GEOPHYSICS, VOL. 72, NO. 6 (NOVEMBER-DECEMBER 2007); P. S231–S248, 14 FIGS.
10.1190/1.2785047

Leading-order seismic imaging using curvelets

Huub Douma[1] and Maarten V. de Hoop[2]

ABSTRACT

Curvelets are plausible candidates for simultaneous compression of seismic data, their images, and the imaging operator itself. We show that with curvelets, the leading-order approximation (in angular frequency, horizontal wavenumber, and migrated location) to common-offset (CO) Kirchhoff depth migration becomes a simple transformation of coordinates of curvelets in the data, combined with amplitude scaling. This transformation is calculated using map migration, which employs the local slopes from the curvelet decomposition of the data. Because the data can be compressed using curvelets, the transformation needs to be calculated for relatively few curvelets only. Numerical examples for homogeneous media show that using the leading-order approximation only provides a good approximation to CO migration for moderate propagation times. As the traveltime increases and rays diverge beyond the spatial support of a curvelet; however, the leading-order approximation is no longer accurate enough. This shows the need for correction beyond leading order, even for homogeneous media.

INTRODUCTION

In high-frequency approximation, seismic waves propagate along rays in the subsurface. The local slopes of reflections in seismic data measured at the surface determine (together with the velocity of the medium at the surface) the directions in which we need to look into the earth from the surface to find the location and orientation of the reflector in the subsurface where the reflection occurred. We generally refer to the determination of a reflector position and orientation from the location of a reflection in the data and the local slope as *map migration* (Kleyn, 1977). For an overview of the literature on this topic we refer to Douma and de Hoop (2006b).

Given the slopes at the source and at the receiver locations, map migration provides one-to-one mapping from the surface seismic measurements, i.e., locations, times, and slopes, to the reflector position and orientation in the image (provided the medium does not allow different reflectors to have identical surface seismic measurements that persist in being identical under small perturbations of the velocity model). This is in contrast to migration techniques that do not make explicit use of the slopes in the data, such as Kirchhoff methods, where the data are summed over diffraction surfaces (see, e.g., Bleistein et al., 2000) so that all points along the diffraction surface are mapped to one diffractor location. The benefit of the explicit use of the local slopes in the data is exploited in several seismic applications such as parsimonious migration (Hua and McMechan, 2001, 2003) controlled directional reception (CDR) (Zavalishin, 1981; Harlan and Burridge, 1983; Sword, 1987; Riabinkin, 1991), and stereotomography (Billette and Lambaré, 1998; Billette et al., 2003). This list is not complete, and more applications exist. All of these methods estimate the slopes from the data using additional processing techniques such as local slant stacking, multidimensional prediction-error filters (Claerbout, 1992, p. 186–201), or plane-wave destruction filters (Fomel, 2002; Claerbout, 1992, p. 93–97).

Curvelets were recently introduced in the field of applied harmonic analysis (Candès and Donoho, 2000; Candès and Guo, 2002; Candès and Donoho, 2004). They allow a sparse representation of objects that are smooth (i.e., twice continuously differentiable) away from discontinuities along smooth edges. Because of the wave character of seismic data, the reflections recorded in seismic data lie predominantly along smooth surfaces (or curves in 2D), just as geologic interfaces in the subsurface lie primarily along smooth surfaces. Therefore, it is plausible that seismic data and their images can be sparsely represented using curvelets. This conjecture was made earlier by Herrmann (2003a, b). We mention that, throughout this paper, we freely interchange sparse representation and compression, because the details of their precise definitions in the mathematical literature are beyond the scope of this work.

Because curvelets have associated local directions, using curvelets as building blocks of seismic data, the local slopes in the data are

Manuscript received by the Editor 26 February 2007; revised manuscript received 1 August 2007; published online 31 October 2007.
[1]Center for Wave Phenomena, Department of Geophysics, Golden, Colorado; presently Princeton University, Department of Geosciences, Princeton, New Jersey. E-mail: hdouma@princeton.edu.
[2]Center for Wave Phenomena, Department of Geophysics, Golden, Colorado; presently Purdue University, Department of Earth and Atmospheric Sciences, West Lafayette, Indiana. E-mail: mdehoop@math.purdue.edu.

built into the representation of the data. Smith (1998b) and later Candès and Demanet (2005) have shown that curvelets sparsify a class of Fourier integral operators from which the seismic imaging operator we treat in this work can be constructed. Realizing that reflections in seismic data lie mainly along smooth curves, it then seems that curvelets are plausible candidates for simultaneous compression of seismic data, their images, and the imaging operator itself. For examples of sparse representations of seismic data and seismic images in the curvelet domain and applications that make use thereof, we refer the reader to Candès et al. (2006), Yarham et al. (2006), Hennenfent and Herrmann (2007), Herrmann et al. (2007), and references therein. For examples of the sparse representation of the normal operator and the demigration/migration operator, we refer to Moghaddam et al. (2007) and Chauris and Nguyen (2007), respectively. The latter work applies the theory that was introduced by Douma and de Hoop (2006a) for common-offset (CO) Kirchhoff time migration, which is here extended to the case of CO Kirchhoff depth migration, except that they apply it to both the demigration and migration operator to obtain the imprint of a local velocity perturbation on a curvelet in the image.

In this work, we aim to investigate the basic characteristics of seismic imaging with curvelets. We show that using curvelets as building blocks of seismic data, the Kirchhoff diffraction stack can, to leading order in angular frequency, horizontal wavenumber, and migrated location, be rewritten as a transformation of coordinates of the curvelets in the data, combined with an amplitude correction. This transformation is calculated using map migration, which uses the local slopes provided by the curvelet decomposition of the data. Our derivation is valid for heterogeneous, smoothly varying media. We verify our method with numerical examples for homogeneous media only and show that, even for such simple media, using the leading-order approximation is accurate only when the rays associated with the center of each curvelet do not diverge beyond the spatial support of a curvelet. If the rays do diverge beyond this support, the leading-order approximation is accurate locally near the map-migrated center of a curvelet, but loses accuracy away from it.

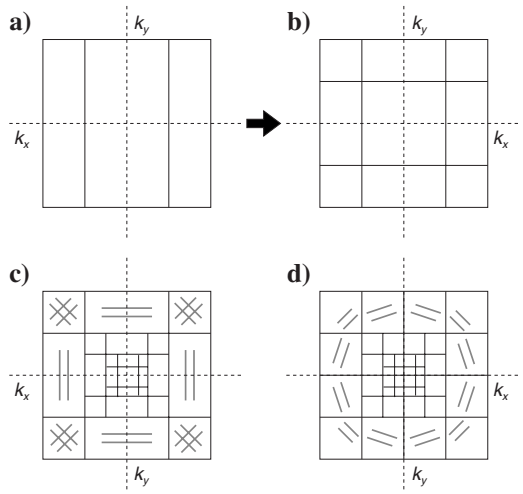

Figure 1. Tiling of the spatial-frequency domain by subband filtering of the horizontal wavenumber axis (a) followed by the vertical wavenumber axis (b). Resulting tiling after three rounds of subband filtering (c) with the dominant directions associated with each tile indicated by the line elements. Tiling of the spatial-frequency domain resulting from the complex wavelet transform (d).

The outline of this paper is as follows. First, we explain what curvelets are, how they are constructed, and what their main properties are. Subsequently, we present examples of digital curvelets from the digital curvelet transform (Candès et al., 2006). Then, we explain the relationship between curvelets and map migration, and proceed to derive our main result that the classical CO Kirchhoff diffraction-stack approach to depth migration can, with respect to leading order, be rewritten as a simple transformation of the input coordinates of curvelets in the data, combined with an amplitude correction. This part extends the previously presented leading-order contribution to CO Kirchhoff migration in constant media by Douma and de Hoop (2006a) to the case of heterogeneous media. Subsequently, we explain how the transformation of coordinates can be calculated from prestack map migration. Finally, we verify the method using numerical examples for homogeneous media and finish with a discussion of the results.

CURVELETS

In this section, we review the main properties of curvelets and how they are constructed. This section has a tutorial character to help the reader unfamiliar with curvelets to understand the material that follows. There are, however, various papers in the field of applied harmonic analysis that provide plenty of details about curvelets and their construction. We refer the reader to the original treatment of Candès and Donoho (2004) for details regarding the continuous transform, while the details of the digital implementation are described by Candès et al. (2006). A short summary of the generally complex-valued curvelets is given by Candès and Demanet (2005). A construction equivalent to that of curvelets was earlier introduced by Smith (1998b), while Do and Vetterli (2003) introduced contourlets that are similar to curvelets in the employed tiling of the frequency plane. Finally, the construction of shearlets (Guo and Labate, 2007) is also closely related to the digital implementation of the curvelet transform.

In 1D, wavelets (Mallat, 1998) are localized in both the independent variable and its Fourier dual, say time and frequency. Their location on the time and frequency axis is uniquely determined by a translation index m and a scale index j, respectively. Wavelets can be extended to higher dimensions by applying the wavelet transform to each dimension successively, such as the rows and columns of a 2D image. The directions associated with multidimensional wavelets obtained in this fashion, are, however, poorly distributed. This is a consequence of the mixing of positive frequencies related to one variable with negative frequencies related to another (Figure 1a–c). To resolve this, complex wavelet transforms were introduced that treat positive and negative frequencies separately (e.g., Kingsbury, 1999, 2001). In this way, the directional selectivity of the transform is improved; instead of three wavelets per subband (with only horizontal, vertical, or a mix of both diagonal orientations; Figure 1c), each subband is subdivided into six wavelets centered around angles $\pm 15°$, $\pm 45°$, and $\pm 75°$ (Figure 1d). Multidimensional extensions to wavelets are characterized by the fact that the width and length of their support in the frequency plane are equal. That is, their support is essentially a square box (see Figure 1c or d for the wavenumber support in 2D). This is referred to as *isotropic* scaling. This isotropic scaling implies that in the limit of infinite frequency, multidimensional wavelets become point-like.

Curvelets are in essence *anisotropic* extensions to wavelets that have better directional selectivity than wavelets (Figure 2a). Instead

of six main directions per subband as for the complex wavelet transform, the number of directions increases as the scale index becomes larger, because curvelets satisfy an anisotropic scaling relation. In the wavenumber domain, the length and width of the support of a curvelet (see Figure 2a) satisfy the parabolic scaling relation

$$\text{length} \propto \text{width}^2 \tag{1}$$

Because of this anisotropic scaling, curvelets become line-like (sometimes referred to as little needles) in the limit of infinite frequency. Note that in the spatial domain, the parabolic scaling relation is width $\propto$ length^2.

The behavior of wavelets and curvelets in the limit of infinite frequency explains, at least intuitively, why wavelets and curvelets allow a sparse representation of point singularities and singularities along smooth curves, respectively. That curvelets allow a sparse (or actually, the sparsest) representation of singularities along smooth curves was shown by Candès and Donoho (2004), who mention the special role of the parabolic scaling. As it turns out, the parabolic scaling is also special with regard to wave propagation through smooth media (Smith, 1998b). For smooth media, Candès and Demanet (2005) demonstrated that under the action of the wave operator, curvelets remain optimally localized in both the spatial and wavenumber domains. This property remains valid for CO Kirchhoff depth migration that we treat in this work.

Curvelets are constructed through the following sequence of operations. First, the spectral domain is band-pass filtered in the radial direction into dyadic annuli (or subbands); this means that the radial widths of two neighboring annuli differ by a factor of 2, the next outer annulus having twice the radial width of the inner one (Figure 2a). Each subband is indexed by a scale index j. Subsequently, each subband is subdivided into angular wedges (or pairs of wedges for real-valued curvelets), and the location of each wedge within a subband is identified by a rotational index l (Figure 2a). The number of wedges in each subband is determined by the frequency content (or the scale index j) of the subband and is proportional to $2^{\lceil j/2 \rceil}$, where the notation $\lceil p \rceil$ denotes the upper integer part of p. This means that the number of wedges in a subband increases only every other scale. This way, the parabolic scaling relation is satisfied. Note that increasing the number of wedges every other scale only is a direct consequence of the dyadic nature of the subband filtering done in the first step, combined with the parabolic scaling. Because the radial length of the subband filters increases with a factor of two at every scale, the number of wedges would have to increase with a factor of $\sqrt{2}$ at every scale to satisfy the parabolic scaling. This would leave a noninteger number of wedges. Hence, the doubling of the number of angular wedges occurs every other scale only. The subband filtering gives curvelets their band-limited nature (just as with wavelets), whereas the subdivision of these subbands into angular wedges provides them with orientation. At the coarsest scale, i.e., the innermost concentric circle in Figure 2a, no angular subdivision is done.

To facilitate translation of curvelets in the spatial domain, each wedge in the wavenumber domain is multiplied by a 2D orthonormal Fourier basis for the rectangle that just covers the support that contains the wedge (indicated by the dashed lines in Figure 2a). According to the discrete Fourier transform, this basis has the fewest members if the area of this rectangle is minimal, as a result the product of both sampling intervals in space is largest. Therefore, the orientation of this rectangle rotates with the angular wedge, and the spatial tiling associated with the local Fourier basis is oriented along the central direction of the angular wedge (Figure 2b); that is, the spatial tiling associated with each angular wedge depends on the particular orientation of the wedge. The translation is indexed by m_1 and m_2 (in two dimensions).

Curvelets are therefore identified by all four indices (j,l,m_1,m_2): j determines their frequency content, l determines their main orientation, while m_1 and m_2 determine their location on the associated spatial grid. The relation between these indices and the location of the curvelet in the spatial and spectral domains is shown in Figure 2a and b.

Roughly speaking, we can think of curvelets as small pieces of band-limited plane waves. The difference between this rough description and the actual interpretation lies, of course, in the fact that a band-limited plane wave has associated with it one $\mathbf{k}$ direction only, whereas a curvelet is associated with a small range of $\mathbf{k}$ vectors. A better description is the term *coherent wave packet*, which was around before curvelets were ever constructed (e.g., Smith, 1998a, b) and dates back to the work of Córdoba and Fefferman (1978). The frequency-domain tiling of curvelets is the same as the dyadic parabolic decomposition or second dyadic decomposition (Gunther Uhlmann, personal communication, 2003) used in the study of Fourier integral operators (Stein, 1993), which was around long before curvelets were introduced (Fefferman, 1973).

