Curvelet-based migration preconditioning and scaling

Felix J. Herrmann[1], Cody R. Brown[1], Yogi A. Erlangga[1], and Peyman P. Moghaddam[1]

(February 2, 2009)

Running head: **migration preconditioning**

ABSTRACT

The extremely large size of typical seismic imaging problems has been one of the major stumbling blocks for iterative techniques to attain accurate migration amplitudes. These iterative methods are important because they complement theoretical approaches that are hampered by difficulties to control problems such as finite-acquisition aperture, source-receiver frequency response, and directivity. To solve these problems, we apply preconditioning, which significantly improves convergence of least-squares migration. We discuss different levels of preconditioning that range from corrections for the order of the migration operator to corrections for spherical spreading, and position and reflector-dip dependent amplitude errors. While the first two corrections correspond to simple scalings in the Fourier and physical domain, the third correction requires phase-space (space spanned by location and dip) scaling, which we carry out with curvelets. We show that our combined preconditioner leads to a significant improvement of the convergence of least-squares 'wave-equation' migration on a line from the SEG AA' salt model.

[1]Seismic Laboratory for Imaging and Modeling, Department of Earth and Ocean Sciences, University of British Columbia, 6339 Stores Road, Vancouver, V6T 1Z4, BC, Canada.

INTRODUCTION

Over the years, extensive research has been done to reduce the computational costs of (least-squares) 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 these methods can be found in the literature (see e.g. Nemeth et al., 1999; Chavent and Plessix, 1999; Hu et al., 2001; Kuhl and Sacchi, 2003; Yu et al., 2006).

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 by Claerbout and Nichols (1994); Rickett (2003); Guitton (2004); Plessix and Mulder (2004), and more recently by Symes (2008) and Herrmann et al. (2008a). These methods vary in degree of sophistication with regard to the estimation of the diagonal, e.g. through migrated-image to remigrated-image matching. They also differ in the way the scaling is applied—i.e., by 'division' in the physical domain or via sparsity promotion in the curvelet domain, as reported recently by Herrmann et al. (2008a). During all these methods, imaged amplitudes are restored by applying a scaling as a post-processing step after migration. Even though our approach and the one of Symes (2008) are similar, the curvelet-domain scaling has the advantage that it is able to handle conflicting dips.

In this paper, we take this line of research a step further by using the above scaling argument to apply the proper preconditioning to the system of equations involved in linearized Born scattering. By "proper" we mean a preconditioner with a computational overhead that justifies its improvement by the increase in convergence rate. Because the system is

solved iteratively, the preconditioner does not need to be very accurate, avoiding unnecessary extra computational overhead. Note that we use the term preconditioning somewhat loosely compared to its formal definition in numerical linear algebra where the solution is not changed with preconditioning. Therefore, we also use this term to denote changes in the forward model that favor least-squares inversion. To illustrate the improvements in the images and in the convergence of least-squares migration, we consider three levels of preconditioning. First, we correct the order of the normal operator by introducing a left preconditioning, consisting of a fractional time integration that corresponds to a scaling in the Fourier domain. This first level of preconditioning follows directly from earlier work on migration-amplitude recovery reported by Herrmann et al. (2008a) and Symes (2008). The next level of preconditioning consists of diagonal scaling in the physical domain that compensates for spherical spreading of seismic waves. As a final step, we also include a curvelet-domain scaling as part of the right preconditioning. This step corrects for the remaining amplitude errors that vary spatially as a function of the reflector dip. We conclude by studying the performance of these different levels of preconditioning on the SEG AA' salt model (O'Brien and Gray, 1996; Aminzadeh et al., 1997), using a reverse-time 'wave-equation' migration code with optimal checkpointing (Symes, 2007). Because our main interest is to study the performance of our preconditioner, we work on data generated by the linearized Born-scattering operator instead of data generated by the solution of the full nonlinear forward model.

PROBLEM FORMULATION

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

$$\mathbf{Ax} \approx \mathbf{b}, \tag{1}$$

where $\mathbf{b} = \mathbf{A}\mathbf{x}_0$ is the multiple-free data with $\mathbf{x}_0$ the true reflector model, $\mathbf{A}$ the linearized Born-scattering operator, and $\mathbf{x}$ the unknown model vector. Here, the symbol $\approx$ refers to "equality" in the least-squares sense. Contrary to many inverse problems, the matrix $\mathbf{A}$, albeit extremely large, is reasonably well behaved and a migrated image can be obtained by applying the adjoint of $\mathbf{A}$ to the data vector—i.e., $\widetilde{\mathbf{x}} = \mathbf{A}^*\mathbf{b}$ with $\mathbf{A}^*$ the migration operator. The symbol * denotes the adjoint, and $\widetilde{\mathbf{x}}$ is the estimate (denoted by the $\widetilde{\ }$) of the image.

