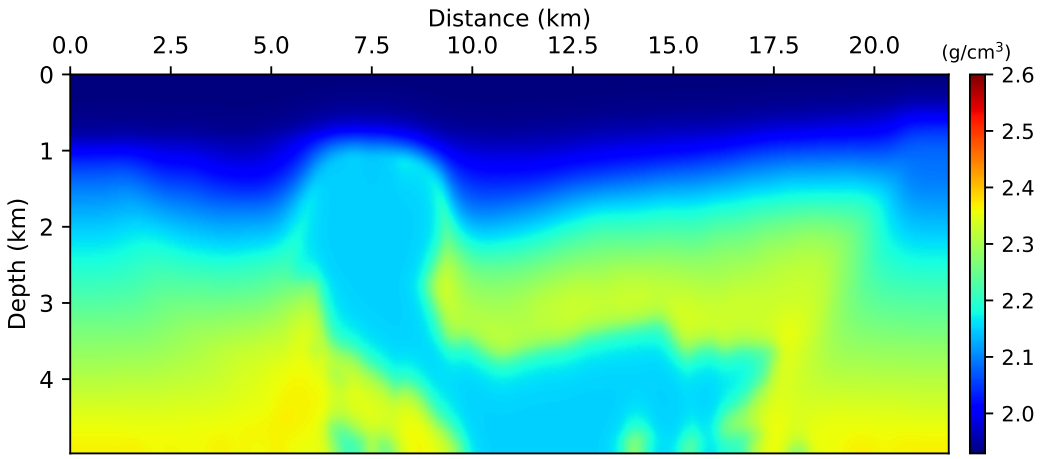


(b)

Figure 4.11: (a) Linearly increasing 1-D velocity model, and (b) conventional frequency-domain FWI results after 500 iterations.



(a)



(b)

Figure 4.12: (a) Initial velocity model acquired from the Laplace FWI, and corresponding (b) density model inferred from the velocity model.

plitude between observed and modeled data is a critical process which can directly contribute to the success of the FWI (Pratt, 1999), we estimated the source wavelet at every iteration based on the updated models. In this research, we implemented the source estimation by separating the amplitude and the phase of a complex source function by following Shin and Min (2006)'s formulation. The major update of the source amplitude is occurred during the first iteration, and Figure 4.13(c) and 4.13(d) show the wavefield after the first round of source estimation. After fitting the source signature, the FWI starts the actual procedure to update the velocity model. The wavefields after 500 iterations are presented in Figure 4.13(e) and 4.13(f). As presented in the synthetic examples, similarly, the GMsFEM with fewer basis functions ($L_i = 5$, black solid line) could simulate the low frequency wave accurately in shorter amount of time (5.81 s) than the reference case (9.45 s). In contrast, there occurs a greater amount of offset between the observed and modeled data using the GMsFEM ($L_i = 5$ and 7) in higher frequency wavefields. Therefore, we apply more basis functions ($L_i = 9$) to simulate 12 Hz wavefield correctly. In this case, the GMsFEM still can compute the wavefield more rapidly (9.09 s) than the reference case (9.44 s).

Cho et al. (2017a) proposed the RTM method combined with the GMsFEM wave modeling, and they demonstrated the seismic imaging examples with the analysis on the wave modeling. In the RTM case, the source of the back-propagation would be a time-reversed observed data. Hence, when we apply the GMsFEM, the inherited numerical artifacts that are generated from the coarse modeling will only have an influence on the forward modeling part. In contrast, in the case of FWI with the GMsFEM, since the misfit of the observed and simulated wavefield would be the source of the back-propagation, we also need to perform the analysis on the errors which might be generated from the wave modeling in the coarse-scale mesh. In this regards, the residual spectra are presented in Figure 4.14 which are obtained in the frequency domain at frequencies of 6 Hz (Figure 4.14(a)) and 12 Hz (Figure 4.14(b)). Since the error values for the two different frequencies are plotted with the same scale, we see that the the magnitude of error is greater in 12 Hz wave example than in 6 Hz one. According to Cho and Gibson (2018)'s work, when the L_2 norm of the numerical error caused by GMsFEM is smaller than 10 %, the GMsFEM error would have little

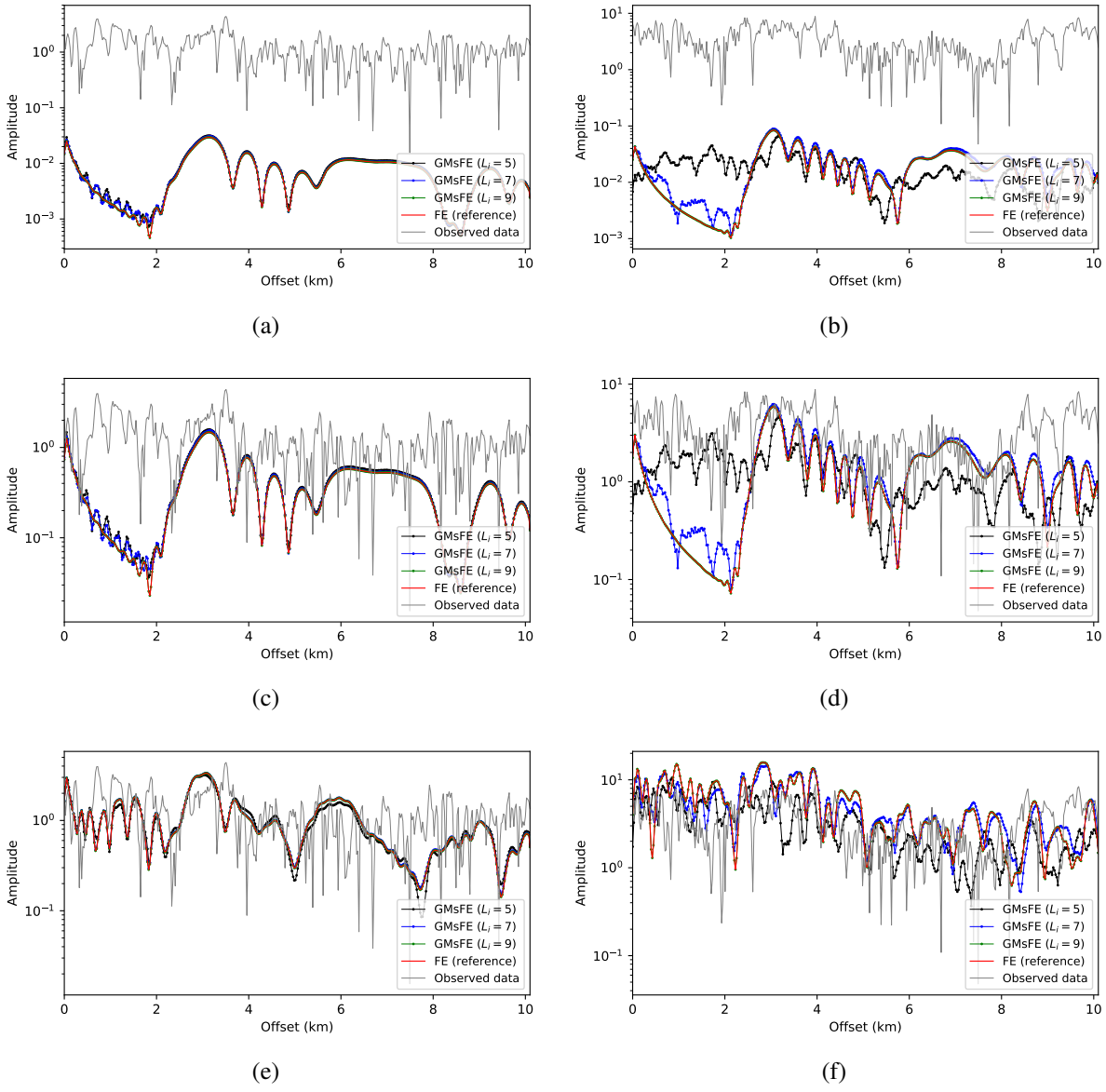


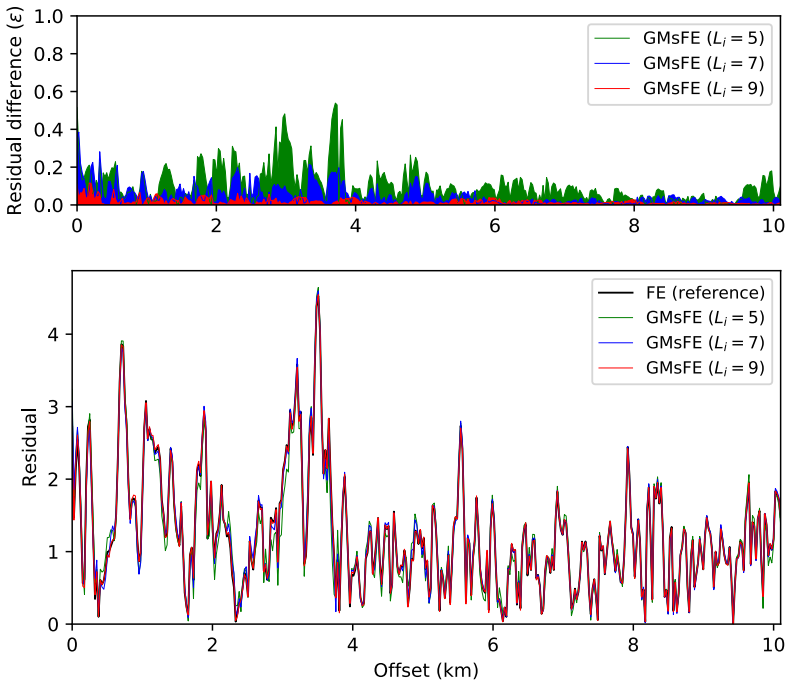
Figure 4.13: Comparisons between the modeled wavefield using the GMsFEM ($L_i = 5, 7, 9$) and observed data (200th shot). Left (a, c, e) and right (b, d, f) panel display 6 Hz and 12 Hz wavefields. Upper panel (a, b) shows the waves before applying the source estimation (after the first iteration), and middle panel (c, d) presents the wavefield after the first round of source estimation. The wavefields after 500 iterations are displayed in lower (e, f) panel.

influence on the final gradient image for the model update due to destructive interference among multiple shot and receiver pairs. To demonstrate the impact of the modeling artifact to the FWI in GMsFEM, the result of multiscale FWI will be presented in the following section.

