A comparative study over improved fast iterative shrinkage-thresholding algorithms: an application to seismic data reconstruction

HAMID REZA KHATAMI[1], MOHAMMAD ALI RIAHI[2], MOHAMMAD MAHDI ABEDI[3] AND AFSHIN AKBARI DEHKHARGANI[1]

1 Petroleum, Mining, and Materials Engineering, Islamic Azad University, Central Tehran Branch, Tehran, Iran

2 Institute of Geophysics, University of Tehran, North Kargar Ave., Tehran 1439955961, Iran (mariahi@ut.ac.ir)

3 BCAM - Basque Center for Applied Mathematics, Alameda de Mazarredo 14, 48009 Bilbao, Spain

Received: April 22, 2023; Revised: July 24, 2023; Accepted: January 9, 2024

ABSTRACT

Seismic data reconstruction is a crucial process involving the restoration of missing or corrupted traces to create a uniform dataset for subsequent data processing. Various factors such as equipment failures, and surface obstacles, result in irregularly located or corrupted traces. The absence of these traces can compromise the quality and accuracy of the resulting image. To address this issue, the Nonuniform Fast Fourier Transform (NUFFT) method is employed to reconstruct missing traces in datasets with non-uniformly sampled data. It works by interpolating the non-uniformly sampled data onto a regular grid, enabling the traditional Fast Fourier Transform application for data recovery. This interpolation process is adjusted using a kernel function to account for non-uniform sampling and reduce aliasing artifacts. The outcome is a collection of Fourier coefficients that can be utilized to reconstruct missing or incomplete parts of data. This problem is transformed into a linear constraint problem, which is efficiently solved using the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA). In this study, we explore various techniques aimed at improving the convergence of FISTA, collectively referred to as improved FISTA methods. To validate the NUFFT+FISTA method for data reconstruction, we conducted numerical tests using 3D and 2D synthetic datasets, as well as field data. These tests show the advantages of the Greedy-FISTA in terms of convergence rate and affirm the accuracy of this approach in filling missing data traces.

Keywords: data reconstruction, FISTA, missing traces, NUFFT

1. INTRODUCTION

Irregular distribution of recorded traces in land operations is mainly due to the presence of urban facilities, lakes, and rivers, and in marine surveys is due to technical issues in

streamers. For many processing steps, the reconstruction of data with missing traces has a particularly important role (*Liu et al., 2015*).

Several reconstruction techniques have been proposed, each with its own strengths and weaknesses. They include filter and transform-based methods (*Abma and Kabir, 2005*), wavefield interpolation (*Mazzucchelli et al., 1998*), and deep learning for data reconstruction (*Abedi and Pardo, 2022*).

Wave equation methods are useful tools as they help to fill in missing observations. However, they require a deep understanding of subsurface velocity distribution and involve complex full wavefield calculations, which limits their practical use. Predictive error filtering methods offer a simpler approach by using low-frequency data components on a regular grid to predict and fill in high-frequency data. For more information on the background of these methods, please refer to *Lan et al. (2022)*.

By applying sparsity regularization to the coefficients in a transform domain, the transform domain methods can reconstruct under-sampled and incomplete data. The data reconstruction based on sparse transforms such as the Fourier, curvelet, contourlet, and seislet transforms has been extensively discussed in the literature (*Eslami and Radha, 2006*; *Hennenfent et al., 2010*; *Gilles et al., 2010*; *Ciabarri et al., 2015*; *Liu et al., 2016*; *Galvis et al., 2022*; *Tian and Qin, 2022*). *Meng et al. (2008)* employed the Nonuniform Fourier Transform (NUFT) in a data reconstruction problem. In order to solve it, they convert the problem into a least squares form. The NUFT is applicable to irregularly sampled signals, in contrast to the classical Fourier transform, which presumes that the signal is sampled at regular intervals.

A fast approach for calculating the NUFT of a signal is called the Nonuniform Fast Fourier Transform (NUFFT) (*Duijndam and Schonewille, 1999*). The NUFFT is an extension of the Fast Fourier Transform (FFT), which is used to compute the Fourier Transform of a uniformly sampled signal. The NUFFT is a multi-step process that includes interpolating the non-uniformly sampled data onto a uniform grid, multiplying the interpolated data by a kernel function, and then applying the FFT to the resulting uniformly sampled data. The NUFFT has many practical applications, including image reconstruction, data compression, and signal processing in fields such as Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) (*Greengard and Lee, 2004*).

