Discretized Adjoint State Time and Frequency Domain Full Waveform Inversion: A Comparative Study

Mahjoobeh Meskaranian [1] and Peyman Pour Moghaddam [2*]

[1] PhD student, Department of Earth Sciences, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran
[2] Assistant Professor, Department of Geology, Faculty of Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

(Received: 13 May 2019, Accepted: 16 November 2019)

Abstract

This study derives the discretized adjoint states full waveform inversion (FWI) in both time and frequency domains based on the Lagrange multiplier method. To achieve this, we applied adjoint state inversion on the discretized wave equation in both time domain and frequency domain. Besides, in this article, we introduce reliability tests to show that the inversion is performing as it should be expected. Reliability tests comprise of objective function descent test and Jacobian test. The influence of data imperfections is also being studied. We define data imperfection as any factor that causes deterioration in FWI results. Some of these factors are coherent and incoherent noises in data, source wavelet inaccuracy in phase and amplitude, and the existence of gaps in the seismic survey. We compare time and frequency domain inversion methods sensitivity to data imperfection. In all cases, we found that time domain full waveform inversion is more sensitive to imperfections in the data. In general, we find that time domain FWI result shows more deterioration than frequency domain FWI. All tests have been done using 2D full waveform inversion codes. We employ the multi-scale inversion and finite difference scheme (FDM) for discretization, and the misfit function is minimized via limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) method.

Keywords: : FWI, Sensitivity Analysis, Discretized Adjoint State Method

*Corresponding author: ppm1407@hotmail.com

1 Introduction

Full waveform is a powerful method to obtain seismic velocity model using the whole wavefield. Foundation of FWI theory is a study published by Tarantola (1984), who demonstrated minimization of the misfit between recorded and modelled data might be computed without calculating the partial derivatives explicitly. In other words, Tarantola (1984) proposed applying an adjoint state method in the seismic inversion context. The adjoint state method is an idea in system control theory that, for the first time, was implemented in geophysical inverse theory by Chavent (1974).

FWI was originally developed in time domain (Tarantola, 1984; Tarantola, 1986; Mora, 1987; Tarantola, 1988), then Pratt and Worthington (1990) and Pratt (1990) introduced frequency domain FWI (FDFWI) approach and successfully applied to 2D cross-hole data.

To mitigate the nonlinearity of FWI and minimize the cycle-skipping effect (ref. Yao et al., 2019), various multiscale strategies have been proposed (e.g., Bunks et al., 1995; Sirgue and Pratt, 2004; and Boonyasiriwat et al., 2009). In time domain, multi-scale FWI method inverts for lowest-frequency sub-bands of seismic data, then fed the inverted model as initial model to higher frequency sub-band. Final result achieved when the result of highest sub-band is calculated. This helps to increase the chances of convergence to the global minimum (Bunks et al., 1995). Furthermore, by applying Fourier transform on the dataset, the frequency content of the waveform is discretized and the multi-scale method can be easily employed in FDFWI. The inversion process is done incrementally adding from low to high-frequency components (Sirgue and Pratt, 2004).

Application of FWI in each domain has some pros and cons. In two dimensions, solving the sparse matrix arising from FDFWI is not so expensive. Besides, various sparse solver approach has the ability to invert multiple shots simultaneously (Brenders and Pratt, 2007). However, in three dimensions, the matrix inversion is time- and memory-consuming. On the other hand, time domain FWI (TDFWI) has the advantage of applying time windowing on data for choosing selected arrivals in data as a function of time (Sheng et al., 2006). In addition, TDFWI is more suitable for implementation on parallel processing computers such as GPGPUs (general-purpose graphics processing unit) (Mao et al., 2012, Yang et al., 2015, Jiang and Zhu, 2018).

Although adjoint state method for inversion is well established in the geophysical community (Oldenburg, 1990, Plessix, 2006), there have been few studies (e.g., Fichtner 2010) to look at full waveform inversion in discrete level, contributing the discretization scheme in the adjoint state formulation. In this study, we are attempting to derive the discretized adjoint state gradient function for both TDFWI and FDFWI using $2K^{th}$ order finite difference scheme (FDM) for discretization. There are some fine details in the calculation of gradient in the discretized adjoint state method, which is not visible if we are only using classical adjoint states.

In this study, we also touch briefly at the important topic of reliability tests for FWI. Any nonlinear inversion scheme should have reliability tests, ensuring that the inversion is performing correctly. This can be done by performing some basic but crucial tests on the value of the calculated gradient and the objective function.

The last topic covered in this study is a comparison between TDFWI and FDFWI. Although there might be quite a few references on either TDFWI or FDFWI (Virieux and Operto, 2009), there has not been many studies on comparing

results from these two methods when there is the imperfection in the data such as random coherent and incoherent noises with different signal to noise ratios, source wavelet inaccuracy in phase and amplitude and existence of gap in the data. The aim of this study is to test and compare the data sensitivity of FWI in time and frequency domain. For this propose we impose the data imperfections such as coherent noise, source wavelet errors, gaps in the survey on synthetic data and present their impact on both TDFWI and FDFWI results. We decide not to add cycle skipping in the comparison study since there have been many studies of cycle skipping for both of these methods. Another reason which we do not add cycle-skipping study is that for a fair comparison, we need fairly good inversion results for noise-free inversions. This is obtained when we have the lowest frequencies in the data. In both FDFWI and TDFWI, we use multi-scale FWI method and the misfit function was minimized via limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) method. In FDFWI we use perfectly matched layer (PML) for absorbing boundary condition (Komatitsch and Tromp, 2003), although sponge is applied for TDFWI (Clayton and Engquist, 1977).

2 Adjoint state method

Inverse problems aim to find model parameters, which can minimize the difference between observed and simulated data. The majority of nonlinear inverse problems utilizes iterative methods to find the solution, such that in each iteration solution is updated for reducing the objective function (misfit function), and the process continues until the misfit reaches to a satisfying value.

$$m_{i+1} = m_i + \gamma_i H_i \ , \quad J(m_{i+1}) < J(m_i), \tag{1}$$

where m is the model parameter, J is the objective function, H_i and γ_i are descent direction and the step length respectively. Gradient methods approximate m by calculating the derivatives of the objective function with respect to the model parameters, which might be computationally expensive in seismic problems.

The adjoint method is an efficient technique to evaluate the derivative of a functional $[p(m)]$, with respect to m. This is a constraint inversion with constraints,

$$L(m)p = S, \tag{2}$$

where $L(m)$ is a linear operator that depends on m , p as wavefield and S is the source. The first derivation of $J[p(m)]$ with respect to m can be defined as:

$$\frac{\partial J}{\partial m} = \frac{\partial J}{\partial p}\frac{\partial p}{\partial m}, \tag{3}$$

The main difficulty of calculating $\frac{\partial J}{\partial m}$ is the computation of the $\frac{\partial P}{\partial m}$ that is often numerically expensive; however, the adjoint method proposes a way of eliminating Frechet derivatives that contains first derivatives by using state equations, $L(m)p = F$ that can be rewritten as:

$$F(p,m) = L(m)p - q = 0, \tag{4}$$

The adjoint state method can be derived in several different ways including, Born approximation (e.g. Tarantola, 1988; Fichtner, 2010), data assimilation (e.g., Chen, 2011) and Lagrange multipliers (e.g., Liu and Tromp, 2006). Lagrange multipliers is a method to find the maxima or minima of a function subject to equality constraints (i.e., minimize $J(m,p)$, subject to $F(p,m) = 0$). The method proposes that instead of considering $J[p(m)]$ as an objective

function, which should be minimized on m, we now consider new function $H(m,p(m),\theta)$ as a function of p, associated with constraint $L(m)p=0$ and Lagrange multiplier θ,

$$H(m,p(m),\theta)=J(p)-\theta(L(m)p-F), \quad (5)$$

$H(m,p,\theta)$ is a new function with higher dimensionally independent variables. Derivatives of H are as follows:

$$\frac{\partial H}{\partial p}=\frac{\partial J}{\partial p}-L^{*}(m)\theta, \quad (6)$$

$$\frac{\partial H}{\partial m}=\frac{\partial J(p)}{\partial m}-\theta\frac{\partial L}{\partial m}p, \quad (7)$$

Setting $\frac{\partial H}{\partial p}=0$, in Equation (6), the adjoint equation is derived.

$$L^{*}(m)\theta=\frac{\partial J}{\partial p} \quad (8)$$

Also $\frac{\partial H}{\partial p}=0$, in Equation (7) gives the gradient equation:

$$\frac{\partial J(p)}{\partial m}=\theta\frac{\partial L}{\partial m}p, \quad (9)$$

Therefore, if we determine p using state Equation (4) and θ using adjoint Equation (8), it is possible to compute the gradient of the objective function using Equation (9), (refer to Plessix, 2006).

