A FAST MODIFIED PARABOLIC RADON TRANSFORM

Brahim Abbad[1,2], Bjørn Ursin[2], and Milton J. Porsani[3]

18th May 2010

[1]Statoil, Drammensveien 264, Vækerø 0283, Oslo, Norway.

[2]Department of Petroleum Engineering and Applied Geophysics,

Norwegian University of Science and Technology,

S.P. Andersens vei 15A, NO-7491, Trondheim, Norway.

Email: abbad@ntnu.no, bjorn.ursin@ntnu.no,

[3]Institute of Geosciences, Federal University of Bahia (UFBA), Salvador, Brazil.

E-mail: porsani@cpgg.ufba.br

ABSTRACT

We present a fast and efficient frequency-domain implementation of a modified parabolic Radon transform (modified PRT) based on a singular value decomposition (SVD) with applications to multiple removal. The problem is transformed into a complex linear system involving a single operator after merging the curvature-frequency parameters into a new variable. A complex SVD is applied to this operator and the forward transform is computed by means of a complex back-substitution that is frequency-independent. The new transform offers a wider curvature range at signal frequencies than the other PRT implementations allowing the mapping in the transform domain of low-frequency events with important residual move-outs (long period multiples). The method is also capable of resolving multiple energy from primaries when they interfer in a small time interval, a situation where most frequency-domain methods fail to discriminate the different wave types. Additionally, the method resists better to AVO effects in the data than does the iteratively reweighted least-squares (IRLS) method. The proposed method was successfully applied to a deep-water seismic line in the Gulf of Mexico to attenuate water-bottom multiples and subsequent peg-legs originating from multiple paths in the water column. Combining the suggested method with the surface-related multiple elimination (SRME) has lead to the best attenuation results in removing residual multiple energy in the stack.

INTRODUCTION

The Parabolic Radon Transform (PRT) is recognized as a popular demultiple tool in seismic processing. The technique relies on the focusing properties of the primary and multiple reflection moveouts in the parabolic Radon domain. The transform can be applied either to common mid-point (CMP) data after stretching the time axis to implement the hyperbolic Radon transform (e.g. Yilmaz, 1989; Sacchi and Ulrych, 1995; Trad *el al.*, 2003), or to seismic data with residual moveouts after the data are corrected for normal moveout corrections (Hampson, 1986; Foster and Mosher, 1992). Hampson (1986) first approximated residual moveouts by parabolic curves and developed the frequency-domain PRT implemented by solving small-size complex-valued inverse problems. Data regularization (interpolation), to fill-in the missing offsets, has been investigated as another application of the transform (Sacchi and Ulrych, 1995; Trad *et al.*, 2002).

Additionally, the smearing problem is common in the standard inversion of the PRT due to limited offset coverage (aperture) and spatial sampling. This directly affects the focusing properties of the transform and degrades the filtering results. Yilmaz (1989), as well as Foster and Mosher (1992), used damped least-squares by means of SVD method to obtain what is known as the regular Radon transform without imposing any constraints in the solution. Methods for constrained inversion were later developed both

in the time (Thorson and Claerbout, 1985; Trad, 2001) and frequency domains (Sacchi and Ulrych, 1995; Cary, 1998; Herrmann *et al.*, 2000; Nowak and Imhof, 2004) to force the sparseness in the Radon domain using diagonal weighting matrices instead of the identity matrix in the least-squares system and yielded encouraging results for a number of applications. While the Radon transforms are often recast as inverse problems, a convolutional approach has been also suggested and the transform can be computed by means of deconvolving the operator that connects to the transform to provide the data (Zhou and Greenhalgh, 1994).

The moveout discrimination between the primaries and multiples in the transform domain is based on the focusing properties of both wave types in separable regions in the PRT domain. When the separation is good, the transform can be successfully applied to suppress the multiple energy and conserve the primaries. However, the presence of noise, amplitude variations with offsets (AVO), and the failure of the parabolic approximation to decompose seismic data properly, all affect the focusing properties in the transform and degrade the demultiple process and affects the reconstructed primaries. This results in multiple energy mapping into the primary energy in the Radon model and vice-versa making any filtering difficult. Amplitude variations with offset along the primaries are often lost after demultiple in the PRT domain.

Sacchi and Ulrych (1995) proposed a nonlinear algorithm with a sparseness constraint based on a minimum entropy along the curvature direction. The algorithm performs well and leads to a good event focus but fails when dealing with closely separated events in a small time window and when the data exhibits significant AVO effects.

Alternatively, others imposed the time-domain sparseness constraints along both curvature and intercept time directions (Cary, 1998; Schonewille and Aaron, 2007). Weighting matrices are built from the result of the regular transform, or from the obtained model at the previous iteration when the algorithm is iterative. Herrmann *et al.* (2000) solved for aliasing and aperture artifacts using a constrained minimum-norm under-determined problem with a weighting matrix constructed from the parabolic model obtained at low frequency components. The benefits of the high-resolution time-domain Radon transform over its frequency-domain counterpart have been reported on synthetic and real data examples by Cary (1998) and Schonewille and Aaron (2007). The gain in resolution is due to the fact that the least-squares Radon model has a higher sparseness in the time-domain than in the frequency domain, allowing for a better choice for the weighting matrix for use in the constrained inversion. However, the computing requirements represent a problem with the time-domain implementations especially on

high-fold marine gathers.

The frequency-domain formulation allows for time savings. Thus, the sparse large-size real inverse problem posed in the time domain is replaced by small-size complex systems to solve for all frequency components within the signal bandwidth in the data. Nevertheless, constraints for sparse inversion in the time-domain offer better focus and enhanced multiple removal than their counterparts implemented in the frequency domain. Examples showing this were recently reported by Schonewille and Aaron (2007). Several algorithms were proposed to solve the problem in the frequency domain to increase the resolution and enhance the filtering process.

We present a novel approach for implementing a modified PRT in the frequency domain that allows for faster inversion compared to the frequency-curvature domain PRT and better performance at low frequencies by offering a wider curvature range. Another advantage is computing efficiency where solving several constrained complex inverse problems are replaced by a similar number of inverse problems involving the same operator, thus allowing a significant gain in computing time for the transform. We include under frequency-curvature domain PRT all methods implemented in the frequency-curvature space to compute the transform using standard inversion (least squares) or constrained inversion (using weighting matrcies to

force the sparseness in the model). The new transform allows for a wider curvature scan at low frequencies without violating the aliasing conditions compared to frequency-curvature methods which have the same curvature range at all processed frequencies. This can be of great utility when dealing with low frequency events having important residual moveouts which can be mapped easily in the new transform domain, while these are beyond the curvature range allowed by the aliasing conditions in the frequency-curvature methods. In addition, the algorithm has a high performance in discriminating several events having comparable curvature parameters which cannot be resolved using most frequency-domain implementations of the PRT. The proposed method is tested against synthetic examples as well as a real deep-marine data set to attenuate water-bottom multiples and peg-legs with multiple paths in the water layer.

THE MODIFIED PARABOLIC RADON TRANSFORM

The inverse PRT is defined as a decomposition of seismic data in the curvature-intercept time domain using a set of curvature values $q_k, k = 1, N_q$ appropriately chosen to satisfy the requirements of sampling and aliasing. The PRT is defined such that the data are expressed as a sum of constant-amplitude reflections with parabolic moveouts. For discrete seismic data

$D(x_n, t)$ recorded at offset x_n and time t, this is defined by a sum along parabolic paths in the $q - \tau$ domain (Hampson, 1986)

$$D(x_n, t) = \sum_{j=1}^{N_q} M(q_j, \tau = t - q_j x_n^2) \qquad n = 1, \cdots, N_x, \tag{1}$$

where $M(q, \tau)$ represents the PRT at curvature q and intercept time τ, and N_q being the number of Radon curvatures while N_x denotes the number of seismic traces.

Using the linearity property of the Fourier transform, equation 1 can be expressed in the frequency domain as

$$d(x_n, f) = \sum_{j=1}^{N_q} m(q_j, f) e^{2\pi i f q_j x_n^2}, \tag{2}$$

where the temporal Fourier transform $d(f)$ for a function $D(t)$ is defined as

$$\begin{cases} d(f) = \int_{-\infty}^{\infty} D(t) e^{+2\pi i f t} dt \\ D(t) = \int_{-\infty}^{\infty} d(f) e^{-2\pi i f t} df \end{cases} . \tag{3}$$

For discrete data, equation 2 is a complex linear system for a given frequency f. In matrix notation, this can be written as

$$\mathbf{d}(f) = \mathbf{L}(f)\mathbf{m}(f), \tag{4}$$

where $\mathbf{d}(f)$ and $\mathbf{m}(f)$ are frequency-dependent vectors. The elements of the matrix $\mathbf{L}(f)$ have the form

$$L_{kj}(f) = e^{2\pi i f q_j x_k^2} \quad k = 1, \cdots, N_x; j = 1, \cdots, N_q. \tag{5}$$

The above formulation of the PRT requires solving the inverse problem in 4 for each frequency component in the signal bandwidth to compute the transform because the complex operator $\mathbf{L}(f)$ is frequency-dependent leading to a different matrix for each spectral component. This system can be solved using a variety of methods (Yilmaz, 1989; Sacchi and Ulrych, 1995; Cary, 1998; Sacchi and Porsani, 1999; Herrmann *et al.*, 2000).

To remove the frequency-dependence in the transform operator $\mathbf{L}(f)$, we introduce a new variable $\lambda = qf$ with unit m^{-2}. Writing 2 in terms of the variable λ leads to the following system of complex equations

$$d(x_n, f) = \sum_{j=1}^{N_q} m(\lambda_j, f) e^{2\pi i \lambda_j x_n^2}. \tag{6}$$

This can be written in a more compact form as

$$\mathbf{d}(f) = \mathbf{L}(\lambda)\mathbf{m}(f), \tag{7}$$

where the $N_x \times \mathrm{N}_\lambda$ complex matrix $\mathbf{L}(\lambda)$ is defined as

$$L_{kj}(\lambda) = e^{2\pi i \lambda_j x_k^2} \quad k = 1, \cdots, N_x; j = 1, \cdots, N_\lambda = N_q. \tag{8}$$

Equations 6 or 7 correspond to a modified PRT which can be computed much faster than the frequency-curvature domain PRT because the inverse of the matrix $\mathbf{L}(\lambda)$ has to be computed only once, given a data acquisition

geometry $[x_n, n = 1, \cdots, N_x]$ and a chosen discretization $[\lambda_j, j = 1, \cdots, N_\lambda]$.

The relations between the different **PRT** transforms are shown in Figure 1 where **FT** denotes the Fourier transform. Time-offset seismic data, $D(x,t)$, are first transformed to the frequency-offset domain, $d(x,f)$, where the inverse problem is formulated in the $\lambda - f$ domain to build the modified PRT $m(\lambda, f)$. Building the PRT panels along the curvature axis, $q - f$ or $q - t$ domains, requires interpolation from the modified PRT space.

SAMPLING AND ALIASING

For a specific seismic data set, the acquisition parameters, $[x_n, n = 1, \cdots, N_x]$, are normally given, and the PRT coordinates, $[\lambda_j, j = 1, \cdots, N_\lambda]$, can be chosen as long as the sampling requirements are fulfilled. Assuming regular sampling, Δx, in the offset domain and $\Delta\lambda$ in the modified PRT domain, we must have (Hugonnet and Canadas, 1995; Schonewille and Duijndam, 2001)

$$\Delta\lambda < \frac{1}{x_{max}^2 - x_{min}^2}, \tag{9}$$

where x_{max} and x_{min} are the maximum and minimum offsets in the data, respectively.

