Seismic Denoising with Nonuniformly Sampled Curvelets

The authors present an extension of the fast discrete curvelet transform (FDCT) to nonuniformly sampled data. This extension not only restores curvelet compression rates for nonuniformly sampled data but also removes noise and maps the data to a regular grid.

Recently introduced, *curvelets* are among the latest members of a growing family of multiscale—and now multidirectional—data expansions.[1,2] The primary aim of these expansions is to find a sparse representation for data. A signal representation is sparse when it can capture a signal as a superposition of a small number of components—in fact, the sparser and more generic the transformation, the more successful its use for signal denoising and separation.

So what makes curvelet decomposition an appropriate transform for seismic data processing, and why generalize this transform to nonuniformly sampled data? To answer these questions, we must first understand seismic data itself. Seismic data volumes are recordings of the amplitudes of transient waves—either human-made or from naturally occurring earthquakes—at the Earth's surface. Each source and receiver pair generates a *trace*, which is a function of time. A seismic data set is the collection of these traces; all the traces together provide a spatio-temporal sampling of the reflected *wave field*, which contains different arrivals that correspond to different interactions of the incident wave field with inhomogeneities in the Earth's subsurface. A common denominator among these arrivals is that they represent *wave fronts*, the main characteristic of which is their relative smoothness in the direction along the front and their relative oscillatory behavior in the "normal" direction.

By virtue of their anisotropic shape, curvelets are well adapted to detect wave fronts because aligned curvelets correlate well with them locally. In this sense, curvelets act as multiscale surfboards riding incoming waves. However, limitations on data acquisition due to the positions of sources and receivers put restrictions on the spatial sampling of seismic wave fields—in seismic exploration on land, for example, there are obstacles such as buildings or lakes, whereas in passive seismology, the seismologist has no control over source position (earthquakes occur irregularly along major plate boundaries).

The current implementations of the fast discrete curvelet transform (FDCT) assume a regular sampling along all axes. If we ignore the nonuniformity of spatial sampling, we can no longer expect to detect wave fronts because of a lack of continuity. We've addressed this issue by extending the FDCT to nonuniformly sampled data. Through this extension, we can detect wave fronts in noise as well as bring the data to a regular grid provided there is at least one datum per grid point. The example in Figure 1 clearly illustrates how continuity along

1521-9615/06/$20.00 © 2006 IEEE
Copublished by the IEEE CS and the AIP

Gilles Hennenfent and Felix J. Herrmann
University of British Columbia

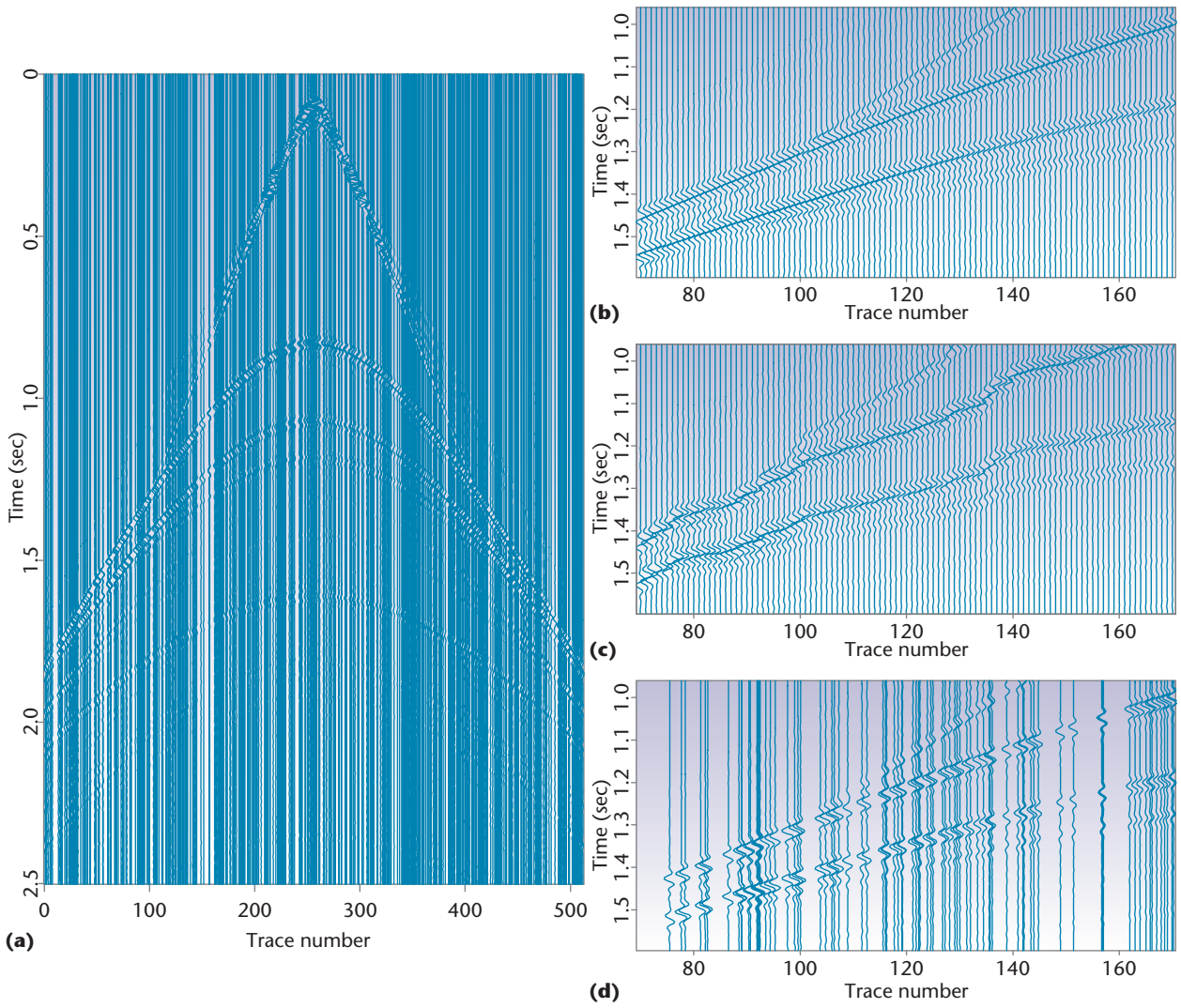

Figure 1. Synthetic seismic data. In examining (a) uniformly (grayscale plot) and nonuniformly sampled (wiggle trace plot) data, (b) windowed regular sampled data, (c) windowed irregular sampled data cast to a regular grid, and (d) windowed data on the nonuniformly sampled grid, we can see continuity along the arriving wave fronts in (b) and (d). Recasting irregular data onto a regular grid destroys this continuity in (c). Here, we've exaggerated the nonuniformly sampled grid's irregularity.

wave fronts is destroyed when casting nonuniformly sampled data to a regular grid, but restored when dealing with the data appropriately. Our denoising and binning algorithm exploits the sparsity of seismic data in the curvelet domain via a nonlinear thresholding on curvelet coefficients. (The term *binning* refers to interpolation toward a regular grid when the number of irregular samples exceeds the regular grid's size.)