2-1 Discretized adjoint state TDFWI

The adjoint state is a mathematically elegant form of calculating the gradient. However, when we use a discretization scheme for the solution of the wave equation, it should also be incorporated in the adjoint state derivation of the gradient. Here we follow the process described in the last section for calculating the gradient in a discrete sense.

Mathematically acoustic homogeneous wave propagation is defined by:

$$\frac{\partial^2 p(t)}{\partial^2 t}=c^2\,\nabla^2 p(t)+q(t) \quad (10)$$

where p, c and q are pressure field, velocity and source, respectively. The finite difference implementation of the wave equation reads:

$$p_{x,y}^{n+1}=2p_{x,y}^{n}-p_{x,y}^{n-1}+\Delta t^2\,c_{x,y}^2\,\nabla^2 p_{x,y}^{n}+q_{x_s,y_s}^{n}, \quad (11)$$

where n is the current time step, x_s, y_s are the source x, y location. $A(2k)^{th}$ order FDM discretization of ∇^2 is defined as:

$$\nabla^2 p_{x,y}^{n}=a_0\,p_{x,y}^{n}+\sum_{i=-k}^{k}a_i^{x}p_{x-i,\,y}^{n}+\sum_{i=-k}^{k}a_i^{y}p_{x,\,y-i}^{n}, \quad (12)$$

where a_i^{x} and a_i^{y}as FDM stencil coefficients. The forward seismic-wave Equation (10) can be rewritten as an equality constraint:

$$F(p_{x,y}^{n},c_{x,y})=p_{x,y}^{n+1}-2p_{x,y}^{n}+p_{x,y}^{n-1}-\Delta t^2\,c_{x,y}^2\,\nabla^2 p_{x,y}^{n}+q_{x_s,y_s}^{n}=0 \quad (13)$$

The full waveform inversion objective function can be written as:

$$J(c_{x,y})=\sum_{x_r,\,y_r}^{all\ n}(d_{x_r,y_r}^{n}-p_{x_r,\,y_r}^{n})^2, \quad (14)$$

where, d_{x_r,y_r}^{n} is data recorded at x_r, y_r location and time sample n. To minimize the objective function over $c_{x,y}$, we use the Lagrangian multiplier method, with the objective function subjected to the equality constraint (Equation (13)),

$$H(\theta_{x,y}^{n},p_{x,y}^{n},c_{x,y})=J(c_{x,y})+\sum_{x,y}^{all\ n}\theta_{x,y}^{n}\,F(p_{x,y}^{n},c_{x,y}), \quad (15)$$

where $\theta_{x,y}^{n}$ is the Lagrange multiplier associated with $F(p_{x,y}^{n},c_{x,y}^{n})$.

As it is mentioned before, state equation is a forward equation that calculates $p_{x,y}^{n}$, and the adjoint equation is derived by setting:

$$\frac{\partial H(\theta_{x,y}^{n},p_{x,y}^{n},c_{x,y})}{\partial p_{x,y}^{n}}=0 \quad (16)$$

$$\theta_{x,y}^{n-1}-2\theta_{x,y}^{n}+\theta_{x,y}^{n+1}-$$

$$\Delta t^2\ c_{x,y}^2\ \nabla^2\theta_{x,y}^n\ - \quad (17)$$
$$2\sum_{x_{r,}\,y_r}^{all\ n}(d_{x_{r,}\,y_r}^n - p_{x_{r,}\,y_r}^n) = 0$$

This is the residual back propagation with values stored in $\theta_{x,y}^n$. Setting $\frac{\partial H(\theta,p,c)}{\partial c} = 0$ will result in gradient equation,

$$\frac{\partial H(\theta_{x,y}^n,\ p_{x,y}^n, c_{x,y})}{\partial c_{x,y}} = 0\ , \quad (18)$$

$$\frac{\partial J(c_{x,y})}{\partial c_{x,y}} + \frac{\partial}{\partial c_{x,y}}\sum_{x_{r,}\,y_r}^{all\ n}\theta_{x,y}^n\ \ F(p_{x,y}^n, c_{x,y}) = 0, \quad (19)$$

$$\frac{\partial J(c_{x,y})}{\partial c_{x,y}} = -\frac{\partial}{\partial c_{x,y}}\sum_{x_{r,}\,y_r}^{all\ n}\theta_{x,y}^n\ \ F(p_{x,y}^n, c_{x,y}) = 2\,\Delta t^2\sum_{x,y}^{all\ n} c_{x,y}\ \theta_{x,y}^n\ \nabla^2 p_{x,y}^n\ . \quad (20)$$

Discretized adjoint state, obtained in Equation (20) shows that for calculation of gradient, we need to cross-correlate the residual back propagated wavefield with 'laplacian' of the forward wavefield. Additional term of $c_{x,y}$ in above equation also needs to be multiplied by cross-correlation.

2-2 Discretized adjoint state FDFWI

The acoustic wave equation mapped to frequency domain can be expressed by:

$$\frac{\omega^2}{c^2}\ p(\omega) + \nabla^2 p(\omega) = -q(\omega), \quad (21)$$

where $p(\omega)$ is pressure field, $q(\omega)$ is the source signature in frequency domain and ω is angular frequency.

For a single frequency $\omega_n = n\Delta\omega$ the discretized state equation using FDM can be expressed by:

$$\omega_n^2\ p_{x,y}^n\ +\ c_{x,y}^2\ \nabla^2 p_{x,y}^n\ +\ q_{x_s,,x_s}^n = 0, \quad (22)$$

where $p_{x,y}^n$ is the wavefield at location $x\Delta x, y\Delta y$ and at an angular frequency of ω_n.

The equality constraint for state equation can be expressed by:

$$F(p_{x,y}^n, c_{x,y}) = \omega_n^2\ p_{x,y}^n\ + c_{x,y}^2\ (A) + q_{x_{s,,}\ x_s}^n = 0. \quad (23)$$

A=$a_0 p_{x,y}^n\ +\ \sum_{i=-k}^{k}\ a_i^x p_{x-i,\ y}^n\ +\ \sum_{i=-k}^{k}\ a_i^y\ p_{x,\ y-i}^n$

The full waveform inversion objective function in frequency domain reads:

$$\min_c J(c_{x,y}) = \sum_{x_{r,}\,y_r}^{n}(d_{x_r,\,y_r}^n - p_{x_r,\,y_r}^n)^2, \quad (24)$$

Similar to the last section, the higher dimensionality objective function H using Equations (23) and (24) using Lagrange multipliers $\theta_{x,y}^n$can be expressed as:

$$H(\theta_{x,y}^n,\ p_{x,y}^n, c_{x,y}) = J(c_{x,y}) + \sum_{x,y}^{all\ n}\theta_{x,y}^n\ \ F(p_{x,y}^n, c_{x,y}), \quad (25)$$

At minimum derivative of the $H(\theta_{x,y}^n,\ p_{x,y}^n, c_{x,y})$ with respect to $p_{x,y}^n$and $c_{x,y}$ should be zero. Adjoint equation is derived by setting $\frac{\partial H(\theta,p,c)}{\partial p} = 0$,

$$\frac{\partial H}{\partial p_{x,y}^n} = -\sum_{x_{r,}\,y_r}^{n} 2(d_{x_{r,}\,y_r}^n - p_{x_{r,}\,y_r}^n)^* + \omega_n^2\,\theta_{x_{r,}\,y_r}^n\ + c_{x,y}^2\,\nabla^2\theta_{x_{r,}\,y_r}^n\ = 0, \quad (26)$$

$$\omega_n^2\,\theta_{x_{r,}\,y_r}^n\ + c_{x,y}^2\,\nabla^2\theta_{x_{r,}\,y_r}^n\ = \sum_{x_{r,}\,y_r}^{n} 2(d_{x_{r,}\,y_r}^n - p_{x_{r,}\,y_r}^n)^*. \quad (27)$$