Curvelets form a *tight frame* for functions in $L^2(\mathbb{R}^2)$. This means that, much as with a basis, we have a reconstruction formula,

$$f = \sum_{\mu \in \mathcal{M}} \langle f, c_\mu \rangle c_\mu, \quad \langle f, c_\mu \rangle = \int_{\mathbb{R}^2} f(\mathbf{x}) c_\mu^*(\mathbf{x}) \mathrm{d}\mathbf{x}, \tag{2}$$

where c_μ denotes a curvelet with multi-index $\mu = (j,l,m_1,m_2)$, the superscript $*$ denotes complex conjugation, $\mathcal{M}$ is a multi-index set, $\mathbf{x}$ is defined as $\mathbf{x} = (x_1\ x_2)^T$, and $f(\mathbf{x}) \in L^2(\mathbb{R}^2)$. Thus, we can express an arbitrary function in $L^2(\mathbb{R}^2)$ as a superposition of curvelets. In the context of this work, it suffices to know that the main difference between a basis and a tight frame is that the frame elements do not need to be linearly independent. This means that a tight frame is overcomplete and that when curvelets are used as building blocks of seismic data, the data can be reconstructed exactly, but with more building blocks than there are samples in the data. The quantity $\langle f, c_\mu \rangle$ is the coefficient of curvelet c_μ, which denotes the projection of the function f on curvelet c_μ. For a more in-depth treatment of frames, we refer the reader to, e.g., Hernández and Weiss (1996). In this reference

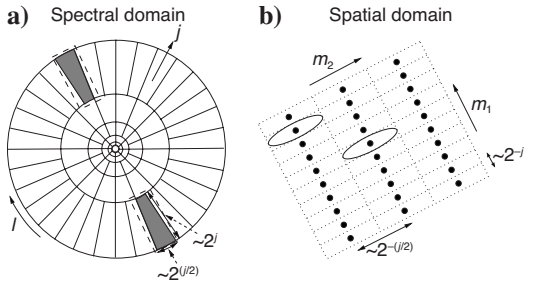

Figure 2. Tilings of the curvelet frame in the spectral domain (a) and the spatial domain (b). In the frequency domain, a curvelet is supported near a wedge on a polar grid (i.e., the actual support extends slightly further than the indicated wedge), where the width of the wedge is proportional to $2^{\lceil j/2 \rceil}$ and its length to 2^j. A local Fourier basis on the support of a wedge provides a Cartesian tiling of the spatial domain (shown schematically in b). The essential support of a curvelet in the spatial domain is indicated by an ellipse (while again the actual support extends beyond this ellipse).

(p. 334–336), it is also explained that the reconstruction formula (equation 2) is equivalent to requiring that Plancherel's formula holds, i.e.,

$$\|f\|^2 = \sum_{\mu \in \mathcal{M}} |\langle f, c_\mu \rangle|^2, \quad \forall f \in L^2. \tag{3}$$

The energy of a signal is conserved through decomposition (equation 2). Considering that curvelets are constructed by windowing the spectral domain, it therefore suffices if the energy of the employed windows sums to one for the resulting curvelets to form a tight frame. Because there are many different windows that satisfy this constraint, many variants of curvelets can be constructed using different windows in the angular or radial direction of the spectral domain (Chauris, 2006). The decay properties of the employed windows in the Fourier domain, however, determine the redundancy of the frame, while their smoothness is closely related to the decay properties of curvelets in the spatial domain, much as with wavelets (Mallat, 1998).

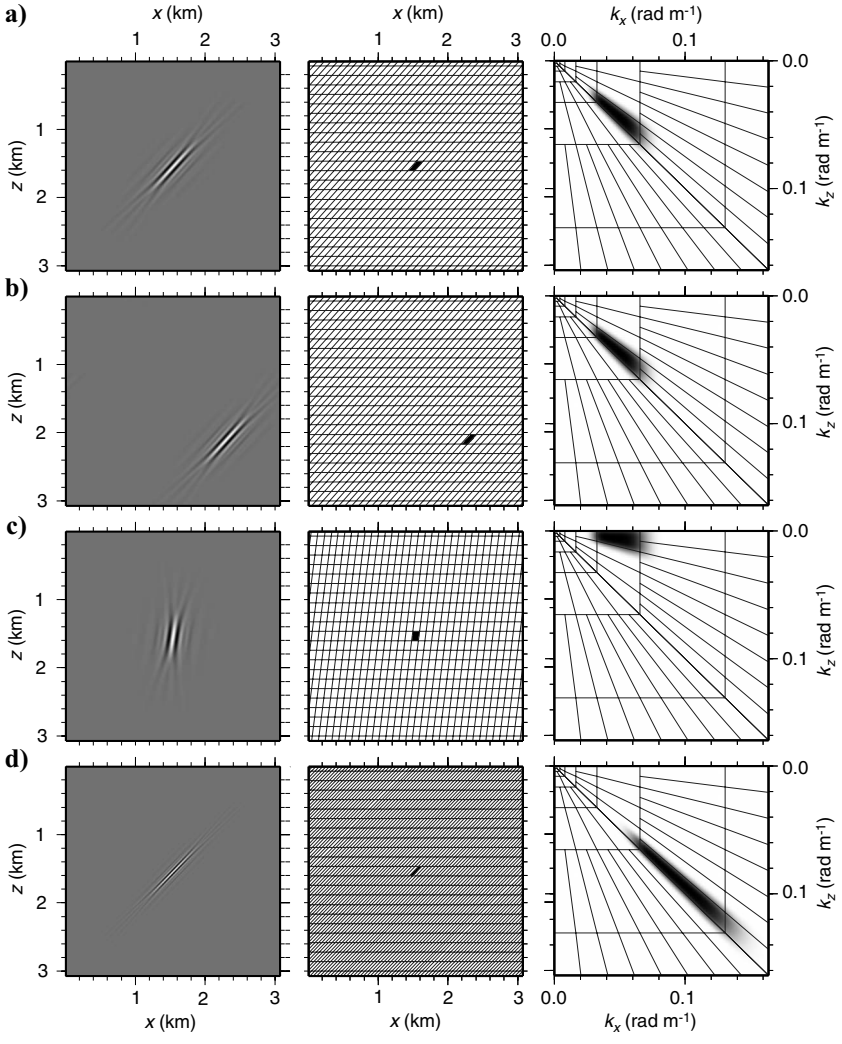

Figure 3. A curvelet (a). Figure b–d shows, respectively, the curvelet when changing the translation indices, the angular index, and the scale index. First column: curvelets in the spatial domain. Second column: associated spatial lattices, and spatial cells colored according to the value of the coefficient (black is one, white is zero). Third column: amplitude spectra and frequency-domain tilings.

Digital curvelets

In the construction of curvelets treated so far, the sampling of the spectral domain is done in polar coordinates, while the sampling of the spatial domain is Cartesian (Figure 2a and b). For the purpose of digital curvelet transforms, the polar coordinates in the spectral domain are replaced with Cartesian coordinates. This allows straightforward application of fast Fourier transform algorithms. To go from polar coordinates to Cartesian coordinates in the spectral domain, the concentric circles in Figure 2a are replaced with concentric squares (Figure 3, right column); hence, the rotational symmetry is replaced with a sheared symmetry. As a consequence, the Cartesian sampling in the spatial domain is a Cartesian grid that is sheared rather than rotated (cf. Figures 2b and the middle column of Figure 3). Here, the centers of the cells are the locations of the centers of the curvelets in space. This construction allows a rapidly computable, digital curvelet transform (Candès et al., 2006). For more details on the implementation of digital curvelet transforms, we refer to Candès et al. (2006). Throughout this work, we use the nonequispaced, fast-Fourier-transform-based curvelet transform that induces a sheared spatial grid, which changes as a function of angular and scale indices l and j, as opposed to the wrapping-based transform, which induces a rectangular grid that is independent of the angular index l and depends only on the scale index j. The wrapping-based transform, moreover, has the advantage that the inverse curvelet transform equals the adjoint curvelet transform, allowing fast computation of the inverse transform. It achieves this, however, at the cost of increasing the number of frame elements and thus the redundancy of the transform (Candès et al., 2006).

With the aim of migration with curvelets in mind and from the intuition that such migration should contain at least a rotation of curvelets (Douma and de Hoop, 2005), Chauris (2006) develops a digital transform that maintains the polar coordinates of the original curvelet transform to be able to implement rotation simply as an interpolation between curvelet coefficients. A rotation of a curvelet can also be written as the combination of a shear and a dilation procedure, two operations that are natural in the existing published digital implementations of the curvelet transform. Therefore, a rotation can similarly be implemented in the curvelet domain using interpolation between coefficients from existing digital curvelet transforms.

Examples of digital curvelets

Figure 3 shows examples of digital curvelets. The left column shows curvelets in the spatial domain, while the right column shows their associated amplitude spectra. Superimposed on the spectra is the spectral tiling of the digital curvelet transform. The middle column shows the associated spatial lattice for each of the curvelets, where the centers of the cells are the locations of the cen-

ters of the curvelets in space. Here, the spatial cells on the spatial lattice are colored according to the magnitude of the curvelet coefficient (here, always unity); black equals one and white equals zero. Figure 3b shows a translated version of the curvelet in Figure 3a; the spectral tile is the same, but the spatial tile has changed, i.e., indices j and l are held constant, but the translation indices m_1 and m_2 are different. Figure 3c shows a curvelet with the same spatial location and the same scale index as that in Figure 3a, but with a different angular index l. That is, the spectral tile has moved within the same filter band, i.e., within the same concentric squares. Note how the spatial lattice changes as we change the angular index l, because we use the nonequispaced fast-Fourier-transform-based curvelet transform as opposed to the wrapping-based transform (Candès et al., 2006). Finally, Figure 3d shows a dilated version of the curvelet shown in Figure 3a; the spatial location is the same, but the spectral tile has moved outward into a neighboring annulus (or subband), i.e., the scale index $j \rightarrow j + 1$. Since the neighboring annulus is subdivided into more wedges, the angular index l has really also changed, but in such a way that the direction of the curvelet is basically the same. Similarly, since the larger scale has a finer spatial sampling associated with it, the translation indices m_1 and m_2 have also changed, but again in such a way that the curvelet location is the same.

CURVELETS AS BUILDING BLOCKS OF SEISMIC DATA

We mentioned in the introduction that it seems intuitive at first sight that curvelets can be used to sparsely represent seismic data, because reflections in seismic data lie mainly along smooth curves. Even though, in this work, we focus on the sparse representation of the imaging operator using curvelets, rather than data compression with curvelets, we illustrate the idea with a simple synthetic example.

Figure 4a shows part of a synthetic common-shot gather, where the wavefront has a cusp. These data relate to a model with a syncline-shaped reflector. Figure 4b shows the reconstructed gather where only the largest curvelet coefficients were used, resulting in a compression factor of 25; i.e., we used 25 times fewer curvelets than the number of sample values in the original gather to reconstruct the data. From Figure 4c, the difference between the original and reconstructed data is negligible. The peak signal-to-noise ratio (PSNR) is almost 40 dB, where the PSNR is defined as $\mathrm{PSNR} = 20\log(\max(O)/\sqrt{\mathrm{MSE}(O,T)})$, with O and T denoting the original and thresholded gathers, respectively, and $\mathrm{MSE}(O,T)$ the mean-squared error given by $\mathrm{MSE}(O,T) = (\sum_{i=1}^{N}\sum_{j=1}^{M}|O(i,j) - T(i,j)|^2)/(N\times M)$, where i and j are indices for the time and offset samples, respectively. Note that the large compression ratio of 25 is partly a result of the synthetic data having many zero or near-zero sample values to start with, as for all other synthetic data examples shown throughout this work; for a field data example, this compression ratio would likely be smaller. Also, in that case, the simple hard thresholding used here would likely not suffice, because the signal-to-noise ratio (SNR) decreases with time and offset. There are, however, several authors who showed using both noisy synthetic and field data that curvelets allow compression of seismic data and seismic images. We refer the reader to Candès et al. (2006), Yarham et al. (2006), Hennenfent and Herrmann (2007), Herrmann et al. (2007), and references therein. The example, here, illustrates the principle and suffices in light of the main focus of this work. We reiterate that by using curvelets as building blocks of seismic data, the local slopes (or directions) in the data are built into the data representation. From this point of view, curvelets provide an appropriate reparameterization of seismic data, and the local (central) slopes in the data can be found, in principle, from the dominant curvelet coefficients resulting from decomposing the data into curvelets.

CURVELETS, WAVE PROPAGATION, AND SEISMIC IMAGING

The action of operators belonging to the class of Fourier integral operators that can be sparsely represented using curvelets, and from which the imaging operator we treat here can be constructed, follows closely the propagation of singularities (Smith, 1998b; Candès and Demanet, 2005). The action of the solution operator of the wave equation on a curvelet of a particular scale can be approximated by moving the curvelet along the ray, associated with the center of the curvelet, through the medium smoothed for that particular scale. This means that, in the appropriately smoothed media, curvelets remain localized in both the spatial domain and the spectral domain. This localization is the direct consequence of the parabolic scaling property of curvelets. Gaussian beams, though somewhat related to curvelets, lack this property, because they do not satisfy the parabolic scaling relation (or any other scaling relation). Hence, the propagated curvelet can be constructed by using neighboring curvelets only, where neighboring is understood in the context of phase space; i.e., a neighboring curvelet is a curvelet that is close in the spatial domain and has orientation close to the orientation of the curvelet that is propagated along the central ray, i.e., the ray associated with the center of the curvelet in phase space. Note that propagation of a curvelet following a ray is an operation in phase space, because it makes explicit use not only of the center location of the curvelet, but of its slope as well. In phase space, caustics are unnoticable, and hence their treatment does not require special care.

In essence, this work aims to exploit the above notion in the context of seismic imaging. The direction in which to propagate curvelets is closely related to the slopes in the data that can be found from the dominant curvelet coefficients resulting from the curvelet decomposition of the data. Seismic imaging that makes explicit use of such slope information in the data is known as map migration (Douma and de Hoop, 2006b, and references therein). We illustrate what map migration is and how it connects to seismic imaging with curvelets, using a simple example.