To solve the reconstruction of the missing trace problem using NUFFT, we utilize the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) method. *Beck and Teboulle (2009)* introduced FISTA to improve the standard Iterative Shrinkage-Thresholding Algorithm (ISTA), enabling quicker convergence rates. FISTA is an iterative method for solving linear convex optimization problems. It is a first-order optimization procedure, which means that each iteration only requires the gradient of the objective function to be computed. The FISTA algorithm operates by iteratively revising a sequence of optimal solution estimates. At each iteration, the algorithm computes a gradient descent step followed by a shrinkage operation that ensures the solution is sparse. The shrinkage operation is carried out by applying a threshold derived from the gradient descent step. The quantity of thresholding is determined by a parameter known as the proximal parameter, which controls the solution sparsity (*Beck and Teboulle, 2009*).

The capacity to handle irregularly recorded data, which is prevalent in land surveys, is one of the benefits of NUFFT-based data reconstruction. In comparison to other reconstruction techniques, it may also reduce computational costs, making it appropriate

for large and 3D datasets. To achieve accurate and effective reconstruction, NUFFT-based reconstruction needs careful parameter selection and optimization (*Liu et al, 2015*). We utilized advanced FISTA algorithms to improve the optimization process related to the Non-Uniform Fast Fourier Transform (NUFFT) missing trace recovery problem. Our approach was tested on both synthetic and real datasets, and it showed remarkable proficiency in reconstructing the missing traces. The results highlight the superior performance achieved by using the improved FISTA algorithms.

Following a brief introduction to data reconstruction, the FISTA and its improved versions are discussed. Then, the newly introduced approaches are evaluated through numerical testing and compared based on various aspects.

2. DATA RECONSTRUCTION

To reconstruct multi-channel data, we use a mathematical model expressed as follows:

$$\boldsymbol{b} = \mathbf{S}\boldsymbol{d}\,, \tag{1}$$

where vector $\boldsymbol{d} \in \mathbb{R}^{n_r \times n_t}$ denotes the recorded field data; vector $\boldsymbol{b} \in \mathbb{R}^{n_s \times n_t}$ represents the incomplete field data, which are often under-sampled; and $\mathbf{S} \in \mathbb{R}^{n_s \times n_r}$ is the sampling operator; n_r, n_t and n_s denote the number of temporal recorded, trace-patch and incomplete samples, respectively, contained in each patch. The sampling operator $\mathbf{S}$ is an under-determined matrix:

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{2}$$

To solve the under-determined problem shown in Eq. (1), it is converted into an optimization problem with a regularization term

$$\arg\min_{d}\left(\frac{1}{2}\left\|\mathbf{S}\boldsymbol{d}-\boldsymbol{b}\right\|_2^2 + \lambda Reg\left(\boldsymbol{d}\right)\right), \tag{3}$$

where $Reg(\cdot)$ denotes the regularization operator that contains prior information on the model, $\|\cdot\|_2^2$ denotes the square of L_2 norm, λ is the regularization parameter weighing the fidelity term $\left\|\mathbf{S}\boldsymbol{d}-\boldsymbol{b}\right\|$ and the model constraint term $Reg(\cdot)$.

The regularization optimization based on sparsity promotion constraints is given as (*Hennenfent et al., 2010*; *Wang et al., 2011*; *Bai et al., 2014*):

$$\begin{cases} \arg\min_{\boldsymbol{x}}\left(\frac{1}{2}\left\|\mathbf{S}\mathbf{F}^{-1}\boldsymbol{x}-\boldsymbol{b}\right\|_2^2 + \lambda\left\|\boldsymbol{x}\right\|_1\right), \\ \boldsymbol{x} = \tilde{\mathbf{F}}\boldsymbol{b}\,, \end{cases} \tag{4}$$

where $\tilde{\mathbf{F}}$ is the non-uniform Fourier transform (NUFFT) operator, $\mathbf{F}^{-1}$ is the inverse of the Fourier transform, $\boldsymbol{x}$ contains the coefficients of the signal in the frequency domain, and $\|\cdot\|_1$ denotes the L_1 norm. Since $\boldsymbol{b}$ can represent irregularly sampled data, the NUFFT is employed to transfer it to the frequency domain. The non-uniform discrete Fourier transform transforms a sequence of N complex numbers $u_0, \ldots, u_{N-1}$ into another sequence of complex numbers $U_0, \ldots, U_{N-1}$ defined by

$$U_k = \sum_{n=0}^{N-1} u_n \mathrm{e}^{-2\pi \mathrm{i} d_n f_k} \quad , \quad 0 \le k \le N-1, \tag{5}$$

where $d_n \in [0,1]$ ($n = 0, \ldots, N-1$) are sample points and $f_k \in [0, N]$ ($k = 0, \ldots, N-1$) are frequencies. The second term in Eq. (4) is in the frequency domain, while the first term is in the time domain. In this way, by solving the minimization problem of Eq. (4), we can reconstruct the full data $\hat{\boldsymbol{d}} \approx \boldsymbol{d}$.