Unfortunately, the output of the above procedure, called migration, produces erroneous results for the amplitudes of the imaged reflectors. To restore these amplitudes, the least-squares solution to Equation 1 can be obtained as the solution of the linear system

$$\mathbf{A}^*\mathbf{A}\mathbf{x} \approx \mathbf{A}^*\mathbf{b}, \tag{2}$$

with $\mathbf{A}^*\mathbf{A}$ the normal or Hessian operator. Solutions to this system are not unique and correspond to solutions of the least-squares method—i.e.,

$$\widetilde{\mathbf{x}}_{LS} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2, \tag{3}$$

which finds image vectors, $\mathbf{x}$, that after modeling fit the data vector, $\mathbf{b}$. However, the matrix $\mathbf{A}^*\mathbf{A}$ is not invertible—i.e., it has zero or small singular values that correspond to shadow zones. Because the system is large, we employ an iterative matrix-free solution method, such as LSQR (Paige and Saunders, 1982). By limiting the number of iterations for this method, we control the energy of the solution vector $\widetilde{\mathbf{x}}$ and we obtain a regularized solution that is equivalent to the solution of a damped least-squares problem (Vogel, 2002). Even though these iterative methods converge relatively quickly (see e.g. Nemeth et al., 1999; Chavent and Plessix, 1999; Hu et al., 2001; Kuhl and Sacchi, 2003; Yu et al., 2006), the sheer size of the imaging problem calls for a further reduction in the number of iterations.

In a perfect world, with infinite computational resources, the ideal preconditioning for the system in Equation 1 corresponds to

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

with the right preconditioning matrix, $\mathbf{M}_R := \left(\mathbf{A}^*\mathbf{A}\right)^{1/2}$, given by the 'square-root' of the normal operator. Here, the symbol := refers to 'defined as'. In this ideal case, migration recovers the image vector $\mathbf{x}$, exactly. Unfortunately, in practice (albeit some recent exciting progress has been made by Demanet and Ying, 2008, using discrete symbol calculus for smooth symbols, a development on which we intend to report in the future) the quantity $\left(\mathbf{A}^*\mathbf{A}\right)^{1/2}$ cannot be computed and we have to resort to appropriate approximations.

In this paper, we propose a combination of left and right preconditioning—i.e., we replace Equation 1 by

$$\overbrace{\mathbf{M}_L^{-1}\mathbf{A}\mathbf{M}_R^{-1}}^{\widehat{\mathbf{A}}}\mathbf{u} \approx \overbrace{\mathbf{M}_L^{-1}\mathbf{b}}^{\widehat{\mathbf{b}}}, \qquad \mathbf{x} := \mathbf{M}_R^{-1}\mathbf{u}, \tag{5}$$

with $\mathbf{M}_L^{-1}$ the left preconditioning matrix. The migrated and least-squares migrated images are given by $\widetilde{\mathbf{x}} = \mathbf{M}_R^{-1}\widetilde{\mathbf{u}}$, with $\widetilde{\mathbf{u}} = \widehat{\mathbf{A}}^*\widehat{\mathbf{b}}$, and by $\widetilde{\mathbf{x}}_{LS} = \mathbf{M}_R^{-1}\widetilde{\mathbf{u}}_{LS}$, with $\widetilde{\mathbf{u}}_{LS} = \arg\min_{\mathbf{u}} \|\widehat{\mathbf{b}} - \widehat{\mathbf{A}}\mathbf{u}\|_2$, respectively. Our preconditioners are derived from the following three observations: (i) under certain conditions—such as the high-frequency limit, smooth background velocity models, and the absence of turning waves (Stolk, 2000)—the normal operator is in d dimensions a $(d-1)$-order pseudodifferential operator (ΨDO, see e.g. recent work by Herrmann et al., 2008a; Symes, 2008, and the references therein), (ii) migration amplitudes decay with depth due to spherical spreading of seismic body waves, and (iii) zero-order ΨDO's can be approximated by a diagonal scaling in the curvelet domain (see e.g. Herrmann et al., 2008a). These observations 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. Finally, we also argue that using curvelets will add a certain robustness to Gaussian noise and modeling errors, an observation substantiated by successful applications of this transform in seismic data processing (Herrmann et al., 2008b; Wang et al., 2008).