Figure 5a shows a CO gather (with an offset of 2000 m) from synthetic data generated from a syncline model with constant velocity above the reflector. Superimposed on the reflections, line elements are drawn tangent to the reflections. Each of these line elements de-

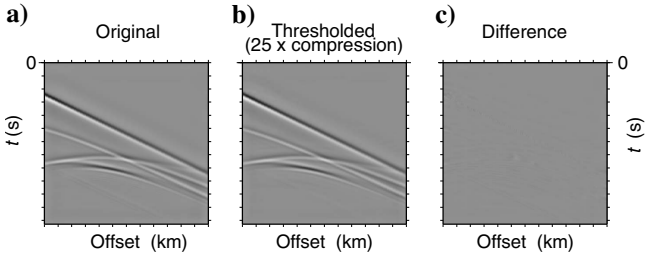

Figure 4. Synthetic common-shot gather with cusped wavefront (a), reconstructed using only the largest curvelet coefficients (b), and the difference (c). The reconstruction is almost identical to the gather in (a) (PSNR = 40 dB).

termines a local slope, p_u, whereas the center of the line determines the two-way traveltime t_u and the common-midpoint (CMP) location x_u. Map migration maps the midpoint location x_u, traveltime t_u, and local slope p_u to their migrated equivalents. The equations for this mapping for a 2D homogeneous medium and a CO acquisition geometry are given in Appendix A. Using the map-migration equations A-1–A-3 and the velocity $v = 2000$ m/s, the migrated location (y_1^m, y_2^m) and the local dip $\phi = \tan^{-1}(vp_m/2)$ can be determined from the midpoint location x_u, traveltime t_u, and local slope p_u, where ϕ is the angle with the horizontal measured clockwise positive and p_m is the local slope after migration. Figure 5b shows the migrated CO counterpart of the data shown in Figure 5a, with the migrated line elements determined using y_1^m, y_2^m, and p_m, superimposed on the image. The migrated-line elements follow the directions in the image, indicating that equations A-1–A-3 indeed capture the kinematics of CO time migration in a high-frequency approximate sense. Note that the line elements from diffractions of the edges of the syncline in Figure 5a are all mapped to the same location, but with different orientations. This is analogous to building a delta function with plane waves from all directions; i.e., a Fourier-transformed delta function has all directions.

Knowing that each curvelet has a central direction associated with it, we aim to replace the line elements in Figure 5 with curvelets. We are thus aiming to lift the applicability of map migration beyond velocity-model building, and show its use for prestack imaging with curvelets. To do this, we derive the leading-order approximation to single-arrival CO Kirchhoff depth migration of a curvelet in the data.

LEADING-ORDER APPROXIMATION TO CO KIRCHHOFF DEPTH MIGRATION USING CURVELETS

We restrict the current treatment to two dimensions for simplicity but emphasize that the extension to three dimensions is straightforward. The CO single-arrival, depth-migration equation in 2D can generally be written as (e.g., Miller et al., 1987; and Bleistein et al., 2000, sec. 5.1 and 5.2)

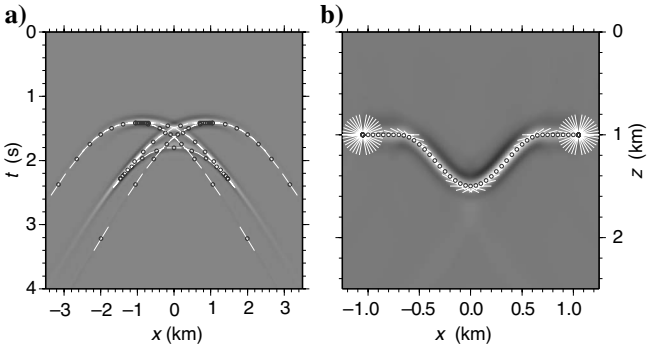

Figure 5. CO ($h = 1000$ m) data (a) and image (b) from a syncline-shaped reflector embedded in a constant velocity ($v = 2000$ m/s) medium. Superposed on both the data and the image are demigrated and migrated line elements calculated using CO map migration.

$$\beta(\mathbf{y}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a(\mathbf{y},x,\omega;h)e^{-i\omega t(\mathbf{y},x;h)}U_s(x,\omega;h)\mathrm{d}x\mathrm{d}\omega, \tag{4}$$

where $\beta(\mathbf{y})$ is the reflectivity function at migrated location $\mathbf{y} = (y_1\, y_2)^T$, ω is the angular frequency, $t(\mathbf{y},x;h)$ is the traveltime from the source to the reflection point to the receiver, and $U_s(x,\omega;h)$ denotes the scattered field, i.e., the data, at midpoint location x and a fixed half-offset h, Fourier-transformed with respect to time t. Figure 6 shows the geometry associated with 2D CO Kirchhoff depth migration. By using equation 4 as our starting point, we exclude the possibility of any multipathing occurring in the medium. Imaging in the presence of multipathing can, however, be brought into a form that allows the same analysis as the one presented here, using methods described by Stolk and de Hoop (2002).

Anticipating that we will decompose the data with curvelets, each of which is localized near a wedge in the spectral domain, we Fourier transform U_s with respect to x (using the convention given in Bleistein et al., 2000) to get

$$\beta(\mathbf{y}) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} a(\mathbf{y},x,\omega;h)e^{-i\{\omega t(\mathbf{y},x;h)-k_x x\}}\mathrm{d}x\right] \times \hat{U}_s(k_x,\omega;h)\mathrm{d}k_x\mathrm{d}\omega. \tag{5}$$

The quantity between the square brackets is an oscillatory integral that can be approximated using the method of stationary phase (e.g., Bleistein et al., 2000, p. 129). Treating $-\omega$ as the formal large parameter and defining Φ as

$$\Phi(\mathbf{y},x,k_x/\omega;h) = t(\mathbf{y},x;h) - \frac{k_x}{\omega}x, \tag{6}$$

we consider the behavior of the integral between brackets in equation 5, for large values of ω, but for a fixed horizontal slowness k_x/ω. Applying the method of stationary phase, we get

$$\beta(\mathbf{y}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a'(\mathbf{y},k_x,\omega;h)e^{-i\omega\Phi(\mathbf{y},x(\mathbf{y},k_x/\omega;h),k_x/\omega;h)} \times \hat{U}_s(k_x,\omega;h)\mathrm{d}k_x\mathrm{d}\omega, \tag{7}$$

where we have defined

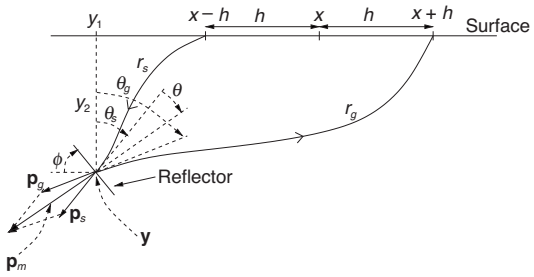

Figure 6. Geometry associated with 2D CO Kirchhoff depth migration.

$$a'(\mathbf{y},k_x,\omega;h) = \frac{a(\mathbf{y},x(\mathbf{y},k_x/\omega;h),\omega;h)}{\sqrt{2\pi|\Phi''(\mathbf{y},x(\mathbf{y},k_x/\omega;h);h)||\omega|}} \times e^{-i(\pi/4)\mathrm{sgn}\,\omega\ \mathrm{sgn}\,\Phi''(\mathbf{y},x(\mathbf{y},k_x/\omega;h);h)} \quad (8)$$

with

$$\Phi''(\mathbf{y},x(\mathbf{y},k_x/\omega;h);h) = \left.\frac{\partial^2\Phi}{\partial x^2}\right|_{x(\mathbf{y},k_x/\omega;h)} \quad (9)$$

being the Hessian evaluated at the stationary midpoint location $x(\mathbf{y},k_x/\omega;h)$. This procedure is reminiscent of plane-wave migration (Akbar et al., 1996). Considering large values of ω while keeping the horizontal slowness k_x/ω fixed is analogous to the method used when calculating the group velocity for dispersive waves (Lighthill, 1978, p. 248); this involves taking the limit as the traveltime $t\to\infty$, although the velocity Δ/t is kept fixed, where Δ is the distance away from the source.

The stationary midpoint location $x(\mathbf{y},k_x/\omega;h)$ satisfies

$$\left.\frac{\partial\Phi(\mathbf{y},x,k_x/\omega;h)}{\partial x}\right|_{x=x(\mathbf{y},k_x/\omega;h)} = 0. \quad (10)$$

This equation determines the midpoint location x as a function of the slope $p = k_x/\omega$ in the CO data and the migrated location $\mathbf{y}$. The stationary midpoint location $x(\mathbf{y},k_x/\omega;h)$ for heterogeneous (smooth) media can, in principle, be found using ray tracing (Billette and Lambaré, 1998). For the purpose of generating the numerical examples for homogeneous media in this work, we use the central midpoint location (x), central traveltime (t), and central slope (p) from the curvelet decomposition of the data to determine the migrated location ($\mathbf{y}$) through the map migration equations A-1–A-3. In this way, for each input curvelet, we know the relation between the midpoint location x on one side, and the migrated location $\mathbf{y}$ and the slope (p) in the data on the other side. For homogeneous media, it turns out that the stationary midpoint location $x(\mathbf{y},k_x/\omega;h)$ is the solution to a quartic equation in x (Douma, 2006, app. J, p. 191).

Let $\beta_\mu(\mathbf{y})$ be the image of one input curvelet c_μ, with multi-index $\mu = (j,l,m_1,m_2)$. That is, we replace the input data $\hat{U}_s(k_x,\omega;h)$ in equation 7 with one curvelet $\hat{c}_\mu(k_x,\omega)$, where the notation $\hat{c}_\mu$ denotes the Fourier transform of the curvelet $c_\mu(x,t)$ in the x-t domain. Therefore, we have

$$\beta_\mu(\mathbf{y}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a'(\mathbf{y},k_x,\omega;h)e^{-iP(\mathbf{y},k_x,\omega;h)}\hat{c}_\mu(k_x,\omega)\mathrm{d}k_x\mathrm{d}\omega, \quad (11)$$

where we have defined the phase function

$$\begin{aligned}P(\mathbf{y},k_x,\omega;h) &= \omega\boldsymbol{\Phi}(\mathbf{y},x(\mathbf{y},k_x/\omega;h),k_x,\omega;h)\\ &= \omega t(\mathbf{y},x(\mathbf{y},k_x/\omega;h);h) - k_x x(\mathbf{y},k_x/\omega;h).\end{aligned} \quad (12)$$

Comparing the integrals in equations 4 and 11, we observe that, in essence, we have extended the Fourier-integration variables from ω to (k_x,ω) by changing coordinates from x to k_x. We note that with equation 11, we have obtained a representation of CO Kirchhoff depth-migration in the same form as the Fourier-integral operator used in Candés and Demanet (2005).

Because for wave propagation with curvelets we need to propagate curvelets through media filtered appropriately for the particular scale of curvelets being propagated (Smith, 1998b), it is natural to assume that we need to do the same for migration with curvelets. This means that the traveltime $t(\mathbf{y},x;h)$ and the amplitude $a(\mathbf{y},x(\mathbf{y},k_x/\omega;h),\omega;h)$ in equation 8 are calculated for appropriately filtered media with the filter depending on the scale of the curvelet being migrated. For more insight into this filtering, we refer to the Discussion section, because this issue is not particular for Kirchhoff depth migration only. For the purpose of this derivation, it suffices to keep in mind that the traveltime and amplitude really depend on the scale index j.

Noticing that $\hat{c}_\mu(k_x,\omega)$ is localized near a wedge with center angular frequency ω_u and center horizontal wavenumber k_x^u, we linearize the dependence of the phase P on ω and k_x around ω_u and k_x^u. This means that, apart from the amplitude a', the oscillatory integral on the right-hand side of equation 11 becomes simply an inverse Fourier transform. Doing this, we have

$$\begin{aligned}P(\mathbf{y},k_x,\omega;h) \approx P(\mathbf{y},k_x^u,\omega_u;h) &+ (\omega-\omega_u)\left.\frac{dP}{d\omega}\right|_{(k_x^u,\omega_u)}\\ &+ (k_x-k_x^u)\left.\frac{dP}{dk_x}\right|_{(k_x^u,\omega_u)},\end{aligned} \quad (13)$$

where the derivatives $dP/d\omega$ and dP/dk_x denote total derivatives. From equation 12, it is immediate that

$$P(\mathbf{y},k_x^u,\omega_u;h) = \omega_u t(\mathbf{y},x(\mathbf{y},p_u;h);h) - k_x^u x(\mathbf{y},p_u;h), \quad (14)$$

where p_u is defined $p_u = k_x^u/\omega_u$. Then, calculating the total derivative, it follows from equation 12 that

$$\begin{aligned}\left.\frac{dP}{d\omega}\right|_{(k_x^u,\omega_u)} &= t(\mathbf{y},x(\mathbf{y},p_u;h);h) + \left.\left(\omega\frac{\partial t}{\partial x}\frac{\partial x}{\partial\omega}\right)\right|_{(k_x^u,\omega_u)}\\ &- \left.\left(k_x\frac{\partial x}{\partial\omega}\right)\right|_{(k_x^u,\omega_u)}.\end{aligned} \quad (15)$$

Because $p = \partial t/\partial x$ and $\omega p = k_x$, it follows that

$$\left.\frac{dP}{d\omega}\right|_{(k_x^u,\omega_u)} = t(\mathbf{y},x(\mathbf{y},p_u;h);h). \quad (16)$$

Similarly, it follows that

$$\begin{aligned}\left.\frac{dP}{dk_x}\right|_{(k_x^u,\omega_u)} &= \left.\left(\omega\frac{\partial t}{\partial x}\frac{\partial x}{\partial k_x}\right)\right|_{(k_x^u,\omega_u)} - x(\mathbf{y},p_u;h)\\ &- \left.\left(k_x\frac{\partial x}{\partial k_x}\right)\right|_{(k_x^u,\omega_u)} = -x(\mathbf{y},p_u;h).\end{aligned} \quad (17)$$

Using equations 14, 16, and 17, in equation 13, we have

$$P(\mathbf{y},k_x,\omega;h) \approx \omega t(\mathbf{y},x(\mathbf{y},p_u;h);h) - k_x x(\mathbf{y},p_u;h). \tag{18}$$

In addition, using that the amplitude $a'(\mathbf{y},k_x,\omega;h)$ varies slowly over the spectral support (i.e., a spectral wedge) of a curvelet, we approximate

$$a'(\mathbf{y},k_x,\omega;h) \approx a'(\mathbf{y},k_x^u,\omega_u;h). \tag{19}$$

Using equations 18 and 19 in equation 11, it follows that

$$\beta_\mu(\mathbf{y}) = a'(\mathbf{y},k_x^u,\omega_u;h)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{c}_\mu(k_x,\omega) \times e^{-i\{\omega t(\mathbf{y},x(\mathbf{y},p_u;h);h)-k_x x(\mathbf{y},p_u;h)\}}\mathrm{d}k_x\mathrm{d}\omega, \tag{20}$$

which is simply an inverse Fourier transform of a curvelet.

