Wave-equation migration velocity analysis by residual moveout fitting

Biondo Biondi

ABSTRACT

Flatness in migrated angle-domain common image gathers is an effective criterion for measuring migration-velocity accuracy. An objective function that measures the power of the stack as a function of residual-moveout parameters directly, and indirectly as a function of migration velocity, can be robustly maximized by using a gradient-based method. This paper presents a method to compute the gradient of this objective function by wave-equation operators. The proposed algorithm has the additional advantage of not requiring the picking of the residual-moveout parameters.

INTRODUCTION

In this paper I build on the framework I presented in Biondi (2010). In that report I presented a tomographic velocity estimation that aims to maximize image focusing using wave-equation operators. In SEP 140 I developed the theory and showed the results of numerical tests for a transmission tomography problem, because transmission tomography is simpler than reflection tomography. In this paper I extend that theory to the broader application of migration velocity analysis (MVA).

Conventional MVA methods are often based on the maximization of the stack power of migrated angle-domain common image gathers. However, direct maximization of the stack power as a function of velocity by wave-equation operators has well-known convergence problems (Chavent and Jacewitz, 1995; Biondi, 2006; Symes, 2008). To overcome these challenges, I propose to maximize stack power as a function of residual-moveout parameters, instead of directly as a function of velocity. In turn, the residual-moveout parameters are defined as solutions of fitting problems that maximize the correlation between the moved-out gathers and the gathers obtained by migrating the recorded data with the given velocity. These fitting problems can be quickly solved by using one-parameter gradient methods and thus do not require the explicit picking of residual-moveout parameters. The avoidance of parameter picking is an important advantage with respect to conventional wave-equation MVA methods (Biondi and Sava, 1999; Sava and Biondi, 2004a,b; Sava, 2004).

This paper presents just the theoretical development without any numerical examples illustrating the proposed method. I plan to present the application and the testing of this theory in upcoming reports.

THEORY

In wave-equation migration, as for example reverse-time migration, the image is computed from the back-propagated receiver wavefield, $P_g(t, \vec{x}, x_s; s)$, and the forward-propagated source wavefield, $P_s(t, \vec{x}, x_s; s)$, where t is the recording time, $\vec{x} = z\vec{z_0} + x\vec{x_0}$ is the model-coordinate vector, x_s is the source position at the surface, and $s(\vec{x})$ is the slowness model.

The prestack image, $I_h(\vec{x}, x_h)$, is computed as the zero lag of the temporal cross-correlation between the spatially-shifted back-propagated receiver wavefield and forward-propagated source wavefield as (Rickett and Sava, 2002):

$$I_h(\vec{x}, \vec{x_h})[P_s(t, \vec{x}), P_g(t, \vec{x})] = \sum_t \sum_{x_s} P_g(t, \vec{x} - \vec{x_h}, x_s) P_s(t, \vec{x} + \vec{x_h}, x_s), \qquad (1)$$

where $\vec{x_h} = x_h\vec{x_0}$ is the *half subsurface offset*, which in this paper I will assume to be horizontal, but it does not need to be in the general case (Biondi and Symes, 2004).

The prestack image as a function of subsurface offset can be transformed to an image as a function of reflection aperture angle, $I_\gamma(\vec{x}, \gamma)$ by using a linear operator $\mathbf{\Gamma}$ (Sava and Fomel, 2003). In matrix notation, if $\mathbf{I}_h$ is a $N_{\vec{x}}N_h \times 1$ matrix and $\mathbf{I}_\gamma$ is a $N_{\vec{x}}N_\gamma \times 1$ matrix, the image transformation from subsurface offset into the angle domain can be expressed as:

$$\mathbf{I}_\gamma = \mathbf{\Gamma}\mathbf{I}_h. \qquad (2)$$

I introduce an objective function that maximizes the flatness of the angle-domain image along the aperture-angle axis at all spatial locations $\vec{x}$. This objective function aims at maximizing image flatness not directly as a function of the slowness, but indirectly through the application of an angle-domain moveout operator $\mathcal{M}_\gamma$, which depends on the column vector of $N_\mu = N_{\vec{x}}$ moveout parameters $\boldsymbol{\mu}_{\vec{x}}$.