PRECONDITIONING

In this section, we introduce different types of preconditioners based on the aforementioned observations. For each preconditioned system, we study the migrated images. Later, we study the convergence of the iterative solver. The examples are computed for the reflectivity and smooth velocity background models plotted in Figure 1. To test our preconditioner, data with 324 shots is generated using Equation 1. Each shot consists of 176 traces of 6.4s and with a trace interval of 24m. The maximum offset of the data is 4224m.

Left preconditioning by fractional differentiation

As stated before, under aforementioned conditions the normal operator corresponds to a $(d-1)$-order ΨDO. In 2-D image space, this operator corresponds to the leading-order behavior of the Laplacian—i.e., the action of $\big(-\Delta\big)\cdot$ in the physical domain, or to $|\xi|^2\cdot$ with ξ the wave vector in the spatial Fourier domain. In data space, this action corresponds to a multiplication by $|\omega|$ in the temporal Fourier domain (Herrmann et al., 2008a). We compensate for this action by defining the following left preconditioning:

$$\mathbf{M}_L^{-1} := \partial_{|t|}^{-1/2}, \tag{6}$$

where $\partial_{|t|}^{-1/2}\cdot := \mathcal{F}^*|\omega|^{-1/2}\mathcal{F}\cdot$ with $\mathcal{F}$ the Fourier transform and $\mathcal{F}^* = \mathcal{F}^{-1}$ its inverse. We define the level I preconditioner with $\mathbf{M}_L^{-1}$ as above and $\mathbf{M}_R^{-1} = \mathbf{I}$ (which implies $\mathbf{x} = \mathbf{u}$).

Comparison of the migrated images before and after left preconditioning (cf. Figures 1(c) and 2(a)) shows that the imprint of the Laplacian is removed—i.e., some of the low-frequency content is restored. However, the migrated image still contains dimming of the amplitudes. Note that this left preconditioning corresponds to the solution of the scaled least-squares problem: $\arg\min_{\mathbf{x}} \frac{1}{2}\|\widehat{\mathbf{b}} - \widehat{\mathbf{A}}\mathbf{x}\|_2^2 = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{M}_L^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x})\|_2^2$.

Right preconditioning by scaling in the physical domain

To further correct the amplitudes, we propose to apply a scaling to compensate for the leading-order amplitude decay. This decay is linear because the reflected waves travel from the source down to the reflector, experiencing an amplitude decay proportional to the square-root of the reflector depth (in 2-D), and back up, experiencing another decay. We correct this linear amplitude decay by defining the right preconditioning matrix:

$$\mathbf{M}_R^{-1} = \mathbf{D}_z := \mathrm{diag}\left(\mathbf{z}\right)^{\frac{1}{2}}, \tag{7}$$

where $z_i = i\Delta z$, $i = 1 \cdots n_z$, with Δz the depth sample interval and n_z the number of samples. Combined with the left preconditioner $\mathbf{M}_L^{-1}$, we call this the level II preconditioner. The results for the migrated image in Figure 2(b) now show further improvement. However, amplitude variations remain, e.g., along the major horizontal reflector just above 3500m.

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 nonstationary dip filter—i.e., we have

$$\big(\Psi f\big)(x) \simeq \int_{\xi \in \mathbb{R}^d} e^{j\xi \cdot x} a(x,\xi) \hat{f}(\xi) \mathrm{d}\xi, \tag{8}$$

with Ψ the Hessian of the preconditioned modeling operator and $a(x,\xi)$ a space- and spatial-frequency dependent filter known as the symbol. We use the symbol $\simeq$ to indicate high-frequency approximation and absence of turning waves. Following ideas that go back to Taylor (1981), a pseudodifferential operator can approximately be diagonalized by linear combinations of localized oscillatory functions, such as curvelets (Candès et al., 2006), wave atoms (Demanet and Ying, 2007), and local Fourier bases (Meyer, 1992). In this paper, we use recent results by Herrmann et al. (2008a) who uses curvelets because they have the additional advantage of being sparse on the model. The action of the ΨDO can, after discretization, be approximated by a scaling in the curvelet domain—i.e., we have the following approximate identity

$$\mathbf{\Psi r} \approx \mathbf{C}^* \mathbf{D}_\Psi^2 \mathbf{C r}, \quad \mathbf{D}_\Psi^2 := \mathrm{diag}\left(\mathbf{d}^2\right), \tag{9}$$

