TIME-IN Variant RADON TRANSFORM BY GENERALIZED FOURIER SLICE THEOREM

ALI Gholami∗
Institute of Geophysics, University of Tehran
Tehran, Iran

MAURICIO D. Sacchi
Department of Physics, University of Alberta
Edmonton, AB, Canada

(Communicated by Naoki Saito)

Abstract. Time-invariant Radon transforms play an important role in many fields of imaging sciences, whereby a function is transformed linearly by integrating it along specific paths, e.g. straight lines, parabolas, etc. In the case of linear Radon transform, the Fourier slice theorem establishes a simple analytic relationship between the 2-D Fourier representation of the function and the 1-D Fourier representation of its Radon transform. However, the theorem can not be utilized for computing the Radon integral along paths other than straight lines. We generalize the Fourier slice theorem to make it applicable to general time-invariant Radon transforms. Specifically, we derive an analytic expression that connects the 1-D Fourier coefficients of the function to the 2-D Fourier coefficients of its general Radon transform. For discrete data, the model coefficients are defined over the data coefficients on non-Cartesian points. It is shown numerically that a simple linear interpolation provide satisfactory results and in this case implementations of both the inverse operator and its adjoint are fast in the sense that they run in \( O(N\log N) \) flops, where \( N \) is the maximum number of samples in the data space or model space. These two canonical operators are utilized for efficient implementation of the sparse Radon transform via the split Bregman iterative method. We provide numerical examples showing high-performance of this method for noise attenuation and wavefield separation in seismic data.

1. Introduction. The Radon transform has a long history in image processing and seismic data processing. In image processing it is often called the Hough transform [13] and it has been used for pattern recognition and feature extraction from images. The Radon transform has also been used in seismic data processing to solve wave propagation and inversion problems in media where the subsurface speed only varies with depth [28]. In the context of our work, we will discuss the Radon transform and a novel implementation of it, for seismic data processing. Radon transforms have been extensively used for random and coherent noise attenuation. Several classes of Radon transforms have been proposed. For instance, the linear Radon transform has been implemented to remove coherent noise [30] from seismic records. The

2010 Mathematics Subject Classification. Primary: 44A12; Secondary: 86A15.
Key words and phrases. Parabolic Radon transform, linear Radon transform, slant stack, generalized Fourier slice theorem, seismic multiple attenuation.

∗ Corresponding author: Ali Gholami.
parabolic Radon transform has been mainly utilized to separate multiples from primary reflections in Normal-Move-Out (NMO) corrected Common-Mid-Point (CMP) gathers [12]. Both linear and parabolic Radon transforms are usually implemented in the frequency-space \((f - x)\) domain via the solution of an inverse problem that entails solving a linear system of equation for each frequency, independently [5].

Radon transforms that utilize hyperbolic integration paths have been also proposed to attenuate multiple reflections in CMP gathers and for seismic data reconstruction. Hyperbolic Radon transforms are time-variant transforms that cannot be implemented in the \((f - x)\) domain. Due to the computational load of the hyperbolic Radon transform, seismic data are usually modified, either by an approximate NMO correction applied to the data [12] or by stretching the data along time axis [29], in order to used the parabolic Radon transform instead. Recently, a fast algorithm based on the butterfly algorithm has been proposed in [17] for generalized Radon transforms. A fast hyperbolic Radon transform has also been represented as convolutions in log-polar coordinates in [21].

[26] proposed to increase the resolution of the Radon transform by incorporating sparsity constraints. A similar idea was also introduced by [23] and [24] to improve the focusing power of the frequency domain Radon transforms. Recently, [10] presented deconvolutive Radon transform as a generalization of the conventional Radon transform and to increase the temporal resolution. Today, high resolution Radon transforms (often called sparse Radon transforms) are part of our arsenal of techniques for de-multiple and seismic data reconstruction [6, 14, 27, 15]. In general, sparsity promoting iterative algorithms are utilized to estimate the Radon coefficients [27, 19]. To a degree, the successful and efficient implementation of such algorithms relies on having access to two efficient canonical operators: the inverse Radon operator and its adjoint, which are needed to estimate the Radon domain coefficients that can model the data.

Traditional solvers for computing linear and parabolic Radon transforms rely on the Toeplitz structure of the time invariant Radon operator and solve the problem for each frequency independently [18, 22]. In this way the solution of original large system is replaced by the solution of smaller systems [12], corresponding to each frequency. Therefore, it is implemented more efficiently than the time domain correspondence due to the scalability in parallel computing. However, in traditional \((f - x)\) algorithms, the coefficients are regularized only in \(p\) direction.

