Stanford Exploration Project, Report 123, October 31, 2005, pages 101–128

Iterative linearized migration and inversion

Huazhong Wang[1]

ABSTRACT

The objective of seismic imaging is to obtain an image of the subsurface reflectors, which is very important for estimating whether a reservoir is beneficial for oil/gas exploration or not. It can also provide the relative changes or absolute values of three elastic parameters: compressional wave velocity V_p, shear wave velocity V_s, and density ρ. Two ways can achieve the objectives. In approach I, the angle reflectivity is given by prestack depth/time migration or linearized inversion, and the relative changes of the three elastic parameters are estimated from the angle reflectivity by AVO/AVA inversion. In, approach II, the relative changes (by linearized inversion) or absolute values (by nonlinear waveform inversion) are obtained directly. I compare non-iterative linearized migration/inversion imaging, iterative linearized migration/inversion imaging, and non-linear waveform inversion. All of these imaging methods can be considered as back-projection and back-scattering imaging. From backscattering imaging, we know that seismic wave illumination has a key influence on so-called true-amplitude imaging, and I give an analysis for the possibility of relative true-amplitude imaging. I also analyze the factors that affect the imaging quality. Finally, I point out that the Born approximation is not a good approximation for linearized migration/inversion imaging, and that the De Wolf approximation is a better choice.

INTRODUCTION

The main objective of migration imaging is to generate an image of the reflectors, that is, to position reflection points and scattering points at their true subsurface locations. The methodology is to downward continuate the observed wavefield to the reflection points or scattering points using a known macro-velocity model with appropriate propagators, and to pick out the focused wavefield with an imaging condition. The focused wavefield displays the image of reflectors or scatterers. Therefore, I give the definition for migration imaging: **Based on some assumptions about the geological medium and with the help of mathematical models, the observed seismic wavefield is extrapolated to the subsurface reflectors using a macro-velocity model with a propagator, and the imaging amplitudes are extracted with an imaging condition**. Generally, the geological medium is assumed to be an acoustic medium, and the mathematical model is either the one-way wave equation or the Kirchhoff integral operator. However, migration imaging has not completely satisfied the needs of oil and gas exploration, since many reservoirs found recently are controlled not only by their

[1]**email:** wang@sep.stanford.edu

geological structures but also by their lithology. Therefore, the lithological parameters are increasingly important to oil and gas exploration. Lithological parameter estimation is typically an inverse problem. In essence, migration imaging is an inverse problem, and it is also ill-posed. However, migration imaging is changed into an apparently well-posed problem by splitting it into two processes: wavefield extrapolation and macro-velocity analysis. The main objective of inversion imaging is to estimate lithological parameters or their disturbances, including reflectivity, P-wave velocity, S-wave velocity, and the density. There are linearized and non-linear inversions. The basis of linearized inversion is to linearize the formula characterizing the scattering wavefield with the Born approximation. The Born approximation is a "physical" approximation, with which only the primaries are modeled. The analytical (for constant background) or formal (for variable background) inversion formulas can be derived from the linearized forward-modeling formulas. This is a non-iterative linearized inversion. Based on least-squares theory, an iterative linearized inversion approach can be derived from linearized forward modeling. For the non-linear waveform inversion, only the wave propagator is linearized at a point in the model space. With the propagator, all of the wave phenomena are characterized. We call this linearization as a "mathematical" approximation, with which both primaries and multiples are simulated. This is the main difference between the two inversion approaches. Theoretically, the non-linear inversion (Tarantola, 1984; Mora, 1987) is superior to the linearized inversion (Bleistein et al., 1987; Bleistein, 1987). In practice, it is very difficult to recover all wavenumber components of the lithological parameters, since the seismic data is frequency-band-limited and aperture-limited and polluted with non-Gaussian noise. Therefore, the linearized migration/inversion is becoming more and more important, especially the iterative linearized migration/inversion approach. Stolt and Weglein (1985) discussed the relation between the migration and the linearized inversion. Gray (1997) gave a comparison of three different examples of true-amplitude imaging. So-called true-amplitude imaging tries to recover the reflectivity of the reflectors.

In this paper, I compare non-iterative linearized migration/inversion imaging, iterative linearized migration/inversion imaging, and nonlinear waveform inversion. All of these imaging methods can be considered as back-projection and backscattering imaging. From backscattering imaging, we know that seismic wave illumination has a key influence on so-called true-amplitude imaging, and I give an analysis for the possibility of relative true-amplitude imaging. I also analyze the factors which affect the image quality. Finally, I point out that the Born approximation is not a good approximation for linearized migration/inversion imaging, and that the De Wolf approximation is a better choice.

WAVE PROPAGATOR AND ITS LINEARIZATION

[1] Acoustic wave equation

Based on inverse theory, the characterization of seismic wave propagation is important for parameter estimation. Here I use the acoustic wave equation with two elastic parameters — bulk modulus and density — to model seismic wave propagation in a geological medium,

though we know that this is a simplification.

$$L\,P \equiv \left(\nabla \cdot \frac{1}{\rho\,(\vec{x})}\nabla + \frac{\omega^2}{\kappa\,(\vec{x})}\right) P\,(\vec{x},\vec{x}_s\omega) = \delta\,(\vec{x}-\vec{x}_s)\,S\,(\omega), \tag{1}$$

where κ is the bulk modulus and ρ is the density. Both parameters vary horizontally as well as vertically. $P\,(\vec{x},\vec{x}_s,\omega)$ is the acoustic pressure wave field, and $S\,(\omega)$ is the monochromatic source function. We can carry out the full waveform inversion with equation (1). Tarantola (1984) gave a detailed theoretical framework. Pratt and Hicks (1998) discussed in detail how to implement seismic waveform inversion in the frequency domain. Now I introduce a background model which is so close to the true model that we can neglect the second and higher-order reflection and transmission effects caused by the interaction between the incident wave and the scattering potential. The background wavefield obeys the following equation:

$$L_0 P \equiv \left(\nabla \cdot \frac{1}{\rho_0}\nabla + \frac{\omega^2}{\kappa_0}\right) P\,(\vec{x},\vec{x}_s,\omega) = \delta\left(\vec{x}-\vec{x}^s\right) S\,(\omega). \tag{2}$$

With the definition $V = L - L_0$, the identity $A = B + B\left(B^{-1} - A^{-1}\right) A$ becomes

$$G = G_0 + G_0 V G, \tag{3}$$

if we associate G with A and G_0 with B. And equation (3) is further rearranged to

$$G = (I - G_0 V)^{-1} G_0. \tag{4}$$

Performing a Taylor expansion on the right term of equation (4) yields

$$G = \left[\sum_{j=0}^{\infty} (G_0 V)^j\right] G_0. \tag{5}$$

Equation (3) is called the Lippmann-Schwinger equation(Clayton and Stolt, 1981). Clearly, if $j \geq 2$, equation (5) depicts second and higher-order scattering terms of wave propagation, which are neglected. The linearized propagator characterizes only the first scattering of wave propagation. That is,

$$G = G_0 + G_0 V G_0. \tag{6}$$

This is the Born approximation, the physical meaning of which is clearly demonstrated by equations (5) and (6). From $L = -\left(\nabla \cdot \frac{1}{\rho}\nabla + \frac{\omega^2}{\kappa}\right)$ and $L_0 = -\left(\nabla \cdot \frac{1}{\rho_0}\nabla + \frac{\omega^2}{\kappa_0}\right)$, the scattering potential V is defined as follows:

$$\begin{aligned} V &= \left(\nabla \cdot \frac{1}{\rho}\nabla + \frac{\omega^2}{\kappa}\right) - \left(\nabla \cdot \frac{1}{\rho_0}\nabla + \frac{\omega^2}{\kappa_0}\right) \\ &= \nabla \cdot \left(\frac{1}{\rho} - \frac{1}{\rho_0}\right)\nabla + \omega^2\left(\frac{1}{\kappa} - \frac{1}{\kappa_0}\right) \\ &= \nabla \cdot \frac{1}{\rho_0}\left(\frac{\rho_0}{\rho} - 1\right)\nabla + \omega^2 \frac{1}{\kappa_0}\left(\frac{\kappa_0}{\kappa} - 1\right) \\ &= \nabla \cdot \frac{a_1}{\rho_0}\nabla + \omega^2 \frac{a_2}{\kappa_0} \end{aligned} \tag{7}$$

where $a_1 = \frac{\rho_0}{\rho} - 1 = \frac{\triangle\rho}{\rho}$ and $\triangle\rho = \rho_0 - \rho$; $a_2 = \frac{\kappa_0}{\kappa} - 1 = \frac{\triangle\kappa_0}{\kappa}$ and $\triangle\kappa = \kappa_0 - \kappa$. Therefore the linearized synthetic wave field is composed of two parts: one is the background wave field described by the background Green's function; the other is the scattering wavefield caused by the scattering potential V. According to equation (3), the total wave field is written as

$$P(\vec{x}_r, \vec{x}_s, \omega) = G_0(\vec{x}_r, \vec{x}_s, \omega) + \omega^2 \int d^3x G_0(\vec{x}_r, \vec{x}, \omega) V(\vec{x}) P(\vec{x}, \vec{x}_s, \omega), \tag{8}$$

and the scattering wavefield after the Born approximation from equation (6) is

$$P_s(\vec{x}_r, \vec{x}_s, \omega) = \omega^2 \int d^3x G_0(\vec{x}_r, \vec{x}, \omega) V(\vec{x}) G_0(\vec{x}, \vec{x}_s, \omega). \tag{9}$$

[2] Scalar wave equation

In seismic wave imaging, the scalar wave equation is much more commonly used. Given the Fourier transform of the scalar wave equation for a point source,

$$L P(\vec{x}, \vec{x}_s, \omega) = \left(\nabla^2 + \frac{\omega^2}{v^2(\vec{x})}\right) P(\vec{x}, \vec{x}_s, \omega) = -\delta(\vec{x} - \vec{x}_s) S(\omega). \tag{10}$$

Equation (10) is a Helmholtz equation. Here $P(\vec{x}, \vec{x}_s, \omega)$ is a total pressure field, $v(\vec{x})$ is the variable acoustic velocity, the density is assumed to be constant, and $S(\omega)$ is a source function. Taking an initial estimation of the medium velocity, $v(\vec{x})$, as the background velocity $v_0(\vec{x})$, $v(\vec{x})$ can be split into the known and unknown parts by the following:

$$\frac{1}{v^2(\vec{x})} = \frac{1}{v_0^2(\vec{x})}(1 + a(\vec{x})), \tag{11}$$

where the background velocity $v_0(\vec{x})$ need not be constant. The variable $a(\vec{x})$ is the unknown velocity perturbation to be determined from the data, which is called the scattering potential of the medium, since it is a measure of the scattering strength at points where the actual medium differs from the background medium. Substituting equation (11) into equation (10) yields

$$L P(\vec{x}, \vec{x}_s, \omega) = \left(\nabla^2 + \frac{\omega^2}{v_0^2(\vec{x})}\right) P(\vec{x}, \vec{x}_s, \omega) = -\delta(\vec{x} - \vec{x}_s) S(\omega) - \omega^2 a(\vec{x}) P(\vec{x}, \vec{x}_s, \omega). \tag{12}$$

For the background medium, the background Green's function satisfies the following equation:

$$L_0 G_0(\vec{x}, \vec{x}_s, \omega) = \left(\nabla^2 + \frac{\omega^2}{v_0^2(\vec{x})}\right) G_0(\vec{x}, \vec{x}_s, \omega) = -\delta(\vec{x} - \vec{x}_s). \tag{13}$$

Therefore, with the help of the Lippman-Schwinger equation of equation (3), the total pressure wave is

$$P(\vec{x}_r, \vec{x}_s, \omega) = G_0(\vec{x}_r, \vec{x}_s, \omega) + \omega^2 \int d^3x G_0(\vec{x}_r, \vec{x}, \omega) a(\vec{x}) P(\vec{x}, \vec{x}_s, \omega), \tag{14}$$

and the scattering wave field caused by the scattering potential $a(\vec{x})$ is

$$P_s(\vec{x}_r,\vec{x}_s,\omega) = \omega^2 \int d^3x G_0(\vec{x}_r,\vec{x},\omega) a(\vec{x}) G_0(\vec{x},\vec{x}_s,\omega). \tag{15}$$

Equations (9) and (15) create a link between the scattered wavefield and the scattering potentials. They are Fredholm integral equations of the first kind and are the bases of the linearized inversion. Given the observed scattered wavefield, the scattering potentials can be solved.

