Probing the extended image volume

Tristan van Leeuwen[1], *Felix Herrmann*[1]

[1] *Dept. of Earth and Ocean sciences University of British Columbia Vancouver, BC, Canada*

SUMMARY

The prestack image volume can be defined as a cross-correlation of the source and receivers wavefields for non-zero space and time lags. If the background velocity is kinematically acceptable, this image volume will have its main contributions at zero lag, even for complex models. Thus, it is an ideal tool for wave-equation migration velocity analysis in the presence of strong lateral heterogeneity. In particular, it allows us to pose migration velocity analysis as a PDE-constrained optimization problem, where the goal is to minimize the energy in the image volume at non-zero lag subject to fitting the data approximately. However, it is computationally infeasible to explicitly form the whole image volume. In this paper, we discuss several ways to reduce the computational costs involved in computing the image volume and evaluating the focusing criterion. We reduce the costs for calculating the data by randomized source synthesis. We also present an efficient way to subsample the image volume. Finally, we propose an alternative optimization criterion and suggest a multiscale inversion strategy for wave-equation MVA.

INTRODUCTION

Full waveform inversion relies on fitting the data in a least-squares sense and may yield velocity models and images with a spectacular amount of detail. This success relies heavily on the availability of low frequencies and long offsets in the data and requires suitable optimization strategy (cf. Bunks et al., 1995; Pratt et al., 1996; Shin and Cha, 2009). Still, a kinematically acceptable starting model has to be provided. Especially for the deeper part of the model that is not sampled properly by diving waves. Such a starting model may in principle be obtained by wave-equation migration velocity analysis. Here, the velocity is selected to optimally focus an extended image volume. This image volume consists of the correlation of a source and receiver wavefield at non-zero space and time lag. The traditional RTM image is obtained by extracting the zero-lag section from the extended image volume according to the traditional imaging condition.

The idea of introducing redundant coordinates in the image domain goes back to the 'survey-sinking' concept of Claerbout (1985), and many variants have been considered over the years (MacKay and Abma, 1992; Berkhout, 1997; Rickett and Sava, 2002; Sava and Fomel, 2006; de Hoop et al., 2006; Yang and Sava, 2010). Stolk and de Hoop (2006) showed that such extended image volumes – in contrast to image volumes obtained by data-binning procedures– are essentially artifact-free, even for complex models. Shen and Symes (2008) formulated a PDE-constrained optimization problem based on these extended images (see also Mulder and van Leeuwen, 2008).

Just as the RTM image is the gradient used in FWI, the extended image volume can be interpreted as the gradient corresponding to an extended modeling operator where the velocity is replaced by a non-stationary convolution operator (Symes, 2008b). This extended velocity operator allows for non-physical interaction over distance. Hence, the physically feasible velocity is a diagonal operator and this may serve as non-linear extension of MVA. However, the formidable computational cost and memory requirements involved in applying and forming the prestack image volume prohibits application of this concept to large scale problems.

In this paper we discuss the use of randomized dimensionality reduction to reduce the costs involved in linearized wave-equation MVA. In conjunction with this we propose a multiscale inversion strategy. The main topics are:

- **Randomized source synthesis.** We randomly superimpose sources to reduce the costs of computing the wavefields that constitute the extended image. (Herrmann et al., 2009; Krebs et al., 2009; Haber et al., 2010; Neelamani et al., 2010)
- **Subsampling the image volume.** To avoid ever forming the image volume, we consider only calculating the action of the image volume *kernel* on a vector, without ever explicitly forming the whole image volume. When the vector is a unit vector corresponding to a point scatterer, we can interpret this as redatuming the source to the location of the scatterer and subsequently redatuming the receivers to all points in the subsurface. This allows us to efficiently subsample the image volume.
- **Multiscale inversion.** We propose an alternative criterion to measure the focusing of the extended images. The criterion is based on a Gaussian weighting function, which was earlier proposed for data-domain velocity analysis (van Leeuwen and Mulder, 2008). The width of the Gaussian can be used to control the scale at which the focusing is measured. In the extreme case of the width going to zero, the approach reduces to the stack-power criterion. Now, we can adapt the parameterization of the model to the degree of subsampling of the extended image and the width of the Gaussian weight. This immediately suggests a multiscale inversion strategy; start out with a very sparse subsampling of the extended image, a large width of the Gaussian and a very smooth medium parameterization, and gradually move to finer scales to add more detail to the velocity model.