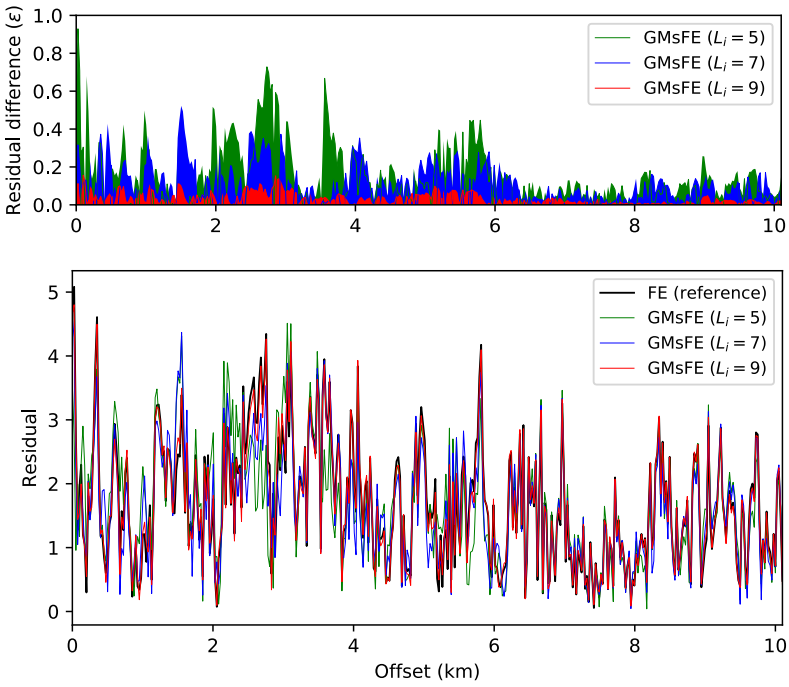
4.4.2 FWI results analysis

For the frequency-domain FWI, we used 50 frequency components with an increment of 0.2 Hz. The minimum and maximum frequencies are 4 Hz and 13.8 Hz, respectively. The fine-scale grid is 25 m, and we superimposed the coarse-scale grid (125 m) on the fine mesh. We applied 500 iterations to acquire the final FWI results and used a fixed step length to update the model ($\alpha = 0.02$). Figure 4.15 presents the comparisons of the gradient images that are calculated from the multiscale modeling scheme with different numbers of basis functions. The left panel shows the gradient of earlier stage of the model update (50 iterations), and the right panel displays the gradient images of the last (500) iterations. Because all the gradient images are plotted under the same color scale, we see that the amount of model update reduces as the FWI proceeds. Figure 4.16 exhibits the error curve as a function of runtime. As observed in the previous Marmousi-2 examples, even though the multiscale FWI with few basis functions ($L_i = 5$, black curve in Figure 4.16) could complete the computation more rapidly than other cases, it could not reduce the error sufficiently due to dispersion error in wave modeling. The dispersive waves generate a chaotic pattern inside of the salt diapir as shown in Figure 4.15(a) and 4.15(b). The chaotic patterns become reduced as we apply a greater number of basis functions. The multiscale FWI results with nine basis functions (Figure 4.15(e)) could produce almost identical gradient images with the reference case (Figure 4.15(g)), and it still provides the FWI results 12 % faster than the reference as displayed in Figure 4.16 with the green curve.

The final results of multiscale FWI are displayed in Figure 4.17. All of the multiscale FWI results show similar macro velocity structures. However, in Figure 4.17(a), due to an insufficient number of basis functions, the location of coarse mesh appears as highlighted with a red dashed circle. Also, the velocity model shows chaotic patterns around the salt body. In these FWI examples, we used the entire frequency band to demonstrate the influence of the number of basis functions



(a)



(b)

Figure 4.14: Residual spectrum of the 200th shot after 500 iterations, which are calculated from the wave examples displayed in Figure 4.13(e) and 4.13(f). In each subfigure, lower panel shows the residual spectrum, and upper panel shows the error of the GMsFEM modeling with different number of basis functions.

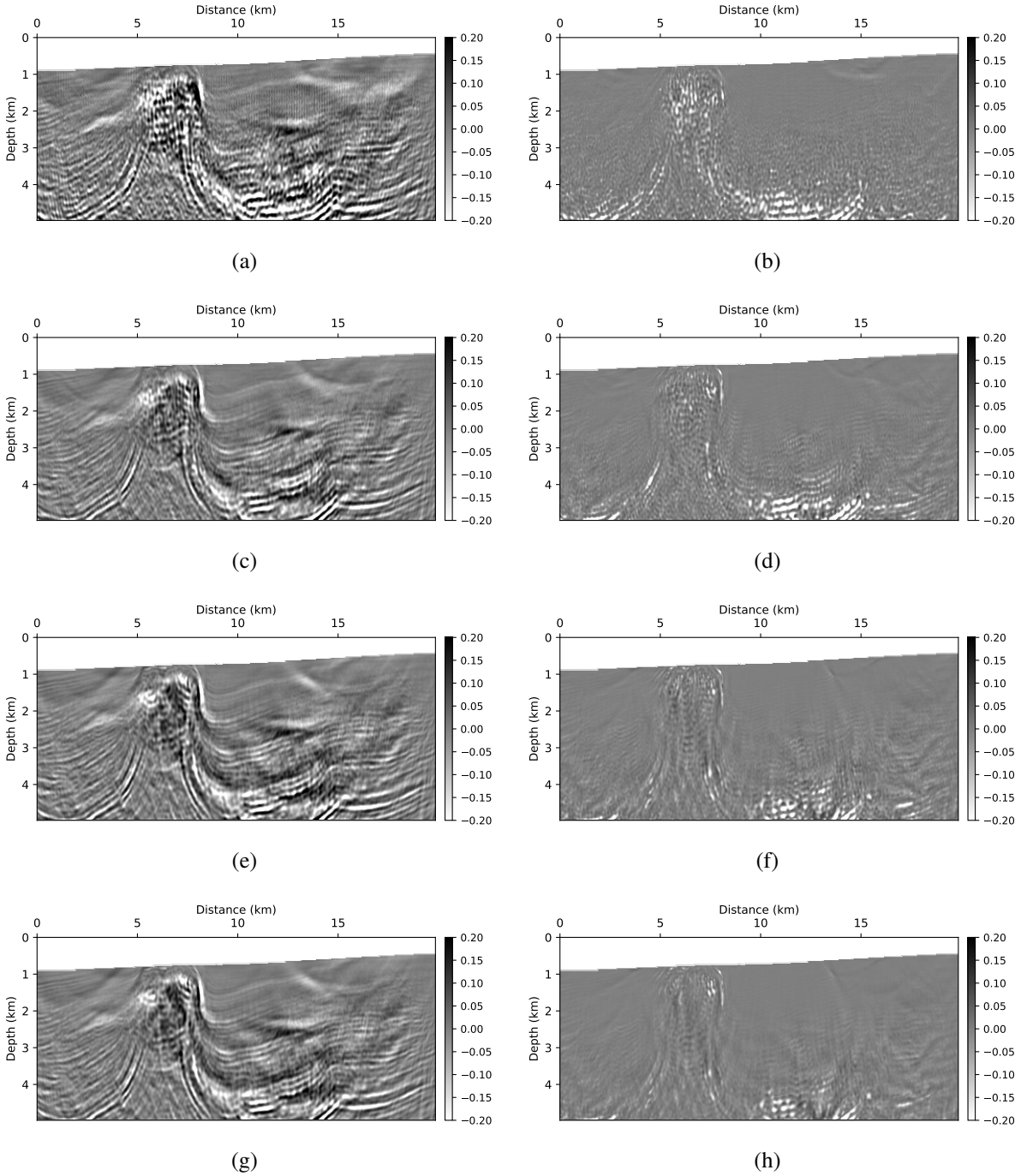


Figure 4.15: Comparisons of the gradient images at different iterations: left and right panels show the gradient of 50 and 500 iterations, respectively. Each row exhibits the gradients computed with different number of basis functions: (a, b) $L_i = 5$, (c, d) $L_i = 7$, (e, f) $L_i = 9$, and (g, h) reference case.

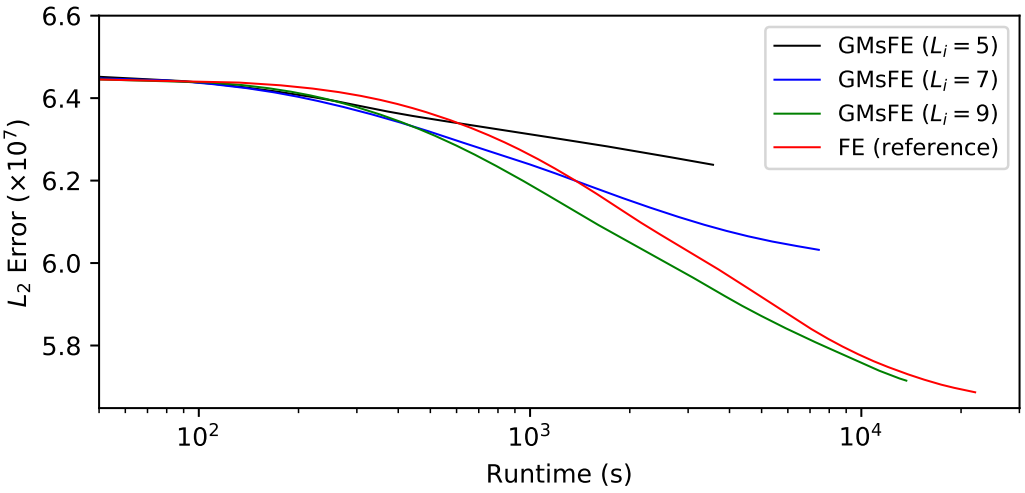


Figure 4.16: Comparison of the error curves from different multiscale FWI (500 iterations).