The Fourier slice theorem [20] provides an efficient way for computing the linear Radon transforms without matrix inversion. However, to our knowledge, this theorem has not been used to compute more general time-invariant Radon transforms such as the parabolic Radon transform which is extensively utilized in the field of applied geophysics. Here, we generalize the Fourier slice theorem to make it applicable to any general time-invariant Radon transform. We derive an analytic expression between the Fourier coefficients of a function and the Fourier coefficients of its general Radon transform. In this way, very fast implementations of the inverse operator and its adjoint are possible via Fast Fourier Transform (FFT) algorithms. These two canonical operators are utilized for efficient implementation of the sparse Radon transform via split Bregman iterative method [11]. The optimal number of algorithm iterations is determined by the discrepancy principle, making the algorithm very efficient for processing large seismic data sets. Numerical results of noise attenuation and wavefield separation from synthetic and field seismic data show high performance of the proposed method.
1.1. Outline. This work is structured as follows. In section 2 we briefly introduce the time-invariant Radon transform in continuous form. Then the Fourier slice theorem and its generalization are described in sections 2.1 and 2.2, respectively. An inverse transformation of the Radon transform by the generalized Fourier slice theorem is formulated in section 2.3. We extent our formulation for treating discrete data in section 3 and analyse the complexity of the algorithm in section 3.1. In section 4 we solve sparse Radon transform by using a Bregman iterative method and the generalized Fourier slice theorem. We give synthetic examples for linear and parabolic Radon transforms and the stability against noise in section 5.1 and compare computational cost and recovery performance with traditional direct methods. An example of spatial interpolation of irregular data is presented in section 5.2. Finally, the method is applied to field data for wavefield separation in section 5.3 and conclusions are given in section 6.

2. The continuous time-invariant Radon transform. Through this paper, we use the terminology adopted in seismic data processing applications. Let \( v(t, x) \) be a seismogram (function) of spatial variable (offset) \( x \) and time \( t \). The variable \( x \) is used to indicate the distance between a source and receiver in a variety of acquisition scenarios. The time-invariant Radon transform of \( v \) can be defined by the following integral

\[
(1) \quad u(\tau, p) = \int_{\text{Dom}(v)} v(t = \tau + p \varphi(x), x) dx,
\]

where \( \tau \) is intercept time, \( p \) is a parameter, and function \( \varphi \) determines the integration path. Specific seismic events contained in the seismogram \( v \) can be modelled by a family of parametric curves \( \tau + p \varphi(x) \) for a predefined function \( \varphi \). For example, linear coherent noise and surface waves can be modelled by the linear Radon transform: \( \varphi(x) = x \). Similarly, reflections (after partial NMO correction) can be modelled by parabolic curves: \( \varphi(x) = x^2 \), [12]. Therefore, some events contained in the seismogram could be focused by (1) in the transform domain. The focused events in the Radon domain can be easily isolated and filtered. Formula (1) is a generalization of the slant stacking technique used in geophysical community. Considering the wide application of Radon transform defined in (1) efficient and accurate implementation of it has always been a goal for geophysicists. Current algorithms solve time-invariant Radon transform (1) via an inverse strategy in the \( f - x \) domain as

\[
(2) \quad v(f, x) = \int u(f, p) \exp(-i2\pi fp\varphi(x)) dp, \quad i^2 = -1,
\]

which is solved for each frequency \( f \), separately. Therefore, in the case of discrete data, one needs to solve a system of complex linear equations in order to recover a frequency slice of the Radon plane [5, 12, 18, 25, 30, 22]:

\[
(3) \quad v(f) = L(f)u(f),
\]

where \( v(f) \) and \( u(f) \) denote frequency slices of data and Radon panels, and \( L(f) \) is a frequency-dependent linear operator. Due to variation of the operator with frequency, processing all frequencies can be a time-consuming task which can limit the usage of the transform for large data sets. Here, we call this method direct Radon transform and use it as a benchmark for comparing the performance of the new method presented in this paper. The Fourier slice theorem is an efficient tool.
for fast computation of Radon transform when $\varphi(x) = x$. Here, we generalize this theorem for fast computation of a general form of Radon transform.

2.1. The Fourier-slice theorem. In the case of linear Radon transform, the following theorem provides a simple way for computing the $f - p$ coefficients of the Radon transform.

**Theorem 2.1** (Fourier slice theorem). Let $v(t, x)$ be a function having two-dimensional (2-D) Fourier transform $\hat{v}(f, k_x)$, and $u(\tau, p)$ be its linear Radon transform defined in (1) for $\varphi(x) = x$, then $\hat{u}(f, p) = \hat{v}(f, k_x = fp)$, where $\hat{u}(f, p)$ is the one-dimensional (1-D) Fourier transform of $u$ along the time axis.

*Proof.* The proof is simple and can be found in many textbooks (e.g., [16]).

This theorem establishes the relation

$$k_x = fp$$

between wave number $k_x$, frequency $f$, and slowness $p$ and states that the Fourier transformation of the projection of $v$ along a slope $p$ is equal to the radial slice taken through the 2-D Fourier transform of $v$ with the same slope (see Figure 1). A fast and invertible algorithm for the linear Radon transform based on this theorem and pseudo-polar Fourier transform is given in [1, 2]. A fast algorithm has also been proposed for 3-D Radon transform via backprojection in [3]. However, this theorem can not be utilized for computing the Radon transform (1) for a general form of $\varphi$. In the next section, we generalize the theorem for this task.