I define the application of the moveout operators $\mathcal{M}_\gamma$ to a prestack image computed by equations 1 and 2 with a background slowness $\bar{s}$, as

$$I_\gamma\left(\vec{x} + \vec{\zeta}(\boldsymbol{\mu}_{\vec{x}}), \gamma; \bar{s}\right) = \mathcal{M}_\gamma[\boldsymbol{\mu}_{\vec{x}}] I_\gamma(\vec{x}, \gamma; \bar{s}), \qquad (3)$$

where $\vec{\zeta}(\boldsymbol{\mu}_{\vec{x}}) = \zeta(\boldsymbol{\mu}_{\vec{x}})\vec{z_0}$ are the moveout shifts, assumed here to be simple depth shifts. The operator $\mathcal{M}_\gamma$ is linear with respect to the input image, but it is nonlinear with respect to the vector of moveout parameters $\boldsymbol{\mu}_{\vec{x}}$. In matrix notation, the application of the moveout operator to the background image $\bar{\mathbf{I}}_\gamma$ can be expressed as $\mathcal{M}_\gamma[\boldsymbol{\mu}_{\vec{x}}]\bar{\mathbf{I}}_\gamma$.

I further define the stacking operator $\mathbf{S}_\gamma$ that sums the image along the aperture-angle axis γ. I can now introduce the objective function that measures the flatness of the image as:

$$J(\boldsymbol{\mu}_{\vec{x}}(\mathbf{s})) = \frac{1}{2}\left\|\mathbf{S}_\gamma\mathcal{M}_\gamma[\boldsymbol{\mu}_{\vec{x}}(\mathbf{s})]\bar{\mathbf{I}}_\gamma\right\|_2^2, \qquad (4)$$

where $\mathbf{s}$ is the slowness vector. This objective function is not a direct function of $\mathbf{s}$, but it depends on it indirectly through the moveout parameters $\boldsymbol{\mu}_{\vec{x}}$. The dependency of the moveout parameters from the slowness function is not defined explicitly; these parameters are defined as the solutions of $N_{\vec{x}}$ independent fitting problems, one for each spatial location in the image.

The fitting problems maximize the zero lag of the cross-correlation between the prestack image computed for a realization of the slowness vector $\mathbf{s}$ and the moved-out image computed with the background slowness $\bar{\mathbf{s}}$. For the sake of keeping the notation as compact as possible, I combine the $N_{\vec{x}}$ independent fitting problems into one by defining the following objective function:

$$J_{\mathrm{F}}\left(\boldsymbol{\mu}_{\vec{x}}\left(\mathbf{s}\right)\right) = \mathbf{S}_{\vec{x}} \left\langle \mathcal{M}_{\gamma}\left[\boldsymbol{\mu}_{\vec{x}}\right] \bar{\mathbf{I}}_{\gamma}, \mathbf{I}_{\gamma}\left(\mathbf{s}\right) \right\rangle_{\gamma}, \tag{5}$$

where with the notation $\langle \mathbf{x}, \mathbf{y} \rangle_{\gamma}$ I indicate the ensemble of inner products between the image vectors $\mathbf{x}$ and $\mathbf{y}$ which spans only the aperture-angle axis γ; the result of these inner products is a vector of length $N_{\vec{x}}$. The stacking operator $\mathbf{S}_{\vec{x}}$ sums the elements of this vector to combine the objective functions into one.