Setting $\frac{\partial H(\theta,p,c)}{\partial p} = 0$ results in gradient equation,

$$\frac{\partial H}{\partial c_{x,y}} = \frac{\partial J}{\partial c_{x,y}} + \frac{\partial}{\partial c_{x,y}}\left(\sum_{x,y}^{all\ n}\theta_{x,y}^n\ \ F(p_{x,y}^n, c_{x,y})\right) = 0, \quad (28)$$

$$\frac{\partial J}{\partial c_{x,y}} = -2\sum_{x,y}^{all\ n}\theta_{x,y}^n\ \ c_{x,y}\ \nabla^2 p_{x,y}^n\ . \quad (29)$$

2-3 Inversion Method

We applied multiscale FWI with the misfit function minimized via limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) method. The velocity model update in each iteration derived as:

$$c_{x,y}^{k+1} = c_{x,y}^{k} - \tau\, H^{k}\, \nabla J\left(c_{x,y}^{k}\right), \tag{30}$$

where H^k is the inverse of the Pseudo-Hessian obtained with LBFGS. The value of line search (τ) is obtained from Equation (31).

$$min_{\tau}\ \frac{1}{2}\|\boldsymbol{d} - R\boldsymbol{p}(c + \tau\,\boldsymbol{\delta c})\|_2 = \frac{1}{2}\|\boldsymbol{d} - R\boldsymbol{p}(c) - \tau\, R\boldsymbol{\delta p}\|_2\ , \tag{31}$$

where R is detection operator which is zero everywhere except at the receivers location and $\delta p = p(c + \delta c) - p(c)$ for small perturbation of δc along the descent direction.

Recalling, $a = d - Rp(c)$ and $b = R\delta p$.

$$\tau = \frac{a^T b}{b^T b}.$$

2-4 Absorbing Boundary

One serious issue in seismic forward modelling is reflections or wrap around waves from the artificial boundaries of the numerical grid. To overcome this effect, different efforts have been undertaken, these methods can be classified in two major groups: absorbing boundary conditions (e.g. Engquist and Majda, 1977, Clayton and Engquist, 1977, Reynolds, 1978, Keys, 1985, Higdon, 1991) and absorbing boundary layers (e.g. Cerjan et al., 1985; Berenger, 1994; Komatitsch and Tromp, 2003; McGarry and Moghaddam; 2009, Pasalic and McGarry, 2010). We use the method of perfectly matched layers (PML) of Bereger (1994) for frequency domain modelling and Clayton and Engquist (1977) paraxial conditions for time domain modelling. Although, in theory, the discretized adjoint state of the absorbing layers needs to be calculated for estimation of the gradient to be highly accurate; however, this calculation is quite complicated and we presume the absorbing layers are self-adjoint.

3 Reliability Test

In order to investigate that time and frequency domains FWI are accurately performed, we need to do some reliability tests. To insure that the inversion is performing correctly, we should make some reliability tests on nonlinear inversion scheme. This can be done by performing some basic but crucial tests on the value of the calculated gradient and the objective function.

3-1 Objective function reduction

As it was previously mentioned, in both time and frequency domains, we applied multi-scale FWI on N frequency sub-bands, each sub-band M iteration, objective function decreasing is one of the reliability tests to evaluate inversion. In each sub-band, the value of objective function per iteration needs to be decreased. A common misconception is that the value of objective function needs to strictly decrease per iteration and when the inversion moves to a new sub-brand, the inversion should start with a smaller value of the objective function. However this is not true, the value of objective function at iteration 1 in each sub-band is totally unrelated to the value of the objective function at previous sub-bands. This is because of the objective function for each sub-band is formed only with the data information in that specific sub-band. However, objective function must be decreased per iteration in this sub-band. In the example section, we show some of the objective function behaviors.

3-2 Jacobian test

The other reliability test that is very important is the Jacobian test. The test is performed on Jacobian of the objective

function of FWI to measure its accuracy. If Jacobian of the objective function is properly implemented, then the following equation at iteration k of the inversion should be held.

$$J(c_{k+1}) - J(c_k) = \frac{1}{2}\left(\nabla J(c_{k+1}) + \nabla J(c_k)\right)^T (c_{k+1} - c_k) \quad (32)$$

In each iteration, we define two functions g_a^k and g_b^k as follows:

$$g_a^k = J(c_{k+1}) - J(c_k), \quad (33)$$

$$g_b^k = \frac{1}{2}\left(\nabla J(c_{k+1}) + \nabla J(c_k)\right)^T (c_{k+1} - c_k), \quad (34)$$

where, *J* refers to Jacobian and c_{k+1} and c_k are velocity models that are obtained from two consecutive iterations, *k + 1* and *k*, respectively. The two functions g_a^k and g_b^k should follow each other closely per iterations. If the two functions significantly deviate from each other, then there is a problem with Jacobian derivation. In example section, we show behavior of these functions.

4 Sensitivity Analysis

The aim of this study is to compare the data imperfections effect on FWI in time and frequency domain. If we compute the relative error based on the exact model when we insert imperfection in the data, it will be difficult to measure the deterioration of the result since TDFWI and FDFWI each gives different inversion result for the noise-free dataset. Instead, we measure the deterioration from the noise-free inverted model rather than the exact model. Here is the step-by-step process:

1. Generate synthetic noise-free data set from the exact model and run FWI in time and frequency domain on this dataset, then run synthetic data from the inverted model and call it 'reference' dataset.
2. Add imperfection to time and frequency domain noise free dataset, run inversion on imperfect data to find the inverted model, then run synthetic on the inverted model, and call it 'erroneous' dataset. The error can be calculated in percentage as follows:

$$E = \frac{\|R - S\|_1}{\|R\|_1} \times 100 \quad (35)$$

In Equation (35), R is reference synthetic dataset (either time or frequency domain) and S is erroneous synthetic dataset.

5 Examples

We generate inversion results by using 2D acoustic FWI on a subset of the North Sea model shown in Figure 1a). The geometry subset is 1600 m long and 2000 m deep with 8 m grid spacing. Initial velocity model was created as a 1D 137 velocity model by averaging over every row of the true velocity model that is shown in Figure 1b). Synthetic data is generated with 32 shots and 189 receivers. Shots are located at a lateral position between 48 m and 1550 m with a 48 m shot spacing and lateral positions of receivers are between 40 m and 1550 m and 8 m receiver spacing. Depth of all shots and receivers is 8 m. The source time function is a Ricker wavelet with a 15 Hz central frequency and total recording time and time sampling are 3 s and 0:001 s respectively. We conducted the inversion process using multi-scale FWI, the frequency band is from 0:33 Hz to 20 Hz, in 0:33 Hz intervals that decomposed into 5 sub-band, with 15 iterations at each sub-band. The 0:33 Hz frequency increment for frequency domain inversion corresponds to 3 seconds of recording time for TDFWI. The first to fifth sub-bands are [0-4] Hz, [4-8] Hz, [8-12] Hz, [12-16] Hz and [16-20] Hz respectively.

The objective functions for noise-free dataset in each band are shown in Figure 2 and 3 for TDFWI and FDFWI, respectively.

Figures 4 and 5 show results of the Jacobian test for TDFWI and FDFWI, respectively.