Furthermore, we make use of the earlier-mentioned fact that curvelets remain localized in the spatial domain under the action of the migration operator. From the concept of map migration, we then infer that a curvelet in the data $c_\mu(x,t)$ with center location (x_u,t_u) and dominant slope p_u, upon imaging, will be localized near the map-migrated location $\mathbf{y}_m$. Therefore, we linearize the dependence of $t(\mathbf{y},x(\mathbf{y},p_u;h);h)$ and $x(\mathbf{y},p_u;h)$ on $\mathbf{y}$ around the map-migrated location $\mathbf{y}_m$. Linearizing $x(\mathbf{y},p_u;h)$ gives

$$x(\mathbf{y},p_u;h) \approx x_u + (\nabla_{\mathbf{y}}x(\mathbf{y},p_u;h))\big|_{\mathbf{y}_m}\cdot(\mathbf{y}-\mathbf{y}_m), \tag{21}$$

where we used $x_u = x(\mathbf{y}_m,p_u;h)$ and where the notation $\nabla_{\mathbf{y}}$ denotes the gradient with respect to $\mathbf{y}$. Linearizing $t(\mathbf{y},x(\mathbf{y},p_u;h);h)$ in $\mathbf{y}$ again involves total derivatives instead of partial derivatives, because both t and x depend on $\mathbf{y}$. Hence, we have

$$\begin{aligned} &t(\mathbf{y},x(\mathbf{y},p_u;h);h)\\ &\quad\approx t(\mathbf{y}_m,x(\mathbf{y}_m,p_u;h);h)\\ &\qquad + \left.\frac{dt(\mathbf{y},x(\mathbf{y},p_u;h);h)}{d\mathbf{y}}\right|_{\mathbf{y}_m}\cdot(\mathbf{y}-\mathbf{y}_m)\\ &\quad= t_u + (\mathbf{y}-\mathbf{y}_m)\cdot\left.\left\{\nabla_{\mathbf{y}}t(\mathbf{y},x(\mathbf{y},p_u;h);h)\right.\right.\\ &\qquad\left.\left. + \frac{\partial t(\mathbf{y},x(\mathbf{y},p_u;h);h)}{\partial x}\nabla_{\mathbf{y}}x(\mathbf{y},p_u;h)\right\}\right|_{\mathbf{y}_m}\\ &\quad= t_u + (\mathbf{y}-\mathbf{y}_m)\cdot\{\nabla_{\mathbf{y}}\,t(\mathbf{y},x(\mathbf{y},p_u;h);h)\big|_{\mathbf{y}_m}\\ &\qquad + p_u\nabla_{\mathbf{y}}x(\mathbf{y},p_u;h)\big|_{\mathbf{y}_m}\}, \end{aligned} \tag{22}$$

since $p_u = (\partial t(\mathbf{y},x(\mathbf{y},p_u;h);h)/\partial x)\big|_{\mathbf{y}_m}$. Using equations 21 and 22 in equation 20, it follows that

$$\beta_\mu(\mathbf{y}) = a'(\mathbf{y},k_x^u,\omega_u;h)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{c}_\mu(k_x,\omega)e^{-i(\omega t_u-k_x x_u)} \times e^{-i\{\omega\,[\nabla_{\mathbf{y}}t(\mathbf{y},x(\mathbf{y},p_u;h);h)+p_u\nabla_{\mathbf{y}}x\,(\mathbf{y},p_u;h);h)]\big|_{\mathbf{y}_m}-k_x\,\nabla_{\mathbf{y}}x(\mathbf{y},p_u;h)\big|_{\mathbf{y}_m}\}\cdot(\mathbf{y}-\mathbf{y}_m)}\mathrm{d}k_x\mathrm{d}\omega. \tag{23}$$

Recognizing the inverse Fourier transform, we have

$$\beta_\mu(\mathbf{y}) = a'(\mathbf{y},k_x^u,\omega_u;h)c_\mu(L\cdot(\mathbf{y}-\mathbf{y}_m)+\mathbf{x}_u), \tag{24}$$

where we have defined the matrix

$$L = \begin{pmatrix} (\nabla_{\mathbf{y}}x(\mathbf{y},p_u;h)\big|_{\mathbf{y}_m})^T \\ (\{\nabla_{\mathbf{y}}t(\mathbf{y},x(\mathbf{y},p_u;h);h) + p_u\nabla_{\mathbf{y}}x(\mathbf{y},p_u;h)\}\big|_{\mathbf{y}_m})^T \end{pmatrix} \tag{25}$$

and

$$\mathbf{x}_u = \begin{pmatrix} x_u \\ t_u \end{pmatrix}. \tag{26}$$

Because $\beta_\mu(\mathbf{y})$ is based on a transformation of coordinates of a single curvelet that is localized in x and t, we further approximate $a'(\mathbf{y},k_x^u,\omega_u;h)\approx a'(\mathbf{y}_m,k_x^u,\omega_u;h)$ in equation 24. This, then, finally gives

$$\beta_\mu(\mathbf{y}) = a'(\mathbf{y}_m,k_x^u,\omega_u;h)c_\mu(L\cdot(\mathbf{y}-\mathbf{y}_m)+\mathbf{x}_u), \tag{27}$$

Therefore, it follows that the *leading-order approximation (in angular frequency, horizontal wavenumber, and migrated location) to CO Kirchhoff depth migration consists of a coordinate transformation applied to a curvelet in the data,* given by

$$\mathbf{y} = L^{-1}\cdot(\mathbf{x}-\mathbf{x}_u)+\mathbf{y}_m, \tag{28}$$

combined with an amplitude correction, where both the matrix L and the amplitude a' are evaluated with the use of prestack map migration.

The matrix L can be written as

$$L = S_{p_u}\cdot T, \tag{29}$$

with

$$S_{p_u} = \begin{pmatrix} 1 & 0 \\ p_u & 1 \end{pmatrix}, \tag{30}$$

$$T = \begin{pmatrix} (\nabla_{\mathbf{y}}x(\mathbf{y},p_u;h)\big|_{\mathbf{y}_m})^T \\ (\nabla_{\mathbf{y}}t(\mathbf{y},x(\mathbf{y},p_u;h);h)\big|_{\mathbf{y}_m})^T \end{pmatrix}. \tag{31}$$

Note that the matrix S_{p_u} defines a unilateral shear along the t-axis; i.e., the matrix S_{p_u} shears the input curvelet with slope p_u along the t-axis to have zero slope.

With the leading-order approximation to one CO migrated curvelet c_μ given by equation 27, we can determine the total image of the (scattered) data u_s using the reconstruction formula in equation 2. Letting M denote the CO migration operator, we have

$$\beta(\mathbf{y}) = [Mu_s](\mathbf{y}) = \sum_{\mu\in\mathcal{M}} \langle u_s, c_\mu\rangle [Mc_\mu](\mathbf{y})$$
$$= \sum_{\mu\in\mathcal{M}} \langle u_s, c_\mu\rangle \beta_\mu(\mathbf{y}), \quad (32)$$

where $\mathcal{M}$ is the index set containing all multi-indices μ represented in the data. Because seismic data can be sparsely represented with curvelets, we can then approximate the total image $\beta(\mathbf{y})$ as

$$\beta(\mathbf{y}) \approx \sum_{\mu\in\tilde{\mathcal{M}}} \langle u_s, c_\mu\rangle \beta_\mu(\mathbf{y}), \quad (33)$$

with $\beta_\mu(\mathbf{y})$ given by equation 27, and $\tilde{\mathcal{M}}$ being the index set that holds the multi-indices of curvelets that survive a certain threshold. Hence, $\tilde{\mathcal{M}}$ is determined by thresholding the curvelets coefficients resulting from the curvelet decomposition of the data u_s.

So far, we have decomposed the data into curvelets and established the leading-order approximation to migration of curvelets in the data. Alternatively, we could decompose the image into curvelets and find the leading-order approximation to the demigration of curvelets in the image, i.e.,

$$\beta(\mathbf{y}) = [Mu_s](\mathbf{y}) = \sum_{\mu'\in\mathcal{M}'} \langle [Mu_s], c_{\mu'}\rangle c_{\mu'}(\mathbf{y})$$
$$= \sum_{\mu'\in\mathcal{M}'} \langle u_s, [M^* c_{\mu'}]\rangle c_{\mu'}(\mathbf{y}), \quad (34)$$

where M^* is the adjoint of the migration operator, i.e., the modeling (or demigration) operator. In this case, a curvelet in the image domain is demigrated, and its projection onto the data is calculated. This procedure would thus allow target-oriented migration by demigrating curvelets contributing to a particular part of the image only.

Higher-order approximations

The leading-order approximation introduced thus far can be extracted from a precise expansion. In the above derivation, we linearized $P(\mathbf{y}, k_x, \omega; h)$ over the support of a curvelet in the spectral domain about its center (k_x^u, ω_u) (cf. equation 18). The remainder of this linearization, however, can, upon exponentiation, be absorbed in the amplitude $a' = a'(\mathbf{y}, k_x, \omega; h)$ (cf. equation 8). Letting $\tilde{a}$ denote this new amplitude, we obtain an oscillatory integral of the equation 11 type, in which the amplitude is $\tilde{a}|\hat{c}_\mu|$ and the phase is linear in accordance with equation 13. It turns out that we can expand $\tilde{a}|\hat{c}_\mu|$, which is independent of the translation indices m_1 and m_2, as an infinite sum of terms of the type $\alpha_{j,l}(\mathbf{y}; h)\hat{\gamma}_{j,l}(k_x, \omega; h)$. This can be done by recognizing that for each $\mathbf{y}$, the latter product can be expanded into a well-defined basis of functions defined on an annulus in (k_x, ω)-space (Taylor, 1991, p. 48) that contains the spectral support of curvelet c_μ. Each term of the series, then, involves convolving the curvelet c_μ by $\gamma_{j,l}$ (which affects the shape of the curvelet), applying a coordinate transformation (cf. equations 16 and 17) to the result, and multiplying by $\alpha_{j,l}$. The generalized-screen algorithm is based upon the same principles. The higher-order approximations are then obtained by (1) expanding the $\gamma_{j,l}$ beyond the linear term in (k_x, ω) about (k_x^u, ω_u), and (2) expanding equations 21 and 22 beyond the linear term in $\mathbf{y}$ about $\mathbf{y}_m$. Once the expansion is found, the application, consisting of Fourier transforms and multiplications only, should be efficient. The computation of the expansions themselves, however, might be a bottleneck. An algorithm that is also based on a separation of variables in phase space has been developed (E. J. Candès, L. Demanet, and L. Ying, personal communication, 2007), but it is not derived from a basis expansion like the one just mentioned.

LEADING-ORDER APPROXIMATION TO CO KIRCHHOFF TIME MIGRATION USING CURVELETS

In this work, all numerical examples will be calculated for homogeneous media only. We do this because the key features and limitations when using the leading-order approximation only are already apparent in such media. The expressions given in this section allow us to calculate the numerical examples shown throughout this work.

The linear coordinate transformation

The case of homogeneous media, i.e., time migration, is obviously a special case of the heterogeneous (smooth) media, i.e., depth migration, treated in the previous section. The derivation of the explicit expression of the transformation L in homogeneous media is given in Appendix B, equation B-25. Using this expression for L in equation 28, we see that the leading-order contribution to 2D CO time migration is explicitly given by the transformation of coordinates

$$\mathbf{y} = R_\phi \cdot S^{-1} \cdot D^{-1} \cdot S_{-p_u} \cdot (\mathbf{x} - \mathbf{x}_u) + \mathbf{y}_m, \quad (35)$$

where we used that $S_{-p_u} = S_{p_u}^{-1}$.

From equation 35, it follows that the leading-order approximation to CO time migration can be described by the following sequence of linear transformations (the translation part of the coordinate transformation is, of course, not linear). First, a curvelet in the data with center location $\mathbf{x}_u$ (i.e., center midpoint location x_u, center two-way traveltime t_u), and center slope p_u, is translated to the origin and sheared along the time axis to have zero slope. Subsequently, the curvelet is dilated in both the vertical and horizontal directions. Then, the curvelet is sheared along the horizontal direction and is rotated to have output dip ϕ. All of these linear transformations have the center location $\mathbf{x}_u$ of the curvelet as their origin. Finally, the resulting transformed curvelet is translated to the migrated output location $\mathbf{y}_m$. This sequence of transformations (except the translation) is depicted for one curvelet in Figure 7, where a parallellogram is added to illustrate the geometric character of these transformations.

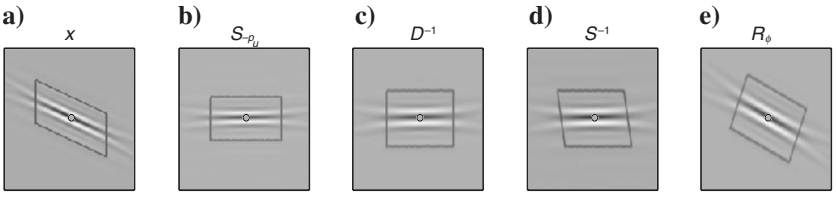

Figure 7. Illustration of the individual components of the linear transformation $L^{-1} = R_\phi \cdot S^{-1} \cdot D^{-1} \cdot S_{-p_u}$ in equation 35: input curvelet (a) after S_{-p_u} (b), D^{-1} (c), S^{-1} (d), and R_ϕ (e). The parallellogram superimposed on the curvelet in Figure 7a and its transformed versions in (b–e) to allow easy identification of the action of each individual component of L^{-1}.

