Seismic wavefield inversion with curvelet-domain sparsity promotion

Felix J. Herrmann[*]*, EOS-UBC and Deli Wang*[†]*, Jilin University*

SUMMARY

Inverting seismic wavefields lies at the heart of seismic data processing and imaging— whether one is applying "a poor man's inverse" by correlating wavefields during imaging or whether one inverts wavefields as part of a focal transform interferrometric deconvolution or as part of computing the 'data inverse'. The success of these wavefield inversions depends on the stability of the inverse with respect to data imperfections such as finite aperture, bandwidth limitation, and missing data. In this paper, we show how curvelet domain sparsity promotion can be used as a suitable prior to invert seismic wavefields. Examples include, seismic data regularization with the focused curvelet-based recovery by sparsity-promoting inversion (fCRSI), which involves the inversion of the primary-wavefield operator, the prediction of multiples by inverting the adjoint of the primary operator, and finally the inversion of the data itself — the so-called 'data inverse'. In all cases, curvelet-domain sparsity leads to a stable inversion.

INTRODUCTION

In this paper, we demonstrate that the discrete curvelet transform (Candes et al., 2006; Hennenfent and Herrmann, 2006) can be used to invert seismic wavefields stably, even in case where the data volumes are sampled incompletely. The crux of our method lies in the combination of the curvelet transform, which attains a fast decay for the magnitude-sorted curvelet coefficients for arbitrary wavefields (see e.g. Candes et al., 2006; Hennenfent and Herrmann, 2006; Herrmann et al., 2008; Herrmann and Hennenfent, 2008, and the references therein), with a sparsity promoting program. By themselves sparsity-promoting programs are not new to the geosciences (Sacchi et al., 1998). However, sparsity promotion with the curvelet transform is relatively new (see e.g. Herrmann et al., 2008, for an overview). The curvelet transform's unparalleled ability to detect wavefront-like events that are locally linear and coherent means it is particularly well suited to seismic data problems. In this paper, we show how this transform can be used to regularize the inversion of seismic wavefields. This type of inversion proves difficult in practice because of the problem size, finite aperture, source/receiver signatures and the presence of noise. By using 3-D curvelets in the shot-receiver-time domain, we leverage continuity along multidimensional wavefronts maximally. As opposed to damped least-squares— a popular method for the regularization of geophysical inverse problems at the expense of additional smoothing— curvelet-domain sparsity promotion preserves wavefronts. This property explains our recent successes applying this strategy to synthetic and real field data with applications ranging from wavefield reconstruction (Herrmann and Hennenfent, 2008; Hennenfent and Herrmann, 2008), wavefield separation, migration amplitude recovery (Herrmann et al., 2008), and compressed wavefield extrapolation (Lin and Herrmann, 2007).

In this paper, we continue to leverage curvelet-domain sparsity promotion towards wavefield inversion with applications including the inversion of primary wavefields part of focusing, the inversion of the adjoint of the primary wavefield part of defocusing for multiple prediction, and finally the stable computation of Berkhout's data inverse Berkhout (2006); Berkhout and Verschuur (2007). First, we briefly introduce the curvelet transform, followed by a common-problem formulation for curvelet-based wavefield inversion (CWI) by sparsity promotion and its application.

CURVELETS

Curvelets are localized 'little plane-waves' (see e.g. Hennenfent and Herrmann, 2006) that are oscillatory in one direction and smooth in the other direction(s). They are multiscale and multi-directional. Curvelets have an anisotropic shape— they obey the so-called parabolic scaling relationship, yielding a width $\propto$ length^2 for the support of curvelets in the physical domain. This anisotropic scaling is necessary to detect wavefronts and explains their high compression rates on seismic data and images, as long as these datasets can be represented as functions with events on piece-wise twice differentiable curves. Then, the events become linear at the fine scales justifying an approximation by the linearly shaped curvelets. Even seismic data with caustics, pinch-outs, faults or strong amplitude variations fit this model, which amounts to a preservation of the sparsity attained by curvelets.