For simplicity, we define $\mathbf{A} = \mathbf{S}\mathbf{F}^{-1}$ in Eq. (4), leading to

$$\arg\min_{\boldsymbol{x}} \left(\|\mathbf{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \lambda \|\boldsymbol{x}\|_1 \right). \tag{6}$$

Equation (6) is a linear constraint problem that can be solved using the FISTA method (*Beck and Teboulle, 2009*).

3. FISTA ALGORITHM

The convex optimization in Eq. (6) can be solved via interior point methods (*Ben-Tal and Nemirovski, 2001*). However, in most applications, the problem is not only large scale (may approach millions of decision variables) but also involves dense matrices, which frequently prevents the usage and potential benefit of complex interior point algorithms. Therefore, inexpensive gradient-based algorithms have been proposed to solve the problem. *Figueiredo et al. (2007)* reformulated Eq. (6) as a box-constrained quadratic problem solved by a gradient projection algorithm. ISTA, which involves matrix-vector multiplication followed by a shrinkage/soft-threshold step in each iteration, is one of the most widely used techniques for solving Eq. (6) (*Figueiredo and Nowak, 2003*; *Daubechies et al., 2004*; *Chambolle et al., 1998*). To be more precise, ISTA's standard procedure is

$$\boldsymbol{x}_{k+1} = \mathcal{T}_\lambda \left[\boldsymbol{x}_k - 2\alpha \mathbf{A}^{\mathrm{T}} \left(\boldsymbol{b} - \mathbf{A}\boldsymbol{x}_k \right) \right], \tag{7}$$

where α is an appropriate step size and $\mathcal{T}$ is the shrinkage operator (e.g., *Mallat, 1999*; *Yang and Gao, 2015*) given by the following expression

$$\mathcal{T}_\gamma \left(\boldsymbol{s} \right) = \max\left(1 - \frac{\gamma}{|s|}, 0 \right) \boldsymbol{s}. \tag{8}$$

Here $\boldsymbol{s}$ is the desired vector that the shrinkage operator is operating on. Since ISTA is a slow method (*Bredies and Lorenz, 2008*), FISTA (Fast ISTA) was proposed by *Beck and Teboulle (2009)* to generate a faster convergence rate. The implementation steps of Eq. (6) using FISTA are shown in the Algorithm 1.

Algorithm 1: FISTA method for data reconstruction in the Fourier domain.

Input:

$\mathbf{A} = \mathbf{S}\mathbf{F}^{-1}$, $\mathbf{A}^{\mathrm{T}} = \tilde{\mathbf{F}}\mathbf{S}^{\mathrm{T}}$, $\boldsymbol{b}$ (under-sampled data), $\boldsymbol{x}_0 = \boldsymbol{x}_1 = \boldsymbol{y}_1 = \tilde{\mathbf{F}}\boldsymbol{b}$,
$\alpha_1 = t_1 = 1$, and L is Lipschitz constant ($1/L$ is step-size).

Output:

$\hat{\boldsymbol{d}} = \mathbf{F}^{-1}(\boldsymbol{x})$ (reconstructed data)

For $k = 1, 2, 3, \ldots$

$$t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$$

$$\alpha_{k+1} = \frac{t_k - 1}{t_{k+1}}$$

$$y_{k+1} = x_k + \alpha_{k+1}(x_k - x_{k-1})$$

$$x_{k+1} = \mathcal{T}_{\lambda/L}\left(\left(\mathbf{I} - \frac{1}{L}\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)y_{k+1} + \frac{1}{L}\mathbf{A}^{\mathrm{T}}\boldsymbol{b}\right)$$

Here, $\mathbf{I}$ is an identity matrix. The main difference between the above algorithm and ISTA is that the iterative shrinkage operator is not employed on the previous point x_{k-1}, but rather at the point y_{k+1}, which uses a specific linear combination of the previous two points $\{x_{k-1}, x_{k-2}\}$. The main computational effort in both ISTA and FISTA remains the same.