Based on our knowledge of map migration (Douma and de Hoop, 2006b), the rotation and translation parts of the sequence of transformations are intuitive. The dilation in the vertical direction with $v/(2\cos\theta)$ (after shearing the curvelet along the time axis to have zero slope) takes care of both the time-depth conversion and of the lowering of the frequency content resulting from the imaging condition (also referred to as NMO stretch [Barnes, 1995] i.e., the frequency content is lowered by a factor $\cos\theta$, because the length of the migrated slowness $\mathbf{p}_m$ is smaller than the sum of the lengths of the input slownesses $\mathbf{p}_{s,r}$ (see Figure 6). Therefore, the translation, rotation, and dilation parts of the transformation conform to our intuition. This intuition was previously used by Douma and de Hoop (2005) for CO time migration with curvelets, where indeed the translation, rotation, and dilation were calculated using map migration. The current transformation thus encompasses this method of imaging with curvelets. However, the unilateral shears along both the vertical and horizontal axes, as well as the dilation of the horizontal axis, are somewhat surprising at first. To understand this, consider the special case of zero-offset migration. In that case, the matrices S_{-p_u} and R_ϕ remain unchanged, while the shear matrix S becomes the identity matrix and the dilation matrix D is given by

$$D|_{h=0} = \begin{pmatrix} 1/\cos\phi & 0 \\ 0 & 2/v \end{pmatrix}. \tag{36}$$

Observe that this matrix includes no $\cos\theta$ term because for zero offset $\theta = 0$ (i.e., $|\mathbf{p}_m| = 2/v$). This matrix does, however, include a squeeze in the horizontal direction equal to $\cos\phi$. Because this squeeze is applied after S_{-p_u}, which shears the input curvelet to have zero slope, the squeeze shortens the long axis of a curvelet. This is the familiar pulse distortion or migration stretch (Tygel et al., 1994; Barnes, 1995). We can understand the presence of this squeeze deformation by considering the extreme case when $\phi = \pi/2$, i.e., the case of a vertical reflector. In that case, the long axis of the curvelet is mapped onto a point. This can be understood by noticing that, assuming a constant velocity, we can record a reflection from a vertical reflector (i.e., $\phi = \pi/2$) in the subsurface only if this reflector is at the surface. Because a zero-offset reflection in this case will be a straight line with slope $2/v$ in the data, migration must indeed map this whole line onto a point at the surface, i.e., the position of the vertical reflector at the surface. Our coordinate transformation indeed achieves this by the squeeze deformation along the long axis of the curvelet.

Calculation of the coordinate transformation

All quantities necessary for the calculation of the coordinate transformation can be found from the curvelet decomposition of the data combined with map migration. First, the central location $\mathbf{x}_u = (x_u\ t_u)^T$ of the curvelet in the data is found from the translation indices m_1 and m_2, while the slope p_u is found from the scale index j and the angular index l. This slope determines the matrix S_{-p_u}. Then, using x_u, t_u, and p_u in the 2D map time-migration equations A-1–A-3, we find the migrated location $\mathbf{y}_m$ and the migrated slope p_m. The migrated dip ϕ can then be calculated from $\phi = \tan^{-1}(vp_m/2)$, which determines the rotation matrix R_ϕ. From $\mathbf{y}_m$ and x_u, the distances $r_{s,g}(\mathbf{y}_m, x_u; h)$ can be calculated using equations B-2 and B-3 for a given half-offset h, and the angles $\theta_{s,g}$ then follow from $\theta_{s,g} = \cos^{-1}(y_2^m/r_{s,g})$, which determine the half-opening angle $\theta = (\theta_g - \theta_s)/2$. Using the calculated values of ϕ, $r_{s,g}(\mathbf{y}_m, x_u; h)$, $\theta_{s,g}$, and θ, we can calculate the matrices D and S using equations B-21, B-23, and B-24. Note that, in digital implementations, the coordinate transform needs to be calculated on a grid. This causes the transformation to change slightly to include the grid-sampling intervals (Douma, 2006, app. L, p.197).

The amplitude a'

Besides an explicit expression for the coordinate transformation in homogeneous media, we need an expression for the amplitude a' in equation 27 to allow calculation of the numerical examples. The amplitude a' in equation 27 is evaluated in Appendix C for the special case of a homogeneous medium (equation C-2). In this evaluation, we choose to use the amplitude in 2.5D as our starting point in order to avoid the presence of any half-derivatives in the leading-order approximation to migration with curvelets.

NUMERICAL EXAMPLES

Figure 8 shows the result of standard sample-by-sample CO Kirchhoff migration of one curvelet (taken from Douma and de Hoop, 2004) in a homogeneous medium, thus not making use of our coordinate transformation. The top row shows the input curvelet in space (the vertical axis was converted to depth using $z = vt_u/2$ for convenience, with z denoting depth, and v and t_u denoting velocity and the two-way traveltime, respectively) and its associated amplitude spectrum. The spatial distribution of the coefficients is shown in the middle panel, just as in Figure 3. The leftmost panel of the bottom row shows the CO Kirchhoff-migrated curvelet. Notice how the migrated curvelet is clearly localized in space and determines only part of the isochron, in sharp contrast to the whole isochron if a single sample (or a spike) would be used as input to the migration. This confirms that in the context of migration (at least for CO time migration) curvelets are indeed a more appropriate choice for building blocks of seismic data than are spikes (currently used to represent seismic data). The spectrum of the migrated curvelet (bottom right) is also clearly localized after the migration, and overlies only four wedges in the curvelet tiling of the spectrum. This exemplifies that, at least for homogeneous media, curvelets remain fairly localized in both the spatial and wavenumber domain. The middle panel shows the coefficients for the spatial area in the lower-left quadrant of the leftmost figure (outlined by the dotted lines) for the wedges labeled 1 through 4. Indeed, this reveals that the migrated curvelet is constructed from several curvelets. We note that no smoothing of the medium as mentioned by Smith (1998b) was necessary, because the medium is here homogeneous.

Figure 9a shows the 2.5D CO Kirchhoff time-migrated curvelet from Figure 8 in more detail, whereas Figure 9d shows the real part of the associated spectrum. Figure 9b shows the result of using the leading-order approximation (equation 27) to image the same curvelet, while Figure 9c shows the difference between the Kirchhoff result and that of the leading-order approximation. Before subtraction, both images were normalized to the same maximum amplitude so that the difference shows only relative amplitude differences between the two images. Figure 9e and f shows the real part of the amplitude spectra of the images shown in Figure 9b and c. Throughout the examples, we have set the Hessian $\Phi'' = 1$ in the calculation of the amplitude a' in equation 8. This was done only to obtain similar amplitudes for both our approach and the implementation of the CO Kirchhoff migration that we used, which lacked full true-amplitude behavior. We emphasize that, in a full implementation of our approach, Φ'' should be taken into account.

For the particular curvelet used in Figure 9, the leading-order approximation (equation 27), based on a simple transformation of coordinates of curvelets, provides a good approximation to the Kirchhoff result. The maximum amplitude of the difference between both methods is 24% of the maximum amplitude in the Kirchhoff image. Note, from the patterns in the real parts of the spectra, that the curvelet is slightly curved because of the migration, whereas the linear coordinate transformation does not take such bending into account (cf. Figure 9d and e). As a consequence, the main difference in the spectrum occurs on the edges of the support of the curvelet in the frequency domain (see Figure 9f). This difference, attributable to the nonlinear deformation of the curvelet that is absent in the present leading-order approximation, can again be considered small.

Although Figure 9 shows a favorable comparison for one particular curvelet, we should also investigate the interference among several curvelets after migration. Figure 10a shows a superposition of eight curvelets, with the same central location in space and the same scale index, but different directions (or angular indices). Figure 10b shows the amplitude spectrum of the superposition of all eight curvelets, revealing that we used rightward-sloping curvelets only. Figure 10c shows the 2.5D CO Kirchhoff migrated result (the offset is here 2 km), while Figure 10d shows the result of the leading-order approximation, i.e., using equation 32, with $\beta_\mu(\mathbf{y})$ calculated using equation 27. Since the data in Figure 10a contain rightward-sloping curvelets only, the Kirchhoff result contains the left part of the band-limited isochron only, as expected. Comparing Figure 10c and d, the leading-order approximation again provides a good approximation to the Kirchhoff-migrated result; the interference of the different curvelets leads to an overall smooth left part of the band-limited isochron that compares favorably with the Kirchhoff result. Away from the isochron, both results show the weak tails of the curvelets. For the Kirchhoff result, these tails are again somewhat curved toward the isochron, while the leading-order approximation does not include such deformation. The difference between these results, however, is small.

We mentioned that our approach encompasses the intuitive approach of imaging with curvelets, presented previously by Douma and de Hoop (2005), consisting of translating and rotating curvelets according to the map-migration equations, dilating the frequency content to account for the change in frequency upon imaging, and correcting the amplitude to correct for geometrical spreading. We refer to this as the TRDA approach. Figure 10e shows the result of applying this approach to the data of Figure 10a. Clearly, the TRDA approach shows the long tails of the curvelets at the steeper parts of the isochron, whereas these tails have been much shortened using our coordinate transformation (cf. Figure 10d). This is caused by the presence of the extra dilation along the main axis of the curvelets that is absent in the TRDA approach.

Figure 11a shows a CO section (with offset equal to 1 km) of synthetic data generated from a syncline model in a constant-velocity background. We use these data to further test the curvelet-based migration and to highlight that triplications in the data are not a problem; crossing events in the data are simply constructed from curvelets that have the same location, but different orientations. In Figure 11a, we used only the largest curvelet coefficients to reconstruct the

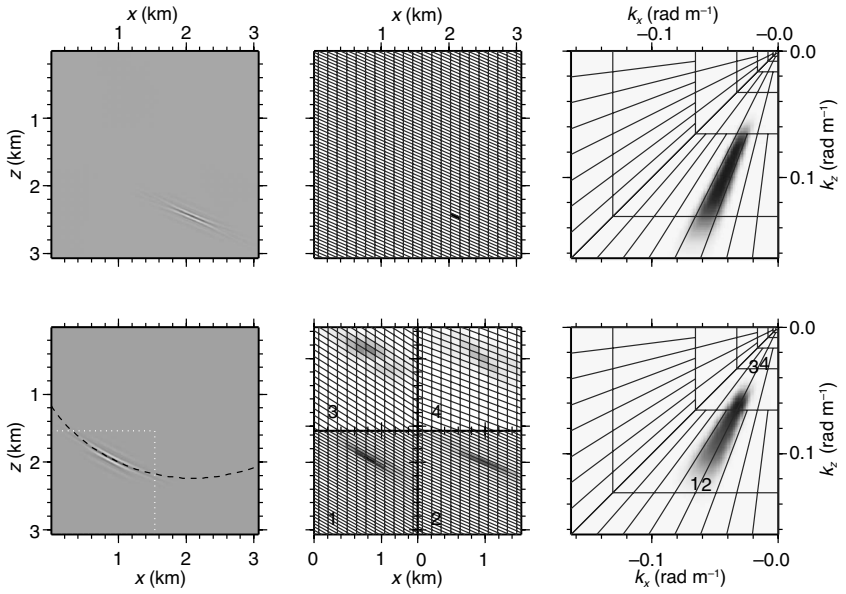

Figure 8. Top row: a curvelet with a dominant frequency of about 30 Hz (left), the normalized absolute value of the coefficient on the spatial lattice (middle), and its amplitude spectrum (right). Bottom row: CO Kirchhoff migration of the curvelet in the top row (left). The middle panel in this row shows the coefficients on the spatial lattice in the lower-left quadrant of the image (indicated with the dotted white lines) for each of the numbered wedges (labeled 1 to 4) in the spectrum (right).

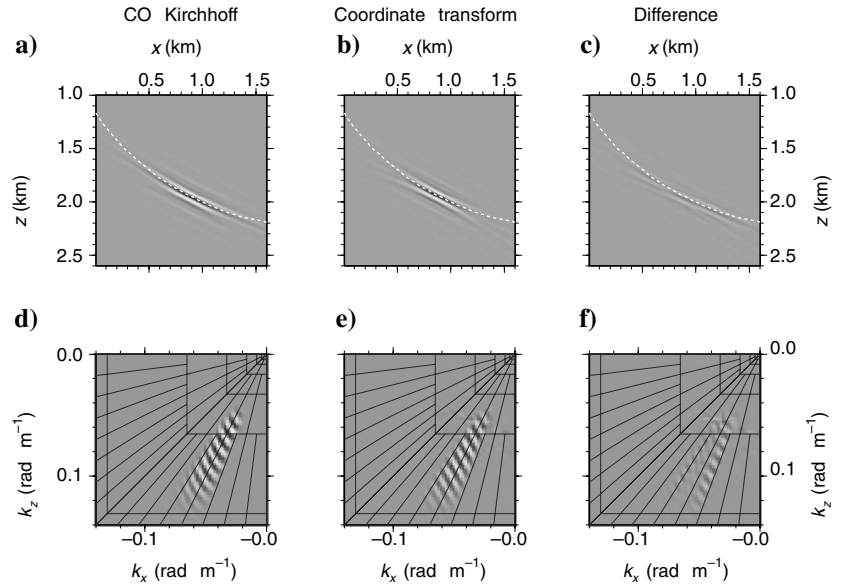

Figure 9. (a) CO Kirchhoff migration of the curvelet in Figure 6(b), coordinate transformation of this curvelet according to equation 28 (c), and the difference (d–f). The real part of the spectrum for the images shown in (a–c), respectively.

data, resulting in approximately 45 times fewer curvelets than input samples in the data with $PSNR = 54$ dB. This thresholding level was determined in the same manner as presented earlier for the data shown in Figure 4, i.e., through visually verifying that the difference between the original and reconstructed data is negligible. The associated amplitude spectrum of the reconstructed data is shown in Figure 11b. Figure 11c shows the 2.5D CO Kirchhoff-migrated result, while Figure 11d shows the curvelet-based migration calculated with the leading-order approximation, i.e, using equation 33. The leading-order, curvelet-based migration provides a good approximation to the Kirchhoff result, with some small artifacts related to the absence of the nonlinear deformation in the leading-order approach.