which is accurate for a reference vector $\mathbf{r}$ close enough to the actual image. In this expression, the matrices $\mathbf{C}$ and $\mathbf{C}^*$ represent the 2-D discrete curvelet transform (see e.g. Candès et al., 2006) for which the adjoint equals the pseudoinverse—i.e., we have $\mathbf{C}^*\mathbf{C} = \mathbf{I}$ with $\mathbf{I}$ the identity matrix. The reciprocal of the curvelet-domain scaling coefficients, $\mathbf{d}^{-2}$, is found by a remigrated image-to-image matched-filtering procedure that involves the reference vector, typically derived from a conventional migrated image, and the remigrated reference vector. Hence, the cost of forming Equation 9 is approximately one modeling and one migration (for further details refer to Herrmann et al., 2008a,c). As we will show below, these additional costs are well offset by the increase in convergence.

By including the above approximation, we define the right preconditioning matrix as

$$\mathbf{M}_R^{-1} = \mathbf{D}_z \mathbf{C}^* \mathbf{D}_\Psi^{-1}, \tag{10}$$

which compensates for the remaining amplitude errors. Because there is no need for high

accuracy, Equation 9 is a good approximation whose accuracy increases with frequency (Herrmann et al., 2008a). Note, however, that this preconditioner by virtue of the underlying assumptions will not help in the removal of nonlocal imaging artifacts. We call the combination of Equations 6 and 10, the level III preconditioner. The image obtained with this system is plotted in Figure 2(c) and shows, as expected, further improvement in the amplitudes of the migrated image. Despite this improvement, image artifacts and amplitude errors are still present and can be attributed to the approximation in Equation 9 and to the fact that migration does not correspond to inversion—i.e., $\mathbf{A}^*\mathbf{A} \neq \mathbf{I}$.

CONVERGENCE OF LEAST-SQUARES MIGRATION

Even though the different preconditioning operators defined so far lead to improvements, problems remain in achieving high-fidelity images. As reported in the literature, the image quality can be further improved by replacing migration with least-squares migration where the (preconditioned) scattering matrix is inverted iteratively. After a limited number of iterations, this method inverts the (preconditioned) system of equations approximately.

The performance of iterative solvers depends on certain properties of the (preconditioned) matrix $\mathbf{A}$, which include its condition number (ratio of the largest to smallest singular value) and the clustering of the singular values (see e.g. De Roeck, 2002, where these quantities are discussed for a relatively small Kirchhoff-based imaging problem). Because LSQR minimizes the residual during each iteration (cf. Equation 3), studying its progress as a function of the number of iterations gives us some way to gauge the performance. For this purpose, we introduce the normalized log-based least-squares residual $\mu_k = 20 \log \|\widehat{\mathbf{A}}\mathbf{u}_k - \widehat{\mathbf{b}}\|_2 / \|\widehat{\mathbf{b}}\|_2$, with $\mathbf{u}_k$ the solution of the (preconditioned) system after k iterations and with $\mathbf{u}_0 = 0$. Following De Roeck (2002), we also track the log-based

least-squares model-space residual—i.e., $\nu_k = 20 \log \|\widehat{\mathbf{A}}^* \big(\widehat{\mathbf{A}}\mathbf{u}_k - \widehat{\mathbf{b}}\big)\|_2 / \|\widehat{\mathbf{A}}^* \widehat{\mathbf{b}}\|_2$. Note that in practice, these latter residuals are never computed because they are not a by-product of LSQR, and require an additional matrix-vector multiply per iteration. Contrary to data-space residuals, which possibly contain unmodeled components that may not be in the range of the modeling operator, model-space residuals typically converge to zero. Both quantities are used to empirically establish the performance of the different levels of preconditioning. The results of this exercise are summarized in Figure 3, which plots the decay of μ_k and ν_k as a function of the number of iterations k. To account for the overhead, plots for the level III preconditioner are offset by one iteration. This is justified because the cost per iteration of evaluating the curvelet transform, and its inverse, is negligible compared to the cost of migration and demigration.

As we move from a single left preconditioner, towards left and right preconditioners, the data residuals decay faster with a significant improvement obtained by the curvelet-domain scaling. For instance, the residual after 10 iterations for level II preconditioning (fractional integration and depth weighting) is attained by only 5 iterations of level III preconditioning (including curvelet-domain scaling), whereas the result after 10 iterations is approximately 2 dB better. Including the computational overhead, level III preconditioning gains 8 dB in four iterations. Even though the picture for the model-space residual is less clear, there is a similar trend for the level III preconditioning. For instance, after 10 iterations, we have an improvement of approximately 4.5 dB.