Curvelets represent a specific tiling of the 2-D/3-D frequency domain into strictly localized wedges. Because the directional sampling increases every-other scale doubling, curvelets become more anisotropic at finer scales. Curvelets compose multi-D data according to $\mathbf{f} = \boldsymbol{C}^T \boldsymbol{C}\mathbf{f}$ with $\boldsymbol{C}$ and $\boldsymbol{C}^T$ the forward and inverse discrete curvelet transform matrices (defined by the fast discrete curvelet transform, FDCT, with wrapping, a type of periodic extenstion, see Candes et al., 2006; Ying et al., 2005). The symbol ^{T} represents the transpose, which is equivalent to the inverse for this choice of curvelet transform. This transform has a moderate redundancy (a factor of roughly 8 in 2-D and 24 in 3-D) and a computational complexity of $\mathscr{O}(n \log n)$ with n the length of $\mathbf{f}$.

COMMON PROBLEM FORMULATION

Curvelet based inversion by sparsity promotion: Our solution strategy is built on the premise that seismic data and images have a sparse representation, $\mathbf{x}_0$, in the curvelet domain. To exploit this property, our forward model reads

$$\mathbf{y} = \boldsymbol{A}\mathbf{x}_0 + \mathbf{n} \tag{1}$$

with $\mathbf{y}$ a vector of noisy and possibly incomplete measurements; $\boldsymbol{A}$ the modeling matrix that includes $\boldsymbol{C}^T$, and $\mathbf{n}$, a zero-centered white Gaussian noise. Because of the redundancy of $\boldsymbol{C}$ and/or the incompleteness of the data, the matrix $\boldsymbol{A}$ can not readily be inverted. However, as long as the data, $\mathbf{y}$, permits a

sparse vector, $\mathbf{x}_0$, the matrix, $\boldsymbol{A}$, can be inverted by a sparsity-promoting program (Candès et al., 2006; Donoho, 2006):

$$\mathbf{P}_\varepsilon : \quad \begin{cases} \widetilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\boldsymbol{A}\mathbf{x} - \mathbf{y}\|_2 \leq \varepsilon \\ \widetilde{\mathbf{f}} = \boldsymbol{S}^T \widetilde{\mathbf{x}} \end{cases} \tag{2}$$

in which ε is a noise-dependent tolerance level, $\boldsymbol{S}^T$ the inverse transform and $\widetilde{\mathbf{f}}$ the solution calculated from the vector $\widetilde{\mathbf{x}}$ (the symbol $\widetilde{\ }$ denotes a vector obtained by nonlinear optimization) minimizing $\mathbf{P}_\varepsilon$. The difference between $\widetilde{\mathbf{x}}$ and $\mathbf{x}_0$ is proportional to the noise level.

Nonlinear programs $\mathbf{P}_\varepsilon$ are not new to seismic data processing as in spiky deconvolution (Taylor et al., 1979; Santosa and Symes, 1986) and Fourier transform-based interpolation (Sacchi et al., 1998). The curvelets' high compression rate makes the nonlinear program $\mathbf{P}_\varepsilon$ perform well when $\boldsymbol{C}^T$ is included in the modeling operator. Despite its large-scale and nonlinearity, the solution of the convex problem $\mathbf{P}_\varepsilon$ can be approximated with a limited (< 250) number of iterations of a threshold-based cooling method derived from work by Figueiredo and Nowak (2003); Daubechies et al. (2005); Elad et al. (2005). At each iteration the descent update ($\mathbf{x} \leftarrow \mathbf{x} + \boldsymbol{A}^T(\mathbf{y} - \boldsymbol{A}\mathbf{x})$), minimizing the quadratic part of Equation 2, is followed by a soft thresholding ($\mathbf{x} \leftarrow T_\lambda(\mathbf{x})$ with $T_\lambda(x) := \mathrm{sgn}(x) \cdot \max(0, |x| - |\lambda|)$) for decreasing threshold levels λ. This soft thresholding on the entries of the unknown curvelet vector captures the sparsity and the cooling, which speeds up the algorithm, allows additional coefficients to fit the data.

