Curvelet-based seismic data processing: a multiscale and nonlinear approach

Felix J. Herrmann[*], **Deli Wang**[†], **Gilles Hennenfent**[*], **Peyman P. Moghaddam**[*]

ABSTRACT

In this letter, the solutions to three seismic processing problems are presented that exploit the multiscale and multi-angular properties of the curvelet transform. Data regularization, multiple removal, and restoration of migration amplitudes are all formulated in terms of a sparsity promoting program that employs the high degree of sparsity attained by curvelets on seismic data and images. For each problem the same nonlinear program is solved, simultaneously minimizing the data misfit and the one norm (ℓ_1) on the desired curvelet-domain solution. Parsimony of curvelets on seismic wavefields and images of the sedimentary crust with wavefront-like features underlies the successful solution of these problems and is a clear indication of the broad applicability of this transform in exploration seismology.

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

[†]Jilin University, 2699 Qianjin street, Changchun, 130012, China. Visiting the University of British Columbia.

INTRODUCTION

In this letter, we demonstrate that the discrete curvelet transform (Candès et al., 2006a; Hennenfent and Herrmann, 2006b) can be used to reconstruct seismic data from incomplete measurements, to separate primaries and multiples and to restore migration amplitudes. The crux of the method lies in the combination of the curvelet transform, which attains a fast decay for the magnitude-sorted curvelet coefficients, with a sparsity promoting program. By themselves sparsity-promoting programs are not new to the geosciences (Sacchi et al., 1998). However, sparsity promotion with the curvelet transform is new. The curvelet transform's unparalleled ability to detect wavefront-like events that are locally linear and coherent means it is particularly well suited to seismic data problems. In this paper, we show examples including data regularization (Hennenfent and Herrmann, 2006a, 2007b), primary-multiple separation (Herrmann et al., 2007a) and migration-amplitude recovery (Herrmann et al., 2007b). Application of this formalism to wavefield extrapolation is presented elsewhere (Lin and Herrmann, 2007).

CURVELETS

Curvelets are localized 'little plane-waves' (see Hennenfent and Herrmann, 2006b, and the on-line ancillary material for an introduction on this topic) that are oscillatory in one direction and smooth in the other direction(s). They are multiscale and multi-directional. Curvelets have an anisotropic shape – they obey the so-called parabolic scaling relationship, yielding a width $\propto$ length^2 for the support of curvelets in the physical domain. This anisotropic scaling is necessary to detect wavefronts and explains their high compression rates on seismic data and images (Candès et al., 2006a; Herrmann et al., 2007b), as long as

2

these datasets can be represented as functions with events on piece-wise twice differentiable curves. Then, the events become linear at the fine scales justifying an approximation by the linearly shaped curvelets. Even seismic data with caustics, pinch-outs, faults or strong amplitude variations fit this model, which amounts to a preservation of the sparsity attained by curvelets.

Curvelets represent a specific tiling of the 2-D/3-D frequency plane into strictly localized wedges. Because the directional sampling increases every-other scale doubling, curvelets become more anisotropic at finer scales. Curvelets compose multi-D data according to $\mathbf{f} = \mathbf{C}^T \mathbf{C} \mathbf{f}$ with $\mathbf{C}$ and $\mathbf{C}^T$ the forward and inverse discrete curvelet transform matrices (defined by the fast discrete curvelet transform, FDCT, with wrapping, a type of periodic extenstion, see Candès et al., 2006a; Ying et al., 2005). The symbol ^{T} represents the transpose, which is equivalent to the inverse for this choice of curvelet transform. This transform has a moderate redundancy (a factor of roughly 8 in 2-D and 24 in 3-D) and a computational complexity of $\mathcal{O}(n \log n)$ with n the length of $\mathbf{f}$. Even though $\mathbf{C}^T \mathbf{C} = \mathbf{I}$, with $\mathbf{I}$ the identity matrix, the converse is not true, i.e., $\mathbf{C}\mathbf{C}^T \neq \mathbf{I}$. This ambiguity can be removed by adding sparsity promotion as a constraint.