3.1. Modified FISTA

The first modification of FISTA is in the update of t_k (Algorithm 1), which replaces the constant values with three parameters p, q and r. The modified FISTA scheme is the same as the Algorithm 1 with a modification as $t_{k+1} = \left(p + \sqrt{q + rt_k^2}\right)\Big/2$. Therefore, there are three hyperparameters to set: p, $q > 0$ and $r \in (0, 4]$. When r is strictly smaller than 4, Algorithm 2 is simply a variant of the inertial forward-backward method (refer to *Liang et al., 2017*) for more details on its convergence properties). A benefit of the parameters p, q and r in FISTA-Mod is faster convergence rate (*Liang et al., 2022*).

3.2. α-FISTA

The degree of freedom provided by the three hyperparameters of modified FISTA allows the design of more advanced schemes, which can make the convergence rate faster in practice. The following modification is an extension of modified-FISTA, where the values of p and q determine the value of r as follows:

$$f(\gamma, \alpha, p, q) = r = 4(1-p) + \frac{4p\left(1-\sqrt{\gamma\alpha}\right)}{1+\sqrt{\gamma\alpha}} + \frac{4\gamma\alpha\left(p^2-q\right)}{\left(1+\sqrt{\gamma\alpha}\right)^2}, \tag{9}$$

where $\gamma \le 1/L$ and $\alpha \ge 0$. This method is called α-FISTA. Since $f(\gamma, \alpha = 0, p, q) = 4$, the α-FISTA provides the $\mathcal{O}\left(1/k^2\right)$ convergence rate for the non-strongly convex case.

3.3. Rada-FISTA

The α-FISTA aims to prevent the oscillating behavior sometimes seen in FISTA applications. This feature works for strong convexity, and it only examines the situation in which strong convexity is available, which is rarely the case in practice. It takes time to estimate the robust convexity in general. Consequently, an efficient method of estimation is also required. To resolve this issue, the next modified version of FISTA has a restarting adaptive scheme (Algorithm 2), incorporating the restarting technique of *O'Donoghue and Candes (2015)* in α-FISTA.

Algorithm 2: Restarting and adaptive α-FISTA (Rada-FISTA).

Input:

$\mathbf{A} = \mathbf{S}\mathbf{F}^{-1}$, $\mathbf{A}^{\mathrm{T}} = \tilde{\mathbf{F}}\mathbf{S}^{\mathrm{T}}$, $\boldsymbol{b}$ (under-sampled data), $\boldsymbol{x}_0 = \boldsymbol{x}_1 = \boldsymbol{y}_1 = \tilde{\mathbf{F}}\boldsymbol{b}$, $\alpha_1 = t_1 = 1$, $p, q \in (0, 1]$, $\xi < 1$, $r \in (0, 4]$ and Lipschitz constant L ($1/L$ is step-size).

Output:

$\hat{\boldsymbol{d}} = \mathbf{F}^{-1}(\boldsymbol{x})$ (reconstructed data)

For k = 1, 2, 3, …

$$t_{k+1} = \frac{p + \sqrt{q + 4\alpha_k^2}}{2}$$

$$\alpha_{k+1} = \frac{t_k - 1}{t_{k+1}}$$

$$y_{k+1} = x_k + \alpha_{k+1}(x_k - x_{k-1})$$

$$x_{k+1} = \mathcal{T}_{\lambda/L}\left(\left(\mathbf{I} - \frac{1}{L}\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)y_{k+1} + \frac{1}{L}\mathbf{A}^{\mathrm{T}}\boldsymbol{b}\right)$$

Restart if $(y_k - x_{k+1})^{\mathrm{T}}(x_{k+1} - x_k) \ge 0$: $r = \xi r$ and $y_k = x_k$ and $t_k = 1$

This strategy can effectively prevent the oscillatory behavior.

3.4. Greedy-FISTA

The FISTA schemes has the oscillation issue when converging to the solution, because $\alpha_k \to 1$. For the restarting scheme, resetting t_k to 1 will cause α_k to increase from 0, approach 1, and cause the subsequent oscillation, after which the scheme will recommence. If the time between two restarts can be cut down with such a loop, further acceleration might be acquired. It turns out that this can be done utilizing a constant α_k (near to or equal 1). Algorithm 3 shows how the Greedy-FISTA is implemented (*Liang et al., 2022*).

Algorithm 3: Greedy-FISTA method.

Input:

$\mathbf{A} = \mathbf{S}\mathbf{F}^{-1}$, $\mathbf{A}^{\mathrm{T}} = \tilde{\mathbf{F}}\mathbf{S}^{\mathrm{T}}$, $\boldsymbol{b}$ (under-sampled data), $\boldsymbol{x}_0 = \boldsymbol{x}_1 = \boldsymbol{y}_1 = \tilde{\mathbf{F}}\boldsymbol{b}$, $s \geq 1$, $\gamma \in \left[1/L, 2/L\right)$, $\xi < 1$, and Lipschitz constant L ($1/L$ is step-size).