Wavefield inversion: Following Berkhout's work on the focal transform (Berkhout and Verschuur, 2006), we introduce seismic data volumes as operators—i.e., we define the following linear wavefield operators $\boldsymbol{V}\cdot = \boldsymbol{F}^H \mathrm{blockdiag}(\widehat{\mathbf{V}})\boldsymbol{F}\cdot$ and $\boldsymbol{V}^H\cdot = \boldsymbol{F}^H \mathrm{blockdiag}(\widehat{\mathbf{V}}^H)\boldsymbol{F}\cdot$ with $\widehat{\mathbf{V}}$ the temporal Fourier-domain data matrix, $\mathrm{blockdiag}(\widehat{\mathbf{V}})$ the multi-frequency data-matrix operator —i.e., data is organized as a block-diagonal matrix with each individual block containing a single monochromatic shot-receiver gather, ^{H} the Hermitian transpose and $\boldsymbol{F}$ the temporal Fourier transform. Applying this linear operator to a wavefield collected in a data matrix (a tall vector with the monochromatic blocks) corresponds to applying the temporal Fourier transform, followed by a matrix-matrix multiplication for each frequency (this implements a multidimensional convolution) and an inverse Fourier transform. To simplify notation, we refer to the block-diagonal matrix and wavefield vector interchangeably. For primary wavefields—i.e., $\boldsymbol{V} = \boldsymbol{\Delta P}$ with $\boldsymbol{\Delta P}$ the primary wavefield (to be more precise, the wavefield without surface related multiples but with internal multiples), $\boldsymbol{\Delta P P}$ (ignoring surface related effects) adds one interaction with the surface and turns primaries into first-order surface-related multiples. Conversely, applying the pseudo inverse of $\boldsymbol{\Delta P}$ to first-order multiples yields primaries. To compute this inverse stably, we propose, with some abuse of notation, to compound the curvelet synthesis matrix with the wavefield operator—i.e.,

$$\begin{aligned} \mathbf{y} &= \overbrace{\boldsymbol{V}\boldsymbol{C}^T}^{\boldsymbol{A}} \mathbf{x}_0 \\ \mathrm{vec}(\boldsymbol{U}) &= \mathrm{vec}\left(\boldsymbol{V}\,\mathrm{vec}^{-1}(\boldsymbol{C}^T\mathbf{x}_0)\right), \end{aligned}$$

where the linear operation vec reorganizes the data matrix into a long vector and vec^{-1} reorganizes a data vector into a data matrix. Since seismic wavefields compress in the curvelet domain, we can by solving $\mathbf{P}_\varepsilon$ find the set of curvelet coefficients whose inverse curvelet transform, acted upon by the wavefield, generates the data—i.e., the wavefield $\boldsymbol{U}$, to some tolerance ε. This solution involves the inversion of the compound operator $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{C}^T$, which corresponds to a curvelet-regularized (through sparsity promotion) inversion of the wavefield $\boldsymbol{V}$, given the wavefield $\boldsymbol{U}$ (reorganized in the vector $\mathbf{y}$). By choosing the wavefields $\boldsymbol{U}$ and $\boldsymbol{V}$ appropriately, we can solve different problems in seismic imaging.

FOCUSED WAVEFIELD RECOVERY

To illustrate the stability of our curvelet-based formulation of wavefield inversion, consider the recovery of seismic wavefields with the curvelet-regularized focal transform with the observed data vector and modeling operator given by $\mathbf{y} = \boldsymbol{R}\boldsymbol{P}$ $= \mathrm{vec}(\boldsymbol{R}_s\boldsymbol{P}\boldsymbol{R}_r^T) = (\boldsymbol{R}_r \otimes \boldsymbol{R}_s)\mathrm{vec}(\boldsymbol{P})$ with $\boldsymbol{R}_{r,s}$ restrictions in the receiver and source coordinates, $\otimes$ the outer product, $\boldsymbol{P}$ the total wavefield and $\boldsymbol{A} = \boldsymbol{R}\boldsymbol{\Delta P}\boldsymbol{C}^T$ the modeling matrix with $\boldsymbol{\Delta P}$ the primary-wavefield operator.