COMMON PROBLEM FORMULATION BY
SPARSITY-PROMOTING INVERSION

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

reads

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

with $\mathbf{y}$ a vector with noisy and possibly incomplete measurements; $\mathbf{A}$ the modeling matrix that includes $\mathbf{C}^T$; and $\mathbf{n}$, a zero-centered white Gaussian noise. Because of the redundancy of $\mathbf{C}$ and/or the incompleteness of the data, the matrix $\mathbf{A}$ can not readily be inverted. However, as long as the data, $\mathbf{y}$, permits a sparse vector, $\mathbf{x}_0$, the matrix, $\mathbf{A}$, can be inverted by a sparsity-promoting program (Candès et al., 2006b; Donoho, 2006):

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

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

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

curvelet vector captures the sparsity and the cooling, which speeds up the algorithm, allows additional coefficients to fit the data.

SEISMIC DATA RECOVERY

The reconstruction of seismic wavefields from regularly-sampled data with missing traces is a setting where a curvelet-based method will perform well. As with other transform-based methods, sparsity is used to reconstruct the wavefield by solving $\mathbf{P}_\epsilon$. It is also shown that the recovery performance can be increased when information on the major primary arrivals is included in the modeling operator.

Curvelet-based recovery

The reconstruction of seismic wavefields from incomplete data corresponds to the inversion of the picking operator $\mathbf{R}$. This operator models missing data by inserting zero traces at source-receiver locations where data is missing passing recorded traces unchanged. The task of the recovery is to undo this operation by filling in the zero traces. Since seismic data is sparse in the curvelet domain, the missing data can be recovered by compounding the picking operator with the curvelet modeling operator, i.e., $\mathbf{A} := \mathbf{R}\mathbf{C}^T$. With this definition for the modeling operator, solving $\mathbf{P}_\epsilon$ corresponds to seeking the sparsest curvelet vector whose inverse curvelet transform, followed by the picking, matches the data at the nonzero traces. Applying the inverse transform (with $\mathbf{S} := \mathbf{C}$ in $\mathbf{P}_\epsilon$) gives the interpolated data. For details on the conditions that determine successful recovery, refer to recent contributions by the first and third named authors (Hennenfent and Herrmann, 2007b,a; Herrmann and Hennenfent, 2007).

An example of curvelet-based recovery is presented in Figure 1 which shows the results of decimating, and then reconstructing, a seismic dataset. The original shot and receiver spacings were 25m, and 80 % of the traces were thrown out at random (see Figure 1(b)). Comparing the 'ground truth' in Figure 1(a) with the recovered data in Figure 1(c) shows a successful recovery in case the high-frequencies are removed. Aside from sparsity in the curvelet domain, no prior information was used during the recovery, which is quite remarkable. Part of the explanation lies in the curvelet's ability to locally exploit the 3-D geometry of the data and this suggests why curvelets are successful for complex datasets where other methods may fail.

Focused recovery

In practice, additional information on the to-be-recovered wavefield is often available. For instance, one may have access to the predominant primary arrivals or to the velocity model. In that case, the recently introduced *focal* transform (Berkhout and Verschuur, 2006), which 'deconvolves' the data with an estimate of the primaries, incorporates this additional information into the recovery process. Application of this primary operator, $\mathbf{\Delta P}$, adds a wavefield interaction with the surface, mapping primaries to first-order surface-related multiples (Verschuur and Berkhout, 1997; Herrmann, 2007). Inversion of this operator, strips the data off one interaction with the surface, focusing primary energy to (directional) sources. This focusing correponds to a collapse of the 3-D primary events to an approximate line source which has a more sparse representation in the curvelet domain.

