Curvelet-based migration preconditioning

Peyman P. Moghaddam, Cody R. Brown and Felix J. Herrmann

SUMMARY

In this paper, we introduce a preconditioner for seismic imaging—i.e., the inversion of the linearized Born scattering operator. This preconditioner approximately corrects for the "square root" of the normal—i.e., the demigration-migration operator. This approach consists of three parts, namely (i) a left preconditoner, defined by a fractional time integration designed to make the migration operator zero order, and two right preconditioners that apply (ii) a scaling in the physical domain accounting for a spherical spreading, and (iii) a curvelet-domain scaling that corrects for spatial and reflector-dip dependent amplitude errors. We show that a combination of these preconditioners lead to a significant improvement of the convergence for iterative least-squares solutions to the seismic imaging problem based on reverse-time migration operators.

INTRODUCTION

Over the years, extensive research has been done to reduce the computational costs of seismic imaging. Improvements in this area are particularly important during iterative least-squares migration, where the linear Born scattering operator is inverted with iterative Lanczos methods, such as LSQR (Paige and Saunders, 1982; De Roeck, 2002). Examples of this method can be found in the literature (see e.g. Nemeth et al., 1999; Chavent and Plessix, 1999; Kuhl and Sacchi, 2003).

The most successful methods to reduce the cost of migration are the so-called scaling methods where the action of the compound linearized modeling-migration operator—known as the Hessian or normal operator—is replaced by a diagonal scaling in some domain, see e.g. contributions from Claerbout and Nichols (1994); Rickett (2003); Guitton (2004), and more recently from Herrmann et al. (2008) and Symes (2008). These methods vary in degree of sophistication with regard to the estimation of the diagonal through migrated-image to remigrated-image matching and in the way the scaling is applied—i.e., by division in the physical or via sparsity promotion in the curvelet domain as reported in Herrmann et al. (2008). Amplitudes are restored during all these methods by applying the scaling as a post-processing step after migration.

In this paper, we take this line of research a step further by using the above scaling argument to define the appropriate preconditioning to the system of equations involved in linearized Born scattering. To illustrate the improvements in the migrated image and in the convergence of least-square migration, we consider three levels of preconditioning. First, we correct for the order of the normal operator by introducing a left preconditioning consisting of a fractional time integration. This first level is consistent with earlier work reported by Herrmann et al. (2008) and Symes (2008). The next level of preconditioning is made off a simple diagonal scaling in the physical domain to compensate for spherical spreading of seismic waves. As a last step, we include a curvelet-domain scaling as part of the right preconditioning. We conclude by studying the performance of these different levels of preconditioning on a synthetic example using a reverse-time 'wave-equation' migration code with optimal checkpointing (Symes, 2007).

PROBLEM FORMULATION

During seismic imaging, the following system of equations needs to be solved

$$\boldsymbol{A}\mathbf{x} = \boldsymbol{b}, \tag{1}$$

where $\mathbf{b}$ is the known data vector, $\boldsymbol{A}$ the linearized Born scattering operator, and $\mathbf{x}$ the unknown model vector. Contrary to many inverse problems, the matrix $\boldsymbol{A}$, albeit extremely large, is reasonably well behaved and a migrated image can be obtained by applying the adjoint of $\boldsymbol{A}$ to the data vector—i.e., $\widetilde{\mathbf{x}} = \boldsymbol{A}^*\mathbf{b}$ with $\boldsymbol{A}^*$ the migration operator, the symbol * denoting the adjoint, and $\widetilde{\mathbf{x}}$ the migrated image. Unfortunately the output of this procedure called migration yields erroneous results for the amplitudes of the imaged reflectors. To restore these amplitudes, the matrix in Equation 1 can be inverted using the method of least-squares—i.e.,

$$\widetilde{\mathbf{x}}_{LS} = \left(\boldsymbol{A}^*\boldsymbol{A}\right)^{-1}\boldsymbol{A}^*\mathbf{b} := \boldsymbol{A}^\dagger\mathbf{b} \tag{2}$$

with $\boldsymbol{A}^*\boldsymbol{A}$ as the normal or Hessian operator. Since $\boldsymbol{A}$ is large, its least-squares inverse needs to be calculated with iterative matrix-free methods such as LSQR (Paige and Saunders, 1982). Even though these methods convergence quickly, the shear size of the imaging problem calls for a reduction of the number of iterations—i.e., the number of matrix-vector multiplies.