The vector of moveout parameters is therefore the solutions of the following maximization problem:

$$\max_{\boldsymbol{\mu}_{\vec{x}}} J_{\mathrm{F}}\left(\boldsymbol{\mu}_{\vec{x}}\left(\mathbf{s}\right)\right). \tag{6}$$

For velocity estimation in the angle domain, an effective parametrization of the moveout is the "curvature" $\mu_{\vec{x}}$, that defines the following moveout equation

$$\zeta\left(\mu_{\vec{x}}\right) = \mu_{\vec{x}} \tan^2 \gamma. \tag{7}$$

Notice that when the slowness $\mathbf{s}$ is equal to the background slowness $\bar{\mathbf{s}}$, the corresponding best-fitting moveout parameters $\bar{\mu}_{\vec{x}}$ are obviously the ones corresponding to no moveout; that is, $\zeta\left(\bar{\mu}_{\vec{x}}\right) = 0$, and consequently $\bar{\mu}_{\vec{x}} = 0$.

Gradient of the objective function

I plan to solve the optimization problem defined in 4 by a gradient-based optimization algorithm. Therefore, the development of an algorithm for efficiently computing the gradient of the objective function with respect to slowness is an essential step to make the method practical. In this section I outline the methodology to compute the gradient, and I leave some of the details to Appendix A.

The gradient is computed using the chain rule. The first term of the chain is the derivative of the objective function in equation 4 with respect the moveout parameters. The second term is the derivatives of the moveout parameters with respect to slowness that are computed from the objective function 5.

Derivatives of objective function (J) with respect to moveout parameters ($\boldsymbol{\mu}_{\vec{x}}$)

The derivatives of 4 with respect to the vector of moveout parameters is easily evaluated using the following expression:

$$\frac{\partial J'}{\boldsymbol{\mu}_{\vec{x}}} = \frac{\partial \mathcal{M}_\gamma{}'}{\partial \boldsymbol{\mu}_{\vec{x}}} \mathbf{S}_\gamma' \mathbf{S}_\gamma \mathcal{M}_\gamma \left[\bar{\boldsymbol{\mu}}_{\vec{x}}\right] \bar{\mathbf{I}}_\gamma. \tag{8}$$

The linear operator $\frac{\partial \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}}$ can be represented as a $N_{\vec{x}} N_\gamma \times N_{\mu_{\vec{x}}}$ matrix. The elements of this matrix are given by:

$$\frac{\partial \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}} \left(\vec{x}, \gamma, \mu_{\vec{x}}\right) = \underbrace{\mathcal{M}_\gamma \left[\boldsymbol{\mu}_{\vec{x}}\right] \dot{I}_\gamma \left(\vec{x}, \gamma; \bar{s}\right)}_{\mathrm{I}} \underbrace{\frac{\partial \zeta}{\partial \mu_{\vec{x}}}}_{\mathrm{II}}. \tag{9}$$

The first term (I) is given by the depth-derivative of the image $\partial I_\gamma \left(\vec{x}, \gamma; \bar{s}\right) / \partial z$ after moveout. This term can be numerically evaluated by applying to the moved-out image a finite-difference approximation of the first-derivative operator. The second term (II) is different from zero only when the spatial coordinate $\vec{x}$ of the image element $I_\gamma \left(\vec{x}, \gamma\right)$ is the same as the coordinate corresponding to the moveout parameter $\mu_{\vec{x}}$. When they do, and for the choice of moveout parameters expressed in equation 7, we have $\partial \zeta / \partial \mu_{\vec{x}} = \tan^2 \gamma$.