The paper is organized as follows. First, we describe the non-linear extended modeling framework and formulate the corresponding constrained optimization problem. From this, we derive a constrained formulation for the traditional wave-equation MVA approach. We then discuss the use of simultaneous sources in this setting and show how we can efficiently compute the action of the extended image on a vector. We illustrate the interferometric interpretation and give an example of extended

images for a simple model. Finally, we discuss the multiscale inversion strategy and illustrate the multiscale behavior of the proposed focusing criterion with a numerical experiment on synthetic data.

THEORY

We represent data and wavefields as monochromatic matrices, where the column index serves as the shot-index (cf. Berkhout, 1985). The data are modeled as:

$$H[m]U = Q, \quad (1)$$

$$D \equiv F[m] = PU, \quad (2)$$

where $H = [\omega^2 m + \nabla^2]$ is the Helmholtz operator for frequency ω and squared-slowness m; Q represents the sources, and P samples the wavefield at the receiver locations. Symes (2008a) proposed to replace the velocity with a non-stationary convolution operator with kernel M, whose elements m_{ij} describe the interaction between gridpoints i and j. The resulting nonlinear formulation of MVA is then given by

$$\min_M ||A \odot M||^2 \text{ s.t. } ||F[M] - D||^2 \leq \sigma, \quad (3)$$

where $||\cdot||$ is used to denote both the ℓ_2 norm for vectors and the Frobenuis norm for matrices, $\odot$ denotes element-wise multiplication (the Hadamard product) and A typically penalizes off-diagonal energy (i.e., $a_{ij} \propto r_{ij}$, where r_{ij} is the Euclidean distance between gridpoints i and j), though other penalties are possible as we will see later on. The costs for modeling the data can be formidable, since M is a full matrix.

Separating M into a diagonal part, m, a perturbation ΔM and linearizing w.r.t m yields:

$$\min_{m,\Delta M} ||A \odot \Delta M||^2 \text{ s.t. } ||F[m] + \nabla F[m]\Delta M - D||^2 \leq \sigma, \quad (4)$$

where $\nabla F[m]$ is the linearized Born modeling operator. We can now turn this into an unconstrained optimization problem by eliminating the constraint approximately via

$$\Delta M[m] = \nabla F[m]^*(F[m] - D) = \sum_\omega \omega^2 \Re(VU^*), \quad (5)$$

where

$$H[m]U = Q, \quad (6)$$

$$H[m]^*V = P^*(PU - D). \quad (7)$$

This leads to the more traditional optimization formulation of wave-equation MVA (Shen and Symes, 2008; Symes, 2008b). Alternatively, we may consider a deconvolution of the wavefields (cf. Herrmann, 2009) or a linearized inversion of the constraint to arrive at a 'true amplitude' extended image.

The computational costs of the approach now lie in computing the wavefields (eq. 6, 7) and computing their outer product to form the extended image (eq. 5). Next, we discuss two strategies to reduce these costs.

Simultaneous sources

We reduce the cost of computing the wavefields by randomized source superposition (Herrmann et al., 2009; Krebs et al., 2009; Neelamani et al., 2010; Haber et al., 2010). In our matrix notation, a randomized supershot is given by

$$\overline{q} = Qw, \quad (8)$$

where w is a vector with i.i.d. random entries with zero mean and unit variance. In the remainder of the paper, dimensionality-reduced quantities will be denoted with a bar $\overline{\cdot}$. We denote a batch of randomized sources as

$$\overline{Q} = QW, \quad (9)$$

where W is a matrix whose columns are the randomized source codes. The number of columns – the batchsize – then determines the number of simultaneous sources used. The extended image volume can now be approximated by

$$\Delta M \approx \overline{\Delta M} = \overline{VU}^* = VWW^*U^*. \quad (10)$$

Note that $\frac{1}{N}WW^* \rightarrow I$, where N is the batchsize, if we increase the batchsize, so we can increase the accuracy of our estimate by increasing the batchsize.