The improvements in convergence for the preconditioned system are also reflected in the least-squares migrated results included in Figure 4. Comparing these images shows a clear enhancement for the preconditioned system, plotted in Figure 4(b), over the least-squares result obtained without preconditioning. Moreover, juxtaposing the preconditioned least-

squares image with the solution for the preconditioned migrated image after one iteration depicted in Figure 2(c) shows a significant enhancement of the overall amplitudes and frequency content (cf. Figure 1(a) and 4(b)).

The bottom line is that each iteration takes 40 and 180 minutes for the modeling and migration on an IBM eServer with 52 processors at 2.2GHz. Aside from one additional modeling and migration, the computation of the diagonal estimation takes approximately 90 minutes on a single CPU. The cost of the 2-D curvelet-transforms part of the preconditioning is less than one minute and is negligible compared to the modeling and migration costs.

EXTENSIONS

The preconditioning methodology presented in this letter constitutes a first step towards a concerted effort to formulate seismic imaging and inversion as an optimization problem. This type of formulation allows us to create high-fidelity images and to make progress towards full-waveform inversion. Having access to appropriate preconditioners is instrumental for this purpose because it makes the computations tractable. We envisage the following extensions of our work:

- generalization to 3-D, which entails a different power for the fractional integration, a different spherical-spreading correction, and a 3-D formulation of our curvelet-domain matched filter. Moreover, in 3-D the theory of approximating the Hessian by a ΨDO is less well developed.

- replacement of the level II physical-space preconditioning by more accurate source Green's function illumination corrections (see e.g., Plessix and Mulder, 2004).

- inclusion of density variations, or in case of elastic wave propagation, the inclusion of

variations in the elastic moduli. This extension requires a multi-parameter formulation.

- regularization by curvelet-domain sparsity promotion, replacing Equation 3 by

$$\widetilde{\mathbf{u}}_{\ell_1} = \arg\min_{\mathbf{u}} \|\mathbf{u}\|_1 \quad \text{subject to} \quad \|\widehat{\mathbf{A}}\mathbf{u} - \widehat{\mathbf{b}}\|_2 \leq \epsilon \tag{11}$$

with ϵ a noise-dependent parameter, and $\mathbf{x}_{\ell_1} := \mathbf{M}_R^{-1}\widetilde{\mathbf{u}}_{\ell_1}$. To benefit from our preconditioning, Equation 11 requires a solver with Newton-type steps, as opposed to most ℓ_1-norm solvers that are based on projected gradients. The advantage of this formulation is that it uses curvelet-domain sparsity, which has proven to be a particularly powerful prior (Wang and Sacchi, 2007; Herrmann et al., 2008b,a; Hennenfent et al., 2008). We will report on the solution of Equation 11 elsewhere.

- incorporation of our preconditioner in sparsity-promoting full-waveform inversion. This problem requires the solution of the unconstrained optimization problem

$$\min_{\mathbf{z}} \frac{1}{2}\|\widehat{\mathbf{b}} - \widehat{\mathbf{F}}[\mathbf{z}]\|_2^2 + \lambda\|\mathbf{z}\|_1, \tag{12}$$

where $\mathbf{z}$ is the curvelet representation of the model, λ a Lagrange multiplier balancing the residual energy and ℓ_1-norm penalty term, and $\widehat{\mathbf{F}}[\mathbf{z}]$ the nonlinear forward map that links the curvelet-domain model to data. Again, the solution of this optimization problem requires a sophisticated solver that can benefit from our preconditioner.

CONCLUSIONS

Because of the size of the seismic imaging problem, preconditioning of least-squares migration is an elusive topic, where traditional approaches from numerical linear algebra have not yet found their way. Lack of direct access to the matrices involved and the cost of

evaluating matrix-free implementations of the operators are both to blame. Nonetheless, the first few iterations of the LSQR algorithm of least-squares migration are known to make significant progress towards the solution. Unfortunately, even for this limited number of iterations, the computational costs are often still prohibitively large for practical problems. The method presented in this paper partly resolves this issue through a combination of left and right preconditioning together with a curvelet-domain scaling. Inclusion of the latter proved particularly important because it restores the amplitudes and leads to faster convergence, at a relatively small computational overhead. Aside from this tangible reduction in computational costs of roughly 50%, the use of curvelet frames opens the enticing perspective to use ℓ_1-norm regularization to improve the quality of images. Preconditioning also plays a pivotal role in making this approach numerically feasible and may extend to a solution of the full-waveform inversion problem with sparsity promotion. Both approaches are justified by ample evidence that curvelets are sparse on the model.

