Wave-equation tomography by beam focusing

Biondo Biondi

ABSTRACT

Velocity can be estimated using a wave-equation operator by maximizing an objective function that measures the flatness of the crosscorrelation computed between a source wavefield and a receiver wavefield. The proposed objective function depends on the parameters of a residual moveout applied to the computed correlation. It is composed of two terms: the first term maximizes the energy of the stack computed on local subarrays as a function of the local curvature. The second term maximizes the power of the stack computed globally as a function of time shifts applied to the stacks of the local subarrays. The first terms is essential to assure global convergence in presence of large velocity errors. The second term plays a role in estimating localized velocity anomalies. Numerical examples of computation of the gradients of the proposed objective function confirm its potential for velocity estimation.

INTRODUCTION

Tomographic velocity estimation based on wave-equation operators can improve seismic imaging in areas where wavefield-continuation migration is needed. However, it is well-known that the straightforward application of waveform inversion to estimate migration velocity fails to converge to an accurate model when the starting model is too far from the correct one. This failure to converge is caused by the non-linear relationship between data amplitudes and velocity. To avoid this failure the velocity-estimation problem can be formulated in the image domain as the maximization (or minimization) of objective functions that are more sensitive to the data kinematics than to the data amplitudes. Two important examples of this approach are the Wave-Equation Migration Velocity Analysis (WEMVA) method (Biondi and Sava, 1999; Sava and Biondi, 2004a,b; Sava, 2004) and the Differential Semblance Optimization (DSO) (Symes and Carazzone, 1991; Shen, 2004; Shen et al., 2005). Luo and Schuster (1991) introduced a method based on a kinematic objective function to solve the problem of transmission tomography. Both their method and the WEMVA method suffer from the drawback that they require the picking of kinematic parameters: correlation lag in one case (Luo and Schuster, 1991), and a residual migration parameter for WEMVA.

In this paper I develop a framework to update migration velocity by maximizing an objective function defined in the image domain. The objective function is defined

as a function of moveout parameters but velocity updating can be performed without explicit picking of the residual moveout parameters. Therefore, it overcomes one of the main difficulties of the WEMVA methods. The methodology is general and can be thus used to optimize the image as a function of arbitrary residual moveout parameters, and possibly of residual migration parameters.

I also introduce a new objective function that overcomes limitations of known methods. This novel objective function has two components: the first term measures the power of the stack over local subarrays (beams) as a function of the moveout curvature. The second term measures the power of the stack across the beams as a function of a bulk-shift of each beam. I then apply the general theory to the computation of the gradient of the proposed objective function with respect to velocity perturbations.

I develop the theory and show the results of numerical tests for a transmission tomography problem because transmission tomography is simpler than reflection tomography, and therefore better suited to the illustration of the basic concepts. I leave to future reports the application to reflection tomography of the method developed for transmission tomography in this report.

THEORY

In this section I develop the general theory for a transmission tomography problem because it is simpler than reflection tomography. In transmission tomography, the data are recorded after only one propagation path through the medium, as opposed to the downgoing and upgoing paths of a typical reflection tomography problems. Furthermore, in transmission tomography there is no need to image and locate reflectors in depth, which is a major hurdle in reflection tomography. However, the application to reflection tomography of the theory presented in this paper should be fairly straightforward. I propose to solve the transmission tomography problem by maximizing an objective function based on the correlation between recorded data and modeled data. This correlation is analogous to the correlation between source and receiver wavefields required by migration imaging condition.

To further simplify the theoretical development, I define an objective function that rewards consistency of the correlation computed independently for each source location. The objective function measures correlation consistency along the receiver axis. However, I use here the receiver axis as a proxy for the offset axis or the aperture-angle axis in reflection tomography. The application of the concepts developed in this paper to objective functions useful in reflection tomography should be straightforward, although it will require more complex notation and result in expressions for the gradients even more complex than the ones presented here.

I define the recorded data as $P_D\left(t, x_g, x_s\right)$, and the modeled data as $\widetilde{P}\left(t, x_g, x_s; s\right)$, where t is the recording time, x_g is the receiver coordinate, x_s is the source coordinate, and $s\left(z, x\right)$ is the slowness model defined in depth z and along the horizontal

coordinate x.

The cross-correlation $C(\tau)$ between the recorded data and the modeled data is defined as a function of the correlation time lag τ as

$$C(\tau)\left[\widetilde{P}(t), P_D(t)\right] = \sum_t \widetilde{P}(t-\tau) P_D(t). \tag{1}$$

I introduce an objective function that maximizes the flatness of the correlation function along the receiver axis for all values of the lag τ. In particular, I aim to maximize local correlation flatness after subdividing the receiver array into local subarrays. To extract the correlation for each subarray centered at $\overline{x_g}$, I apply a local beam-decomposition operators $\mathbf{B}_{\overline{x}}$. Within in each subarray, traces are defined by the local offset Δx_g. The dimensions of each $\mathbf{B}_{\overline{x}}$ are thus $\left(N_{\Delta x_g} N_\tau \times N_{x_g} N_\tau\right)$.

Given a background slowness s_0 we can compute the correlation in equation 1. In each subarray, the correlation can be flattened by the application of $N_{\overline{x_g}}$ moveout operators $\mathcal{M}_{\overline{x}}$; that is

$$C(\tau + \theta(\boldsymbol{\mu}_{\overline{x}}))\left[\widetilde{P}(t; s_0), P_D(t)\right] = \mathcal{M}_{\overline{x}}\left[\theta(\boldsymbol{\mu}_{\overline{x}}), \mathbf{B}_{\overline{x}} C(\tau; s_0)\right], \tag{2}$$

where $\boldsymbol{\mu}_{\overline{x}}$ are the moveout parameters and θ are the corresponding time shifts. I further define the local stacking operator $\mathbf{S}_{\overline{x}}$ that sums the correlation traces along the local offset axis Δx_g.