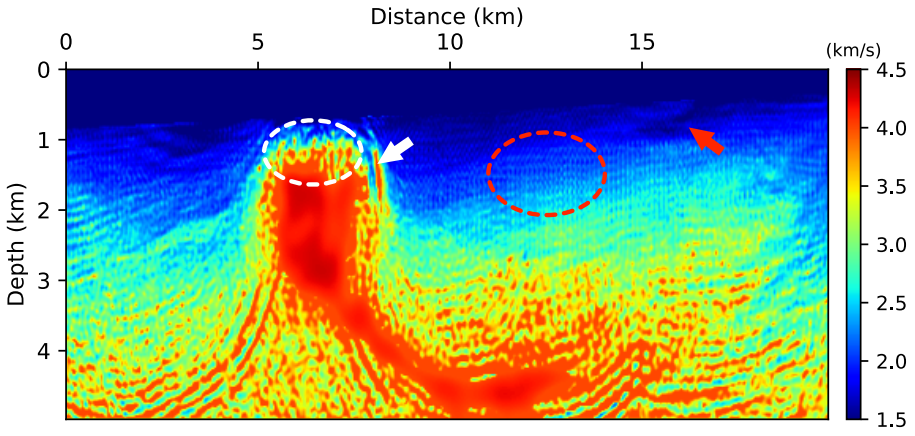
on the final FWI results. Hence, the results contain incomplete features such as migration artifacts (pointed with white arrow), high-frequency noise around the salt top, and cyclic skipping effect as pointed with the red arrow. Regarding the noisy patterns, all the noise and aliasing errors are attributed to the reason that we tried to simulate high frequency waves by using a coarse mesh with few basis functions which is not sufficient to capture the detailed variation of the high frequency waves. To put it differently, in the multiscale FWI, we can compute the low frequency gradient images more rapidly without sacrificing the accuracy. Therefore, we propose a strategy to optimize the computation speed of the FWI through the multiscale modeling scheme. For example, we divide the target frequency band (4$\sim$13.8 Hz) into multiple bands and tune the number of basis functions according to the frequency components. In detail, we apply more basis functions as we move on to the other band with higher frequencies. Table 4.1 shows the strategy that we used for applying multiple frequency groups which accumulate the frequency components by maintaining the low frequency band. By applying more iterations to the groups with low frequency components (4$\sim$8 Hz), the multiscale FWI builds the macro velocity structures at early stages of the model update. Also, at low frequency group, since we can use few basis functions, we achieve higher speedup compared to the high frequency group. Figure 4.18 displays the normalized error

Frequency group	Temporal frequencies (Hz)	Basis functions (L_i)	Iterations
1	$4.0 \cdots 5.8$	5	150
2	$4.0 \cdots 7.8$	6	150
3	$4.0 \cdots 9.8$	7	100
4	$4.0 \cdots 11.8$	8	100
5	$4.0 \cdots 13.8$	9	300

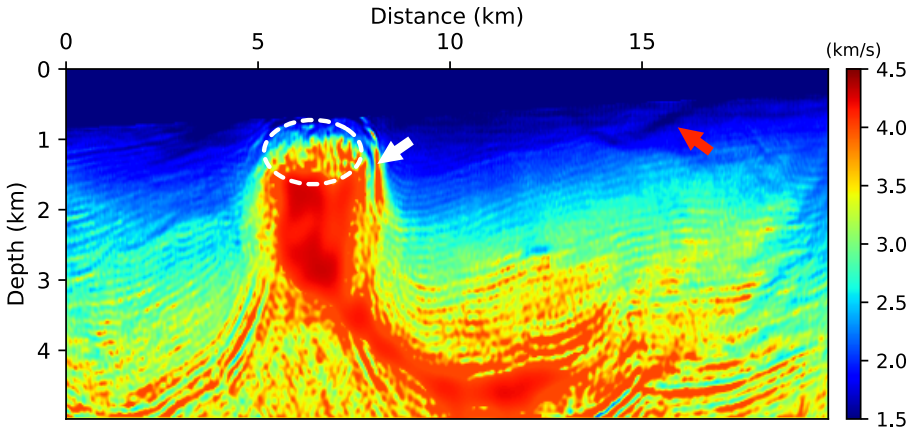
Table 4.1: Strategy for utilizing complex frequencies in the multiscale FWI.

curve of the frequency-adaptive multiscale FWI. At the transition point of each frequency group, there occurs a jump on the error curve since we newly incorporate more frequencies in inverse problem. Also, as we move on to the last frequency group, the rate of convergence becomes slow. We displayed the error curve as a function of iterations (Figure 4.18(a)) to compare the accuracy of multiscale FWI with the reference FWI. According to error curve shown in Figure 4.18, the L_2-error diminishes more rapidly at the first and second frequency groups with the multiscale FWI (red line) than the reference case (black line). The final FWI result of the frequency-adaptive approach is presented in Figure 4.19. As we put more weight on the lower frequency, all the issues that are raised in the examples shown in Figure 4.17 such as noisy salt top and migration artifact are resolved.

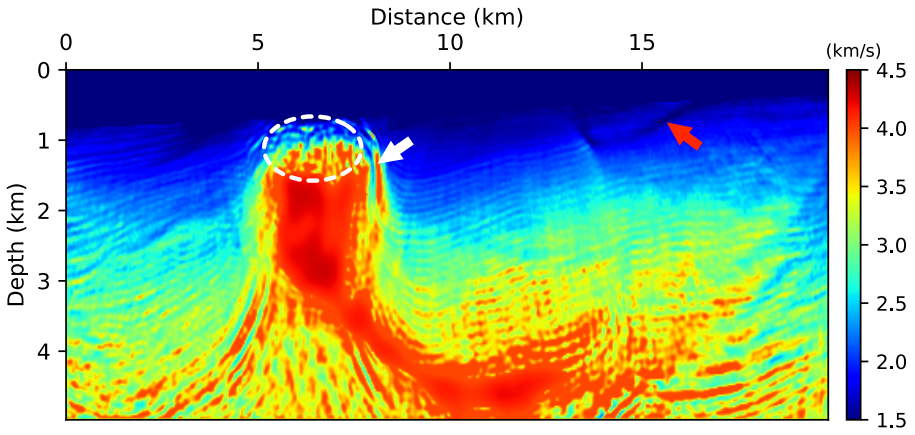
For the validation of the FWI results, we performed the Kirchhoff depth migration using final FWI result (Figure 4.19) as a background velocity. Comparing the results with the near-offset gather (Figure 4.10), the Kirchhoff depth migration (Figure 4.20(a)) could resolve the salt boundary and the strata with steep slope which terminate around the salt flank. To make comparisons between the location of reflectors and the velocity structure, we superposed the FWI results on the migration section as presented in Figure 4.20(b)). For further validation of the velocity model, we sampled a number of common-image-gathers (CIG) as displayed in Figure 4.21. Provided that the third CIG gather from the left penetrates the salt diapir, some of the reflectors and multiples associated with the salt body and sea bottom are not well flattened compared to other reflectors. Nevertheless, most of the reflectors are flattened, and it represents that the velocity acquired from



(a)

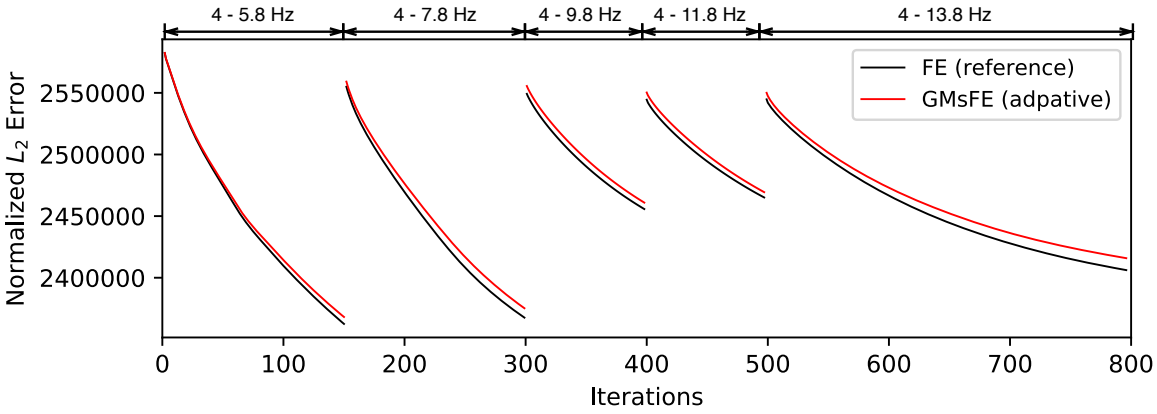


(b)

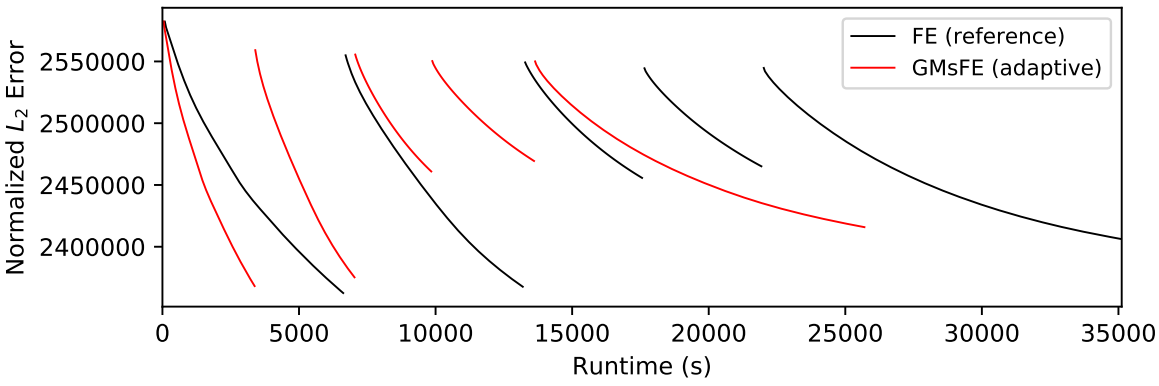


(c)

Figure 4.17: The multiscale FWI results varying the number of basis functions: (a) $L_i = 5$, (b) $L_i = 7$, and (c) $L_i = 9$.