In equation 6, the phase difference between the two last expressions must

satisfy, for $\lambda > 0$,

$$2\pi\lambda x_{max}^2 - 2\pi\lambda\left[x_{max} - \Delta x\right]^2 < 2\pi. \tag{10}$$

This is the case when

$$\lambda < \frac{1}{2x_{max}\Delta x}. \tag{11}$$

For $\lambda < 0$, we get similarly

$$\lambda > \frac{-1}{2x_{max}\Delta x}, \tag{12}$$

so that

$$|\lambda| < \frac{1}{2x_{max}\Delta x}. \tag{13}$$

When the two bounds in inequality 9 and equation 13 are exactly satisfied we have, for $x_{min} = 0$,

$$N_\lambda = 2\frac{x_{max}^2}{2x_{max}\Delta x} = \frac{x_{max}}{\Delta x} \simeq N_x. \tag{14}$$

Thus the matrix $\mathbf{L}(\lambda)$ in equation 7 can, in principle, be inverted. It is, however, numerically ill-conditioned, and a stabilized solution (as shown later) must be used.

In standard seismic data, the curvature values are normally positive. In order to reduce the $\lambda-$ range we must partly NMO-correct the data, so there are reflections with negative and positive curvatures. Then by observing

the minimum moveout in the data, ΔT_{min} (which maybe negative), and the maximum moveout, ΔT_{max}, we can limit the range of $\lambda-$values to

$$\lambda_{min} = \frac{\Delta T_{min} f_{max}}{x_{max}^2} < \lambda < \frac{\Delta T_{max} f_{max}}{x_{max}^2} = \lambda_{max}, \tag{15}$$

where f_{max} is the maximum signal frequency in the data. λ_{min} and λ_{max} must, of course, satisfy inequality 13 above.

For a high resolution modified PRT, we choose a relatively small sampling interval, $\Delta\lambda$, satisfying inequality 9 (Herrmann *et al.*, 2000; Schonewille and Aaron, 2007). The number of values in the PRT domain,

$$N_\lambda = \frac{\lambda_{max} - \lambda_{min}}{\Delta\lambda}, \tag{16}$$

is then normally larger than the number of seismic channels N_x.

A vertical line in the $q - f$ domain corresponds to an event with a parabolic moveout parameter ($q = constant$) in the time-offset domain. In the $\lambda - f$ domain this gives a radial line defined by the equation $\lambda = qf$. Keeping this in mind, all operations in the $q - f$ domain can also be done in the $\lambda - f$ domain. We also note that from $\lambda = qf$, the range in $q-$ values for low frequencies is larger than for high frequencies. This means that the modified PRT has better curvature coverage at low frequencies than the frequency-curvature domain PRT. Low frequency events with important parabolic moveouts may violate aliasing conditions in the frequency-

curvature domain PRT, whereas these can be appropriately mapped in the $\lambda - f$ domain thanks to a wide curvature range allowed by the transform (Figure 2b). The standard and modified PRT's have the same curvature range only at the maximum signal frequency f_{max}. For smaller frequencies, the modified PRT has a wider curvature range providing better coverage in the curvature domain at most signal frequencies (Figure 2). The red rectangle in Figure 2a represents the covered curvature zone in the frequency-curvature domain PRT according to the sampling rules (Turner, 1991; Schonewille and Duijndam, 2001). The larger zone shows the curvature range covered by the sampling relations for the modified PRT method expressed by equation 13. This shows a curvature range that is inversely proportial to the processed frequency, thus allowing the mapping of low frequency events in the $\lambda - f$ domain even for important curvature (Figure 2b). The red area in this figure shows the equivalent $\lambda-$ space highlighted according to the sampling rules in the standard PRT. In fact, the two red regions in the $q - f$ and the $\lambda - f$ domains are equivalent and show the limited covergae in the standard PRT at most signal frequencies. Correspondingly, the $\lambda-$sampling interval, $\Delta\lambda$, is inversely proportional to the processed frequency, thus keeping the same number of $\lambda-$values for each frequency. This does not affect the performance of the modified PRT, but artefacts will appear at low frequencies when computing the standard transforms, either $q - f$ or $q - t$ domains, is required. In fact, only a limited $\lambda-$range is used to recover the correspond-

ing $q-f$ domain in response to the aliasing conditions. The benefits from this additional curvature range are important when multiples with large residual moveouts are input for removal. In such a case, the standard PRT may not map all the events in the frequency-curvature domain, while this can be handled by the modified PRT. Examples include multiple reflections with several paths in the water layer for deep water seismic surveys which exhibit large residual moveouts compared to primaries flattened along the offset axis. For most cases, the new domain, $\lambda - f$, allows to perform demultiple based on moveout discrimination between primaries and multiples as well as data interpolation avoiding the need to the abovementioned interpolation (see the flowchart in Figure 1). This interpolation is required if separation of parabolic events in a seismic gather requires a dynamic $q-f$ or $q-t$ rejection zone that cannot be implemented directly in the $\lambda-f$ space.

The least-squares solution

We want to solve the linear problem in equation 7 when the number of seismic channels, N_x, is larger than the number of values in the $\lambda-$domain, N_λ. Then the problem is overdetermined and the least-squares solution can be written as

$$\mathbf{m_{LS}}(f) = \left[\mathbf{L^H}(\lambda)\mathbf{L}(\lambda)\right]^{-1}\mathbf{L^H}(\lambda)\mathbf{d}(f), \tag{17}$$

where $\mathbf{H}$ denotes the complex conjugate transpose (the adjoint operator). It is well known that computing this expression as written may give numerical

problems due to the large condition number of the matrix $\mathbf{L^H}(\lambda)\mathbf{L}(\lambda)$.

Instead we use the singular value decomposition (SVD) (Björck, 1996)

$$\mathbf{L}(\lambda) = \mathbf{U} \begin{bmatrix} \mathbf{\Sigma} \\ \\ 0 \end{bmatrix} \mathbf{V^H}, \tag{18}$$

where $\mathbf{U}$ and $\mathbf{V}$ are complex unitary matrices and $\mathbf{\Sigma} = diag[\sigma_1, \sigma_2, \cdots, \sigma_{N_\lambda}]$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{N_\lambda} \geq 0$.

A stable least-squares estimate is given by

$$\mathbf{m_{LS}}(f) = \mathbf{V} \begin{bmatrix} \mathbf{\Sigma}^\dagger & \mathbf{0} \end{bmatrix} \mathbf{U^H d}(f), \tag{19}$$

where $\mathbf{\Sigma}^\dagger$ is the pseudo-inverse of $\mathbf{\Sigma}$ with components

$$\sigma_k^\dagger = \begin{cases} \sigma_k^{-1} & \text{if } \sigma_k \geq \epsilon \\ \\ 0 & \text{if } \sigma_k < \epsilon \end{cases}, \quad k = 1, \cdots, N_\lambda. \tag{20}$$

Truncated SVD is used to exclude small singular values while computing the pseudo-inverse. The numerical rank is often greater than the mathematical rank of $\mathbf{L}(\lambda)$ due to rounding errors during computations. We define the mathematical rank of a given matrix as the number of rows or columns linearly independent. Using the decomposition 18, the mathematical rank denotes the number of non-zero singular values of the matrix. The null singular values (numerically very small) reside in the null-space domain of the matrix. Equation 20 assignes a numerical rank to $\mathbf{\Sigma}$ and hence to $\mathbf{L}(\lambda)$ by

moving the small singular values into the null space of $\mathbf{L}(\lambda)$ (Björck, 1996).

Alternatively, one may use a damped least-squares stabilization with

$$\sigma_k^\dagger = \frac{\sigma_k}{\sigma_k^2 + \epsilon}, \quad k = 1, \cdots, N_\lambda, \tag{21}$$

where ϵ now is a damping parameter. This stabilization reduces the condition number from $\sigma_1/\sigma_{N_\lambda}$ to a smaller value.

With $\tilde{\mathbf{m}}_{\mathbf{LS}}(f) = \mathbf{V}^{\mathbf{H}}\mathbf{m}_{\mathbf{LS}}(f)$ and $\tilde{\mathbf{d}}(f) = \mathbf{U}^{\mathbf{H}}\mathbf{d}(f)$, equation 19 gives

$$\tilde{\mathbf{m}}_{\mathbf{LS},k}(f) = \begin{cases} \sigma_k^\dagger \tilde{\mathbf{d}}_k(f) & k = 1, \cdots, N_\lambda \\ 0 & k = N_\lambda + 1, \cdots, N_x \end{cases}. \tag{22}$$

It is seen that the components $[\tilde{\mathbf{d}}_k, k = N_{\lambda+1}, \cdots, N_x]$ do not contribute to the solution.

The minimum-norm solution

In a high-resolution modified PRT, we choose the number of $\lambda-$values, N_λ, to be larger than the number of seismic channels. Then the problem is underdetermined and we choose the minimum-norm solution (Björck, 1996) which can be written as

$$\mathbf{m}_{\mathbf{MN}}(f) = \mathbf{L}^{\mathbf{H}}(\lambda)\left[\mathbf{L}(\lambda)\mathbf{L}^{\mathbf{H}}(\lambda)\right]^{-1}\mathbf{d}(f). \tag{23}$$

We use the SVD decomposition for the operator $\mathbf{L}(\lambda)$

$$\mathbf{L}(\lambda) = \mathbf{U}[\mathbf{\Sigma} \quad \mathbf{0}]\mathbf{V}^{\mathbf{H}}, \tag{24}$$

where $\mathbf{\Sigma} = diag[\sigma_1, \sigma_2, \cdots, \sigma_{N_x}]$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{N_x} \geq 0$, to obtain a stable estimate

$$\mathbf{m_{MN}}(f) = \mathbf{V} \begin{bmatrix} \mathbf{\Sigma}^{\dagger} \\ 0 \end{bmatrix} \mathbf{U}^{\mathbf{H}}\mathbf{d}(f), \tag{25}$$

where $\mathbf{\Sigma}^{\dagger}$ is the pseudo-inverse of $\mathbf{\Sigma}$ as defined in equation 20 or, alternatively, in equation 21.

The transformations $\tilde{\mathbf{m}}_{\mathbf{MN}}(f) = \mathbf{V}^{\mathbf{H}}\mathbf{m_{MN}}(f)$ and $\tilde{\mathbf{d}}(f) = \mathbf{U}^{\mathbf{H}}\mathbf{d}(f)$ gives

$$\tilde{\mathbf{m}}_{\mathbf{MN},k}(f) = \begin{cases} \sigma_k^{\dagger}\tilde{\mathbf{d}}_k(f) & k = 1, \cdots, N_x \\ 0 & k = N_x + 1, \cdots, N_\lambda \end{cases}. \tag{26}$$

We see that the last components of $\tilde{\mathbf{m}}_{\mathbf{MN}}$ are zero, corresponding to the minimum-norm properties of the solution $\mathbf{m_{MN}}$.