In a perfect world with infinite computational resources, the ideal preconditioning for the algorithm in Equation 1 would correspond to

$$\boldsymbol{A}\boldsymbol{M}_R^{-1}\boldsymbol{u} = \boldsymbol{b} \qquad \boldsymbol{x} := \boldsymbol{M}_R^{-1}\mathbf{u} \tag{3}$$

with $\boldsymbol{M}_R := \left(\boldsymbol{A}^*\boldsymbol{A}\right)^{1/2}$, the square-root of the normal operator. In this ideal case, migration would recover the image vector $\mathbf{x}$. Unfortunately, in practice the quantity $\left(\boldsymbol{A}^*\boldsymbol{A}\right)^{1/2}$ cannot be computed and we have to resort to the appropriate approximations. In this paper, we propose a combination of left and right preconditioning—i.e., we replace Equation 1 by

$$\boldsymbol{M}_L^{-1}\boldsymbol{A}\boldsymbol{M}_R^{-1}\mathbf{u} = \boldsymbol{M}_L^{-1}\boldsymbol{b} \qquad \boldsymbol{x} := \boldsymbol{M}_R^{-1}\mathbf{u} \tag{4}$$

with $\boldsymbol{M}_L^{-1}$ being the left preconditioning operator. In this case, the migrated and least-squares migrated image are given by $\widetilde{\mathbf{x}} = \boldsymbol{M}_R^{-1}\widetilde{\mathbf{u}}$ with $\widetilde{\mathbf{u}} = \boldsymbol{F}^*\mathbf{y}$ and $\widetilde{\mathbf{x}}_{LS} = \boldsymbol{M}_R^{-1}\widetilde{\mathbf{u}}_{LS}$ with $\widetilde{\mathbf{u}}_{LS} = \boldsymbol{F}^\dagger\mathbf{y}$ respectively. Here, $\mathbf{y} = \boldsymbol{M}_L^{-1}\boldsymbol{b}$ which is the left preconditioned data. Our preconditioners derive from the following three observations:

1. the normal operator is in d dimensions a $(d-1)$-order pseudo differential operator (ΨDO, see e.g. recent work by Symes, 2008; Herrmann et al., 2008, and the references therein).

2. migration amplitudes decay with depth due to spherical spreading of waves.
3. zero-order ΨDO's can be approximated by a diagonal scaling in the curvelet domain (see e.g. Herrmann et al., 2008).

These findings allow us to define a series of increasingly more accurate approximations to the square-root of the normal operator leading to better and better preconditioners.

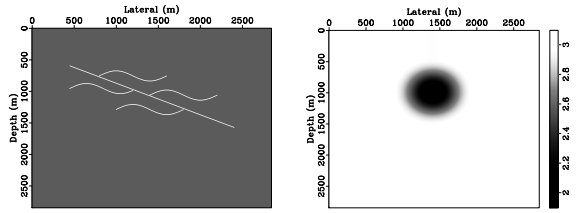

Figure 1: Synthetic example. **(a)** Reflectivity. **(b)** Low-velocity lens model used as a background velocity for both migration and modeling operators.

PRECONDITIONING

In this section, we introduce different types of preconditioners based on the aforementioned observations. For each definition of the preconditioning operators, migrated images—i.e., $\widetilde{\mathbf{x}}$, are shown as well. The examples are computed for the reflectivity and smooth velocity background models plotted in Figure 1. The Born scattering and migration are performed with the reverse-time 'wave-equation' finite-difference code developed by Symes (2007).

Left preconditioning by fractional differentiation
As stated before, the normal operator corresponds under certain conditions—such as the high-frequency limit and a smooth background velocity model—to a $(d-1)$-order ΨDO. In two-dimensional image space, this operator corresponds to a leading-order behavior of a Laplacian—i.e., the action of $\left(-\Delta\right)\cdot$ in the physical domain, or to $|\xi|^2\cdot$ with ξ the wave vector in the spatial Fourier domain. It can be shown that in the data space this action corresponds to a multiplication by $|\omega|$ in the temporal Fourier domain, which can be compensated for by the following left preconditioning