Probing the image volume

Instead of forming the image volume kernel explicitly, we consider computing only its action on a vector. The action of the image volume kernel on a vector, $y = \Delta Mx$, can be efficiently computed as follows.

1. solve the forward equation (eq. 6) for U and calculate $z = U^*x$.
2. use z as stacking weights for the rhs ($\Delta D = (PU - D)$) of the adjoint equation (eq. 7): $H^*[m]y = P^*\Delta Dz$

This means that, given the source wavefield U, we can calculate the action of the image volume on a vector for *one* additional PDE solve (per frequency). Of particular practical value is the case $x = e_i$, the i^{th} unit vector. U^*e_i then represents the Green's function from the sources to a single scattering point, indicated by i. The vector ΔDz can then be interpreted as the response from a redatumed virtual source located at the subsurface scattering point. The vector y is now the image of the source computed by time-reversal, as is done in source localization (Mantzel et al., 2010, for example). The difference with source-localization is that one wants to determine the source location, given the medium velocity. Here, we want to do the opposite: given the source location, determine the medium velocity. Note that the time-reversal may also be interpreted as a redatuming of the receivers in accordance with the 'survey sinking' concept introduced by Claerbout (1985). A schematic depiction is given in figure 1. Such an interferometric interpretation of the image volume is also presented by Vasconcelos et al. (2010). Note that it also bears resemblance to the CFP concept, where ΔM would be the 'reflectivity matrix' (Berkhout, 1997). In practice, we would pick a number of subsurface points, possibly informed by a migrated image, and calculate the image volume for these points.

Example

We generate 'observed' data for the model depicted in figure 2 (a) for 61 equispaced sources, 301 equispaced receivers and frequencies [3:0.5:25]Hz . The extended image for a single scatterpoint is depicted in figure 3 for the velocity: $m_{\text{true}} + \alpha\delta m$ for $\alpha = \{-500, 0, 500\}$ m/s. The velocity perturbation, δm, is depicted in figure 2 (b). The corresponding simultaneous source-images ($\overline{\Delta M}$) for 10 simultaneous sources are depicted in figure 4. Although the use of simultaneous sources introduces some cross-talk artifacts in the image volume, the focusing is still distinct.

MULTISCALE INVERSION STRATEGY

The traditional formulation of MVA measures the amount of focusing with an *annihilator* that penalizes off-diagonal energy. Alternatively, this may be translated into taking a derivative with respect to transform-domain redundant coordinates such as angle (cf. de Bruin et al., 1990; Rickett and Sava, 2002; Sava and Fomel, 2006; Stolk and de Hoop, 2006). However, these annihilators are derived under idealized high-frequency assumptions under which the extended image will collapse to a diagonal matrix for the correct velocity. As seen in the previous examples some artifacts are present in these images, however. These are due to finite-frequency and finite-aperture effects and, in case of simultaneous sources, cross-talk between the different sources. To mitigate the influence of these artifacts somewhat we propose to measure the amount of focusing as follows:

$$\phi[m] = ||G_\sigma \odot \Delta M[m]||^2, \tag{11}$$

where $g_{ij} \propto \exp[-r_{ij}^2/\sigma^2]$, where r_{ij} denotes the Euclidean distance between gridpoints i and j. The correct velocity model will then maximize this functional. This approach can be seen as a generalization of the stack-power criterion, which arises if we let $\sigma \downarrow 0$. Such a weighted norm was earlier found to be useful for data-domain velocity analysis (van Leeuwen and Mulder, 2008).