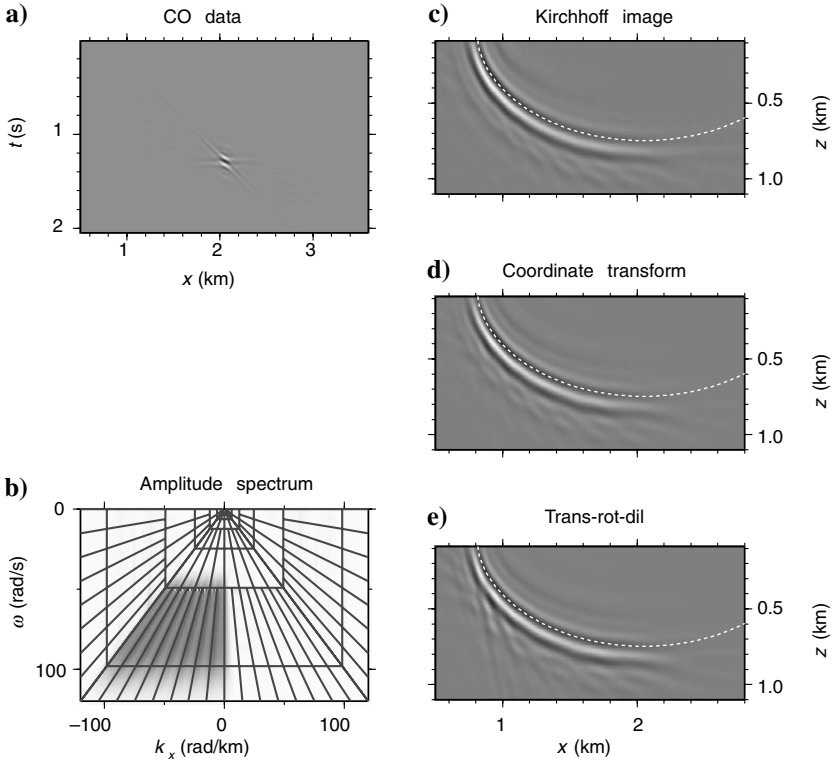

Figure 10. Superposition of eight curvelets (a), the associated amplitude spectrum (b), the CO Kirchhoff-migrated result (c), the result from the coordinate transformation of equation 28, combined with amplitude scaling (cf. equation 27) (d), and the intuitive TRDA approach from Douma and de Hoop (2005) (e).

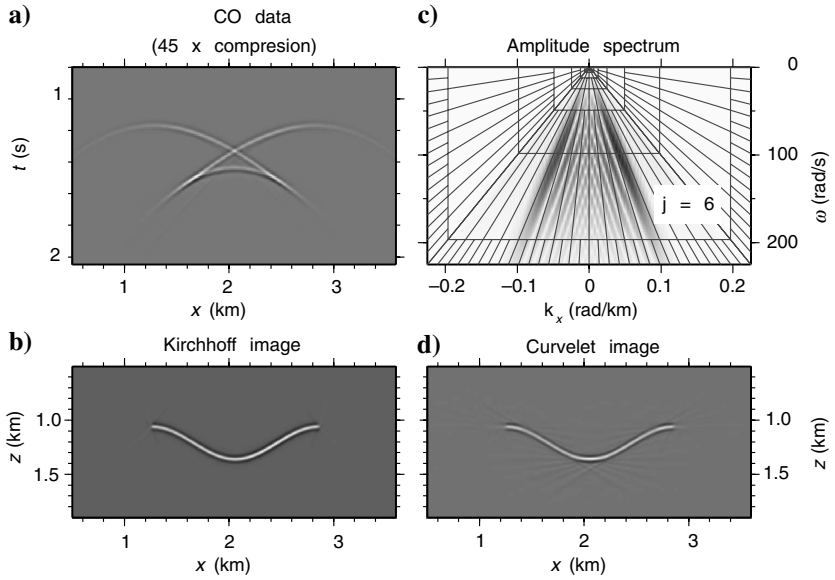

Figure 11. Sparse representation of CO data ($h = 1000$ m) from a syncline model (a) using only the largest curvelet coefficients (resulting in 45 times fewer curvelets than input samples; PSNR = 54 dB) and (b) its associated amplitude spectrum. The full data, i.e., without thresholding, were migrated using a standard CO Kirchhoff migration (c), while the thresholded data were migrated using equations 27 and 33 (d).

The accuracy illustrated in the numerical examples so far should be interpreted with care. All of our numerical examples were generated for moderate traveltimes only. The imaging of a data sample at later times results in propagation in the subsurface over longer distances. Here, the rays diverge and eventually the eight curvelets subjected to the transformation of coordinates separate. This is illustrated in Figure 12, which is as Figure 10, but now for a longer traveltime. It is our current understanding that if the rays diverge beyond the spatial support of a curvelet, the leading-order approximation is no longer accurate enough away from the center of the imaged curvelet; accuracy is then obtained only in the neighborhood of this center. This happens, for example, when the ratio of the length of the isochron to the length of a curvelet of a particular scale along its main ridge is large compared to the number of angles present in the curvelet decomposition at that scale. In this case, approximations beyond the leading order in $\mathbf{y}$ (cf. equations 21 and 22) and ω and k_x (cf. equation 18), like the one we described above, become relevant. The holes in Figure 12b are bigger than in the Kirchhoff image in Figure 12a, particularly for the steeper dips, precisely because of the lack of accuracy of the leading-order approximation for large traveltimes. The leading-order approximation is accurate enough in the neighborhood of the centers of the imaged curvelets (white dots). Note that the Kirchhoff image shows holes in the isochron in Figure 12a, because not all curvelets in the data had the same central time and central midpoint as is the case in Figure 10.

OTHER APPLICATIONS

Waves of different frequencies are sensitive to the earth's structure at different scales. Using curvelets in seismic imaging naturally allows for such analysis by simply choosing to image curvelets with selected frequency content (i.e., scale index) only. Although for constant velocity media the medium is the same at every scale, Figure 13 illustrates the principle by showing the resulting images as a function of scale index j (Figure 13a–c). Figure 13d–f shows the associated cumulative

images; the images for $j<4$ are not shown because the data have essentially no energy for $j<4$ (see Figure 11). Following such per-scale imaging, the associated image gathers can then be submitted to a velocity analysis, hence establishing a migration velocity analysis as a function of frequency. Note that per-scale imaging also allows the study of the finite-frequency interaction with the interfaces, i.e., the frequency dependence of the reflection coefficient.

Curvelet-based seismic imaging can have several potential benefits over traditional imaging algorithms. The associated local orientation of curvelets allows controlled illumination (Rietveld and Berkhout, 1992) of the subsurface. Figure 14a shows the amplitude spectrum of the data in Figure 11a, but reconstructed with rightward-sloping curvelets only. Hence, dip filtering is here achieved by simply reconstructing the data using rightward-sloping curvelets only. Figure 14b shows the resulting image in which the right part of the syncline is suppressed. Note that a partial image of the discontinuity on the right edge of the syncline is also visible, but now imaged with rightward-sloping curvelets only. Here, the controlled illumination is achieved by dip filtering the data with curvelets. More generally, one can synthesize a virtual data set by manipulating the observed data set with a curvelet transform. Dip filtering the image with curvelets, in contrast, leads naturally to the focusing-in-dip procedure of Brandsberg-Dahl et al. (2003).

Curvelets can also be directly applied in downward-continuation-based migration. The double-square-root equation is an evolution equation in depth. Its solution can be constructed, for example, via the composition of thin-slab propagators whose action on curvelets is of the form in equation 11, except that t and ω are replaced with z and k_z, respectively. We anticipate that propagation over a small depth interval will then be both advantageous and quite accurate. However, because such a procedure involves propagation over small depth steps only, at each step the downward-continued data need to be decomposed into curvelets again. If the curvelet transform was orthogonal, this would not be a problem. But because this transform is redundant, the number of curvelets is growing with each step. Hence, such a scheme would rapidly involve many curvelets. To reduce the number of curvelets again, a thresholding procedure could be applied at each step. This is currently a topic of ongoing research.

DISCUSSION

For depth migration in heterogeneous media, it appears natural to calculate the relevant coordinate transformation for a medium smoothed for the particular scale of the curvelet that is migrated, just as in the case of constructing an approximate solution to the wave equation in media with limited smoothness (Smith, 1998b). In this case, a curvelet with scale index j, i.e., a wavelength proportional to 2^{-j}, should be migrated through a smoothed medium that has essentially no heterogeneity beyond a scale proportional to $\sqrt{2^{-j}} = 2^{-j/2}$, i.e., beyond the width of a curvelet. From a physical point of view, this is reminiscent of the first Fresnel zone of a wave being proportional to $\sqrt{\lambda}$, with λ the wavelength. A wave indeed roughly averages the medium properties upon propagation over the first Fresnel zone, which is proportional to $\sqrt{\lambda}$, with λ the wavelength, and thus to $1/\sqrt{\omega}$ (Kravtsov, 1988; Jannane et al., 1989; Spetzler and Snieder, 2004). This means that for a curvelet of scale 2^j, the traveltime $t(\mathbf{y},x;h)$ in equation 6 and the amplitude $a(\mathbf{y},x(\mathbf{y},k_x/\omega;h),\omega;h)$ in equation 8 are evaluated using asymptotic-ray theory based on a Hamiltonian obtained from the wave operator

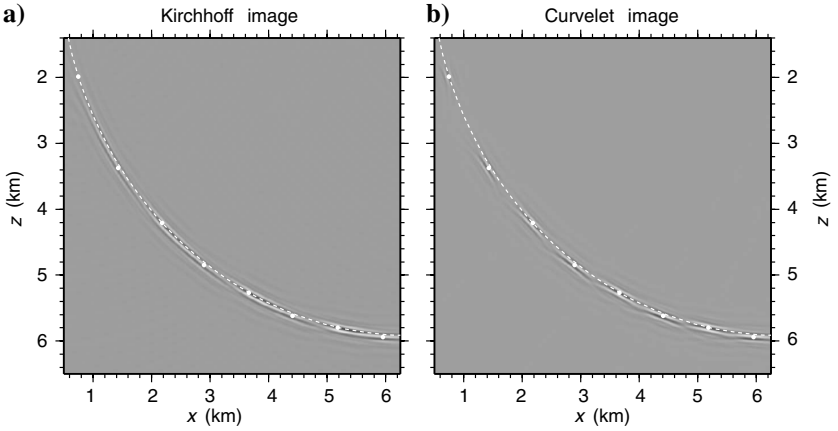

Figure 12. Essentially the same test as in Figure 8, but for a larger traveltime: CO Kirchhoff-migrated result (a) versus the result obtained from the coordinate transformation combined with amplitude scaling (b). The centers of the coordinate-transformed curvelets are indicated by the white circles. In the curvelet image, accuracy is obtained only in neighborhoods of the centers of the transformed curvelets.

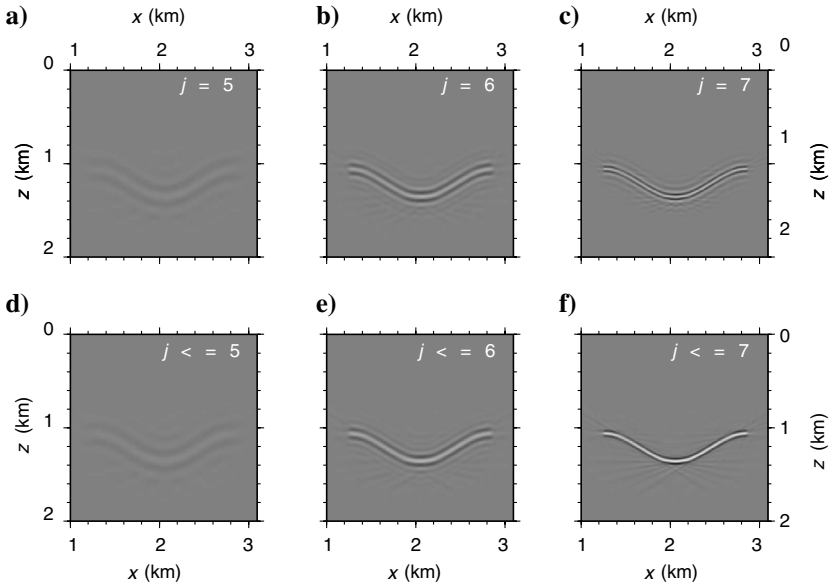

Figure 13. Images obtained using equation 27 from the thresholded CO data shown in Figure 9a, for scales $j = 5$ (a), $j = 6$ (b), $j = 7$ (c), and for the cumulative scales $j \le 5$ (d), $j \le 6$ (e), and $j \le 7$ (f). This illustrates the potential to image the subsurface at different frequencies by simply choosing to use curvelets with certain frequencies (i.e., scale indices) only.

for a band-pass-filtered medium with wavenumbers up to magnitude $2^{j/2}$ only.

We have calculated the coordinate transformation in equation 28 using an approach in the spatial domain. For each significant curvelet coefficient, we apply an inverse curvelet transform, and transform the resulting curvelet in the spatial domain according to equation 28. The current implementation does not provide an efficient algorithm to perform seismic imaging with curvelets. One could calculate our coordinate transformation in the curvelet frame itself. In this case, letting M denote the imaging operator, we have

$$\beta(\mathbf{y}) = [Mu_s](\mathbf{y}) = \sum_{\mu' \in \mathcal{M}'} \langle Mu_s, c_{\mu'} \rangle c_{\mu'}(\mathbf{y})$$

$$= \sum_{\mu \in \mathcal{M}} \sum_{\mu' \in \mathcal{M}'} \langle u_s, c_{\mu} \rangle \langle Mc_{\mu}, c_{\mu'} \rangle c_{\mu'}(\mathbf{y}), \quad (37)$$

and imaging thus becomes a mapping of indices μ in the data to μ' in the image, combined with the calculation of the coefficients $\langle Mc_{\mu}, c_{\mu'} \rangle$. In the context of our transformation of coordinates, this transformation would need to be projected onto the curvelet frame to give the matrix elements $\langle Mc_{\mu}, c_{\mu'} \rangle$. Chauris and Nguyen (2007) mention they calculate the TRDA approach to migration (Douma and de Hoop, 2005) in the curvelet domain based on a Shannon interpolation scheme, but the details of their implementation have not been published so far.