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 Tamas Nemeth and Michael Friedlander for insightful discussions and 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 in part carried out within the SINBAD project with support, secured through ITF, from BG Group, BP, Chevron, ExxonMobil and Shell. We finally thank the anonymous reviewer and W. W. Symes for their constructive

comments and suggestions that greatly improved our paper.

REFERENCES

Aminzadeh, F., J. Brac, and T. Kunz, 1997, 3-D Salt and Overthrust Model. SEG/EAGE 3-D Modeling Series, No. 1: Society of Exploration Geophysicists, Tulsa.

Candès, E. J., L. Demanet, D. L. Donoho, and L. Ying, 2006, Fast discrete curvelet transforms: SIAM Multiscale Modeling and Simulation, **5**, 861–899.

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 SEP-82, 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.

Demanet, L. and L. Ying, 2007, Wave atoms and sparsity of oscillatory patterns: Journal of Applied and Computational Harmonic Analysis, **23**, 368–387.

———, 2008, Discrete symbol calculus. (Submitted for publication. Available at `http://math.stanford.edu/~laurent/papers/DSC.pdf`).

Guitton, A., 2004, Amplitude and kinematic corrections of migrated images for nonunitary imaging operators: Geophysics, **69**, 1017–1024.

Hennenfent, G., E. van den Berg, M. P. Friedlander, and F. J. Herrmann, 2008, New insights into one-norm solvers from the Pareto curve: Geophysics, **73**, no. 4, A23–A26.

Herrmann, F. J., P. P. Moghaddam, and C. C. Stolk, 2008a, Sparsity- and continuity-promoting seismic imaging with curvelet frames: Journal of Applied and Computational Harmonic Analysis, **24**, 150–173.

Herrmann, F. J., D. Wang, G. Hennenfent, and P. P. Moghaddam, 2008b, Curvelet-based seismic data processing: a multiscale and nonlinear approach: Geophysics, **73**, no. 1,

A1–A5.

Herrmann, F. J., D. Wang, and D. J. E. Verschuur, 2008c, Adaptive curvelet-domain primary-multiple separation: Geophysics, **73**, A17–A21.

Hu, J., G. T. Schuster, and P. A. Valasek, 2001, Poststack migration deconvolution: Geophysics, **66**, 939–952.

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

Meyer, Y., 1992, Wavelets and operators: Cambridge University Press.

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

O'Brien, M. and S. Gray, 1996, Can we image beneath salt?: The Leading Edge, **15**, 17–22.

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

Plessix, R. and W. Mulder, 2004, Frequency-domain finite difference amplitude-preserving migration: Geophysical Journal International, **157**, 975–987.

Rickett, J. E., 2003, Illumination-based normalization for wave-equation depth migration: Geophysics, **68**, 1371–1379.

Stolk, C. C., 2000, Microlocal analysis of a seismic linearized inverse problem: Wave Motion, **32**, 267–290.

Symes, W. W., 2007, Reverse time migration with optimal checkpointing: Geophysics, **72**, SM213–SM221.

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

Taylor, M., 1981, Pseudodifferential operators: Princeton University Press.

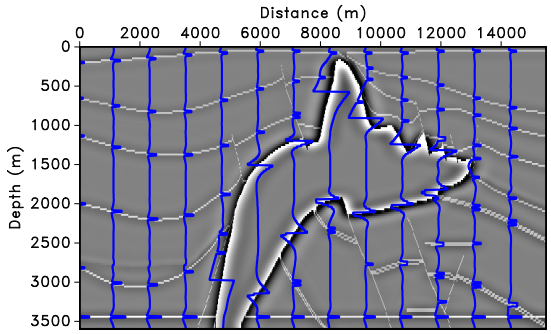
Vogel, C., 2002, Computational Methods for Inverse Problems: SIAM.

Wang, D., R. Saab, O. Yilmaz, and F. J. Herrmann, 2008, Bayesian wavefield separation by transform-domain sparsity promotion: Geophysics, **73**, A33–A38.