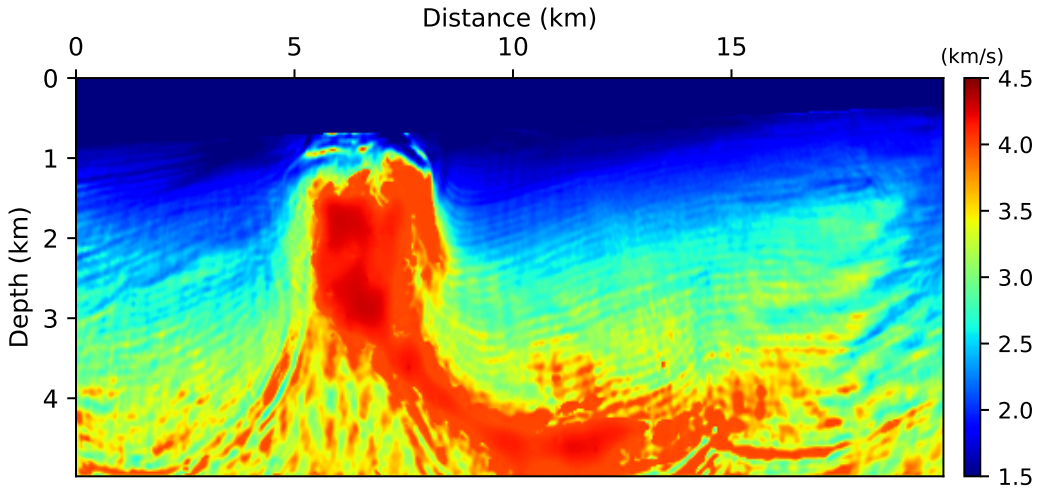


(a)

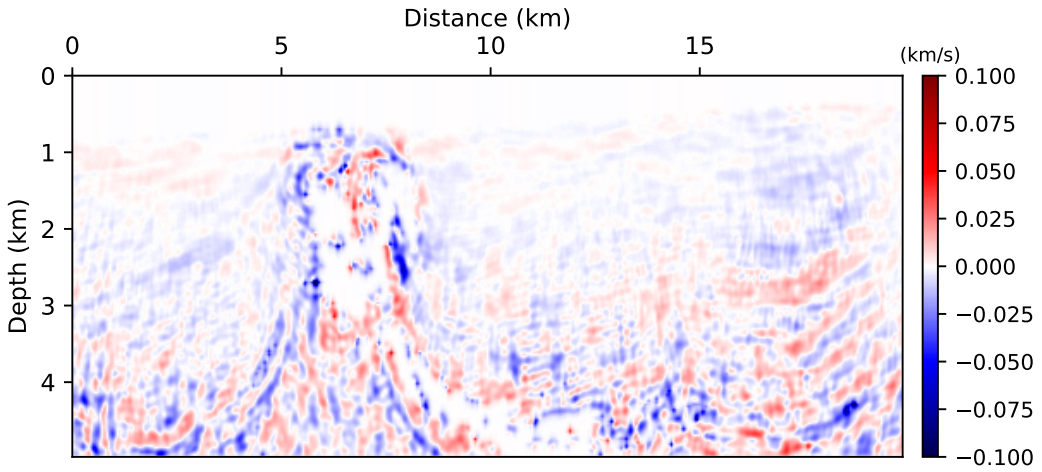


(b)

Figure 4.18: Normalized error curves as a function of (a) iterations and (b) runtime.

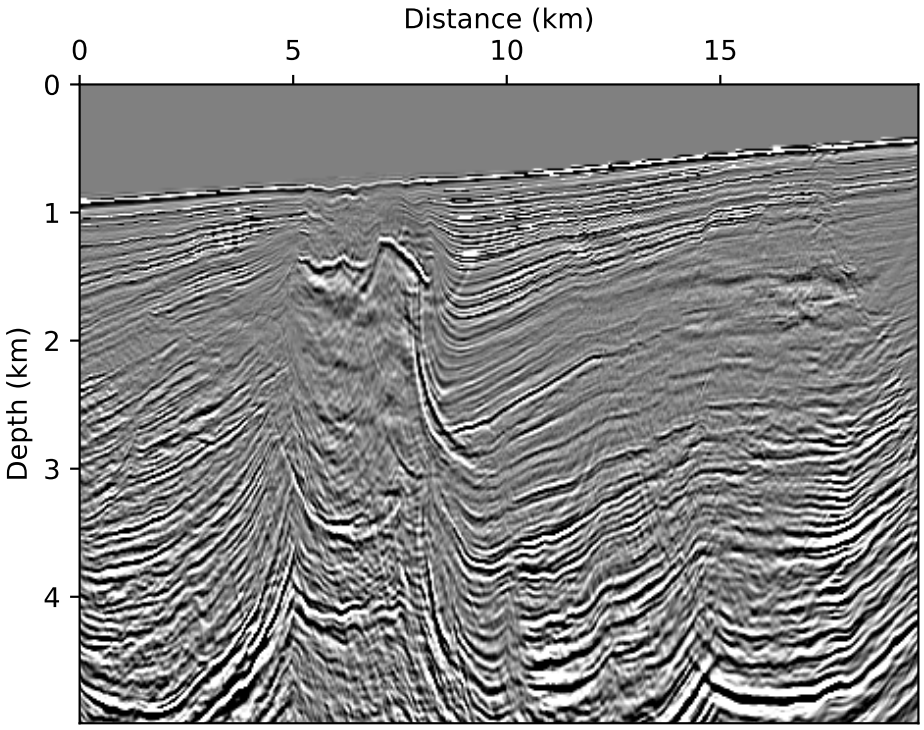


(a)

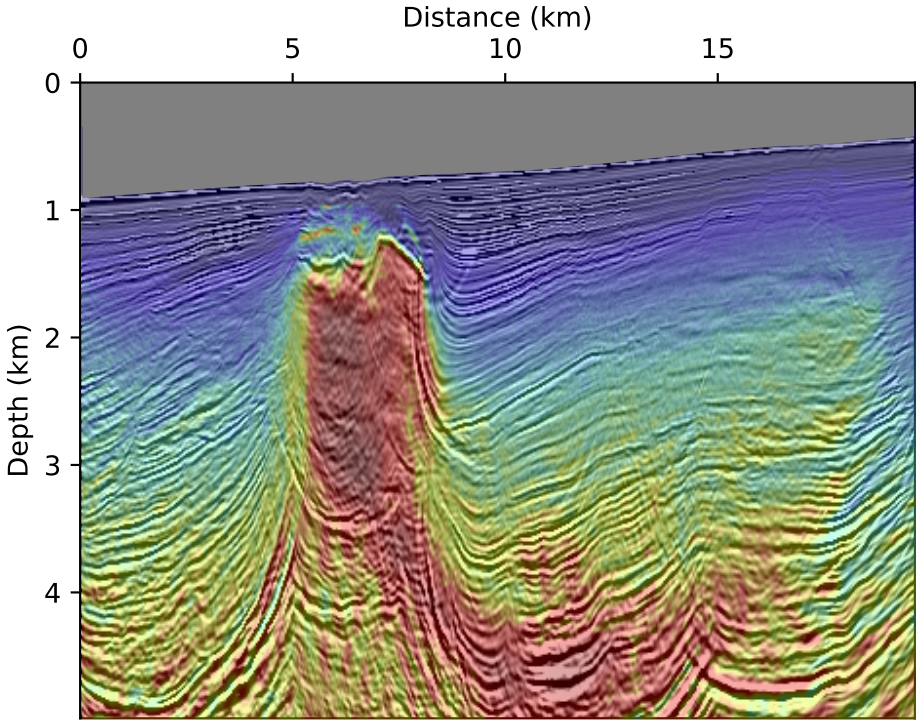


(b)

Figure 4.19: (a) Velocity model acquired from the multiscale FWI with multiple frequency loop and adaptive multiscale basis functions, and (b) difference of velocity model between the multiscale FWI and reference FWI.



(a)



(b)

Figure 4.20: (a) Kirchhoff depth migration section, and (b) the migration section superimposed on the FWI result shown in Figure 4.19.