I can now introduce the first, and local, term of the objective function that measures the flatness of the correlation within each subarray as:

$$J_{\text{Local}}(\boldsymbol{\mu}_{\overline{x}}(s)) = \frac{1}{2} \sum_{x_s} \sum_{\overline{x_g}} \left\| \mathbf{S}_{\overline{x}} \mathcal{M}_{\overline{x}}\left[\theta(\boldsymbol{\mu}_{\overline{x}}(s)), \mathbf{B}_{\overline{x}} C(\tau; s_0)\right] \right\|_2^2. \tag{3}$$

This objective function is not a direct function of the slowness s, but it depends indirectly from it through the moveout parameters $\boldsymbol{\mu}_{\overline{x}}$. These parameters are the solutions of $N_{x_s} \times N_{\overline{x_g}}$ independent fitting problems, one for each subarray and source location. These *auxiliary* objective functions measure the zero lag of the cross-correlation between the correlation computed for a realization of the slowness function s and and the moved-out correlation computed with the background slowness s_0,

$$\begin{aligned} J_{\text{FL}}(\boldsymbol{\mu}_{\overline{x}}) &= C(\tau = 0)\left[\mathcal{M}_{\overline{x}}\left[\theta(\boldsymbol{\mu}_{\overline{x}}), \mathbf{B}_{\overline{x}} C(\tau; s_0)\right], \mathbf{B}_{\overline{x}} C(\tau; s)\right] \\ &= \left\langle \mathcal{M}_{\overline{x}}\left[\theta(\boldsymbol{\mu}_{\overline{x}}), \mathbf{B}_{\overline{x}} C(\tau; s_0)\right], \mathbf{B}_{\overline{x}} C(\tau; s) \right\rangle, \end{aligned} \tag{4}$$

where with the notation $\langle \mathbf{x}, \mathbf{y} \rangle$ I indicate the inner product of the vectors $\mathbf{x}$ and $\mathbf{y}$. This inner product spans the time-lag axis τ and the local offset axis Δx_g. The local moveout parameters are the solutions of the following maximization problem:

$$\max_{\boldsymbol{\mu}_{\overline{x}}} J_{\text{FL}}(\boldsymbol{\mu}_{\overline{x}}). \tag{5}$$

For velocity estimation, the most effective parametrization of the moveout within each beam is the curvature μ_C, that defines the following moveout equation

$$\theta\left(\boldsymbol{\mu}_{\overline{x}}\right) = \mu_C {\Delta x_g}^2. \tag{6}$$

Notice that when the slowness is equal to the background slowness s_0, the corresponding best-fitting moveout parameters $\bar{\mu}_{\overline{x}}$ are obviously the ones corresponding to no moveout; that is, $\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right) = 0$.

As the numerical examples I show in the next section demonstrate, the beam curvature is effective to capture the long-wavelength perturbations in the velocity model, but is less effective to capture the short-wavelength perturbations. Accordingly, a wave-equation tomography based solely on the objective function 3 may have difficulties to estimate short-wavelength velocity perturbations.

To address this shortcoming I introduce a second, and global, term to the objective function. This term measures flatness across the subarrays, after the local moveouts have been applied, and is defined as,

$$J_{\text{Global}}\left(\boldsymbol{\mu}\left(s\right)\right) = \frac{1}{2}\sum_{x_s} \left\| \mathbf{S}\mathcal{M}\left\{\theta\left(\boldsymbol{\mu}\left(s\right)\right), \boldsymbol{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}C\left(\tau; s_0\right)\right]\right\}\right\|_2^2, \tag{7}$$

where $\boldsymbol{\Sigma}_{\overline{x}}$ assembles all the results of the stacking over the subarrays into a global array, $\mathbf{S}$ is a global stacking operator, and $\mathcal{M}$ is a global moveout operator function of the vector of parameter $\boldsymbol{\mu}$. They both operate on the result of the local stacking of the subarrays. As in the previous case, the moveout parameters are solutions of N_{x_s} independent fitting problems, one for each source location. Similarly, these auxiliary objective functions measure the zero lag of the cross-correlation between the local stack of the correlation computed using the current slowness function and the local stack of the moved-out correlation computed using the background slowness; that is,

$$\begin{aligned} &J_{\text{FG}}\left(\boldsymbol{\mu}\right) = \\ &\quad \left\langle \mathcal{M}\left\{\theta\left(\boldsymbol{\mu}\right), \boldsymbol{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}C\left(\tau; s_0\right)\right]\right\}, \boldsymbol{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}C\left(\tau; s\right)\right]\right\rangle. \end{aligned} \tag{8}$$

In this case the inner product spans only the time-lag axis τ.

The global moveout parameters are the solutions of the following N_{x_s} maximization problems

$$\max_{\boldsymbol{\mu}} J_{\text{FG}}\left(\boldsymbol{\mu}\right). \tag{9}$$

I chose to parametrize the global moveout as simple time shifts for each beam center $\overline{x}$ that is, the moveout equation is

$$\theta\left(\boldsymbol{\mu}\right) = \mu_\theta. \tag{10}$$

Notice that with this choice of moveout parameters each maximization problem in 9 is an ensemble of $N_{\overline{x_g}}$ independent problems. This consideration becomes important when computing the gradient of the objective function.

Combining the objective function in 3 and in 7 we define the maximization problem that we solve to estimate slowness:

$$\max_{s} \left[J_{\text{Local}} \left(\boldsymbol{\mu}_{\overline{x}} \left(s \right) \right) + \epsilon J_{\text{Global}} \left(\boldsymbol{\mu} \left(s \right) \right) \right], \tag{11}$$

where the parameter ϵ can be tuned to find an optimal relative scaling between the local and global components, although in principle $\epsilon = 1$ should be effective.

Gradient of the objective function

I plan to solve the optimization problem defined in 11 by a gradient-based optimization algorithm. Therefore, the development of an algorithm to efficiently compute the gradient of the objective function with respect to slowness is an essential step to make the method practical. In this section I introduce the basic methodology to compute the gradients, and I leave some of the details to Appendix A.

The gradient of both the local objective function 3 and the global one 7 are computed using the chain rule. The first terms of the chains are the derivatives of the objective function with respect the moveout parameters. The second terms are the derivatives of the moveout parameters with respect to slowness; they are computed from the fitting objective functions 4 and 8.

Derivatives with respect to moveout parameters

The computation of the derivatives of 3 with respect to each vector of local-moveout parameters is easily evaluated using the following expression:

$$\frac{\partial J_{\text{Local}}}{\boldsymbol{\mu}_{\overline{x}}} = \frac{\partial \mathcal{M}_{\overline{x}}{}'}{\partial \boldsymbol{\mu}_{\overline{x}}} \mathbf{S}'_{\overline{x}} \mathbf{S}_{\overline{x}} \mathcal{M}_{\overline{x}} \left[\theta \left(\bar{\boldsymbol{\mu}}_{\overline{x}} \right), \mathbf{B}_{\overline{x}} C \left(\tau; s_0 \right) \right]. \tag{12}$$

The linear operator $\frac{\partial \mathcal{M}_{\overline{x}}}{\partial \boldsymbol{\mu}_{\overline{x}}}$ has the dimensions $\left(N_{\Delta x_g} N_\tau \times N_{\mu_{\overline{x}}} \right)$ and is given by

$$\frac{\partial \mathcal{M}_{\overline{x}}}{\partial \boldsymbol{\mu}_{\overline{x}}} = \mathcal{M}_{\overline{x}} \left[\theta \left(\boldsymbol{\mu}_{\overline{x}} \right), \mathbf{B}_{\overline{x}} \, \dot{C} \left(\tau; s_0 \right) \right] \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\overline{x}}}, \tag{13}$$

where $\dot{C} \left(\tau; s_0 \right) = C \left(\tau \right) \left[\widetilde{P} \left(t; s_0 \right), \dot{P}_D \left(t \right) \right]$, with $\dot{P}_D$ being the time derivative of the recorded-data traces. For the choice of moveout parameters expressed in equation 6 we have $\partial \theta / \partial \boldsymbol{\mu}_{\overline{x}} = \partial \theta / \partial \mu_C = \Delta {x_g}^2$.