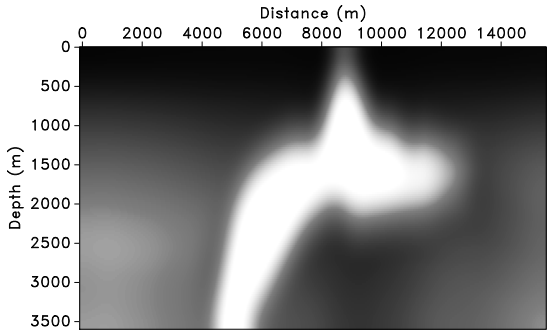
Wang, J. and M. D. Sacchi, 2007, High-resolution wave-equation amplitude-variation-with-ray-parameter (avp) imaging with sparseness constraints: Geophysics, **72**, S11–S18.

Yu, J., J. Hu, G. T. Schuster, and R. Estill, 2006, Prestack migration deconvolution: Geophysics, **71**, S53–S62.

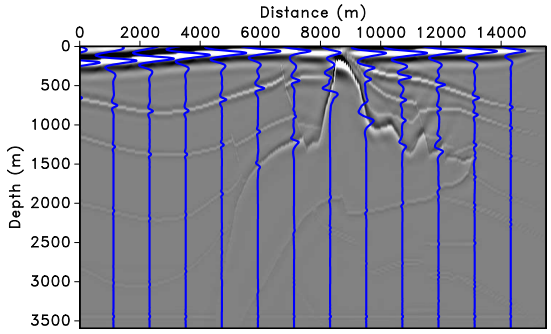
LIST OF FIGURES



(a)

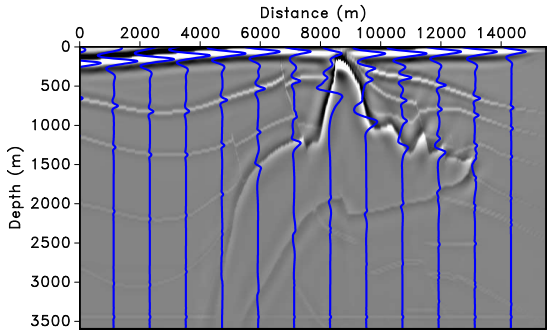


(b)

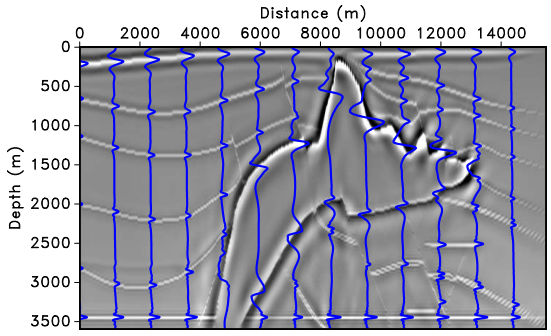


(c)

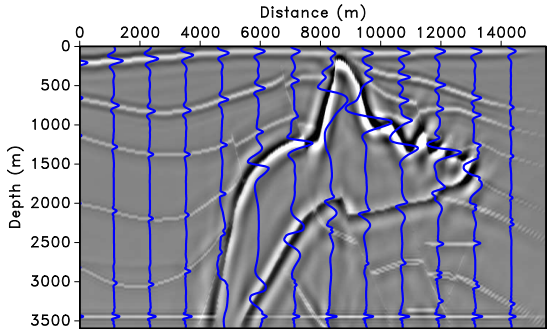
Figure 1: The SEG/EAGE AA′ salt model. **(a)** Reflectivity defined by the high-pass filtered velocity model. **(b)** Smoothed velocity model. **(c)** The migrated image according to Eq. 1. This image suffers from deteriorated amplitudes, especially under the high-velocity salt and for steep reflectors and faults.



(a)

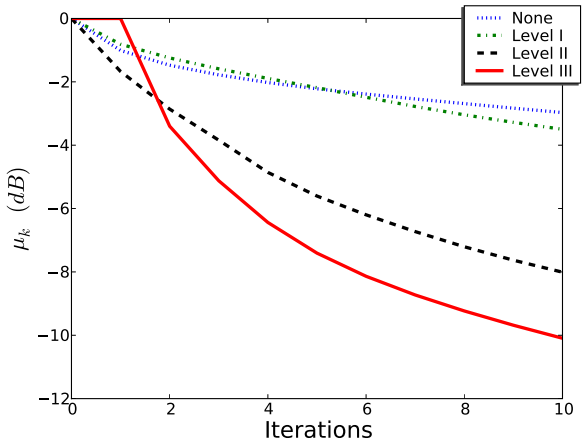


(b)