In this article, we discuss curvelets and demonstrate their sparseness on seismic data. We also describe our extension to the curvelet transform via the nonequally sampled fast Fourier transform (NFFT).[3] Our extension restores the transform's performance for nonuniformly sampled data.

The Curvelet Transform

Since their introduction, curvelet transforms[2] have received increasing interest in the seismic research community, primarily because curvelets can detect wave fronts. Here, we examine their main properties and investigate curvelets' approximation error on seismic data.

Main Properties

It comes as no surprise that their original construc-

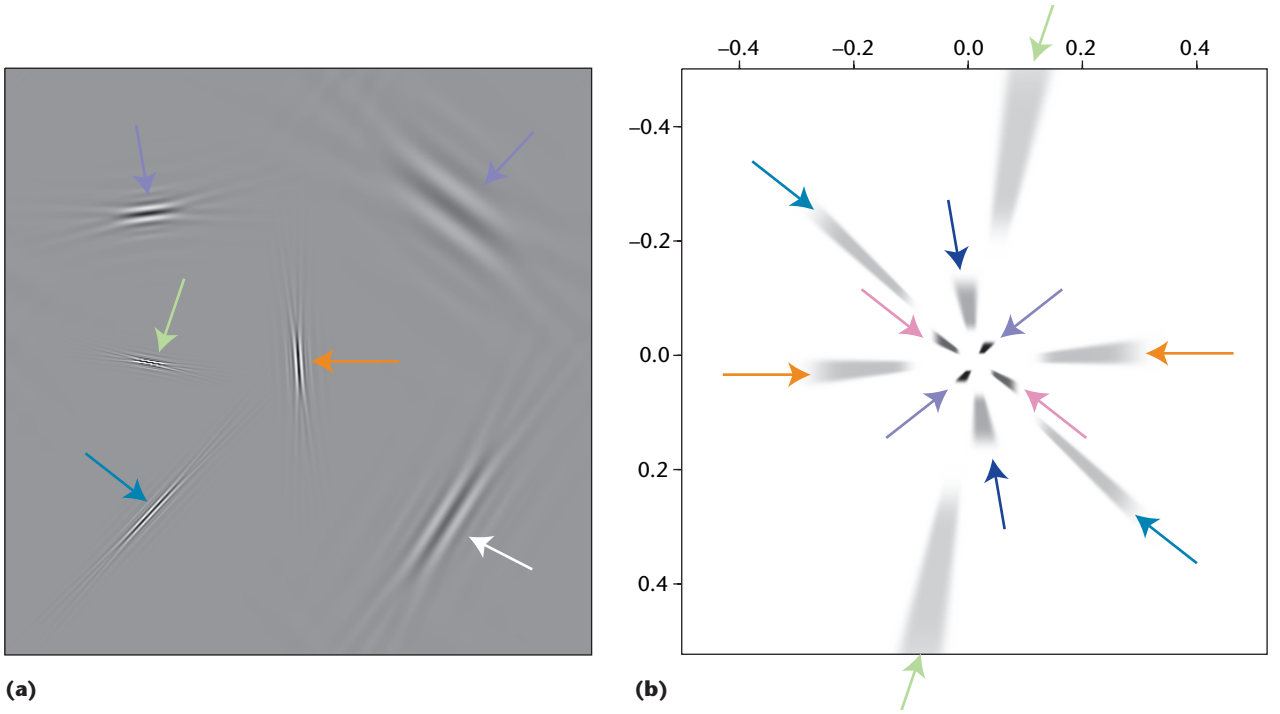

Figure 2. Spatial (left) and frequency (right) viewpoints of six real curvelets at different scales and angles. As opposed to complex curvelets, real curvelets live in two angular wedges symmetric about the origin. Comparing the curvelets in the two domains also shows their microlocal correspondence,[6] relating the orientation of curvelets in both domains. Because of their rapid decay in the physical space and compact support in the Fourier space, curvelets localize in phase space.

tion, through the so-called second dyadic partitioning, came from the field of harmonic analysis,[4] in which they were used as expansions for asymptotic solutions of wave equations. FDCT developers[2] have also recognized this connection, which has resulted in important contributions to the compression of Green's functions[5] as well as to nonlinear approximations of functions with intermittent regularity.[1] These functions are assumed to be piece-wise smooth with singularities (regions in which the derivative diverges) on piece-wise smooth curves. Within the Earth's crust, these singularities correspond to geologic unconformities at which waves reflect; in seismic data, these singularities correspond to wave fronts. Geologic boundaries as well as wave fronts contain points of intermittent regularity, such as faults along sedimentary layers or *caustics* in wave fronts. A caustic is a region where rays concentrate and is typically observed when an incident wave field reflects at curved reflectors.

Our purpose isn't to compress operators—rather, we're interested in separating different seismic data components, which, except for possible incoherent measurement noise, consist of components that are the solution of a wave equation. For this purpose, we employ the curvelet transform as a vehicle that

- is rich enough to account for the multiscale and multidirectional properties of seismic data with intermittent regularity;
- is local in phase space (the space spanned by space and spatial frequency);
- exploits smoothness along (and oscillatory behavior across) arriving wave fronts;
- differentiates among different signal components on the basis of location, angle, and frequency content;
- obtains fast decay of nonlinear approximation errors for seismic data; and
- permits a fast ($\mathcal{O}(K\log K)$ operations, where K is the data size) multidimensional (2D or 3D) implementation.