According to these definition, the solution of $\mathbf{P}_\varepsilon$ corresponds to a curvelet-regularized focal transform during which the restricted primary operator is inverted, given incomplete data. As such, we 'deconvolve' the incomplete data with the (incomplete) primaries, yielding an additional focusing of the energy by converting first-order multiples to primaries and primaries to a directional line source. This focusing corresponds to a collapse of 3-D primary events onto an approximate line source, which has a sparser representation in the curvelet domain. Applying the inverse curvelet transform, followed by 'convolution' with $\boldsymbol{\Delta P}$, yields the interpolation, i.e. $\boldsymbol{S}^T := \boldsymbol{\Delta P}\boldsymbol{C}^T$. Comparing the curvelet recovery with the focused curvelet recovery (Figure 1(c) and 1(d)) shows an overall improvement in the recovered details.

DEFOCUSSED MULTIPLE PREDICTION

A second example where the inversion of wavefields may be useful is in multiple prediction. During standard Surface-Related Multiple Elimination (SRME, Verschuur et al., 1992), multiples are predicted by multidimensional convolutions of the data matrix with an estimate for the primaries (or the data itself). As a result, the 'source wavelet' appears twice making a subsequent global wavelet matching necessary to remove this wavelet from the prediction. By choosing the wavefield in terms of the total data, $\boldsymbol{U} = \boldsymbol{P}$ and the operator $\boldsymbol{V}$ as the adjoint of the primary operator—i.e., $\boldsymbol{V} = \boldsymbol{\Delta P}^H$, solving $\mathbf{P}_\varepsilon$ with $\boldsymbol{S}^T = \boldsymbol{C}^T$ yields an estimate for the multiples obtained by 'deconvolving' the data matrix with the adjoint of the primaries. Since this procedure involves inverting $\boldsymbol{\Delta P}^H$, we can expect a deconvolution of the 'source wavelet'. Indeed, we observe an increase in the frequency content of multiples predicted by solving $\mathbf{P}_\varepsilon$, compared to calculating the multiples via $\boldsymbol{P P}$. The observed artifacts are likely caused by remnant multiple energy and 3-D effects present in the primary wavefield, which we used to define the operator used to predict the multiples. Also notice the improved amplitudes for the far offsets.

DATA INVERSE

As a final example, we show how the presented methodology can be used to calculate the inverse data space. As shown by Berkhout (2006), the data inverse leads to a natural separation of the primary wavefield and the surface related effects—i.e. the multiple-generation boundary condition at the surface and the source and receiver characteristics. Mathematically, the data 'inverse' (see also Sheng, 1995), which represents the inverse of the forward scattering series, can be written as

$$\boldsymbol{P}^{\dagger} = \boldsymbol{\Delta P}^{\dagger} - \mathscr{A}, \tag{3}$$

where the symbol † denotes the pseudo inverse rather than an ordinary inverse and where $\mathscr{A}$ contains the boundary condition at the surface and the inverse of the source and receiver signatures (see e.g. Berkhout and Verschuur, 2006). With this example, we illustrate how curvelet regularization overcomes practical difficulties related to computing wavefield inverses on (real) data.

Again, our curvelet-based formulation comes to our rescue by setting $\boldsymbol{V} = \boldsymbol{P}$ and $\boldsymbol{U} = \boldsymbol{I}\mathbf{d}_{\Psi}$ with $\boldsymbol{I}\mathbf{d}_{\Psi}$ the bandwidth- and dip-limited delta-Dirac line source. As can be seen in Figure. 3, solving $\mathbf{P}_{\varepsilon}$ for this setting (with $\boldsymbol{S}^T = \boldsymbol{C}^T$), yields a stable estimate for the data inverse of real data (Figure 3(a) contains a shot of this data volume). As we can see, the data inverse contains mostly acausal energy (as expected, see Figure 3(a)) with a strong bandwidth- and dip-limited pulse (actually a line source) at zero time. This latter contribution corresponds to the term $\mathscr{A}$ in Equation 3, which contains the surface-related multiples. To verify whether this property holds, we also invert our SRME-estimate for the primaries, $\boldsymbol{\Delta P}$. For this primary wavefield, the large contribution of the directional 'source' $\mathscr{A}$ is removed, which means that the multiple-generating boundary condition was successfully removed. This removal confirms the validity of the concept of the data inverse on real data. However, artifacts are present in these results and these are mostly due to remnant multiple energy and 3-D effects.