$$\begin{aligned} \boldsymbol{M}_L^{-1}\boldsymbol{A}\mathbf{x} &= \boldsymbol{M}_L^{-1}\mathbf{b} \quad \text{with} \quad \boldsymbol{M}_L^{-1} := \partial_t^{-1/2} \\ \boldsymbol{F}\mathbf{x} &= \mathbf{y}, \end{aligned} \tag{5}$$

where $\partial_t^{-1/2}\cdot = F^{-1}|\omega|^{-1/2}\cdot$ with F^{-1} the inverse Fourier transform. Comparing the migrated images before and after preconditioning (cf. Figures 3(a) and 2(b))—i.e., $\boldsymbol{A}^*\mathbf{b}$ compared to $\boldsymbol{F}^*\mathbf{y}$—shows that the imprint of the Laplacian on the reflectors in the image has been removed. This observation can be understood since $\boldsymbol{M}_L^*\boldsymbol{M}_L \longleftrightarrow |\omega|$. However, the migrated image still contains a dimming of the amplitudes.

Right preconditioning by scaling in the physical domain
To further correct the amplitudes, we propose to apply a scaling that corrects for the amplitude changes induced by the action of the Hessian. On physical grounds, we can safely state that the amplitudes of the migrated image will decay quadratically with depth. This depth-dependent amplitude decay is quadratic because reflected waves travel from the source at the surface down to the reflector, experiencing an amplitude decay proportional to the depth of the reflector, and back up, experiencing another decay—hence the quadratic dependence. We can compensate for this amplitude decay by imposing the following right preconditioning:

$$\begin{aligned} \boldsymbol{M}_L^{-1}\boldsymbol{A}\boldsymbol{M}_R^{-1}\mathbf{u} &= \mathbf{y}, \quad \boldsymbol{M}_R^{-1} = \boldsymbol{D}_z := \mathrm{diag}\left(\mathbf{z}\right) \\ \boldsymbol{F}\mathbf{u} &= \mathbf{y}, \quad \boldsymbol{x} := \boldsymbol{M}_R^{-1}\mathbf{u} \end{aligned} \tag{6}$$

where $z_i = i\Delta z\, i = 1\cdots n_z$, with Δz the vertical sample interval and n_z the number of depth samples. The left preconditioner $\boldsymbol{M}_L$ is defined as before. The results for the migrated image in Figure 2(c) for this preconditioning show an improvement as expected. However, there are remaining amplitude differences between the original band-width limited reflectivity and the migrated reflectivity.

Right preconditioning by scaling in the curvelet domain
After applying the left preconditioning, the Hessian can be modeled by a zero-order ΨDO whose action corresponds to that of a non-stationary dip filter—i.e., we have

$$\left(\Psi f\right)(x) = \int_{x\in\mathbb{R}^d} e^{jk\cdot x} a(x,\xi)\hat{f}(k)\mathrm{d}\xi \tag{7}$$

with Ψ the Hessian for the precondioned modeling operator and $a(x,\xi)$ a space- and spatial-frequency dependent filter known as the symbol. As shown in Herrmann et al. (2008), the action of the ΨDO can after discretization be approximated by a scaling in the curvelet domain—i.e., we have the following

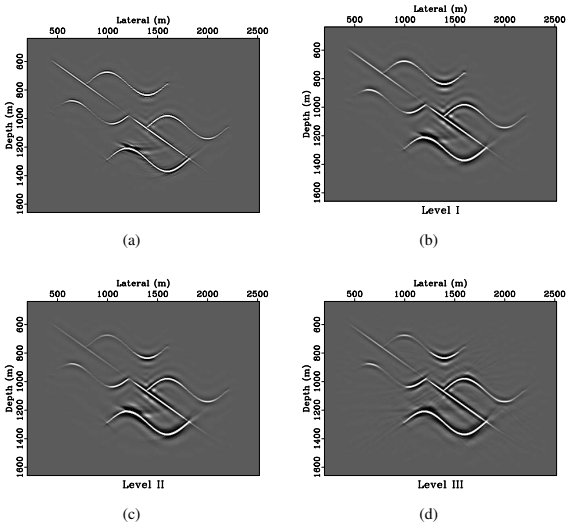

Figure 2: Different levels of preconditioning. **(a)** Migrated image without preconditioning—i.e., $\widetilde{\mathbf{x}} = \boldsymbol{A}^*\mathbf{b}$. **(b)** Migrated result for left preconditioning (cf. Equation 5). **(c)** Migrated result for left-right (depth-correction only) preconditioning (cf. Equation 6). **(d)** The same but now including curvelet-domain scaling (cf. Equation 9).