MULTIPLE ATTENUATION AND DATA REGULARIZATION

We may attenuate multiple reflections in the $\lambda - f$ domain by partly correcting the seismic data for normal moveout effects such that the primary reflections appear with less moveout than multiples. The main assumption

on seismic data input to the transform is that residual moveouts of both primary and multiple reflections have parabolic behavior as a function of offset (Hampson, 1986). The data are transformed to the $\lambda - f$ domain using an inverse operator

$$\hat{\mathbf{m}}(f) = \mathbf{L}^{\dagger}(\lambda)\mathbf{d}(f), \tag{27}$$

defined in equation 19 or 25. An estimate of the primary reflections **is** obtained by filtering the data $\hat{\mathbf{m}}_{\mathbf{P}}(f) = \mathbf{F}(\lambda, f)\hat{\mathbf{m}}(f)$ where the filter is

$$\mathbf{F}(\lambda, f) = \begin{cases} 1 & \text{for } \lambda \leq q_0 f \\ \\ 0 & \text{for } \lambda > q_0 f \end{cases}, \tag{28}$$

where q_0 is a curvature parameter that separates primaries and multiples in the $\lambda - f$ space. The primary reflections are then obtained by transforming back to the $x - f$ domain

$$\mathbf{d}_{\mathbf{P}}(f) = \mathbf{L}(\lambda)\mathbf{F}(\lambda, f)\mathbf{L}^{\dagger}(\lambda)\mathbf{d}(f) = \mathbf{G}(\lambda, f)\mathbf{d}(f). \tag{29}$$

The matrix $\mathbf{G}(\lambda, f) = \mathbf{L}(\lambda)\mathbf{F}(\lambda, f)\mathbf{L}^{\dagger}(\lambda)$ has dimensions $N_x \times N_x$. It behaves as a filter applied to seismic data in the frequency-offset domain to remove multiples directly from the data. The multiples can be also estimated using the following relationship

$$\mathbf{d}_{\mathbf{M}}(f) = \left[\mathbf{I}_{N_x \times N_x} - \mathbf{G}(\lambda, f)\right] \mathbf{d}(f). \tag{30}$$

where $\mathbf{I}_{N_x \times N_x}$ is the identity matrix.

For a given frequency, the operator $\mathbf{G}$ depends only on the data acquisition parameters, our choice of parameters in the $\lambda-$ domain, and the choice of parameters for the pseudo-inverse.

Radon transforms are an attractive domain for data interpolation to fill in the missing offsets or to output a data set with a regular offset sampling from an input gather where acquisition conditions didn't allow to achieve such a purpose (Trad *et al.*, 2002). Regularization is also required when data are input to some processing programs such as the frequency-wavenumber filtering, some migration algorithms, and surface-related multiple elimination (SRME). This can be done by the modified PRT from

$$\bar{\mathbf{d}}(f) = \bar{\mathbf{L}}(\lambda)\hat{\mathbf{m}}(f) = \bar{\mathbf{L}}(\lambda)\mathbf{L}^{\dagger}(\lambda)\mathbf{d}(f), \tag{31}$$

where $\bar{\mathbf{L}}(\lambda)$ is defined as in equation 8, but the points $[x_n, n = 1, \cdots, \bar{N_x}]$ are the desired offsets after interpolation. The points $[\lambda_j, j = 1, \cdots, N_\lambda]$ are the same in $\bar{\mathbf{L}}(\lambda)$ as in $\mathbf{L}(\lambda)$. $\bar{N_x}$ being the number of channels in the interpolated data set. We may even combine the two algorithms to attenuate multiples and interpolate data at the same time

$$\bar{\mathbf{d_P}}(f) = \bar{\mathbf{L}}(\lambda)\mathbf{F}(\lambda, f)\mathbf{L}^{\dagger}(\lambda)\mathbf{d}(f). \tag{32}$$

In all these cases there is, for each frequency, a multiplication by a fixed matrix which transforms data from the $x - f$ domain to the $\lambda - f$ domain.

NUMERICAL RESULTS

Synthetic data

1. Constant-amplitude events

To test the efficiency of the novel approach over popular methods, we investigated a complex synthetic data example that consists of a large number of events (20 parabolas) interfering in a narrow time window as illustrated in Figure 3a. This model contains 12 primaries (Figure 3b) and 8 multiples (Figure 3c). The primaries are characterized by negative curvature values whereas the multiples have positive curvatures. This situation can be produced by applying normal moveout corrections using a velocity law between the stacking velocities of the primaries and those of the multiples. There are $N_x = 100$ seismic channels with offset spacing $\Delta x = 0.02km$ and a maximum offset $x_{max} = 2.0km$. From equations 9 and 13 we obtain the bounds $\Delta\lambda < 0.25km^{-2}$ and $|\lambda| < 12.5km^{-2}$.

The moveout at the maximum offset used to compute the transform is in the range $\Delta T \in [-0.3s, 0.3s]$. With a maximum signal frequency $f_{max} = 60Hz$, this gives from equation 15 $\lambda \in [-4.5km^{-2}, 4.5km^{-2}]$ which is within the maximum value computed above. In Figure 4, we have computed the singular values for the matrix $\mathbf{L}(\lambda)$ for different values of N_λ (or equivalently $\Delta\lambda$) with a fixed range of λ.

For $N_\lambda = 50$ (the curve in black), the inverse problem in 7 is overdetermined ($N_x > N_\lambda$) and the condition number is small ($\sigma_1/\sigma_{50} = 3.11,$). In this case, the least-squares inverse operator $\mathbf{L^{-1}}(\lambda)$ can be computed, but the transform is not able to map appropriately the primary and multiple modes for accurate filtering due to poor sampling along the λ -axis. For $N_\lambda = N_x = 100$ (the curve in grey), the operator $\mathbf{L}(\lambda)$ is a square matrix, but the condition number is large ($\sigma_1/\sigma_{100} = 1.34 \times 10^9$). In this case, damping or truncation is required to estimate a pseudo-inverse for the operator $\mathbf{L}(\lambda)$ and compute the transform. The curve for a finer sampling rate, $N_\lambda = 250$, is shown in the same figure in dashed line and corresponds to an under-determined problem. The general tendency for the three curves is that the decrease in singular values of $\mathbf{L}(\lambda)$ is slow up to about half the number of seismic traces before a rapid drop to vanishing singular values [$\sigma_k \simeq 0$ for $k = N_x/2 + 1, \cdots, min(N_x, N_\lambda)$]. We used the minimum-norm solution which converts the problem into an inverse problem of dimension $N_x \times N_x$ as provided by equation 25. The SVD of the minimum-norm operator gives an infinite condition number for $N_\lambda = 250$. To compensate for this, the pseudo-inverse in equation 20 was used with $\epsilon = 10^{-3}$. The complex operator $\mathbf{L}(\lambda)$ presents the property of rank defficiency regardless the size of the operator, the range of curvatures and offsets. About half of the singular values of $\mathbf{L}(\lambda)$ are mathematically zero (numerically very small) and do not contribute into the solution as they belong to the null-space of

$\mathbf{L}(\lambda)$. When $N_\lambda < N_x/2$, the operator $\mathbf{L}(\lambda)$ has a small condition number and no truncation is required to compute the transform, but the expected resolution in the $\lambda - f$ domain is not achieved leading to poor demultiple and data interpolation results. The parameter ϵ in equation 20 defines the limit between singular values used to compute the new transform and those belonging to the null space and assumed to be mathematically zero. The presence of the null space explains a linear dependence between rows and columns of $\mathbf{L}(\lambda)$ making the computation of the transform an ill-conditioned inversion problem when $N_\lambda > N_x/2$.

The mapping of the whole gather in Figures 3a in the $\lambda - f$ domain is shown in Figures 3d. Primary and multiple energy are mapped in distinct areas of the $\lambda - f$ domain. Each event with a parabolic moveout in the time-offset domain is transformed along a radial line in the $\lambda - f$ domain passing through the origin and characterized by the respective event curvature as the slope value and with a frequency content shaping the wavelet spectrum. According to this, events with positive curvatures ($q > 0$) are mapped into the positive λ−space while negative λ−space contains events with $q < 0$. Events in the time-offset domain having the same curvature value overlap within their spectra and their transforms add constructively in the $\lambda - f$ space.

The filtering is applied following a constant q-value to separate the events in the original gather according to their curvatures as provided by the relation 28. Since primaries have negative curvatures in the example, a mute zone was established at $q = 0$ (corresponding also to $\lambda = 0$) to separate primaries and multiples, as shown in Figures 3e and 3f, respectively. The reconstructed gathers after inverse transform are shown in Figures 3g, 3h, and 3i, they show the events in the original gather with their true curvature and amplitude characteristics. The reconstruction errors for the three gathers are depicted in Figures 3j, 3k, and 3l. This error is very small for the whole gather and is mainly due to the use of truncation in estimating the pseudo-inverse of $\mathbf{L}(\lambda)$. All PRT methods suffer such residual errors after reconstruction due to the ill-conditioning of the inverse problem. The residual errors from separating primaries and multiples show opposite signs on the gathers. This occurs mainly along the axis $\lambda = 0$ thus generating horizontal low-amplitude artifacts in the filtered gathers.

2. Events exhibiting AVO effects

To examine the relative merits of the novel approach when the data exhibit AVO effects, we included amplitude variations with offset in the example of Figure 3. The introduced amplitude functions are second-order functions in normalized offset given by

$$A(y) = a_0 + a_1 y + a_2 y^2; \quad y = \frac{|x|}{x_{max}}, \tag{33}$$

where y is the normalized offset. a_0, a_1, and a_2 are real coefficients that describe the amplitude functions along the events. The parameter a_0 represents the intercept of the AVO function, while a_1 and a_2 denote the amplitude gradient and curvature in the offset domain, respectively.

Figures 5a, 5b, and 5c depict the gathers in Figures 3a, 3b, and 3c after AVO effects were included along the parabolic moveouts on individual events according to equation 33. The amplitude effects on the 12 primary events (in Figure 5b) are plotted in Figure 6 (curve in black). The amplitude plots of the events are ordered with respect to increasing zero-offset time of the primaries, row by row and column by column in Figure 6. These show significant amplitude changes with normalized offset and presenting several zero-crossings. The event interference between primaries and multiples and between primaries alone leads to amplitude distorsion on individual events in the gather. Thus, the amplitudes of the events being interferred add to each other within the crossing area, hence altering the AVO curve and making the amplitude inversion inaccurate. The green curves in Figure 6 show the practical amplitude functions for each event of the primaries that accounts for the interference with the other primaries. Outside the interference zone, the theoretical and practical amplitude curves are equal.

The results of applying the modified PRT algorithm to the gather in-

cluding such AVO effects are given in Figure 5 and the recovered amplitude functions of the primaries after demultiple are also drawn in Figure 6 (curves in blue). Figure 5d shows the $\lambda - f$ mapping of the gather in Figure 5a. With a mute zone applied at $\lambda = 0$, the primary and multiple mappings in the modified PRT domain are plotted in Figures 5e and 5f, respectively. The reconstructed gathers, depicted in Figures 5g, 5h, and 5i, preserve the amplitude variations, showing that this method is an alternative to other frequency-domain algorithms requiring removal of the AVO effects (using sliding gain functions for instance) before the transform is applied. The residual errors after reconstruction are very small for the entire gather (Figure 5j) and show again the low-amplitude horizontal artifacts (Figures 5k and 5l) similar also to the first example (constant-amplitude events).

Schonewille and Zwartjes (2001) proposed a sparseness-constrained frequency-domain algorithm resisting well to AVO effects using the iteratively reweighted least-squares (IRLS) algorithm combined with a complex version of the LSQR solver. The approach proposed by Schonewille and Zwartjes (2002) is explained in the Appendix. Both algorithms give comparable results, but the modified PRT algorithm has the advantage to be much faster.