We emphasize that our procedure does not resort to a matrix representation of the migration operator using the diagonal components of the matrix only (Candès and Demanet, 2005). By *diagonal*, we mean here the components of the matrix that exactly coincide with the migrated local slopes calculated using map migration. Doing this would mean that one curvelet would be mapped by migration to one other curvelet only. This procedure is based on estimates of the sparsity of the migration-operator matrix that follow an asymptotic argument, while in migration of seismic data, we do not generally encounter such fine scales. We now know that decomposing $\beta(\mathbf{y})$ into curvelets actually leads to quite a number of significant expansion coefficients. Hence, one curvelet needs to be mapped to a collection of other curvelets (see Figure 8 for an example in homogeneous media). In our approach, one curvelet is transformed by our coordinate transformation. This transformed curvelet can be built from a collection of other curvelets. Therefore, in our approach, one curvelet is in principle mapped to several other curvelets that together form the leading-order approximation.

Combined with the fact that curvelets allow a sparse representation of the data, our work could possibly lead to a gain in computational efficiency over existing depth-migration algorithms. Our numerical examples show, however, that even for CO migration in homogeneous media, a correction beyond the leading-order is necessary when the rays diverge beyond the spatial support of a curvelet. We currently do not know how much additional computational overhead is needed for such correction.

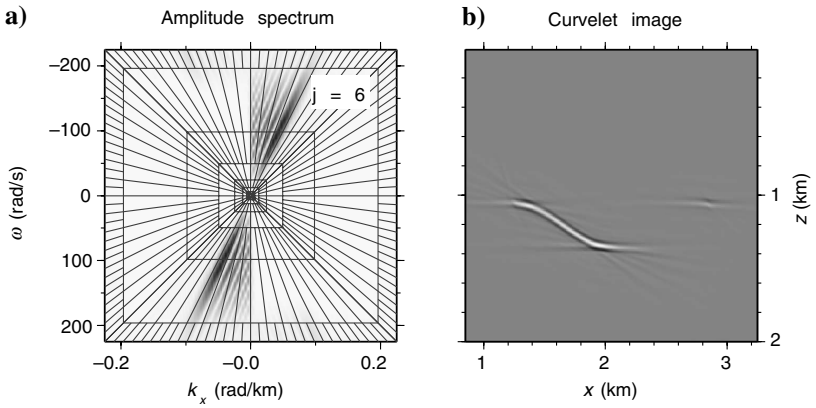

Figure 14. Amplitude spectrum of the thresholded data in Figure 9a, but now using only rightward-sloping curvelets (a), and the resulting image of these curvelets using equations 27 and 33 (b). By using curvelets with certain slopes (i.e., certain angular indices) only, we achieve a controlled illumination of the subsurface.

CONCLUSIONS

We have shown that, using curvelets, the classical CO Kirchhoff diffraction-stack approach to depth migration can to leading order be rewritten as a simple transformation of the input coordinates of curvelets in the data, combined with an amplitude correction. This transformation can be calculated using map migration, which uses the local slopes provided by the curvelet decomposition of the data. Indeed, the leading-order approximation to curvelet-based seismic imaging, carried out using map migration, suppresses the need for integration over diffraction surfaces. The coordinate transformation we derive encompasses the intuitive method of imaging using curvelets based on translation, rotation, and dilation of curvelets (i.e., the TRDA approach) that we proposed previously. Considering that the data can be sparsely represented using curvelets, the coordinate transformation needs to be calculated for many fewer curvelets than there are samples in the data.

For moderate reflection times, we numerically verified that our transformation of coordinates provides a good approximation to CO migration in homogeneous media. For longer traveltimes, however, rays eventually diverge beyond the spatial support of a curvelet. In this case, the leading-order approximation is accurate only in a neighborhood of the center of the imaged curvelet but loses accuracy away from this center. This happens, for example, when the ratio of the length of the isochron to the length of a curvelet of a particular scale along its main ridge is large compared to the number of angles present in the curvelet decomposition at that scale. In this case, it seems that, even for homogeneous media, we need to resort to corrections beyond our current leading-order approximation. In this context, we mention that higher order approximations can, in principle, be obtained from an exact representation of the migration operator by expanding the integrand of the integral representation acting on a curvelet into (infinitely many) terms, each of which reveals a separation of variables in phase space. The question of whether curvelet-based seismic imaging will lead to a gain in computational efficiency over existing algorithms remains open.

ACKNOWLEDGMENTS

We thank Emmanuel Candès for providing us with an early version of his Matlab implementation of the digital curvelet transform via unequally spaced fast Fourier transforms. The detailed comments of several anonymous reviewers

helped improve the original manuscript. H.D. thanks Roel Snieder and John Etgen for several helpful discussions.

APPENDIX A

2D CO MAP TIME MIGRATION

Douma and de Hoop (2006b) present explicit expressions for CO map time migration (i.e., migration in a medium with effectively constant velocity) that use only the slope in a CO gather (and the velocity), rather than the slopes in both a CO gather and a CMP gather (and the velocity), such as the equations presented by Sword (1987, p. 22). The expressions in 3D from Douma and de Hoop (2006b) simplify in 2D to

$$y_1^m = x_u - \left(\frac{vt_u}{2}\right)^2 \frac{\Lambda_u}{h}, \tag{A-1}$$

$$y_2^m = \sqrt{\left[1 - \left(\frac{vt_u\Lambda_u}{2h}\right)^2\right]\left\{\left(\frac{vt_u}{2}\right)^2 - h^2\right\}}, \tag{A-2}$$

$$p_m = \frac{p_u t_u |\Lambda_u - 1||\Lambda_u + 1|}{\sqrt{\left[1 - \left(\frac{vt_u\Lambda_u}{2h}\right)^2\right]\left\{t_u^2 - \left(\frac{2h}{v}\right)^2\right\}}}, \tag{A-3}$$

where we defined

$$\Lambda_u = \Lambda_u(p_u,\Theta_u,h) = \frac{1}{\sqrt{2}p_u h}\sqrt{\Theta_u\left(1 - \sqrt{1 - \frac{4(p_u h)^4}{\Theta_u^2}}\right)} \tag{A-4}$$

with

$$\begin{aligned}\Theta_u &= \Theta_u(t_u,p_u,h)\\ &= t_u^2 + \left(\frac{2h}{v}\right)^4 \frac{1}{t_u^2} - 2\left(\frac{2h}{v}\right)^2\left(1 - \left(\frac{vp_u}{2}\right)^2\right).\end{aligned} \tag{A-5}$$

In these expressions, x_u, t_u, and $p_u = \partial t_u/\partial x_u$ are the midpoint location, two-way (unmigrated) traveltime, and the (unmigrated) slope in a CO section, respectively, while y_1^m, $y_2^m = vt_m/2$, and $p_m = \partial t_m/\partial y_1^m$ are the migrated horizontal and vertical location and the migrated slope, respectively, with t_m the migrated two-way traveltime. Furthermore, h denotes the half-offset, and v is the medium velocity. Equations A-1–A-3 are explicit expressions that determine the migrated reflector coordinates (y_1^m, y_2^m, p_m) from the specular-reflection coordinates (x_u, t_u, p_u), given h and v. Note that Douma and de Hoop (2006b) used the definitions $p_u = 1/2\partial t_u/\partial x_u$ and $p_m = 1/2\partial t_m/\partial y_1^m$ instead of $p_u = \partial t_u/\partial x_u$ and $p_m = \partial t_m/\partial y_1^m$, as is done throughout this work.

At first sight, one might expect these map-migration equations to use both the local slopes at the source and the receiver, or equivalently the local slopes in a CO gather and in a CMP gather, because these slopes, after all, determine the directions in which to look into the earth for the reflector location. It turns out, however, that in the absence of caustics, only the slope in the CO gather needs to be known (Douma and de Hoop, 2006b). Equations A-1–A-3, indeed, do not use the offset horizontal slowness defined as $p_h = \partial t_u/\partial h$, so that, in practice, only p_u needs to be estimated, and the slope in a CMP gather can be ignored. It is possible to derive map-migration equations that use the offset horizontal slowness p_h instead of the velocity (Fomel, 2005). In this way, the need to estimate the velocity from CMP gathers is replaced with the need to estimate the local slopes in a CMP gather. This idea dates back to the work of Ottolini (1983). In the context of 2D prestack time migration with curvelets, the additional slopes p_h can be estimated by decomposition of the full data volume (i.e., time, midpoint, and offset) with 3D extensions of curvelets (Ying et al., 2005). In this way, both p_u and p_h can be obtained. Hence, 2D prestack time migration without velocity picking would require 3D curvelets. In the spatial domain, 3D equivalents of curvelets look like circular disks that are smooth along the disk and oscillatory orthogonal to the disk. Roughly speaking, they are smoothed, circular pieces of a band-limited plane wave in 3D.

APPENDIX B

EXPLICIT EXPRESSION FOR THE LINEAR TRANSFORMATION L FOR CO KIRCHHOFF TIME MIGRATION WITH CURVELETS

To find an explicit expression for the linear transformation L given in equations 28 and 29 for the special case of a homogeneous medium, we consider the linear transformation T in equation 31 and in particular the term $\nabla_{\mathbf{y}} t(\mathbf{y},x(\mathbf{y},p_u;h);h)|_{\mathbf{y}_m}$. For the derivation of this expression, it is important to realize that in calculating the gradient of t with respect to $\mathbf{y}$, the midpoint location x is fixed. Using that

$$t(\mathbf{y},x;h) = (r_s(\mathbf{y},x;h) + r_g(\mathbf{y},x;h))/v, \tag{B-1}$$

with r_s and r_g for a homogeneous medium given by

$$r_s = r_s(\mathbf{y},x;h) = \sqrt{(x - h - y_1)^2 + y_2^2}, \tag{B-2}$$

$$r_g = r_g(\mathbf{y},x;h) = \sqrt{(x + h - y_1)^2 + y_2^2}, \tag{B-3}$$

it follows that

$$\begin{aligned}&\nabla_{\mathbf{y}} t(\mathbf{y},x(\mathbf{y},p_u;h);h)|_{\mathbf{y}_m}\\ &= \frac{1}{v}\begin{pmatrix} -\left(\dfrac{x_u - h - y_1^m}{r_s} + \dfrac{x_u + h - y_1^m}{r_g}\right)\\ \dfrac{y_2^m}{r_s} + \dfrac{y_2^m}{r_g}\end{pmatrix}\\ &= \frac{1}{v}\begin{pmatrix} -(\sin\theta_s + \sin\theta_g)\\ \cos\theta_s + \cos\theta_g\end{pmatrix},\end{aligned} \tag{B-4}$$

where $\mathbf{y}_m = (y_1^m\ y_2^m)^T$, and where the angles θ_s and θ_g are defined in Figure 6. Then, defining the slowness vectors $\mathbf{p}_s$ and $\mathbf{p}_g$ at the source and receiver locations, respectively, as

$$\mathbf{p}_{s,g} = \frac{1}{v}\begin{pmatrix} -\sin\theta_{s,g}\\ \cos\theta_{s,g}\end{pmatrix}, \tag{B-5}$$

we can rewrite equation B-4 as

$$\nabla_{\mathbf{y}} t(\mathbf{y},x(\mathbf{y},p_u;h);h)|_{\mathbf{y}_m} = \mathbf{p}_s + \mathbf{p}_g = \mathbf{p}_m. \tag{B-6}$$

Here, $\mathbf{p}_m$ is the slowness vector associated with the dip covector $\xi_m = \omega\mathbf{p}_m$ (i.e., the wave vector associated with the reflector). Defining

$\phi = (\theta_s + \theta_g)/2$ as the dip (i.e., the angle with the horizontal measured clockwise positive), it follows that

$$\nabla_{\mathbf{y}} t(\mathbf{y}, x(\mathbf{y}, p_u; h); h)|_{\mathbf{y}_m} = \frac{2\cos\theta}{v}\begin{pmatrix} -\sin\phi \\ \cos\phi \end{pmatrix}, \tag{B-7}$$

where $\theta = (\theta_g - \theta_s)/2$ is the half-opening angle.

Next, consider the term $\nabla_{\mathbf{y}} x(\mathbf{y}, p_u; h)|_{\mathbf{y}_m}$. The stationary midpoint location satisfies

$$\left(\frac{\partial}{\partial x}\left[\frac{(r_s + r_g)}{v} - px\right]\right) = 0, \tag{B-8}$$

with r_s and r_g given in equations B-2 and B-3, respectively. Explicitly writing out this equation gives

$$\frac{x - h - y_1}{r_s} + \frac{x + h - y_1}{r_g} + pv = 0. \tag{B-9}$$

Now, treating x as the dependent variable and $\mathbf{y}$ and $p = k_x/\omega$ as the independent variables, we can take the partial derivative with respect to y_1 on both sides of equation B-9 to get

$$\left(\frac{\partial x}{\partial y_1} - 1\right)\left\{\left(\frac{1}{r_s} + \frac{1}{r_g}\right) + \frac{(x - h - y_1)^2}{r_s^3} + \frac{(x + h - y_1)^2}{r_g^3}\right\} = 0. \tag{B-10}$$

Therefore, we have

$$\left.\frac{\partial x(\mathbf{y}, p; h)}{\partial y_1}\right|_{\mathbf{y}_m} = 1. \tag{B-11}$$

This result is intuitive, because changing the horizontal component y_1 of the image location $\mathbf{y}$, while keeping the slope p constant, should result in a simple translation of the whole geometry shown in Figure 6 along the horizontal axis. Similarly, taking the partial derivative with respect to y_2 on both sides of equation B-9, we get

$$\frac{1}{r_s}\frac{\partial x}{\partial y_2} - \frac{x - h - y_1}{r_s^2}\frac{\partial r_s}{\partial y_2} + \frac{1}{r_g}\frac{\partial x}{\partial y_2} - \frac{x + h - y_1}{r_g^2}\frac{\partial r_g}{\partial y_2} = 0. \tag{B-12}$$

Using that,

$$\frac{\partial r_s}{\partial y_2} = \frac{x - h - y_1}{r_s^2}\frac{\partial x}{\partial y_2} + \frac{y_2}{r_s}, \tag{B-13}$$

$$\frac{\partial r_g}{\partial y_2} = \frac{x + h - y_1}{r_g^2}\frac{\partial x}{\partial y_2} + \frac{y_2}{r_g}, \tag{B-14}$$

and substituting these expressions in equation B-12, it follows that

$$\left.\frac{\partial x(\mathbf{y}, p)}{\partial y_2}\right|_{\mathbf{y}_m} = \frac{\tan\theta_s r_g^3 + \tan\theta_g r_s^3}{r_g^3 + r_s^3}. \tag{B-15}$$