Similarly, the evaluation of the derivatives of 7 with respect to each shift parameter $\boldsymbol{\mu}$ is easily carried out by the following:

$$\frac{\partial J_{\text{Global}}}{\partial \boldsymbol{\mu}} = \frac{\partial \mathcal{M}'}{\partial \boldsymbol{\mu}} \mathbf{S}' \mathbf{S} \boldsymbol{\Sigma}_{\overline{x}} \mathbf{S}_{\overline{x}} \mathcal{M}_{\overline{x}} \left[\theta \left(\bar{\boldsymbol{\mu}}_{\overline{x}} \right), \mathbf{B}_{\overline{x}} C \left(\tau; s_0 \right) \right], \tag{14}$$

where the linear operator $\frac{\partial \mathcal{M}}{\partial \boldsymbol{\mu}}$ is given by

$$\frac{\partial \mathcal{M}}{\partial \boldsymbol{\mu}} = \mathcal{M}\left\{\theta\left(\boldsymbol{\mu}\right), \boldsymbol{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}\,\dot{C}\left(\tau; s_0\right)\right]\right\}\frac{\partial \theta}{\partial \boldsymbol{\mu}}. \tag{15}$$

When the moveout parameters are simple trace-by-trace phase shifts, as defined in equation 10, it results that $\partial\theta/\partial\boldsymbol{\mu} = 1$.

On a practical note, the preceding expressions look more daunting than they are in practice. They greatly simplify in the important case when the gradient is evaluated for $\bar{\boldsymbol{\mu}}_{\overline{x}} = 0$ and $\bar{\boldsymbol{\mu}} = 0$. This simplifying condition is actually always fulfilled unless the optimization algorithm includes inner iterations for fitting the moveout parameters using a linearized approach. Under these conditions, equations 12 and 13 become, respectively,

$$\left.\frac{\partial J_{\text{Local}}}{\boldsymbol{\mu}_{\overline{x}}}\right|_{\boldsymbol{\mu}_{\overline{x}}=0} = \frac{\partial \mathcal{M}_{\overline{x}}{}'}{\partial \boldsymbol{\mu}_{\overline{x}}}\mathbf{S}'_{\overline{x}}\mathbf{S}_{\overline{x}}\mathbf{B}_{\overline{x}}C\left(\tau; s_0\right), \tag{16}$$

and

$$\frac{\partial \mathcal{M}_{\overline{x}}}{\partial \boldsymbol{\mu}_{\overline{x}}} = \mathbf{B}_{\overline{x}}\,\dot{C}\left(\tau; s_0\right)\frac{\partial \theta}{\partial \boldsymbol{\mu}_{\overline{x}}}. \tag{17}$$

Similarly, equations 14 and 15 become

$$\left.\frac{\partial J_{\text{Global}}}{\partial \boldsymbol{\mu}}\right|_{\boldsymbol{\mu}_{\overline{x}}=0, \boldsymbol{\mu}=0} = \frac{\partial \mathcal{M}'}{\partial \boldsymbol{\mu}}\mathbf{S}'\mathbf{S}\boldsymbol{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathbf{B}_{\overline{x}}C\left(\tau; s_0\right), \tag{18}$$

and

$$\frac{\partial \mathcal{M}}{\partial \boldsymbol{\mu}} = \boldsymbol{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathbf{B}_{\overline{x}}\,\dot{C}\left(\tau; s_0\right)\frac{\partial \theta}{\partial \boldsymbol{\mu}}. \tag{19}$$

Derivatives with respect to slowness

The evaluation of the derivatives of the moveout parameters with respect to slowness follows a slightly different procedure from the one above because the moveout parameters are solutions of the optimization problems 5 and 9. We take advantage of the fact that we need to evaluate the derivatives only at the solution points, where the objective functions are maximized and thus their derivatives with respect to the moveout parameters are zero. We can therefore write:

$$\left.\frac{\partial J_{\text{FL}}\left(\boldsymbol{\mu}_{\overline{x}}\right)}{\partial \boldsymbol{\mu}_{\overline{x}}}\right|_{\boldsymbol{\mu}_{\overline{x}}=\bar{\boldsymbol{\mu}}_{\overline{x}}} = \dot{J_{\text{FL}}}\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right) = 0 = \left\langle \left.\frac{\partial \mathcal{M}_{\overline{x}}\left[\theta\left(\boldsymbol{\mu}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}C\left(\tau; s_0\right)\right]}{\partial \boldsymbol{\mu}_{\overline{x}}}\right|_{\boldsymbol{\mu}_{\overline{x}}=\bar{\boldsymbol{\mu}}_{\overline{x}}}, \mathbf{B}_{\overline{x}}C\left(\tau; s\right)\right\rangle,$$

and

$$\left.\frac{\partial J_{\text{FG}}\left(\boldsymbol{\mu}\right)}{\partial \boldsymbol{\mu}}\right|_{\boldsymbol{\mu}=\bar{\boldsymbol{\mu}}} = \dot{J_{\text{FG}}}\left(\bar{\boldsymbol{\mu}}\right) = 0 = \left\langle \left.\frac{\partial \mathcal{M}\left\{\theta\left(\boldsymbol{\mu}\right), \boldsymbol{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}C\left(\tau; s_0\right)\right]\right\}}{\partial \boldsymbol{\mu}}\right|_{\boldsymbol{\mu}=\bar{\boldsymbol{\mu}}}, \boldsymbol{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}C\left(\tau; s\right)\right]\right\rangle.$$

Using the rule for differentiating implicit functions, and taking advantage that the fitting problems are all independent from each other (i.e. the cross derivatives with respect to the moveout parameters are all zero), we can formally write:

$$\left.\frac{\partial \boldsymbol{\mu}_{\overline{x}}}{\partial s}\right|_{\mu_{\overline{x}}=\bar{\mu}_{\overline{x}}} = -\frac{\frac{\partial \dot{J_{\mathrm{FL}}}(\boldsymbol{\mu}_{\overline{x}})}{\partial s}}{\frac{\partial \dot{J_{\mathrm{FL}}}(\boldsymbol{\mu}_{\overline{x}})}{\partial \boldsymbol{\mu}_{\overline{x}}}}, \tag{20}$$

and

$$\left.\frac{\partial \boldsymbol{\mu}}{\partial s}\right|_{\mu=\bar{\mu}} = -\frac{\frac{\partial \dot{J_{\mathrm{FG}}}(\boldsymbol{\mu})}{\partial s}}{\frac{\partial \dot{J_{\mathrm{FG}}}(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}}. \tag{21}$$

