A comparison between the behavior of objective functions for waveform inversion in the frequency and Laplace domains

Changsoo Shin¹ and Wansoo Ha¹

ABSTRACT

In the frequency domain, gradient-based local-optimization methods of waveform inversions have been unsuccessful at inverting subsurface parameters without an accurate starting model. Such methods could not correct automatically for poor starting models because multiple local minima made it difficult to approach the true global minimum. In this study, we compared the behavior of objective functions in the frequency and Laplace domains. Wavefields in the Laplace domain correspond to the zero-frequency component of a damped wavefield; thus, the Laplace-domain waveform inversion can image smooth velocity models. Objective functions in the Laplace-domain inversion have a smoother surface and fewer local minima than in the frequency-domain inversion. We applied the waveform inversion to a 2D slice of the acoustic SEG/EAGE salt model in the Laplace domain and recovered smooth velocity models from inaccurate initial velocity conditions. We also successfully imaged velocities of the salt, SEG overthrust, and Institut Français du Pétrole Marmousi models with the frequency-domain inversion method by using the inverted velocity model of the Laplace-domain inversion as the initial model.

INTRODUCTION

Geophysicists have devoted considerable effort to seismic-inversion methods during the past 20 years. Most waveform-inversion methods are based on the local optimization approach introduced by Lailly (1984) and Tarantola (1984). The objective function of this classic waveform-inversion method is the $l_2$-norm of differences between modeled and observed wavefields.

The method back-propagates residual data errors in a manner similar to reverse-time migration, and the use of back propagation avoids calculating the computationally intensive Fréchet derivatives. Instead, the steepest descent direction is found efficiently by multiplying back-propagated residuals with virtual sources (Pratt et al., 1998). Then one can find the solution through a suite of iterations. Kolb et al. (1986) and Gauthier et al. (1986) present numerical results of this method in the time domain, and Geller and Hara (1993), Pratt et al. (1998), and Pratt (1999) present results in the frequency domain.

However, iterative inversion methods have been unsuccessful at inverting the seismic data obtained from complex earth models. One of the main difficulties for nonlinear seismic inversion is the presence of numerous local minima in the objective function (Bunks et al., 1995). Obviously, an accurate starting-velocity model is needed to avoid this problem.

Global optimization algorithms have advantages over local optimization approaches because they can escape local minima and do not require an accurate initial model or function derivatives to find a global minimum (Ma, 2001). Indeed, global optimization methods such as simulated annealing and genetic algorithms were introduced to avoid the local minima problem of nonlinear least-squares inversions. Simulated annealing, a sophisticated version of a Monte Carlo technique, is a random search that samples the objective function in the space of independent variables (Landa et al., 1989; Sen and Stoffa, 1991; Ma, 2002). Genetic algorithms progressively modify a set of initial solutions by mimicking the evolutionary behavior of biological systems until an acceptable result is achieved (Stoffa and Sen, 1991; Wilson et al., 1994; Boschetti et al., 1996). However, global optimization is limited by the large dimensionality involved in most geophysical optimization problems and by the large number of iterations required to find the global minimum (Clément et al., 2001). Otherwise, reducing the computational cost by using a sparse parameterization of model space introduces biases.

Several approaches have been developed that attempt to reduce the number of local minima in the objective function. In a Bayesian approach, the inversion can be formulated as an estimation problem in which the a priori probability distribution is the Gibbs distribution. Then solutions are obtained by maximizing the posterior probability of model parameters using a Monte Carlo method (Rothman,
The layer-based waveform-inversion procedure was designed to determine interval velocities, impedance contrasts, and interface positions for a limited number of layers to explain major seismic events (Jurado et al., 1995). Sirgue and Pratt (2004) introduce the frequency-domain multiscale approach, which avoids many local minima by inverting frequencies sequentially from low to high. The multigrid method improves the performance of iterative inversions by decomposing the problem by scale (Bunks et al., 1995). The reduced least-squares approach involves nested optimization with a linear inner loop for high-wavenumber components and a nonlinear outer loop for low-wavenumber components (Jin and Madariaga, 1994). There also have been attempts to reduce the number of local minima in the objective function by adopting some measure of coherency (Al-Yahya, 1987; Landa et al., 1989; Symes and Carazzone, 1991). Brenders and Pratt (2007) use complex-valued frequencies for waveform tomography to ease the local minima problem. Sheng et al. (2006) developed early arrival waveform tomography, which fits the waveforms of early arrivals. An early arrival misfit function generally contains fewer local minima than that of the full-waveform tomography.