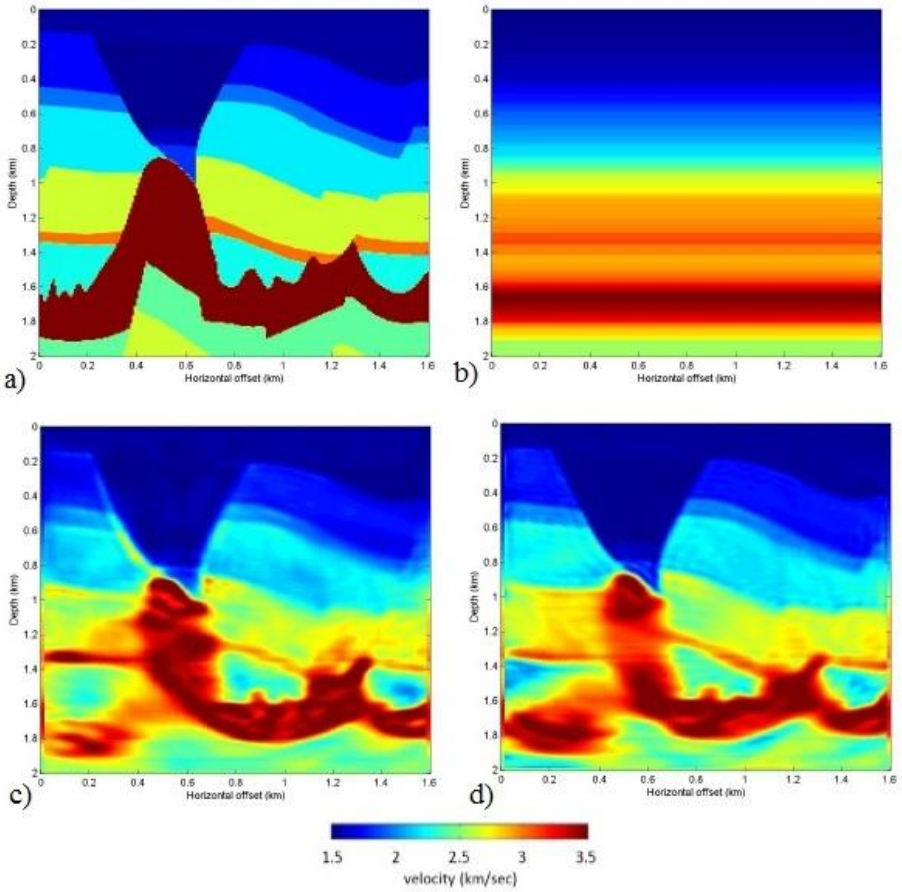

Figure 1: a) True velocity model, b) Initial velocity model, c) FDFWI result, d) TDFWI result.

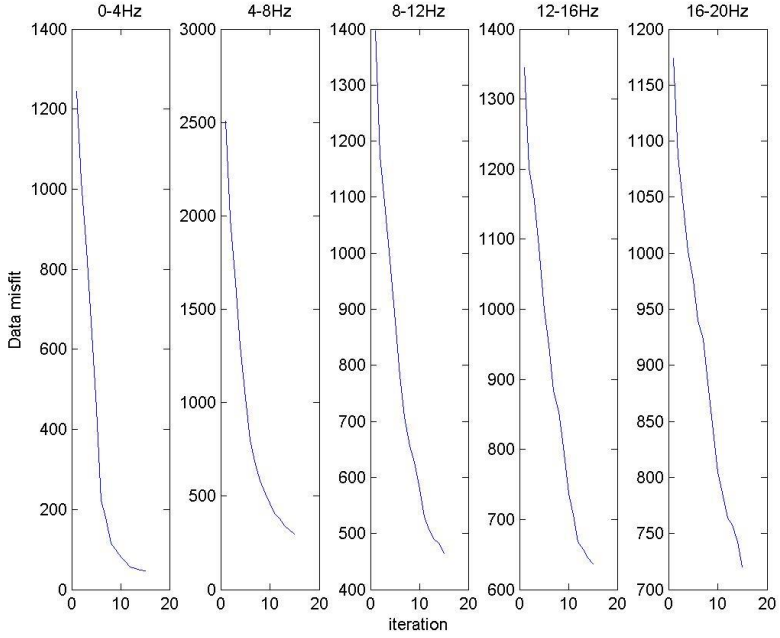

Figure 2: Objective function decreasing test for noise free TDFWI.

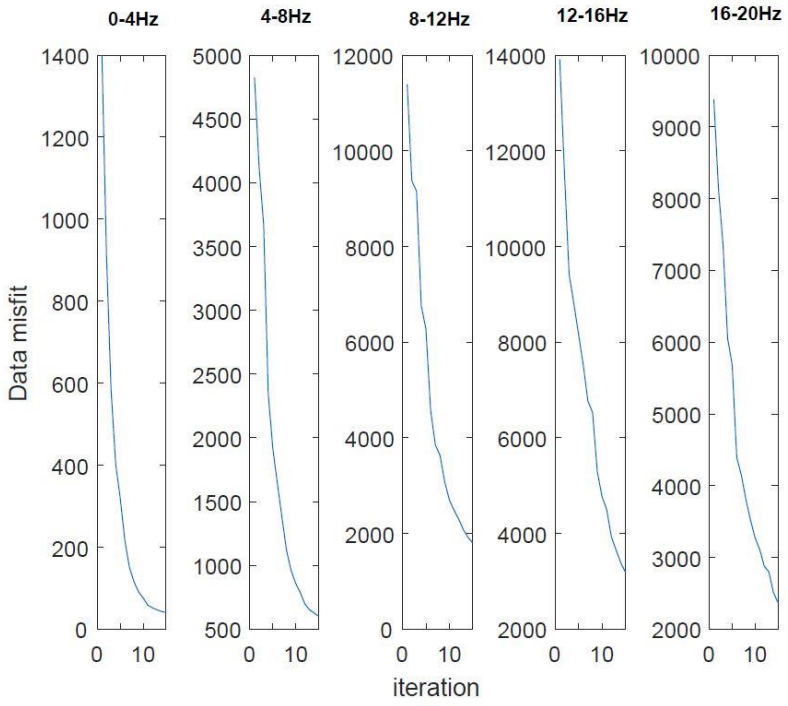

Figure 3. Objective function decreasing test for noise free FDFWI.

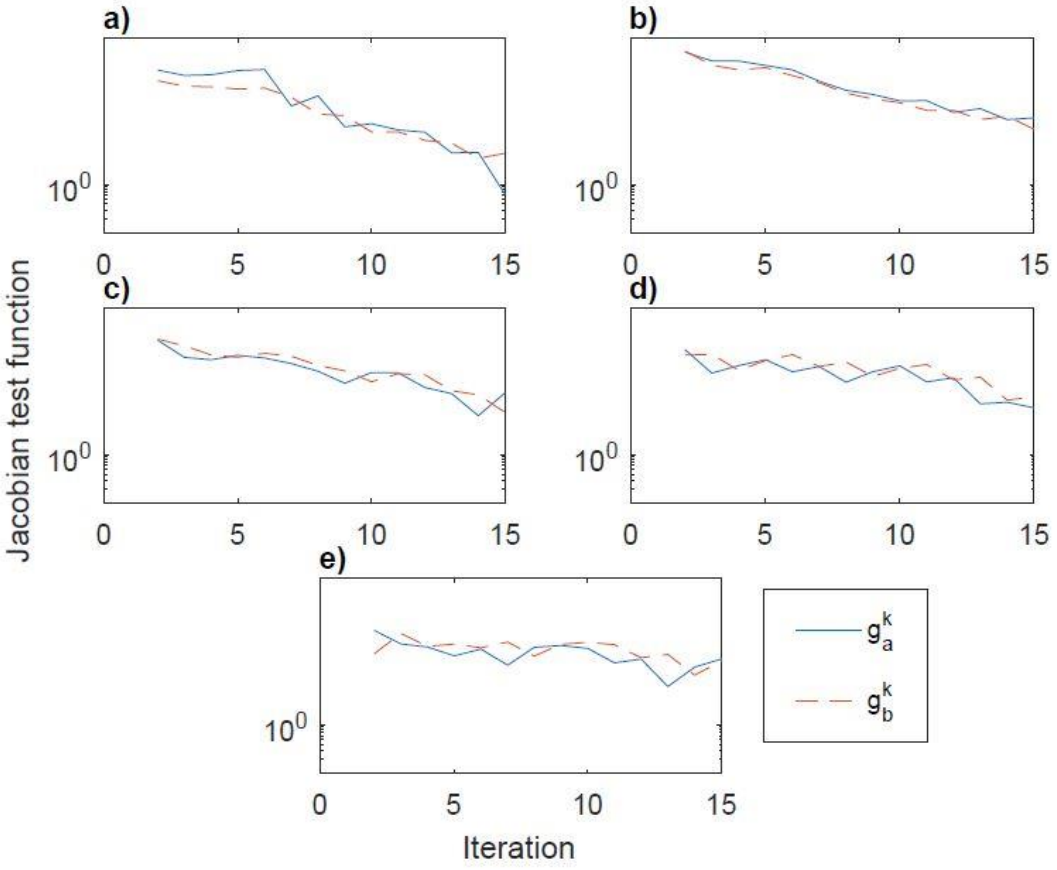

Figure 4. Jacobian test for noise free TDFWI are, (a) first sub-band, (b) second sub-band, (c) third sub-band, (d) fourth sub-band, (e) fifth sub-band.