Appendix A presents the analytical development of these expressions to compute the derivatives of the moveout parameters with respect to slowness. As for the derivatives of the main objective function with respect to moveout parameters, the final results for the special case of $\bar{\boldsymbol{\mu}}_{\overline{x}} = 0$ and $\bar{\boldsymbol{\mu}} = 0$ have a fairly simple analytical expression. The derivative of the local moveout parameters are (A-1):

$$\left.\frac{\partial \boldsymbol{\mu}_{\overline{x}}}{\partial s}\right|_{\mu_{\overline{x}}=0} = -\frac{\left\langle \mathbf{B}_{\overline{x}}\, \dot{C}\,(\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\overline{x}}}, \mathbf{B}_{\overline{x}} \mathbf{P_D} \frac{\partial \widetilde{P}}{\partial s} \right\rangle}{\left\langle \mathbf{B}_{\overline{x}}\, \ddot{C}\,(\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\overline{x}}}, \mathbf{B}_{\overline{x}} C\,(\tau; s) \right\rangle}, \tag{22}$$

and the derivative of the global moveout parameters are (A-2):

$$\left.\frac{\partial \boldsymbol{\mu}}{\partial s}\right|_{\mu_{\overline{x}}=0,\mu=0} = -\frac{\left\langle \boldsymbol{\Sigma}_{\overline{x}} \mathbf{S}_{\overline{x}} \mathbf{B}_{\overline{x}}\, \dot{C}\,(\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}}, \boldsymbol{\Sigma}_{\overline{x}} \mathbf{S}_{\overline{x}} \mathbf{B}_{\overline{x}} \mathbf{P_D} \frac{\partial \widetilde{P}}{\partial s} \right\rangle}{\left\langle \boldsymbol{\Sigma}_{\overline{x}} \mathbf{S}_{\overline{x}} \mathbf{B}_{\overline{x}}\, \ddot{C}\,(\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}}, \boldsymbol{\Sigma}_{\overline{x}} \mathbf{S}_{\overline{x}} \mathbf{B}_{\overline{x}} C\,(\tau; s) \right\rangle}, \tag{23}$$

in which $\ddot{C}\,(\tau) = C\,(\tau)\left[\widetilde{P}\,(t)\,, \ddot{P}_D\,(t)\right]$. In both equations 22 and 23 the operator $\mathbf{P_D}$ represents a convolution with the recorded data, whereas the operator $\partial \widetilde{P}/\partial s$ is the basic wave-equation tomography operator that that links perturbations in the slowness model to perturbations in the modeled data.

Combining the derivatives in equation 22 with the derivatives in equations 16-17 we can compute the gradient of the local objective function 3 with respect to slowness as:

$$\nabla J_{\mathrm{Local}} = -\underbrace{\frac{\partial \widetilde{P}'}{\partial s} \mathbf{P_D}' \mathbf{B}'_{\overline{x}}}_{\mathrm{I}} \underbrace{\mathbf{B}_{\overline{x}}\, \dot{C}\,(\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\overline{x}}}}_{\mathrm{II}} \underbrace{\frac{\frac{\partial J_{\mathrm{Local}}}{\boldsymbol{\mu}_{\overline{x}}}}{\left\langle \mathbf{B}_{\overline{x}}\, \ddot{C}\,(\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\overline{x}}}, \mathbf{B}_{\overline{x}} C\,(\tau; s) \right\rangle}}_{\mathrm{III}}. \tag{24}$$

I will now examine the effects of each of the terms in equation 24 starting from the rightmost one. The third term (III) produces a scalar for each local curvature parameter μ_C. This scalar multiplies the traces in each beam, after they have been differentiated in time and scaled by $\partial\theta/\partial\boldsymbol{\mu}_{\overline{x}}$, as described by the second term (II).

Notice that the phase introduced by the time derivative of the correlation function in (II) is crucial for the successful backprojection into the slowness model that is accomplished by the first term (I). In this term, first $\mathbf{B}'_{\overline{x}}$ projects the traces of each individual beam into the space of the global array, then the convolution with the recorded data $\mathbf{P_D}'$ time shifts the correlation function by the time delay of the events. Finally, the adjoint of the operator $\partial\widetilde{P}/\partial s$ backprojects the perturbation in the wavefields at the receiver array into the slowness model.

The expression of the gradient of the global objective function 7 with respect to slowness is similarly derived by combining the derivatives in equation 23 with the derivatives in equations 18-19 and is the following three-terms expression:

$$\nabla J_{\text{Global}} = \underbrace{-\frac{\partial \widetilde{P}'}{\partial s}\,\mathbf{P_D}'\mathbf{B}'_{\overline{x}}\mathbf{S}'_{\overline{x}}\mathbf{\Sigma}'_{\overline{x}}}_{\text{I}}\;\underbrace{\mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathbf{B}_{\overline{x}}\,\dot{C}\,(\tau;s_0)\,\frac{\partial\theta}{\partial\boldsymbol{\mu}}}_{\text{II}}\;\underbrace{\frac{\frac{\partial J_{\text{Global}}}{\partial\boldsymbol{\mu}}}{\left\langle \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathbf{B}_{\overline{x}}\,\ddot{C}\,(\tau;s_0)\,\frac{\partial\theta}{\partial\boldsymbol{\mu}},\,\mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathbf{B}_{\overline{x}}C\,(\tau;s)\right\rangle}}_{\text{III}}. \tag{25}$$

The structure of equation 25 is similar to the structure of equation 24 and the terms have similar explanations. The only important difference is that in term II the chain $\mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}$ performs the stack over the local arrays and the assemblage of the stacked traces into the global array, whereas its adjoint in term I spreads the stacked traces back into the local arrays reforming the local beams.

NUMERICAL COMPUTATION OF SEARCH DIRECTIONS

To test the method presented in the previous section, I computed the gradient of the objective functions for two synthetic data sets. The first assumes a uniform slowness error, whereas the second assumes a localized velocity error. In both cases the correct velocity was constant and equal to 1 km/s. Both sets of experiments were computed assuming 50 shots at the surface (actually at 50 meters depth) spaced 80 meters apart, and the receiver array at a depth of 950 meters. The data modeled with a uniform slowness error were recorded by a split-spread array of 512 receivers spaced 10 meters apart, whereas the data modeled with a localized velocity error were recorded by a split-spread array of 720 receivers spaced 10 meters apart. In both cases, the receiver array was moving along with the shots, to maintain a uniform offset coverage for each shot. For the sake of clarity, for all the cases the figures display the search directions instead of the gradients. Also notice that I am showing separately the search directions computed for each term of the proposed objective function, without