In this paper, we examine the behavior of objective functions for waveform inversion in the Laplace domain. Waveform inversions in the Laplace domain have been introduced by Shin and Cha (2008). A wavefield in the Laplace domain can be thought of as a zero-frequency component of a damped wavefield. It can recover a smooth velocity model from data containing no low-frequency information, provided a suitable objective function is used (Cha, 2008; Ha, 2008; Shin and Cha, 2008).

The method considered here uses a complex frequency, similar to the work of Brenders and Pratt (2007). Using positive values in the imaginary part of the frequency is equivalent to damping the time-domain signals. By damping the time signals, our approach places more weight on early arrivals than on late arrivals. Accordingly, this method is similar to the early arrival waveform tomography method (Sheng et al., 2006). However, we use zero for the real part of the frequency and use several damping constants in the imaginary part. When the damping constant is large, only the first arrival of the data is used and the method is similar to traveltime tomography. However, our method can image the velocity of deep structures by inverting at several damping constants ranging from low to high.

Using a synthetic salt model, we show that the objective functions in the Laplace domain have fewer local minima than in the frequency domain. We can avoid local minima yet still attain an efficient local-minimization method by conducting the waveform inversion in the Laplace domain. We confirm that smooth background velocity models can be obtained from inaccurate initial velocity models by presenting numerical examples of a 2D SEG/EAGE salt model. A smooth velocity model obtained by the Laplace-domain inversion can then be used as an initial velocity model in the frequency-domain inversion to obtain a more accurate result. Finally, we show that the inverted velocity model obtained in the Laplace domain can be used successfully for the frequency-domain inversion using the salt, SEG overthrust, and IFP Marmousi models.

**THEORY**

First, we review the relationship between frequency-domain and Laplace-domain wavefields (Cha, 2008; Ha, 2008; Shin and Cha, 2008). The acoustic-wave equation can be expressed as a matrix equation using the finite-element method, given by

\[ \mathbf{M}u + C\dot{u} + Ku = f, \]  

(1)

where \( \mathbf{M} \) is a mass matrix, \( \mathbf{K} \) is a stiffness matrix, \( \mathbf{C} \) is a damping matrix, \( \mathbf{u} \) is a wavefield in the time domain, and \( \mathbf{f} \) is a source vector (Marfurt, 1984; Zienkiewicz and Taylor, 1991). Taking the Laplace transform, we obtain

\[ \mathbf{S}\tilde{u} = \tilde{f}, \]

(2)

where

\[ \mathbf{S} = \mathbf{K} + sC + s^2\mathbf{M}, \]

\[ \tilde{\mathbf{u}}(s) = \int \mathbf{u}(t)e^{-st}dt, \]

\[ \tilde{\mathbf{f}}(s) = \int \mathbf{f}(t)e^{-st}dt, \]

(3)

and \( s \) is a positive, real damping constant. Because the wavefield \( \mathbf{u}(t) \) is a causal signal and therefore vanishes when \( t < 0 \), \( \tilde{\mathbf{u}}(s) \) in equation 3 can be expressed as

\[ \tilde{\mathbf{u}}(s) \approx \lim_{\omega \to 0} \int_{0}^{\infty} \mathbf{u}(t)e^{-st}dt = \lim_{\omega \to 0} \int_{-\infty}^{\infty} \mathbf{u}(t)e^{-st}e^{-i\omega t}dt. \]

(4)

Here we introduced the complex damping constant \( \text{s} = s + i\omega \). The right-hand side of equation 4 can also be obtained from the Fourier transform by using a complex frequency \( \omega^* = \omega - is \) and allowing \( \omega \) to approach zero, or

\[ \lim_{\omega \to 0} \tilde{\mathbf{u}}(\omega) \approx \lim_{\omega \to 0} \int_{-\infty}^{\infty} \mathbf{u}(t)e^{-i\omega s}dt \]

\[ = \lim_{\omega \to 0} \int_{-\infty}^{\infty} \mathbf{u}(t)e^{-st}e^{-i\omega t}dt. \]

(5)

The imaginary part in the complex frequency introduces damping to the time-domain signal. Equations 4 and 5 imply that a Laplace-transformed wavefield corresponds to the zero-frequency component of a damped wavefield. We can obtain the zero-frequency component of the damped wavefield by modeling in the frequency domain with complex frequencies whose real parts are zero and imaginary parts are \(-s\). This fact greatly simplifies the implementation of waveform inversions in the Laplace domain; we can use an algorithm for frequency-domain inversion without further modification.