[3] The scattering potential and the reflectivity

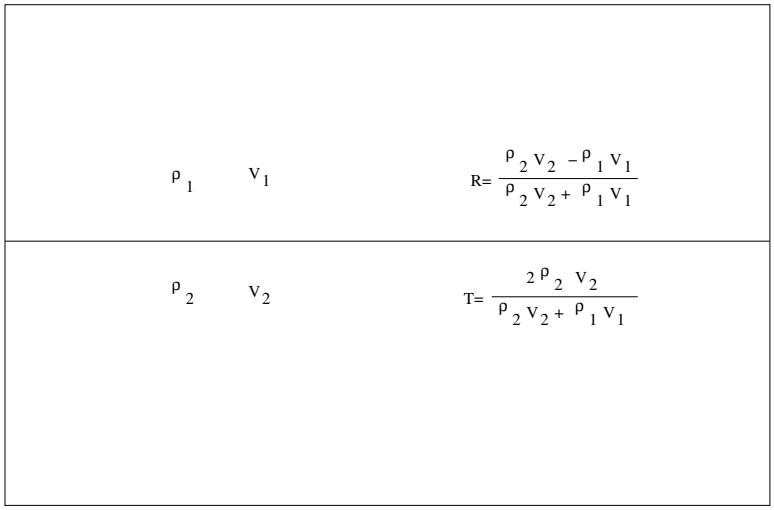

Figure 1: The acoustic wave reflectivity and the transmission of a planar reflector in the case of zero incident angle huazhong1-reflectivity [CR]

From Fig.1, and assuming that the density is constant, the normal reflectivity is defined as

$$R|_{\theta=0} = \frac{v_2 - v_1}{v_2 + v_1}, \tag{16}$$

and the transmission coefficient is

$$T|_{\theta=0} = \frac{2v_2}{v_2 + v_1}, \tag{17}$$

where θ is the incident angle. Therefore, defining the scattering potential as $1 + a(\vec{x}) = \left(\frac{v_0(\vec{x})}{v(\vec{x})}\right)^2$, equation (16) and (17) can be rewritten as

$$R|_{\theta=0} = \frac{1-\sqrt{1+a}}{1+\sqrt{1+a}}, \tag{18}$$

and

$$T|_{\theta=0} = \frac{2}{1+\sqrt{1+a}}, \tag{19}$$

respectively. If waves meet a reflector with a non-zero incident angle, the reflectivity and transmission coefficient are

$$R(\theta_1) = \frac{v_2 cos\theta_1 - v_1 cos\theta_2}{v^2 cos\theta_1 + v_1 cos\theta_2}, \tag{20}$$

and

$$T(\theta_1) = \frac{2v_2 cos\theta_1}{v_2 cos\theta_1 + v_1 cos\theta_2}. \tag{21}$$

Similarly, they can be expressed with the scattering potential as

$$R(\theta_1) = \frac{cos\theta_1 - \sqrt{1+a}cos\theta_2}{cos\theta_1 + \sqrt{1+a}cos\theta_2}, \tag{22}$$

and

$$T(\theta_1) = \frac{2cos\theta_1}{cos\theta_1 + \sqrt{1+a}cos\theta_2}, \tag{23}$$

respectively. The angle reflectivity has a close relation to the scattering potentials. Generally, the angle reflectivity is estimated by amplitude-preserved imaging, and lithological parameter disturbances are evaluated from them by AVO/AVA inversion.

ITERATIVE INVERSION IMAGING ALGORITHMS

[1] Operator linearizeation in non-linear inversion

A linear operator L depicts a physics process. It can be written as

$$L\mathbf{m} = \mathbf{d}, \tag{24}$$

where $\mathbf{d}$ is the synthetic wavefield and $\mathbf{m}$ is a medium model. The linear operator L can be seen as a function, which can be expanded into a Taylor series near a known model $\mathbf{m}_0$ as follows:

$$L\mathbf{m} = L\mathbf{m}_0 + \frac{\partial L}{\partial \mathbf{m}}\triangle\mathbf{m} + \frac{\partial^2 L}{\partial \mathbf{m}^2}(\triangle\mathbf{m})^2 + \cdots + \frac{\partial^n L}{\partial \mathbf{m}^n}(\triangle\mathbf{m})^n + \cdots. \tag{25}$$

Omitting all the terms that are higher than second-order yields a linearized equation:

$$\frac{\partial L}{\partial \mathbf{m}}\triangle\mathbf{m} \approx L\mathbf{m} - L\mathbf{m}_0. \tag{26}$$

If $L\mathbf{m}$ stands for the observed data, and if $L\mathbf{m}_0$ synthesizes a wavefield with a known background model and a given operator, equation (26) can be rewritten as

$$\frac{\partial L}{\partial \mathbf{m}}\triangle\mathbf{m} \approx \mathbf{d}^{obs} - \mathbf{d}^{cal}. \tag{27}$$

Equation (27) can be regarded as a matrix equation, which may be ill-conditioned. The model disturbance can be solved by many linear algebraic algorithms. If the background model $\mathbf{m}_0$ is very close to the true model, the true model can be approached by some iterative algorithms. This idea is meaningful, but impractical. In fact, equation (27) can be simplified to

$$A\delta\mathbf{m} = \delta\mathbf{d}, \tag{28}$$

where $A = \partial L/\partial\mathbf{m}$ and $\delta\mathbf{d} = \mathbf{d}^{obs} - \mathbf{d}^{cal}$. Least-squares methods are then used to solve the inverse problem. Equations (9) and (15) can also be expressed in the form of equation (28). Comparing equation (28) with equations (9) and (15), it is clearly seen that the main difference between linearized inversion and non-linear waveform inversion consists in the forward modeling operator. The operator after Born approximation models only the primaries; however, the Frechèt derivative $A = \partial L/\partial\mathbf{m}$ models all the wave phenomena. The Born approximation should be replaced by the De Wolf approximation. The non-linear inversion incurs much higher calculation costs.

[2] L_2 norm or cost function definition

Based on least-squares theory, a minimizing problem can be defined, which aims to find δm^* in order to minimize the cost function. The L_2 norm or cost function definition is given by

$$f(\delta\mathbf{m}) = \|A\delta\mathbf{m} - \delta\mathbf{d}\|_2^2. \tag{29}$$

In order to constrain the inverse problem, or to use some prior information to bound the solution of the inverse problem, regularization is commonly used. In this case, the cost function needs to be modified.

[3] Iterative algorithms

Many algorithms can be chosen to solve the minimizing problem which is determined by the property of matrix A. The matrix A is hoped to be positive definite. Here only Newton's iterative algorithms are listed. **[A] Initial Newton's approach** Performing a Taylor expansion of $f(\delta\mathbf{m})$ near the point $\delta\mathbf{m}^{(k)}$ yields

$$\begin{aligned} f(\delta\mathbf{m}) &\approx \phi(\delta\mathbf{m}) \\ &= f\left(\delta\mathbf{m}^{(k)}\right) + \nabla f\left(\delta\mathbf{m}^{(k)}\right)\left(\delta\mathbf{m} - \delta\mathbf{m}^{(k)}\right) \\ &+ \frac{1}{2}\left(\delta\mathbf{m} - \delta\mathbf{m}^{(k)}\right)^T \nabla^2 f\left(\delta\mathbf{m}^{(k)}\right)\left(\delta\mathbf{m} - \delta\mathbf{m}^{(k)}\right). \end{aligned} \tag{30}$$

Letting $\partial\phi(\delta\mathbf{m})/\partial\delta\mathbf{m} = 0$ yields $\nabla f\left(\delta\mathbf{m}^{(k)}\right) + \nabla^2 f\left(\delta\mathbf{m}^{(k)}\right)\left(\delta\mathbf{m} - \delta\mathbf{m}^{(k)}\right) = 0$. If the Hessian $\nabla^2 f\left(\delta\mathbf{m}^{(k)}\right)$ is invertible, the Newton iterative algorithm is

$$\delta\mathbf{m}^{(k+1)} = \delta\mathbf{m}^{(k)} - \left[\nabla^2 f\left(\delta\mathbf{m}^{(k)}\right)\right]^{-1} \nabla f\left(\delta\mathbf{m}^{(k)}\right). \tag{31}$$

Clearly, the simple Newton iterative algorithm lacks 1D searching. The Newton iterative algorithm with 1D searching is called the damping Newton algorithm. Algorithm procedure:

(a) Assign the initial model $\delta\mathbf{m}^{(1)}$ and the acceptable error $\varepsilon < 0$ and set the iterative number $k = 1$.

(b) Calculate $\nabla f\left(\delta\mathbf{m}^{(k)}\right)$ and $\left[\nabla^2 f\left(\delta\mathbf{m}^{(k)}\right)\right]^{-1}$.

(c) If $\|\nabla f\left(\delta\mathbf{m}^{(k)}\right)\| \le \varepsilon$, then stop iteration; otherwise, $\mathbf{d}^{(k)} = -\left[\nabla^2 f\left(\delta\mathbf{m}^{(k)}\right)\right]^{-1}\nabla f\left(\delta\mathbf{m}^{(k)}\right)$.

(d) Starting from $\delta\mathbf{m}^{(k)}$, carry out 1D searching along the searching direction $\mathbf{d}^{(k)}$ for $\lambda^{(k)}$ satisfying $f\left[\left(\delta\mathbf{m}^{(k)}\right)+\lambda_k\mathbf{d}^{(k)}\right] = min\left\{f\left[\left(\delta\mathbf{m}^{(k)}\right)+\lambda\mathbf{d}^{(k)}\right]\right\}_{\lambda\ge 0}$.