Output:

$\hat{\boldsymbol{d}} = \mathbf{F}^{-1}\left(\boldsymbol{x}\right)$ (reconstructed data)

For $k = 1, 2, 3, \ldots$

$$y_{k+1} = x_k + \left(x_k - x_{k-1}\right)$$

$$x_{k+1} = \mathcal{T}_{\gamma\lambda}\left(\left(\mathbf{I} - \gamma\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)y_{k+1} + \gamma\mathbf{A}^{\mathrm{T}}\boldsymbol{b}\right)$$

Restart if $\left(y_k - x_{k+1}\right)^{\mathrm{T}}\left(x_{k+1} - x_k\right) \geq 0 : y_k = x_k$

Safeguard if $\left\|x_{k+1} - x_k\right\| \geq s\left\|x_1 - x_0\right\| 0 : \ \gamma = \max\left\{\xi\gamma, 1/L\right\}$

A larger step-size (larger than $1/L$) is chosen for γ, which can further shorten the oscillation period. A safeguard step is provided to ensure convergence because a big step size can cause divergence. The safety feature eventually leads to $\gamma = 1/L$ if it is used frequently. In practice, $\gamma \in \left[1/L, 1.3/L\right]$ provides faster performance than Rada-FISTA and restarting FISTA (*Liang et al., 2022*).

3.5. Modified APG

The following optimization technique is based on the accelerated gradient method, known as the accelerated proximal gradient (*Nesterov, 2004*). To solve Eq. (6), one can use modified APG, which combines the APG and modified-FISTA methods. We recommend interested readers to Section 6 of *Liang et al. (2022)* for a more thorough explanation of the modified-APG approach.

Algorithm 4: Modified APG procedure.

Input:

$\mathbf{A} = \mathbf{S}\mathbf{F}^{-1}$, $\mathbf{A}^{\mathrm{T}} = \tilde{\mathbf{F}}\mathbf{S}^{\mathrm{T}}$, $\boldsymbol{b}$ (under-sampled data), $\boldsymbol{x}_0 = \boldsymbol{x}_1 = \boldsymbol{y}_1 = \tilde{\mathbf{F}}\boldsymbol{b}$, $\theta_0 \in [0,1]$, $\tau = \sigma\alpha/L$, $\sigma \in [0,1]$, and Lipschitz constant L ($1/L$ is step-size).

Output:

$\hat{\boldsymbol{d}} = \mathbf{F}^{-1}(\boldsymbol{x})$ (reconstructed data)

For k = 1, 2, 3, …

$$\theta_k = \frac{2}{\left(\sigma - \frac{\tau}{\theta_{k-1}^2}\right) + \sqrt{\left(\sigma - \frac{\tau}{\theta_{k-1}^2}\right)^2 + 4}}$$

$$\alpha_k = \frac{\theta_{k-1}(1-\theta_{k-1})}{\theta_{k-1}^2 + \theta_k}$$

$$y_k = x_k + \alpha_k (x_k - x_{k-1})$$

$$x_{k+1} = \mathcal{T}_{\gamma\lambda}\left(\left(\mathbf{I} - \frac{1}{L}\mathbf{A}^{\mathrm{T}}\mathbf{A}\right) y_{k+1} + \frac{1}{L}\mathbf{A}^{\mathrm{T}}\boldsymbol{b}\right)$$

Table 1 compares different aspects of the studied methods. The hyperparameters indicate the parameters that the user must define prior to the algorithm execution. Even though the modified version contains more hyperparameters, it can deliver much more accurate results and faster convergence if the parameters are properly tuned. A decent parameter selection requires trial and error. The choice in Table 1 was initially made by trial and error, while, we had no a priori information in hand about the optimal value of the parameters.

Table 1. Comparison of the presented methods. See text for the meaning of symbols.