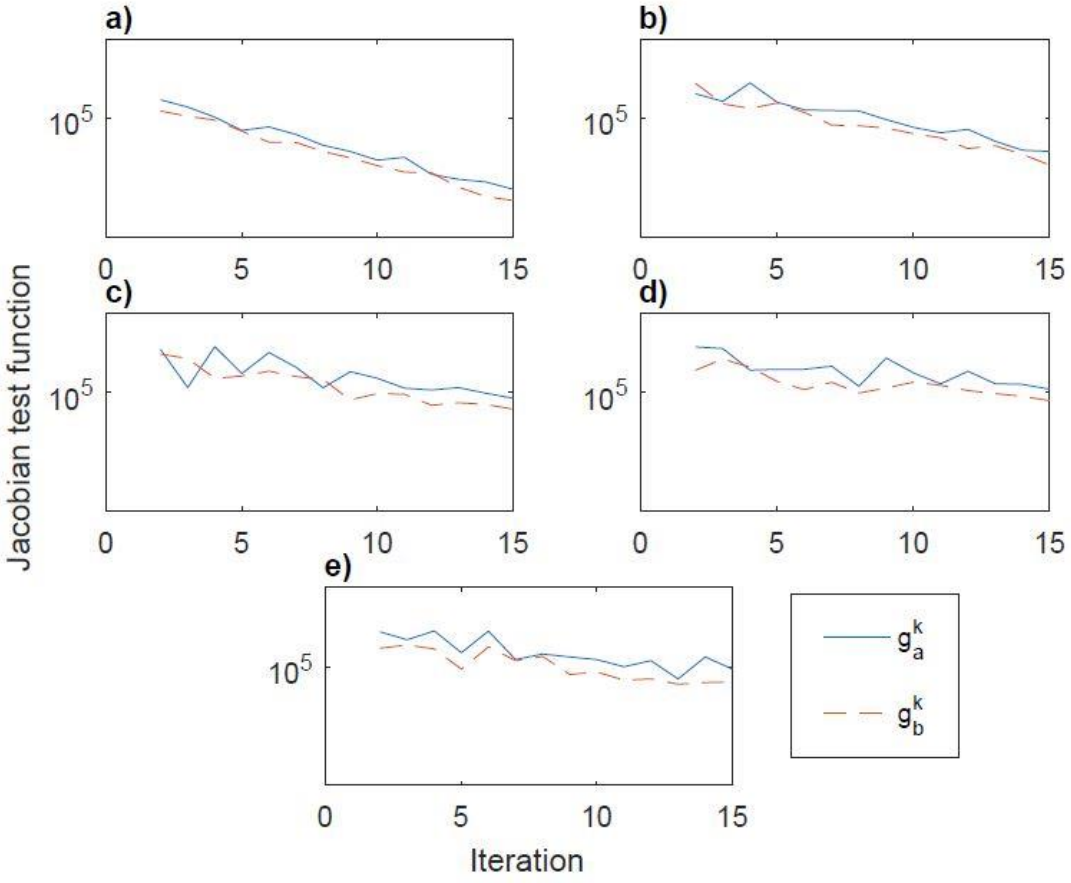

Figure 5. Jacobian test for noise free FDFWI are, (a) first sub-band, (b) second sub-band, (c) third sub-band, (d) fourth sub-band, (e) fifth sub-band.

4-1 Gap in survey

In practical FWI, there are several challenges for collecting data, due to natural phenomena and equipment limitations. We assume there are gaps in the survey (i.e., regions without any sources and receivers) and investigate 157 of their effects on inversion results. We impose two types of gap formations to create syntactic data. The first gap is a continuous gap at the center of the model and the second gap is two equally separated gaps with centers are located in one-third and two-thirds of the model. The length of the total gap is equal for both types of a gap, for example, a 100 m gap of the first type could be comparable with a 50-50 m gap of the second type. Figure 6 shows the receiver data for both types of gaps with 400 m width. Figure 7 shows FWI result of gap errors. As it can be seen in this figure, for FDFWI increasing continuous gap has a visible effect on salt body and its surrounding while the level of deterioration is not very visible in discrete gap. By looking at these figures, we can also say continuous gap for both FD and TD FWI looks to have more deteriorating effect that discrete gap with the same size.

4-2 Coherent noise

We cannot deny the presence of noise in seismic data, wind motion or cable vibrations can generate random noise similar to incoherent noise. Some of the noise sources Such as pumpjack noise, 60 Hz powerline noise, ground roll, reverberating refractions and multiples (Chopra and Marfurt, 2014) create more

coherent energy on the data and can be misinterpreted as a true signal. Therefore, the sensitivity of FWI to coherent and incoherent noise has been examined both in frequency and time domain. To create coherent noise, we first generated a random Gaussian noise, then apply a smoothing filter on it with σ, that is a control parameter defined as Equation (36).

$$\sigma = \frac{L}{10\,S_L}, \quad (36)$$

In Equation (36), L is the length of the model and SL is the smoothing distance. FWI results of noise errors are shown in Figure 8. As it can be seen in this figure, coherent noise when SNR is low has destructing effect on the results but the main salt body can somehow be seen in low SNR with FDFWI method while the salt body is deteriorated too much to be seen as a single body in TDFWI.

4-3 Source wavelet error

We also examine the effect of phase and amplitude errors on source wavelet for different FWI methods. Here we test four different shifted phases in degree (0, 10, 20, and 30) and for each phase, we use two amplitudes (0.8, 1). Figure 9 shows FWI result of wavelet errors. As it can be seen from these figures, error in source wavelet phase estimation has much deteriorating effect on results than error in amplitudes.

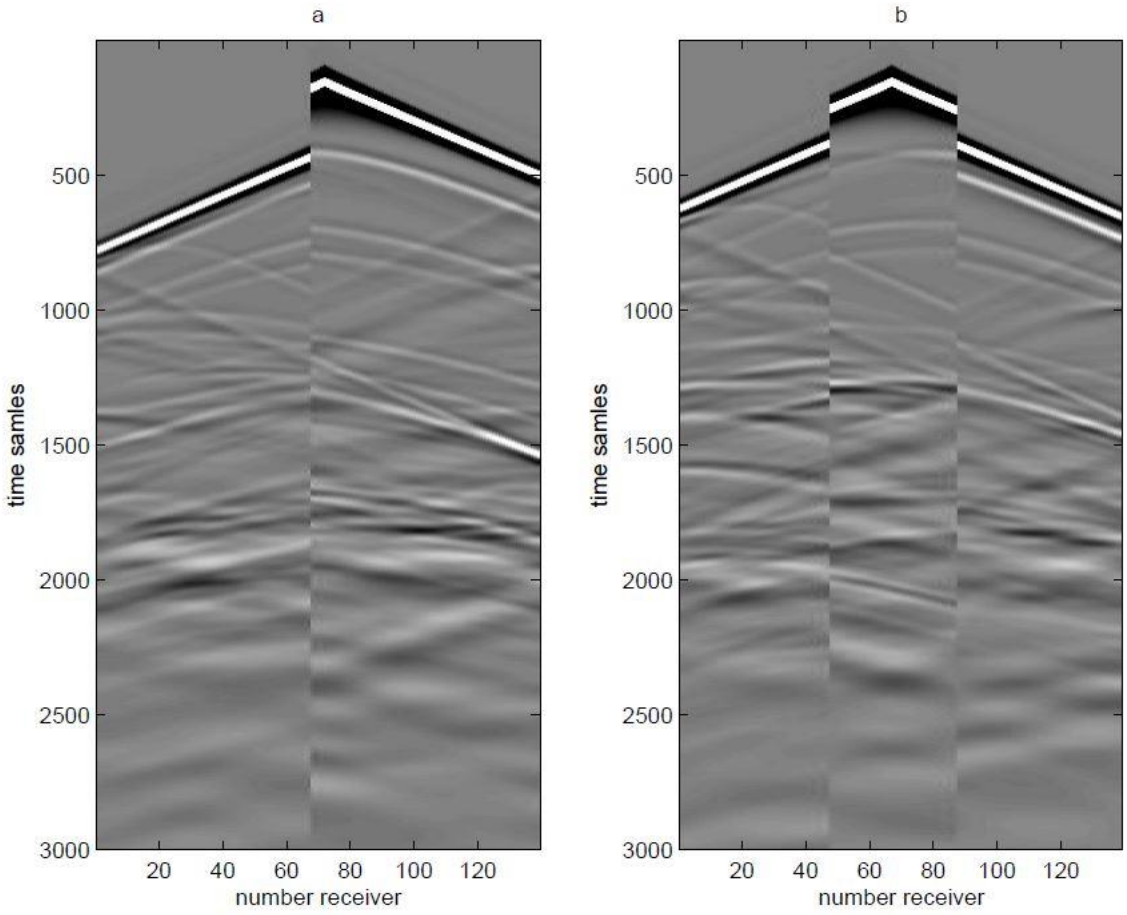

Figure 6. Receivers' data for, a) 400 meter continuous gap. b) 400 meter discrete gap