(e) Letting $\delta\mathbf{m}^{(k+1)} = \delta\mathbf{m}^{(k)}+\lambda_k\mathbf{d}^{(k)}$ and $k := k+1$, go to step (b). If the Hessian $\nabla^2 f\left(\delta\mathbf{m}^{(k)}\right)$ is not positive definite, the Newton algorithm should be modified further. That means $\nabla^2 f\left(\delta\mathbf{m}^{(k)}\right)$ is replaced with $\nabla^2 f\left(\delta\mathbf{m}^{(k)}\right)+\varepsilon_k I$. If the ε_k is chosen suitably, the matrix $\nabla^2 f\left(\delta\mathbf{m}^{(k)}\right)+\varepsilon_k I$ will be positive definite. **[B] Quasi-Newton algorithm:** The main feature of this algorithm is that the inverse of the Hessian matrix is not explicitly calculated. Further implementing the differential operation on both sides of equation (30) yields

$$\nabla f\left(\delta\mathbf{m}^{(k)}\right) \approx \nabla f\left(\delta\mathbf{m}^{(k+1)}\right)+\nabla^2 f\left(\delta\mathbf{m}^{(k+1)}\right)\left(\delta\mathbf{m}^{k}-\delta\mathbf{m}^{(k+1)}\right), \tag{32}$$

and defining $\mathbf{p}^{(k)} = \delta\mathbf{m}^{k}-\delta\mathbf{m}^{(k+1)}$ and $\mathbf{q}^{(k)} = \nabla f\left(\delta\mathbf{m}^{(k+1)}\right)-\nabla f\left(\delta\mathbf{m}^{(k)}\right)$ yields the Quasi-Newton condition,

$$\mathbf{p}^{(k)} = H_{k+1}\mathbf{q}^{(k)}, \tag{33}$$

where $H_{k+1} = \left[\nabla^2 f\left(\delta\mathbf{m}^{(k+1)}\right)\right]^{-1}$. A series of formulas for calculating H_{k+1} is listed below: Formula 1: $H_{k+1} = H_k + \frac{\left(\mathbf{p}^{(k)}-H_k\mathbf{q}^{(k)}\right)\left(\mathbf{p}^{(k)}-H_k\mathbf{q}^{(k)}\right)^T}{\left(\mathbf{q}^{(k)}\right)^T\left(\mathbf{p}^{(k)}-H_k\mathbf{q}^{(k)}\right)}$. Formula 2: $H_{k+1}^{DFP} = H_k + \frac{\mathbf{p}^{(k)}\left(\mathbf{p}^{(k)}\right)^T}{\left(\mathbf{p}^{(k)}\right)^T\mathbf{q}^{(k)}} - \frac{\left(H_k\mathbf{q}^{(k)}\right)\left(H_k\mathbf{q}^{(k)}\right)^T}{\left(\mathbf{q}^{k}\right)^T H_k\mathbf{q}^{(k)}}$, which is the DFP (Davidon-Fletcher-Powell) algorithm. Formula 3: $H_{k+1}^{BFGS} = H_k + \left(1+\frac{\left(\mathbf{q}^{(k)}\right)^T H_k\mathbf{q}^{(k)}}{\left(\mathbf{p}^{(k)}\right)^T\mathbf{q}^{(k)}}\right)\frac{\mathbf{p}^{(k)}\left(\mathbf{p}^{(k)}\right)^T}{\left(\mathbf{p}^{(k)}\right)^T\mathbf{q}^{(k)}} - \frac{\mathbf{p}^{(k)}\left(\mathbf{q}^{(k)}\right)^T H_k + H_k\mathbf{q}^{(k)}\left(\mathbf{p}^{(k)}\right)^T}{\left(\mathbf{p}^{k}\right)^T q^{(k)}}$, which is the BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm. Formula 4: $H_{k+1}^{\phi} = (1-\phi)H_{k+1}^{DFP}+\phi H_{k+1}^{BFGS}$, where ϕ is an parameter. The algorithm procedure is as follows:

(a) Assign the initial model $\delta\mathbf{m}^{(1)}$ and the acceptable error $\varepsilon < 0$.

(b) Setting $H_1 = I_n$ and the iterative number $k = 1$, calculate $\mathbf{g}_1 = \nabla f\left(\delta\mathbf{m}^{(1)}\right)$.

(c) Let $\mathbf{d}^{(k)} = -H_k\mathbf{g}_k$.

(d) Starting from $\delta\mathbf{m}^{(k)}$, carry out 1D searching along the searching direction $\mathbf{d}^{(k)}$ for $\lambda^{(k)}$ satisfying $f\left[\left(\delta\mathbf{m}^{(k)}\right)+\lambda_k\mathbf{d}^{(k)}\right] = min\left\{f\left[\left(\delta\mathbf{m}^{(k)}\right)+\lambda\mathbf{d}^{(k)}\right]\right\}_{\lambda\ge 0}$.

(e) If $\|\nabla f\left(\delta\ \mathbf{m}^{(k)}\right)\| \le \varepsilon$, then stop iteration; otherwise, go to Step (f).

(f) If $k = n$, then let $\|\delta\mathbf{m}^{(1)} = \delta\mathbf{m}^{(k+1)}\|$, go to Step(b); Otherwise, go to Step (g).

(g) Letting $\mathbf{g}_{k+1} = \nabla f\left(\delta\mathbf{m}^{(k+1)}\right)$, $\mathbf{p}^{(k)} = \delta\mathbf{m}^{(k)}-\delta\mathbf{m}^{(k+1)}$ and $\mathbf{q}^{(k)} = \nabla f\left(\delta\mathbf{m}^{(k+1)}\right)-\nabla f\left(\delta\mathbf{m}^{(k)}\right)$, calculate H_{k+1} with any of Formula 1-4. Setting $k := k+1$, go to Step (c).

COMPARISON AMONG MIGRATION/INVERSION METHODS

(A) Non-iterative linearized migration/inversion

(1) Wave equation prestack migration/inversion

Near the scattering point $\vec{x}$, we can define an error function or a norm as

$$E(R(\vec{x})) = \sum_{\omega_{min}}^{\omega_{max}} \left(U_S(\vec{x},\omega) - U_I(\vec{x},\omega)\, R(\vec{x})\right)^2 d\omega, \tag{34}$$

where $R(\vec{x})$ is the reflectivity, $U_S(\vec{x},\omega)$ is the upcoming wavefield, which is downward extrapolated to a reflector, and $U_I(\vec{x},\omega)$ is the incident wavefield propagated to the reflector. At the scattering point $\vec{x}$, the scattering wavefield $U_S(\vec{x},\omega)$ should be equal or close to the convolution between the incident wavefield $U_I(\vec{x},\omega)$ and the reflectivity function. From equation (34), the imaging condition of the migration/inversion is as follows:

$$R(\vec{x}) = \frac{\sum_{\omega_{min}}^{\omega_{max}} U_S(\vec{x},\omega)\, U_I^*(\vec{x},\omega)}{\sum_{\omega_{min}}^{\omega_{max}} \left(U_I(\vec{x},\omega)\, U_I^*(\vec{x},\omega) + \varepsilon\right)}. \tag{35}$$

The term in the numerator is a correlation imaging for prestack migration. The term in the denominator expresses the illumination of the scattering points. Fig.2 geometrically explains the imaging condition, which says that the imaging occurs at the arrival time of the incident wave which equals the take-off time of the upcoming wave. In the frequency domain, the

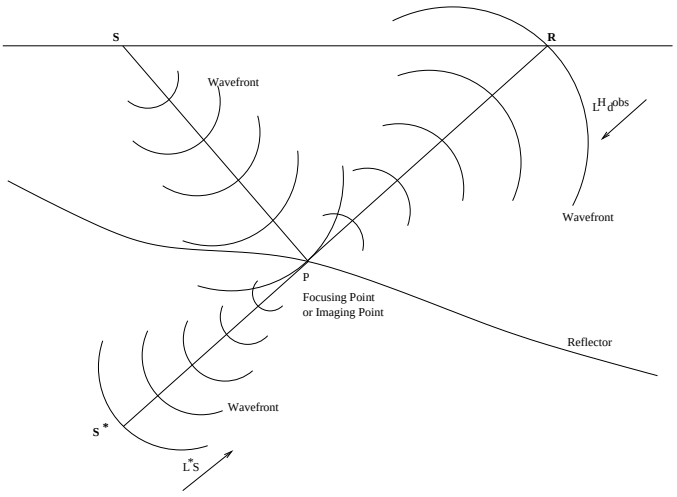

Figure 2: The geometry explication of the cross-correlation imaging condition. S^* is a virtual source of the real source S. The propagator L^* is the conjugate of the downward propagator L. Therefore, both the propagator L^H and L^* collapse the wavefronts into a point—the imaging point P. huazhong1-Imaging_fig [CR]

extrapolated upcoming wavefield at the scattering point is

$$U_S(\vec{x},\omega) = G^H(\vec{x}_r,\vec{x},\omega)\, \mathbf{d}^{obs}(\vec{x}_r,\vec{x}_s,\omega), \tag{36}$$

and the incident wavefield at the same point is

$$U_I(\vec{x},\omega) = G(\vec{x},\vec{x}_s,\omega). \tag{37}$$

Substituting equations (36) and (37) into equation (35) and applying the WKBJ approximation to the Green's functions, we can rewrite equation (35) as follows:

$$\begin{aligned} R(\vec{x}) &= \frac{\sum_{\omega_{min}}^{\omega_{max}} G(\vec{x},\vec{x}_s,\omega) G^H(\vec{x}_r,\vec{x},\omega)\mathbf{d}^{obs}(\vec{x}_r,\vec{x}_s,\omega)}{\sum_{\omega_{min}}^{\omega_{max}} \left[G(\vec{x},\vec{x}_s,\omega)(G(\vec{x},\vec{x}_s,\omega))^* + \varepsilon\right]} \\ &= \frac{\sum_{\omega_{min}}^{\omega_{max}} A(\vec{x},\vec{x}_s,\omega) e^{i\omega\tau(\vec{x},\vec{x}_s,\omega)} A(\vec{x}_r,\vec{x},\omega) e^{i\omega\tau(\vec{x}_r,\vec{x},\omega)}\mathbf{d}^{obs}(\vec{x}_r,\vec{x}_s,\omega)}{\sum_{\omega_{min}}^{\omega_{max}} |A(\vec{x},\vec{x}_s,\omega)|^2} \\ &= \frac{\sum_{\omega_{min}}^{\omega_{max}} A(\vec{x}_r,\vec{x},\vec{x}_s,\omega) e^{i\omega\tau(\vec{x}_r,\vec{x},\vec{x}_s,\omega)}\mathbf{d}^{obs}(\vec{x}_r,\vec{x}_s,\omega)}{\sum_{\omega_{min}}^{\omega_{max}} |A(\vec{x},\vec{x}_s,\omega)|^2}, \end{aligned} \tag{38}$$

where $A(\vec{x}_r,\vec{x},\vec{x}_s,\omega) = A(\vec{x},\vec{x}_s,\omega)\,A(\vec{x}_r,\vec{x},\omega)$ and $\tau(\vec{x}_r,\vec{x},\vec{x}_s,\omega) = \tau(\vec{x},\vec{x}_s,\omega) + \tau(\vec{x}_r,\vec{x},\omega)$. From equation (38), it is clear that the seismic illumination plays a key role in migration/inversion imaging. The possibility of relative true-amplitude imaging will be discussed later.

(2) Wave theory tomography

(a) Fourier Diffraction Tomography for constant background Wu and Toksoz (1987) gave the plane-wave response in the direction $\vec{r}$ from an incident wave $\vec{i}$:

$$P_S^{pl}\left(\vec{i},\vec{r}\right) = -k^2 \tilde{O}\left[k\left(\vec{r}-\vec{i}\right)\right] \tag{39}$$

where $\tilde{O}\left[k\left(\vec{r}-\vec{i}\right)\right]$ is the 3D Fourier transform of the object function $O(\vec{r})$. $P_S^{pl}\left(\vec{i},\vec{r}\right)$ is some kind of projection. Comparing this to linear Radon transform, we know that the object function can be accurately restored if the angles of the plane waves continuously change around the object. Fig.3 shows the projection from the real plane wave source and from the virtual plane wave source.