There is no zero frequency in a wavefield. However, the damped wavefield generally has the zero-frequency information. For in-

---

**Figure 1.** Sine function as defined by equation 6.
stance, consider a function $f(t)$ defined as

$$f(t) = \begin{cases} 
0, & t < 0 \\
\sin \omega_0 t, & 0 \leq t \leq 2N\pi/\omega_0 \\
0, & t > 2N\pi/\omega_0 
\end{cases} \quad (6)$$

where $\omega_0 > 0$ and $N$ is very large (Figure 1). This function has no zero-frequency value, i.e.,

$$\int_{-\infty}^{\infty} f(t)e^{-i0^+t}dt = \int_{0}^{2N\pi/\omega_0} \sin \omega_0 t dt = 0.$$

(7)

However, the zero frequency of the damped function is not zero, as seen by

$$\int_{-\infty}^{\infty} f(t)e^{-st}e^{-i0^+t}dt$$

$$= \int_{0}^{2N\pi/\omega_0} \sin \omega_0 t e^{-st} dt$$

$$= \int_{0}^{\infty} \sin \omega_0 t e^{-st} dt = L(\sin \omega_0 t)$$

$$= \frac{\omega_0}{s^2 + \omega_0^2} \neq 0,$$

(8)

where $s > 0$ and $L$ denotes the Laplace transform.

The damped signal $u(t)e^{-st}$ contains information on shallow structures and early arrivals when the damping constant $s$ is large, and information on shallow and deep structures, and early and late arrivals, when $s$ is small (Figure 2).

In the frequency domain, we used a fixed damping constant to generate the forward-modeled data $\tilde{u}(\omega)$. In the Laplace domain, we set the angular frequency to zero and varied the damping constant to generate the forward-modeled data $\tilde{u}(s)$. In the frequency domain, the field data were transformed using a fast Fourier transform (FFT). In the Laplace domain, the field data were transformed by a simple numerical integration, such as the trapezoidal rule (Figure 3). We calculated the residual vector between the forward-modeled data and observed data for each damping constant (or frequency) and each shot. Then we computed the gradient direction by multiplying a virtual source vector by the back-propagated residual vector. Dividing the gradient direction by the pseudo-Hessian (Shin et al., 2001), we then obtained the velocity-update direction. We updated the velocity and iterated until the mean of squared residuals met our convergence criterion. Note
that the only difference between algorithms in the two domains is the choice of frequency and damping constant. The detailed frequency-domain algorithm is presented by Shin et al. (2007).

When the damping constant is zero, the wave equation in the Laplace domain is the same as Poisson’s equation found in electrical prospecting. Therefore, this method resembles the inversion of Poisson’s equation that usually provides smooth inversion results (Shin and Cha, 2008).

CONFIGURATIONS OF OBJECTIVE FUNCTIONS

Velocity inversion by gradient methods requires that the objective function be smooth and unimodal over a large range of velocities (Chauris and Noble, 2001). We examined configurations of objective functions numerically in the frequency and Laplace domains. A summary of our selected functions is given in Table 1; it includes the $l_2$-norm, logarithmic objective function, amplitude and phase objective function (Shin and Min, 2006), Huber norm and $l_1$-norm with two types of implementations for complex values (Claerbout and Muir, 1973; Bube and Langan, 1997; Guitton and Symes, 2003; Chung et al., 2007; Pyun, 2007), and power and integral objective functions (Ha, 2008). The value $u_{ijk}$ is the modeled data, $d_{ijk}$ is the observed data, $N_f$ is the number of frequencies in the frequency domain or the number of damping constants in the Laplace domain, $N_r$ is the number of shots, and $N_s$ is the number of receivers.

These objective functions are tested widely except for the power and integrals. The logarithmic, power, and integral objective functions take the log, power, and integral of the wavefield, respectively. They produce a natural scaling of the residual wavefield by the modeled wavefield (Shin et al., 2007) and can be thought of as a kind of $l_2$-norm with preconditioning on the modeled and observed data.

We tested the power and integral objective function with various exponents and intervals ranging from 0.0 to 1.0. The steepest descent directions of the power and integral objective functions are given in Appendix A. The $l_1$-norm takes the absolute error and is more robust to large errors than the $l_2$-norm (Claerbout and Muir, 1973). The Huber norm approximates the $l_1$-norm when residuals between the modeled and observed data are small, and it approximates the $l_2$-norm when residuals are large (Bube and Langan, 1997; Guitton and Symes, 2003).