As we can see in Figures 2 and 3, curvelets are local in both space and spatial frequency; they correspond to a partitioning of the 2D Fourier plane by highly anisotropic elements (for the high frequencies) that obey the paramount parabolic scaling principle:[4] width $\propto$ length^2. As opposed to discrete wavelets, which are designed to provide sparse representations of functions with point singularities, curvelets provide sparse representations for functions with singularities on curves. Moreover,

whereas multiscale wavelets consist of collections of location- and scale-indexed basis functions, curvelets represent a family of functions formed from translations, rotations, and parabolic scalings. As such, curvelets form a frame with moderate redundancy. The elements in this transform, which we call *prototype waveforms*, are

- *multiscale*, with frequency support on dyadic coronae in the 2D Fourier plane;
- *multidirectional*, with angles that correspond to the centers of the wedges (for every other resolution doubling, the number of angles doubles);
- *anisotropic*, obeying the scaling law width ∝ length^2; and
- *local*, allowing for thresholding, which locally adapts to the nonstationary signal.

Frames differ from orthonormal bases. Orthonormal transforms (or matrices) compose an arbitrary finite-energy discretized signal vector $\mathbf{f} \in \mathbb{R}^K$ of length K ($\mathbf{f}$ is a discretization of the multivariate function $f(s, t)$: $\mathbb{R}^2 \rightarrow \mathbb{R}$, with s and t the spatial and temporal coordinates, respectively) according to

$$\mathbf{f} = \mathbf{B}^{-1}\mathbf{B}\mathbf{f} = \mathbf{B}^H\mathbf{B}\mathbf{f} := \sum_{m \in \mathcal{M}} \langle \mathbf{f}, \varphi_m \rangle \varphi_m \ , \qquad (1)$$

with $\mathbf{B}^H$ denoting the matrix adjoint of the decomposition matrix $\mathbf{B}$, and the brackets $\langle \ . \ , \ . \ \rangle$ denoting the standard discrete inner product $\langle \mathbf{f}, \varphi_m \rangle = \mathbf{f}^H \varphi_m$ of $\mathbf{f}$ with the mth column vector of $\mathbf{B}^H$. Because $\mathbf{B}$ is an orthonormal basis, its adjoint matrix corresponds to its inverse (the inverse transform). The summation in Equation 1 runs over the index set $\mathcal{M}$ of size $M = K$. As opposed to an orthonormal transform, a redundant frame expansion decomposes a K-length signal into a frame expansion with $M > K$ elements. Consequently, the composition matrix is rectangular, with the number of columns exceeding the number of rows.

The regularly sampled FDCT is a frame represented by the matrix $\mathbf{C}$. Applying this matrix to a vector $\mathbf{f}$ creates a multi-index coefficient vector $\mathbf{x} = \mathbf{C}\mathbf{f}$ with $\mathbf{x} := \{x_m\}_{m \in \mathcal{M}}$; the multi-index m runs over all the locations, orientations, and scales (details about the FDCT's discrete constructions appear elsewhere[2]). We chose the numerically tight FDCT via wrapping as our curvelet transform. For this transform, the pseudo-inverse (denoted by the symbol †) equals the adjoint, and we have $\mathbf{f} = \mathbf{C}^\dagger\mathbf{x} = \mathbf{C}^H\mathbf{x}$, which implies $\mathbf{C}^H\mathbf{C} = \mathbf{I}$.

Nonlinear Approximation Rates

The nonlinear approximation rate expresses the asymptotic decay of the l_2-difference between the original data and the partial reconstruction from the largest L coefficients. In dimension two, Fourier attains an asymptotic decay rate of only $\mathcal{O}(L^{-1/2})$ for data consisting of twice-differentiable functions with singularities on piece-wise twice-differentiable curves, whereas curvelets asymptotically obtain the optimal rate $\mathcal{O}(L^{-2})$, ignoring log-like factors.[1] Although wavelets improve on Fourier, their approximation rate of $\mathcal{O}(L^{-1})$ is suboptimal.

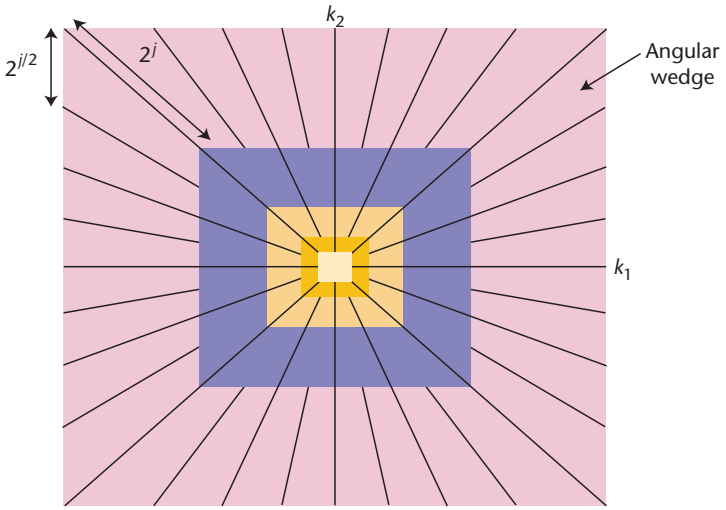

Figure 3. Discrete curvelet partitioning of the 2D Fourier plane into second dyadic coronae and subpartitioning of the coronae into angular wedges. As frequency increases, the wedges become more anisotropic, yielding needle-like curvelets that align better and better with curved wave fronts.

By virtue of their multiscale and multidirectional construction, curvelets sparsely represent seismic data. Not only do individual curvelets capture the main characteristics of wave fronts locally—they look like little waves—but they also jointly capture the seismic energy effectively. We can see this performance in Figure 4, which shows the nonlinear approximation rates for representative seismic synthetic data. We computed the rates for each of the following cases:

- curvelets and wavelets on regularly sampled data,
- curvelets on nonuniformly sampled data (treated as uniformly sampled data), and
- our extension of the curvelet transform on nonuniformly sampled data.