(b) Inverse Generalized Radon Transform for variable background The scattered wavefield after Bron and WKBJ approximation is of the following form:

$$\begin{aligned} P_S(\vec{r},\vec{s},t) &= -\frac{\partial^2}{\partial t^2}\int_\Omega A(\vec{r},\vec{x},\vec{s})\,\delta\left[t-\tau(\vec{r},\vec{x},\vec{s})\right] f(\vec{x})\,d^3\vec{x} \\ &= -\int_\Omega A(\vec{r},\vec{x},\vec{s})\,\delta^{"}\left[t-\tau(\vec{r},\vec{x},\vec{s})\right] f(\vec{x})\,d^3\vec{x}, \end{aligned} \tag{40}$$

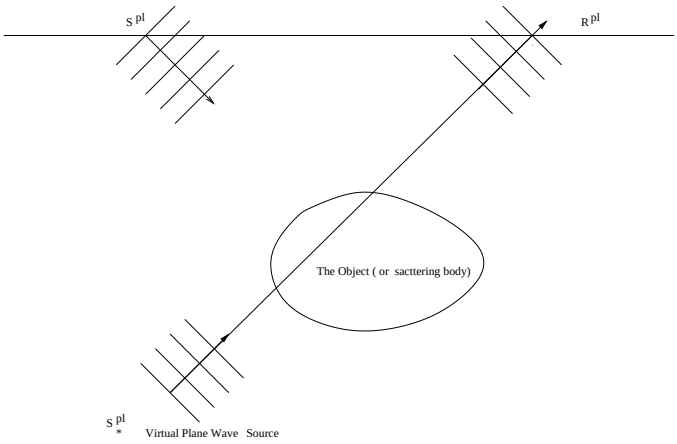

Figure 3: The geometry of plane wave propagation. huazhong1-planewave_tomography [CR]

where $A(\vec{r},\vec{x},\vec{s}) = A(\vec{r},\vec{x})\,A(\vec{x},\vec{s})$ and $\tau(\vec{r},\vec{x},\vec{s}) = T(\vec{s},\vec{x}) + T(\vec{r},\vec{x})$. It is known that the diffraction-time surface $R_x = \{\mathbf{d} : t = T(\vec{s},\vec{x}) + T(\vec{r},\vec{x})\}$ in the data space is a counterpart of the isochron surface $I_d = \{\mathbf{x} : t = T(\vec{s},\vec{x}) + T(\vec{r},\vec{x})\}$ in the model space. These dual geometric associations naturally give rise to a corresponding pair of projection operators. Equation (40) can be written as

$$P_S(\mathbf{d}) = -\frac{\partial^2}{\partial t^2}\int_{I_d} A(\vec{r},\vec{x},\vec{s})\, f(\mathbf{x}). \tag{41}$$

The diffraction curve in the data space is a projection of an isochron in the model space. This is a kind of Radon transform (Miller et al., 1987; Hubral et al., 1996). The standard Radon transform and inverse Radon transform in three dimensions are given by

$$f^{\triangle}\left(\vec{\xi},p\right) = \int \delta\left(p - \vec{\xi}\cdot\vec{x}\right) f(\vec{x})\, d^3\vec{x} \tag{42}$$

and

$$f(\vec{x}_0) = -\frac{1}{8\pi^2}\int\left[\frac{\partial^2}{\partial p^2} f^{\triangle}\left(\vec{\xi},p\right)|_{p=\vec{\xi}\cdot\vec{x}_0}\right] d^2\xi \tag{43}$$

respectively, where p is the distance from the origin to a plane which cuts through the object body, ξ is the unity direction vector which is normal to the plane, and $\vec{x}$ is a point on the plane. Comparing this with the classical Radon transform and its inverse, the final 3D inversion formula can be given as

$$f(\vec{x}_0) = \frac{1}{\pi^2}\int d^2\xi\,(\vec{r},\vec{x}_0,\vec{s})\,\frac{|cos^3\alpha(\vec{r},\vec{x}_0,\vec{s})|}{v_0^3(\vec{x}_0)\,A(\vec{r},\vec{x}_0,\vec{s})} P_S(\vec{r},\vec{s},t=\tau_0). \tag{44}$$

In equation (44), the angle variable $\xi(\vec{r},\vec{x}_0,\vec{s})$ near the imaging point $\vec{x}_0$ is used, rather than the measurement configuration at the surface. Fig.4 illustrates this. The angle variable is related to the measurement configuration and reflects the seismic wave illumination aperture. Only if the aperture is large can a high resolution image be obtained. The relative true-amplitude imaging is severely affected by the angle variable. Bleistein and Stockwell (2001); Zhang (2004) gave some similar true-amplitude migration/inversion formulas.

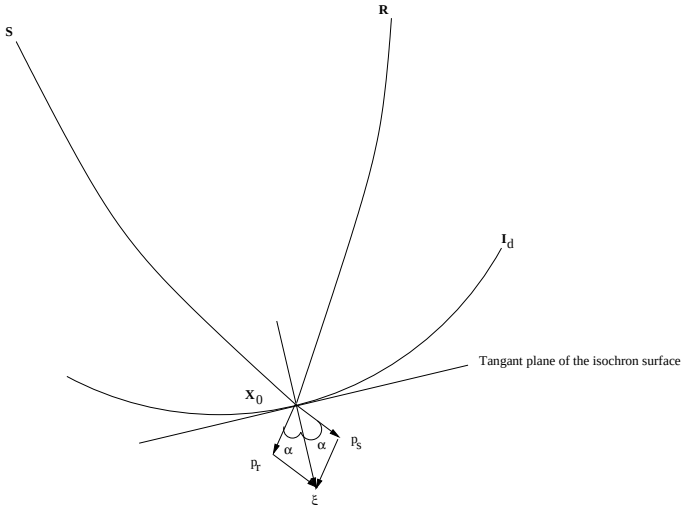

Figure 4: The geometry between the incident and scattering rays near the scattering point, or imaging point $\mathbf{x}_0$. huazhong1-GRT_fig [CR]

(3) Least-squares migration/inversion

The conventional prestack migration can be characterized as

$$\delta\mathbf{m}_{mig} = A^H \delta\mathbf{d}, \tag{45}$$

where A^H is a conjugate transpose matrix, which is a back-propagator of the wavefield. The least-squares prestack migration/inversion imaging can be carried out by the following equation:

$$\delta\mathbf{m}_{inv} = \left(A^H A\right)^{-1}\left(A^H \delta\mathbf{d}\right) = H^{-1}\left(A^H \delta\mathbf{d}\right) = H^{-1}\delta\mathbf{m}_{mig}, \tag{46}$$

where $H = A^H A$ is a Hessian matrix. The meaning of the Hessian will be discussed in detail. Equation (46) says that the deconvolution of the conventional prestack migration by the inverse of the Hessian produces the migration/inversion results.

(B) Iterative linearized migration/inversion

We can define a minimizing problem that aims at finding δm^* by minimizing the following cost function:

$$f\left(\delta\mathbf{m}\right) = \| A\delta\mathbf{m} - \delta\mathbf{d}\|_2^2. \tag{47}$$

The Newton iterative algorithms can be used for solving the minimizing problem. The standard Newton iterative algorithm is

$$\delta\mathbf{m}^{(k+1)} = \delta\mathbf{m}^{(k)} - \left[\nabla^2 f\left(\delta\mathbf{m}^{(k)}\right)\right]^{-1} \nabla f\left(\delta\mathbf{m}^{(k)}\right). \tag{48}$$

However, the inverse of the Hessian is difficult to calculate. The Quasi-Newton algorithms are used commonly. The inverse of the Hessian matrix can be calculated with the DFP formula:

$$H_{k+1}^{DFP} = H_k + \frac{\mathbf{p}^{(k)}\left(\mathbf{p}^{(k)}\right)^T}{\left(\mathbf{p}^{(k)}\right)^T \mathbf{q}^{(k)}} - \frac{\left(H_k\mathbf{q}^{(k)}\right)\left(H_k\mathbf{q}^{(k)}\right)^T}{\left(\mathbf{q}^{k}\right)^T H_k\mathbf{q}^{(k)}} \tag{49}$$

where $\mathbf{p}^{(k)} = \delta\mathbf{m}^k - \delta\mathbf{m}^{(k+1)}$, and $\mathbf{q}^{(k)} = \nabla f\left(\delta\mathbf{m}^{(k+1)}\right) - \nabla f\left(\delta\mathbf{m}^{(k)}\right)$.

The Quasi-Newton iterative algorithm is

$$\delta\mathbf{m}^{(k+1)} = \delta\mathbf{m}^{(k)} - H_{k+1}\nabla f\left(\delta\mathbf{m}^{(k)}\right). \tag{50}$$

(C) Non-linear waveform inversion

The basic procedure for non-linear waveform inversion is similar to the linearized iterative migration/inversion. It is worth mentioning again that the Frechèt derivative in the non-linear waveform inversion is quite different from the propagator after the Born approximation in the linearized migration/inversion.

ASPECTS OF LINEARIZED MIGRATION/INVERSION

[1] Numerical calculation of Green's function

The forward and backward propagations of the wavefield play a key role in the migration/inversion imaging. The Helmholtz equation is commonly used for depicting wave propagation in acoustic media. Based on it, I derive the formulas for traveltime and amplitude calculation corresponding to the main seismic wave energy. In spherical coordinates, the Helmholtz equation is of the following form:

$$\begin{aligned}\nabla^2\tilde{U}(\theta,\varphi,r;\omega) &= \left\{\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{1}{r^2}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\}\tilde{U}(\theta,\varphi,r;\omega)\\ &= -\frac{\omega^2}{v^2}\tilde{U}(\theta,\varphi,r;\omega).\end{aligned} \tag{51}$$

Equation (51) can be rewritten as

$$\left(\frac{\partial}{\partial r}+\frac{1}{r}\right)^2\tilde{U}+\left[\frac{1}{r^2\sin^2\theta}\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]\tilde{U}=\left(-\frac{\omega^2}{v^2}+\frac{1}{r^2}\right)\tilde{U}(\theta,\varphi,r;\omega), \tag{52}$$

The outward one-way wave equation can be derived from equation (52) as follows:

$$\left(\frac{\partial}{\partial r}+\frac{1}{r}\right)\tilde{U}(\theta,\varphi,r;\omega)=i\sqrt{\alpha}\sqrt{1+\frac{1}{r^2\sin^2\theta}\left[\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]}\tilde{U}(\theta,\varphi,r;\omega), \tag{53}$$

where α is defined as $\alpha = -\frac{\omega^2}{v^2}+\frac{1}{r^2}$. Equation (53) can be expanded as

$$\left(\frac{\partial}{\partial r}+\frac{1}{r}\right)\tilde{U}(\theta,\varphi,r;\omega)=i\sqrt{\alpha}\tilde{U}(\theta,\varphi,r;\omega)+\frac{\frac{ia}{\sqrt{\alpha}r^2\sin^2\theta}\left[\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{\partial^2}{\partial\varphi^2}\right]}{1+\frac{b}{\alpha r^2\sin^2\theta}\left[\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{\partial^2}{\partial\varphi^2}\right]}\tilde{U}(\theta,\varphi,r;\omega), \tag{54}$$

where a and b are the optimal coefficients. Then, equation (54) is split into two equations:

$$\frac{\partial}{\partial r}\tilde{U}(\theta,\varphi,r;\omega)=\left(-\frac{1}{r}+i\sqrt{\alpha}\right)\tilde{U}(\theta,\varphi,r;\omega), \tag{55}$$

$$\begin{aligned}&\left\{1+\frac{b}{\alpha r^2\sin^2\theta}\left[\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{\partial^2}{\partial\varphi^2}\right]\right\}\frac{\partial}{\partial r}\tilde{U}(\theta,\varphi,r;\omega)\\&=\frac{ia}{\sqrt{\alpha}r^2\sin^2\theta}\left[\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{\partial^2}{\partial\varphi^2}\right]\tilde{U}(\theta,\varphi,r;\omega).\end{aligned} \tag{56}$$

Obviously, equation (55) can be solved analytically, and equation (56) can be solved by finite-differences. The finite-difference scheme can be written in the following form:

$$[I-(\alpha_\theta-i\beta_\theta)T_\theta]\left[I-\left(\alpha_\varphi-i\beta_\varphi\right)T_\theta\right]\tilde{U}^n_{i,j}=[I-(\alpha_\theta+i\beta_\theta)T_\theta]\left[I-\left(\alpha_\varphi+i\beta_\varphi\right)T_\theta\right]\tilde{U}^n_{i,j}, \tag{57}$$

where $\tilde{U}^n_{i,j}=U(i\triangle\theta,j\triangle\varphi,n\triangle r)$, $\alpha_\theta=\frac{b}{r^2\alpha\triangle\theta^2}$, $\alpha_\varphi=\frac{b}{r^2\alpha\triangle\varphi^2}$, $\beta_\theta=\frac{a\triangle r}{2\sqrt{\alpha}r^2\triangle\theta^2}$ and $\beta_\varphi=\frac{a\triangle r}{2\sqrt{\alpha}r^2\triangle\varphi^2}$. The one-way wave extrapolation in the spherical coordinate system can be implemented by solving equation (55) and (57) in the frequency-space domain. The traveltime and amplitude corresponding to the maximum energy can be picked out in the frequency domain or time domain. With the picked traveltime and amplitude, we carried out a 3D integral prestack depth migration which gave a high-quality imaging result (Huazhong, 2003). This demonstrates that the method can be used for constructing the Green's functions in the migration/inversion imaging.