Using this together with equation B-11, it follows that

$$\nabla_{\mathbf{y}} x(\mathbf{y}, p_u; h)|_{\mathbf{y}_m} = \left(1 \quad \frac{\tan\theta_s r_g^3 + \tan\theta_g r_s^3}{r_g^3 + r_s^3}\right)^T. \tag{B-16}$$

Finally, using the results from equations B-7 and B-16 in equation 31, we find the resulting explicit expression for the linear transformation T to be

$$T = \begin{pmatrix} 1 & \dfrac{\tan\theta_s r_g^3 + \tan\theta_g r_s^3}{r_s^3 + r_g^3} \\ \dfrac{-2\cos\theta\sin\phi}{v} & \dfrac{2\cos\theta\cos\phi}{v} \end{pmatrix}. \tag{B-17}$$

Inspection of the second row of this matrix reveals that this transformation contains a rotation with angle $\phi = (\theta_s + \theta_g)/2$, which is the migrated dip. Making this rotation explicit, we rewrite T as

$$T = T' \cdot R_{-\phi}, \tag{B-18}$$

with

$$T' = \begin{pmatrix} \cos\phi + X\sin\phi & X\cos\phi - \sin\phi \\ 0 & \dfrac{2\cos\theta}{v} \end{pmatrix}, \tag{B-19}$$

$$R_{-\phi} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}, \tag{B-20}$$

where we introduced

$$X = \frac{\tan\theta_s r_g^3 + \tan\theta_g r_s^3}{r_s^3 + r_g^3} \tag{B-21}$$

for notational convenience. The matrix $R_{-\phi}$ describes a rotation with angle ϕ. Because ϕ is clockwise positive and since t increases downward while x increases to the right, $R_{-\phi}$ describes an anticlockwise rotation. Further, inspecting equation B-19, it follows that T' can be written as the matrix product of a dilation matrix D and a unilateral shear matrix S. That is, we can write T' as

$$T' = D \cdot S, \tag{B-22}$$

with

$$D = \begin{pmatrix} \cos\phi + X\sin\phi & 0 \\ 0 & \dfrac{2\cos\theta}{v} \end{pmatrix}, \tag{B-23}$$

$$S = \begin{pmatrix} 1 & \dfrac{X\cos\phi - \sin\phi}{\cos\phi + X\sin\phi} \\ 0 & 1 \end{pmatrix}. \tag{B-24}$$

Therefore, using equations B-18 and B-22 in equation 29, it follows that

$$L = S_{p_u} \cdot D \cdot S \cdot R_{-\phi}. \tag{B-25}$$

APPENDIX C

THE AMPLITUDE FOR 2.5D CO KIRCHHOFF TIME MIGRATION USING CURVELETS

For the special case of 2.5D CO time migration in homogeneous media the amplitude $a(\mathbf{y},x,\omega;h)$ in equation 4 is given by (Bleistein et al., 2000, p. 309, equation 6.3.25)

$$a(\mathbf{y},x,\omega;h) = 4y_2\sqrt{\frac{|\omega|}{2\pi v^3}}\frac{\sqrt{r_s + r_g}(r_s^2 + r_g^2)}{(r_s r_g)^{3/2}} \times\cos\,\theta(x,\mathbf{y};h)e^{i(\pi/4)\mathrm{sgn}\,\omega}, \qquad \text{(C-1)}$$

with v the velocity, $\theta(x,\mathbf{y};h)$ half the opening angle between the slowness vectors at the source and the receiver locations (Figure 6), and r_s and r_g given by equations B-2 and B-3, respectively. Following the same procedure as outlined in the main text for CO depth migration, but now with the amplitude $a(\mathbf{y},x,\omega;h)$ given by equation C-1, we find that the amplitude $a'(\mathbf{y},k_x,\omega;h)$ is given by

$$a'(\mathbf{y},k_x,\omega;h) = \frac{2y_2}{\pi\sqrt{v^3|\Phi''|}}\frac{\sqrt{r_s + r_g}(r_s^2 + r_g^2)}{(r_s r_g)^{3/2}}\cos\,\theta(x,\mathbf{y};h), \qquad \text{(C-2)}$$

where it is understood that r_s, r_g, θ, and Φ'' are evaluated at the stationary midpoint location $x(\mathbf{y},k_x/\omega;h)$. Note that the term $\sqrt{|\omega|}$ has disappeared, because the amplitude $a(\mathbf{y},x,\omega;h)$ has $\sqrt{|\omega|}$ in the numerator. Calculating the Hessian explicitly gives

$$\Phi'' = \frac{y_2^2}{v}\left(\frac{1}{r_s^3(\mathbf{y},x;h)} + \frac{1}{r_g^3(\mathbf{y},x;h)}\right). \qquad \text{(C-3)}$$

Because $r_s(\mathbf{y},x;h)$ and $r_g(\mathbf{y},x;h)$ represent distances (that are always positive) and because we have $y_2 \geq 0$, it follows that the Hessian is positive definite, whence $\mathrm{sgn}(\Phi'') = 1$. In obtaining equation C-2, therefore, the term $e^{i(\pi/4)\mathrm{sgn}\,\omega}$ in equation C-1 was canceled by the term $e^{-i(\pi/4)\mathrm{sgn}\,\omega\,\mathrm{sgn}\,\Phi''} = e^{-i(\pi/4)\mathrm{sgn}\,\omega}$ in equation 8.

REFERENCES

Akbar, F., M. Sen, and P. Stoffa, 1996, Prestack plane-wave Kirchhoff migration in laterally varying media: Geophysics, **61**, 1068–1079.

Barnes, A., 1995, Discussion on: "Pulse distortion in depth migration," M. Tygel, J. Schleicher, and P. Hubral, authors: Geophysics, **60**, 1942–1944.

Billette, F., and G. Lambaré, 1998, Velocity macro-model estimation from seismic reflection data by stereotomography: Geophysical Journal International, **135**, 671–690.

Billette, F., S. Le Bégat, P. Podvin, and G. Lambaré, 2003, Practical aspects and applications of 2D stereotomography: Geophysics, **68**, 1008–1021.

Bleistein, N., J. Cohen, and J. Stockwell, 2000, Mathematics of multidimensional seismic imaging, migration, and inversion: Springer-Verlag, New York, Inc.

Brandsberg-Dahl, S., M. de Hoop, and B. Ursin, 2003, Focusing in dip and AVA compensation on scattering-angle/azimuth common image gathers: Geophysics, **68**, 232–254.

Candès, E. J., and L. Demanet, 2005, The curvelet representation of wave propagators is optimally sparse: Communications on Pure and Applied Mathematics, **58**, 1472–1528.

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

Candès, E. J., and D. L. Donoho, 2000, Curvelets — A surprisingly effective nonadaptive representation for objects with edges, *in* C. Rabut, A. Cohen, and L. L. Schumacher, eds., Curves and surfaces: Vanderbilt University Press, 105–120.

———, 2004, New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities: Communications on Pure and Applied Mathematics, **57**, 219–266.

Candès, E. J., and F. Guo, 2002, New multiscale transforms, minimum total variation synthesis: Applications to edge-preserving image reconstruction: Signal Processing, **82**, 1519–1543.

Chauris, H., 2006, Seismic imaging in the curvelet domain and its implications for the curvelet design: 76th Annual International Meeting, SEG, Expanded Abstracts, 2406–2410.

Chauris, H., and T. Nguyen, 2007, Towards interactive seismic imaging with curvelets: Presented at the 69th Annual Conference and Exhibition, EAGE, Session WO8.

Claerbout, J. F., 1992, Earth soundings analysis: Processing versus inversion: Blackwell Scientific Publications, Inc.

Córdoba, A., and C. Fefferman, 1978, Wave packets and Fourier integral operators: Communications in Partial Differential Equations, **3**, 979–1005.

Do, M., and M. Vetterli, 2003, Contourlets, *in* J. Stoeckler and G. V. Welland, eds., Beyond wavelets: Academic Press, Inc., 1–27.

Douma, H., 2006, A hybrid formulation of map migration and wave-equation-based migration using curvelets: Ph.D. dissertation, Colorado School of Mines.

Douma, H., and M. de Hoop, 2004, Wave-character preserving pre-stack map migration using curvelets: 74th Annual International Meeting, SEG, Expanded Abstracts, 961–964.

———, 2005, On common-offset pre-stack time-migration with curvelets: 75th Annual International Meeting, SEG, Expanded Abstracts, 2009–2014.

———, 2006a, Leading-order seismic imaging using curvelets: 76th Annual International Meeting, SEG, Expanded Abstracts, 2411–2415.

———, 2006b, Explicit expressions for prestack map time-migration in isotropic and VTI media and the applicability of map depth-migration in heterogeneous anisotropic media: Geophysics, **71**, no. 1, S13–S28.

Fefferman, C., 1973, A note on spherical summation multipliers: Israel Journal of Mathematics, **15**, 44–52.

Fomel, S., 2002, Applications of plane-wave destruction filters: Geophysics, **67**, 1946–1960.

———, 2005, Velocity-independent time-domain seismic imaging using local event slopes: 75th Annual International Meeting, SEG, Expanded Abstracts, 2269–2273.

Guo, K., and D. Labate, 2007, Optimally sparse multidimensional representation using shearlets: Society for Industrial and Applied Mathematics Journal on Mathematical Analysis, **39**, 298–318.

Harlan, W., and R. Burridge, 1983, A tomographic velocity inversion for untracked data: Stanford Exploration Project report SEP-37: Stanford University, 1–7.

Hennenfent, G., and F. Herrmann, 2007, Curvelet reconstruction with sparsity-promoting inversion: Successes and challenges: Presented at the 69th Annual Conference and Exhibition, EAGE, Session WO8.

Hernández, E., and G. Weiss, 1996, A first course on wavelets: CRC Press, Inc.

Herrmann, F., 2003a, Multifractional splines: Application to seismic imaging, *in* M. A. Unser, A. Aldroubi, and A. F. Laine, eds., Proceedings of SPIE technical conference on wavelets: Applications in signal and image processing X, 5207, 240–258.

———, 2003b, Optimal imaging with curvelets: 73rd Annual International Meeting, Expanded Abstracts, SEG, 997.

Herrmann, F., G. Hennenfent, and P. Moghaddam, 2007, Seismic imaging and processing with curvelets: Presented at the 69th Annual Conference 2nd Exhibition, EAGE, Session WO8.

Hua, B., and G. A. McMechan, 2001, Parsimonious 2D poststack Kirchhoff depth migration: Geophysics, **66**, 1497–1503.

———, 2003, Parsimonious 2D prestack Kirchhoff depth migration: Geophysics, **68**, 1043–1051.

Jannane, H., W. Beydoen, E. Crase, D. Cao, Z. Koren, E. Landa, M. Menses, A. Pica, M. Noble, G. Roeth, S. Singh, R. Snieder, A. Tarantola, D. Trezeguet, and M. Xie, 1989, Wavelengths of earth structures that can be resolved from seismic reflection data: Geophysics, **54**, 906–910.

Kingsbury, N., 1999, Image processing with complex wavelets: Philosophical Transactions of the Royal Society of London, **A357**, 2543–2560.

———, 2001, Complex wavelets for shift invariant analysis and filtering of signals: Journal of Applied and Computational Harmonic Analysis, **10**, 234–253.

Kleyn, A., 1977, On the migration of reflection time contour maps: Geophysical Prospecting, **25**, 125–140.

Kravtsov, Y., 1988, Rays and caustics as physical objects, *in* E. Wolf, ed., Progress in optics, XXVI: Elsevier Science Publishing Company, Inc., 227–348.

Lighthill, M., 1978, Waves in fluids: Cambridge University Press.

Mallat, S. G., 1998, A wavelet tour of signal processing: Academic Press, Inc.

Miller, D., M. Oristaglio, and G. Beylkin, 1987, A new slant on seismic imaging — Migration and integral geometry: Geophysics, **52**, 943–964.

Moghaddam, P., F. Herrmann, and C. Stolk, 2007, Seismic amplitude recovery with curvelets: Presented at the 69th Annual Conference and Exhibition, EAGE, Session WO8.

Ottolini, R., 1983, Velocity independent seismic imaging: Stanford Exploration Project report SEP-37: Stanford University, 59–68.
Riabinkin, L., 1991, Fundamentals of resolving power of controlled directional reception (CDR) of seismic waves, *in* G. Gardner and L. Lu, eds., Slant stack processing: SEG Geophysics Reprints Series, 36–60.
Rietveld, W., and A. Berkhout, 1992, Depth migration combined with controlled illumination: 62nd Annual International Meeting, SEG, Expanded Abstracts, 931–934.
Smith, H., 1998a, A Hardy space for Fourier integral operators: Journal of Geometric Analysis, **8**, 629–654.
——–, 1998b, A parametrix construction for wave equations with $C^{1,1}$ coefficients: Annales de L'Institut Fourier **48**, 797–835.
Spetzler, J., and R. Snieder, 2004, The Fresnel volume and transmitted waves: Geophysics, **69**, 653–663.
Stein, E. M., 1993, Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals, 2nd ed.: Princeton University Press.
Stolk, C. C., and M. V. de Hoop, 2002, Microlocal analysis of seismic inverse scattering in anisotropic, elastic media: Communications on Pure and Applied Mathematics, **55**, 261–301.
Sword, C. H., 1987, Tomographic determination of interval velocities from reflection seismic data: The method of controlled directional reception: Ph.D. dissertation, Stanford University.
Taylor, M., 1991, Pseudodifferential operators and nonlinear PDE: Birkhäuser, Boston, Inc.
Tygel, M., J. Schleicher, and P. Hubral, 1994, Pulse distortion in depth migration: Geophysics, **59**, 1561–1569.
Yarham, C., U. Boeniger, and F. Herrmann, 2006, Curvelet-based ground roll removal: 76th Annual International Meeting, SEG, Expanded Abstracts, SPNA 1–7.
Ying, L., L. Demanet, and E. Candés, 2005, 3D discrete curvelet transform, http://www.curvelet.org, accessed August 16, 2007.
Zavalishin, B., 1981, Perfection of methods for constructing seismic images using controlled directional reception: Soviet Geology and Geophysics, **22**, 98–104.