For uniformly sampled data, the nonlinear approximation rate of curvelets outperforms Daubechies 6 wavelets by a wide margin. The figure also shows the importance of treating nonuni-

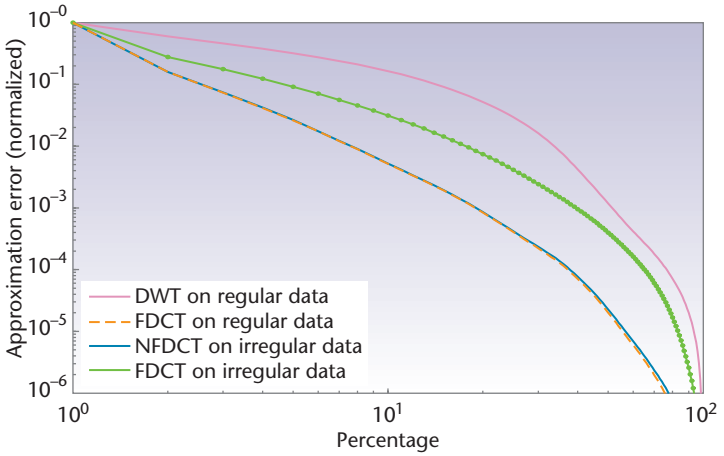

Figure 4. Decays of the nonlinear approximation error. Curvelets on the regular grid (orange dashed line) clearly outperform discrete wavelets (pink line). Our extension of the curvelet transform for nonuniformly sampled data (blue line) retains the performance of the regularly sampled curvelet transform on uniformly sampled data, as opposed to the inferior performance obtained when irregular data is treated as regular (green line with dots).

Table 1. Binning and denoising errors measured by signal-to-noise ratio.* The SNR is 0 dB for the initial (nonuniformly) noisy data.

Operations performed	Signal-to-noise ratio (dB)
Linear binning	–1.96
Nonequally sampled fast Fourier transform (NFFT) binning	9.04
Denoising	13.35
NFFT binning and denoising	8

** SNR is defined as $10\log_{10} \|\mathbf{f}\|_2^2/\|\mathbf{f} - \tilde{\mathbf{f}}\|_2^2$, where $\mathbf{f}$ is the original function and $\tilde{\mathbf{f}}$ its estimate after binning and/or denoising.*

formly sampled data correctly in the curvelet transform—treating it as uniformly sampled data seriously deteriorates performance.

To address the nonuniformly sampled data issue, the geophysicist commonly uses binning to bring nonuniformly sampled data to the regular grid. To compare the reconstructions, we use space-domain linear interpolation for wavelets and include NFFT binning as our extension to the FDCT. Figure 5 shows reconstructions for nonuniformly sampled data with binning for 1 percent of the coefficients. The partial reconstruction with the nonuniformly sampled curvelet transform performs nearly as well as the uniformly sampled transform and outperforms the result obtained with wavelets. Table 1 lists detailed performance measures. Notice that even for this bad signal-to-noise ratio (SNR), we lost only 1 dB between noise-free NFFT binning and noisy NFFT binning combined with denoising.

The NFDCT: A Curvelet Frame for Seismic Processing

As Figure 4 illustrates, treating nonuniformly sampled data as regular data can seriously deteriorate the performance of curvelet approximations and, hence, signal denoising and separation. Because seismic data is usually acquired irregularly, failure to account for nonuniformly sampled data can adversely affect seismic imaging. Our work's main contribution is to extend the FDCT toward nonuniformly sampled grids. The FDCT $\mathbf{C}$ on an arbitrary uniformly sampled vector $\mathbf{f}$ factors as $\mathbf{TF}$, with $\mathbf{F}$ the orthonormal Fourier transform and $\mathbf{T}$ the curvelet tiling matrix (that is, $\mathbf{Cf} := \mathbf{TFf}$). Later, we'll replace the ordinary Fourier transform with its nonuniformly sampled counterpart, which is a natural choice because the curvelet construction is defined in the Fourier domain.

From this point on, we denote nonuniformly sampled N-vectors $\underline{\mathbf{f}} \in \mathbb{R}^N$ with an underbar, and $\underline{\mathbf{f}} := \{f(x_p)\}_{p=1,\ldots,N}$ at the nodes $x_p \in \mathcal{X}$, where $\mathcal{X} := \{x_p = (s_p, t_p) \in \mathbb{R} \times \mathbb{N}: -1/2 \le s_p < 1/2 \text{ and } 0 \le t_p < N_t\}_{p=1,\ldots,N}$, with N representing the total number of nodes and N_t the number of regular time samples. Here, we consider the number of source and receiver positions larger than the size of the corresponding regular spatial grid.

At the heart of nonequally sampled Fourier transforms of bandwidth-limited functions lies the fast evaluation of the following sum:[3,7]

$$\underline{\mathbf{f}} := \{\sum_{k \in K} \hat{f}_k e^{-2\pi i k x_p}\}_{p=1,\ldots,N} \quad (2)$$

This expression corresponds to the discrete inverse Fourier transform from a uniformly sampled grid $\mathcal{K} := \{k_j = (k_j^s, k_j^t) \in \mathbb{Z}^2: -K_{s,t}/2 \le k^{s,t}_j < K_{s,t}/2\}_{j=1,\ldots,K}$ in the Fourier domain—denoted by the hat symbol $\hat{}$ and consisting of $K = K_s \times K_t$ samples with $K_t = N_t$—toward the nonuniformly sampled grid $\mathcal{X}$. In matrix-vector notation, this expression becomes

$$\underline{\mathbf{f}} = \mathbf{A}\hat{\mathbf{f}} . \quad (3)$$

The NFFT is an implementation that approximately evaluates the sum of Equation 2 with a fast algorithm based on a previous study.[7] By replacing the regular FFT in the FDCT's implementation with the NFFT's pseudo-inverse, we arrive at a