Method	Hyperparameters	Restart Option	Safeguard Option
FISTA	λ, L	---	---
Modified FISTA	$p > 0, q > 0, r \in (0,4], \lambda, L$	---	---
α-FISTA	$p > 0, q > 0, \alpha \geq 0, \gamma \leq 1/L, \lambda, L$	---	---
Rada-FISTA	$p \in (0,1], q \in (0,1], \xi < 1, r \in (0,4], \lambda, L$	✓	---
Greedy-FISTA	$\gamma \in [1/L, 2/L), \xi < 1, s > 1, \lambda, L$	✓	✓
Modified APG	$\theta_0 \in [0,], \sigma \in [0,], \lambda, L$	---	---

4. NUMERICAL TEST

We apply the introduced and proposed methods to the reconstruction of data using the NUFFT method in order to compare them. Initially, a simple 2D synthetic test is examined, followed by another more complex synthetic data from wide azimuth SEG Salt Model C3 (*https://wiki.seg.org/wiki/SEG_C3_WA*). A set of synthetic 3D data is also evaluated to examine the performance of the methods. The mentioned 2D data are shown in Fig. 1.

We compare the methodologies for recovering the synthetic data's missing traces. Figure 2 depicts the three instances in with 30, 50, and 70% randomly missing traces, respectively. Missing traces are recovered using the proposed NUFFT method along with the aforementioned optimization techniques. The convergence criteria of the algorithm are determined by the square root of the mean error between the actual and reconstructed data $MSE < 2.5 \times 10^{-5}$. Figure 3 illustrates the convergence rates of all approaches in Fig. 2. All parameters are set through trial and error, and the optimal configuration for each method is utilized. Our procedure to obtain the optimal parameter is to start with a simple FISTA, then fix its optimal value to be used in the improved FISTA schemes. By implementing the improved FISTA using optimal values obtained from simple FISTA, a suitable value for the new parameter will be achieved. The values in Table 2, which displays the parameters values for each test, seem nearly optimal for the reconstruction problem and are a good starting point for any new dataset.

Figure 3 demonstrates the convergence rate of the introduced methodologies. Original FISTA algorithm has the slowest convergence, while Greedy-FISTA outperforms other methods. Figure 4 illustrates the reconstructed data from Fig. 2 using the NUFFT and Greedy-FISTA methods. Even in the rough case, where 70% percent of traces are lacking (Fig. 2c), the missing traces are perfectly reconstructed (Fig. 4c).

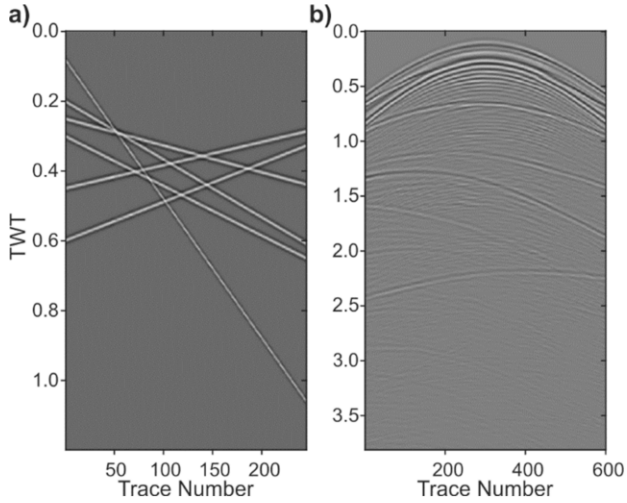

Fig. 1. **a)** Synthetic 2D data, **b)** SEG synthetic data; *TWT* stands for two-way travel time.

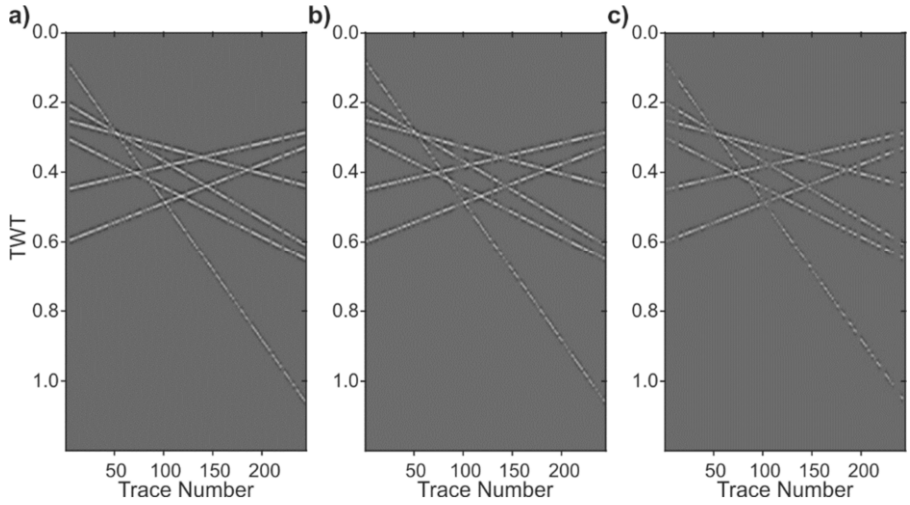

Fig. 2. Synthetic 2D data, shown in Fig. 1a, with randomly removed **a)** 30%, **b)** 50% and **c)** 70% of the traces; *TWT* stands for two-way travel time.