By compounding the non-adaptive, independent of the data, curvelet transform with the data-adaptive *focal* transform, i.e., $\mathbf{A} := \mathbf{R\Delta P C}^{T}$, the recovery can be improved by solving

$\mathbf{P}_\epsilon$. The solution of $\mathbf{P}_\epsilon$ now entails the inversion of $\mathbf{\Delta P}$, yielding the sparsest set of curvelet coefficients that matches the incomplete data when 'convolved' with the primaries. Applying the inverse curvelet transform, followed by 'convolution' with $\mathbf{\Delta P}$ yields the interpolation, i.e. $\mathbf{S}^T := \mathbf{\Delta P C}^T$. Comparing the curvelet recovery with the focused curvelet recovery (Figure 1(c) and 1(d)) shows an overall improvement in the recovered details.

SEISMIC SIGNAL SEPARATION

Predictive multiple suppression involves two steps, namely multiple prediction and primary-multiple separation. In practice, the second step appears difficult and adaptive least-squares ℓ_2-matched-filtering techniques are known to lead to residual multiple energy, high frequency jitter and deterioration of the primaries (Herrmann et al., 2007a). By employing the curvelet's ability to detect wavefronts with conflicting dips (e.g. caustics), a non-adaptive, independent of the total data, separation scheme can be defined that is robust with respect to moderate errors in the multiple prediction. The nonlinear program, $\mathbf{P}_\epsilon$, with $\mathbf{y}$ defined by the total data, can be adapted to separate multiples from primaries by replacing the ℓ_1 norm by a weighted ℓ_1 norm, i.e., $\|\mathbf{x}\|_1 \mapsto \|\mathbf{x}\|_{1,\mathbf{w}} = \sum_\mu |w_\mu x_\mu|$ with μ running over all curvelets and $\mathbf{w}$ a vector with positive weights. By defining these weights proportional to the magnitude of the curvelet coefficients of the 2-D SRME-predicted multiples, the solution of $\mathbf{P}_\epsilon$ with $\mathbf{A} := \mathbf{C}^T$ removes multiples. Primaries and multiples naturally separate in the curvelet domain and the weighting further promotes this separation while solving $\mathbf{P}_\epsilon$. The weights that are fixed during the optimization penalize the entries in the curvelet vector for which the predicted multiples are significant. The emphasis on the weights versus the data misfit (the proportionality constant) is user defined. The estimate for the primaries is obtained by inverse curvelet transforming the curvelet vector that minimizes $\mathbf{P}_\epsilon$ for the

7

weighted ℓ_1 norm ($\mathbf{A} = \mathbf{S}^T := \mathbf{C}^T$).

Figure 2 shows an example of 3-D curvelet-based primary-multiple separation of a North Sea dataset with the weights set according to the curvelet-domain magnitudes of the SRME-predicted multiples multiplied by 1.25. Comparison between the estimates for the primaries from adaptive subtraction by ℓ_2-matched filtering (Verschuur and Berkhout, 1997) and from our nonlinear and non-adaptive curvelet-based separation shows an improvement in (i) the elimination of the focused multiple energy below shot location 1000 m, induced by out-of-plane scattering due to small 3-D variations in the multiple-generating reflectors and (ii) an overall improved continuity and noise reduction. This example demonstrates that the multiscale and multi-angular curvelet domain can be used to separate primaries and multiples given an inaccurate prediction for the multiples. However, the separation goes at the expense of a moderate loss of primary energy which compares favorably compared to the loss associated with ℓ_2-matched filtering (see also Herrmann et al., 2007a).

MIGRATION-AMPLITUDE RECOVERY

Restoring migration amplitudes is another area where curvelets can be shown to play an important role. In this application, the purpose is to replace computationally expensive amplitude recovery methods, such as least-squares migration (Nemeth et al., 1999; Kuhl and Sacchi, 2003), by an amplitude scaling (Guitton, 2004). This scaling can be calculated from a demigrated-migrated reference vector close to the actual reflectivity.