[2] The matrix expression of linearized migration/inversion

The linearized migration/inversion can be formulated from the integral expressed in equation (15). It can be regarded as an inverse generalized Radon transform. Equation (15) can be expressed as a matrix equation. The process for solving the equation set is just the migration/inversion imaging. Following Berkhout (1997), we first give a matrix expression of wave propagation from a source to a scatterer and then to a receiver:

$$W^U R(\theta) W^D=\begin{bmatrix} g^U_{11} & g^U_{12} & \cdots & g^U_{1P}\\ g^U_{21} & g^U_{22} & \cdots & g^U_{2P}\\ \cdots & \cdots & \cdots & \cdots\\ g^U_{M1} & g^U_{M2} & \cdots & g^U_{MP}\end{bmatrix}\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q}\\ r_{21} & r_{22} & \cdots & r_{2Q}\\ \cdots & \cdots & \cdots & \cdots\\ r_{P1} & r_{P2} & \cdots & r_{PQ}\end{bmatrix}\begin{bmatrix} g^D_{11} & g^D_{12} & \cdots & g^D_{1N}\\ g^D_{21} & g^D_{22} & \cdots & g^D_{2N}\\ \cdots & \cdots & \cdots & \cdots\\ g^D_{Q1} & g^D_{Q2} & \cdots & g^D_{QN}\end{bmatrix}, \tag{58}$$

where W^U is a discretized Green's function for upward wave propagation, W^D is a discretized Green's function for downward wave propagation, and $R(\theta)$ is a reflectivity matrix, which is related to the incident angle. If the variation of reflectivity with angle is neglected, $R(\theta)$ becomes a diagonal matrix. The reflectivity in this case is assumed to be the normal reflectivity. In practice, the reflectivity of a reflector varies with the incident angle. This is called

an AVO/AVA phenomenon in seismology. The prestack migration/inversion aims at estimating the angle reflectivity to evaluate lithological variations. On the other hand, the residual moveout of the angle reflectivity indicates whether the macro migration/inversion velocity is reasonable or not. The synthetic wave field can be written as follows:

$$P(\vec{x}_r,\vec{x}_s,\omega)=\sum_{i_z=Z_1}^{Z_N} W^U R(\theta) W^D|_{i_z}=\sum_{i_z=Z_1}^{Z_N}\begin{bmatrix} (g^U r g^D)_{11} & (g^U r g^D)_{12} & \cdots & (g^U r g^D)_{1N} \\ (g^U r g^D)_{21} & (g^U r g^D)_{22} & \cdots & (g^U r g^D)_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ (g^U r g^D)_{M1} & (g^U r g^D)_{M2} & \cdots & (g^U r g^D)_{MN} \end{bmatrix}_{i_z}. \tag{59}$$

In equation (59), each element of the matrix P is a recorded seismic trace in the time domain and a recorded amplitude value for a shot-receiver pair in the frequency domain. Each column is a shot gather, and each row is a common receiver gather. Therefore, equation (59) can be regarded as the matrix expression of equation (15). The classical prestack migration can be formulated as the following:

$$\left[W^U(z_0,z_1)\right]^H P(\vec{x}_r,\vec{x}_s,z_0,\omega)\left[W^D(z_1,z_0)\right]^H = R(z_1). \tag{60}$$

The detailed matrix expression of equation (60) is

$$\begin{bmatrix} \tilde{g}^U_{11} & \tilde{g}^U_{21} & \cdots & \tilde{g}^U_{M1} \\ \tilde{g}^U_{12} & \tilde{g}^U_{22} & \cdots & \tilde{g}^U_{M2} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}^U_{1P} & \tilde{g}^U_{2P} & \cdots & \tilde{g}^U_{MP} \end{bmatrix}\begin{bmatrix} P^{rs}_{11} & P^{rs}_{12} & \cdots & P^{rs}_{1N} \\ P^{rs}_{21} & P^{rs}_{22} & \cdots & P^{rs}_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ P^{rs}_{M1} & P^{rs}_{M2} & \cdots & P^{rs}_{MN} \end{bmatrix}\begin{bmatrix} \tilde{g}^D_{11} & \tilde{g}^D_{21} & \cdots & \tilde{g}^D_{Q1} \\ \tilde{g}^D_{12} & \tilde{g}^D_{22} & \cdots & \tilde{g}^D_{Q2} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}^D_{1N} & \tilde{g}^D_{2N} & \cdots & \tilde{g}^D_{QN} \end{bmatrix}$$
$$= \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q} \\ r_{21} & r_{22} & \cdots & r_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1} & r_{P2} & \cdots & r_{PQ} \end{bmatrix}, \tag{61}$$

where $R(z_1)$ is **the image of the first layer**. In $R(z_1)$, each row is an angle gather at an imaging point, and each column is a common angle gather. The multiplication of the pth row in the matrix $\left[W^U(z_0,z_1)\right]^H$ by any column in the matrix $P(\vec{x}_r,\vec{x}_s,\omega)$ corresponds to a detection focusing of a shot gather; the multiplication of the qth column in the matrix $\left[W^D(z_1,z_0)\right]^H$ by any row in the matrix $P(\vec{x}_r,\vec{x}_s,\omega)$ corresponds to an emission focusing. Then, **the image of the second layer** can be obtained with

$$\left[W^U(z_0,z_2)\right]^H P(\vec{x}_r,\vec{x}_s,z_0,\omega)\left[W^D(z_2,z_0)\right]^H = R(z_2). \tag{62}$$

Generally, **the image of the z_ith layer** is

$$\left[W^U(z_0,z_i)\right]^H P(\vec{x}_r,\vec{x}_s,z_0,\omega)\left[W^D(z_i,z_0)\right]^H = R(z_i). \tag{63}$$

Here, the matrices $\left[W^U\right]^H$ and $\left[W^D\right]^H$ are non-recursive. Otherwise, equation (62) and (63) will be of the following forms:

$$\left[W^U(z_1,z_2)\right]^H\left[W^U(z_0,z_1)\right]^H P(\vec{x}_r,\vec{x}_s,z_0,\omega)\left[W^D(z_1,z_0)\right]^H\left[W^D(z_2,z_1)\right]^H = R(z_2), \tag{64}$$

or

$$\left[W^U(z_1,z_2)\right]^H P(\vec{x}_r,\vec{x}_s,z_1,\omega)\left[W^D(z_2,z_1)\right]^H = R(z_2), \tag{65}$$

and

$$\left[W^U(z_{i-1},z_i)\right]^H \cdots \left[W^U(z_0,z_1)\right]^H P(\vec{x}_r,\vec{x}_s,z_0,\omega)\left[W^D(z_1,z_0)\right]^H \cdots \left[W^D(z_i,z_{i-1})\right]^H = R(z_i), \tag{66}$$

or

$$\left[W^U(z_{i-1},z_i)\right]^H P(\vec{x}_r,\vec{x}_s,z_{i-1},\omega)\left[W^D(z_i,z_{i-1})\right]^H = R(z_i). \tag{67}$$

Defining the cost function as

$$E(R(\theta)) = \|W^U R(\theta) W^D - P(\vec{x}_r,\vec{x}_s,\omega)\|_2^2 \tag{68}$$

yields the formula of the linearized migration/inversion:

$$R(\theta) = \frac{\left[W^U\right]^H P(\vec{x}_r,\vec{x}_s,\omega)\left[W^D\right]^H}{\left[W^U\right]^H\left[W^U\right]\left[W^D\right]\left[W^D\right]^H}. \tag{69}$$

The matrix expression of the migration/inversion in equation (69) is

$$\frac{\begin{bmatrix} \tilde{g}^U_{11} & \tilde{g}^U_{21} & \cdots & \tilde{g}^U_{M1} \\ \tilde{g}^U_{12} & \tilde{g}^U_{22} & \cdots & \tilde{g}^U_{M2} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}^U_{1P} & \tilde{g}^U_{2P} & \cdots & \tilde{g}^U_{MP} \end{bmatrix}\begin{bmatrix} P^{rs}_{11} & P^{rs}_{12} & \cdots & P^{rs}_{1N} \\ P^{rs}_{21} & P^{rs}_{22} & \cdots & P^{rs}_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ P^{rs}_{M1} & P^{rs}_{M2} & \cdots & P^{rs}_{MN} \end{bmatrix}\begin{bmatrix} \tilde{g}^D_{11} & \tilde{g}^D_{21} & \cdots & \tilde{g}^D_{Q1} \\ \tilde{g}^D_{12} & \tilde{g}^D_{22} & \cdots & \tilde{g}^D_{Q2} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}^D_{1N} & \tilde{g}^D_{2N} & \cdots & \tilde{g}^D_{QN} \end{bmatrix}}{[M]}$$

$$= \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q} \\ r_{21} & r_{22} & \cdots & r_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1} & r_{P2} & \cdots & r_{PQ} \end{bmatrix}, \tag{70}$$

where the denominator term $[M]$ is

$$[M] = \begin{bmatrix} \tilde{g}^D_{11} & \tilde{g}^D_{21} & \cdots & \tilde{g}^D_{Q1} \\ \tilde{g}^D_{12} & \tilde{g}^D_{22} & \cdots & \tilde{g}^D_{Q2} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}^D_{1N} & \tilde{g}^D_{2N} & \cdots & \tilde{g}^D_{QN} \end{bmatrix}\begin{bmatrix} \tilde{g}^U_{11} & \tilde{g}^U_{21} & \cdots & \tilde{g}^U_{M1} \\ \tilde{g}^U_{12} & \tilde{g}^U_{22} & \cdots & \tilde{g}^U_{M2} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}^U_{1P} & \tilde{g}^U_{2P} & \cdots & \tilde{g}^U_{MP} \end{bmatrix}$$
$$\begin{bmatrix} g^U_{11} & g^U_{12} & \cdots & g^U_{1P} \\ g^U_{21} & g^U_{22} & \cdots & g^U_{2P} \\ \cdots & \cdots & \cdots & \cdots \\ g^U_{M1} & g^U_{M2} & \cdots & g^U_{MP} \end{bmatrix}\begin{bmatrix} g^D_{11} & g^D_{12} & \cdots & g^D_{1N} \\ g^D_{21} & g^D_{22} & \cdots & g^D_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ g^D_{Q1} & g^D_{Q2} & \cdots & g^D_{QN} \end{bmatrix}. \tag{71}$$

From equation (61) and (70), the migration/inversion can be locally implemented, because all elements in the matrices W^U, $\left[W^U\right]^H$, W^D and $\left[W^D\right]^H$ relate only to a given layer. If the matrix $\left[W^U\right]^H$ is the inverse of the matrix W^U, it can be expressed as

$$\left[W^U\right]^H W^U = E, \tag{72}$$

where the matrix E is an identity matrix. Similarly, if the matrix $\left[W^D\right]^H$ is the inverse of the matrix W^D, we have

$$W^D\left[W^D\right]^H = E. \tag{73}$$

In practice, $\left[W^U\right]^H$ and $\left[W^D\right]^H$ are the conjugates of W^U and W^D respectively. Therefore, the matrix $[M]$ is a band-width-limited diagonal matrix. The velocity structure and the acquisition geometry affect the inner structure of the matrix. In fact, M is a Hessian which will be discussed in detail later.