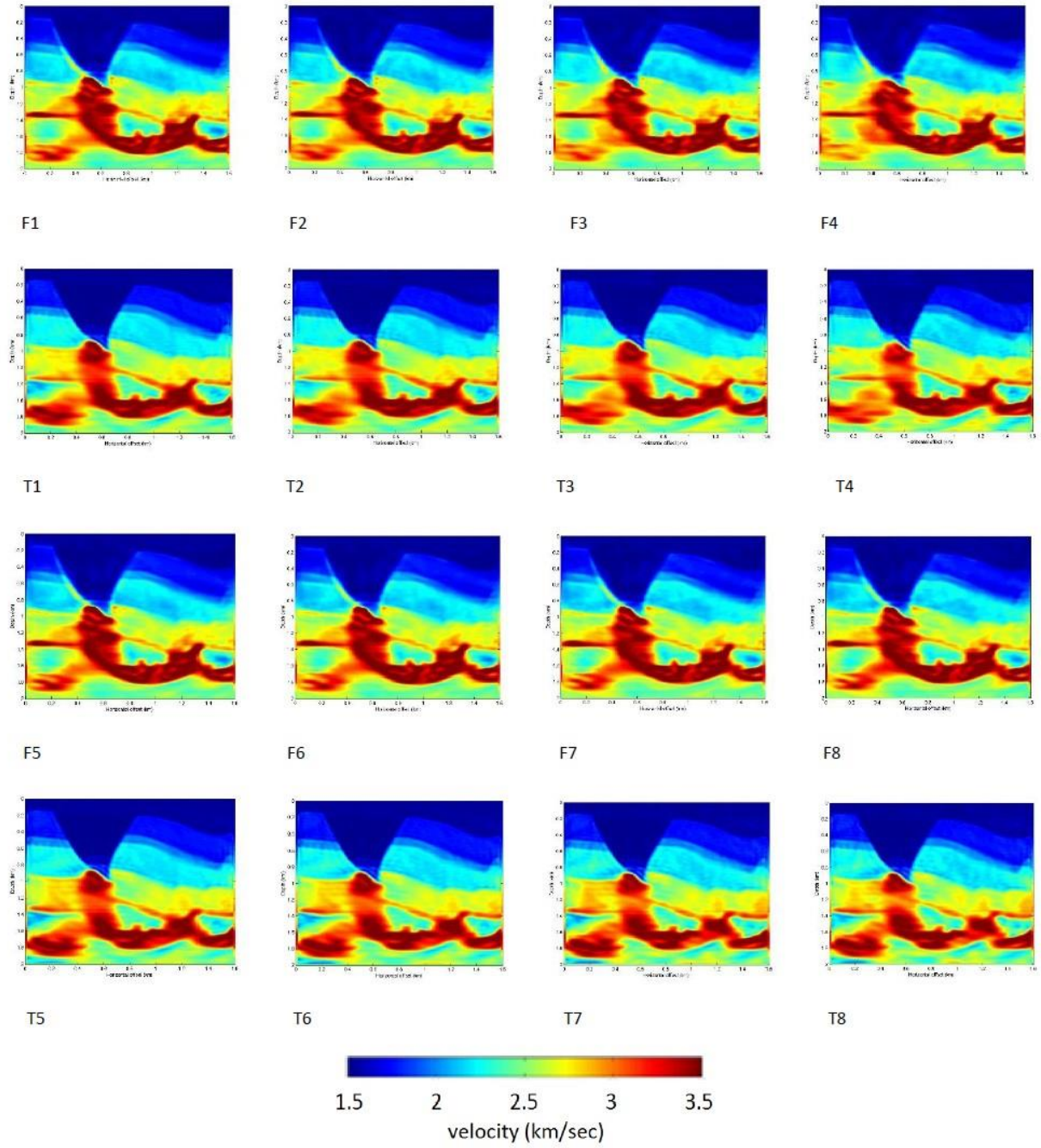

Figure 7: FWI results in presence of gap error. The horizontal and vertical axes are respectively horizontal offset (km) and depth (km). FDFWI results are, (F1) 100m continuous gap, (F2) 200m continuous gap, (F3) 300m continuous gap, (F4) 400m continuous gap. TDFWI results are, (T1) 100m continuous gap error, (T2) 200m continuous gap error, (T3) 300m continuous gap error, (T4) 400m continuous gap error, FDFWI results are, (F5) 100m discrete gap,(F6) 200m discrete gap, (F7) 300m discrete gap, (F8) 400m discrete gap. TDFWI results are, (T5) 100m discrete gap error, (T6) 200m discrete gap error, (T7) 300m discrete gap error, (T8) 400m discrete gap error.

5 Conclusions

In this paper, we first derive discretized adjoint state full waveform inversion in both time and frequency domains. The adjoint state method applied to discretized wave equation with FDM scheme rather than continuous version of wave equation which most predecessor

works has been done. This has merit of extreme accuracy in discretized level, which has been confirmed by reliability tests for inversion schemes.