The recovered amplitude curves are shown by the red curve in Figure 6 for the 12 events after a single iteration. The recovered amplitudes are very

accurate at most offsets. At medium offsets, $0.2 \leq y \leq 0.9$, the suggested approach allows for more accurate conservation of the AVO curve than does the IRLS method. At near offsets ($y < 0.2$), the IRLS method has generally smaller amplitude errors but the error is larger than the modified PRT for some primaries (events 6 and 7 for instance). The same observation can be made at very far offsets ($y > 0.9$) where the IRLS method performs slightly better. The proposed method proves higher accuracy in amplitude recovery than the IRLS approach at most offsets. This observation is further confirmed on Figure 7 which depicts the average absolute amplitude error for the 12 events in Figure 5b. The superiority in performance of our approach (blue curve) is again demonstrated except at the extreme offsets in the data. The black and green curves in Figure 7 show the absolute errors of the AVO function using 2 and 5 iterations in the IRLS approach, respectively. With further iterations in the IRLS method, increased accuracy is achieved at most offsets. However, a loss of accuracy can also be observed at intermediate offset ranges when comparing results after 2 and 5 iterations. There is a significant reduction of the errors associated with most offsets than the first iteration (red curve) but these errors still larger than the modified PRT errors at most offsets which has the added advantage to be very fast.

Figure 8 illustrates the method when the data exhibit AVO effects and contain random noise. The results show that the modified PRT method

stills perform well in such a complicated model. In the $\lambda - f$ domain, the recognition of the radial lines mapping the events is still possible. The filtered primaries exhibit AVO responses approximating those in the original model, and the reconstruction error has only residual noise and the artifact with a zero curvature (at $q = 0$). As the residual noise resides in the small singular values of the operator $\mathbf{L}(\lambda)$, this is altered after reconstruction either by damping or truncation. This explains why the reconstruction error of the gather in Figure 8a is dominated by noise (Figure 8j).

The Mississipi Canyon seismic line

The modified PRT method was tested on a marine seismic line from the Gulf of Mexico. The area is characterized by a deep sea-bottom with important amplitude variations with offset and water-bottom multiples. The data was tested by several demultiple programs due to the complexity of the model (Dragoset, 1999; Guitton and Cambois, 1999; Hadidi et al., 1999; Lamont *et al.*, 1999; Lokshtanov, 1999; Verschuur and Prein, 1999; Trad *et al.*, 2003). Figure 9b shows the gather at location CMP 500 and illustrates the strong interference in the area and the large number of events recorded. Nonhyperbolic velocity analysis was performed on the data to generate post-NMO data to be input for the modified PRT algorithm. Due to the thick water layer, ($1.4km$ to$1.5km$), and the thin layers following, moveout velocities are

very low (mostly ranging from 1500 to 2000 m/s). A velocity spectrum computed using the semblance coefficient is given in Figure 9d for this gather. The decrease of stacking velocities after $4s$ is an obvious indicator of the presence of water-bottom multiple reflections and peg-legs originating from the layers just below the sea-bed. No primary energy can be recognized on the velocity spectrum at late times ($> 4s$) due to the presence of multiple reflections masking any primaries from a deep salt body in the model. To flatten primary reflections at far offsets, we scanned for both the moveout velocity and the effective anellipticity parameters to obtain more accurate NMO corrections. AVO variations were accounted for in nonhyperbolic velocity analysis using sliding gain functions which allowed the picking of the reflected energy in the gathers. Figure 9c shows the gather after applying nonhyperbolic NMO corrections with parameters derived from automatic nonhyperbolic velocity analysis (Abbad *et al.*, 2009). The primary energy corrected with the exact moveout parameters (moveout velocity and effective anellipticity) is well flattened in the full offset range, whereas multiple energy has positive residual moveouts that can be well approximated using parabolic curves as observed in the lower part of the gather after $4s$ (Figure 9c). The stretched area around time $2.7s$ in the NMO-corrected gather is due to the strong interference in the original gather at far offsets which can be observed also at a larger time on the gather before moveout corrections depicted in Figure 9b.

The data acquisition parameters are described in detail in Verschuur and Prein (1999). The offsets range from $0.1km$ to $4.874km$ with a maximum CMP fold equal to $N_x = 90$. The data were processed in the time range $[1.6s - 7.0s]$, limited to the region below the sea bottom where reflections are recorded. The modified PRT was applied to the NMO-corrected data in Figure 9c with a moveout range $\Delta T \in [-0.05s, 0.7s]$ at the maximum data offset. Negative moveout values were included to account for primaries possibly overcorrected at far offsets. Using a maximum signal frequency $f_{max} = 60Hz$, this gives $\lambda \in [-0.127km^{-2}, 1.79km^{-2}]$. The last value is just below the limit $|\lambda| < 1.89km^{-2}$ according to equation 13. Figure 10 shows the distribution of singular values of the matrix $\mathbf{L}(\lambda)$ in equation 8 for different values $N_\lambda = 45, 90$ and 225. For $N_\lambda = 45$ the range of variations in the singular values is small, but for $N_\lambda = 90$ and $N_\lambda = 225$ the singular values become very small after the 45 first singular values. Again, there is a sharp drop in the singular values around $N_x/2$ $[(\sigma_k \simeq 0$ for $k = N_x/2 + 1, \cdots, N_\lambda)]$. We use $N_\lambda = 225$ and a minimum-norm solution as described in equation 25 with a cut-off value $\epsilon = 5 \times 10^{-2}$ in equation 20 to compute the modified PRT. This gives the transformed data shown in Figure 9a. The multiple reflections are filtered out with a cut-off line defined by $\lambda > q_0 f$ chosen such that

$$\lambda > \frac{\Delta T_0 f}{x_{max}^2}, \tag{34}$$

where $\Delta T_0 = 0.1s$ is the cut-off value for the moveout at the maximum offset, x_{max}. This gives the transformed primaries and multiples shown in Figures 9e and 9i, respectively.

The data gathers corresponding to these domains are computed through equation 7 and are illustrated in Figures 9g and 9k. Inverse nonhyperbolic moveout corrections are applied to reconstruct the original gathers filtered into primary (Figure 9f) and multiple (Figure 9j) reflections. These gathers show that multiples are concentrated in the lower part of the gather due to multiple paths in the water-column and peg-legs in subsequent thin layers below the sea-bed. The velocity spectra for the filtered primaries and multiples are respectively depicted in Figures 9h and 9l. After demultiple, significant primary energy is recovered after $4s$ and can be clearly recognized in the primary gather (Figure 9f) and also in the corresponding velocity spectrum (Figure 9h).

Figure 11 shows the stack obtained from automatic nonhyperbolic velocity analysis without any demultiple. The stack illustrates the main features of the area mainly a deep salt body at around recording time $5s$. The lower part of the stack is dominated by important multiple reflections after $3.5s$ with multiple paths in the water layer and peg-legs in the layers beneath the sea-bottom. The application of the surface-related multiple elimination

(SRME) algorithm to this seismic line is illustrated in Figure 12a (from Verschuur and Prein, 1999) with the estimated multiples in Figure 12b. It is obvious that the SRME method removes multiples after the first water-layer multiple and doesn't attenuate any multiples arriving before like internal multiples. As this method requires data regularization, the PRT method was used as part of the SRME method to fill in the near missing offsets in the data which increases the implementation time of the method. The obtained stack after demultiple via the SRME shows improved quality in the lower part of the section with most multiple energy being removed. However, residual multiples can be clearly observed between 4 and $5s$.

The application of the modified PRT method to this data is illustrated in Figure 13. The demultiple result in Figure 13a shows efficient removal of the associated multiple energy in the data. In the upper part of the section, the result is better than the SRME method, because the method allowed the removal of internal multiples coming after the reflection from the water bottom (Figure 13b) which enhanced the continuity of the shallow reflectors in the section. However, residual multiples still observable around time $4.5s$ in the stack. The modified PRT was applied with the parameters derived for the single CMP gather shown in Figure 9. As suggested by Hadidi *et al.* (1999), the modified PRT method was also applied to the data after the SRME method (as shown in Figure 12). The resulting stack shown in

Figure 14 is clearly the best in terms of removing multiple reflections in the data.

COMPUTATIONAL COST

Both the modified PRT and the constrained PRT based on the IRLS method work in the $x - f$ domain. This requires $2N_x$ FFT's of length N_f $(= 2^k \geq N_t, k:$ integer) for both methods. For multiple attenuation with the modified PRT as given in equation 29 there will be additional N_{f_0} (number of discrete frequencies) multiplications with the $N_x \times N_x$ complex matrices $\mathbf{G}(\lambda, f)$. For the IRLS approach there will additionally be N_{f_0} complex inversions times the number of iterations of the system given in (A-3) which increases significantly the computing time since each system involves a different matrix for any processed signal frequency. The computing times of the two algorithms depend mainly on the number of processed frequencies and on the CMP fold.

Table 1 summarizes the CPU timing for the two methods according to tests on a Pentium IV platfrom. The reported times confirm the considerable gain in computing time in the modified PRT algorithm. The IRLS method has larger computing requirements becoming heavy if several iterations are necessary. This is the case when the amplitude functions need to be conserved in view of a subsequent AVO inversion. For the whole line, the demultiple process requires about $51min$ for the modified PRT and almost 6 hours for a single iteration of IRLS if amplitude conservation is a relevant issue of

the processing sequence. The timing results show that the modified PRT is about six to eight times faster than the IRLS method.

CONCLUSIONS

A fast implementation of a modified PRT has been proposed. The new transform is linked to the frequency-curvature domain PRT by means of interpolation and serves as a fast and efficient tool for multiple removal and data interpolation. The method shows also a better conservation of AVO variations along the primaries after demultiple, compared to the popular frequency-domain implementations of the PRT. The algorithm is very fast for marine seismic gathers having the same streamer geometry. If the processing parameters for the parabolic Radon filtering are the same for these gathers, the pseudo-inverse operator for the minimum-norm solution can be saved in memory and the modified PRT panel is obtained for each CMP gather by means of complex back-substitution only. This has advantages over frequency-domain algorithms requiring saving the same operators in memory for each spectral component to expect similar time saving (Trad *et al.*, 2003).

The synthetic and real data examples for multiple attenuation showed that the new method is six to eight times faster than a PRT method based on the iteratively reweighted least-squares (IRLS) inversion method. The Mississipi Canyon data example showed that optimum multiple suppression was obtained by applying the SRME algorithm followed by the modified PRT method.

ACKNOWLEDGMENTS

The authors thank WesternGeco for for making available the Gulf of Mexico data set and Eric Verschuur for sharing the results from the SRME algorithm on this data set. This project received financial support from the Norwegian University of Science and Technology (NTNU), PetroMax, the Federal University of Bahia (UFBA), CNPq and National Institute of Science and Technology of Petroleum Geophysics INCT-GP, VISTA, and the Norwegian Research Council through the ROSE project.