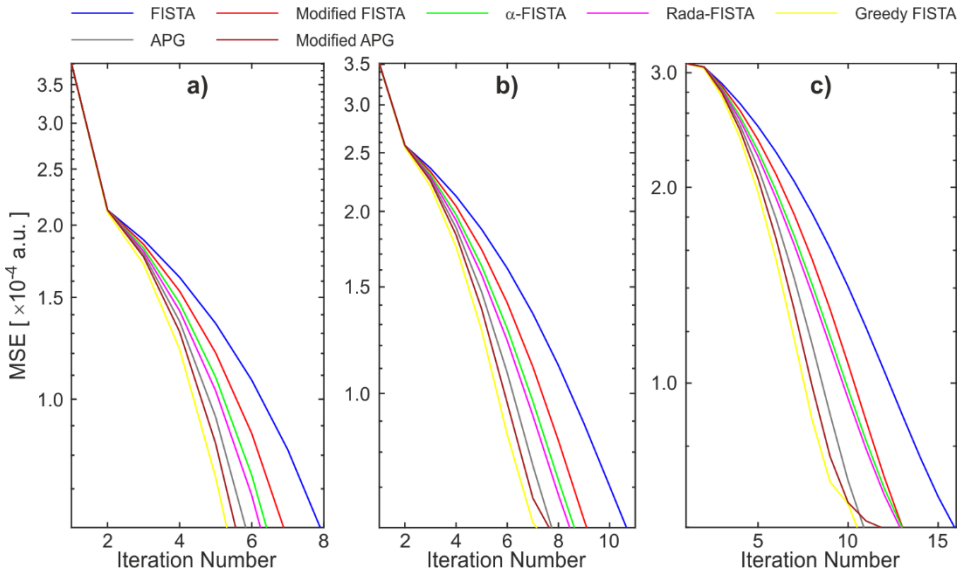

Fig. 3. Respective convergence rates for reconstruction of data shown in Fig. 2a–c, expressed using mean square error (*MSE*, convergence criterium: $MSE = 2.5 \times 10^{-5}$).

It is necessary to mention that the Fourier transform can handle linear or quasi-linear events. In the presence of curved or very complex events, the data is divided into small windows (Fig. 5, synthetic data and 10 field data). Fourier transform can be obtained by

Table 2. Optimal parameters obtained using the improved FISTA method on synthetic 2D data (Test 1) and complex synthetic data (Test 2) (obtained by trial and error).

Parameter	Test 1	Test 2	Parameter	Test 1	Test 2
λ	10	10	r	4	4
α	1	1	ξ	1.1	1.1
L	1	1	s	100	100
p	5	10	σ	0.0001	0.0001
q	10	100	θ_0	0.001	0.001

constructing subdivided windows in advance. The generalized Gabor transform of a signal is regarded as a windowed inverse Fourier transform (*Rebollo-Neira and Fernandez-Rubio, 1999*).

The second test is performed on complex synthetic data shown in Fig. 1b. In Fig. 5, 30%, 50% and 70% of traces are eliminated to generate the input data for the reconstruction schemes. Figure 6 shows the results from modified and enhanced optimization methods. The convergence criteria are determined by $MSE < 2.5 \times 10^{-5}$. Again, the Greedy-FISTA algorithm outperforms the other algorithm, while the original FISTA algorithm has the slowest convergence rate. The NUFFT resulted well in interpolating and not extrapolating. This issue is seen in most of our tests. Therefore, it can be considered a drawback of this method. However, we have to admit that the noisy trace reconstruction in the boundary of the model exists in cases where many traces are missing which is far from real cases for the field data, as shown later.