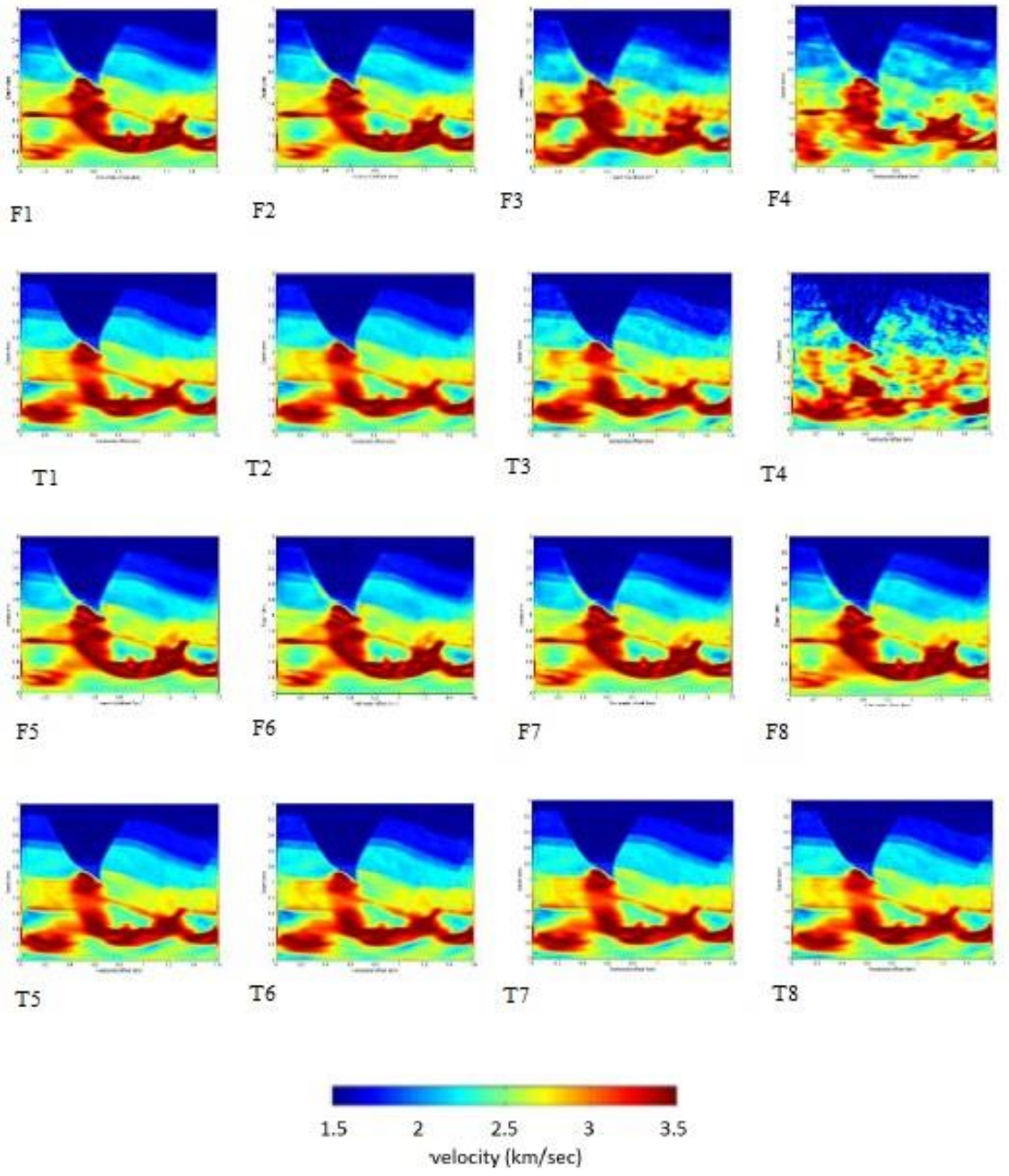

Figure 8. FWI results in presence of different errors with coherent noise. The horizontal and vertical axes are respectively horizontal offset (km) and depth (km). Frequency domain results are, (F1) sigma=0 and SNR=5dB, (F2) sigma=1 and SNR=5dB. (F3) sigma=0.3 and SNR=5dB. (F4) sigma=0.1 and SNR=5dB, (F5) sigma=0 and SNR=20dB. (F6) sigma=1 and SNR=20dB, (F7) sigma=0.3 and SNR=20dB,(F8) sigma=0.1 and SNR=20dB. Time domain results are, (T1) sigma=0 and SNR=5dB, (T2) sigma=1 and SNR=5dB, (T3) sigma=0.3 and SNR=5dB, (T4) sigma=0.1 and SNR=5dB, (T5) sigma=0 and SNR=20dB, (T6) sigma=1 and SNR=20dB, (T7) sigma=0.3 and SNR=20dB, (T8) sigma=0.1 and SNR=20dB.

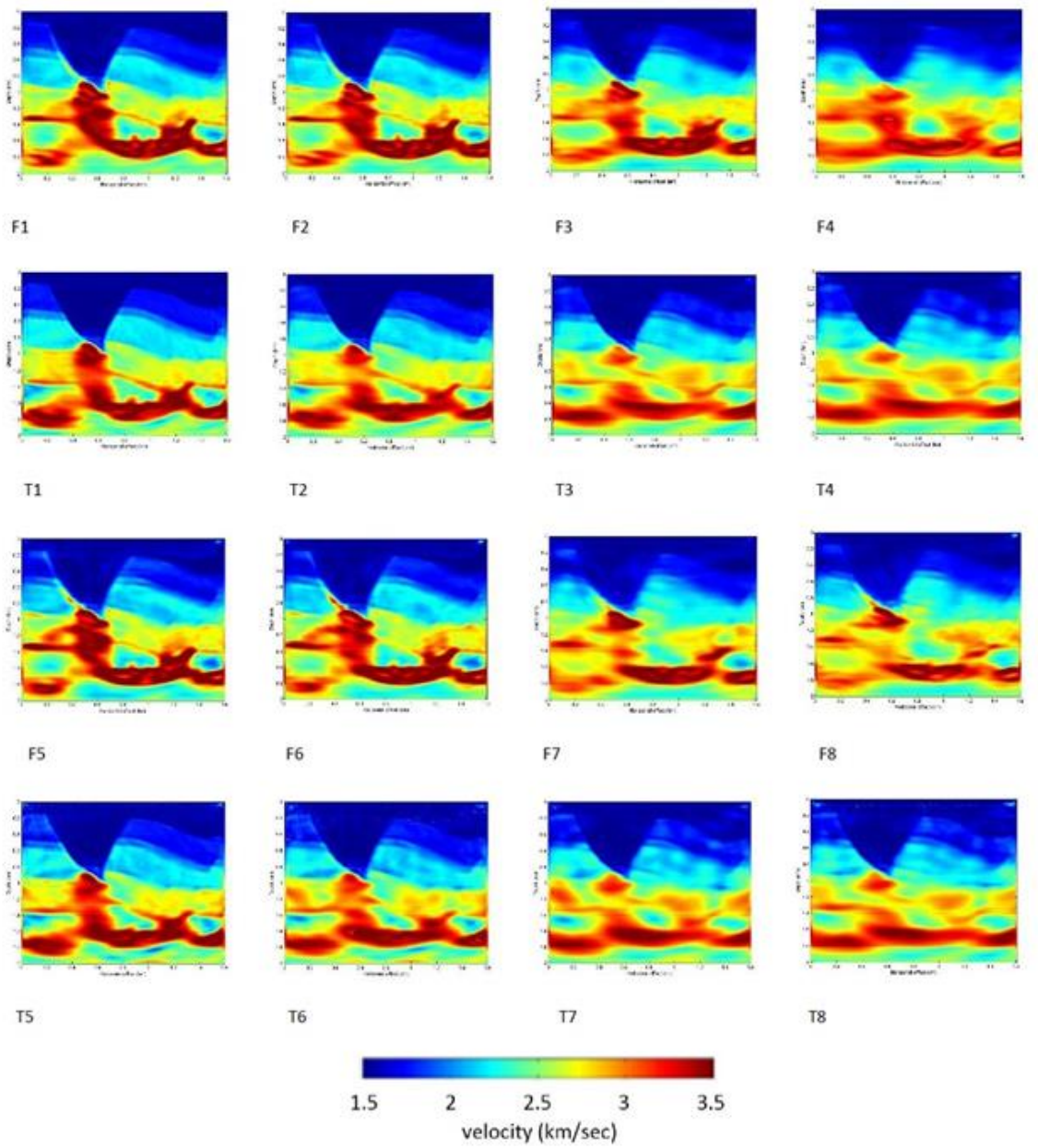

Figure 9. FWI results in presence of different source wavelet errors. The horizontal and vertical axes are respectively horizontal offset (km) and depth (km). Frequency domain results are, (F1) Phase=0 and AMP=1, (F2) Phase=10 and AMP=1,(F3) Phase=20 and AMP=1,(F4) Phase=30 and AMP=1,(F5) Phase=0 and AMP=0.8,(F6)Phase=10 and AMP=0.8, (F7) Phase=20 and AMP=.8, (F8) Phase=30 and AMP=.8. Time domain results are, (T1) Phase=0 and AMP=1, (T2) Phase=10 and AMP=1, (T3) Phase=20 and AMP=1, (T4) Phase=30 and AMP=1, (T5) Phase=0 and AMP=0.8, (T6) Phase=10 and AMP=0.8, (T7) Phase=20 and AMP=.8, (T8) Phase=30 and AMP=.8

In this paper, we also introduce reliability tests that comprise of objective function descent test and Jacobian test. Objective function test states that objective function needs to decrease per iteration no matter which kind of optimization or methodology has been used for inversion. Jacobian test is a test to verify that the gradient of the objective function, which is a key element of the inversion, is properly calculated. In another word "gradient is gradient".