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Graphical representation of the Fourier slice theorem. The seismogram $v$ includes a linear event. The projection of $v$ along the slope of the event is the red line shown in $u$ which is the IFT of the radial slice taken from $\hat{v}$ using relation $k_x = fp$. 
2.2. Generalized Fourier-slice theorem.

**Theorem 2.2** (generalized Fourier-slice theorem). Let \( v(t, x) \) be a function having 1-D Fourier transform \( \hat{v}(f, x) \), and \( u(\tau, p) \) be its general Radon transform defined in (1) with 2-D Fourier transform \( \hat{u}(f, k_p) \), then there is the simple relation

\[
k_p = f \varphi(x)
\]

between \( k_p, f, \) and \( x \).

**Proof.** Applying the 2-D Fourier transform to both sides of (1) gives

\[
\hat{u}(f, k_p) = \iint v(t = \tau + p \varphi(x), x) e^{-i2\pi(f \tau + k_p p)} d\tau dp dx
\]

(6)

\[
= \iint \hat{v}(f, x) e^{i2\pi f p \varphi(x)} e^{-i2\pi k_p p} dp dx
\]

(7)

\[
= \int \hat{v}(f, x) \left( \int e^{-i2\pi[k_p - f \varphi(x)] p} dp \right) dx
\]

(8)

which confirms the relation (5).

According to the new theorem the Fourier transformation of a trace at offset \( x \) is the same as the radical slice through the 2-D Fourier transformation of the Radon transform with the slope \( \varphi(x) \). Figure 2 schematically shows the relation between computation of the linear Radon transform via (1) and via the generalized Fourier-slice theorem. The difference between the Fourier-slice theorem and its generalized form is that the former relates the 2-D Fourier transformation of the data to the 1-D Fourier transformation of its Radon plane while the latter relates the 1-D Fourier transformation of the data to 2-D Fourier transformation of its Radon plane.

2.3. **Inverse transform.** The inverse Fourier transform recovers \( v(t, x) \) from \( \hat{v}(f, x) \) via

\[
v(t, x) = \int \hat{v}(f, x) e^{i2\pi f t} df.
\]

(9)

Using \( \hat{v}(f, x) = \hat{u}(f, k_p = f \varphi(x)) \) from the theorem 2.2 we have

\[
v(t, x) = \int \hat{u}(f, k_p = f \varphi(x)) e^{i2\pi f t} df
\]

(10)

and using the following expression

\[
\hat{u}(f, k_p) = \int \hat{u}(f, p) e^{-i2\pi k_p p} dp
\]

(11)

we arrive to the following results

\[
v(t, x) = \iint \hat{u}(f, p) e^{-i2\pi f p \varphi(x)} e^{i2\pi f t} df dp
\]

(12)

\[
= \iint \hat{u}(f, p) e^{i2\pi f[t - p \varphi(x)]} df dp
\]

(13)

\[
= \iint \hat{u}(f, p) e^{i2\pi f \tau} df dp
\]

(14)

\[
= \iint u(\tau = t - p \varphi(x), p) dp.
\]

(15)
Figure 2. Graphical representation of the generalized Fourier slice theorem. The seismogram \( v \) includes a linear event. The Fourier transform of a trace at offset \( x \) is the same as the radical slice taken through the 2-D Fourier transform of its linear Radon transform \( u \) using the relation \( k_p = f x \).

3. The discrete time-invariant Radon transform. The discrete time-invariant Radon transforms are of particular interest in practical seismic applications when dealing with discrete data in time and space. We assume that the sampling of the data domain and the Radon domain are uniform. We designate \( \Delta x, \Delta t, \Delta p, \) and \( \Delta \tau \) the sampling interval along \( x, t, p, \) and \( \tau \) axes, respectively. We also define \( n_x, n_t, n_p, \) and \( n_\tau \) the corresponding discrete index parameters, then the correspondance between the discrete and continuous variables is given by

\[
\begin{align*}
x(n_x) &= x_{\text{min}} + n_x \Delta x, & n_x = 0, \ldots, N_x - 1, \\
t(n_t) &= t_{\text{min}} + n_t \Delta t, & n_t = 0, \ldots, N_t - 1, \\
p(n_p) &= p_{\text{min}} + n_p \Delta p, & n_p = 0, \ldots, N_p - 1, \\
\tau(n_\tau) &= \tau_{\text{min}} + n_\tau \Delta \tau, & n_\tau = 0, \ldots, N_\tau - 1.
\end{align*}
\]

Usually, the lower limits \( t_{\text{min}}, x_{\text{min}} \) and \( \tau_{\text{min}} \) are zero. However, parameter \( p_{\text{min}} \) can get negative values corresponding to, e.g., negative slopes or curvatures. We use \( \ell_1 \) and \( \ell_2 \) to define uniformly sampled frequency and wavenumber of the Radon plane:

\[
f(\ell_1) = \ell_1 \Delta f \quad \text{and} \quad k_p(\ell_2) = \ell_2 \Delta k_p,
\]

for \( \ell_1 = 0, \ldots, N_\tau - 1 \) and \( \ell_2 = 0, \ldots, N_p - 1 \). Therefore, according to theorem 2.2

\[
\hat{u}(f(\ell_1), k_p(\ell_2)) = \hat{v}(f(\ell_1), x = \varphi^{-1}(k_p(\ell_2)/f(\ell_1))).
\]