The preceding expression simplifies when the gradient is evaluated for $\boldsymbol{\mu}_{\vec{x}} = 0$. This simplifying condition is actually always fulfilled unless the optimization algorithm includes inner iterations for fitting the moveout parameters using a linearized approach. Under this condition, equation 8 becomes

$$\left.\frac{\partial J}{\boldsymbol{\mu}_{\vec{x}}}\right|_{\boldsymbol{\mu}_{\vec{x}}=0}^{\prime} = \left.\frac{\partial \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}}\right|_{\boldsymbol{\mu}_{\vec{x}}=0}^{\prime} \mathbf{S}_\gamma' \mathbf{S}_\gamma \bar{\mathbf{I}}_\gamma, \tag{10}$$

and equation 9 becomes

$$\frac{\partial \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}} \left(\vec{x}, \gamma, \mu_{\vec{x}} = 0\right) = \dot{I}_\gamma \left(\vec{x}, \gamma; \bar{s}\right) \frac{\partial \zeta}{\partial \boldsymbol{\mu}_{\vec{x}}}. \tag{11}$$

Derivatives of moveout parameters ($\boldsymbol{\mu}_{\vec{x}}$) with respect to slowness ($\mathbf{s}$)

The evaluation of the derivatives of the moveout parameters with respect to slowness takes advantage of the fact that we need to evaluate the derivatives only at maxima for the objective function in equation 5. At the maxima, the objective function is stationary and thus its derivatives with respect to the moveout parameters are zero, and we can write:

$$\left.\frac{\partial J_{\mathrm{F}} \left(\boldsymbol{\mu}_{\vec{x}}\right)}{\partial \boldsymbol{\mu}_{\vec{x}}}\right|_{\boldsymbol{\mu}_{\vec{x}}=\bar{\boldsymbol{\mu}}_{\vec{x}}} = \dot{J}_{\mathrm{F}} \left(\bar{\boldsymbol{\mu}}_{\vec{x}}\right) = \mathbf{S}_{\vec{x}} \left\langle \left.\frac{\partial \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}}\right|_{\boldsymbol{\mu}_{\vec{x}}=\bar{\boldsymbol{\mu}}_{\vec{x}}}, \mathbf{I}_\gamma \right\rangle_\gamma = 0. \tag{12}$$

As discussed above, the derivatives in the second term (II) of equation 9 are different from zero only when the moveout coefficient $\mu_{\vec{x}}$ and the image element share the same spatial coordinate. Consequently, for each $\mu_{\vec{x}}$ there is only one $\vec{x}$ for which the inner products above are different from zero. Equation 12 can thus be simplified into:

$$\dot{J}_{\rm F}\left(\bar{\boldsymbol{\mu}}_{\vec{x}}\right) = \left\langle \left.\frac{\partial \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}}\right|_{\boldsymbol{\mu}_{\vec{x}}=\bar{\boldsymbol{\mu}}_{\vec{x}}}, \mathbf{I}_\gamma \right\rangle_\gamma = 0. \tag{13}$$

Using the rule for differentiating implicit functions, and taking advantage of the fact 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:

$$\left.\frac{\partial \boldsymbol{\mu}_{\vec{x}}}{\partial \mathbf{s}}\right|_{\mu_{\vec{x}}=\bar{\mu}_{\vec{x}}} = -\frac{\frac{\partial \dot{J}_{\rm F}(\boldsymbol{\mu}_{\vec{x}})}{\partial \mathbf{s}}}{\frac{\partial \dot{J}_{\rm F}(\boldsymbol{\mu}_{\vec{x}})}{\partial \boldsymbol{\mu}_{\vec{x}}}}. \tag{14}$$

From equation 13 and 14, the derivative of the local moveout parameters with respect to slowness is:

$$\left.\frac{\partial \boldsymbol{\mu}_{\vec{x}}}{\partial \mathbf{s}}\right|_{\mu_{\vec{x}}=\bar{\mu}_{\vec{x}}} = -\frac{\left\langle \left.\frac{\partial \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}}\right|_{\boldsymbol{\mu}_{\vec{x}}=\bar{\boldsymbol{\mu}}_{\vec{x}}}, \frac{\partial \mathbf{I}_\gamma}{\partial \mathbf{s}} \right\rangle_\gamma}{\left\langle \left.\frac{\partial^2 \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}^2}\right|_{\boldsymbol{\mu}_{\vec{x}}=\bar{\boldsymbol{\mu}}_{\vec{x}}}, \mathbf{I}_\gamma \right\rangle_\gamma}. \tag{15}$$

Appendix A presents the development for the expressions to compute the terms $\partial^2\mathcal{M}_\gamma/\partial\boldsymbol{\mu}_{\vec{x}}^2$ (A-3), and $\partial\mathbf{I}_\gamma/\partial\mathbf{s}$ (A-5).