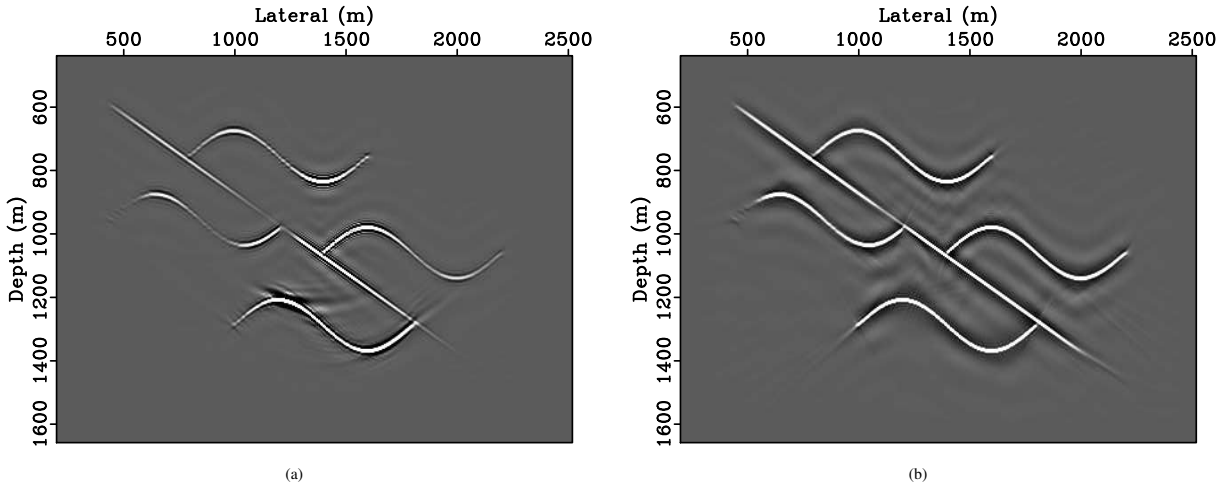

Figure 3: Migration versus least-squares migration. **(a)** Conventional migration. **(b)** Least-squares migration with level III preconditioning. Notice the improvement in the recovered image.

approximate identity

$$\mathbf{\Psi r} \approx \boldsymbol{C}^* \boldsymbol{D}_\Psi^2 \boldsymbol{C}\mathbf{r}, \quad \boldsymbol{D}_\Psi^2 = \mathrm{diag}\left(\mathbf{d}^2\right) \tag{8}$$

accurate for a reference vector $\mathbf{r}$ close enough to the actual image. For details on how to estimate the curvelet-domain scaling coefficients refer to Herrmann et al. (2008) or to another contribution on curvelet-domain matched filtering by the authors to the proceedings of this meeting. The cost for calculating the diagonal scaling in Equation 8 is roughly one migration and one remigration.

Again, we can compensate for this unwanted non-stationary contribution to the amplitudes of the migrated image by defining the following preconditioned system

$$\begin{aligned} \boldsymbol{M}_L^{-1} \boldsymbol{A} \boldsymbol{M}_R^{-1} \mathbf{u} &= \mathbf{y}, \quad \boldsymbol{M}_R^{-1} = \boldsymbol{D}_z \boldsymbol{C}^* \boldsymbol{D}_\psi^{-1} \\ \boldsymbol{F}\mathbf{u} &= \mathbf{y}, \quad \boldsymbol{x} := \boldsymbol{M}_R^{-1}\mathbf{u}. \end{aligned} \tag{9}$$

The corresponding image obtained with this system is plotted in Figure 2(d) and shows a further improvement in the amplitudes for the migrated image. However, some image artifacts and amplitude errors remain.