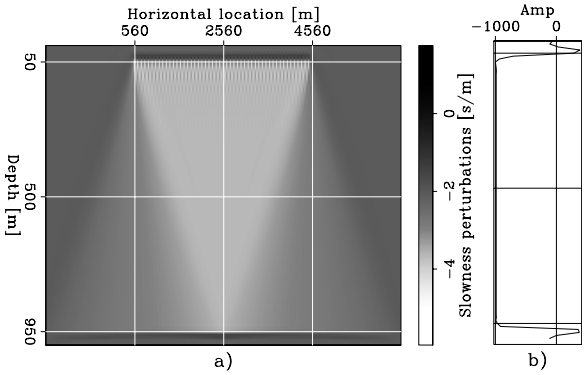

Figure 1: Search direction computed using conventional full-waveform inversion with a positive small (0.1%) uniform slowness error: a) slowness perturbations and b) slowness perturbations averaged over the horizontal direction. [**CR**]

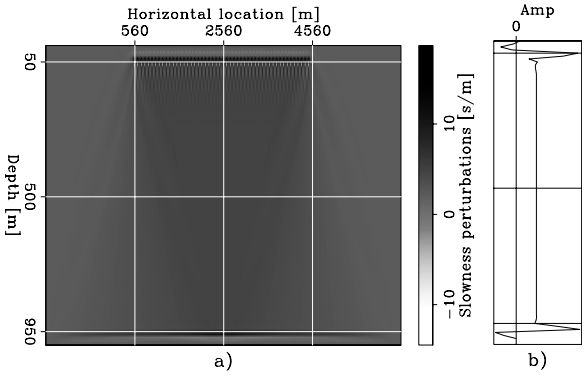

Figure 2: Search direction computed using conventional full-waveform inversion with a positive large (8.0%) uniform slowness error: a) slowness perturbations and b) slowness perturbations averaged over the horizontal direction. [**CR**]

showing their sum. At the moment of writing this report, I am not confident of the relative scaling between the gradient computed for the local objective function and the gradient computed for the global one.

Uniform velocity error

I compare the results obtained for the uniform velocity error case with the corresponding gradient obtained from conventional full-waveform inversion. To compute the search directions Δs for full-waveform inversion I applied the following expression:

$$\Delta s = -\sum_{x_s}\sum_{x_g} \frac{\partial \widetilde{P}'}{\partial s} \left[\widetilde{P}\left(t, x_g, x_s; s_0\right) - P_D\left(t, x_g, x_s\right)\right]. \tag{26}$$

Figures 1–2 show the search direction obtained by applying expression 26. Panels a) show the velocity perturbations for each z and x of the model; panels b) show the velocity perturbation averaged over the horizontal direction. As expected, the search direction computed using expression 26 is pointing in the correct direction when the velocity error is small (0.1%), Figure 1. However, when the velocity error is large the search direction has the wrong polarity (8.0%), Figure 2.

Figure 2 illustrates the limitations of full-waveform inversion when applied to estimating the background velocity. I therefore applied the proposed method to the

case with a large velocity error, to demonstrate that it overcomes these limitations of full-waveform inversion.

Figure 3 shows the derivatives of the objective function with respect to the moveout parameters computed using equations 16-17 and 18-19. The plot in panel a) displays the derivatives with respect to the local curvature, μ_C; the plot in panel b) displays the derivatives with respect to the time shifts of the beam centers, μ_θ. In both cases the derivatives are plotted as a function of the beam center coordinate, $\overline{x}$, for the source location at $x_s = 2.56$ km; that is, in the middle of the model. As visible from the figure, I clipped to zero the derivatives for both parameters outside of the -1.7 km $\leq \overline{x} \leq 1.7$ km range to avoid edge effects.

Figure 3a clearly shows that the local curvature is an appropriate parameters to measure the effects of large-scale velocity errors. In contrast, the alternating signs of the time shift derivatives shown in Figure 3b demonstrates that such perturbations are not well captured by the global moveout parameters.

These observations are confirmed by Figures 4 and 5. Figure 4 shows the search direction computed using equation 24. In contrast with the search direction computed with conventional waveform inversion (Figure 2), the search direction shown in Figure 4 has the correct polarity. It provides a good search direction, similar to the one provided by full-waveform inversion with the small velocity error that is shown in Figure 1.

The search direction obtained by computing the gradient of the global objective function using equation 25 is shown in Figure 5. It shows strong edge effects and its horizontal average (Figure 5b)) has the wrong polarity.

Localized velocity error

To analyze the interplay between the local and the global objective function, I computed the search directions in the case of a spatially localized slowness error. As for the previous examples, I computed the search directions provided by the gradient of both the local and the global objective functions. Figure 6 shows the slowness error that was assumed for the background slowness model s_0. In addition to the horizontal average, this figure (and the ones that follow) shows the vertical average in panel c) at the bottom of the figure.

Similarly to Figure 3, Figure 7, shows the derivatives of the objective function with respect to the moveout parameters computed using equations 16-17 and 18-19. The plot in panel a) displays the derivatives with respect to the local curvature, μ_C; the plot in panel b) displays the derivatives with respect to the time shifts of the beam centers, μ_θ. In both cases the derivatives are plotted as a function of the beam center coordinate, $\overline{x}$, for the source location at $x_s = 2.56$ km; that is, in the middle of the model. As visible from the figure, I clipped to zero the derivatives for both parameters outside of the -2.5 km $\leq \overline{x} \leq 2.5$ km range to avoid edge effects.

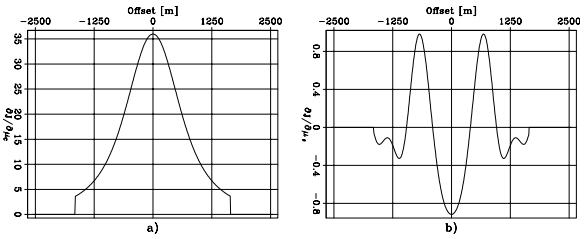

Figure 3: Derivatives of the objective functions with respect to the moveout parameters plotted as a function of beam center coordinate, $\overline{x}$, for the source location at $x_s = 2.56$ km: a) derivatives of J_{Local} with respect to to the local beam curvatures, μ_C, b) derivatives of J_{Global} with respect to the time shifts of the beam centers, μ_θ. [**CR**]

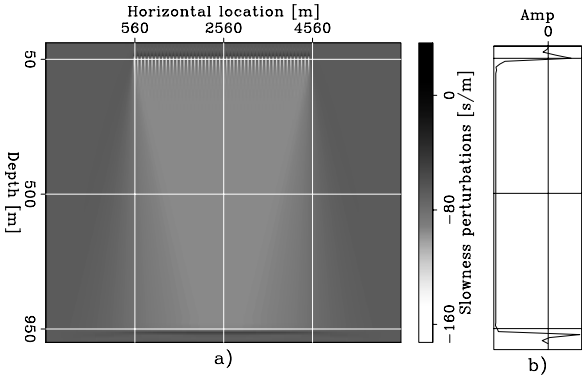

Figure 4: Search direction computed using the gradient of the local objective function J_{Local} with a positive large (8.0%) uniform slowness error: a) slowness perturbations and b) slowness perturbations averaged over the horizontal direction. [**CR**]