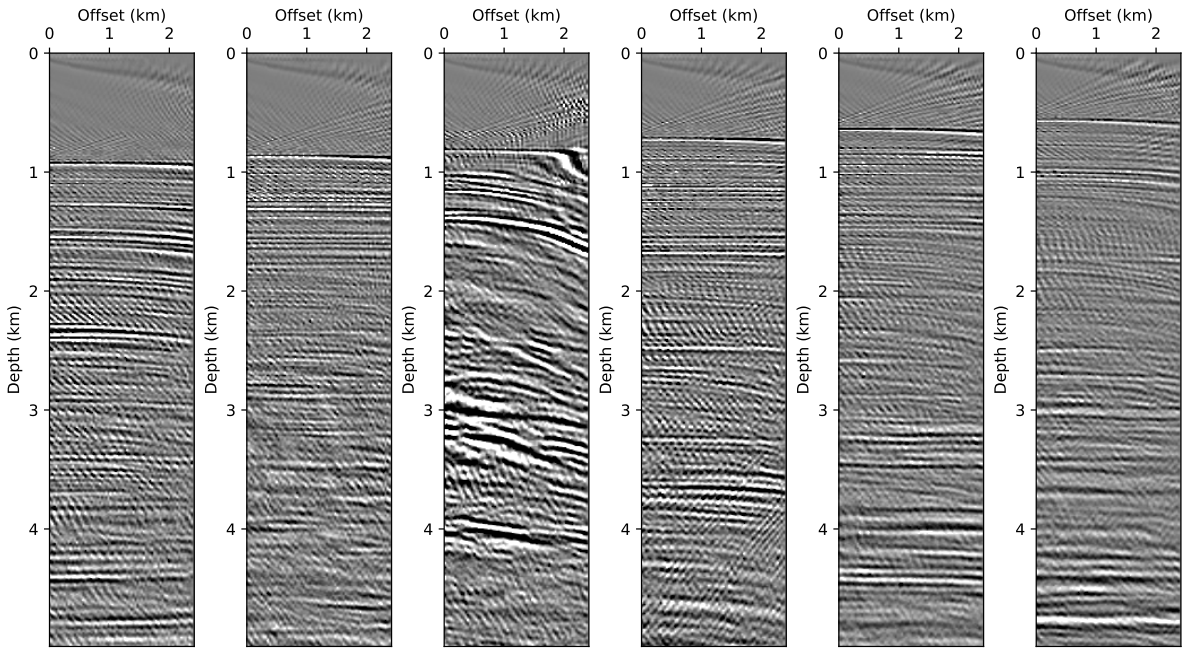
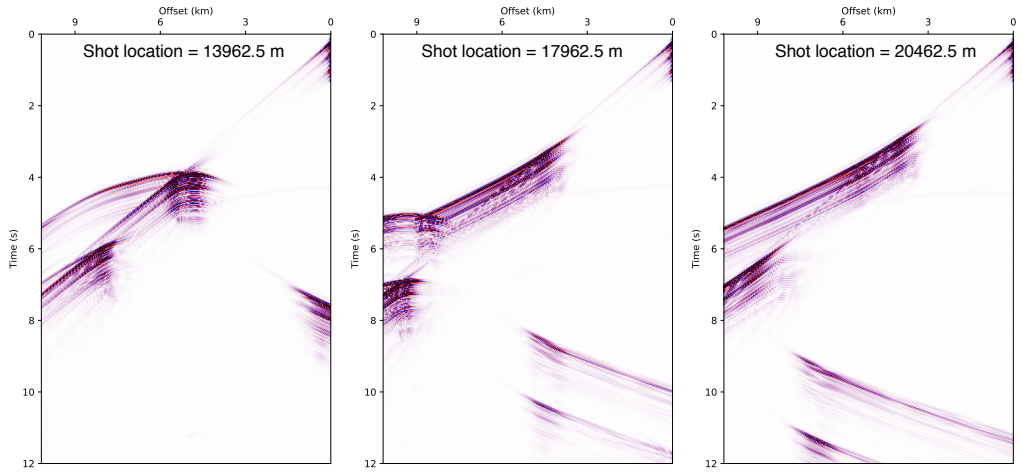


Figure 4.21: Common image gathers (CIG) from the Kirchhoff depth migration. The sampled points are as follows: 1.25, 3.75, 6.25, 11.25, 13.75, 16.25 (km) from the left.

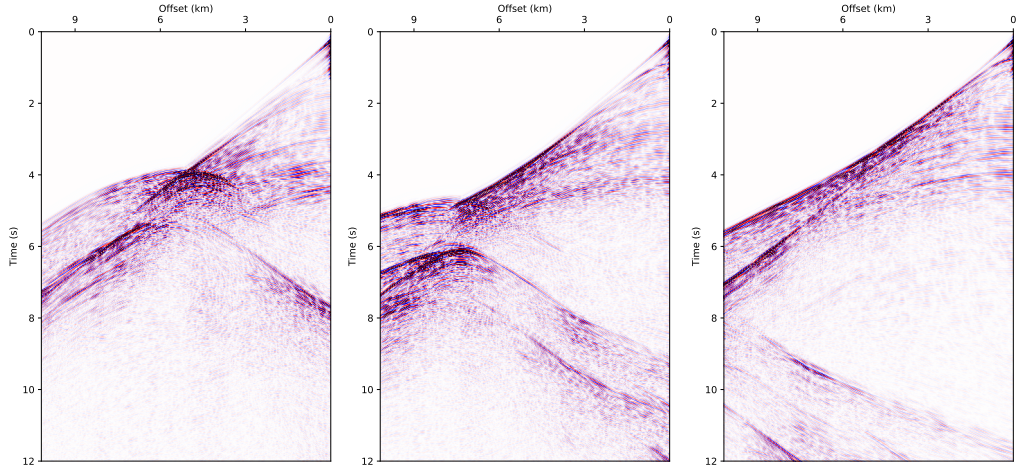
the multiscale FWI is reliable. Since the flattened CIGs do not always guarantee that the velocity is accurate, we also generated shot gathers at three different points (13962.5; 17962.5; 20462.5 m) as presented in Figure 4.13. Although the shot gathers (Figure 4.22(a)) that are calculated from the initial velocity model (Figure 4.12(a)) could generate the major events which might originated from the salt body, it could not create all the detailed reflections. In contrast, the shot gathers (Figure 4.22(b)) from the multiscale FWI result (Figure 4.19(a)) could simulate the wavefields which are kinematically close to the observed data (Figure 4.22(c)). Therefore, we can conclude that the GMsFEM could not only accelerate the wave modeling procedure but also could provide a reliable earth structure by applying it to the frequency-domain FWI.

4.5 Conclusions

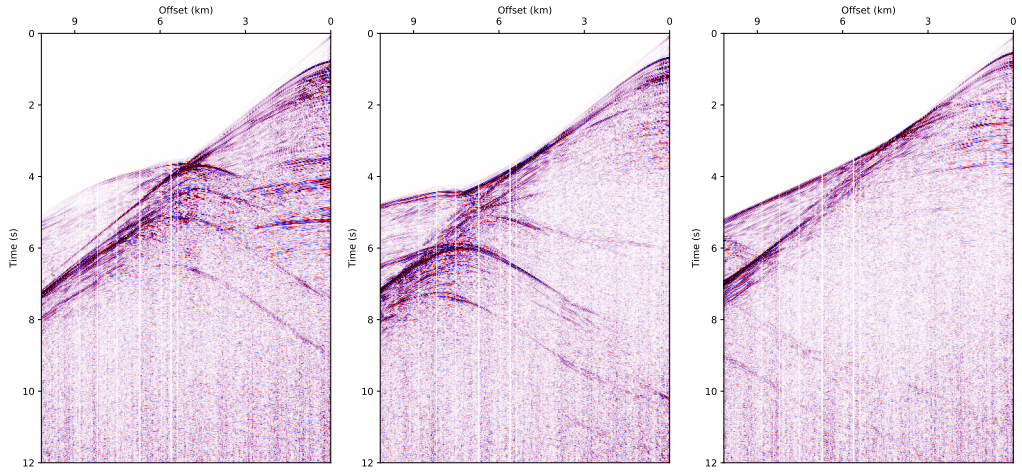
FWI, which requires many iterations of model updates, can benefit from the application of GMsFEM to enhance its efficiency by selecting the number of basis functions appropriately. We introduced the first example of applying the GMsFEM wave modeling engine to the frequency



(a)



(b)



(c)

Figure 4.22: Wave modeling results from the (a) initial velocity model (Figure 4.12(a)) and (b) inverted velocity model (Figure 4.19(a)) with corresponding (c) observed data.

domain FWI, and demonstrated the effect of wave modeling in a coarse grid through an analysis of data matrices and sensitivity kernels. We demonstrated the multiscale FWI using the Marmousi-2 model, and the FWI results show how varying the number of basis functions can control the tradeoff between the accuracy and computational speed. In addition, we showed FWI examples of the field data which is obtained in Gulf of Mexico region. In the field data examples, we demonstrated that the tradeoff between the accuracy and the speedup can be tuned by analyzing the forward, backward wavefield, and corresponding FWI results. As applying the proposed multsicale FWI with a relatively small number of basis functions can quickly construct a macro velocity model using low frequency, we proposed a strategy to optimize the FWI procedure by utilizing frequency-adaptive multiscale basis functions based on the target frequency group. Through this approach, we use few basis functions at low frequency group from which we could achieve a greater amount of speedup than in the high frequency groups. Thus, we could obtain the final velocity model which can be used for a background model for subsequent depth migration. We also validated the final multsicale FWI by investigating the CIGs of the Kirchhoff depth migration. The multsicale FWI could provide the results 78% and 8% faster than the reference case at the lower and highest frequency group, respectively. In overall, we could reduce the runtime up to 30% through the application of the GMsFEM wave modeling to the FWI.

5. SUMMARY AND CONCLUSIONS

5.1 Conclusions

In the preceding sections, I investigated the application of the multiscale wave modeling method using GMsFEM to seismic depth imaging and full-waveform inversion in the frequency-domain.

I first proposed a method to efficiently simulate seismic waves in the frequency domain by reducing the number of unknowns, and this enhancement on the computational efficiency directly contribute to the acceleration of the depth seismic imaging via RTM. In the first application of the GMsFEM to the RTM, tuning the number of multiscale basis functions plays an important role to manipulate the trade-off between the accuracy and the speedup for the best performance of the multiscale RTM. In the proposed method, I applied the multiscale mesh consisting of fine- and coarse-scale meshes. In the multiscale method, I first built the basis functions which incorporate fine-scale heterogeneity in the offline stage, then applied these multiscale basis functions to simulate the acoustic waves in the coarse-scale mesh. By doing so, I was able to achieve significant computational speedup when solving the inversion of the large impedance matrix. Although the multiscale RTM with larger coarse-grid with few basis functions could not resolve all the detailed subsurface structure, it could provide the RTM images showing macro structures more rapidly than the reference case, which uses the CG FEM for the wave modeling engine. To put it differently, the proposed RTM with a small number of basis functions helps to accelerate updating of velocity model by computing a large-scale structural image within a short amount of time. Applying more basis functions enabled the multiscale RTM to resolve the fine-scale strata without sacrificing accuracy. In addition, the proposed multiscale RTM algorithm still obtained images more rapidly than when performing the computations on the original fine-scale grid.