DISCUSSION AND CONCLUSIONS

In this paper, we presented three different examples that involve the inversion of seismic wavefields. We showed that curvelet-based sparsity promotion leads to inverses that are stable with respect to missing data, finite aperture and bandwidth limitation. Wavefield reconstruction was improved by inverting the primary operator, which leads to an increased recovery because of focusing towards the source. This focusing is induced by inverting the primary-wavefield operator. Conversely, inverting the adjoint of the primary-wavefield operator restores the frequency content and far-offset amplitudes of the predicted multiples. Finally, we also showed that our curvelet-sparsity promoting formulation can be used on real data to compute the data inverse, arguably the most challenging of the three examples. In that case, the surface related multiples are focused to a directional source while, as expected, most of the primaries (and internal multiples) are mapped to negative times.

Our findings are encouraging and may have profound implications on how wavefields are inverted during (interferrometic) imaging. The fact that the focal transform images 'the source'—i.e. the primaries are mapped to a directional line source, which corresponds to prestack reflectivity when applying the focal transform after redatuming, supports this claim. This means that the presented formulation can be used to replace the current 'poor man's' inverse— through multidimensional correlation— by wavefield deconvolution. The examples on real data presented in this paper show that further application of this methodology on real data is well within reach.

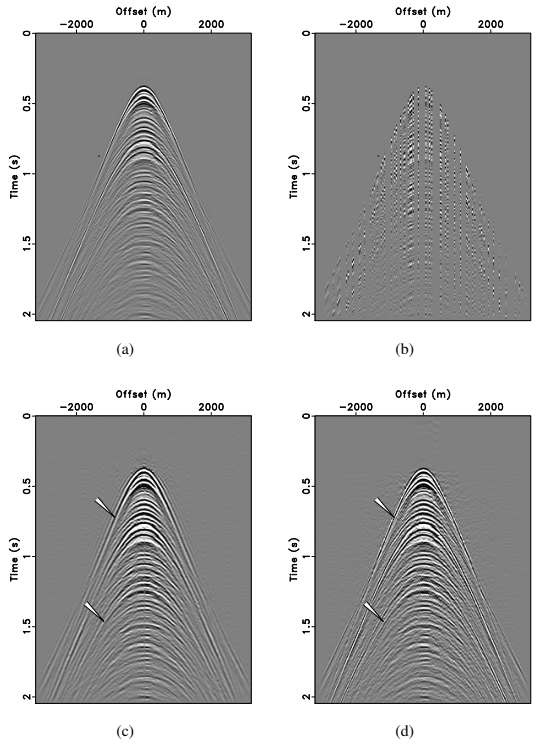

Figure 1: Comparison between 3-D curvelet-based recovery by sparsity-promoting inversion with and without focusing. **(a)** Fully sampled real North Sea field data shot gather. **(b)** Randomly subsampled shot gather from a 3-D data volume with 80% of the traces missing in the receiver and shot directions. **(c)** Curvelet-based recovery. **(d)** Curvelet-based recovery with focusing. Notice the improvement (denoted by the arrows) from the focusing with the primary operator.

ACKNOWLEDGMENTS

The authors would like to thank Eric Verschuur for providing us with the SRME-primaries. We also would like to thank the authors of CurveLab for making their codes available. The examples presented were prepared with Madagascar supplemented by SLIMPy operator overloading, developed by Sean Ross-Ross and Henryk Modzelewski. Norsk Hydro is thanked for making the field dataset available. In addition, this work was in part financially supported by the NSERC Discovery Grant (22R81254) of F.J.H. and CRD Grant DNOISE (334810-05), and was carried out as part of the SINBAD project with support, secured through ITF, from BG Group, BP, Chevron, ExxonMobil and Shell.