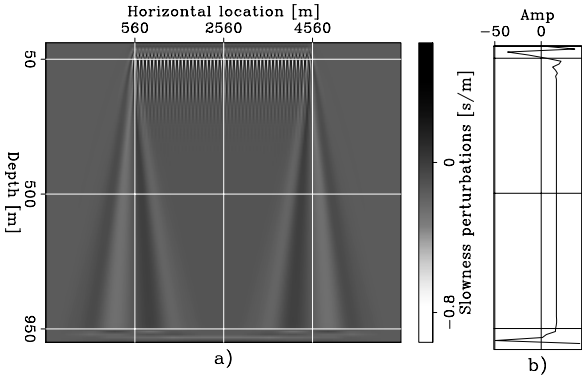

Figure 5: Search direction computed using the gradient of the global objective function J_{Global} with a positive large (8.0%) uniform slowness error: a) slowness perturbations and b) slowness perturbations averaged over the horizontal direction. [**CR**]

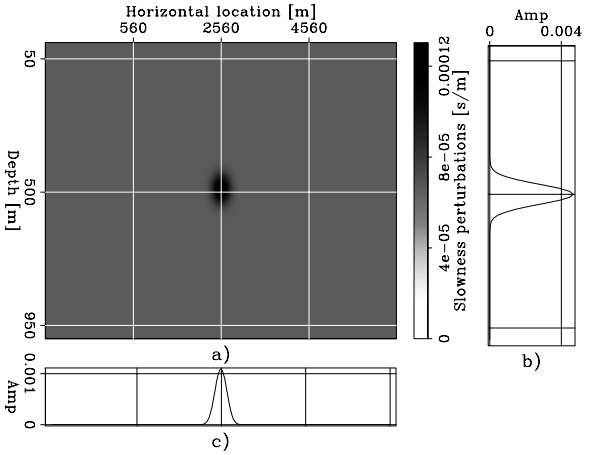

Figure 6: The slowness error that was assumed for the background slowness model s_0: a) slowness perturbations, b) slowness perturbations averaged over the horizontal direction, and c) slowness perturbations averaged over the vertical direction. [**ER**]

Figure 7a clearly shows that local curvature is not as an appropriate parameter for measuring the effects of small-scale velocity errors as it is for large-scale ones. The alternating signs of the objective-function derivatives causes the search direction to be highly oscillating as well. In contrast, the time-shifts derivatives shown in Figure 7b shows a large anomaly corresponding the localized velocity error and will provide useful slowness updates to localize the velocity anomaly.

These observations are confirmed by Figure 8 and 9. Figure 8 shows the search direction computed using equation 24. As expected it oscillating around the velocity anomaly. Figure 9 shows instead a nicely localized anomaly with the correct sign. It is useful to notice that the horizontal averages of the search directions shown in Figure 8 and 9 have opposite polarity. The average of the search direction provided by the global component has the wrong polarity, except at the depth of the anomaly. Whereas the average of the search direction provided by the local component has the correct polarity. This observation confirms the analysis that local curvature carries more reliable information for the long-wavelength component of the velocity updates than the global time shifts, as observed when analyzing the uniform velocity error example.

Another interesting observation can be made by computing the ratio between the amplitudes of the slowness updates in the two cases. In the case of uniform error, the update computed from the local curvature is larger than the other by approximately a factor of 200. In contrast, in the case of the localized anomaly, the ratio between amplitudes is only about 10. This difference in relative amplitudes confirms that the two components of the objective function switch in relative importance between the two cases.

CONCLUSIONS

To reliably estimate velocity using wavefield operators, I introduce a new objective function that rewards flatness of the correlation between source wavefield and receiver

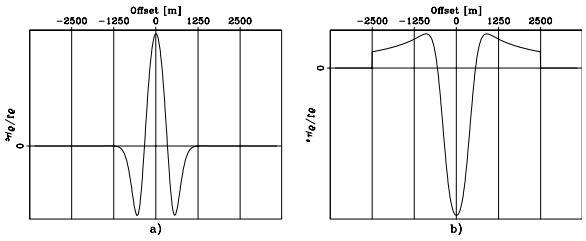

Figure 7: Derivatives of the objective functions with respect to the moveout parameters plotted as a function of beam center coordinate, $\overline{x}$, for the source location at $x_s = 2.56$ km: a) derivatives of J_{Local} with respect to to the local beam curvatures, μ_C, b) derivatives of J_{Global} with respect to the time shifts of the beam centers, μ_θ. [**CR**]

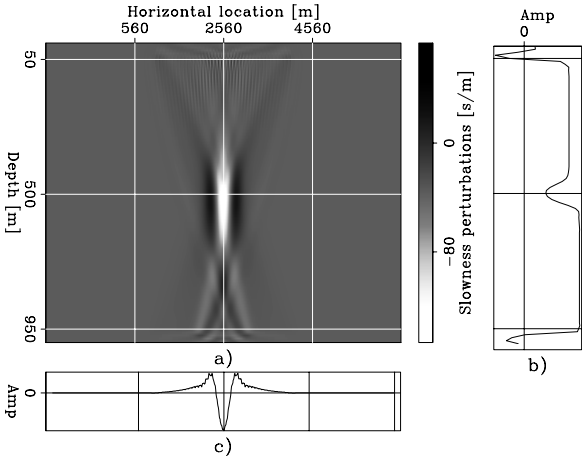

Figure 8: Search direction computed using the gradient of the local objective function J_{Local} with the localized velocity error shown in Figure 6: a) slowness perturbations, b) slowness perturbations averaged over the horizontal direction, and c) slowness perturbations averaged over the vertical direction. [**CR**]

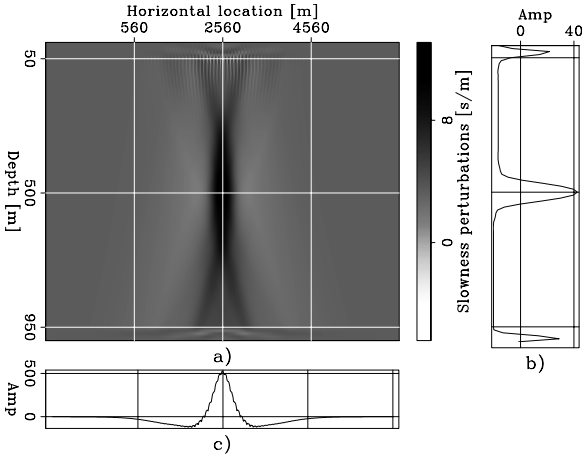

Figure 9: Search direction computed using the gradient of the global objective function J_{Global} with the localized velocity error shown in Figure 6: a) slowness perturbations, b) slowness perturbations averaged over the horizontal direction, and c) slowness perturbations averaged over the vertical direction. [**CR**]

wavefield. The proposed objective function is maximized as a function of the slowness model through the application of residual moveout operators to the correlation. The first term of the objective function measures the power of the stack over local beams as a function of the local beam curvature. Maximization of this first term ensures global convergence in presence of large velocity errors. The second term maximizes the global power of the stack as a function of time shifts applied to the local stack over the beams. Maximization of this second term helps the estimation of localized velocity errors.