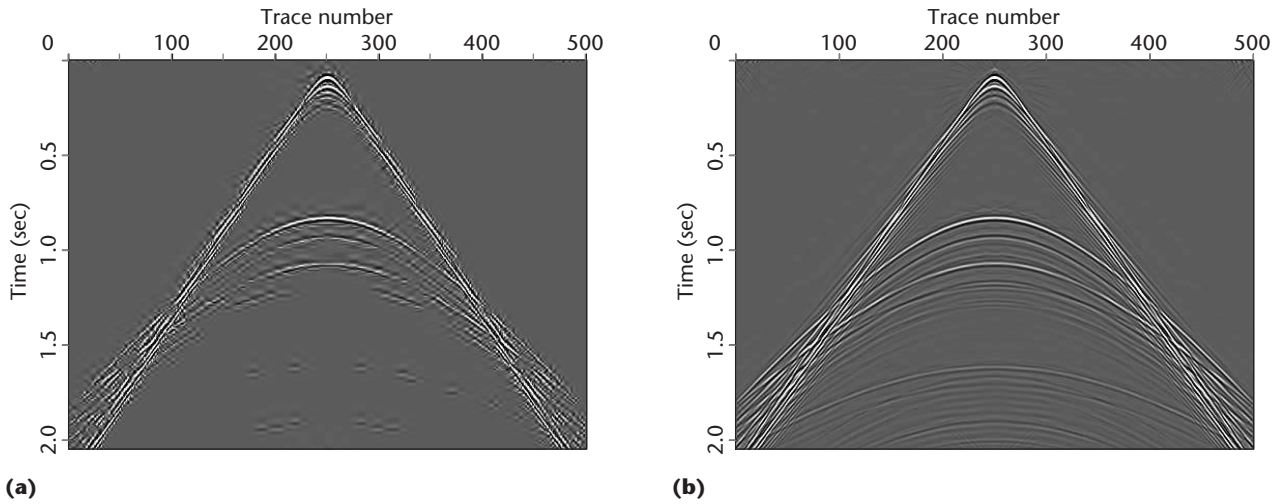

Figure 5. Partial reconstruction using 1 percent of the wavelet and curvelet coefficients for nonuniformly sampled data. In examining (a) linear binning and reconstruction with 1 percent of the wavelet coefficients, and (b) curvelet binning and reconstruction with 1 percent of the curvelet coefficients, we can see a drastic improvement using curvelets rather than wavelets for partial reconstruction.

transform that takes irregularly sampled data to the regularly sampled Fourier domain.

By limiting the maximum distance between the nodes to K_s^{-1} and using more irregular than regular samples ($N > K$), the pseudo-inverse of $\mathbf{A}$ is well conditioned when including an additional diagonal weighting $\mathbf{W}$, which is proportional to the number of sources and receivers per unit on the interval.[3] The NFDCT is thus defined as

$$\mathbf{x} = \mathbf{C}\underline{\mathbf{f}} := \mathbf{T}\mathbf{A}^{\dagger}\underline{\mathbf{f}}, \tag{4}$$

where $\mathbf{A}^{\dagger} := (\mathbf{A}^{H}\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^{H}\mathbf{W}$. Under the previously described irregular sampling conditions, the NFDCT produces curvelet coefficients that pertain to a regular Fourier grid. Hence, applying the regular inverse curvelet transform to these curvelet coefficients yields data on the regular grid. This process corresponds to NFFT-based binning.

Signal Estimation and Separation by Thresholding

Successful denoising and signal separation depend largely on a transform's ability to sparsely represent a particular type of image. Discrete wavelet transforms and curvelets accomplish (near) optimal nonlinear approximation rates for certain classes of images.[8] As we argued earlier, for example, curvelets appear to be the appropriate choice for seismic data. Let's review some estimation techniques for orthonormal wavelets and overcomplete curvelets.

Denoising by Shrinkage

Thresholding on an expansion's coefficients with respect to a collection of prototype waveforms is a key component in solving denoising problems.[8,9] The signal model is

$$\mathbf{d} = \mathbf{m} + \mathbf{n}, \tag{5}$$

where $\mathbf{m}$ is the unknown deterministic signal component and $\mathbf{n}$ is zero-centered white Gaussian noise with standard deviation σ. The Gaussian assumption is fundamental in this work, but whiteness isn't a prerequisite.

Soft thresholding[8,9] on each element of the noisy data coefficient vector solves for the model $\mathbf{m}$ through

$$\tilde{\mathbf{m}} = \mathbf{S}^{\dagger}S_{\mathbf{w}}(\mathbf{S}\mathbf{d}). \tag{6}$$

In this expression, $\mathbf{S}$ stands for an arbitrary sparse signal expansion and $S_{\mathbf{w}}$ for soft thresholding, defined element-wise as

$$S_w(x) := \begin{cases} x - sign(x)w & |x| \geq w \\ 0 & |x| < w \end{cases}, \tag{7}$$

where $w \geq 0$—a real-valued threshold. The vector $\mathbf{w}$ contains the thresholds for each coefficient, and this shrinkage operation via thresholding forms the basis for our denoising and signal separation. Figure 6 illustrates the estimation by shrinkage as de-

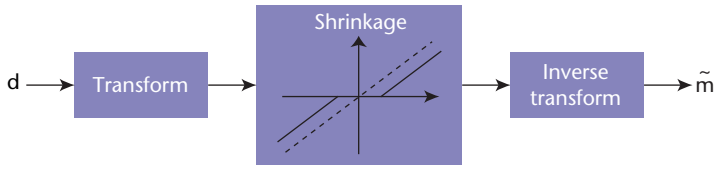

Figure 6. Three-step estimation by shrinkage on transformed domain coefficients. We first bring the noisy data $\mathbf{d}$ to a transformed domain, where we apply soft thresholding on the coefficients. Finally, we obtain a denoised estimate $\tilde{\mathbf{m}}$ by applying the corresponding inverse transform to the thresholded coefficients.

scribed in Equation 6.

Denoising with Orthonormal Bases

For arbitrary orthonormal transforms $\mathbf{S} := \mathbf{B}$, we have $\mathbf{S}^\dagger = \mathbf{B}^{-1}$, and Equation 6 solves the minimization problem

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{x}\|_2^2 + \|\mathbf{x}\|_{1,\mathbf{w}}, \quad (8)$$

where $\{\mathbf{y}, \mathbf{x}\} := \{\mathbf{Bd}, \mathbf{Bm}\}$ are the transformed coefficients, and $\|\mathbf{x}\|_{1,\mathbf{w}}$ is a weighted l_1-penalty functional given by

$$\|\mathbf{x}\|_{1,\mathbf{w}} = \sum_{m \in \mathcal{M}} w_m |x_m|. \quad (9)$$

Given that the noise is zero-centered white Gaussian with standard deviation σ, its maximum amplitude has a high probability of being below $\sigma(2\log_e m)^{1/2}$. This bound is rather conservative.[9] By setting each weight to $w_m = 3\sigma$, Equation 6 yields a satisfactory estimate for $\mathbf{m}$. During this estimation, the quadratic mismatch between the data and model is minimized jointly with the weighted l_1-penalty functional; the quadratic term is known as the *log-likelihood*. We assume the model is a superposition of prototype waveforms, with coefficients drawn independently from a probability function $Pr\{x_m\} \propto \exp(-\text{Const} \cdot w_m |x_m|)$ that corresponds to the Laplace distribution, which enhances sparsity.[10,11]

Denoising with Tight Frames

The FDCT with wrapping is a tight frame with a synthesis matrix $\mathbf{C}^\dagger = \mathbf{C}^H$ that has more columns than rows. The coefficient vector exceeds the data size by a factor of roughly eight. In this case, $\mathbf{CC}^H \neq \mathbf{I}$, and Equation 6 is no longer equivalent to the minimization problem in Equation 8. However, for a tight frame with an l_2-norm for the columns of the synthesis matrix close to unity, shrinkage still provides a good approximation[11] to minimization problem's solution of Equation 8.