[3] The meaning and calculation of $\nabla f(\delta m)$ and $\nabla^2 f(\delta m)$

The iterative formula of the least-squares migration/inversion is:

$$\delta\mathbf{m}^{(k+1)} = \delta\mathbf{m}^{(k)} - H^{k+1}\nabla f\left(\delta\mathbf{m}^{(k)}\right), \tag{74}$$

where H^{k+1} is the inverse of the Hessian. The first-order derivative $\nabla f\left(\delta\mathbf{m}^{(k)}\right)$ of the cost function with respect to the medium parameters is

$$\nabla f(\delta\mathbf{m}) = 2A^H\left(A\delta\mathbf{m} - \delta\mathbf{d}^{obs}\right). \tag{75}$$

If the residual wavefield is defined as

$$\mathbf{P}^{residual} = A\delta\mathbf{m} - \delta\mathbf{d}^{obs}, \tag{76}$$

and equation (75) is rewritten as

$$\nabla f(\delta\mathbf{m}) = 2A^H\mathbf{P}^{residual}, \tag{77}$$

then the first-order derivative means that the residual wavefield is back-propagated. It is further equivalent to the classical prestack migration if the parameter disturbance $\delta\mathbf{m}$ is set to zero at the first iteration. The residual wavefield $\mathbf{P}^{residual}$ belongs to the data space $D(\vec{x}_r, \vec{x}_s, t)$, and $\nabla f(\delta\mathbf{m})$ pertains to the image space $I(\vec{x})$. Calculating the first-order derivative requires one-time modeling $A\delta\mathbf{m}$, which can be implemented by a prestack demigration, and one-time classical prestack migration of $2A^H\mathbf{P}^{residual}$:

$$\nabla f(\delta\mathbf{m}) = 2 \begin{bmatrix} \tilde{g}^U_{11} & \tilde{g}^U_{21} & \cdots & \tilde{g}^U_{M1} \\ \tilde{g}^U_{12} & \tilde{g}^U_{22} & \cdots & \tilde{g}^U_{M2} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}^U_{1P} & \tilde{g}^U_{2P} & \cdots & \tilde{g}^U_{MP} \end{bmatrix} \begin{bmatrix} P^{residual-rs}_{11} & P^{residual-rs}_{12} & \cdots & P^{residual-rs}_{1N} \\ P^{residual-rs}_{21} & P^{residual-rs}_{22} & \cdots & P^{residual-rs}_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ P^{residual-rs}_{M1} & P^{residual-rs}_{M2} & \cdots & P^{residual-rs}_{MN} \end{bmatrix} \begin{bmatrix} \tilde{g}^D_{11} & \tilde{g}^D_{21} & \cdots & \tilde{g}^D_{Q1} \\ \tilde{g}^D_{12} & \tilde{g}^D_{22} & \cdots & \tilde{g}^D_{Q2} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}^D_{1N} & \tilde{g}^D_{2N} & \cdots & \tilde{g}^D_{QN} \end{bmatrix}. \tag{78}$$

In the first iterative step, $\delta\mathbf{m} = 0$, equation (78) is rewritten as

$$\begin{aligned}\nabla f(\delta\mathbf{m}) &= -2\begin{bmatrix} \tilde{g}^U_{11} & \tilde{g}^U_{21} & \cdots & \tilde{g}^U_{M1} \\ \tilde{g}^U_{12} & \tilde{g}^U_{22} & \cdots & \tilde{g}^U_{M2} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}^U_{1P} & \tilde{g}^U_{2P} & \cdots & \tilde{g}^U_{MP} \end{bmatrix}\begin{bmatrix} \delta\mathbf{d}^{obs-rs}_{11} & \delta\mathbf{d}^{obs-rs}_{12} & \cdots & \delta\mathbf{d}^{obs-rs}_{1N} \\ \delta\mathbf{b}^{obs-rs}_{21} & \delta\mathbf{d}^{obs-rs}_{22} & \cdots & \delta\mathbf{d}^{obs-rs}_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ \delta\mathbf{d}^{obs-rs}_{M1} & \delta\mathbf{d}^{obs-rs}_{M2} & \cdots & \delta\mathbf{d}^{obs-rs}_{MN} \end{bmatrix} \\ &\quad \begin{bmatrix} \tilde{g}^D_{11} & \tilde{g}^D_{21} & \cdots & \tilde{g}^D_{Q1} \\ \tilde{g}^D_{12} & \tilde{g}^D_{22} & \cdots & \tilde{g}^D_{Q2} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}^D_{1N} & \tilde{g}^D_{2N} & \cdots & \tilde{g}^D_{QN} \end{bmatrix} \\ &= -2\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q} \\ r_{21} & r_{22} & \cdots & r_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1} & r_{P2} & \cdots & r_{PQ} \end{bmatrix}. \end{aligned} \tag{79}$$

The Hessian is the second-order derivative of the cost function with respect to the medium parameters. It is of the following form:

$$\nabla^2 f(\delta\mathbf{m}) = 2A^H A. \tag{80}$$

In the least-squares migration/inversion, the Hessian is a deconvolution operator. It is used for de-blurring the image of the classical prestack migration. Physically, the Hessian is an indicator of the illumination. The energy of the wave propagating through a certain medium is expressed as follows:

$$E(\delta\mathbf{m}) = \|\delta\mathbf{d}\|^2 = \delta\mathbf{d}^H\delta\mathbf{d} = \delta\mathbf{m}^H A^H A\delta\mathbf{m} = \left[W^D\right]^H[\delta\mathbf{m}]^H\left[W^U\right]^H\left[W^U\right][\delta\mathbf{m}]\left[W^D\right]. \tag{81}$$

For a given layer and from the modeling equation (58), equation (81) can be rewritten as

$$\begin{aligned} E(\delta\mathbf{m}) &= \begin{bmatrix} g^D_{11} & g^D_{12} & \cdots & g^D_{1N} \\ g^D_{21} & g^D_{22} & \cdots & g^D_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ g^D_{Q1} & g^D_{Q2} & \cdots & g^D_{QN} \end{bmatrix}^H\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q} \\ r_{21} & r_{22} & \cdots & r_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1} & r_{P2} & \cdots & r_{PQ} \end{bmatrix}^H\begin{bmatrix} g^U_{11} & g^U_{12} & \cdots & g^U_{1P} \\ g^U_{21} & g^U_{22} & \cdots & g^U_{2P} \\ \cdots & \cdots & \cdots & \cdots \\ g^U_{M1} & g^U_{M2} & \cdots & g^U_{MP} \end{bmatrix}^H \\ &\quad \begin{bmatrix} g^U_{11} & g^U_{12} & \cdots & g^U_{1P} \\ g^U_{21} & g^U_{22} & \cdots & g^U_{2P} \\ \cdots & \cdots & \cdots & \cdots \\ g^U_{M1} & g^U_{M2} & \cdots & g^U_{MP} \end{bmatrix}\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q} \\ r_{21} & r_{22} & \cdots & r_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1} & r_{P2} & \cdots & r_{PQ} \end{bmatrix}\begin{bmatrix} g^D_{11} & g^D_{12} & \cdots & g^D_{1N} \\ g^D_{21} & g^D_{22} & \cdots & g^D_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ g^D_{Q1} & g^D_{Q2} & \cdots & g^D_{QN} \end{bmatrix}. \end{aligned} \tag{82}$$

Clearly, for a horizontal reflector with an even reflectivity and only the zero-offset reflectivity considered, $A^H A$ determines the energy of the wave which propagates to the layer. Equation

(82) can be rewritten as follows:

$$
\begin{aligned}
A^H A &= \left[W^D\right]^H \left[W^U\right]^H \left[W^U\right]\left[W^D\right] \\
&= \begin{bmatrix} g^D_{11} & g^D_{12} & \cdots & g^D_{1N} \\ g^D_{21} & g^D_{22} & \cdots & g^D_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ g^D_{Q1} & g^D_{Q2} & \cdots & g^D_{QN} \end{bmatrix}^H \begin{bmatrix} g^U_{11} & g^U_{12} & \cdots & g^U_{1P} \\ g^U_{21} & g^U_{22} & \cdots & g^U_{2P} \\ \cdots & \cdots & \cdots & \cdots \\ g^U_{M1} & g^U_{M2} & \cdots & g^U_{MP} \end{bmatrix}^H \\
&\quad \begin{bmatrix} g^U_{11} & g^U_{12} & \cdots & g^U_{1P} \\ g^U_{21} & g^U_{22} & \cdots & g^U_{2P} \\ \cdots & \cdots & \cdots & \cdots \\ g^U_{M1} & g^U_{M2} & \cdots & g^U_{MP} \end{bmatrix} \begin{bmatrix} g^D_{11} & g^D_{12} & \cdots & g^D_{1N} \\ g^D_{21} & g^D_{22} & \cdots & g^D_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ g^D_{Q1} & g^D_{Q2} & \cdots & g^D_{QN} \end{bmatrix} \\
&= \begin{bmatrix} \tilde{g}^D_{11} & \tilde{g}^D_{21} & \cdots & \tilde{g}^D_{Q1} \\ \tilde{g}^D_{12} & \tilde{g}^D_{22} & \cdots & \tilde{g}^D_{Q2} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}^D_{1N} & \tilde{g}^D_{2N} & \cdots & \tilde{g}^D_{QN} \end{bmatrix} \begin{bmatrix} \tilde{g}^U_{11} & \tilde{g}^U_{21} & \cdots & \tilde{g}^U_{M1} \\ \tilde{g}^U_{12} & \tilde{g}^U_{22} & \cdots & \tilde{g}^U_{M2} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}^U_{1P} & \tilde{g}^U_{2P} & \cdots & \tilde{g}^U_{MP} \end{bmatrix} \\
&\quad \begin{bmatrix} g^U_{11} & g^U_{12} & \cdots & g^U_{1P} \\ g^U_{21} & g^U_{22} & \cdots & g^U_{2P} \\ \cdots & \cdots & \cdots & \cdots \\ g^U_{M1} & g^U_{M2} & \cdots & g^U_{MP} \end{bmatrix} \begin{bmatrix} g^D_{11} & g^D_{12} & \cdots & g^D_{1N} \\ g^D_{21} & g^D_{22} & \cdots & g^D_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ g^D_{Q1} & g^D_{Q2} & \cdots & g^D_{QN} \end{bmatrix}.
\end{aligned} \tag{83}
$$

The row in the matrix $\left[W^U\right]^H$ multiplied by the column of the matrix $\left[W^U\right]$ and the row in the matrix $\left[W^D\right]$ multiplied by the column of the matrix $\left[W^D\right]^H$ are the cross-correlation between the conjugate of the Green's function and the Green's function at different receiver or shot positions respectively. The auto-correlation has a peak value, and the cross-correlation decreases rapidly as the distance increases between the receiver and shot positions. The auto-correlation values are on the diagonal. Therefore, the Hessian is a band-width-limited diagonal matrix. Its inverse is also a kind of band-width-limited diagonal matrix. In the extreme case, where only the elements on the diagonal of the Hessian are left, with non-diagonal set to zero, the elements on the diagonal of the inverse of the Hessian are the reciprocals of the elements on the diagonal of the Hessian. Therefore, the inverse of the Hessian plays the role of decreasing strong illumination and enhancing poor illumination. The Hessian itself reflects the illumination of each imaging point. The matrix expression of migration/inversion can be

summarized as follows:

$$\begin{aligned}
&\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q} \\ r_{21} & r_{22} & \cdots & r_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1} & r_{P2} & \cdots & r_{PQ} \end{bmatrix}_{P\times Q}^{k+1} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q} \\ r_{21} & r_{22} & \cdots & r_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1} & r_{P2} & \cdots & r_{PQ} \end{bmatrix}_{P\times Q}^{k} \\
-\ 2\ &\left\{ \begin{bmatrix} \tilde{g}_{11}^U & \tilde{g}_{21}^U & \cdots & \tilde{g}_{M1}^U \\ \tilde{g}_{12}^U & \tilde{g}_{22}^U & \cdots & \tilde{g}_{M2}^U \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}_{1P}^U & \tilde{g}_{2P}^U & \cdots & \tilde{g}_{MP}^U \end{bmatrix} \begin{bmatrix} g_{11}^U & g_{12}^U & \cdots & g_{1P}^U \\ g_{21}^U & g_{22}^U & \cdots & g_{2P}^U \\ \cdots & \cdots & \cdots & \cdots \\ g_{M1}^U & g_{M2}^U & \cdots & g_{MP}^U \end{bmatrix} \right\}_{P\times P}^{-1} \\
&\begin{bmatrix} \tilde{g}_{11}^U & \tilde{g}_{21}^U & \cdots & \tilde{g}_{M1}^U \\ \tilde{g}_{12}^U & \tilde{g}_{22}^U & \cdots & \tilde{g}_{M2}^U \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}_{1P}^U & \tilde{g}_{2P}^U & \cdots & \tilde{g}_{MP}^U \end{bmatrix} \begin{bmatrix} P_{11}^{rs} & P_{12}^{rs} & \cdots & P_{1N}^{rs} \\ P_{21}^{rs} & P_{22}^{rs} & \cdots & P_{2N}^{rs} \\ \cdots & \cdots & \cdots & \cdots \\ P_{M1}^{rs} & P_{M2}^{rs} & \cdots & P_{MN}^{rs} \end{bmatrix} \begin{bmatrix} \tilde{g}_{11}^D & \tilde{g}_{21}^D & \cdots & \tilde{g}_{Q1}^D \\ \tilde{g}_{12}^D & \tilde{g}_{22}^D & \cdots & \tilde{g}_{Q2}^D \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}_{1N}^D & \tilde{g}_{2N}^D & \cdots & \tilde{g}_{QN}^D \end{bmatrix}_{P\times Q} \\
&\left\{ \begin{bmatrix} \tilde{g}_{11}^D & \tilde{g}_{21}^D & \cdots & \tilde{g}_{Q1}^D \\ \tilde{g}_{12}^D & \tilde{g}_{22}^D & \cdots & \tilde{g}_{Q2}^D \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}_{1N}^D & \tilde{g}_{2N}^D & \cdots & \tilde{g}_{QN}^D \end{bmatrix} \begin{bmatrix} g_{11}^D & g_{12}^D & \cdots & g_{1N}^D \\ g_{21}^D & g_{22}^D & \cdots & g_{2N}^D \\ \cdots & \cdots & \cdots & \cdots \\ g_{Q1}^D & g_{Q2}^D & \cdots & g_{QN}^D \end{bmatrix} \right\}_{Q\times Q}^{-1}.
\end{aligned} \tag{84}$$

Substituting the residual imaging matrix into equation (84), it can be rewritten as follows:

$$\begin{aligned}
&\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q} \\ r_{21} & r_{22} & \cdots & r_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1} & r_{P2} & \cdots & r_{PQ} \end{bmatrix}_{P\times Q}^{k+1} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q} \\ r_{21} & r_{22} & \cdots & r_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1} & r_{P2} & \cdots & r_{PQ} \end{bmatrix}_{P\times Q}^{k} \\
-\ 2\ &\left\{ \begin{bmatrix} \tilde{g}_{11}^U & \tilde{g}_{21}^U & \cdots & \tilde{g}_{M1}^U \\ \tilde{g}_{12}^U & \tilde{g}_{22}^U & \cdots & \tilde{g}_{M2}^U \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}_{1P}^U & \tilde{g}_{2P}^U & \cdots & \tilde{g}_{MP}^U \end{bmatrix} \begin{bmatrix} g_{11}^U & g_{12}^U & \cdots & g_{1P}^U \\ g_{21}^U & g_{22}^U & \cdots & g_{2P}^U \\ \cdots & \cdots & \cdots & \cdots \\ g_{M1}^U & g_{M2}^U & \cdots & g_{MP}^U \end{bmatrix} \right\}_{P\times P}^{-1} \\
&\begin{bmatrix} r_{11}^{residual} & r_{12}^{residual} & \cdots & r_{1Q}^{residual} \\ r_{21}^{residual} & r_{22}^{residual} & \cdots & r_{2Q}^{residual} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1}^{residual} & r_{P2}^{residual} & \cdots & r_{PQ}^{residual} \end{bmatrix}_{P\times Q} \\
&\left\{ \begin{bmatrix} \tilde{g}_{11}^D & \tilde{g}_{21}^D & \cdots & \tilde{g}_{Q1}^D \\ \tilde{g}_{12}^D & \tilde{g}_{22}^D & \cdots & \tilde{g}_{Q2}^D \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{g}_{1N}^D & \tilde{g}_{2N}^D & \cdots & \tilde{g}_{QN}^D \end{bmatrix} \begin{bmatrix} g_{11}^D & g_{12}^D & \cdots & g_{1N}^D \\ g_{21}^D & g_{22}^D & \cdots & g_{2N}^D \\ \cdots & \cdots & \cdots & \cdots \\ g_{Q1}^D & g_{Q2}^D & \cdots & g_{QN}^D \end{bmatrix} \right\}_{Q\times Q}^{-1}.
\end{aligned} \tag{85}$$

With the quasi-Newton condition $\mathbf{p}^{(k)} = H_{k+1}\mathbf{q}^{(k)}$, and $\mathbf{p}^{(k)} = \delta\mathbf{m}^k - \delta\mathbf{m}^{(k+1)}$, $\mathbf{q}^{(k)} = \nabla f\left(\delta\mathbf{m}^{(k+1)}\right) - \nabla f\left(\delta\mathbf{m}^{(k)}\right)$, the DFP algorithm for calculating the inverse of the Hessian matrix is

$$H_{k+1}^{DFP} = H_k + \frac{\mathbf{p}^{(k)}\left(\mathbf{p}^{(k)}\right)^T}{\left(\mathbf{p}^{(k)}\right)^T\mathbf{q}^{(k)}} - \frac{\left(H_k\mathbf{q}^{(k)}\right)\left(H_k\mathbf{q}^{(k)}\right)^T}{\left(\mathbf{q}^k\right)^T H_k\mathbf{q}^{(k)}}, \tag{86}$$

where

$$\mathbf{p}^{(k)} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q} \\ r_{21} & r_{22} & \cdots & r_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1} & r_{P2} & \cdots & r_{PQ} \end{bmatrix}^{k} - \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q} \\ r_{21} & r_{22} & \cdots & r_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1} & r_{P2} & \cdots & r_{PQ} \end{bmatrix}^{k+1} \tag{87}$$

and

$$\begin{aligned} \mathbf{q}^{(k)} &= \left[2A^H\mathbf{P}^{residual}\right]^{k+1} - \left[2A^H\mathbf{P}^{residual}\right]^{k} \\ &= \begin{bmatrix} r_{11}^{residual} & r_{12}^{residual} & \cdots & r_{1Q}^{residual} \\ r_{21}^{residual} & r_{22}^{residual} & \cdots & r_{2Q}^{residual} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1}^{residual} & r_{P2}^{residual} & \cdots & r_{PQ}^{residual} \end{bmatrix}^{k+1} \\ &- \begin{bmatrix} r_{11}^{residual} & r_{12}^{residual} & \cdots & r_{1Q}^{residual} \\ r_{21}^{residual} & r_{22}^{residual} & \cdots & r_{2Q}^{residual} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1}^{residual} & r_{P2}^{residual} & \cdots & r_{PQ}^{residual} \end{bmatrix}^{k}, \end{aligned} \tag{88}$$

where $r_{ij}^{residual}$ is the image with the residual wavefield.

[4] Analysis of some factors that influence imaging

It is worthwhile to analyze the factors which affect the imaging quality and how they do so. We list four main factors: (1) too coarse sampling, (2) uneven sampling intervals or missing data, (3) illumination deficiency, (4) migration operator. All these factors will slow the convergence of iterative migration/inversion algorithms. It is obvious that too coarse sampling produces aliasing. Fig.5 demonstrates that uneven sampling intervals or missing data will cause imaging noise. Because the reflections from the vicinity of a reflecting point can not cancel each other, imaging noise appears. The illumination greatly affects the imaging quality in the case of a complex medium. From the view of the Radon transform and its inverse, if the seismic data set is band-width-unlimited, and the acquisition geometry is continuously distributed around an object body, the object can be uniquely restored with the inverse of the Radon transform. This is theoretically true. According to ray theory tomography, if rays are missing which should pass through the region of interest, the image of the region will be blurred. In fact, this is an amplitude distortion of the image. Therefore, the true medium parameters can never be recovered from real seismic data. The possibility of relative true-amplitude imaging is analyzed later. Propagators describe wave propagation in the background medium. Given the seismic data and the background parameters, only the propagators affect the iterative algorithms. Propagators should characterize wave propagation as accurately as possible. However, commonly used propagators are not so accurate and many wave phenomena are neglected. This can cause amplitude and phase distortion, as well as errors in Hessian calculation. Further, these will slow down the convergence of migration/inversion.

[5] The imaging gather of migration/inversion

Equation (84) or (85) gives the imaging matrix:

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1Q} \\ r_{21} & r_{22} & \cdots & r_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ r_{P1} & r_{P2} & \cdots & r_{PQ} \end{bmatrix}_{P\times Q}, \tag{89}$$

where P is the number of scatterers and Q is the number of incident angles for a scatterer. A row of the reflectivity matrix is an angle gather. However, the angle is not evenly sampled in a constant angle interval in complex medium. Physically, a reflection point or scatterer is not illuminated with a constant incident angle interval; mathematically, an angle gather evenly sampled with a constant angle interval for the point can be yielded with Fourier transform approaches. Weglein and Stolt (1999) and Sava and Fomel (2003) gave an approach for creating the angle gathers from the imaged data set. However, in our mind, we should know the difference. The uneven-incident-angle illumination will cause imaging noises. And this will also cause amplitude distortions in angle gathers. Fig.5 shows that a reflection point or scatterer is not illuminated with a constant incident angle interval.

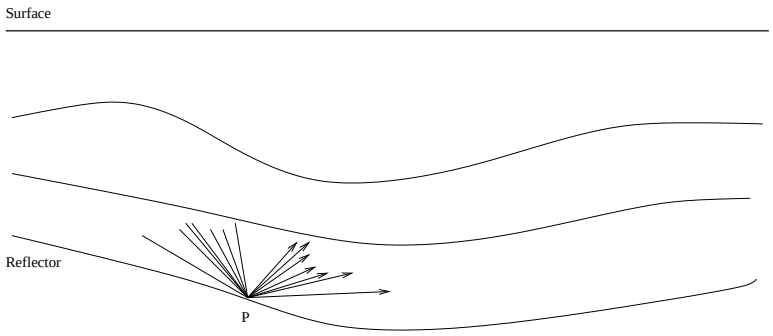

Figure 5: the illumination with the uneven incident angles huazhong1-even_incident_angle [CR]

[6] The resolution of migration/inversion

To understand which factors influence the resolution of the migration/inversion, we analyze the following matrices:

$$\left\{\left[W^U\right]^H\left[W^U\right]\right\}^{-1}\left\{\left[W^U\right]^H\left[W^U\right]\right\}[R(\theta)]\left\{\left[W^D\right]^H\left[W^D\right]\right\}\left\{\left[W^D\right]^H\left[W^D\right]\right\}^{-1}. \tag{90}$$

Ideally, all the matrices $\left\{\left[W^U\right]^H\left[W^U\right]\right\}^{-1}$, $\left\{\left[W^U\right]^H\left[W^U\right]\right\}$, $\left\{\left[W^D\right]^H\left[W^D\right]\right\}$ and $\left\{\left[W^D\right]^H\left[W^D\right]\right\}^{-1}$ are the identity matrix E, and at this time the image $[R(\theta)]$ is of the highest resolution. From ray theory tomography, resolution is closely related to acquisition aperture. The larger the aperture, the higher the resolution. If the matrices in equation (90) are required to be an

identity, besides the aperture, the spatial and temporal sampling and the propagator also have obvious effects. Theoretically, the Hessian blurs the true image of reflectivity, and the inverse of the Hessian deblurs the blurred image. In practice, the calculation of the Hessian is affected by many factors. As a result, the imaging quality is not improved distinctly.