The corresponding optimization problem is then given by

$$\max_{m\in B} \sum_{i\in\mathscr{I}} ||(G_\sigma e_i) \odot (\Delta M[m] e_i)||^2, \tag{12}$$

where $\mathscr{I}$ is the set of points for which we want to evaluate the extended image and B is an appropriate basis that will ensure smoothness of the model (cubic splines, for example). The freedom to choose the set $\mathscr{I}$, the basis and the scale immediately suggests a multiscale inversion strategy:

1. Choose initial set of points $\mathscr{I}_0$, initial scale σ_0 and initial basis B_0
2. Solve $m^{k+1} = \text{argmin}_{m\in B_k} \sum_{i\in\mathscr{I}_k} ||(G_{\sigma_k} e_i) \odot (\Delta M[m] e_i)||^2$
3. Choose a finer grid $\mathscr{I}_{k+1}$, a smaller scale σ_{k+1} and a 'less smooth' basis B_k
4. Repeat 2-4

As the algorithm proceeds to finer scales, the behavior of the misfit approaches the stackpower criterion, which has been shown to be equivalent to the least-squares criterion (cf. Stolk and Symes, 2003). Thus, this gives us a natural way to proceed from velocity analysis (for large σ) to FWI (as $\sigma \downarrow 0$).

In contrast to the annihilator, A, the domain of influence of G_σ is relatively small. This has two advantages: *i)* the influence of the cross-talk introduced by using simultaneous sources is automatically reduced, and *ii)* we may consider stacking the extended images for several subsurface points and weighing them with the stacked weights. The computational gain here lies in the fact that we would only have to evaluate the action of the extended image on one vector to evaluate the misfit for several subsurface points.

Example

We evaluate the focusing power for the three subsurface points indicated in figure 2 for the model $m = m_{\text{true}} + \alpha\delta m$ for various scales. We normalize the focusing power with the norm of the extended image itself. The result for all sources and 10 simultaneous sources is shown in figure 5. The focusing power for the simultaneous sources shows very similar behavior as the one for all the sources. This suggests that the computational cost of forming the extended images can be reduced significantly by employing simultaneous sources without affecting their sensitivity to velocity errors. The scale parameter controls the width of the basin of attraction of the focusing criterion.

CONCLUSIONS

The non-linear extended-modeling formulation would in principle allow us to combine velocity analysis and waveform inversion in a consistent manner. In particular, it promises a formulation of wave-equation MVA that incorporates nonlinear effects, such as multiple scattering, that are usually ignored in traditional MVA. However, the formidable computational cost involved has so far hampered its application to large scale problems. The linearization of the full non-linear extended modeling problem leads to a constrained formulation of the more traditional wave-equation MVA approach. The constrained formulation tries to optimally focus the extended image subject to fitting the data. The computational costs here lie in modeling the forward and adjoint wavefields that constitute the extended image and forming the whole image volume. We apply randomized source encoding techniques to reduce the costs of modeling the wavefields. The corresponding extended image exhibits very reasonable focusing behavior even though the randomized sources introduce some cross-talk artifacts. We then consider spatially subsampling the image volume. The extended image for one point can be efficiently calculated at the cost of one extra PDE solve, given the source wavefield. Since the energy in the extended images is localized, we may calculate the images for several points simultaneously. In particular we may apply randomized source encoding to the subsurface sources. Finally, we propose an alternative focusing criterion that may be seen as a generalization of the stack-power criterion. The focusing penalty is a Gaussian with a prescribed width, which allows us to determine the scale at which we measure the misfit. The Gaussian weight also automatically reduces the influence of the crosstalk introduced by using simultaneous sources. The freedom to choose the set of points at which we evaluate the extended image and the scale allows for a multiscale/layerstripping inversion strategy. We investigate the behavior of the focusing criterion for various scales for a simple model and find that 'region of convexity' is proportional to the scale.

ACKNOWLEDGMENTS

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BP, Chevron, ConocoPhillips, Petrobras, Total SA, and WesternGeco.