In order to exploit curvelet sparsity, we propose to scale in the curvelet domain. This choice seems natural because migrated images suffer from spatially varying and dip-dependent amplitude deterioration that can be accommodated by curvelets. The advantages of this

8

approach are manifold and include (i) a correct handling of reflectors with conflicting dips and (ii) a stable curvelet sparsity-promoting inversion of the diagonal that restores the amplitudes and removes the clutter by exploiting curvelet sparsity on the model.

The method is based on the approximate identity: $\mathbf{K}^T\mathbf{K}\mathbf{r} \approx \mathbf{C}^T\mathbf{D}_r\mathbf{C}\mathbf{r}$ with $\mathbf{K}$ and $\mathbf{K}^T$ the demigration, migration operators and $\mathbf{D}_r$ a reference-model specific scaling (Herrmann et al., 2007b). By defining the modeling matrix as $\mathbf{A} := \mathbf{C}^T\sqrt{\mathbf{D}_r}$, $\mathbf{P}_\epsilon$ can be used to recover the migration amplitudes from the migrated image. Possible spurious side-band effects and erroneously detected curvelets (Candès and Guo, 2002) are removed by supplementing the ℓ_1 norm in $\mathbf{P}_\epsilon$ with an anisotropic diffusion norm (Fehmers and Höcker, 2003). This norm enhances the continuity along the imaged reflectors and removes spurious artifacts.

Results for the SEG AA' dataset (O'Brien and Gray, 1996; Aminzadeh et al., 1997) are summarized in Figure 3. These results are obtained with a reverse-time 'wave-equation' finite-difference migration code. To illustrate the recovery performance, idealized seismic data is generated by demigration, followed by adding white Gaussian noise, yielding a signal-to-noise ratio (SNR) of only 3 dB. This data is subsequently migrated and used as input. Despite the poor SNR, the image in Figure 3(a) contains most reflectors, which can be explained by the redundancy of the data, the migration operator's sophistication (diffractions at the bottom of the salt are handled correctly) and the perfect 'match' between the demigration and migration operators. However, the noise gives rise to clutter and there is dimming of the amplitudes, in particular for steep dips under the salt. Nonlinear recovery removes most of this clutter and more importantly the amplitudes for the sub-salt steep-dipping events are mostly restored. This idealized example shows how curvelets can be used to recover the image amplitudes. As long as the background velocity model is sufficiently smooth and the reflectivity sufficiently sparse, this recovery method can be expected to

perform well even for more complex images.

DISCUSSION AND CONCLUSIONS

The presented examples show that problems in data acquisition and imaging can be solved with a common problem formulation during which sparsity in the curvelet domain is promoted. For curved wavefront-like features that oscillate in one direction and that are smooth in the other direction(s), curvelets attain high compression rates while other transforms do not necessarily achieve sparsity for these geometries. Seismic images of sedimentary basins and seismic wave arrivals in the data both behave in this fashion, so that curvelets are particularly valuable for compression. It is this compression that underlies the success of our sparsity promoting formulation. First, we showed on real data that missing data can be recovered by solving a nonlinear optimization problem where the data misfit and the ℓ_1-norm on the curvelet coefficients are simultaneously minimized. This recovery is improved further with a combined curvelet-focal transform. Sparsity also proved essential during the primary-multiple separation. In this case, it leads to a form of decorrelation of primaries and multiples, reducing the probability of having large overlapping curvelet entries between these different events. Finally, the sparsity of curvelets on the image itself was exploited to recover the migration amplitudes of the synthetic subsalt imaging example. Through these three examples, the successful application of curvelets, enhanced with sparsity-promoting inversion, opens new perspectives on seismic data processing and imaging. The ability of curvelets to detect wavefront-like features is key to our success and opens an exciting new outlook towards future developments in exploration seismology.