[7] The regularization of the migration/inversion algorithms

In probability theory, the cost function is defined as (Tarantola, 1984)

$$2S(\mathbf{m}) = (g(\mathbf{m}) - \mathbf{d}_{obs})^T C_D^{-1} (g(\mathbf{m}) - \mathbf{d}_{obs}) + \left(\mathbf{m} - \mathbf{m}_{prior}\right) C_M^{-1} \left(\mathbf{m} - \mathbf{m}_{prior}\right) \tag{91}$$

where C_D and C_M are the covariance matrices for the data sets and models respectively. It is relatively easy to analyze the role of the covariance matrix C_M. The second term of equation (91) can be expressed in the following matrix form:

$$\begin{pmatrix} m_1 & m_2 & m_3 & \cdots & m_N \end{pmatrix} \begin{Bmatrix} c_{11} & c_{12} & c_{13} & \cdots & c_{1N} \\ c_{21} & c_{22} & c_{23} & \cdots & c_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{N1} & c_{N2} & c_{N3} & \cdots & c_{NN} \end{Bmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \\ \vdots \\ m_N \end{pmatrix}. \tag{92}$$

The vector $\mathbf{m} - \mathbf{m}_{prior}$ represents a stochastic process. The covariance matrix C_M stands for the linear relation of the stochastic process at two different times. If the value of the covariance is very small, the stochastic process at the two different times changes a little. Therefore, the uncertainty of the stochastic process is also low. Small values are chosen for the elements in the matrix C_M if we have enough a priori information of the model. The similarity between two images produced by two successive iterative migration/inversion steps are used to generate the covariance matrix. The correlation coefficients of the two images can be used to fill the covariance matrix C_M. Of course, the reciprocal of the correlation coefficients of the two images can be used to fill the covariance matrix C_M^{-1}. The local singular values (from SVD) of an image reflect the local continuity of an event, which can be used for constraints.

[8] The convergence of the iterative migration/inversion

Again, the iterative formula for migration/inversion imaging is

$$\delta\mathbf{m}^{(k+1)} = \delta\mathbf{m}^{(k)} - H^{k+1} \nabla f\left(\delta\mathbf{m}^{(k)}\right). \tag{93}$$

The choice of background parameters and the method for linearizing the propagators determine whether the cost function is (or approaches) a quadratic function. Therefore, the Born approximation should be replaced by De Wolf approximation or other more accurate approximations. On the other hand, to produce the image at the first iterative step with the so-called true-amplitude imaging approaches will accelerate the convergence of the iterative migration/inversion.

THE POSSIBILITY OF RELATIVE TRUE-AMPLITUDE IMAGING

Beylkin (1985) gave a formula for estimating the (locations of) discontinuities of the unknown function describing the medium, rather than for estimating the function itself. The generalized back-projection operator R^* dual to the generalized Radon transform R is defined as

$$\left(R^*u\right)(y)=\int_{\partial X}u\left(t,\xi\right)|_{t=\phi(y,\xi,\eta)}b\left(y,\xi\right)d\xi, \tag{94}$$

where $b(y,\xi)\in C^{\infty}(X\times\partial X)$ is a weighting function which is a smooth, non-negative function on $X\times\partial X$. It is defined as

$$b\left(y,\xi\right)=\frac{h\left(y,\xi\right)}{a\left(y,\xi\right)}\chi\left(y,\xi\right), \tag{95}$$

where $h(y,\xi)$ is a Jacobian determinant, $\chi(y,\xi)$ is a cut-off function, $u(t,\xi)$ can be regarded as a generalized projection, which is the observed scattered wavefield on the surface. With the help of a Fourier integral operator (FIO), the following formula can be obtained:

$$\left(I_{\partial X_{\eta}^{0}f}\right)(y)=\frac{1}{8\pi^{3}}\int_{\Omega_{\eta}(y)}e^{-ip\cdot y}\,f^{\wedge}(p)\,dp, \tag{96}$$

where $p=k\nabla_{y}\phi\left(y,\xi,\eta\right)$, which is the direction normal to the reflector at the imaging point. ∂X_{η}^{0} is the limited aperture on the surface, and $\Omega_{\eta}(y)$ is the corresponding limited aperture near the imaging point. And $f^{\wedge}(p)$ is the Fourier transform of the object function. From **Theorem 1** in Beylkin (1985), the inverse transform of the Fourier transform of the object function, $\left(I_{\partial X_{\eta}^{0}f}\right)(y)$, approximately equals the generalized back-projection defined by equation (94). Clearly, if the observation aperture ∂X_{η}^{0} or $\Omega_{\eta}(y)$ is unlimited, the discontinuities of the object function can be accurately estimated with equation (94) under some assumptions. That is to say, true-amplitude imaging can be theoretically implemented under some conditions. Here, we point that the relative true-amplitude imaging can be obtained in practice even though the observation aperture is always limited. Equation (96) can be considered as a summation. If the distribution of the Fourier transform of the object function on $\Omega_{\eta}(y)$ is even, and the range of the distribution is the same at every imaging point (or scattering point), the summation will be same. Even though the summation is not the true estimation for the discontinuities of the object function, the relative relation among these discontinuities at each imaging point is correct. This is the essence of the relative true-amplitude imaging or true-amplitude imaging. What prevent relative true-amplitude imaging? Missing data and limited aperture are the two main factors, assuming that the the back-projection operator is suitable. Beylkin (1985); Bleistein and Stockwell (2001); Zhang (2004) discussed how to construct a suitable back-projection operator or true-amplitude imaging operator. From the above point, we can further understand the meaning of the illumination analysis. The objective of the illumination analysis is to make clear how to map the observations on ∂X_{η}^{0} to $\Omega_{\eta}(y)$. If, on $\Omega_{\eta}(y)$, the distribution of the Fourier transform of the object function is even, and the range of the distribution is the same at every imaging point (or scattering point), then relative true-amplitude imaging can be achieved. It is worth to point out that the distribution should be symmetrical about the normal to the reflector at the imaging point.

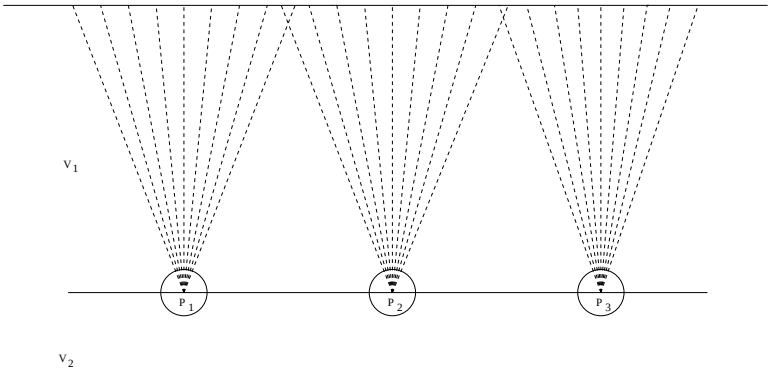

Figure 6: the possibility interpretation of the relative true-amplitude imaging in the case of simple medium structure. huazhong1-true_amplitude_and_illumination [CR]

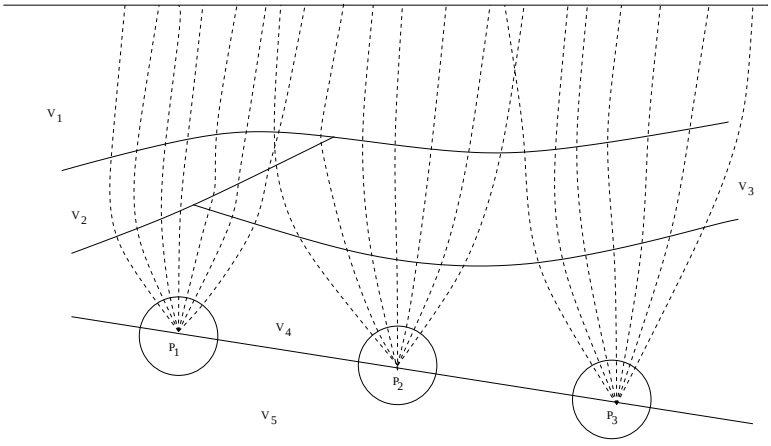

Figure 7: the possibility interpretation of the relative true-amplitude imaging in the case of complex medium structure. huazhong1-true_amplitude_and_illum22 [CR]

CONCLUSION AND DISCUSSION

Comparing non-iterative linearized migration/inversion, iterative linearized migration/inversion, and nonlinear waveform inversion, linearized migration/inversion has been carefully reviewed. The following statements reflect my opinions about linearized migration/inversion.

(1) There exists a paradox: linearized migration/inversion requires that the background velocity is as close as possible to the true velocity distribution; however the Born approximation needs a very smooth background velocity, so that the incident wavefield only includes downward propagating waves. The Born approximation is not a good linearization method and it should be replaced by more accurate approximations, such as De Wolf approximation.

(2) In essence, non-iterative linearized migration/inversion, iterative linearized migration/inversion, and nonlinear waveform inversion, are all inverse scattering imaging methods.

(3) Relative true-amplitude imaging is possible, the condition is that the distribution of the Fourier transform of the object function on $\Omega_\eta(y)$ is even, and the range of the distribution is the same at every imaging point (or scattering point).

(4) The Hessian matrix is band-limited and its inverse is also band-limited, its inner structure has close relation with the complexity of the velocity and the acquisition geometry. It reflects the energy illuminating imaging points.

ACKNOWLEDGMENT

The first author thanks the SEP Group at Stanford University for providing an opportunity for him to serve as a visiting scholar.

REFERENCES

Berkhout, A. J., 1997, Pushing the limits of seismic imaging, part I: Prestack migration in terms of double dynamic focusing: Geophysics, **62**, 937–953.

Beylkin, G., 1985, Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized radon transform: J. Math. Phys. 26(1), 99–108.

Bleistein, N. Cohen, J., and Stockwell, J., 2001, Mathematics of multidimensional seismic imaging, migration, and inversion Springer, 299.

Bleistein, N., 1987, On the imaging of reflectors in the Earth: Geophysics, **52**, 931–942.

Bleistein, N., Cohen, J. K., and Hagin, F. G., 1987, Two and one-half dimensional Born inversion with an arbitrary reference: Geophysics, **52**, 26–36.

Clayton, R. W., and Stolt, R. H., 1981, A born-WKBJ inversion method for acoustic reflection data: Geophysics, **46**, 1559–1567.

Gray, S. H., 1997, True-amplitude seismic migration: A comparison of three approaches: Geophysics, **62**, 929–936.

Huazhong, W., 2003, Seisimic imaging research report: Reflection Seismology Research Group, School of Ocean and Earch Science, Tongji University.

Hubral, P., Schleicher, J., and Tygel, M., 1996, A unified approach to 3-D seismic reflection imaging, part I: Basic concepts: Geophysics, **61**, 742–758.

Miller, D., Oristaglio, M., and Beylkin, G., 1987, A new slant on seismic imaging - Migration and integral geometry: Geophysics, **52**, 943–964.

Mora, P. R., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **52**, 1211–1228.

Pratt, R. G., S. C., and Hicks, G. J., 1998, Gauss-newton and full newton methods in frequency-space seismic waveform inversion: Geophys. J. Int., **133**, 341–362.

Sava, P. C., and Fomel, S., 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074.

Stolt, R. H., and Weglein, A. B., 1985, Migration and inversion of seismic data: Geophysics, **50**, 2458–2472.

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

Weglein, A., and Stolt, R., 1999, Migration-inversion revisited: The Leading Edge, August, 950–952,975.

Wu, R. S., and Toksoz, M. N., 1987, Diffraction tomography and multisource holography applied to seismic imaging: Geophysics, **52**, 11–25.

Zhang, Y., 2004, True amplitude migration based on one-way wave equations: Lecture Notes for UCSC WTOPI 2004 Workshop, 1–24.