Figure 4 shows a velocity model developed to examine configurations of objective functions. In the true velocity model, the velocity of the salt body is 4.5 km/s and the background is a linearly increasing velocity model varying from 1.5 to 3.5 km/s. We simulated the observed data from this model using a finite-element method in both domains. Trial velocity models were produced by changing the velocity of the salt $V_s$ from 1 to 6 km/s and the maximum background velocity $V_{max}$ from 1.5 to 5.5 km/s. Then we obtained the modeled data and calculated each objective function. There were 77 shots and 391 receivers, the shot interval was 200 m, the receiver interval was 40 m, and the grid size was 40 m. The maximum offset was 15.6 km and the recording time was 10 s. We placed shots and receivers one grid point below the surface so we could apply free-surface boundary conditions in this and following examples.

**Table 1. Objective functions used in this research.**

| Objective functions | $E(m) = \sum_{i=1}^{N_f} \sum_{j=1}^{N_r} \left| u_{ijk} - d_{ijk} \right|^2$ |
|---------------------|-----------------------------------------------------------------|
| $l_2$               | $E(m) = \sum_{i=1}^{N_f} \sum_{j=1}^{N_r} \left| |r_{ijk}| \right| \leq \epsilon $ |
| Logarithmic         | $E(m) = \sum_{i=1}^{N_f} \sum_{j=1}^{N_r} \left[ \left| \ln u_{ijk} \left[ r_{ijk} \right] \right| \leq \epsilon \right. \left. \left| \ln r_{ijk} \right| \right] $ |
| Amplitude           | $E(m) = \sum_{i=1}^{N_f} \sum_{j=1}^{N_r} \left| \left| \ln u_{ijk} \left[ r_{ijk} \right] \right| \right| \left| \ln r_{ijk} \right| $ |
| Phase               | $E(m) = \sum_{i=1}^{N_f} \sum_{j=1}^{N_r} \sum_{k=1}^{N_s} \left| M(r_{ijk} = u_{ijk} - d_{ijk}) \right|^2$ |
| Huber               | $E(m) = \sum_{i=1}^{N_f} \sum_{j=1}^{N_r} \sum_{k=1}^{N_s} \left| M(r_{ijk} = u_{ijk} - d_{ijk}) \right|^2$ |
| Integral            | $E(m) = \sum_{i=1}^{N_f} \sum_{j=1}^{N_r} \sum_{k=1}^{N_s} \left| \left| \left| \ln u_{ijk} \left[ r_{ijk} \right] \right| \right| \left| \ln r_{ijk} \right| \right| \left| \ln r_{ijk} \right| $ |

| 1. $M(r_{ijk}) = \left\{ \right.$ |
| 2. $M(r_{ijk}) = \left\{ \right.$ |

**Figure 4.** Two-parameter salt model. The first parameter is the velocity of the salt body. The second parameter is the maximum velocity of the background velocity. The background velocity is a linearly increasing velocity ranging from 1.5 km/s to the maximum velocity.
indicates that objective functions are sensitive to the traveltime of the wavefield, which is controlled by the background velocity.

**Configurations of objective functions in the Laplace domain**

In Figures 6 and 7, we show contours of some representative objective functions in the Laplace domain. These functions are unimodal over a wide range of velocities and damping constants and have no imaginary parts in the Laplace domain. Accordingly, the amplitude objective function is the same as the logarithmic objective function, and we cannot use the phase objective function.

The two implementations of the Huber norm and those of the $l_1$-norm in Table 1 reduce to the same function in the Laplace domain. In this domain, we can separate objective functions into two categories based on their configurations: the $l_2$-norm, $l_1$-norm, and Huber norm comprise type 1 functions, and the logarithmic objective function comprises a type 2 function. The power and integral objective functions showed both types of configurations, depending on the exponent and integral interval, respectively. The power objective function showed a type 2 configuration when the exponent was equal to or smaller than 0.1 and a type 1 configuration when the exponent was equal to or larger than 0.25. The integral objective function showed a type 2 configuration when the integral interval began at zero ($p_1 = 0.0$) and a type 1 configuration when $p_1 \geq 0.25$. Type 1 objective functions lost their sensitivity to variations in the salt velocity at high damping constants, making it impossible to find salt velocities using them. On the other hand, type 2 functions are very smooth and desirable for local optimization methods.