ACKNOWLEDGMENTS

REFERENCES

Aminzadeh, F., J. Brac, and T. Kunz, 1997, 3-D Salt and Overthrust Model. Number 1 *in* SEG/EAGE 3-D Modeling Series: Society of Exploration Geophysicists, Tulsa.

Berkhout, A. J. and D. J. Verschuur, 2006, Focal transformation, an imaging concept for signal restoration and noise removal: Geophysics, **71**, A55–A59.

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

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

Candès, E. J., J. K. Romberg, and T. Tao, 2006b, Stable signal recovery from incomplete and inaccurate measurements: Communications on Pure and Applied Mathematics, **59**, 1207–1223.

Daubechies, I., M. Defrise, and C. de Mol, 2005, An iterative thresholding algorithm for linear inverse problems with a sparsity constraints: Communications on Pure and Applied Mathematics, 1413–1457.

Donoho, D. L., 2006, Compressed sensing: IEEE Transactions on Information Theory, **52**, 1289–1306.

Elad, M., J. L. Starck, P. Querre, and D. L. Donoho, 2005, Simultaneous Cartoon and Texture Image Inpainting using Morphological Component Analysis (MCA): Journal of Applied and Computational Harmonic Analysis, **19**, 340–358.

Fehmers, G. C. and C. F. W. Höcker, 2003, Fast structural interpretation with structure-oriented filtering: Geophysics, **68**, 1286–1293.

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

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

Hennenfent, G. and F. Herrmann, 2007a, Random sampling: new insights into the reconstruction of coarsely-sampled wavefields: Presented at the SEG International Exposition and 77th Annual Meeting.

Hennenfent, G. and F. J. Herrmann, 2006a, Application of stable signal recovery to seismic interpolation: Presented at the SEG International Exposition and 76th Annual Meeting.

——, 2006b, Seismic denoising with non-uniformly sampled curvelets: IEEE Computing in Science and Engineering, **8**, 16–25.

——, 2007b, Irregular sampling: from aliasing to noise: Presented at the EAGE 69th Conference & Exhibition.

Herrmann, F. J., 2007, Surface related multiple prediction from incomplete data: Presented at the EAGE 69th Conference & Exhibition.

Herrmann, F. J., U. Boeniger, and D. J. Verschuur, 2007a, Nonlinear primary-multiple separation with directional curvelet frames: Geophysical Journal International, **170**, 781–799.

Herrmann, F. J. and G. Hennenfent, 2007, Non-parametric seismic data recovery with curvelet frames. Technical report.

Herrmann, F. J., P. P. Moghaddam, and C. Stolk, 2007b, Sparsity- and continuity-promoting seismic imaging with curvelet frames: Journal of Applied and Computational Harmonic Analysis. Accepted for publication.

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

Lin, T. and F. J. Herrmann, 2007, Compressed wavefield extrapolation: Geophysics. Ac-

cepted for publication.

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

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

Sacchi, M., T. Ulrych, and C. Walker, 1998, Interpolation and extrapolation using a high-resolution discrete Fourier transform: IEEE Transactions on Signal Processing, **46**, 31–38.

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

Taylor, H. L., S. Banks, and J. McCoy, 1979, Deconvolution with the ℓ_1 norm: Geophysics, **44**, 39–52.

Verschuur, D. J. and A. J. Berkhout, 1997, Estimation of multiple scattering by iterative inversion, part II: practical aspects and examples: Geophysics, **62**, 1596–1611.