Combining the derivatives in equation 15 with the derivatives in equations 10-11 we can easily compute the gradient of the objective function 4 with respect to slowness that can be written, when $\bar{\boldsymbol{\mu}}_{\vec{x}} = 0$, as:

$$\nabla J = -\underbrace{\left(\frac{\partial \mathbf{P}_s{}'}{\partial \mathbf{s}}\bar{\mathbf{P}}'_g + \frac{\partial \mathbf{P}_g{}'}{\partial \mathbf{s}}\bar{\mathbf{P}}'_s\right)\boldsymbol{\Gamma}'}_{\rm I}\underbrace{\frac{\partial \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}}}_{\rm II}\underbrace{\frac{\frac{\partial \mathcal{M}_\gamma{}'}{\partial \boldsymbol{\mu}_{\vec{x}}}\mathbf{S}'_\gamma\mathbf{S}_\gamma\bar{\mathbf{I}}_\gamma}{\frac{\partial^2 \mathcal{M}_\gamma{}'}{\partial \boldsymbol{\mu}_{\vec{x}}^2}\bar{\mathbf{I}}_\gamma}}_{\rm III}. \tag{16}$$

I will now examine the effects of each of the terms in equation 16 starting from the rightmost one. The third term (III) produces a scalar for each local curvature parameter $\mu_{\vec{x}}$. This scalar multiplies the image for each physical location after it has been differentiated in depth and scaled by $\partial\zeta/\partial\mu_{\vec{x}}$, as described by the second term (II). Notice that the phase introduced by the depth derivative of the image in (II) is crucial for the successful backprojection into the slowness model that is accomplished by the first term (I). In this term, first $\boldsymbol{\Gamma}'$ transforms the image from the aperture-angle domain into the subsurface-offset domain. The transformed image is scaled, horizontally shifted, and spread across the shot axis by the application of

$\bar{\mathbf{P}}'_s$ and $\bar{\mathbf{P}}'_g$. Finally, the adjoint of operators $\frac{\partial \mathbf{P}_s}{\partial \mathbf{s}}'$ and $\frac{\partial \mathbf{P}_g}{\partial \mathbf{s}}'$ backproject the image perturbations into the slowness model along the source wavepaths and the receiver wavepaths, respectively.

Computational cost

The computational cost and storage overheads for evaluating terms II and III in the gradient expression 16 are limited because only operations on the prestack image are required. On the contrary, the computation of term I is computationally more demanding. It requires the forward propagation and backward propagation of wavefields. The application of $\bar{\mathbf{P}}'_s$ and $\bar{\mathbf{P}}'_g$ requires the storage, and retrieval, of the source wavefield and receiver wavefield. Furthermore, to apply $\frac{\partial \mathbf{P}_s}{\partial \mathbf{s}}'$ and $\frac{\partial \mathbf{P}_g}{\partial \mathbf{s}}'$ we need to correlate the source and receiver wavefields with the wavefields generated by the image derivatives. In summary, the computational cost of one gradient computation of the proposed method can be roughly estimated to be twice the computational cost of one gradient computation of a full-waveform inversion algorithm. The factor of two occurs because two propagations are needed to backproject the image perturbations into the slowness model along both the source wavepaths and the receiver wavepaths.