**Effects of maximum offset and receiver interval**

Figure 8 shows slices of the logarithmic objective function across true values of $V_s$ and $V_{\text{max}}$ in the frequency domain with varying maximum offset. Figure 9 shows the same slices with varying receiver interval. Both figures contain local minima regardless of the maximum offset distance and receiver interval. However, the section obtained in the Laplace domain shows no local minima when the offset distance changes from 2.8 to 5.2 km (Figure 10). It also
Figure 6. Configurations of several objective functions in the Laplace domain with damping constants of 10.0. Objective functions are (a) conventional $l_2$, (b) logarithmic, (c) power ($p = 0.1$), (d) power ($p = 0.5$), (e) integral ($p_1 = 0.0, p_2 = 0.5$), and (f) integral ($p_1 = 0.5, p_2 = 1.0$).
Laplace-domain waveform inversion

Figure 7. Configurations of several objective functions in the Laplace domain with damping constants of 40.0. Objective functions are (a) conventional $l_2$, (b) logarithmic, (c) power ($p = 0.1$), (d) power ($p = 0.5$), (e) integral ($p_1 = 0.0, p_2 = 0.5$), and (f) integral ($p_1 = 0.5, p_2 = 1.0$).
shows no local minima when the receiver interval is 80, 120, and 160 m but does develop local minima when the receiver interval is 200 m (Figure 11).

These smooth configurations indicate a waveform inversion with the logarithmic objective function in the Laplace domain is better suited for gradient-based minimization methods than in the frequency domain. Thus, we can use smaller offsets and larger receiver intervals than in the frequency-domain inversion to reduce acquisition or computation costs. However, larger maximum offset and finer shot spacing lead to a better-conditioned inverse problem. Although the global minimum in the Laplace domain is less well defined than that in the frequency domain, we can avoid local minima and obtain smooth velocities by inverting in the Laplace domain.

Discussions on configurations of objective functions

We can attribute differences in configurations of objective functions to differences in seismograms of each domain. Figures 12 and 13 show seismogram sections in both domains using the true velocity stated above and trial velocities of $V_{max} = 3.1 \text{ km/s}$, $V_s = 4.0 \text{ km/s}$, $V_{max} = 2.7 \text{ km/s}$, and $V_s = 4.0 \text{ km/s}$. Each figure shows the original and logarithmic seismogram of shot 39. Figure 12 uses an angular frequency of 20 rad/s (3.2 Hz) in the frequency domain; Figure 13 uses a damping constant of 20 in the Laplace domain. In Figure 12, residuals between the modeled and observed data oscillated because of seismogram oscillations in the frequency domain. At high frequency, these oscillations generate many local minima. On the other hand, the amplitude of the seismogram in the Laplace domain decreased as the offset between a shot and a receiver

![Figure 8](image1.png)

Figure 8. Configurations of the logarithmic objective function with varying maximum offset distances in the frequency domain at an angular frequency of 20.0 rad/s (3.2 Hz). Section crossings (a) $V_{max} = 3.5 \text{ km/s}$ and (b) $V_s = 4.5 \text{ km/s}$.

![Figure 9](image2.png)

Figure 9. Configurations of the logarithmic objective function with varying receiver interval in the frequency domain at an angular frequency of 20.0 rad/s (3.2 Hz). Section crossings (a) $V_{max} = 3.5 \text{ km/s}$ and (b) $V_s = 4.5 \text{ km/s}$.

![Figure 10](image3.png)

Figure 10. Configurations of the logarithmic objective function with varying maximum offset distances in the Laplace domain at a damping constant of 20.0. Section crossings (a) $V_{max} = 3.5 \text{ km/s}$ and (b) $V_s = 4.5 \text{ km/s}$.
increased (Figure 13). This made residuals between the modeled and observed data increase as the offset increased. Accordingly, fewer local minima were formed.

Although this is the case, the objective functions using seismograms in Figure 13a still generated type 1 configurations and cannot invert velocity models. Amplitudes of seismograms decreased very rapidly as the offset between shot and receiver increased. This made most of the information in the seismogram unusable. Seismograms in Figure 13b generated type 2 configurations and can invert smooth velocity models. The logarithmic, power, and integral objective function with appropriate exponents and integral intervals generated seismograms similar to Figure 13b. These objective functions have a scaling effect on seismograms and residual wavefields (Shin et al., 2007; see Appendix A). This effect changed objective functions to have a type 2 configuration.