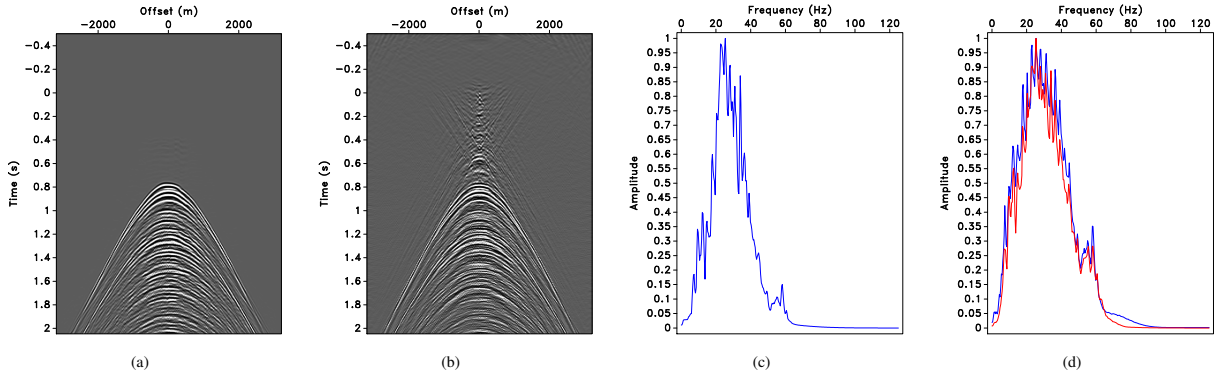

Figure 2: Comparison between 3-D curvelet-based multiple prediction (via $\mathbf{P}_{\mathcal{E}}$ and $\boldsymbol{A} = \boldsymbol{\Delta P}^H \boldsymbol{C}^T$ and $\mathbf{y} = \mathrm{vec}(\boldsymbol{P})$). **(a)** A shot from the conventional multiple prediction according to $\boldsymbol{PP}$. **(b)** The same but not now by inverting $\boldsymbol{\Delta P}^T$. **(c)** The amplitude-normalized stacked temporal Fourier spectrum of the multiple prediction plotted in **(a)**. **(d)** The amplitude-normalized stacked temporal Fourier spectra of the total data (see Fig. 3(a)) and of the multiple prediction obtained by defocusing. Notice the improvement in the frequency content (e.g. compare the red line for the spectrum of the total data with the blue line for the spectrum of our estimate) and in the large-offset amplitudes for the multiple predictions obtained by defocusing. The artifacts in the defocused result are due to remnant multiple energy in the primary wavefield whose adjoint is used to calculate the multiples.