The data-handling task could be simplified if the frequency-domain downward-continuation migration is used instead of reverse-time migration, because computations can be performed independently for each frequency. The adaptation of the theory presented in this paper to downward-continuation migration is fairly straightforward. It would only require to exchange expressions 1 and A-4 with the corresponding frequency-domain expressions.

CONCLUSIONS

The theoretical framework I presented in Biondi (2010) can be extended from transmission-tomography problems to MVA problems. The computational cost of the proposed method can be high, though the cost of one iteration is comparable with the cost of one iteration of full waveform inversion. Numerical tests of the gradient operator and of the complete velocity-estimation method are needed.

ACKNOWLEDGMENTS

I would like to thank Yang Zhang for useful suggestions to make my mathematical notation a little bit easier to follow.

REFERENCES

Almomin, A. and Y. Tang, 2010, Migration velocity analysis based on linearization of the two-way wave equation: SEP-Report, **142**, X–Y.

Biondi, B., 2006, 3D Seismic Imaging: Society of Exploration Geophysicists.

———, 2010, Wave-equation tomography by beam focusing: SEP-Report, **140**, 23–38.

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

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

Chavent, G. and C. A. Jacewitz, 1995, Determination of background velocities by multiple migration fitting: Geophysics, **60**, 476–490.

Rickett, J. and P. Sava, 2002, Offset and angle-domain common image-point gathers for shot-profile migration: Geophysics, **67**, 883–889.

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.

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

Symes, W. W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **55**, 765–804.

APPENDIX A

DETAILS OF GRADIENT COMPUTATION

In this appendix I present the analytical development needed to compute all the terms in equation 16. Equations 14 and 15 provide the expression for computing the derivatives of the moveout parameters with respect to slowness as:

$$\left.\frac{\partial \boldsymbol{\mu}_{\vec{x}}}{\partial \mathbf{s}}\right|_{\mu_{\vec{x}}=\bar{\mu}_{\vec{x}}} = -\frac{\frac{\partial \dot{J}_{\mathrm{F}}(\boldsymbol{\mu}_{\vec{x}})}{\partial \mathbf{s}}}{\frac{\partial \dot{J}_{\mathrm{F}}(\boldsymbol{\mu}_{\vec{x}})}{\partial \boldsymbol{\mu}_{\vec{x}}}} = -\frac{\left\langle \left.\frac{\partial \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}}\right|_{\boldsymbol{\mu}_{\vec{x}}=\bar{\boldsymbol{\mu}}_{\vec{x}}}, \frac{\partial \mathbf{I}_\gamma}{\partial \mathbf{s}} \right\rangle_\gamma}{\left\langle \left.\frac{\partial^2 \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}^2}\right|_{\boldsymbol{\mu}_{\vec{x}}=\bar{\boldsymbol{\mu}}_{\vec{x}}}, \mathbf{I}_\gamma \right\rangle_\gamma}, \tag{A-1}$$

where the elements of the matrix $\frac{\partial \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}}$ are computed using either equation 9 or equation 11, and the elements of the matrix $\frac{\partial^2 \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}^2}$ are given by

$$\frac{\partial^2 \mathcal{M}_\gamma}{\partial \boldsymbol{\mu}_{\vec{x}}^2}(\vec{x},\gamma,\mu_{\vec{x}}) = \mathcal{M}_\gamma\left[\boldsymbol{\mu}_{\vec{x}}\right] \dot{I}_\gamma\left(\vec{x},\gamma;\bar{\mathbf{s}}\right) \frac{\partial^2 \zeta}{\partial \mu_{\vec{x}}^2} + \mathcal{M}_\gamma\left[\boldsymbol{\mu}_{\vec{x}}\right] \ddot{I}_\gamma\left(\vec{x},\gamma;\bar{\mathbf{s}}\right) \frac{\partial \zeta}{\partial \mu_{\vec{x}}}.$$