In this study, we also compare TDFWI and FDFWI by designing different tests

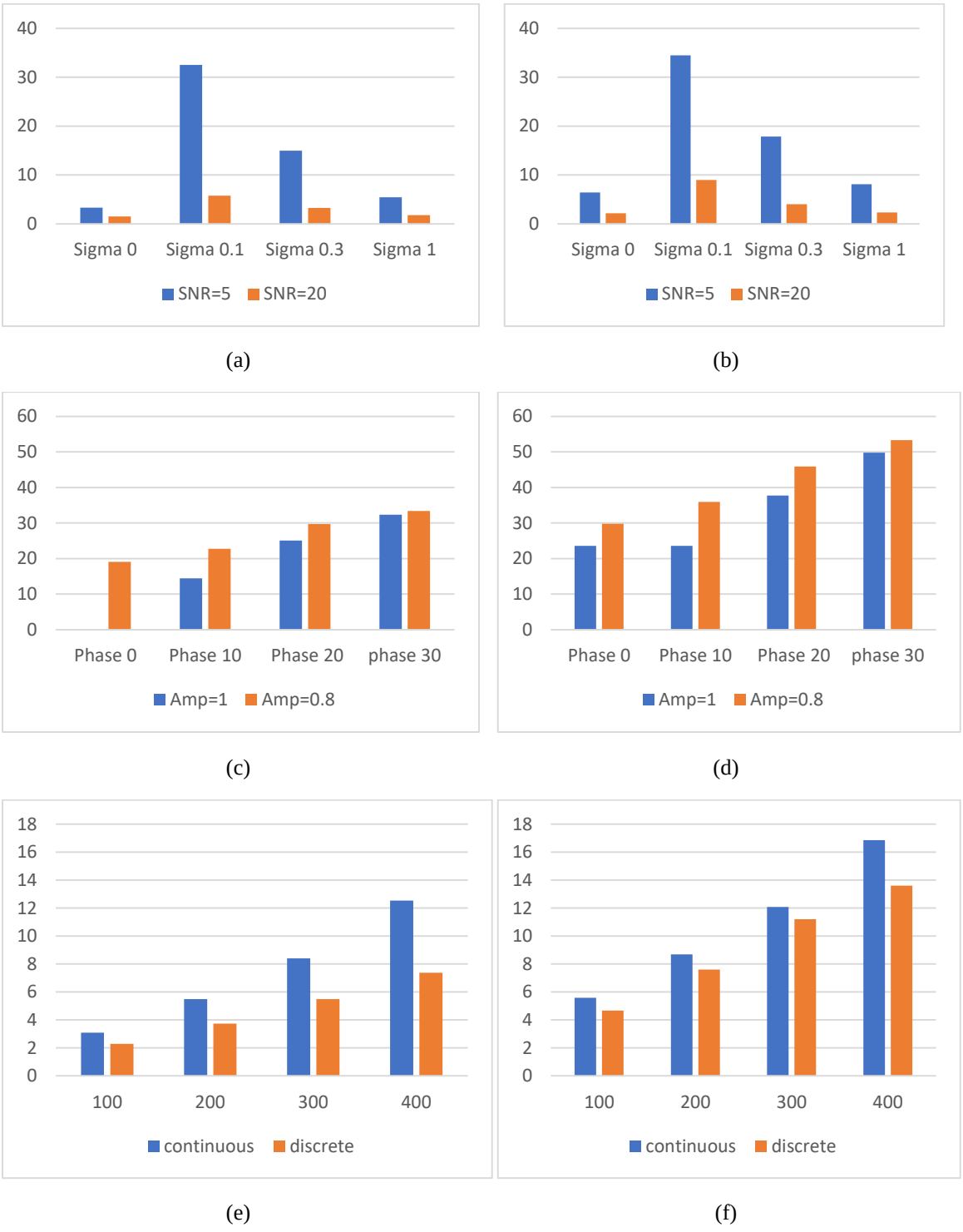

Figure 10. Error calculated from Equation (3), (a) Coherent noise error in frequency domain, (b) Coherent noise error in time domain, (c) Wavelet source error in frequency domain, (d) Wavelet source error in time, (e) Gap error in frequency domain, (f) Gap error in time domain.

to analyze the effect of various errors in data which we call it data imperfections. Data imperfections include random and coherent noises in data, errors in estimation of source wavelet and existence of gap in the seismic survey.

The results of noise tests shown in Figure 10 prove that coherent noise has a great destructive impact on TDFWI and FDFWI, especially when SNR is low. The gap tests demonstrate that one big gap in seismic survey cause more deterioration in FWI results than series of dis-joint gaps with sum of equal size. The source wavelet results show that both TDFWI and FDFWI are sensitive to errors in amplitude and phase of source wavelet with more sensitivity on phase error. The result of all tests shows more deterioration in TDFWI than that of FDFWI. In other words, for every test which we conducted, we observed that FDFWI is more preferable than TDFWI in term of resistance to data imperfection.

References

Berenger, J.P., 1994, A perfectly matched layer for the absorption of electromagnetic waves: Journal of computational physics, **114**(2), 185-200.

Boonyasiriwat, C., Valasek, P., Routh, P., Cao, W., Schuster, G.T. and Macy, B., 2009, An efficient multiscale method for time-domain waveform tomography: Geophysics, **74**(6), WCC59-WCC68.

Brenders, A.J. and Pratt, R.G., 2007. Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model: Geophysical Journal International, **168**(1), 133-151.

Bunks, C., Salek, F. M., Zaleski, S., and Chavent, G., 1995, Multiscale seismic waveform inversion: Geophysics, 60, 1457-1473.

Cerjan, C., Kosloff, D., Kosloff, R. and Reshef, M., 1985, A nonreflecting boundary condition for discrete acoustic and elastic wave equations: Geophysics, **50**(4), 705-708.

Chavent, G., 1974, Identification of functional parameters in partial differential equations: Joint Automatic Control Conference, 12, 155-156.

Chen, P., 2011, Full-wave seismic data assimilation: theoretical background and recent advances: Pure and applied geophysics, **168**(10), 1527-1552.

Chopra, S. and Marfurt, K., 2014, Causes and appearance of noise in seismic data volumes: Search and Discovery Article, 41476.

Clayton, R. and Engquist, B., 1977, Absorbing boundary conditions for acoustic and elastic wave equations: Bulletin of the seismological society of America, **67**(6), 1529-1540.

Engquist, B. and Majda, A., 1977, Absorbing boundary conditions for numerical simulation of waves. Proceedings of the National Academy of Sciences, **74**(5), 1765-1766.

Fichtner, A., 2010, Full seismic waveform modelling and inversion: Springer Science & Business Media.

Higdon, R.L., 1991, Absorbing boundary conditions for elastic waves: Geophysics, **56**(2), 231-241.

Jiang, J. and Zhu, P., 2018. Acceleration for 2D time-domain elastic full waveform inversion using a single GPU card. Journal of Applied Geophysics, **152**,173-187.

Keys, R.G., 1985, Absorbing boundary conditions for acoustic media: Geophysics, **50**(6), 892-902.

Komatitsch, D. and Tromp, J., 2003, A perfectly matched layer absorbing boundary condition for the second-order seismic wave equation: Geophysical Journal International, **154**(1), 146-153.

Liu, Q. and Tromp, J., 2006, Finite-frequency kernels based on adjoint methods: Bulletin of the Seismological Society of America, **96**(6), 2383-2397.

Mao, J., Wu, R.S. and Wang, B., 2012. Multiscale full waveform inversion using GPU. In SEG Technical Program Expanded Abstracts, 1-7. Society of Exploration Geophysicists.

McGarry, R. and Moghaddam, P., 2009, NPML boundary conditions for second-order wave equations: 79th Annual Meeting SEG, Expanded Abstracts, 3590-3594. Society of Exploration Geophysicists.

Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multi offset seismic data: Geophysics, **52**(9), 1211-1228.

Oldenburg, D., 1990, Inversion of electromagnetic data: An overview of new techniques: Surveys in Geophysics, **11**(2-3), 231-270.

Pasalic, D. and McGarry, R., 2010, Convolutional perfectly matched layer for isotropic and anisotropic acoustic wave equations: 80th Annual Meeting SEG, Expanded Abstracts, 2925-2929. Society of Exploration Geophysicists.

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**(2), 495-503.

Pratt, R.G., 1990, Inverse theory applied to multi-source cross-hole tomography. Part 2: Elastic wave-equation method: Geophysical Prospecting, **38**(3), 311-329.

Pratt, R.G. and Worthington, M.H., 1990, Inverse theory applied to multi-source cross-hole tomography. Part 1: Acoustic wave-equation method: Geophysical prospecting, **38**(3), 287-310.

Reynolds, A.C., 1978. Boundary conditions for the numerical solution of wave propagation problems: Geophysics, **43**(6), 1099-1110.

Sheng, J., Leeds, A., Buddensiek, M. and Schuster, G.T., 2006, Early arrival waveform tomography on near-surface refraction data: Geophysics, **71**(4), U47-U57.

Sirgue, L. and Pratt, R.G., 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**(1), 231-248.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**(8), 1259-1266.

Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**(10), 1893-1903.

Tarantola, A., 1988, Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation: Scattering and Attenuations of Seismic Waves, Part I, 365-399. Birkhäuser, Basel.

Virieux, J. and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**(6), WCC1-WCC26.

Yang, P., Gao, J. and Wang, B., 2015, A graphics processing unit implementation of time-domain full-waveform inversion: Geophysics, **80**(3), F31-F39.

Yao, G., da Silva, N.V., Warner, M., Wu, D. and Yang, C., 2019. Tackling cycle skipping in full-waveform inversion with intermediate data. Geophysics, **84**(3), R411-R427.