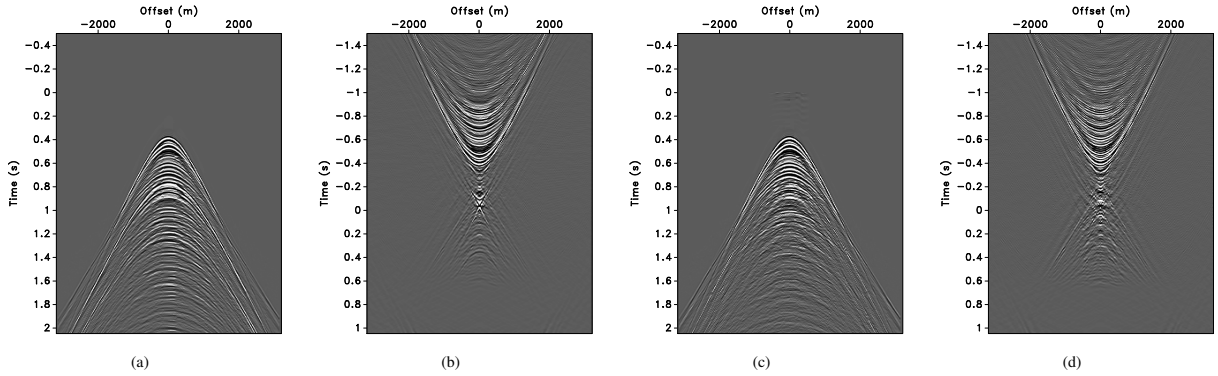

Figure 3: Example of computing the 'data inverse' with 3-D curvelet-based sparsity promotion. (via $\mathbf{P}_{\mathcal{E}}$ with $\boldsymbol{A} = \boldsymbol{PC}^T$ and $\mathbf{y} = \mathrm{vec}(\boldsymbol{I}\mathbf{d}_{\Psi})$). **(a)** A shot from the total data volume. **(b)** Corresponding shot from the estimate for the data inverse of $\boldsymbol{P}$. **(c)** A shot from the SRME-estimate for the primaries. **(d)** Corresponding shot for the 'data inverse' of the SRME-predicted primaries. Notice that the directional line source at time zero for the 'data inverse' of the total data is more or less completely absent in the 'data inverse' for the SRME-primaries, an observation consistent with the fact that surface-related multiple energy maps to $\mathscr{A}$.

REFERENCES

Berkhout, A. and E. Verschuur, 2007, Seismic processing in the inverse data space, removal of surface-related and internal multiples: Presented at the EAGE 69th Conference & Exhibition, London, B036.

Berkhout, A. J., 2006, Seismic processing in the inverse data space: Geophysics, **71**.

Berkhout, A. J. and D. J. Verschuur, 2006, Focal transformation, an imaging concept for signal restoration and noise removal: Geophysics, **71**, 1596–1611.

Candès, E., J. Romberg, and T. Tao, 2006, Stable signal recovery from incomplete and inaccurate measurements: Comm. Pure Appl. Math., **59**, 1207–1223.

Candes, E. J., L. Demanet, D. L. Donoho, and L. Ying, 2006, Fast discrete curvelet transforms: SIAM Multiscale Model. Simul., **5**, 861–899.

Daubechies, I., M. Defrise, and C. de Mol, 2005, An iterative thresholding algorithm for linear inverse problems with a sparsity constrains: CPAM, 1413–1457.

Donoho, D. L., 2006, Compressed sensing: IEEE Trans. Inform. Theory, **52**, 1289–1306.

Elad, M., J. L. Starck, P. Querre, and D. L. Donoho, 2005, Simulataneous Cartoon and Texture Image Inpainting using Morphological Component Analysis (MCA): Appl. Comput. Harmon. Anal., **19**, 340–358.

Figueiredo, M. and R. Nowak, 2003, An EM algorithm for wavelet-based image restoration: IEEE Trans. Image Processing, **12**, 906–916.

Hennenfent, G. and F. J. Herrmann, 2006, Seismic denoising with non-uniformly sampled curvelets: IEEE Comp. in Sci. and Eng., **8**, 16–25.

———, 2008, Simply denoise: wavefield reconstruction via jittered undersampling: Geophysics, **73**.

Herrmann, F., D. Wang, G. Hennenfent, and P. Moghaddam, 2008, Curvelet-based seismic data processing: a multiscale and nonlinear approach: Geophysics, **73**, A1–A5. (doi:10.1190/1.2799517).

Herrmann, F. J. and G. Hennenfent, 2008, Non-parametric seismic data recovery with curvelet frames: Geophysical Journal International, **173**, 223–248. (doi: 10.1111/j.1365-246X.2007.03698.x).

Lin, T. and F. J. Herrmann, 2007, Compressed wavefield extrapolation: Geophysics, **72**, SM77–SM93.

Sacchi, M. D., T. J. Ulrych, and C. Walker, 1998, Interpolation and extrapolation using a high-resolution discrete Fourier transform: **46**, 31–38.

Santosa, F. and W. Symes, 1986, Linear inversion of band-limited reflection seismogram: SIAM J. of Sci. Comput., **7**.

Sheng, P., 1995, Introduction to wave scattering, localization, and mesoscopic phenomena: Academic Press.

Taylor, H. L., S. Banks, and J. McCoy, 1979, Deconvolution with the ℓ_1 norm: Geophysics, **44**, 39.

Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, 1166–1177.

Ying, L., L. Demanet, and E. J. Candés, 2005, 3D discrete curvelet transform: Wavelets XI, Expanded Abstracts, 591413, SPIE.