Laplace-transformed seismograms of the real field data can oscillate. However, oscillations are smaller than those of seismograms in the frequency domain, and most of these oscillations vanish at high damping constants.

NUMERICAL EXAMPLES

We inverted a 2D section of the SEG/EAGE salt model (Figure 14a) in the Laplace domain. The grid spacing was 40 m, and the number of shots was 381. We used 42 damping constants ranging from 0.5 to 21.0. We inverted the velocity at those damping constants simultaneously. Initial velocity models of inversions were homogeneous models with velocities of 3.0 and 4.45 km/s and a linearly increasing velocity model varying from 1.5 to 4.5 km/s (Figure 14b).

Figure 15 shows inverted velocity models obtained from the logarithmic objective function after 100 iterations. All inverted velocity models are similar in shape and each recovered the smooth velocity models successfully. Inversion in the frequency domain did not converge with homogeneous initial velocity models. We cannot say that the logarithmic objective function in the Laplace domain is unimodal with a large number of parameters as it was for the two-parameter model. However, we can say that the logarithmic objective function in the Laplace domain is more robust to the choice of initial velocity model than in the frequency domain.

Figure 16 shows the mean square error defined as

$$E = \frac{1}{N_f N_r N_e} \sum_{i}^{N_f} \sum_{j}^{N_r} \sum_{k}^{N_e} \left( \ln \frac{u_{ijk}}{d_{ijk}} \right)^2. \quad (9)$$

Error curves of the inversion using homogeneous initial velocity models started at large values but converged to 0.02%–0.13% of their initial error after 150 iterations.

Figures 17 and 18 show inverted velocity models of the power objective function ($p = 0.05$) and integral objective function ($p_1 = 0.0, p_2 = 0.1$). These inverted velocity models also are similar, regardless of initial velocity models.
Successive waveform inversion in Laplace and frequency domains

The inverted velocity model of waveform inversion in the Laplace domain is a smooth velocity model that can be used for migration (Versteeg, 1993; Gray, 2000; Pacheco and Larner, 2005). However, we can obtain a more accurate velocity model by using the inverted velocity model in the Laplace domain as the initial velocity model for the waveform inversion in the frequency domain. Laplace-domain inversion can recover smooth velocity models from the data containing no low-frequency information, and frequency-domain inversion can recover additional information on velocity models.

Figure 19 shows two inverted velocity models of the frequency-domain waveform inversion. We generated the forward data in the frequency domain by using the finite-element method. The grid size was 20 m, number of shots was 778, number of receivers was 780, and maximum offset was 15.58 km. We simultaneously used 66 fre-
quencies ranging from 4.8 to 17.8 Hz. Figure 19a was obtained by using the linearly increasing velocity as the initial velocity model. It imaged reflectors and salt-top boundaries; however, it could not recover background velocity and salt-bottom boundaries. Figure 19b was obtained by using an inverted velocity of the Laplace-domain inversion (Figure 15c) as the initial velocity model. Although the left flank of the salt is imaged poorly, it could recover reflectors and salt-top and salt-bottom boundaries.

Figure 20a shows the SEG overthrust model. We made the forward data in the Laplace domain using a grid spacing of 100 m. The number of shots and receivers were both 202. The maximum offset was 20.1 km. Intervals between shots and receivers were both 100 m. We simultaneously inverted at 20 damping constants ranging from 1.0 to 20.0 and used Figure 20b as the initial velocity model for the inversion.

Figure 21a shows the inverted velocity model using the logarithmic objective function after 200 iterations. It is notably smooth and shows large-scale structures of the SEG overthrust velocity model. We inverted the model in the frequency domain using Figure 21a as the starting velocity model. We used the conventional $l_2$-norm for the inversion and made the forward data in the frequency domain using the finite-element method. The grid spacing was 25 m and number of shots was 800. The number of receivers was 802 and maximum offset was 20 km. The recording time was 4 s. We simultaneously inverted 61 frequencies ranging from 5 to 20 Hz.

Figure 21b shows the inverted velocity model after 100 iterations. Although it has some artifacts in the deep part caused by insufficient offset, it shows stratified layers and faults very clearly. On the other hand, the inversion result of the frequency domain by using the linearly increasing velocity model as the initial velocity model could not image substructures properly.