Having computed \( \hat{u} \) the Radon transform \( u \) is obtained by 2-D inverse Fourier transformation. Figure 3 shows the points of \( \hat{u}(f, k_p) \) over imposed to the Cartesian grids of \( \hat{v}(f, x) \) for the linear and parabolic Radon transforms. Clearly, the computation
of \( \hat{u} \) from the samples of \( \hat{v} \) requires an interpolation stage. The latter is implemented in the frequency domain and discussed with more detail as we progress with our exposition. The inverse Radon transform is performed by computing \( \hat{v}(f,x) \) from \( \hat{u}(f,k) \) via the expression

\[
\hat{v}(f(\ell_1), x(n_x)) = \hat{u}(f(\ell_1), k_p = f(\ell_1)\varphi(x(n_x))),
\]
followed by inverse Fourier transform along the frequency axis.

Let \( f_{\text{max}} \) be the maximum frequency in the trace at largest offset, \( x_{\text{max}} \), then according to the relation (5) the maximum value of \( k_p \) is defined as \( k_{p\text{max}} = f_{\text{max}}\varphi(x_{\text{max}}) \), and thus

\[
\Delta k_p = \frac{1}{k_{p\text{max}}} = \frac{1}{f_{\text{max}}\varphi(x_{\text{max}})}
\]

which is in accordance with the result of [25]. Furthermore, since \( k_{p\text{max}} = (N_p - 1)\Delta k_p \) we obtain

\[
\Delta k_p = \frac{f_{\text{max}}\varphi(x_{\text{max}})}{N_p - 1}.
\]

On the other hand, according to the Nyquist sampling theory the maximum value of \( p \) recoverable from the sampled spectrum is defined as

\[
p_{\text{max}} = \frac{1}{2\Delta k_p}
\]

and aliasing occurs for the events with \( p \) values above this. For the computed spectrum of length \( N_p \) we need to perform an inverse DFT of length, at least, \( N_p \) and naturally the resulting Radon coefficients will be \( N_p \)-periodic. The left and right halves of the output sequence are thus swapped to get the Radon coefficients corresponding to the range \([-p_{\text{max}}, p_{\text{max}}]\).

3.1. Computational complexities. The forward transform is obtained first by performing a 1-D fast Fourier transform (FFT) along the time axis followed by interpolation. The 1-D FFT can be computed in \( O(N_x N_t \log N_t) \) operations while the cost of the (linear) interpolation is negligible. Then one performs a 2-D inverse FFT (IFFT), with total operations count \( O(N_x N_p \log N_p N_p) \), in order to get the estimated Radon coefficients. The same mechanism with reversed order can also be used for the inverse Radon transform to reconstruct the data from the Radon coefficients. Therefore, the implementation of both the forward and inverse Radon transformations has a complexity given by \( O(N\log N) \) where \( N \) is the maximum number of samples in the data space or model space. Generally, the number of samples in the data space is larger than or equal to the number of samples in the Radon space. Thus, \( N \) can be considered as the number of data samples.

4. Sparse Radon transforms. Time-invariant Radon transforms have been extensively used for noise attenuation and wavefield separation. The method is based on the focusing power of the Radon transform [27]. However, a practical implementation of the transform requires to consider the finiteness of the data and discretization effects that often deteriorate the resolution of the resulting Radon coefficients. Therefore, regularization strategies are required in order to estimate highly resolved
Figure 3. Relation between the grid points of the Radon plane $\hat{u}(f, k_p)$ (in green) and those of the data $\hat{v}(f, x)$ (in black) corresponding to the forward linear (a) and parabolic (c) Radon transforms. (b) and (d) show the grid points for the inverse transformations.

(focused) distribution of the Radon coefficients. Using the notation from linear algebra and based on the steps of the algorithm discussed above, the inverse Radon transform via the generalized Fourier slice theorem can be written as

$$v = F_1^{-1} S F_2 u$$

where $F_1$ and $F_2$ denote 1-D FFT and 2-D FFT operators of appropriate size. The operator $S$ re-samples the Fourier coefficients of the Radon transform from the Cartesian grid to the grid points defined by (17). Figure 3 shows the relation between the gridpoints of $\hat{u}$ and those of $\hat{v}$ for forward/inverse linear and parabolic Radon transforms. The figure is shown for only the positive frequencies and wavenumbers. In the case of the linear transformation, for each frequency $f$, the distance between successive points along the wavenumber axis remains constant. These points form a so called pseudopolar grid and [1, 2] used the chirp z-transform to calculate the coefficients value at these points. In the case of parabolic Radon
transform, however, the distance between successive points are non-uniform and demands another interpolation scheme. As seen from Figure 3, in the inverse transformation, the first column and the diagonal coefficients of $\hat{u}$ correspond to the first and last traces, respectively, and thus there is no need for interpolation. One needs to interpolate the coefficients of the intermediate traces. The algorithm in this paper employs linear interpolation for its efficiency and acceptable accuracy. The more smooth the data under analysis is the better will be the performance of linear interpolation. Generally, seismic data are band-limited both in time and in spatial directions. Therefore, the method works satisfactorily in most situations of seismic studies. However, one may use more accurate interpolation methods for complicated data.