I tested the application of the proposed objective function by computing its gradients for two simple problems: the estimation of a large and spatially uniform velocity error and the estimation of a spatially localized velocity error. The computed search directions confirm the potential of the proposed method and illustrate the different roles played by the the local and the global terms of the objective function.

REFERENCES

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.

Luo, Y. and G. T. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.

Sava, P., 2004, Migration and velocity analysis by wavefield extrapolation: PhD thesis, Stanford University.

Sava, P. and B. Biondi, 2004a, Wave-equation migration velocity analysis—I: Theory: Geophysical Prospecting, **52**, 593–623.

———, 2004b, Wave-equation migration velocity analysis—II: Examples: Geophysical Prospecting, **52**, 607–623.

Shen, P., 2004, Wave equation migration velocity analysis by differential semblance optimization: PhD thesis, Rice University.

Shen, P., W. W. Symes, S. Morton, and H. Calandra, 2005, Differential semblance velocity analysis via shot profile migration: SEG, Expanded Abstracts, **24**, 2249–2252.

Symes, W. W. and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.

APPENDIX A

DETAILS OF GRADIENT COMPUTATION

In this appendix I present the analytical development needed to derive equations 22-23 from equations 20-21.

Equation 20 can be rewritten as

$$\left.\frac{\partial \boldsymbol{\mu}_{\overline{x}}}{\partial s}\right|_{\mu_{\overline{x}}=\bar{\mu}_{\overline{x}}} = -\frac{\frac{\partial \dot{J_{\mathrm{FL}}}(\boldsymbol{\mu}_{\overline{x}})}{\partial s}}{\frac{\partial \dot{J_{\mathrm{FL}}}(\boldsymbol{\mu}_{\overline{x}})}{\partial \boldsymbol{\mu}_{\overline{x}}}}$$
$$= -\frac{\left\langle \left.\frac{\partial \mathcal{M}_{\overline{x}}[\theta(\boldsymbol{\mu}_{\overline{x}}),\mathbf{B}_{\overline{x}}C(\tau;s_0)]}{\partial \boldsymbol{\mu}_{\overline{x}}}\right|_{\boldsymbol{\mu}_{\overline{x}}=\bar{\boldsymbol{\mu}}_{\overline{x}}}, \mathbf{B}_{\overline{x}}\mathbf{P_D}\frac{\partial \widetilde{P}}{\partial s}\right\rangle}{\left\langle \left.\frac{\partial^2 \mathcal{M}_{\overline{x}}[\theta(\boldsymbol{\mu}_{\overline{x}}),\mathbf{B}_{\overline{x}}C(\tau;s_0)]}{\partial \boldsymbol{\mu}_{\overline{x}}^2}\right|_{\boldsymbol{\mu}_{\overline{x}}=\bar{\boldsymbol{\mu}}_{\overline{x}}}, \mathbf{B}_{\overline{x}}C\left(\tau;s\right)\right\rangle},$$

where,

$$\begin{aligned}\frac{\partial^2 \mathcal{M}_{\overline{x}}\left[\theta\left(\boldsymbol{\mu}_{\overline{x}}\right),\mathbf{B}_{\overline{x}}C\left(\tau;s_0\right)\right]}{\partial \boldsymbol{\mu}_{\overline{x}}^2} &= \mathcal{M}_{\overline{x}}\left[\theta\left(\boldsymbol{\mu}_{\overline{x}}\right),\mathbf{B}_{\overline{x}}\,\dot{C}\left(\tau;s_0\right)\right]\frac{\partial^2\theta}{\partial \boldsymbol{\mu}_{\overline{x}}^2}\\ &+ \mathcal{M}_{\overline{x}}\left[\theta\left(\boldsymbol{\mu}_{\overline{x}}\right),\mathbf{B}_{\overline{x}}\,\ddot{C}\left(\tau;s_0\right)\right]\frac{\partial\theta}{\partial \boldsymbol{\mu}_{\overline{x}}},\end{aligned}$$

and in which $\ddot{C}\left(\tau\right) = C\left(\tau\right)\left[\widetilde{P}\left(t\right),\ddot{P}_D\left(t\right)\right]$. Given the moveout parametrization expressed in 6, $\partial^2\theta/\partial\boldsymbol{\mu}_{\overline{x}}^2 = 0$ and the previous expression simplifies into the following:

$$\frac{\partial^2 \mathcal{M}_{\overline{x}}\left[\theta\left(\boldsymbol{\mu}_{\overline{x}}\right),\mathbf{B}_{\overline{x}}C\left(\tau;s_0\right)\right]}{\partial \boldsymbol{\mu}_{\overline{x}}^2} = \mathcal{M}_{\overline{x}}\left[\theta\left(\boldsymbol{\mu}_{\overline{x}}\left(s\right)\right),\mathbf{B}_{\overline{x}}\,\ddot{C}\left(\tau;s_0\right)\right]\frac{\partial\theta}{\partial \boldsymbol{\mu}_{\overline{x}}}.$$

Consequently, the general expression for the gradient of the local moveout parameters with respect to the slowness model is:

$$\left.\frac{\partial \boldsymbol{\mu}_{\overline{x}}}{\partial s}\right|_{\mu_{\overline{x}}=\bar{\mu}_{\overline{x}}} = -\frac{\left\langle \mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right),\mathbf{B}_{\overline{x}}\,\dot{C}\left(\tau;s_0\right)\right]\frac{\partial\theta}{\partial \boldsymbol{\mu}_{\overline{x}}},\mathbf{B}_{\overline{x}}\mathbf{P_D}\frac{\partial \widetilde{P}}{\partial s}\right\rangle}{\left\langle \mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right),\mathbf{B}_{\overline{x}}\,\ddot{C}\left(\tau;s_0\right)\right]\frac{\partial\theta}{\partial \boldsymbol{\mu}_{\overline{x}}},\mathbf{B}_{\overline{x}}C\left(\tau;s\right)\right\rangle}.$$