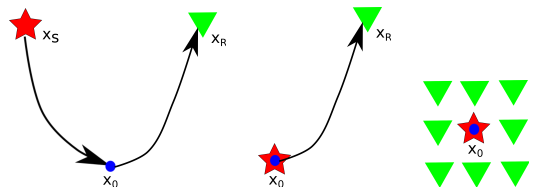

Figure 1: Schematic depiction of interferometric interpretation of the extended image volume. The data are (kinematically) equivalent to the Green's function from x_s to x_r (a). Upon correlating with the Green's function from x_s to x_0 (the source wavefield) we obtain the response from a virtual source located at x_0 (b). Finally, correlation with the Greens function from x_r to every point in the subsurface (the receiver wavefield) yields the response of the virtual experiment depicted in (c).

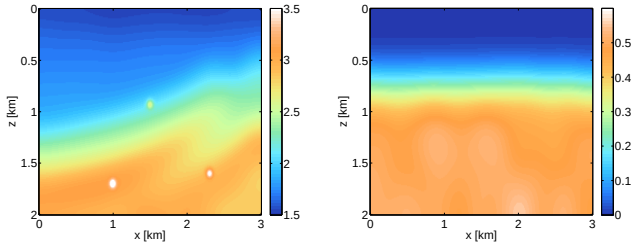

Figure 2: (a) Smooth velocity [km/s] with three point-scatterers, used in the examples. (b) Velocity perturbation [km/s].

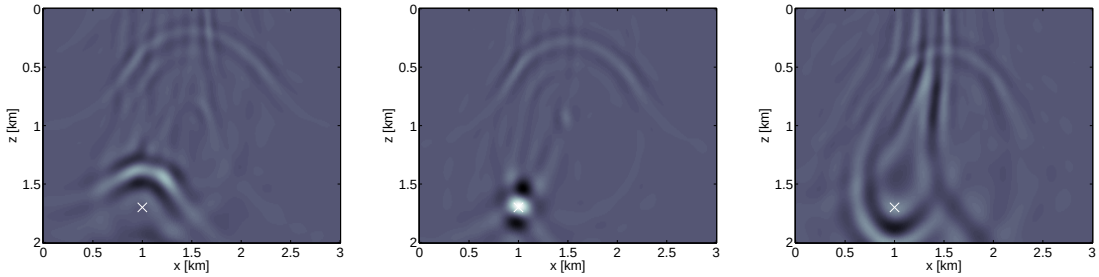

Figure 3: Image volume for single scattering point, indicated by $\times$ for (a) low velocity, (b) correct velocity and (c) high velocity.

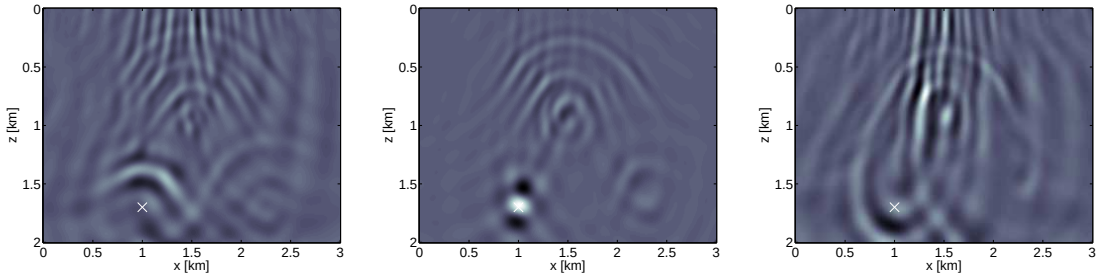

Figure 4: Image volume computed with 10 simultaneous sources for single scattering point, indicated by $\times$ for (a) low velocity, (b) correct velocity and (c) high velocity.