Let $f(\ell_1)$ and $k_p(\ell_2)$ for $\ell_1 = 0, \ldots, N_x - 1$ and $\ell_2 = 0, \ldots, N_p - 1$, be the Cartesian grid points where $\hat{u}$ is defined, then $\hat{v}(f(\ell_1), x(n_x))$ is calculated as

$$\hat{v}(f(\ell_1), x(n_x)) = (1 - \delta)\hat{u}(f(\ell_1), k_p(\ell)) + \delta\hat{u}(f(\ell_1), k_p(\ell + 1))$$

where $\ell = \lfloor 2p_{\max}f(\ell_1)/\phi(x(n_x)) \rfloor$ and $\delta = 2p_{\max}f(\ell_1)/\phi(x(n_x)) - \ell$. The values of $\ell$ and $\delta$ can simply be computed and stored prior to interpolation. This allows a fast implementation of our algorithm via iterative solvers.

Using the formulation given by (22), the Radon coefficients can be regularized in order to increase its resolution by using an appropriate sparsity promoting regularization. Introducing a new variable $\hat{b} = F_1 v$ (which is indeed the $f - x$ domain representation of the data) reduces problem (22) to a simple interpolation problem $S\hat{u} = \hat{b}$, where $\hat{u} = F_2^{-1}u$. Therefore, the problem of finding a sparse solution can be defined as $u = F_2^{-1}u$ where $\hat{u}$ is a solution to the basis pursuit program [8]

$$\arg\min_{\hat{u}} ||F_2^{-1}\hat{u}||_1 \quad \text{subject to} \quad S\hat{u} = \hat{b}.$$  

where $|| \cdot ||_1$ denotes the sum of absolute values. By this program we look for the sparsest Radon panel, in the one-norm sense, whose 2-D FFT at the specified grid points (see Figures 3b and d) matches the $f - x$ coefficients of the data, $b$. However, the presence of uncertainties in the data or inaccuracy in the interpolation may cause $b$ out of range and hence (24) may not have a solution. A better choice is

$$\arg\min_{\hat{u}} ||F_2^{-1}\hat{u}||_1 \quad \text{subject to} \quad ||S\hat{u} - \hat{b}||_2^2 \leq \varepsilon$$

where $\varepsilon$ is the error bound. For linear interpolation, the matrix of $S$ is sparse (there will be two non-zero entries at each row) and when combined with efficient applications of $F_2$/inverse $F_2$ via FFT’s make efficient solution of problem (25) possible via split Bregman iterations [11]. The details of the algorithm is fully analysed in [11]. We just provide the algorithm (Algorithm 1) for the solution of problem (25). In the algorithm shrink is the soft shrinkage function and $d$ is the diagonal elements of $\alpha S^T S + \beta I$. The advantage of the split Bregman iterations employed for solving our problem is that $\alpha S^T S + \beta I$ can be well approximated by its diagonal, where $I$ denotes identity matrix. Therefore, the run time for application of each Bregman iteration is mostly due to 2-D FFT/IFFT’s and hence the total cost at each iteration is $O(N \log N)$ where $N$ is the number of coefficients. Aside from fast convergence of the Bregman iteration a main advantage of it for solving our problem is that here the number of Bregman iterations has the role of regularization parameter. We allow the iterations to proceed until the misfit between predicted and observed data reach a desired predefined tolerance.
Algorithm 1: The split Bregman iterations for sparsity based time-invariant Radon transform by the generalized Fourier slice theorem.

1. Initialize: choose the Bregman parameters $\alpha, \beta > 0$ and set $k = 0$, $\hat{b}_0 = \hat{b} = F_1v, p^0 = q^0 = 0$ and $d = \text{diag}(\alpha S^T S + \beta I)$

2. while $\|S \hat{u} - \hat{b}\|^2_2 > \varepsilon$ do

3. $\hat{u}^{k+1} = \alpha S^T \hat{b}^k + \beta F_2(p^k - q^k)$

4. $p^{k+1} = \text{shrink}(F_2^{-1} \hat{u}^{k+1} + q^k, 1/\beta)$

5. $q^{k+1} = q^k + (F_2^{-1} \hat{u}^{k+1} - p^{k+1})$

6. $b^{k+1} = b^k + \hat{b} - S \hat{u}^{k+1}$

7. $k = k + 1$

8. end

9. $u = F_2^{-1} \hat{u}^{k+1}$

5. Numerical examples. In this section, we present numerical examples from synthetic and field data to show the performance of the proposed method. In all examples tested, we set $N_r = N_t$ and determine $N_p$ from (21) for an upper limit of variables.