Appendix

AMPLITUDE-COMPENSATED SPARSENESS-CONSTRAINED TRANSFORM IN THE FREQUENCY DOMAIN

The cost function for a constrained problem in the least-squares sense can be expressed as

$$S(\mathbf{m}) = \| \mathbf{Lm} - \mathbf{d} \| + \epsilon \mathbf{m}^{\mathbf{H}} \mathbf{W}_{\mathbf{m}}^{-\mathbf{1}} \mathbf{m}, \tag{A-1}$$

where $\mathbf{W_m}$ is a weighting matrix used for preconditioning the model space, chosen to impose some desired features in the model (sparseness here), and ϵ is a damping factor. The minimization of the cost function in (A-1) is translated into solving the complex system of equations

$$\begin{bmatrix} \mathbf{L} \\ \xi^{1/2} \mathbf{W}_{\mathbf{m}}^{-\mathbf{1/2}} \end{bmatrix} \mathbf{m} = \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix}. \tag{A-2}$$

We apply a left-preconditioning to the model $\mathbf{m}$ and a right-preconditioning to the operator $\mathbf{L}$ with the weighting matrix $\mathbf{W_m}$, so the system in (A-2) can be written as

$$\begin{bmatrix} \bar{\mathbf{L}} \\ \xi^{1/2} \mathbf{I} \end{bmatrix} \bar{\mathbf{m}} = \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \text{ where } \begin{array}{c} \bar{\mathbf{L}} = \mathbf{L} \mathbf{W}_{\mathbf{m}}^{1/2} \\ \bar{\mathbf{m}} = \mathbf{W}_{\mathbf{m}}^{1/2} \mathbf{m} \end{array}. \tag{A-3}$$

The problem in (A-3) is first solved for the $\bar{\mathbf{m}}$ before recovering the true model $\mathbf{m}$ by removing the effect of $\mathbf{W_m}$. The weighting matrix $\mathbf{W_m}$ is chosen such that

$$\mathbf{W}_{\mathbf{m},k} = |\mathbf{m}_k|^{(2-p)/2}. \tag{A-4}$$

where p is the norm of the problem. $\mathbf{W_m} = \mathbf{I}$ for the least-squares solution ($p = 2$). Several choices for designing the weighting matrices $\mathbf{W_m}$ for the model $\mathbf{m}$ have been proposed (Sacchi and Ulrych, 1995; Cary, 1998; Hermann *et al.*, 2000; Trad *et al.*, 2001; Nowak and Imhof, 2004; Ji, 2006). $\mathbf{W_m}$ **can be** defined from the computed model at the previous frequency component. Schonewille and Zwartjes (2001) proposed an iterative process with a number of outer loops to refine the solution. For the first iteration, the weighting matrix we used in the paper examples is given by

$$\mathbf{W}_{\mathbf{m},k}(f_j) = \frac{|\mathbf{m}_k(f_{j-1})|}{max|\mathbf{m}(f_{j-1})|} \quad \text{for} f_j \geq f_{min}, \quad k = 1, \cdots, N_q. \tag{A-5}$$

No weighting is used for frequencies outside this domain ($\mathbf{W_m} = \mathbf{I}$) which corresponds to the least-squares solution. For an iteration $i > 1$, $\mathbf{W_m}$ is the following average (Schonewille and Zwartjes, 2001)

$$\mathbf{W}^{(i)}_{\mathbf{m},k} = \frac{\sum_{j=1}^{n_f} |\mathbf{m}_k(f_j)|^{(i-1)}}{n_f \max|\mathbf{m}_k|^{(i-1)}}, \tag{A-6}$$

where i denotes the iteration index. The system (A-3) is solved iteratively by updating $\mathbf{W_m}$ at each iteration to impose more sparseness in the solution. We used the complex LSQR to solve the preconditioned problem in (A-3) given $\mathbf{W_m}$. Except for the first iteration, $\mathbf{W_m}$ is frequency-independent and is obtained by the averaging in (A-6). Other choices can be made for building the matrix $\mathbf{W_m}$ in equations (A-5) and (A-6) giving several variants to this algorithm.

REFERENCES

Abbad, B., B. Ursin, and D. Rappin, 2009, Automatic nonhyperbolic velocity analysis: Geophysics, **74**, Issue **2**, U1-U12.

Björck, Å, 1996, Numerical methods for least squares problems, Society of Industrial and Applied Mathematics.

Cary, P.W., 1998, The simplest discrete Radon transform, The 68^{th} Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 1999-2002.

Dragoset, B., 1999, A practical approach to surface multiple attenuation: The Leading Edge, **18**, 104-108.

Foster, D.J., and C.C. Mosher, 1992, Suppression of multiple reflections using the Radon transform: Geophysics, **57**, 386-395.

Guitton A., and G. Cambois, 1999, Multiple elimination using a pattern-recognition technique: The Leading Edge, **18**, 92-98.

Hadidi, M.T., M. Sabih, D.E. Johnston, and C. Calderon-Macias, 1999, Mobil's results for the 1997 workshop on multiple attenuation: The Leading Edge, **18**, 100-103.

Hampson, D., 1986, Inverse velocity stacking for multiple elimination, Journal of the Canadian Society of Exploration Geophysicists, **22**, 44-55.

Herrmann, P., T. Mojesky, M., Mageson, and P. Hugonnet, 2000, De-aliased, high-resolution Radon transforms, The 70^{th} Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 1953-1956.

Hugonnet, P., and G. Canadas, 1995, Aliasing in the parabolic Radon transform: The 65^{th} Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 1366-1369.

Ji, J., 2006, CGG method for robust inversion and its application to velocity-stack inversion: Geophysics, **71**, Issue **4**, R59-R67.

Lamont, M.G., B.M. Hartley, and N.F. Uren, 1999, Multiple attenuation using the NMO and ISR preconditioning transforms: The Leading Edge, **18**, 110-114.

Lokshtanov, D., 1999, Multiple suppression by data-consistent deconvolution: The Leading Edge, **18**, 115-119.

Nowak, E.J., and M.G. Imhof, 2004, Diffractor localization via weighted Radon transform, The 74^{th} Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 2108-2111.

Sacchi, M.D., and M.J. Porsani, 1999, Fast high resolution parabolic Radon transform, The 69^{th} Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 1477-1480.

Sacchi, M.D., and T.J. Ulrych, 1995, High resolution velocity gathers and offset space reconstruction: Geophysics, **60**, 1169-1177.

Schonewille, M.A., and P. A., Aaron, 2007, Applications of time-domain high-resolution Radon demultiple, The 69^{th} EAGE Conference and Exhibition.

Schonewille, M.A., and A. J. W. Duijndam, 2001, Parabolic Radon trans-

form, sampling and efficiency: Geophysics, **66**, 667-678.

Schonewille, M.A., and P. Zwartjes, 2001, High-resolution transforms and amplitude preservation: The 72^{th} Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 2066-2069.

Thorson, J.R., and J.F. Claerbout, 1985, Velocity-stack and slant-stack stochastic inversion, Geophysics, **50**, 2727-2741.

Trad D., 2001, Computation and implementation of the Radon transforms. PhD thesis, University of British Columbia, Canada.

Trad, D., T. Ulrych, and M.D. Sacchi, 2002, Accurate interpolation with high-resolution time-variant Radon transforms: Geophysics, **67**, 644-656.

Trad, D., T. Ulrych, and M.D. Sacchi, 2003, Latest views of the sparse Radon transform: Geophysics, **68**, 386-399.

Turner, G., 1990, Aliasing in the tau-p transform and the removal of spatially aliased coherent noise: Geophysics, **55**, 1496-1503.

Verschuur, D. J., and R. J. Plein, 1999, Multiple removal results from Delft University: The Leading Edge, **18**, 86-91.

Yilmaz, Ö, 1989, Velocity stack processing: Geophysical Prospecting, **37**, 357-382.

Zhou, B., and S.A. Greenhalgh, 1994, Linear and parabolic $\tau - p$ transforms revisited: Geophysics, **59**, 1133-1149.

LIST OF FIGURES

LIST OF TABLES

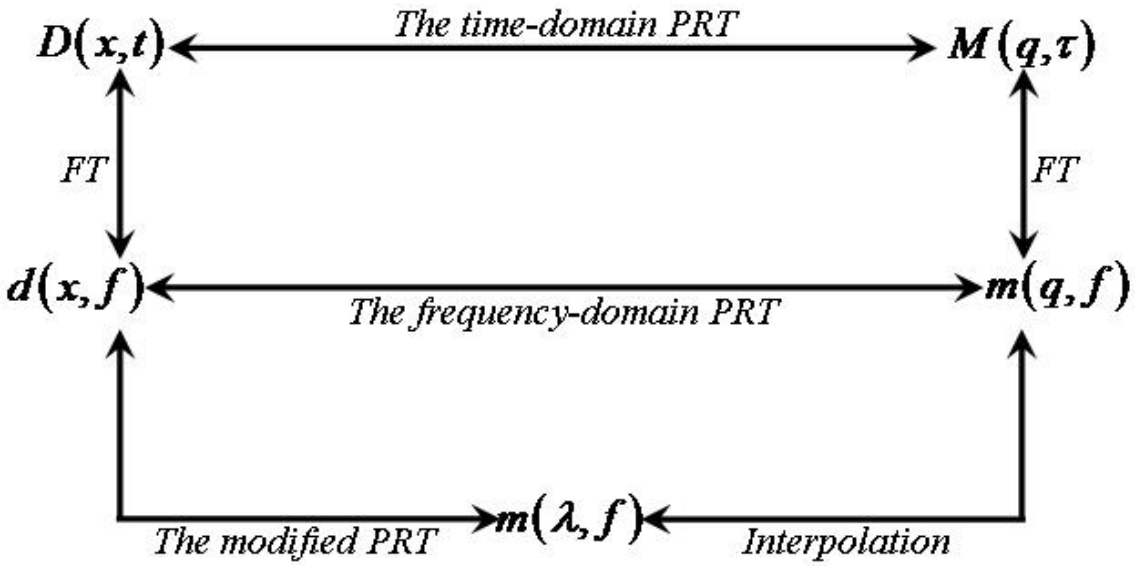

Figure 1. Relations between the different transforms. **PRT** stands for the standard parabolic Radon transform, **FT** for the Fourier transform. t, f and q are time, frequency and curvature, respectively. $\lambda = qf$ is introduced to define the modified PRT with a frequency-independent operator. The frequency-domain PRT and the modified PRT are connected by means of interpolation in the $\lambda - f$ domain.

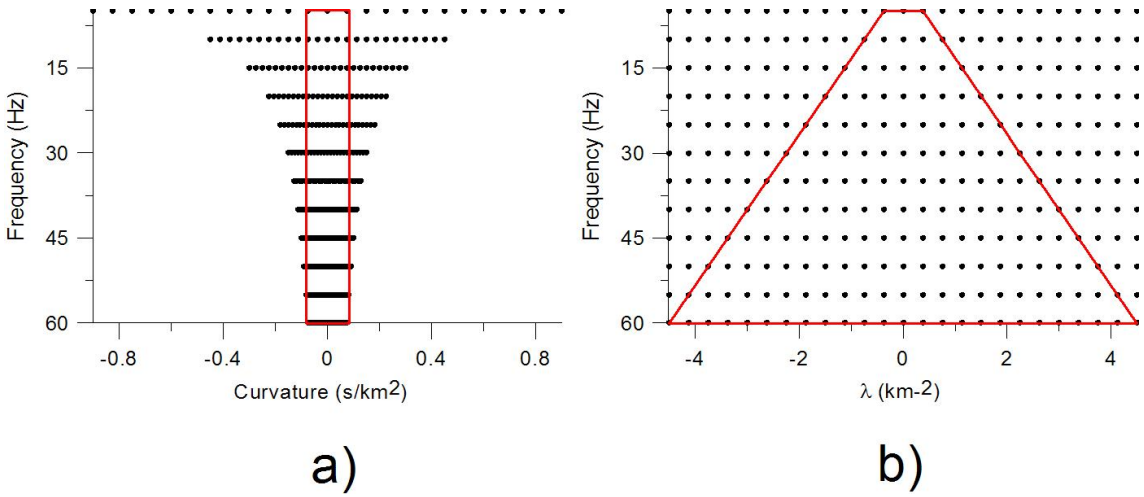

Figure 2. Relation between the curvature range for the standard a) and **the** modified PRT's b). The regions in red show the covered curvature- (in a)) and $\lambda-$ regions (in b)) for the standard PRT.