We used the IFP Marmousi model (Figure 22a) to simulate the real data. We made the forward data in the time domain using the fourth-
order finite-difference method. The recording time was 10 s. Figure 3a shows a shot gather. We applied a trapezoidal filter to the Ricker source wavelet, with a source wavelet main frequency of 8.9 Hz. Figure 3b presents the frequency spectrum of the shot gather and shows that it does not contain low-frequency information. The grid spacing was 16 m, number of shots was 576, and interval between shots was 16 m. We used 576 receivers with a maximum offset of 9.2 km. We then transformed the data to the Laplace domain for inversion in the Laplace domain. We simultaneously inverted at 10 damping constants ranging from 2.0 to 20.0 using the finite-element method and used the logarithmic objective function for the inversion. We used only 144 shots with an interval of 64 m for the inversion.

Figure 22b shows the linearly increasing initial velocity model, and Figure 23a shows the inverted velocity model after 25 iterations. The inversion recovered the large-scale structures of the Marmousi velocity model.
model. After inverting the velocity model in the Laplace domain, we transformed the time-domain data into the frequency domain and inverted using Figure 23a as the initial velocity model, using every shot and receiver for the inversion. We simultaneously inverted at 68 frequencies ranging from 5.0 to 18.4 Hz using the finite-element method. The objective function was the conventional $l_2$-norm.

Figure 23b shows the inverted velocity model after 300 iterations. Although there are small artifacts in the deep part caused by insufficient offset, it clearly demonstrates recovery of the model faults and target reservoir. On the other hand, the inversion result obtained by using the linearly increasing initial velocity model failed to image the faults and reservoir (Figure 23c).

Figures 24 and 25 show estimated source wavelets and error curves in the Laplace-domain and frequency-domain inversions, respectively. We estimated the source wavelet by using the method of Shin et al. (2007). The mean square error of the Laplace-domain inversion is defined using equation 9; that of the frequency-domain inversion is defined as

$$E = \frac{1}{N_fN_rN_t} \sum_{f=1}^{N_f} \sum_{r=1}^{N_r} \sum_{t=1}^{N_t} (u_{ijk} - d_{ijk})^2,$$

Error curves and inverted velocity models confirm the inversion was successful.

Examples using a 2D slice of the SEG/EAGE salt, SEG overthrust, and IFP Marmousi models show that successive waveform inversions in the Laplace and frequency domains can be used for complex models. This approach successfully inverted complex models without using low frequencies in the frequency domain. We speculate that this also can be applied to complex real data. More synthetic and real data examples are found in Cha (2008), Ha (2008), and Shin and Cha (2008).

**CONCLUSIONS**

We have examined configurations of various objective functions in the frequency and Laplace domains. Our numerical calculations using the two-parameter salt model show that objective functions in the frequency domain develop many local minima at high frequencies. On the other hand, the logarithmic, power, and integral objective functions with appropriate exponents and integral intervals in the Laplace domain show no local minima across a large range of velocities and damping constants.

We argue that these characteristics are caused by seismograms in the Laplace domain. As the offset between a shot and receiver increases, the amplitude of the seismogram decreases rapidly. Seismograms in the Laplace domain are much simpler than in the frequency domain and make objective functions much smoother. Accordingly, logarithmic, power, and integral objective functions in the Laplace domain are much more robust with respect to the initial velocity model than objective functions in the frequency domain.

The waveform inversion in the frequency domain using homogeneous initial velocity models fails to invert velocity models. However, the inversion using logarithmic, power, and integral objective functions in the Laplace domain is robust enough to invert smooth velocity models from inaccurate initial conditions. We confirm the
robustness of the objective function using the SEG/EAGE salt model. The inverted velocity model of the Laplace-domain inversion can be used successfully as an initial velocity model of the frequency-domain inversion to obtain a high-resolution velocity model. The waveform inversion of a 2D slice of the SEG/EAGE salt, SEG overthrust, and IFP Marmousi models in the frequency domain, using the inverted velocity model of the Laplace-domain inversion as the initial velocity model, could recover the velocity of complex structures.

We speculate that this method can also be applied to real data. One would still need to find the optimal range of Laplace damping constants and the optimal range of the exponent and integral interval of the power and integral objective functions. The method also will have to be extended to 3D elastic seismic problems. It is possible to apply successive inversion of a few damping constants (ranging from low to high) to reduce computational costs. With these advances, the strategy outlined here can be implemented for eventual imaging of the interior of the earth.