| $N_t$ | Backslash | Levinson | PCG-Chan | GFST | speed-up | Error |
|-------|------------|----------|----------|------|----------|-------|
| $2^9$ | 4.6480     | 6.3021   | 1.2668   | 0.0181 | 256      | 348   | 69    | 43e-7 |
| $2^{10}$ | 49.485    | 32.273   | 8.3074   | 0.0615 | 804      | 524   | 135   | 24e-7 |
| $2^{11}$ | 595.83    | 179.41   | 33.672   | 0.2447 | 2434     | 733   | 137   | 11e-7 |
| $2^{12}$ | 7707.0    | 1384.2   | 290.30   | 1.0561 | 7297     | 1310  | 274   | 60e-8 |

5.1. Synthetic data. Two data sets have been simulated in order to examine the functionality of the proposed Radon transforms. The first (Figure 4a) includes linear events and the second (Figure 4b) consists of parabolic events. Both data sets are of dimension $1024 \times 1024$. The absolute value of the $f - x$ representation of the data are shown below each data set. We analyzed the data for $f_{max} = 50$ Hz and the resulting paths of Radon coefficients are shown in green in Figure 4.

Both non-sparse and sparse Radon transforms have been applied to the simulated data using the parameters defined above and the results are shown in Figures 5 and 6 for the linear and parabolic transforms, respectively. The predicted data from the estimated Radon transform is also shown below each figure. It can be seen from our results that for both non-sparse and sparse methods the estimated Radon
coefficients can predict the data with a high degree of accuracy. The non-sparse coefficients, as expected, clearly show energy leakage from the main peaks along two crossing paths. On the other hand, the sparse coefficients are focused and free of smearing. From a computational time viewpoint, however, the non-sparse coefficients have been obtained in 28e-3 seconds while the sparse coefficients have been calculated by 20 Bregman iterations with a total time of 1.2 sec.

We continued by applying the parabolic Radon transform to data of different sizes. Both forward (sparse and non-sparse) and inverse transforms were examined. In order to perform the new method, we employed linear interpolation to calculate the coefficients at desired points as in Figure 3. To show the effect of problem size, all data sets examined were square, $N_t = N_x$, with equal number of samples in the data and model space and the analysis was done for all frequencies up to the Nyquist. We compared the new method with the traditional $f - x$ algorithm to evaluate the least-squares forward transform and employed three different fast linear solvers to solve the corresponding complex Toeplitz systems in 3 for each frequency: the MATLAB backslash operator, the Levinson recursion, and the preconditioned conjugate gradient method with the Chan preconditioner (PCG-Chan) [7]. A constant value has been added to the main diagonal of the Toeplitz matrices to stabilize the inversion in the forward transform. Table 1 shows computational times, speed-up of the proposed method over these three methods, and the mean-squared-error of the results of the proposed method with respect to the results of

| $N_t$ | Number of iterations | Average time per iteration | Total time |
|-------|----------------------|---------------------------|------------|
| $2^9$ | 83                   | 0.0332                    | 2.7516     |
| $2^{10}$ | 70                  | 0.1589                    | 11.122     |
| $2^{11}$ | 65                  | 0.6638                    | 43.150     |
| $2^{12}$ | 58                  | 2.4840                    | 144.07     |

Table 2. Number of iterations, average time per iteration, and total computational time (in sec) for the sparse parabolic Radon transform via Algorithm 1. $N_t$ refers to the time length of the square data matrices for the experiments.

| $N_t$ | Time | Speed-up | Error |
|-------|------|----------|-------|
|       | Direct sums | GFST |       |
| $2^9$ | 0.2844          | 0.0112 | 25  | 12e-5 |
| $2^{10}$ | 2.5050       | 0.0497 | 50  | 11e-6 |
| $2^{11}$ | 17.926      | 0.2139 | 84  | 33e-7 |
| $2^{12}$ | 137.45      | 0.8475 | 162 | 34e-7 |

Table 3. Computational time (in sec) for the inverse parabolic Radon transform via direct sums and the generalized Fourier slice theorem (GFST), speed-up, and mean-squared-error compared to the direct method. $N_t$ refers to the time length of the square data matrices for the experiments.
Figure 4. Simulated 2-D data sets including linear (a) and parabolic (b) events. For both data sets \( N_t = N_x = 1024 \). (c) and (d) are the \( f - x \) representation of the data (absolute value). The green paths show constant wavenumber of the corresponding Radon panel. The proposed non-sparse and sparse Radon transforms have been applied to these data and the results are shown in Figures 5 and 6.

The data has also been transformed by sparse parabolic Radon transform using Algorithm 1. Table 2 shows the number of Bregman iterations, average time per iteration, and total computational time (in sec) required for convergence to a given misfit of \( 10^{-3} \). Here again all results shown were averaged over 20 runs. It is seen that the number of Bregman iteration does not increase with the size of data. Furthermore, a comparison of the run times in Table 2 with those reported in Table 1 shows that the proposed algorithm for sparse transform is even faster than the traditional non-sparse solvers.

The new method has also been tested for inverse parabolic Radon transform and compared with the traditional direct calculations defined in 3. The results are shown in Table 3 for different sizes of data. The inverse transform has been calculated significantly faster by the generalized Fourier slice theorem.