After demonstrating the RTM using the multiscale wave modeling engine, I proposed a method which utilizes multiple coarse-meshes and corresponding projection matrices to improve the efficiency of the wave modeling in the frequency-domain. The key strategy of the proposed method

was to apply different modeling parameters such as coarse-grid size and the number of basis functions according to the desired frequency bands. The GMsFEM algorithm combined with flexible multiscale grids efficiently simulated waves in the frequency domain. Also, I showed that the GMsFEM using the newer local spectral problem including the partition of unity term provides the multiscale basis functions which can enhance the accuracy of the solutions. I then presented the relationship between the size of the coarse cell and the number of basis functions in GMsFEM to maximize the speedup. The results show that the coarser mesh required more basis functions to capture the influence of fine-scale heterogeneity, but one can use a smaller number of basis functions as the coarse mesh becomes finer. I also applied a different sizes of coarse mesh and the corresponding number of multiscale basis to enhance the computational efficiency of the wave modeling in the RTM. When the multiscale RTM results were compared with a fixed mesh and a flexible mesh, the RTM result with the frequency-adaptive meshes produces better images than the other one.

One of the advantages of using frequency-domain RTM is that it can be converted to the FWI with ease by replacing the back-propagation source from the observed data to residual vector. Since the FWI requires multiple iterations of model updates, it is expected that we can save computational cost by considering the GMsFEM wave modeling. In this dissertation, I introduced the first example of applying the GMsFEM wave modeling engine to the frequency domain FWI, and demonstrated the effect of wave modeling in a coarse grid through an analysis of data matrices and sensitivity kernels.

To demonstrate the robustness and efficiency of the multiscale FWI, I presented two different examples: one synthetic example using Marmousi-2 model and one field dataset which is obtained from the Gulf of Mexico area. From the synthetic examples, the FWI results show how varying the number of basis functions can control the tradeoff between the accuracy and computational speed. In the field data examples, I demonstrated that the tradeoff between the accuracy and the speedup can be tuned by analyzing the forward, backward wavefield, and corresponding FWI results. As applying the proposed multiscale FWI with a relatively small number of basis functions

can quickly construct a macro-velocity model using low frequency, I proposed a strategy to optimize the multiscale FWI procedure by utilizing frequency-adaptive multiscale basis functions based on the target frequency group. Through this approach, I used few basis functions at the low-frequency group from which we could achieve a greater amount of speedup than in the high-frequency groups. Thus, I was able to obtain the final velocity model which can be used for a background model for subsequent depth migration.

Given all the numerical examples which are presented in this paper, it can be concluded that the GMsFEM can reduce the computational burden of the wave-equation based seismic imaging techniques RTM and FWI by dramatically reducing the dimension of the impedance matrix. Furthermore, in the multiscale approach, it is flexible in terms of applying the modeling parameters according to the desired frequency components that one needs to simulate, which makes the method an attractive tool for other practical applications.

5.2 Further Study

There are a number of seismic problems which could take benefit from the multiscale method by reducing the computational burden.

First, applying the GMsFEM wave modeling engine to the imaging tool incorporating multiple parameters such as elastic, poro-elastic, or anisotropy waves modeling could provide significant enhancement on the computational efficiency. Also, for more accurate simulation of the phenomena in the real world, considering more modeling parameters (i.e., attenuation, viscosity, or fluid-solid interaction) would be helpful to enhance the result of seismic imaging and waveform inversion.

Second, the multiscale approach will be more attractive when it applies an unstructured mesh. When we only deal with the structured mesh as shown in this paper, the advantage of FEM could not exceed that of FDM due to its speed and easy implementation. For example, by employing the unstructured triangular mesh, we can include the complicated topography to the wave simulation, which may be critical when we handle the elastic wave problem. In addition, utilizing unstructured mesh enable us to delineate complex geological structures (i.e., the edge of salt, shale diapir, or karst), which might create a large impedance contrast on the seismic response. Therefore, defining

an accurate geological interface (or structural boundary) can directly contribute to the accurate simulation of the seismic waves.

Finally, expanding the current 2-dimensional RTM and FWI code to the 3-dimensional problem can provide practical perspectives for defining the subsurface structure. Even in the 3-dimensional case, we can still solve the same local spectral problem to construct the multiscale basis functions in a more straightforward manner based on the theory which is presented in this dissertation. Nevertheless, in 3-dimensional problems, especially in the frequency-domain, the size of global impedance matrix would be significantly larger than 2-dimension case. Therefore, one might need to consider a memory efficient sparse-matrix solver such as Multifrontal massively parallel sparse direct solver (MUMPS) (Amestory et al., 2000; Amestoy et al., 2001, 2003).

REFERENCES

Amestory, P., I. Duff, and J.-Y. L'Excellent, 2000, Multifrontal parallel distributed symmetric and unsymmetric solvers: Computer Methods in Applied Mechanics and Engineerting, **184**, 501–520.

Amestoy, P. R., I. S. Duff, J.-Y. L. Excellent, and J. Koster, 2001, A fully asynchronous multifrontal solver using distributed dynamic scheduling: SIAM Journal on Matrix Analysis and Applications, **23**, 15–41.

Amestoy, P. R., I. S. Duff, J.-Y. L. Excellent, and X. S. Li, 2003, Impact of the implementation of mpi point-to-point communications on the performance of two general sparse solvers: Parallel Computing, **29**, 833–849.

Arbogast, T., 2003, An overview of subgrid upscaling for elliptic problems in mixed form: Contemporary Mathematics, **329**, 21–32.

Arbogast, T., S. L. Bryant, et al., 2002, A two-scale numerical subgrid technique for waterflood simulations: SPE Journal, **7**, 446–457.

Artemyev, M., H. Baek, and R. L. Gibson, 2015, Fast generalized multiscale FEM for complex media: Effortless modeling of topography and heterogeneity: 2015 SEG Annual Meeting.

Babuška, I., F. Ihlenburg, E. T. Paik, and S. A. Sauter, 1995, A generalized finite element method for solving the Helmholtz equation in two dimensions with minimal pollution : Computer Methods in Applied Mechanics and Engineering, **128**, 325–359.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse-time migration: Geophysics, **48**, 1514–1524.

Berenger, J.-P., 1994, A perfectly matched layer for the absorption of electromagnetic waves: Journal of Computational Physics, **114**, 185–200.

Billette, F. J., and S. Brandsberg-Dahl, 2004, The 2004 BP Velocity Benchmark: Presented at the 67th EAGE Conference & Exhibition, EAGE.

Biondi, B., and G. Shan, 2002, Prestack imaging of overturned reflections by reverse time migra-

tion: 82nd SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 1284–1287.

Biondi, B., and W. W. Symes, 2004, Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging: Geophysics, **69**, 1283–1298.

Brigham, E. O., 1988, The fast fourier transform and its applications: Presented at the , Prentice-Hall.

Brossier, R., J. Virieux, and S. Operto, 2008, Parsimonious finite-volume frequency-domain method for 2-dp–sv-wave modelling: Geophysical Journal International, **175**, 541–559.

Chavent, G., and R.-E. Plessix, 1999, An optimal true-amplitude least-squares prestack depth-migration operator: Geophysics, **64**, 508–515.

Chen, J.-B., 2012, An average-derivative optimal scheme for frequency-domain scalar wave equationan average-derivative method: Geophysics, **77**, T201–T210.

Chiu, S., E. Kanasewich, and S. Phadke, 1986, Three-dimensional determination of structure and velocity by seismic tomography: Geophysics, **51**, 1559–1571.

Cho, Y., and R. L. Gibson, 2018, Reverse-time migration via frequency-adaptive multiscale spatial grids: SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 4493–4497.

Cho, Y., R. L. Gibson, and S. Fu, 2017a, Frequency-domain reverse-time migration using generalized multiscale forward modeling: SEG Technical Program Expanded Abstracts 2017, Society of Exploration Geophysicists, 4583–4588.

———, 2018, A model reduction approach for full-waveform inversion via generalized multiscale finite elements: SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 1113–1117.

Cho, Y., R. L. Gibson, M. Vasilyeva, and Y. Efendiev, 2017b, Generalized multiscale finite elements for simulation of elastic-wave propagation in fractured media: Geophysics, **83**, WA9–WA20.

Cho, Y., W. Ha, Y. Kim, C. Shin, S. Singh, and E. Park, 2016, Laplace-fourier domain full-wave

form inversion of deep-sea seismic data acquired with limited offsets: Pure and Applied Geophysics, **173**, 749–773.

Chung, E. T., Y. Efendiev, and R. L. Gibson, 2011, An energy-conserving discontinuous multiscale finite element method for the wave equation in heterogeneous media: Advances in Adaptive Data, **03**, 251–268.

Chung, E. T., Y. Efendiev, R. L. Gibson, and M. Vasilyeva, 2016, A generalized multiscale finite element method for elastic wave propagation in fractured media: GEM-International Journal on Geomathematics, 1–20.

Chung, E. T., Y. Efendiev, and W. T. Leung, 2014, Generalized multiscale finite element methods for wave propagation in heterogeneous media: Multiscale Modeling & Simulation, **12**, 1691–1721.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 341–362.

Claerbout, J. F., and S. M. Doherty, 1972, Downward continuation of moveout-corrected seismograms: Geophysics, **37**, 741–768.

Clayton, R., and B. Engquist, 1977, Absorbing boundary conditions for acoustic and elastic wave equations: Bulletin of the Seismological Society of America, **67**, 1529–1540.

Collino, F., and C. Tsogka, 2001, Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media: Geophysics, **66**, 294–307.

Danecek, P., and G. Seriani, 2008, An efficient parallel chebyshev pseudo-spectral method for large scale 3d seismic forward modelling: Presented at the 70th EAGE Conference and Exhibition incorporating SPE EUROPEC 2008.

Dodson, D., R. Grimes, and J. Lewis, 1991, Sparse extensions to the Fortran basic linear algebra subprograms: ACM Transactions on Mathematical Software, **17**, 253–263.

Dongarra, J., J. D. Croz, S. Hammarling, and R. Hanson, 1988, An extended set of Fortran basic linear algebra subprograms: ACM Transactions on Mathematical Software, **14**, 1–32.