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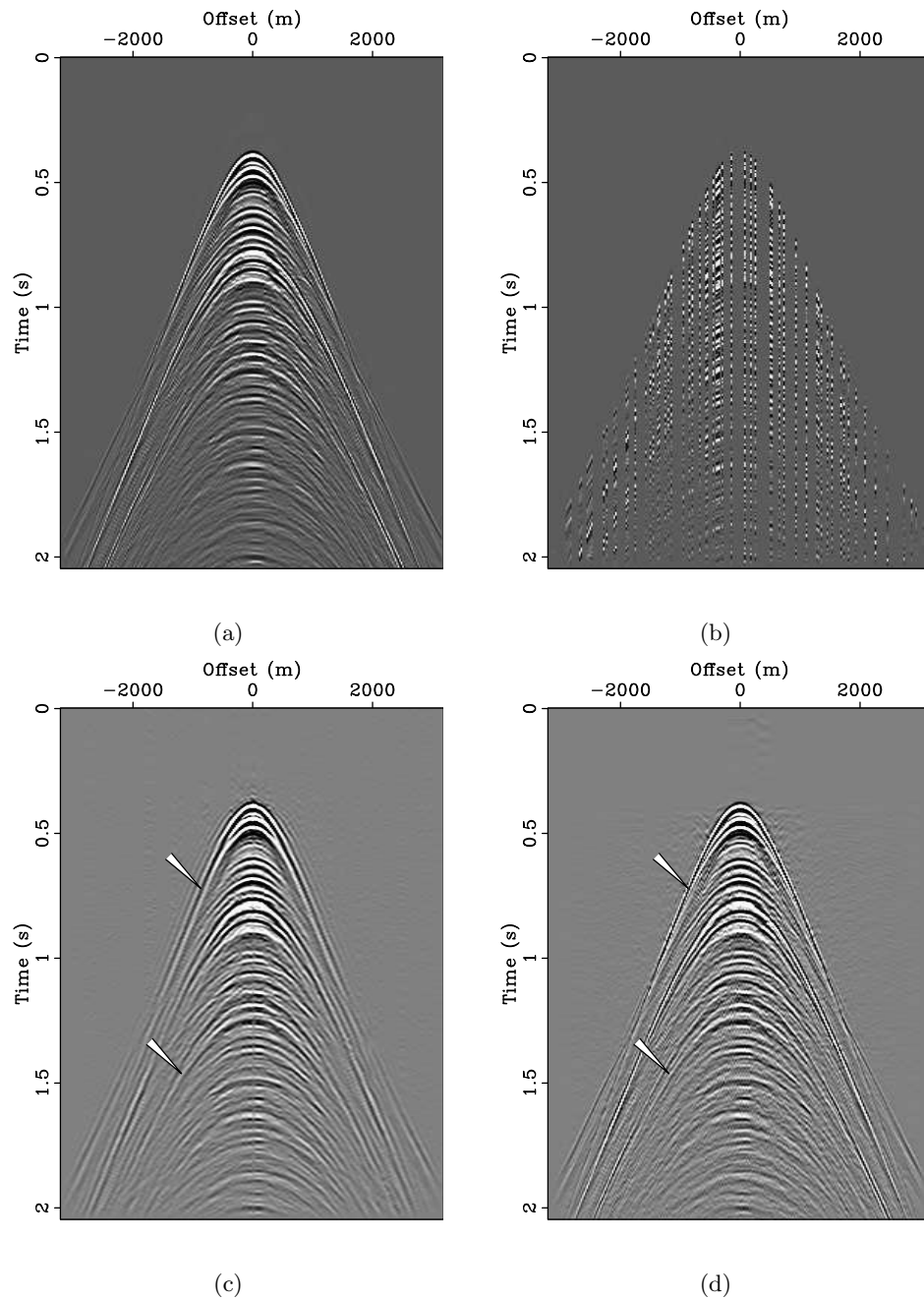
LIST OF FIGURES