We continued by performing the new method in the presence of noise. The gathers (Figure 4) were contaminated by additive random noise of signal-to-noise-ratio (SNR) -5 dB (Figures 7a and b). Here again we analysed the data for \( f_{max} = \)}
Figure 5. Non-sparse (a) and sparse (b) linear Radon transform of the data shown in Figure 4(a). (c) and (d) show the predicted data. The sparse solution has been obtained via 20 Bregman iterations with computational time = 1.2 seconds.

50 Hz and stopped the Bregman iterations according to the discrepancy principle by assuming known standard deviation of the added noise. The resulting predictive (denoised) data by sparse Radon transform are shown in Figures 7c and d. As seen the algorithm treated noisy data very stably.

5.2. Interpolation of irregular data. Seismic data sets are generally irregularly sampled in the spatial direction due to variable operating conditions during seismic data acquisition. This irregular sampling can limit the effectiveness of data processing and imaging algorithms. In general, acquired traces need to be interpolated to a regular grid prior to start classical processes for signal enhancement and data inversion [9]. The Radon transform is a useful and efficient tool for seismic data regularization. The sampling operator $S$ in Algorithm 1 explicitly depends upon the spacial coordinates of the traces $x$, therefore, when the Radon coefficients have been estimated, the interpolated data can easily be constructed by the inverse Radon transformation to any desired grid points. We use 24 traces of the data shown in Figure 4 collected at random locations (see Figures 8a and b) and used Algorithm 1 to perform sparse Radon transforms. Then, the inverse transformation
Figure 6. Non-sparse (a) and sparse (b) parabolic Radon transforms of the data shown in Figure 4(b). (c) and (d) show the predicted data. The sparse solution has been obtained by 20 Bregman iterations with computational time = 1.2 seconds.

was used to synthesize data at all grid points. The original data before and after the process of reconstruction are shown in Figure 8. It can be seen that the original data sets have been accurately constructed.

5.3. Field data. As a final example, we tested our method on a real CMP gather and compared it with the direct method. The parabolic Radon transform is known as a successful tool for the attenuation of multiple reflections in CMP gathers [12]. Figure 9 shows a field CMP gather after NMO correction applied using the velocity of primary reflections. As can be seen from this figure primary reflections are flatted but multiple reflections show parabolic moveout. Figure 10a shows the sparse parabolic Radon panel generated by the proposed method (300 iterations of Algorithm 1) and Figure 10b shows the corresponding predicted data. As can be seen, the energy of primary reflections are well clustered around $p$ value zero and separated form the multiples energy, enclosed by the rectangle in Figure 10a. We just transformed back the multiples energy (Figure 10c) and subtracted it from the original data to get the primary only wavefield as shown in Figure 10d. The same procedure was applied to the data using traditional direct method and 300 iterations.
of fast iterative shrinkage-thresholding algorithm (FISTA) [4] and the results are shown in Figure 11. As can be seen from Figures 10 and 11, the obtained results by the two methods are very similar. However, the average time per iteration for the proposed method was 0.01 sec while it was 0.6 sec for direct method.

6. Conclusions. We generalized the Fourier slice theorem for computing a general time-invariant Radon transform. The new theorem establishes an analytic expression between the 1-D Fourier coefficients of a function and the 2-D Fourier coefficients of its general Radon transform. This allows efficient computation of Radon integrals via FFT’s and interpolation. Based on linear interpolation scheme, fast implementations of the forward and inverse operators were developed in the sense that they run in $O(N\log N)$ flops, where $N$ is the maximum number of samples in the data space or model space. These operators were utilized for efficient implementation of the sparse Radon transform via a Bregman iterative method. The efficiency and high-accuracy of the proposed algorithm for seismic data processing were shown by using numerical examples from linear and parabolic Radon transforms. The linear interpolation employed here provided satisfactory results for seismic data. However, its applicability to other types of data is questionable and needs further research.
Figure 8. (a) and (b) show decimated seismic data shown in Figure 4, generated by randomly removing 98 percent of the traces (1000 out of 1024). (c) and (d) show reconstructed data by the proposed sparse Radon transforms.

REFERENCES

[1] A. Averbuch, R. Coifman, D. L. Donoho, M. Israeli and Y. Shkonisky, A framework for discrete integral transformation I- the pseudo-polar Fourier transform, *SIAM J. of Scientific Computing*, 30 (2008), 764–784.

[2] A. Averbuch, R. Coifman, D. L. Donoho, M. Israeli, Y. Shkonisky and I. Sedelnikov, A framework for discrete integral transformation II- the 2D discrete Radon transform, *SIAM J. of Scientific Computing*, 30 (2008), 785–803.

[3] S. Basu and Y. Bresler, $O(N^3 \log N)$ backprojection algorithm for the 3D Radon transform, *IEEE Trans. Medical Imaging*, 21 (2002), 76–88.

[4] A. Beck and M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM J. Imaging Sciences*, 2 (2009), 183–202.