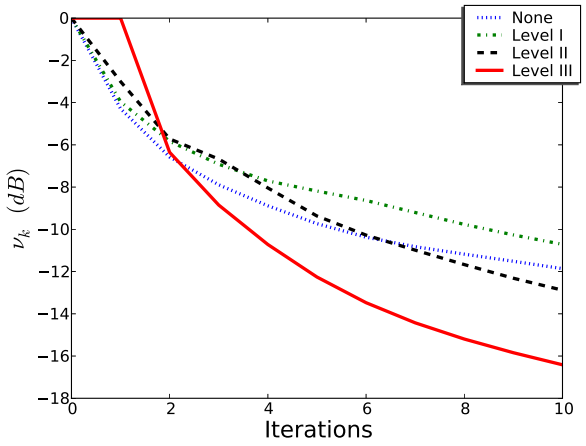


(c)

Figure 2: Migrated images for different levels of preconditioning. **(a)** Result for left preconditioning (level I, cf. Equation 6). **(b)** Result for left-right (including depth-correction) preconditioning (level II, cf. Equation 7). **(c)** The same but now including curvelet-domain scaling (level III, cf. Equation 10).

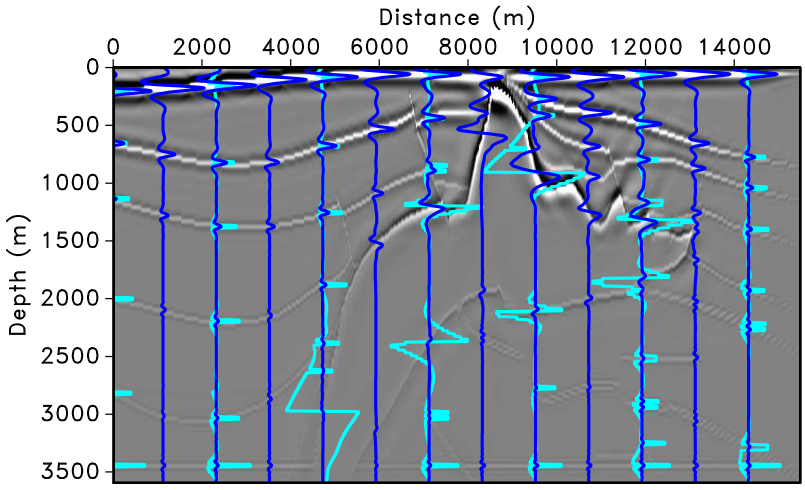


(a)

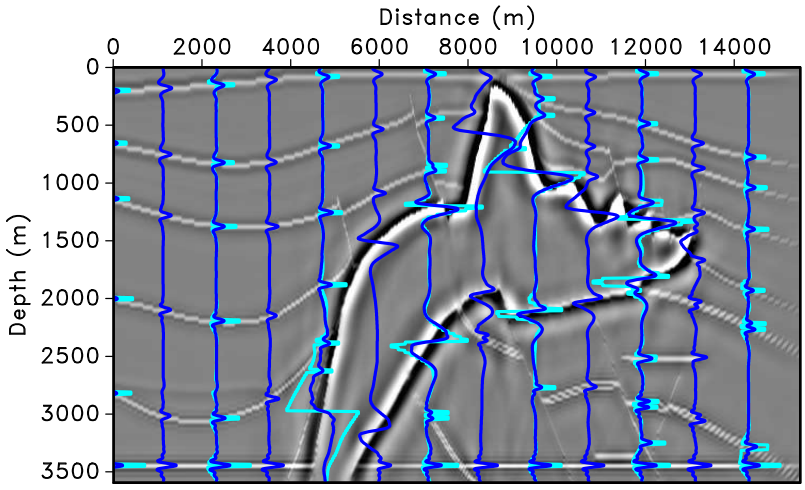


(b)

Figure 3: Residual decays for different levels of preconditioning. The dotted blue lines corresponds to least-squares migration without preconditioning, the dash-dotted lines to level I preconditioning, the dashed black lines to level II preconditioning, and the red solid lines to level III preconditioning. This is offset by one iteration to account for the overhead. **(a)** Plot for the decay of the data-space normalized residues μ_k as a function of the number of LSQR iterations. **(b)** The same but now for the model-space normalized residuals ν_k.



(a)



(b)

Figure 4: Least-squares migration without and with preconditioning juxtaposed with the original reflectivity (in light blue). **(a)** Least-squares migrated image. **(b)** Least-squares image with level III preconditioning. Notice the improvement in the recovered reflectivity.