Figure 1: Comparison between 3-D curvelet-based recovery by sparsity-promoting inversion with and without focusing. **(a)** Fully sampled real SAGA data shot gather. **(b)** Randomly subsampled shot gather from a 3-D data volume with 80 % of the traces missing in the receiver and shot directions. **(c)** Curvelet-based recovery. **(d)** Curvelet-based recovery with focusing. Notice the improvement (denoted by the arrows) from the focusing with the primary operator.
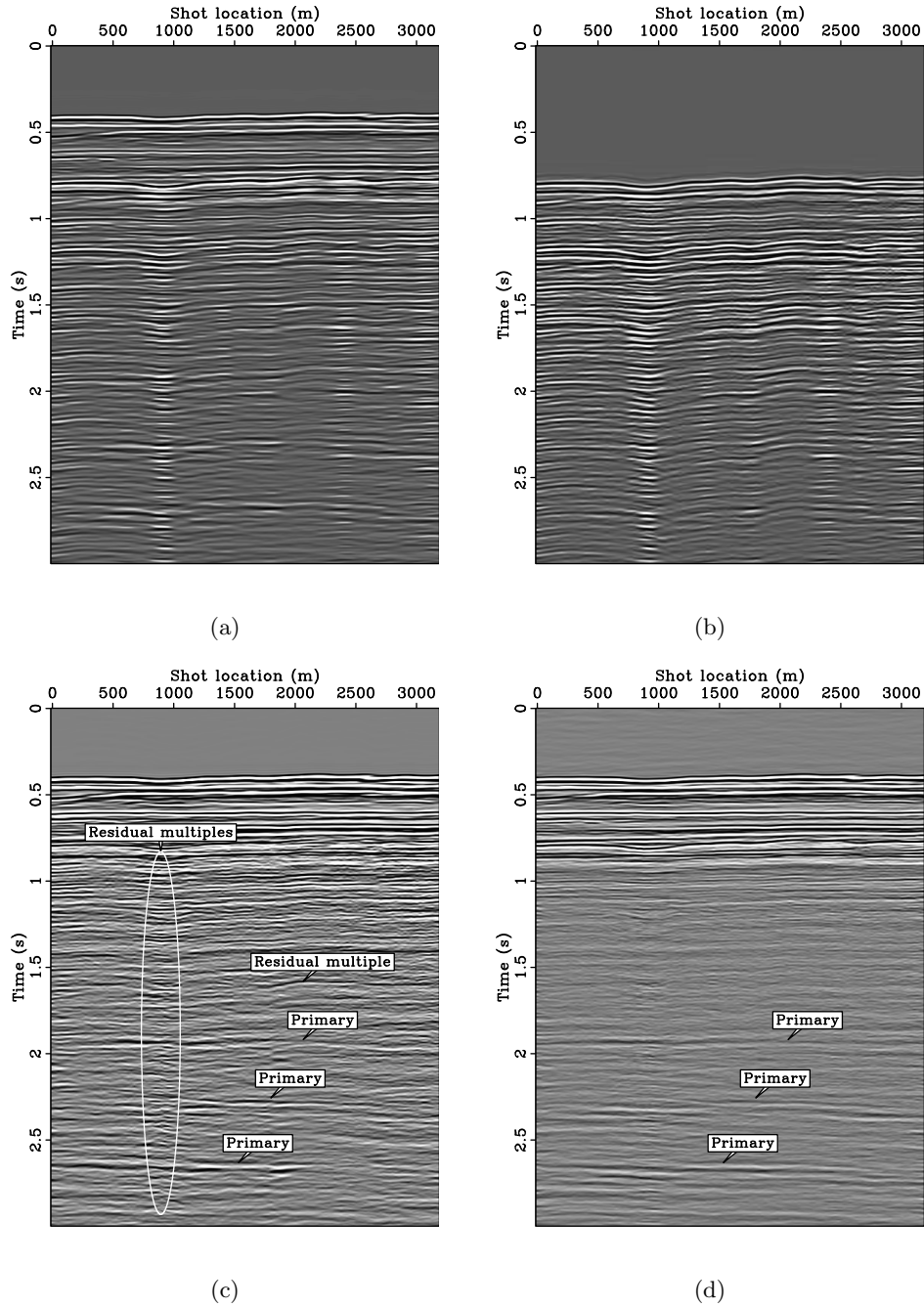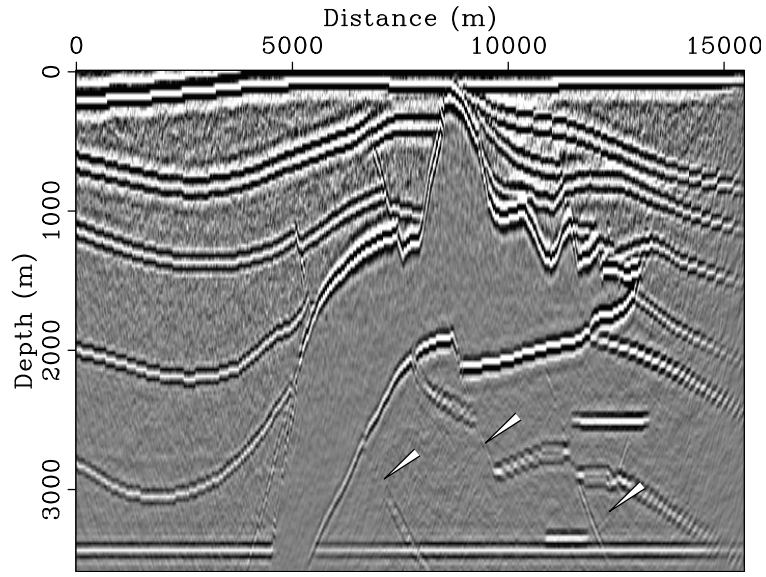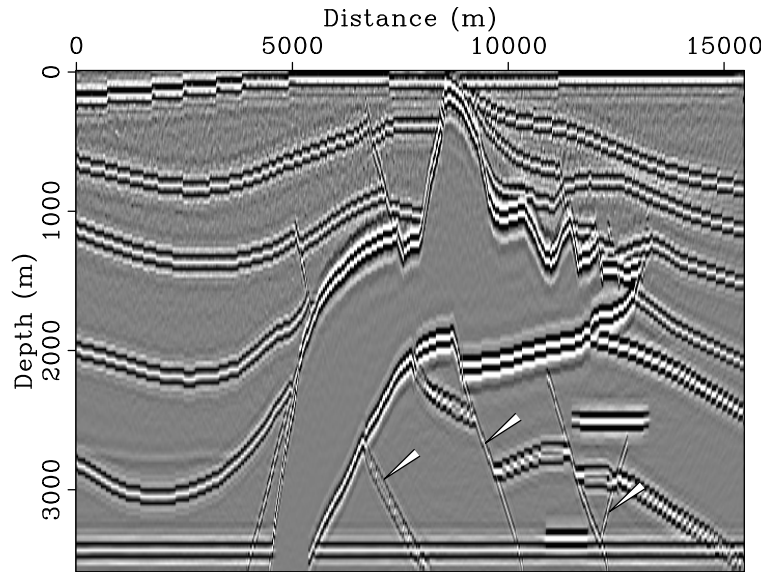
16

Figure 2: 3-D Primary-multiple separation with $\mathbf{P}_\epsilon$ for the SAGA dataset. **(a)** Near-offset section including multiples. **(b)** The SRME-predicted multiples. **(c)** The estimated primaries according to ℓ_2-matched filtering. **(d)** The estimated primaries obtained with $\mathbf{P}_\epsilon$. Notice the improvement, in areas with small 3-D effects (ellipsoid) and residual multiples.

Herrmann *et. al.* –

(a)

(b)

Figure 3: Image amplitude recovery for a migrated image calculated from noisy data (SNR 3 dB). **(a)** Image with clutter. **(b)** Image after nonlinear recovery. The clearly visible non-stationary noise in **(a)** is mostly removed during the recovery while the amplitudes are also restored. Steeply dipping reflectors (denoted by the arrows) under the salt are also well recovered.

Herrmann *et. al.* –