[5] G. Beylkin, Discrete Radon transform, *IEEE Transactions on Acoustics, Speech and Signal Processing*, 35 (1987), 162–172.

[6] P. W. Cary, The simplest discrete Radon transform, *68th Annual International Meeting, SEG, Expanded Abstracts*, (1998), 1999–2002.

[7] R. H. Chan and M. K. Ng, Conjugate gradient methods for Toeplitz systems, *SIAM Review*, 38 (1996), 427–482.

[8] S. Chen, D. L. Donoho and M. Saunders, Atomic decomposition by basis pursuit, *SIAM J. of Scientific Computation*, 20 (1998), 33–61.
Figure 9. A field CMP gather, NMO corrected using the velocity of primaries. Note that the primary waves are flatted (with $p$ values clustered around zero) but multiples show parabolic moveouts.

[9] A. Gholami, Nonconvex compressed sensing with frequency mask for seismic data reconstruction and denoising, *Geophysical Prospecting*, 62 (2014), 1389–1405.
[10] A. Gholami, Deconvolutive Radon transform, *Geophysics*, 82 (2017), V117–V125.
[11] T. Goldstein and S. Osher, The split Bregman method for $l_1$ regularized problems, *SIAM J. Imag. Sci.*, 2 (2009), 323–343.
[12] D. Hampson, Inverse velocity stacking for multiple elimination, *56th Annual International Meeting, SEG, Expanded Abstracts*, (1986), 422–424.
[13] P. E. Hart, How the Hough transform was invented, *IEEE Signal Processing Magazine*, 26 (2008), 18–22.
[14] P. Herrmann, T. Mojesky, M. Magesan and P. Hugonnet, De-aliased, high-resolution Radon transforms, *70th Annual International Meeting, SEG, Expanded Abstracts*, 19 (2000), 1953–1956.
[15] K. Hokstad and R. Sollie, 3D surface-related multiple elimination using parabolic sparse inversion, *Geophysics*, 71 (2006), V145–V152.
[16] J. Hsieh, *Computed Tomography Principles, Design, Artifacts, and Recent Advances*, 2nd Edition, 2nd revised ed., SPIE Publications, Bellingham, WA, 2015.
[17] J. Hu, S. Fomel, L. Demanet and L. Ying, A fast butterfly algorithm for generalized Radon transforms, *Geophysics*, 78 (2013), U41–U51.
[18] C. Kostov, Toeplitz structure in slant-stack inversion, *SEG Technical Program Expanded Abstracts*, (1990), 1618–1621.
[19] W. Lu, An accelerated sparse time-invariant Radon transform in the mixed frequency-time domain based on iterative 2D model shrinkage, *Geophysics*, 78 (2013), V147–V155.
[20] R. M. Mersereau, Recovering multidimensional signals from their projections, *Computer Graphics and Image Processing*, 2 (1973), 179–195.
[21] V. Nikitin, F. Andersson, M. Carlsson and A. Duchkov, Fast hyperbolic Radon transform by log-polar convolutions, *SEG Technical Program Expanded Abstracts*, (2016), 4534–4539.
[22] M. D. Sacchi and M. Porsani, Fast high resolution parabolic Radon transform, *SEG Technical Program Expanded Abstracts*, (1999), 1477–1480.
[23] M. Sacchi and T. Ulrych, High-resolution velocity gathers and offset space reconstruction, *Geophysics*, 60 (1995), 1169–1177.
[24] M. Sacchi and T. Ulrych, Improving resolution of Radon operators using a model re-weighted least squares procedure, *Journal of Seismic Exploration*, 4 (1995), 315–328.
Figure 10. (a) Sparse parabolic Radon transform of the field data (Figure 7), obtained by generalized Fourier slice theorem, and (b) the corresponding predicted data. (c) Separated multiples, corresponding to the coefficients enclosed by the rectangle in (a). (d) Separated primaries, obtained by subtraction of (c) from the original wavefield.

[25] M. Schonewille and A. Duijndam, Parabolic Radon transform, sampling and efficiency, Geophysics, 66 (2001), 667–678.
[26] J. R. Thorson and J. F. Claerbout, Velocity-stack and slant-stack stochastic inversion, Geophysics, 50 (1985), 2727–2741.
[27] D. Trad, T. Ulrych and M. Sacchi, Latest views of the sparse Radon transform, Geophysics, 68 (2003), 386–399.
[28] S. Treitel, P. Gutowski and D. Wagner, Plane-wave decomposition of seismograms, Geophysics, 47 (1982), 1375–1401.
[29] O. Yilmaz, Velocity stack processing, SEG Technical Program Expanded Abstracts, (1988), 1013–1016.
[30] O. Yilmaz, Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data, 2nd edition, Investigations in geophysics, Society of Exploration Geophysicists, Tulsa, OK, 2001.
Figure 11. (a) Sparse parabolic Radon transform of the field data (Figure 7), obtained by direct method and FISTA, and (b) the corresponding predicted data. (c) Separated multiples, corresponding to the coefficients enclosed by the rectangle in (a). (d) Separated primaries, obtained by subtraction of (c) from the original wavefield.