ACKNOWLEDGMENTS

This work was supported financially by the National Laboratory Project of the Ministry of Science and Technology, Korea Research Foundation grant KRF-2007-314-D00320 funded by the Korean government (MOEHRD), and Brain Korea 21 Project of the Ministry of Education.

APPENDIX A

STEPEST DESCENT DIRECTION FOR POWER AND INTEGRAL OBJECTIVE FUNCTIONS

The first step in minimizing an objective function $E$ is to find the gradient $\nabla E$, or equivalently $\partial E/\partial m_l$, for every parameter $m_l$. Differentiating equation 2 with respect to a model parameter $m_l$, we obtain

$$ \frac{\partial S}{\partial m_l} \hat{u} + S \frac{\partial \hat{u}}{\partial m_l} = 0, \quad l = 1, 2, \ldots, N_p $$

or

$$ \frac{\partial \hat{u}}{\partial m_l} = S^{-1} v_l, \quad l = 1, 2, \ldots, N_p, $$

where $N_p$ is the number of parameters and

$$ v_l = - \frac{\partial S}{\partial m_l} \hat{u}. $$

In equation A-2, $v_l$ is the virtual source vector required to perturb the $l$th parametric element (Pratt et al., 1998). Equation A-2 defines the partial derivative wavefield.

Our problem can be solved on frequency by frequency. Once the gradient direction for each frequency has been calculated, it is necessary to sum them over all frequencies to obtain the final parameter update. If we define $\delta r_{ijk}$ as the residual error at the $k$th receiver of the $j$th shot of the $l$th frequency, then the objective function of interest is

$$ E(m) = \frac{1}{2} \sum N_f \sum N_r \sum N_p \delta r_{ijk} \delta r_{jik}. \quad (A-4) $$

The gradient direction of the objective function with respect to the $l$th parameter $m_l$ can be expressed as

$$ \frac{\partial E}{\partial m_l} = \sum N_f \text{Re} \left[ \sum N_r \sum N_p \frac{\partial \delta r_{ijk}}{\partial m_l} \delta r_{jik} \right], \quad l = 1, 2, \ldots, N_p, $$

where the subscript $l$ is the parameter number. In the power objective function (Ha, 2008), the residual is $\delta r_{ijk} = u^p_{ijk} - d^p_{ijk}$ and the gradient direction is

$$ \frac{\partial E}{\partial m_l} = \sum N_f \text{Re} \left[ \sum N_r \sum N_p (p \cdot u^p_{ijk} - 1) \frac{\partial u^p_{ijk}}{\partial m_l} (u^p_{ijk} - d^p_{ijk}) \right]. $$

(A-5)

In the integral objective function (Ha, 2008), the residual is $\delta r_{ijk} = \int_{t_1}^{t_2} u^p_{ijk} dp - \int_{t_1}^{t_2} d^p_{ijk} dp$. Hence, the gradient direction is

$$ \frac{\partial E}{\partial m_l} = \sum N_f \text{Re} \left[ \sum N_r \sum N_p \left( \frac{p_2 \ln u_{ijk} - 1}{u^p_{ijk} - 1} - \frac{p_1 \ln u_{ijk} - 1}{u^p_{ijk} - 1} \right) \right], $$

where $N_r$ is the number of receivers and $N_p$ is the number of parameters.

Transposing the partial derivative wavefield, expressed in terms of the complex-impedance matrix $S$ with the virtual source $v_l$ in equation A-2 and substituting it into equations A-6 and A-7 yields

$$ \frac{\partial E}{\partial m_l} = \text{Re}[(v_l)^T S^{-1} r]. $$

(A-8)

For all parameters, equation A-8 can be rewritten as

$$ \nabla E = \text{Re}[V^T S^{-1} r], $$

(A-9)

where $V = [v_1, v_2, \ldots, v_{N_p}]$ is the virtual-source matrix consisting of all virtual source vectors. Equation A-9 is a frequency-domain matrix-vector expression of Tarantola’s (1984) original time-domain formula for computing the steepest descent direction in a classic seismic inverse problem. It says that to obtain the gradient of $E$ and, consequently, the steepest descent direction, we must back propagate the residual data in a time-reversed order with a full two-way wave equation and then compute the zero-lag convolution between the virtual source and back-propagated wavefield (Shin et al., 2007).

REFERENCES


Brenders, A. J., and R. G. Pratt, 2007, Full waveform tomography for lithospheric imaging: Results from a blind test in a realistic crustal model: Geo-