Efendiev, Y., J. Galvis, and X.-H. Wu, 2011, Multiscale finite element methods for high-contrast problems using local spectral basis functions: Journal of Computational Physics, **230**, 937–955.

Foltinek, D., D. Eaton, J. M. andP Moghaddam, and R. McGarry, 2009, Industrial-scale reverse time migration on GPU hardware: 2009 SEG Annual Meeting, 1–5.

Fomel, S., L. Ying, and X. Song, 2013, Seismic wave extrapolation using low-rank symbol approximation: Geophysical Prospecting, **61**, 526–536.

Fu, S., and K. Gao, 2017, A fast solver for the helmholtz equation based on the generalized multiscale finite-element method: Geophysical Journal International, **211**, 819–835.

Gao, K., S. Fu, and E. T. Chung, 2018, A high-order multiscale finite-element method for time-domain acoustic-wave modeling: Journal of Computational Physics, **360**, 120–136.

Gao, K., S. Fu, R. L. Gibson, E. T. Chung, and Y. Efendiev, 2015, Generalized Multiscale Finite-Element Method (GMsFEM) for elastic wave propagation in heterogeneous, anisotropic media: Journal of Computational Physics, **295**, 161–188.

Gardner, G., L. Gardner, and A. Gregory, 1974, Formation velocity and density: The diagnostic basics for stratigraphic traps: Geophysics, **39**, 770–780.

Gazdag, J., 1978, Wave equation migration with the phase-shift method: Geophysics, **43**, 1342–1351.

Ghysels, P., X. S. Li, F.-H. Rouet, S. Williams, and A. Napov, 2016, An efficient multicore implementation of a novel hss-structured multifrontal solver using randomized sampling: SIAM Journal on Scientific Computing, **38**, S358–S384.

Gibson, R. L., and S. Fu, 2015, Reverse time migration based on generalized multiscale finite element forward modeling: SEG Technical Program Expanded Abstracts 2015, Society of Exploration Geophysicists, 4137–4142.

Gibson, R. L., K. Gao, E. Chung, and Y. Efendiev, 2014, Multiscale modeling of acoustic wave propagation in 2D media: Geophysics, **79**, T61–T75.

Gilbert, F., and A. M. Dziewonski, 1975, An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra: Phil. Trans. R. Soc. Lond. A, **278**, 187–269.

Ha, W., Y. Cha, and C. Shin, 2010, A comparison between Laplace domain and frequency domain

methods for inverting seismic waveforms: Exploration Geophysics, **41**, 189–197.

Higdon, R., 1986, Absorbing boundary conditions for difference approximations to the multi-dimensional wave equation: Mathematic of Computation, **47**, 437–459.

Hou, T. Y., and X.-H. Wu, 1997, A multiscale finite element method for elliptic problems in composite materials and porous media: Journal of Computational Physics, **134**, 169–189.

Hustedt, B., S. Operto, and J. Virieux, 2004, Mixed-grid and staggered-grid finite-difference methods for frequency-domain acoustic wave modelling: Geophysical Journal International, **157**, 1269–1296.

Jo, C.-H., C. Shin, and J. H. Suh, 1996, An optimal 9-point, finite-difference, frequency-space, 2-d scalar wave extrapolator: Geophysics, **61**, 529–537.

Komatitsch, D., and J. Tromp, 1999, Introduction to the spectral element method for three-dimensional seismic wave propagation: Geophysical journal international, **139**, 806–822.

Komatitsch, D., and J.-P. Vilotte, 1998, The spectral element method: an efficient tool to simulate the seismic response of 2d and 3d geological structures: Bulletin of the seismological society of America, **88**, 368–392.

Koo, N.-H., C. Shin, D.-J. Min, K.-P. Park, and H.-Y. Lee, 2011, Source estimation and direct wave reconstruction in Laplace-domain waveform inversion for deep-sea seismic data: Geophysical Journal International, **187**, 861–870.

Kwon, J., I. Miller-Evans, and C. Shin, 2017, Efficient laplace constant selection strategy for the laplace-domain waveform inversion, *in* SEG Technical Program Expanded Abstracts 2017: Society of Exploration Geophysicists, 1617–1621.

Lailly, P., and J. Bednar, 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on inverse scattering: theory and application, Siam Philadelphia, PA, 206–220.

Lawson, C., R. Hanson, D. Kincaid, and F. Krough, 1979, Basic linear algebra subprograms for Fortran usage: ACM Transactions on Mathematical Software, **5**, 308–325.

Levander, A. R., 1988, Fourth-order finite-difference p-sv seismograms: Geophysics, **53**, 1425–

1436.

Lines, L., and S. Treitel, 1984, A review of least-squares inversion and its application to geophysical problems: Geophysical prospecting, **32**, 159–186.

Liu, H., B. Li, H. Liu, X. Tong, Q. Liu, X. Wang, and W. Liu, 2012, The issues of prestack reverse time migration and solutions with Graphic Processing Unit implementation: Geophysical Prospecting, **60**, 906–918.

Liu, Y., and M. K. Sen, 2012, A hybrid absorbing boundary condition for elastic staggered-grid modeling: Geophysical Prospecting, **60**, 1114–1132.

Marfurt, K. J., 1984, Accuracy of finite–difference and finite–element modeling of the scalar and elastic wave equations: Geophysics, **49**, 533–549.

Martin, G. S., R. Wiley, and K. J. Marfurt, 2006, Marmousi2: An elastic upgrade for marmousi: The Leading Edge, **25**, 156–166.

Martin, R., and D. Komatitsch, 2009, An unsplit convolutional perfectly matched layer technique improved at grazing incidence for the viscoelastic wave equation: Geophysical Journal International, **179**, 333–344.

McMechan, G., 1983, Migration by extrapolation of time-dependent boundary values: Geophysical prospecting, **31**, 413–420.

Min, D.-J., C. Shin, R. G. Pratt, and H. S. Yoo, 2003, Weighted-averaging finite-element method for 2d elastic wave equations in the frequency domain: Bulletin of the Seismological Society of America, **93**, 904–921.

Murphy, G. E., and S. H. Gray, 1999, Manual seismic reflection tomography: Geophysics, **64**, 1546–1552.

Nunes, V., and S. E. Minkoff, 2014, Imaging via subgrid upscaling and reverse time migration: 84th SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 4008–4013.

Oldham, R. D., 1906, The constitution of the interior of the earth, as revealed by earthquakes: Quarterly Journal of the Geological Society, **62**, 456–475.

Operto, S., J. Virieux, P. Amestoy, J.-Y. L?Excellent, L. Giraud, and H. B. H. Ali, 2007, 3d finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study: Geophysics, **72**, SM195–SM211.

Pan, W., K. A. Innanen, and W. Liao, 2017, Accelerating hessian-free gauss-newton full-waveform inversion via l-bfgs preconditioned conjugate-gradient algorithm: Geophysics, **82**, R49–R64.

Pestana, R. C., and P. L. Stoffa, 2010, Time evolution of the wave equation using rapid expansion method: Geophysics, **75**, T121–T131.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503.

———, 2007, A helmholtz iterative solver for 3d seismic-imaging problems: Geophysics, **72**, SM185–SM194.

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.

Pratt, R. G., C. Shin, and G. J. Hick, 1998, Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Rogers, A., 1914, über erdebenwellen. viia beobachtungen an registrierungen von ferubeben in göttingen und folgerungen über die konstitution des erdkörpers: By b. gutenberg. nachrichten der k. gesellschaft der wissenschaften zu göttingen. mathematischphysikalische klasse. 1914. pp. 1-52.

Saad, Y., 2003, Iterative methods for sparse linear systems: siam, **82**.

Shin, C., and Y. Cha, 2009, Waveform inversion in the laplace-fourier domain: Geophysical Journal International, **177**, 1067–1079.

Shin, C., and Y. H. Cha, 2008, Waveform inversion in the laplace domain: Geophysical Journal International, **173**, 922–931.

Shin, C., S. Jang, and D.-J. Min, 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: Geophysical prospecting, **49**, 592–606.

Shin, C., N.-H. Koo, Y. H. Cha, and K.-P. Park, 2010, Sequentially ordered single-frequency 2-d

acoustic waveform inversion in the laplace-fourier domain: Geophysical Journal International, **181**, 935–950.

Shin, C., and D.-J. Min, 2006, Waveform inversion using a logarithmic wavefield: Geophysics, **71**, R31–R42.

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

Sirgue, L., and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248.

Štekl, I., and R. G. Pratt, 1998, Accurate viscoelastic modeling by frequency-domain finite differences using rotated operators: Geophysics, **63**, 1779–1794.

Stolt, R., 1978, Migration by fourier transform: Geophysics, **43**, 23–48.

Sun, J., S. Fomel, and L. Ying, 2015, Low-rank one-step wave extrapolation for reverse time migration: Geophysics, **81**, S39–S54.

Symes, W. W., 2007, Reverse time migration with optimal checkpointing: Geophysics, **72**, SM213–SM221.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Turkel, E., D. Gordon, R. Gordon, and S. Tsynkov, 2013, Compact 2d and 3d sixth order schemes for the helmholtz equation with variable wave number: Journal of Computational Physics, **232**, 272–287.

Vdovina, T., S. E. Minkoff, and S. M. Griffith, 2009, A two–scale solution algorithm for the elastic wave equation: SIAM Journal on Scientific Computing, **31**, 3356–3386.

Virieux, J., 1986, P-sv wave propagation in heterogeneous media: Velocity-stress finite-difference method: Geophysics, **51**, 889–901.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.

Wang, S., M. V. de Hoop, and J. Xia, 2011, On 3d modeling of seismic wave propagation via a

structured parallel multifrontal direct helmholtz solver: Geophysical Prospecting, **59**, 857–873.

Whitmore, N. D., 1983, Iterative depth migration by backward time propagation: 53rd SEG Technical Program Expanded Abstracts, Society of Exploration Geophysicists, 382–385.

Woodhouse, J. H., and A. M. Dziewonski, 1984, Mapping the upper mantle: Three-dimensional modeling of earth structure by inversion of seismic waveforms: Journal of Geophysical Research: Solid Earth, **89**, 5953–5986.