Denoising and Binning with the NFDCT

By combining the nonuniformly sampled curvelet transform with shrinkage (see Equation 7), we arrive at our main result:

$$\tilde{\mathbf{m}} = \mathbf{C}^\dagger S_{\mathbf{w}}(\mathbf{C}\underline{\mathbf{d}}), \quad (10)$$

which accomplishes the joint task of binning and (in-) coherent signal separation on nonuniformly sampled data $\underline{\mathbf{d}}$. In this expression, the nonuniform data vector $\underline{\mathbf{d}}$ is curvelet-transformed with the NFDCT, followed by a thresholding and the regular inverse curvelet transform (FDCT). Under the assumptions we stated earlier about the signal's bandwidth limitation and unequal sampling, we can stably compute the pseudo-inverse used to Fourier-transform the unequally sampled points. As such, we can safely assume that the regular sampled Fourier data is still close to the Fourier transform of the corresponding uniformly sampled data. We proceed as if we were dealing with the uniformly sampled case, by thresholding and applying the uniformly sampled inverse curvelet transform (IFDCT). The result of this operation is a combined denoising and binning, in which irregular bandwidth-limited noisy data is denoised and mapped to a regular grid. Figures 7 and 8 demonstrate this technique.

Coherent Signal Separation

Although thresholding estimators are used primary to separate incoherent noise from deterministic signal components, our work extends thresholding estimators to the separation of coherent signal components (for example, primary-multiple separation in seismic exploration[12]).

In this case, the signal model becomes slightly more complicated:

$$\mathbf{s} = \mathbf{s}_1 + \mathbf{s}_2 + \mathbf{n}, \quad (11)$$

where $\mathbf{s}_1, \mathbf{s}_2$ are the two coherent signal components. Given a prediction $\breve{\mathbf{s}}_2$ for the second component, we can estimate the first component via Equation 6, where the weighting is defined as

$$\mathbf{w} := \max(3\sigma, \delta |\breve{\mathbf{x}}_2|), \quad (12)$$

with $\breve{\mathbf{x}}_2 := \mathbf{C}\breve{\mathbf{s}}_2$. This weighting corresponds to a varying threshold defined in terms of the curvelet transform for the predicted signal component. The

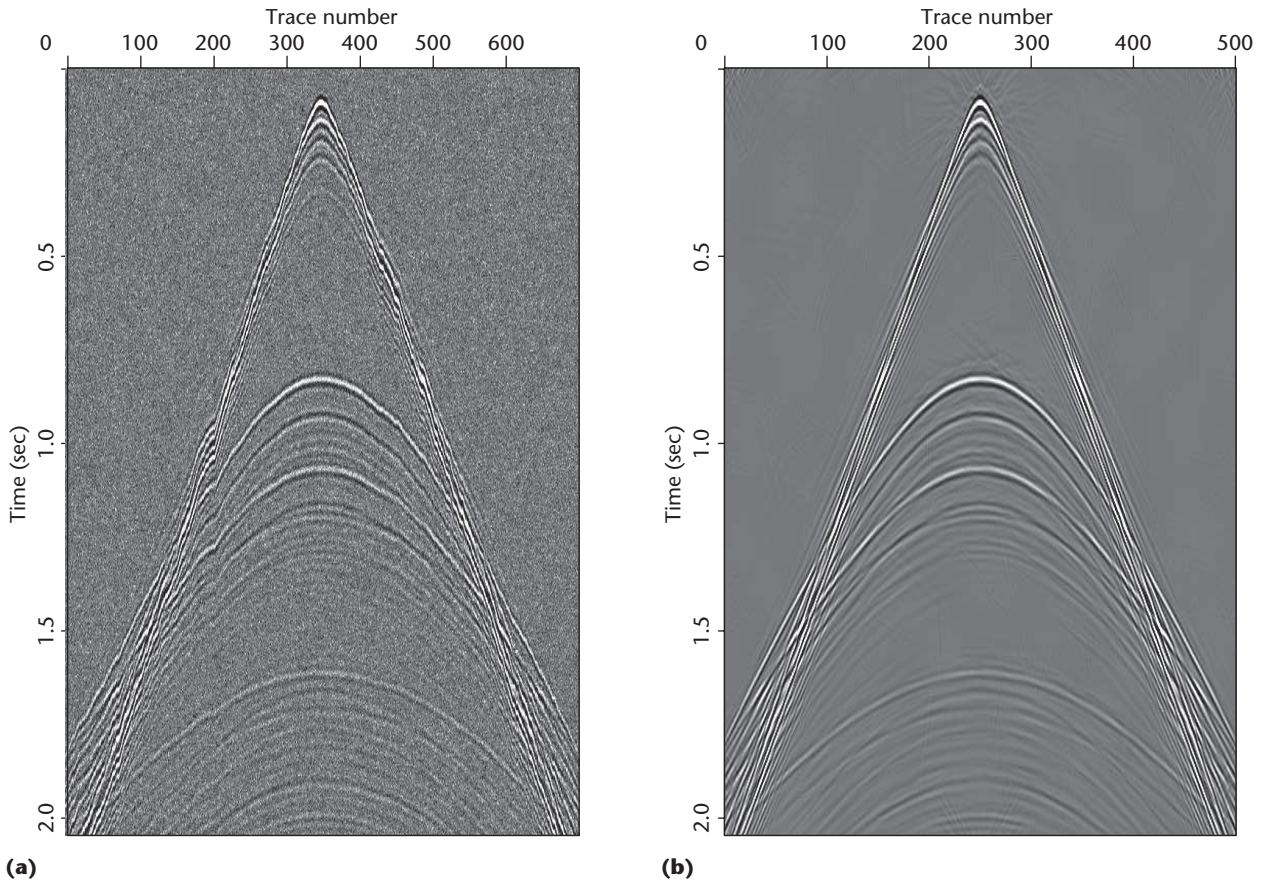

Figure 7. Incoherent noise removal through shrinkage (Equations 6 and 10). We plotted (a) noisy nonuniformly sampled data in a regular grid and with a signal-to-noise ratio (SNR) of 0 dB, and (b) binned and denoised data (see Equation 10). Notice the significant improvement over the SNR listed in Table 1.