When $\bar{\boldsymbol{\mu}}_{\overline{x}} = 0$, the general expression further simplifies into:

$$\left.\frac{\partial \boldsymbol{\mu}_{\overline{x}}}{\partial s}\right|_{\mu_{\overline{x}}=0} = -\frac{\left\langle \mathbf{B}_{\overline{x}}\,\dot{C}\left(\tau;s_0\right)\frac{\partial\theta}{\partial \boldsymbol{\mu}_{\overline{x}}},\mathbf{B}_{\overline{x}}\mathbf{P_D}\frac{\partial \widetilde{P}}{\partial s}\right\rangle}{\left\langle \mathbf{B}_{\overline{x}}\,\ddot{C}\left(\tau;s_0\right)\frac{\partial\theta}{\partial \boldsymbol{\mu}_{\overline{x}}},\mathbf{B}_{\overline{x}}C\left(\tau;s\right)\right\rangle}. \tag{A-1}$$

Similar derivation can be developed for the derivative of the global moveout parameters with respect to slowness. Equation 21 can be rewritten as:

$$\left.\frac{\partial \boldsymbol{\mu}}{\partial s}\right|_{\mu=\bar{\mu}} = -\frac{\frac{\partial \dot{J_{\mathrm{FG}}}(\boldsymbol{\mu})}{\partial s}}{\frac{\partial \dot{J_{\mathrm{FG}}}(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}}$$
$$= -\frac{\left\langle \left.\frac{\partial \mathcal{M}\{\theta(\boldsymbol{\mu}),\boldsymbol{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}[\theta(\bar{\boldsymbol{\mu}}_{\overline{x}}),\mathbf{B}_{\overline{x}}C(\tau;s_0)]\}}{\partial \boldsymbol{\mu}}\right|_{\boldsymbol{\mu}=\bar{\boldsymbol{\mu}}}, \boldsymbol{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right),\mathbf{B}_{\overline{x}}\mathbf{P_D}\frac{\partial \widetilde{P}}{\partial s}\right]\right\rangle}{\left\langle \left.\frac{\partial^2 \mathcal{M}\{\theta(\boldsymbol{\mu}),\boldsymbol{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}[\theta(\bar{\boldsymbol{\mu}}_{\overline{x}}),\mathbf{B}_{\overline{x}}C(\tau;s_0)]\}}{\partial \boldsymbol{\mu}^2}\right|_{\boldsymbol{\mu}=\bar{\boldsymbol{\mu}}}, \boldsymbol{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right),\mathbf{B}_{\overline{x}}C\left(\tau;s\right)\right]\right\rangle},$$

where,

$$\frac{\partial^2 \mathcal{M}\left\{\theta\left(\boldsymbol{\mu}\right), \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}C\left(\tau; s_0\right)\right]\right\}}{\partial \boldsymbol{\mu}^2} = \mathcal{M}\left\{\theta\left(\boldsymbol{\mu}\right), \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}\,\dot{C}\left(\tau; s_0\right)\right]\right\}\frac{\partial^2 \theta}{\partial \boldsymbol{\mu}^2} + \mathcal{M}\left\{\theta\left(\boldsymbol{\mu}\right), \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}\,\ddot{C}\left(\tau; s_0\right)\right]\right\}\frac{\partial \theta}{\partial \boldsymbol{\mu}}.$$

Given the moveout parametrization in expressed in 10, $\partial^2\theta/\partial\boldsymbol{\mu}^2 = 0$ and the previous expression simplifies into:

$$\frac{\partial^2 \mathcal{M}\left\{\theta\left(\boldsymbol{\mu}\right), \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}C\left(\tau; s_0\right)\right]\right\}}{\partial \boldsymbol{\mu}^2} = \mathcal{M}\left\{\theta\left(\boldsymbol{\mu}\right), \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}\,\ddot{C}\left(\tau; s_0\right)\right]\right\}\frac{\partial \theta}{\partial \boldsymbol{\mu}}.$$

The general expression for the gradient of the global moveout parameters with respect to the slowness model is:

$$\left.\frac{\partial \boldsymbol{\mu}}{\partial s}\right|_{\mu=\bar{\mu}} = -\frac{\left\langle \mathcal{M}\left\{\theta\left(\bar{\boldsymbol{\mu}}\right), \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}\,\dot{C}\left(\tau; s_0\right)\right]\right\}\frac{\partial \theta}{\partial \boldsymbol{\mu}}, \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}\mathbf{P}_{\mathbf{D}}\frac{\partial \widetilde{P}}{\partial s}\right]\right\rangle}{\left\langle \mathcal{M}\left\{\theta\left(\bar{\boldsymbol{\mu}}\right), \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}\,\ddot{C}\left(\tau; s_0\right)\right]\right\}\frac{\partial \theta}{\partial \boldsymbol{\mu}}, \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathcal{M}_{\overline{x}}\left[\theta\left(\bar{\boldsymbol{\mu}}_{\overline{x}}\right), \mathbf{B}_{\overline{x}}C\left(\tau; s\right)\right]\right\rangle}.$$

When $\bar{\boldsymbol{\mu}}_{\overline{x}} = 0$ and $\bar{\boldsymbol{\mu}} = 0$ the general expression further simplifies into:

$$\left.\frac{\partial \boldsymbol{\mu}}{\partial s}\right|_{\mu_{\overline{x}}=0,\mu=0} = -\frac{\left\langle \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathbf{B}_{\overline{x}}\,\dot{C}\left(\tau; s_0\right)\frac{\partial \theta}{\partial \boldsymbol{\mu}}, \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathbf{B}_{\overline{x}}\mathbf{P}_{\mathbf{D}}\frac{\partial \widetilde{P}}{\partial s}\right\rangle}{\left\langle \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathbf{B}_{\overline{x}}\,\ddot{C}\left(\tau; s_0\right)\frac{\partial \theta}{\partial \boldsymbol{\mu}}, \mathbf{\Sigma}_{\overline{x}}\mathbf{S}_{\overline{x}}\mathbf{B}_{\overline{x}}C\left(\tau; s\right)\right\rangle}. \tag{A-2}$$

Research Personnel

Ali Almomin graduated from Texas A&M University in 2007 with a BS in Geophysics. Then, he joined Saudi Aramco and worked in several exploration and research departments with a focus on 3D seismic processing and near surface imaging. He joined Stanford Exploration Project in 2009 to pursue a PhD in Geophysics and is currently working on seismic tomography. He is a member of SEG, EAGE, and SPE.

Gboyega Ayeni received his B.Sc. in Applied Geophysics from Obafemi Awolowo University, Nigeria in 2004. He was a Shell sholar at University of Leeds, where he obtained an M.Sc with Distinction in Exploration Geophysics. Gboyega joined SEP in September 2006 to work towards his Ph.D in Geophysics. He is a member of SEG, EAGE, AGU, SPE and AAPG.

Ohad Barak received a B.Sc. (2006) and an M.Sc. (2009) in Geophysics from Tel-Aviv University. In 2008 he joined Paradigm Geophysical and worked there as a developer. He joined SEP in 2009 and is currently pursuing a Ph.D. in geophysics at Stanford University, and a longer biography.