Wu, Z., and T. Alkhalifah, 2018, A highly accurate finite-difference method with minimum dispersion error for solving the helmholtz equation: Journal of Computational Physics, **365**, 350–361.

Xue, Z., J. Sun, S. Fomel, and T. Zhu, 2017, Accelerating full-waveform inversion with attenuation compensation: Geophysics, **83**, A13–A20.

Zhang, J., and M. N. Toksöz, 1998, Nonlinear refraction traveltime tomography: Geophysics, **63**, 1726–1737.

APPENDIX A

PML FOR THE MULTISCALE HELMHOLTZ SOLVER

For most of the wave propagation models, we assume an infinite (or half infinite with free surface) space. Thus, an absorbing boundary condition plays an important role in wave equation based seismic imaging by suppressing the outgoing wavefield at the boundary of computational domain $\partial\Omega$. There are several ways to implement the absorbing boundary numerically, such as the one-way wave equation method (Clayton and Engquist, 1977; Higdon, 1986), Hybrid absorbing boundary (Liu and Sen, 2012), and Perfectly Matched Layer (PML) (Berenger, 1994; Collino and Tsogka, 2001; Martin and Komatitsch, 2009). Effective absorbing boundaries need to suppress the outgoing wavefield without a large amount of memory consumption. In this regard, we employed PML due to its superior absorption effect and ease of implementation in the frequency-domain. The PML damping function $g(x)$ can be defined as:

$$g(x) = \left(1 + i\frac{d(x)}{\omega}\right)^{-1}, \tag{A.1}$$

where x denotes the horizontal and vertical direction of spatial variables. Given the domain size as $\Omega = (0, p)^2$, the damping parameters in the PML zone can be:

$$d(x) = \begin{cases} \frac{C}{\xi}\left(\frac{x-\xi}{\xi}\right)^2, & x \in [0, \xi], \\ \frac{C}{\xi}\left(\frac{x-1+\xi}{\xi}\right)^2, & x \in [p-\xi, p], \end{cases} \tag{A.2}$$

where C is a positive constant independent of the frequency ω. Collino and Tsogka (2001) proposed a way to define an appropriate value of the constant. The constant C can be written as a function of model as: $\log\left(\frac{1}{R}\right)\frac{3v_p}{2\xi}$, where ξ means the thickness of the PML zone and R is a small number (≈ 0.001) which is applied to stabilize the calculation. The thickness of PML zone (ξ) varies with the velocity model; however, it approximates to a wavelength which is similar to one

coarse cell size. We applied 10 elements (200 m) on the edge of the density and velocity models to pad the PML zone. The Helmholtz equation can be rewritten by including the damping factors as follows:

$$\left[-\frac{\omega^2}{\rho v^2 g_1 g_2}u - \frac{\partial}{\partial x_1}\left(\frac{g_1}{g_2}\frac{1}{\rho}\frac{\partial}{\partial x_1}\right) - \frac{\partial}{\partial x_2}\left(\frac{g_2}{g_1}\frac{1}{\rho}\frac{\partial}{\partial x_2}\right)\right]u = f \tag{A.3}$$

The above Helmholtz equation can be expressed in a weak formulation through finding $u_H \in V_0^H$ such that

$$\int_\Omega \left[-\frac{\omega^2}{\rho v^2 g_1 g_2}u_H w_H + -\frac{1}{\rho}\frac{g_1}{g_2}\frac{\partial u_H}{\partial x_1}\frac{\partial w_H}{\partial x_1}w_H - \frac{1}{\rho}\frac{g_2}{g_1}\frac{\partial u_H}{\partial x_2}\frac{\partial w_H}{\partial x_2}w_H\right] d\mathbf{x} = \int_\Omega f w_H d\mathbf{x}, \tag{A.4}$$

where $\forall w_h \in V_0^H$. The corresponding impedance $\mathbf{S}$ and source $\mathbf{f}$ matrices can be calculated with this equation, and we then incorporate them into the discrete system to acquire the frequency-domain wavefields without artifacts reflected from the model boundaries.

APPENDIX B

LAPLACE-DOMAIN FWI FOR BUILDING AN INITIAL MODEL

B.1 Logarithmic objective function & gradient

The key concept of the FWI involves reducing the residual between the observed data and the modeled data based on the objective function. The approach introduced in the main text of this paper used the L_2 objective function, which can be expressed as

$$E(\mathbf{m}) = \frac{1}{2}\sum_{i=1}^{n_s}\sum_{j=1}^{n_r} \| \left(u_{ij} - d_{ij}\right)^2 \|. \tag{B.1}$$

However, the sensitivity to noise and the nonlinearity of the wave equation hinders practical application of the FWI to the field data without having a good priori information. As an alternative, we can perform the FWI in the Laplace-domain, which has distinctive features in terms of the convergence rate to the global minimum through the damped wavefield. After applying the damping functions, the Green's function exhibits singularity around the source location. Except for the area nearby the source position, the magnitude of the wavefields becomes too small to compute accurate residuals. Therefore, in the Laplace-domain FWI, we employ the logarithmic objective function (Shin and Min, 2006), which can be presented as

$$E(\mathbf{m}) = \frac{1}{2}\sum_{i=1}^{n_s}\sum_{j=1}^{n_r} \left(\ln\frac{|u_{ij}|}{|d_{ij}|}\right)^2. \tag{B.2}$$

Above logarithmic objective function at an arbitrary angular frequency ω can be written as

$$E(\mathbf{m}) = \frac{1}{2}\sum_{i=1}^{n_s}\sum_{j=1}^{n_r} \left(\ln\frac{|u_{ij}|}{|d_{ij}|}\right)\left(\ln\frac{|u_{ij}|}{|d_{ij}|}\right)^*, \tag{B.3}$$

where the asterisk $*$ means a complex conjugate, and the term $\ln\frac{|u_{ij}|}{|d_{ij}|}$ can be named as the residual $\mathbf{r}$. The gradient term to reduce the amount of residual can be obtained by applying the following

equation:

$$\frac{\partial E(\mathbf{m})}{\partial m_k} = \Re\left[\sum_{i=1}^{n_s}\sum_{j=1}^{n_r}\frac{1}{u_{ij}}\frac{\partial u_{ij}}{\partial m_k}\left(\ln\frac{|u_{ij}|}{|d_{ij}|}\right)^*\right]. \tag{B.4}$$

As presented in equation 3.19, the general form of the gradient involving the virtual source $\mathbf{v}$ can be written as

$$\frac{\partial E(\mathbf{m})}{\partial m_k} = \int_0^{\omega_{\max}}\sum_{i=1}^{n_s}\Re\left[\mathbf{v}^{\mathrm{T}}\mathbf{S}_h^{-1}\mathbf{r}_i^*(\omega)\right]d\omega, \tag{B.5}$$

where the residual $\mathbf{r}$ with logarithmic objective function can be expressed as

$$\mathbf{r} = \left[\frac{1}{u_1}\ln\frac{u_1}{d_1} \quad \frac{1}{u_2}\ln\frac{u_2}{d_2} \quad \cdots \quad \frac{1}{u_{n_r}}\ln\frac{u_{n_r}}{d_{n_r}} \quad 0 \quad 0 \quad \cdots \quad 0\right]^{\mathrm{T}}. \tag{B.6}$$

Similarly, as shown in the previous case, we can obtain the final model update term by applying the pseudo-Hessian matrix (equation 4.25).

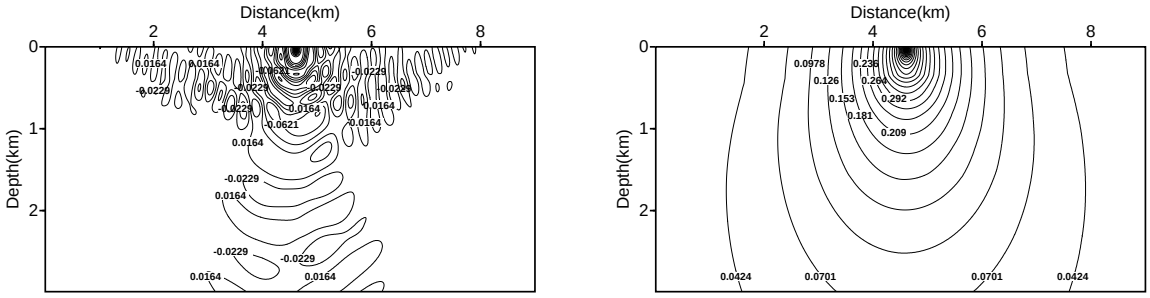
B.2 Laplace-domain vs. Frequency-domain wavefield

By employing the Laplace-domain for waveform inversion, we obtain several advantages: 1) avoiding convergence to local minimum, 2) reducing dependency on an initial model, 3) obtaining smoothed macro-velocity model that can be directly used as a migration velocity model. These advantages are attributed to the characteristics of Laplace-domain which uses a zero frequency component (Shin and Cha, 2008). Also, we can implement the Laplace-transform by using complex Fast-Fourier Transform (FFT). Equation B.7 proves the characteristic of Laplace-domain: zero-frequency component and relationship with frequency-domain.

$$\int f(t)e^{-i\omega^* t}\,dt = \int f(t)e^{-st}\,dt, \qquad \omega^* = r - si \tag{B.7}$$

where ω^* means complex angular frequency. In frequency-domain FWI, we use only real part of the angular frequency; however, when we work in Laplace-domain we replace the real part r to zero and only use the imaginary part s for calculating damped wavefield. This is the reason why Laplace-domain can be considered as a zero-frequency wavefield, and it helps FWI converge to

global minimum by generating extremely smooth wavefields (Figure B.1).



(a) Frequency-domain wavefield ($f = 5Hz$)

(b) Laplace-domain wavefield ($s = 1$)

Figure B.1: Comparison of wavefields in different domain. Wavefields in Laplace-domain shows extremely smoothed waves, and it contributes to converge to global minimum and build a long-wavelength velocity model.