δ expresses confidence in the prediction. The estimator of Equation 6 using weights as defined in Equation 12 again corresponds to a maximum a posteriori (MAP) estimator that minimizes the log-likelihood function with coefficients selected from a cross-correlation-weighted probability function $Pr\{x_m\} \propto \exp(-\text{Const} \cdot w_m |x_m|)$ for $m \in \mathcal{M}$. This function is weighted by the prediction for the second signal component.

Applications to Seismic Data

Among seismic data's most striking features is that it contains wave fronts possibly contaminated with bandwidth-limited Gaussian noise. As we've shown, we can remove this random component by forward transforming the (ir)regular data with the (N)FDCT, followed by a simple shrinkage and reconstruction with the IFDCT. Figure 7 illustrates the performance of curvelet denoising by shrinkage. Because denoising's performance is (aside from the binning error) as good as regular denoising, we show the results only for nonuniformly sampled data in which denoising is combined with binning. We can extend these methods to the case of coherent signal removal according to the threshold defined in Equation 12. To emphasize the NFDCT's added value, we include an example in which signal separation is performed on irregular data cast into a regular grid and on the irregular data itself with the NFDCT.

The removal of ghost events related to multiple interactions of the wave field with the surface is paramount to the success of seismic imaging based on linearized inverse scattering. These ghosts—also known as multiples—violate the linearization and create artifacts in the image. Removing these artifacts has proven to be difficult due to position and amplitude errors in the multiple prediction. Researchers have developed adaptive subtraction techniques based on matched filtering[13] to counter these prediction errors and robustly separate the two signal components. Unfortunately, matched

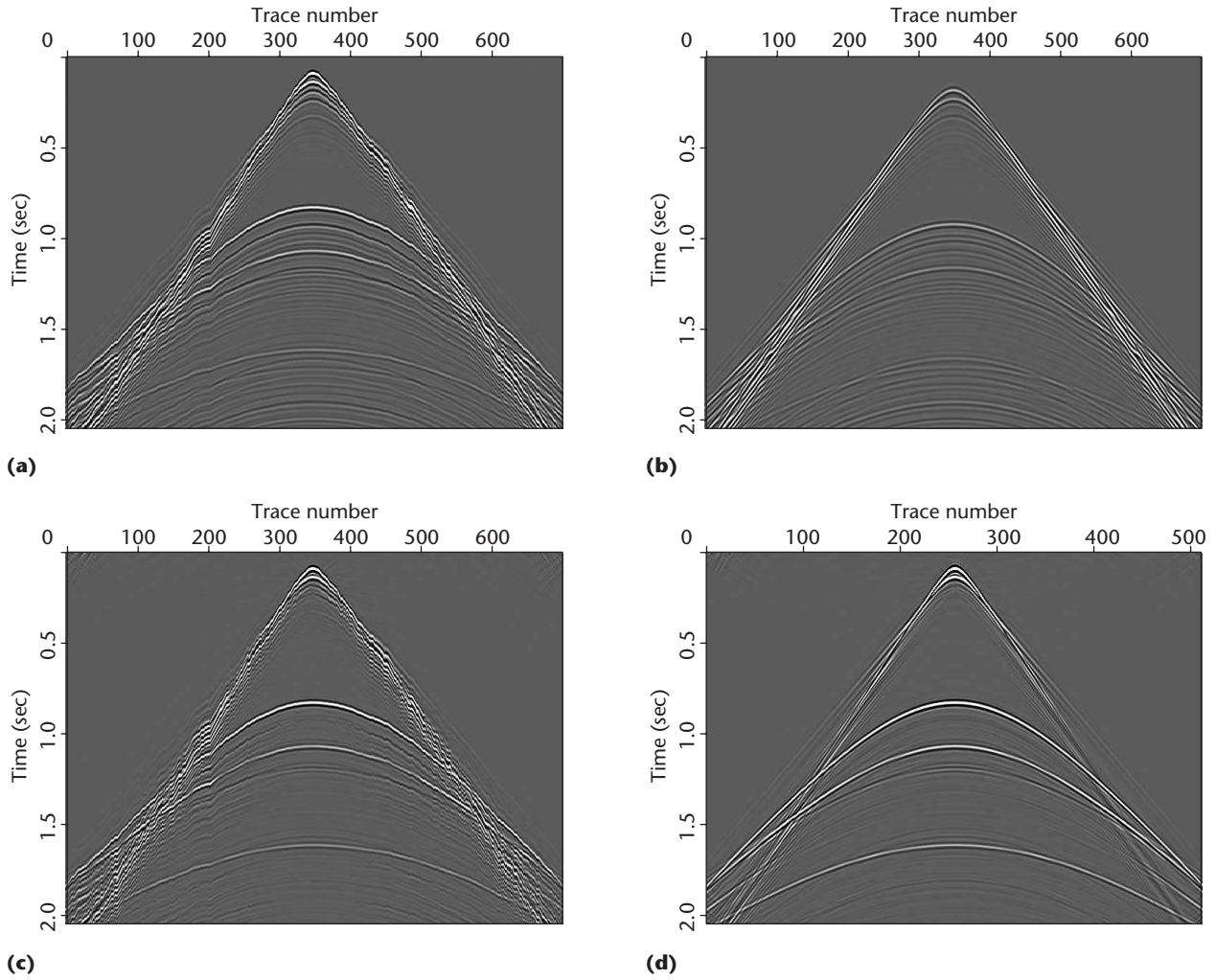

Figure 8. Removal of ghost events related to multiple interactions of the wave field with the surface. We can compare (a) synthetic nonuniformly sampled data that contains primary and multiple reflections treated as regular data, (b) predicted multiples, (c) estimated primaries using the fast discrete curvelet transform (FDCT) on (a) and weights as defined in Equation 12, and (d) estimated primaries using the nonuniformly sampled FDCT (NFDCT) on (a) and weights as defined in Equation 12.