CONVERGENCE OF LEAST-SQUARES MIGRATION

Even though the different preconditioning operators defined in the previous section lead to an improvement for the migrated image, problems still remain in getting a high-fidelity image. In practice, for cases where the background velocity model is accuratly known, these problems can be avoided by replacing migration with least-squares migration. In this case, the scattering operator is inverted using an iterative method such as LSQR, which solves the following optimization problem

$$\begin{cases} \widetilde{\mathbf{u}} = \arg\min_{\mathbf{u}} \|\mathbf{u}\|_2 \quad \text{s.t.} \quad \boldsymbol{F}\mathbf{u} = \mathbf{y} \\ \widetilde{\mathbf{x}} = \boldsymbol{M}_R^{-1}\widetilde{\mathbf{u}}. \end{cases} \tag{10}$$

The performance of iterative solutions to this problem depends on the properties of the matrix $\boldsymbol{F}$. Unfortunately, the size of matrices involved in a typical migration precludes extensive studies, such as computing the singular values, looking for clustering of simular values or estimating their largest and smallest values. These properties determine convergence of an iterative solver but are out of reach for operators defined in terms of reverse-time 'wave-equation' migration. For a more detailed discussion on this topic with Kirchoff migration on a small model, the reader is referred to recent work by De Roeck (2002).

Because of the sheer size of the problem, we limited ourselves to study the progress made during the first fifteen iterations of the iterative solver LSQR (Paige and Saunders, 1982). Because LSQR aims to minimize the residual at each iteration, we introduce the normalized log-based least-squares residual after k iterations as

$$\mu_k = 20 \log \|\boldsymbol{F}\mathbf{u}_k - \mathbf{y}\|_2 / \|\mathbf{y}\|_2. \tag{11}$$

As pointed out legitimately by De Roeck (2002), we cannot expect this residue to go to zero because of the presence of unmodeled components, such as noise in raw data, which may not be in the range of the modeling operator. Therefore, we also study the normalized log-based least-squares model-space residual—i.e.,

$$\nu_k = 20 \log \|\boldsymbol{F}^*\left(\boldsymbol{F}\mathbf{u}_k - \mathbf{y}\right)\|_2 / \|\boldsymbol{F}^*\mathbf{y}\|_2, \tag{12}$$

which can become small.

To empirically establish the performance of our preconditioners, we include in Figure 4 plots for the decay of μ_k and ν_k as a function of the number of iterations k of the LSQR algorithm. The cost of each iteration of this algorithm is dominated by one matrix-vector multiply by $\boldsymbol{F}$ and $\boldsymbol{F}^*$. We compare the decay of these two quantities for the different preconditioning levels introduced earlier. Both residual errors decay faster as we move from a single left preconditioner, towards left and right preconditioners with a significant improvement obtained

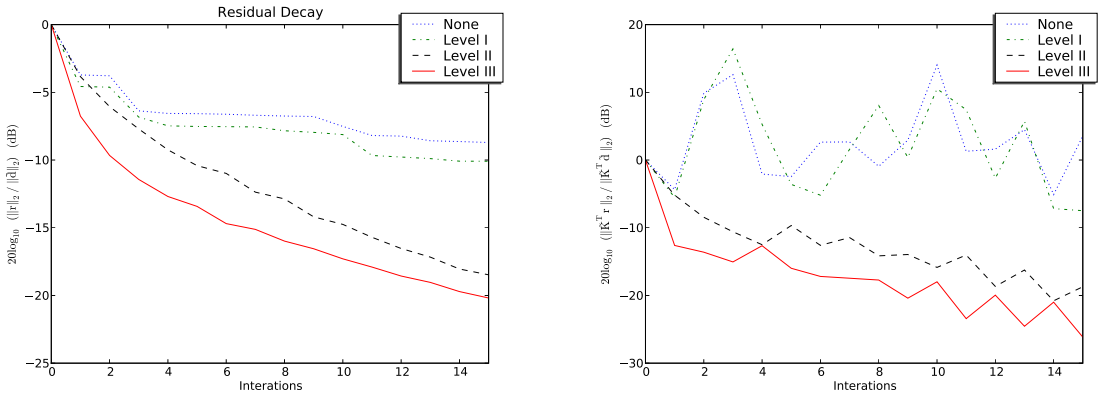

Figure 4: Residual decays for different levels of preconditioning. The dotted line corresponds to migration without preconditioning—i.e., $\widetilde{\mathbf{x}} = \boldsymbol{A}^*\mathbf{b}$. The level I migrated result corresponds to the left preconditioning (cf. Equation 5). Level II migrated result for left-right (depth-correction only) preconditioning (cf. Equation 6). Level III the same but now including curvelet-domain scaling (cf. Equation 9). **(a)** Plot for the decay of the data-space normalized residues (cf. Equation 11) as a function of the number of LSQR iterations. **(b)** The same but now for the model-space normalized residuals (cf. Equation 12).