In this last expression $\ddot{I}_\gamma\left(\vec{x},\gamma;\bar{\mathbf{s}}\right)=\partial^2 I_\gamma\left(\vec{x},\gamma;\bar{\mathbf{s}}\right)/\partial z^2$. Given the moveout parametrization expressed in 7, $\partial^2\zeta/\partial\mu_{\vec{x}}^2=0$ and the previous expression simplifies into the following:

$$\frac{\partial^2\mathcal{M}_\gamma}{\partial\boldsymbol{\mu}_{\vec{x}}^2}=\mathcal{M}_\gamma\left[\boldsymbol{\mu}_{\vec{x}}\right]\ddot{I}_\gamma\left(\vec{x},\gamma;\bar{s}\right)\frac{\partial\zeta}{\partial\mu_{\vec{x}}}. \tag{A-2}$$

Furthermore, when $\bar{\boldsymbol{\mu}}_{\vec{x}}=0$, equation A-2 further simplifies into:

$$\frac{\partial^2\mathcal{M}_\gamma}{\partial\boldsymbol{\mu}_{\vec{x}}^2}=\ddot{I}_\gamma\left(\vec{x},\gamma;\bar{s}\right)\frac{\partial\zeta}{\partial\mu_{\vec{x}}}. \tag{A-3}$$

The derivative of the image vector with respect to slowness, $\partial\mathbf{I}_\gamma/\partial\mathbf{s}$ are evaluated by applying the conventional wave-equation tomography operator that links perturbations in the slowness model to perturbations in the propagated wavefields by a first-order Born linearization of the wave equation.

Applying the chain rule to equation 1, and taking into account the offset-to-angle transformation 2, we can write

$$\frac{\partial I_\gamma\left(\vec{x},\gamma;s\right)}{\partial s}=$$
$$\boldsymbol{\Gamma}\sum_t\sum_{x_s}\left[\bar{P}_g\left(t,\vec{x}-\vec{x_h},x_s\right)\frac{\partial P_s\left(t,\vec{x}+\vec{x_h},x_s\right)}{\partial s}+\bar{P}_s\left(t,\vec{x}-\vec{x_h},x_s\right)\frac{\partial P_g\left(t,\vec{x}-\vec{x_h},x_s\right)}{\partial s}\right], \tag{A-4}$$

where the wavefields $\bar{P}_s$ and $\bar{P}_g$ are computed with the background slowness, and the wavefield derivatives $\partial P_s/\partial s$ and $\partial P_g/\partial s$ are computed by the conventional adjoint-state methodology that is at the basis of full waveform inversion and wave-equation tomography.

In more compact matrix notation the previous expression can be written as

$$\frac{\partial\mathbf{I}_\gamma}{\partial\mathbf{s}}=\boldsymbol{\Gamma}\left(\bar{\mathbf{P}}_g\frac{\partial\mathbf{P}_s}{\partial\mathbf{s}}+\bar{\mathbf{P}}_s\frac{\partial\mathbf{P}_g}{\partial\mathbf{s}}\right), \tag{A-5}$$

where the matrices $\bar{\mathbf{P}}_s$ and $\bar{\mathbf{P}}_g$ are composed of the wavefields for every source and depth level, and properly shifted in space by the subsurface offset. For the computation of the gradient, we need to apply the adjoint operator that is:

$$\frac{\partial\mathbf{I}_\gamma{}'}{\partial\mathbf{s}}=\left(\frac{\partial\mathbf{P}_s{}'}{\partial\mathbf{s}}\bar{\mathbf{P}}_g'+\frac{\partial\mathbf{P}_g{}'}{\partial\mathbf{s}}\bar{\mathbf{P}}_s'\right)\boldsymbol{\Gamma}'. \tag{A-6}$$

Almomin and Tang (2010) present an equivalent, but different, derivation of an algorithm to compute the application of the operator $\frac{\partial\mathbf{I}_\gamma}{\partial\mathbf{s}}$, (or its adjoint) to a vector of slowness perturbations (or image perturbations).