filtering can inadvertently remove primary energy and leave an unwanted remainder of multiple energy. As a consequence, primaries can accidentally be deteriorated and ghosts may locally remain. As Figure 8 shows, we've seen some good results from formulating this signal-separation problem as a weighted shrinkage in the curvelet domain. We obtained these results by using $\tilde{\mathbf{s}}_1 = \mathbf{C}^{\dagger} S_{\mathbf{w}}(\mathbf{C}\underline{\mathbf{s}})$, where the weights $\mathbf{w}$ are defined as in Equation 12 and $\tilde{\mathbf{s}}_2$ represents predicted multiples. We set the constants to $\delta = 1.6$ and σ according to the noise level; we left the predicted multiples as they were. By virtue of the NFDCT, the result for the nonuniformly sampled case is almost as good as the result for the uniformly sampled case.

Our extension makes curvelets applicable in other fields where nonuniform sampling is a concern (for example, charge-coupled device cameras and medical imaging). In a future article, we'll report on an extension of our method to the case in which the interpolation grid's size exceeds the number of unequally sampled data points and on robust sampling criteria for seismic data. This research aims at answering two main questions:

- What's the performance of seismic data recovery given a certain acquisition geometry?
- What's the "optimal" acquisition geometry for recovering seismic data to within a desired accuracy?

We're also working on other curvelet-based techniques for seismic data separation and seismic image amplitude recovery. As these applications indicate, curvelets will change seismic data processing and imaging. Promoting sparsity in the curvelet domain as part of the seismic workflow will increase image quality, allowing us to obtain more details for deeper targets.

Acknowledgments

We thank the authors of the Fast Discrete Cosine Transformation (FDCT)[2] *and the nonequally sampled fast Fourier transform (NFFT),*[3] *as well as Colin Russell for his coding. We performed this work as part of the Seismic Imaging by Next-Generation Basis Functions Decomposition (SINBAD) project, with financial support through the Industry Technology Facilitator (ITF) from the following organizations: BG Group, British Petroleum, Chevron, ExxonMobil, and Shell Oil. Additional funding came from the Natural Sciences and Engineering Research Council (NSERC) of Canada Discovery Grant 22R81254 and from the Pseudodifferential Operator Theory and Seismic Imaging (POTSI) project, funded through the Mathematics of Information Technology and Complex Systems (MITACS).*

References

1. J. Candes and D.L. Donoho, "New Tight Frames of Curvelets and Optimal Representations of Objects with C2 Singularities," *Comm. Pure and Applied Mathematics*, vol. 57, no. 2, 2004, pp. 219–266.
2. E.J. Candes et al., *Fast Discrete Curvelet Transforms*, tech. report, Calif. Inst. of Tech., 2005; www.curvelet.org.
3. S. Kunis and D. Potts, *NFFT 2.0*, tech. report, Univ. of Lubeck, 2005; www.math.uni-luebeck.de/potts/nfft/.
4. H.F. Smith, "A Hardy Space for Fourier Integral Operators," *J. Geometric Analysis*, vol. 8, no. 4, 1998, pp. 629–654.
5. E.J. Candes and L. Demanet, "Curvelets and Fourier Integral Operators," *Royal Academy of Science, Paris*, vol. 336, 2003, pp. 395–398.
6. E.J. Candes and D.L. Donoho, "Recovering Edges in Ill-Posed Problems: Optimality of Curvelet Frames," *Annals of Statistics*, vol. 30, no. 3, 2002, pp. 784–842.
7. G. Beylkin, "On the Fast Fourier Transforms of Functions with Singularities," *Applied and Computational Harmonic Analysis*, vol. 2, no. 4, 1995, pp. 363–381.
8. D.L. Donoho and I.M. Johnstone, "Minimax Estimation via Wavelet Shrinkage," *Annals of Statistics*, vol. 26, no. 3, 1998, pp. 879–921.
9. S.G. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1997.
10. J.L. Starck, M. Elad, and D.L. Donoho, "Redundant Multiscale Transforms and Their Application for Morphological Component Analysis," *Advances in Imaging and Electron Physics*, vol. 132, 2004, pp. 288–348.
11. M. Elad, "Why Simple Shrinkage Is Still Relevant for Redundant Representations?" *IEEE Trans. Information Theory*, submitted 2005.
12. F.J. Herrmann, U. Boeniger, and D.J. Verschuur, "Non-Linear Primary-Multiple Separation with Directional Curvelet Frames," submitted to *Geophysical J. Int'l*, 2006.
13. D.J. Verschuur, A.J. Berkhout, and C.P.A. Wapenaar, "Adaptive Surface-Related Multiple Elimination," *Geophysics*, vol. 57, no. 9, 1992, pp. 1166–1177.

Gilles Hennenfent *is a PhD student in the Seismic Laboratory for Imaging and Modeling at the University of British Columbia. His research interests include fast approximate algorithms and multiscale methods applied to stable seismic signal recovery. Hennenfent has an engineering diploma in applied physics from the Ecole Nationale Supérieure de Physique de Strasbourg, France, and an MSc in photonics, image, and cybernetics from the Louis Pasteur University, France. He is a student member of the European Association of Geoscientists and Engineers (EAGE) and the Society of Exploration Geophysicists (SEG). Contact him at ghennenfent@eos.ubc.ca.*

Felix J. Herrmann *is an assistant professor in the Department of Earth and Ocean Sciences at the University of British Columbia, where he also heads the Seismic Laboratory for Imaging and Modeling. His research is directed toward creating a fundamental understanding of seismic imaging and inversion, as well as establishing a direct link between local aspects of seismic reflectivity and major events in the geological and rock-physical processes responsible for rapid changes in the Earth's elastic properties. Herrmann has a PhD in engineering physics from the Delft University of Technology, the Netherlands. He is member of EAGE, SEG, the Society for Industrial and Applied Mathematics (SIAM), and the American Geophysical Union (AGU). Contact him at fherrmann@eos.ubc.ca.*