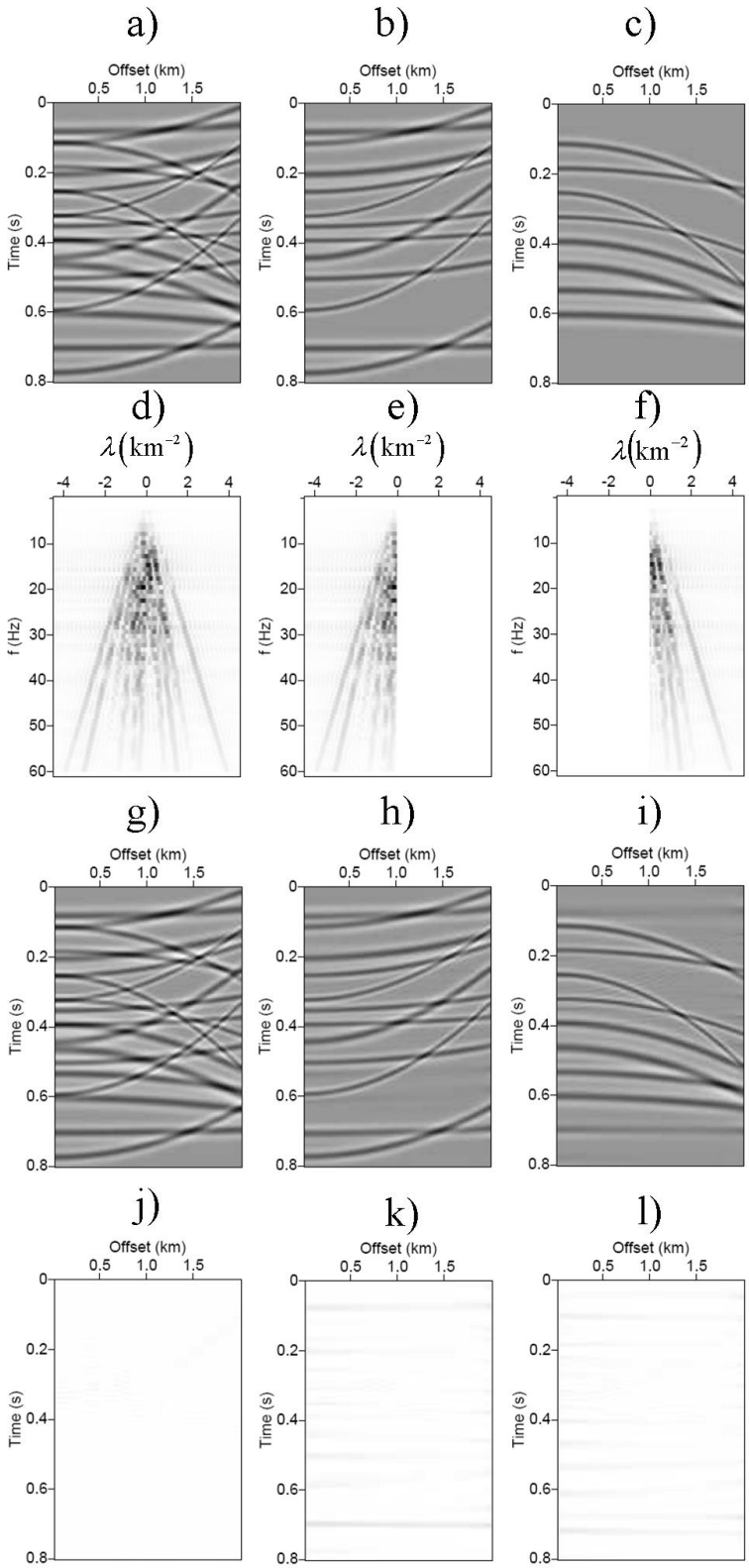
a)
b)
c)
Offset (km)
0.5 1.0 1.5
Time (s)
0 0.2 0.4 0.6 0.8
d)
e)
f)
λ(km⁻²)
-4 -2 0 2 4
f (Hz)
10 20 30 40 50 60
g)
h)
i)
j)
k)
l)

Figure 3. The modified PRT for a synthetic gather containing 20 noise-free events (12 primaries and 8 multiples) in a narrow time window without AVO changes. a) The total gather; b) The gather for primaries; c) The gather for multiples. d), e), and f) are the $\lambda - f$ domains mapping the events in a), b), and c), respectively. g), h), and i) are reconstructed gathers after inverse transforms. j), k), and l) are reconstruction errors for the respective gathers.

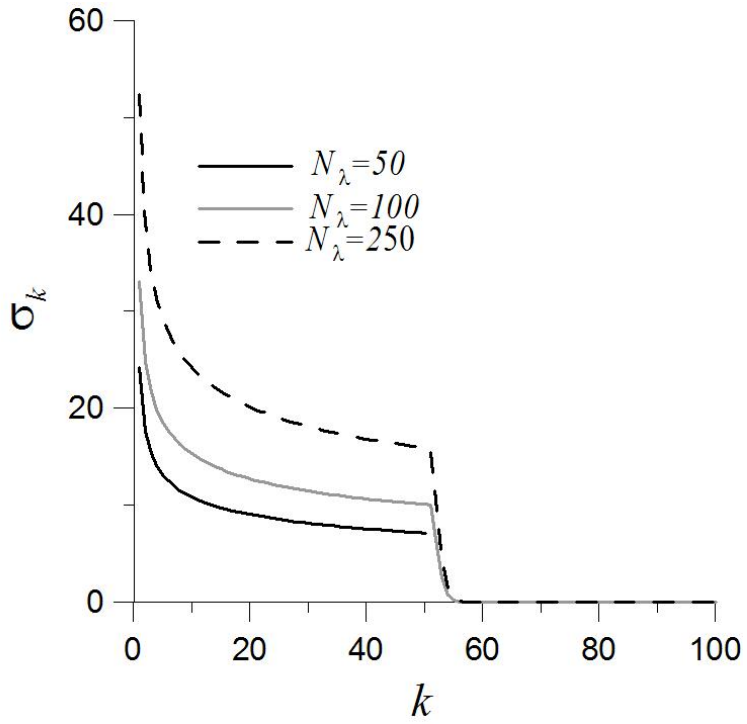

Figure 4. Singular values σ_k for the operator $\mathbf{L}(\lambda)$ for the synthetic data example ($N_x = 100$) as function of the eigenvalue index k.

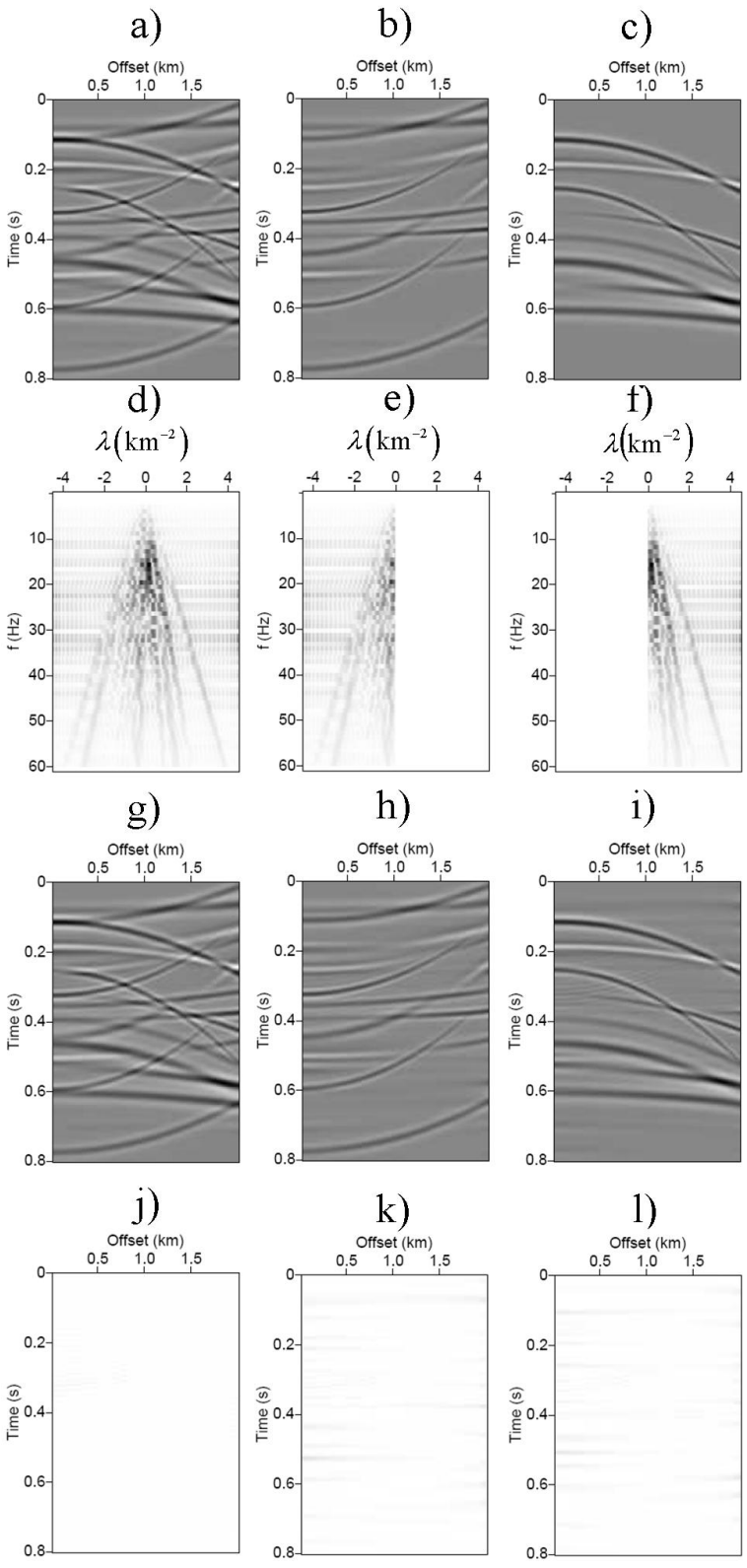
a)
b)
c)
Offset (km)
0.5 1.0 1.5
Time (s)
0 0.2 0.4 0.6 0.8
d)
e)
f)
λ(km⁻²)
-4 -2 0 2 4
f (Hz)
10 20 30 40 50 60
g)
h)
i)
j)
k)
l)

Figure 5. The modfied PRT for the gather containing the events in Figure 3 with AVO effects included along the events. a) The total gather; b) The gather for primaries; c) The gather for multiples. d), e), and f) are the $\lambda - f$ domains mapping the events in a), b), and c), respectively. g), h), and i) are the reconstructed gathers after inverse transforms. j), k), and l) denote reconstruction errors for the respective gathers.