by including the curvelet-domain scaling in the right preconditioning. Even though, the picture for the model space residual is less clear—because of more significant iteration-to-iteration variations—it is evident that our preconditioning leads to significant improvements. For instance, after two iterations with the full-preconditioned system, we already get a residual that is better than the residual obtained by the least-squares solution of the original system (cf. Equation 1). Furthermore the curvelet domain system residue after only four iterations is roughly equivalent to the residue of the left-right preconditioned system without curvelet preconditioning after eight interations. The residual error in the model space, Figure 4(b), clearly show the benefits of left-right preconditioning over the original and the left preconditioned system. Finally, Figure 3(b) show significant—albeit at the cost of a relatively small number of additional matrix-vector multiplies—improvements of a least-squares migration over a single migration with the preconditioned system (cf. 2(b) and 4(a)).

CONCLUSIONS

Migration preconditioning has somewhat been an illusive topic since traditional methods from linear algebra are not applicable. This is because of the sheer size of the migration problem and other complicating factors such as the presence of a source function. Nonetheless, the first few iterations of a least-squares migration are known to make significant progress towards the solution in cases where the background-velocity model is accurate. Unfortunately, the computational costs involved with least-squares migration are often still prohibitive in practice. The method presented in this paper partly resolves this issue through appropriate preconditioning. The presented example shows the best convergence for the data and model space residues for a preconditioning that consists of a combination of left and right preconditioning. The computational overhead for this preconditioning is relatively minor and mainly consists of an additional migration-remigration as part of calculating the curvelet domain approximation of the normal operator. This combined preconditioning leads to a significantly faster decay for the first few iterations, translating to a practical reduction of computational cost of roughly 50%. Finally, the presented formalism also lends itself for a sparsity-promoting inversion, during which the one-norm of the model-space curvelet coefficients are minimized.

ACKNOWLEDGMENTS

The authors would like to thank C. C. Stolk and W. W. Symes for their input on migration-amplitude recovery and for the use of the reverse-time migration code. We also would like to thank the authors of CurveLab (`www.curvelet.org`). The examples were prepared with Madagascar (`rsf.sf.net`), supplemented by SLIMpy operator overloading, developed by S. Ross-Ross. This work was in part financially supported by the NSERC Discovery (22R81254) and CRD Grants DNOISE (334810-05) of F.J.H. and was carried out as part of the SINBAD project with support, secured through ITF, from BG Group, BP, Chevron, ExxonMobil and Shell.

REFERENCES

Chavent, G., and R. E. Plessix, 1999, An optimal true-amplitude least-squares prestack depth-migration operator: Geophysics, **64**, 508–515.

Claerbout, J., and D. Nichols, 1994, Spectral preconditioning: Technical report, Stanford Exploration Project.

De Roeck, Y., 2002, Sparse linear algebra and geophysical migration: A review of direct and iterative methods: Numerical Algorithms, **29**, 283–322.

Guitton, A., 2004, Adaptive subtraction of multiples using the l^1-norm: Geophys Prospect, **52**, 27–27.

Herrmann, F. J., P. P. Moghaddam, and C. C. Stolk, 2008, Sparsity- and continuity-promoting seismic imaging with curvelet frames: Journal of Applied and Computational Harmonic Analysis, **24**, 150–173. (doi:10.1016/j.acha.2007.06.007).

Kuhl, H., and M. D. Sacchi, 2003, Least-squares wave-equation migration for avp/ava inversion: Geophysics, **68**, 262–273.

Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: **64**, 208–221.

Paige, C. C., and M. A. Saunders, 1982, LSQR: An algorithm for sparse linear equations and sparse least squares: ACM TOMS, **8**, 43–71.

Rickett, J. E., 2003, Illumination-based normalization for wave-equation depth migration: Geophysics, **68**, no. 4.

Symes, W. W., 2007, Reverse time migration with optimal checkpointing: Geophysics, **72**, SM213–SM221. (10.1190/1.2742686).

———, 2008, Approximate linearized inversion by optimal scaling of prestack depth migration: Geophysics, **73**, R23–R35. (10.1190/1.2836323).