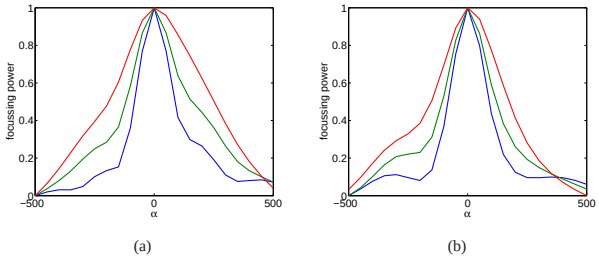

Figure 5: Focusing power for various scales: $\sigma = \{50, 100, 200\}$ m (blue,green,red), for (a) all sources. (b) 10 simultaneous sources.

REFERENCES

Berkhout, A. J., 1985, Seismic migration: Imaging of acoustic energy by wave field extrapolation: Elsevier.

———, 1997, Pushing the limits of seismic imaging, Part I: Prestack migration in terms of double dynamic focusing: Geophysics, **62**, 937.

Bunks, C., F. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473.

Claerbout, J., 1985, Imaging the earth's interior: Blackwell Scientific Publishers.

de Bruin, C. G. M., C. P. A. Wapenaar, and A. J. Berkhout, 1990, Angle-dependent reflectivity by means of prestack migration: Geophysics, **55**, 1223–1234.

de Hoop, M., R. van der Hilst, and P. Shen, 2006, Wave-equation reflection tomography: annihilators and sensitivity kernels: Geophysical Journal International, **167**, 1332–1352.

Haber, E., M. Chung, and F. J. Herrmann, 2010, An effective method for parameter estimation with pde constraints with multiple right hand sides: Technical Report TR-2010-4, UBC-Earth and Ocean Sciences Department.

Herrmann, F. J., 2009, Compressive imaging by wavefield inversion with group sparsity: SEG Technical Program Expanded Abstracts, SEG, SEG, 2337–2341.

Herrmann, F. J., Y. A. Erlangga, and T. Lin, 2009, Compressive simultaneous full-waveform simulation: Geophysics, **74**, A35.

Krebs, J. R., J. E. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M.-D. Lacasse, 2009, Fast full-wavefield seismic inversion using encoded sources: Geophysics, **74**, WCC177–WCC188.

MacKay, S., and R. Abma, 1992, Imaging and velocity analysis with depth-focusing analysis: Geophysics, **57**, 1608–1622.

Mantzel, W., J. Romberg, and K. Sabra, 2010, Randomized group testing for acoustic source localization: SPIE Computational Imaging VIII.

Mulder, W., and T. van Leeuwen, 2008, Automatic migration velocity analysis and multiples: SEG Expanded Abstracts, **27**, 3128–3132.

Neelamani, R. N., C. E. Krohn, J. R. Krebs, J. K. Romberg, M. Deffenbaugh, and J. E. Anderson, 2010, Efficient seismic forward modeling using simultaneous random sources and sparsity: Geophysics, **75**, WB15–WB27.

Pratt, R., Z. Song, P. Williamson, and M. Warner, 1996, Two-dimensional velocity models from wide-angle seismic data by wavefield inversion: Geophysical Journal International, **124**, 232–340.

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

Sava, P., and S. Fomel, 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, S209–S217.

Shen, P., and W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, VE49–VE59.

Shin, C., and Y. Cha, 2009, Waveform inversion in the Laplace-Fourier domain: Geophysical Journal International, **177**, 1067–1079.

Stolk, C., and M. de Hoop, 2006, Seismic inverse scattering in the downward continuation approach: Wave Motion, **43**, 579–598.

Stolk, C., and W. Symes, 2003, Smooth objective functionals for seismic velocity inversion: Inverse Problems, **19**, 73–89.

Symes, W., 2008a, Approximate linearized inversion by optimal scaling of prestack depth migration: Geophysics, **73**, R23–R35.

———, 2008b, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790.

van Leeuwen, T., and W. Mulder, 2008, Velocity analysis based on data correlation: Geophysical Prospecting, **56**, 791–803.

Vasconcelos, I., P. Sava, and H. Douma, 2010, Nonlinear extended images via image-domain interferometry: Geophysics, **75**, SA105–SA115.

Yang, T., and P. Sava, 2010, Moveout analysis of wave-equation extended images: Geophysics, **75**, S151–S161.