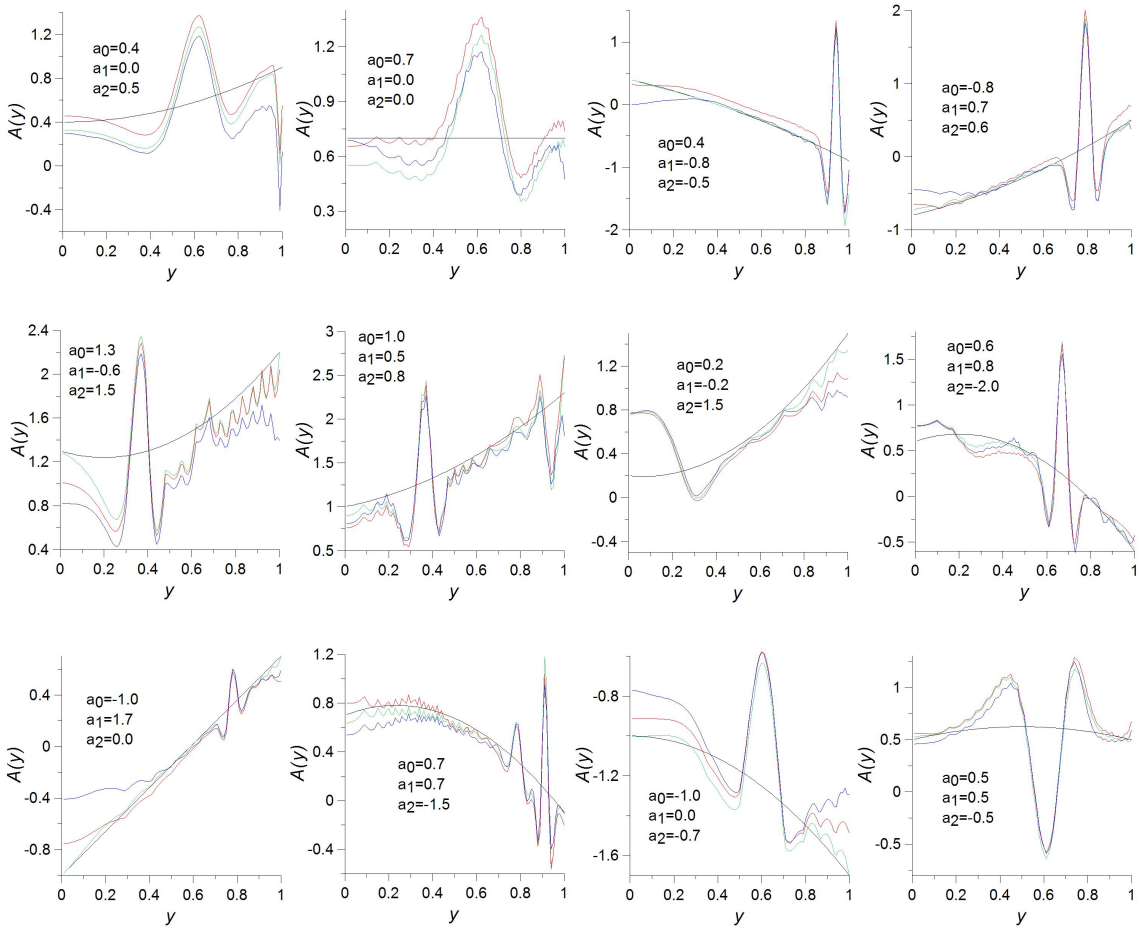

Figure 6. Theoretical and inverted AVO curves with respect to the normalized offset ,y, in the example of Figure 5. The amplitude curves for the modified PRT (blue), the IRLS method (red), the exact AVO response (black), and the practical AVO response that accounts for event interference (green) are all compared for the 12 primary events.

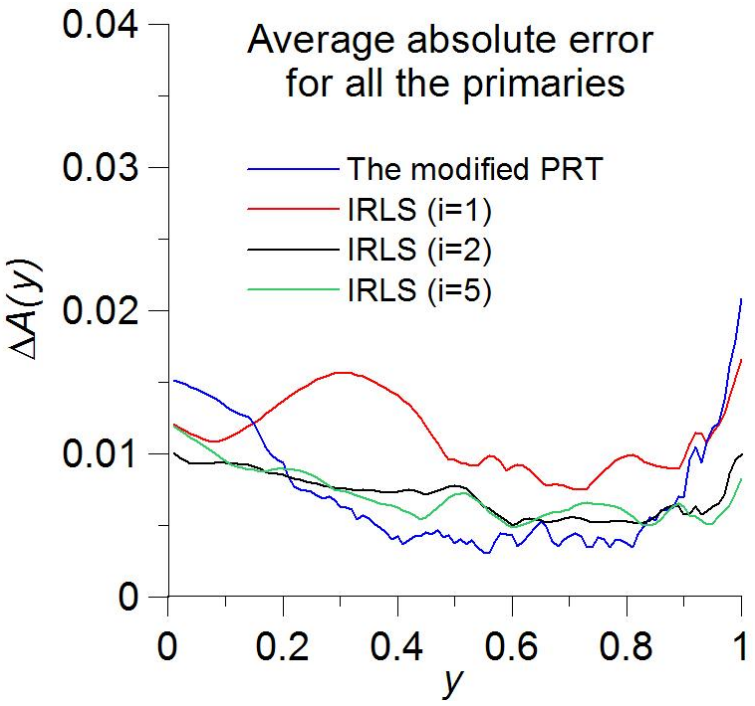

Figure 7. Average absolute error on the AVO inverted curves with respect to normalized offset, y, in the example of Figure 4 over the 12 primary events using the modified PRT (blue curve) and the IRLS methods using one (red curve), two (black curve), and five iterations (green curve).

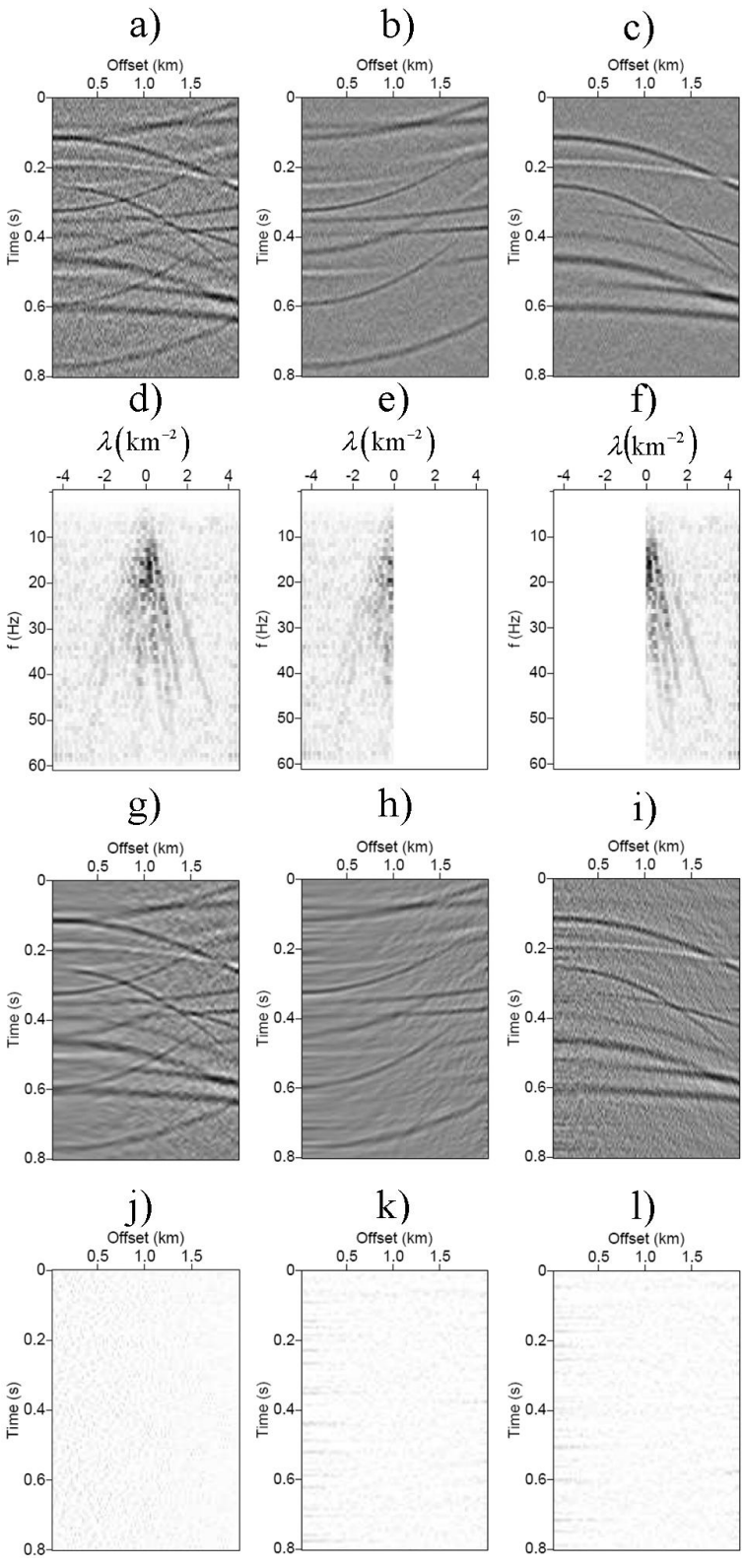
a)
b)
c)
Offset (km)
0.5 1.0 1.5
Time (s)
0 0.2 0.4 0.6 0.8
d)
e)
f)
$\lambda(\mathrm{km}^{-2})$
-4 -2 0 2 4
f (Hz)
10 20 30 40 50 60
g)
h)
i)
j)
k)
l)

Figure 8. The modified PRT for the gather in Figure 3 with AVO effects and noise added. a) The total gather; b) the gather for primaries; c) The gather for multiples. d), e), and f) are the $\lambda - f$ domains mapping the events in a), b), and c), respectively. g), h), and i) are the reconstructed gathers after inverse transforms. j), k), and l) are reconstruction errors for the respective gathers.

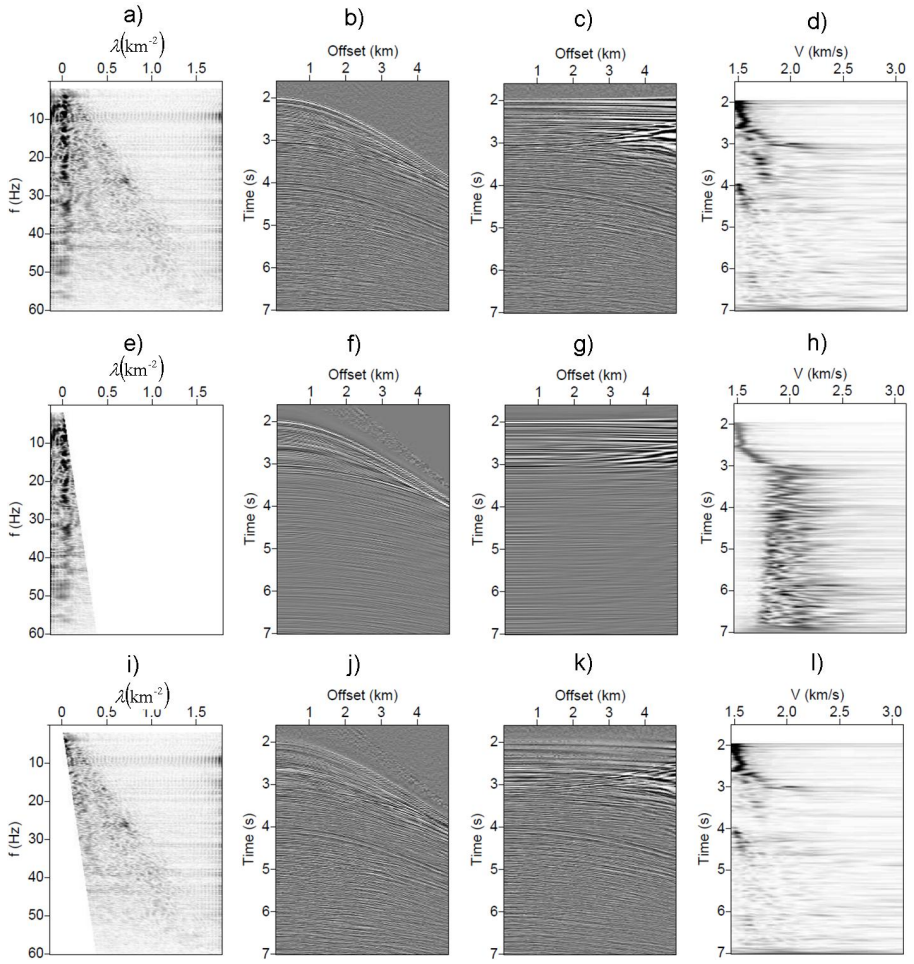

Figure 9. The modified PRT for a real data gather from the Gulf of Mexico. a) The $\lambda - f$ domain for the gather in c). b) The input gather. c) The gather in b) after nonhyperbolic NMO corrections. d) Semblance map for the gather in b). e) and i) The $\lambda - f$ domain for the primaries and multiples, respectively. f) and j) Filtered primaries and multiples after inverse moveout corrections. g) and k) recovered primary and multiple gathers after inverse transform. h) and l) are velocity spectra for the separated primaries and multiples.

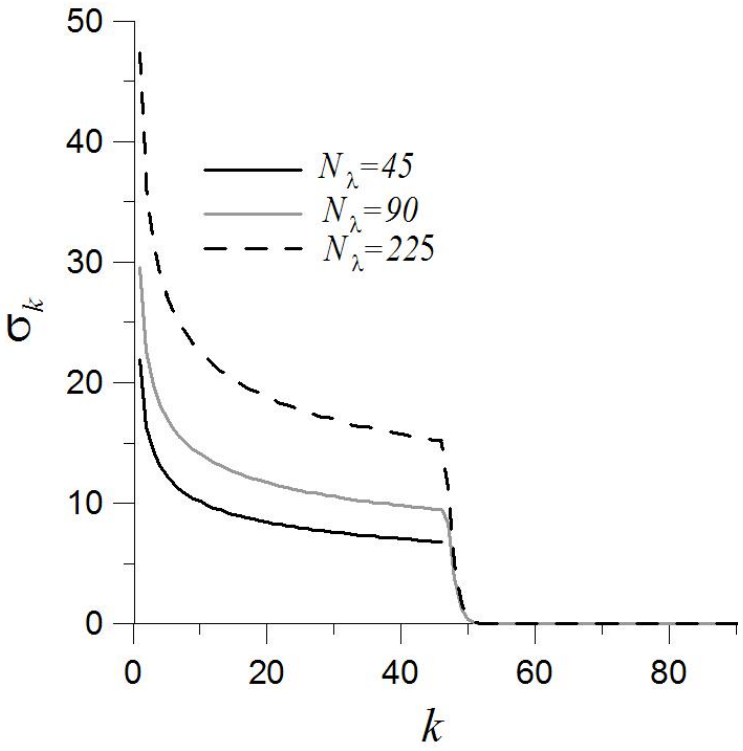

Figure 10. Singular values σ_k for the operator $\mathbf{L}(\lambda)$ for the Gulf of Mexico data ($N_x = 90$) as function of the eigenvalue index k.

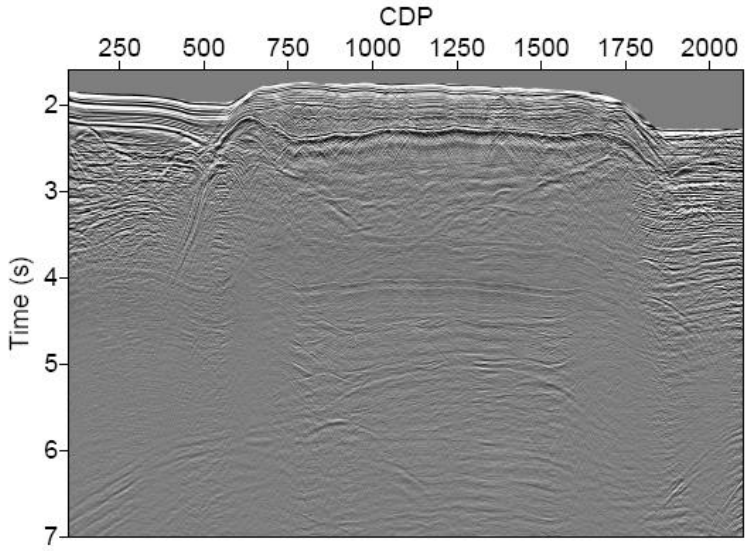

Figure 11. Raw stack obtained through nonhyperbolic moveout corrections and stack. No demultiple was applied.

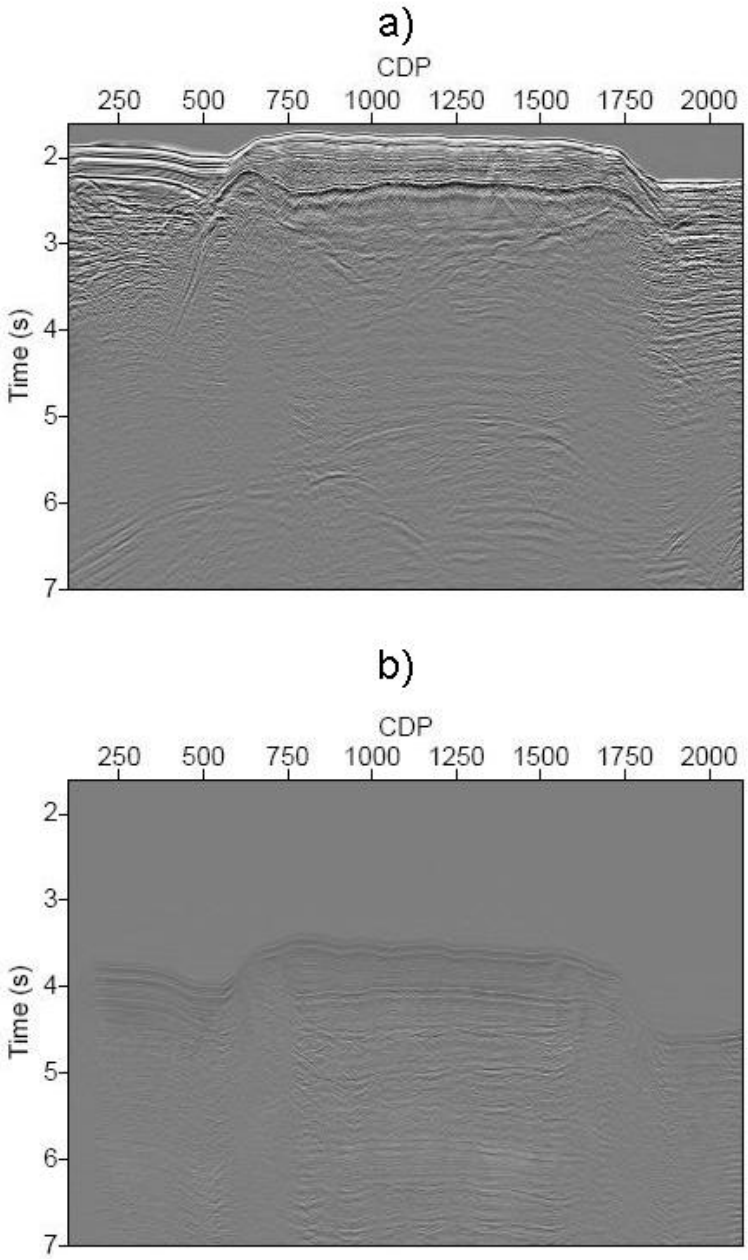

Figure 12. Demultiple results using the surface related multiple elimination (SRME) algorithm (from Verschuur and Plein, 1999). a) Obtained stack. b) Estimated multiples.

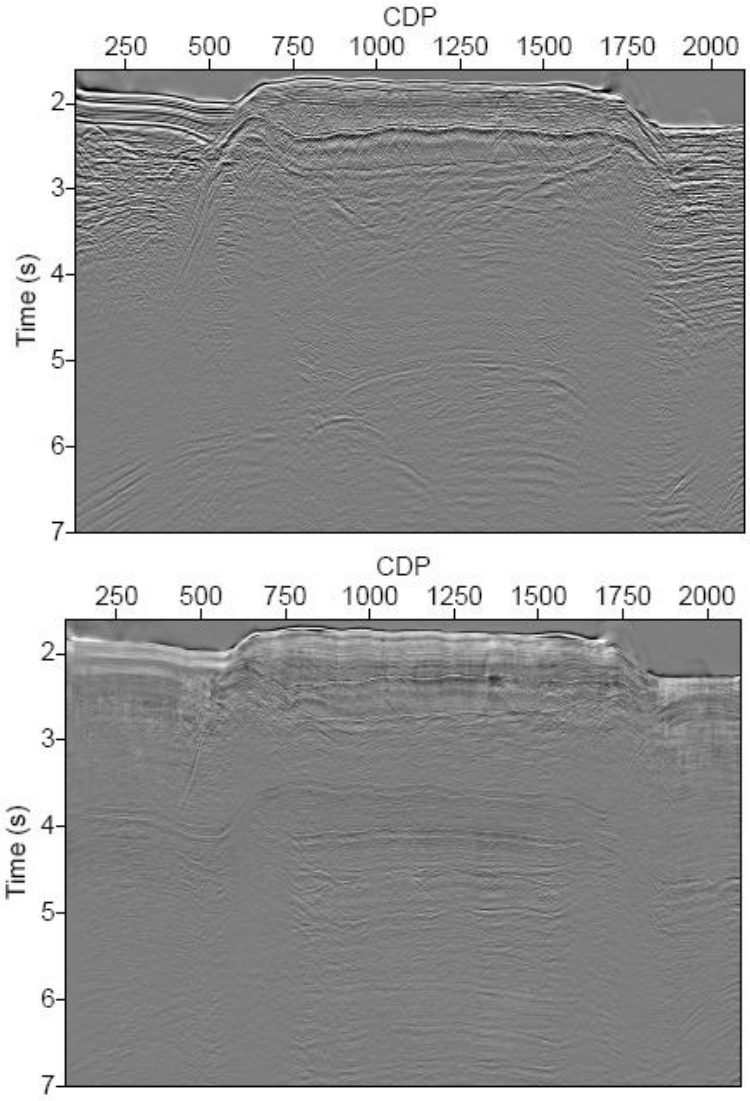

Figure 13. Demultiple results using the modified PRT algorithm. a) Obtained stack. b) Estimated multiples.

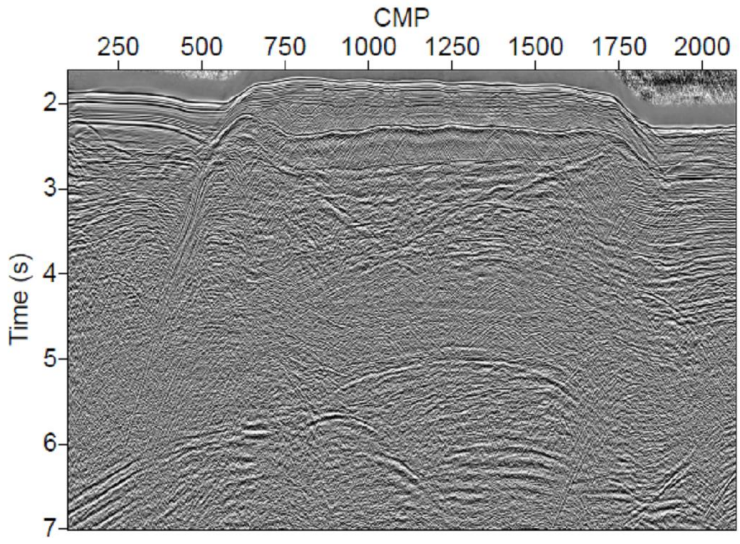

Figure 14. Stack obtained with multiple removal using the SRME method followed by the modified PRT algorithm. Automatic gain control is applied after stack.

CPU time	Modified PRT	IRLS (i=1)	IRLS (i=2)	IRLS (i=3)	IRLS (i=5)
Synthetic example	0.517 s	1.870 s	3.113 s	4.293 s	6.516 s
Single CMP	2.865 s	13.525 s	19.199 s	25.089 s	35.637 s
Real data (2180 CMP's)	51 min 03 s	291 min 05 s	582 min 26 s	1153 min	1438 min

Table 1. CPU time for the different algorithms on synthetic and real data examples.