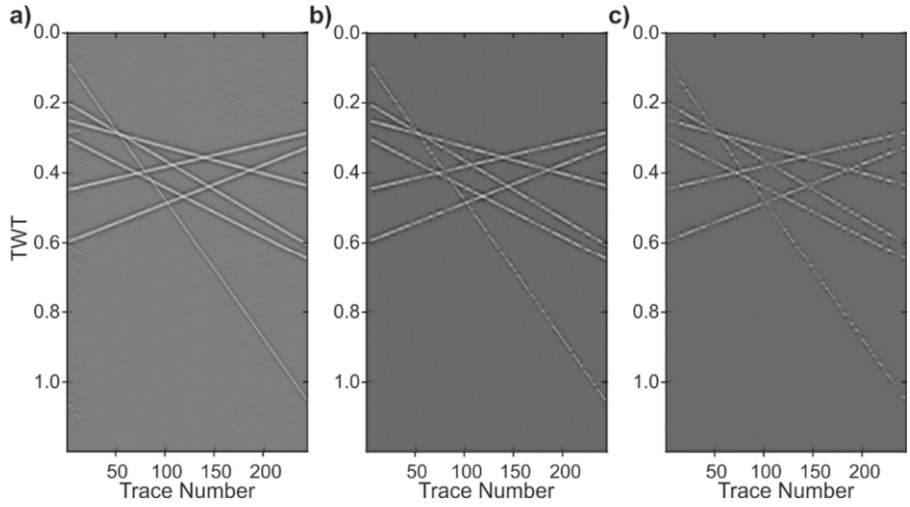

Fig. 4. **a)** Reconstruction of data in Fig. 2c (70% of traces are lacking), and results of reconstruction using **b)** NUFFT and **c)** Greedy-FISTA method; *TWT* stands for two-way travel time.

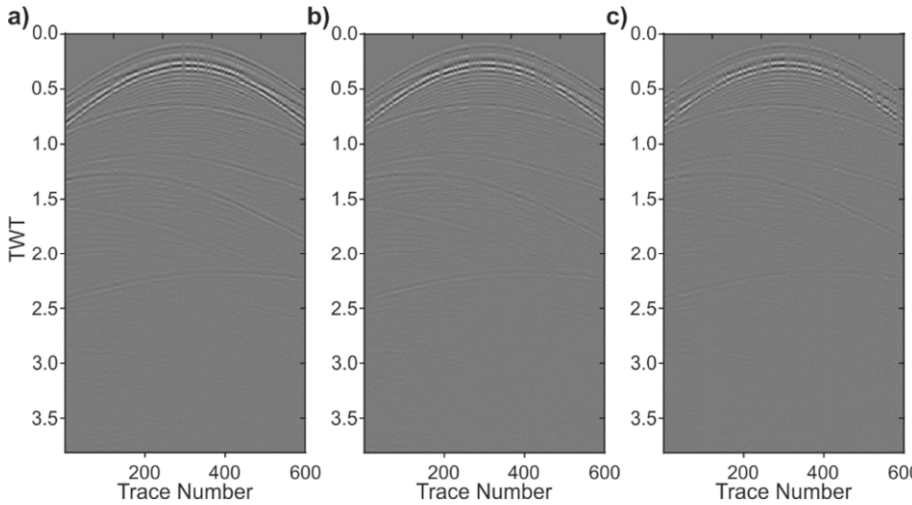

Fig. 5. Synthetic SEG data, shown in Fig. 1b, and 10 field data added, with randomly removed **a)** 30%, **b)** 50% and **c)** 70% of the traces; *TWT* stands for two-way travel time.

Figure 7 shows the results of reconstruction of data shown in Fig. 5 using the NUFFT and Greedy-FISTA methods, and the improved FISTA method demonstrates a satisfactory recovery of lost traces. For both experiments, the original FISTA algorithm has the slowest convergence, while the Greedy-FISTA algorithm outperforms other methods (Figs 3 and 6). Every modification introduced in the previous section accelerates the convergence rate. Both Greedy-FISTA and modified-APG converge in nearly half the number of iterations compared to the original FISTA algorithm. Plots of Fig. 8 show, respectively, zoomed upper sections of Figs 5 and 7.

Next, we apply the Greedy-FISTA to a 3D synthetic dataset to demonstrate the efficacy. The 3D P-velocity model is illustrated in Fig. 9a. Figure 9b shows the shot record corresponding to a source (at 0.5×0.5 km) on the surface, and receivers are distributed uniformly on the surface with a 10-m interval. Figure 9c depicts the shot record with 50% of the total traces removed. The output of Greedy-FISTA is in Fig. 9d. It accurately and efficiently reconstructs the missing traces in a 3D dataset.

To evaluate the effectiveness of the presented approach on real data, we use the Greedy-FISTA for the reconstruction of real-field data displayed in Fig. 10. The field data consists of about 500 traces with a 50-m trace interval and a time sampling rate of 2 ms. It shows that Greedy-FISTA can effectively reconstruct the real-field data using the NUFFT approach.

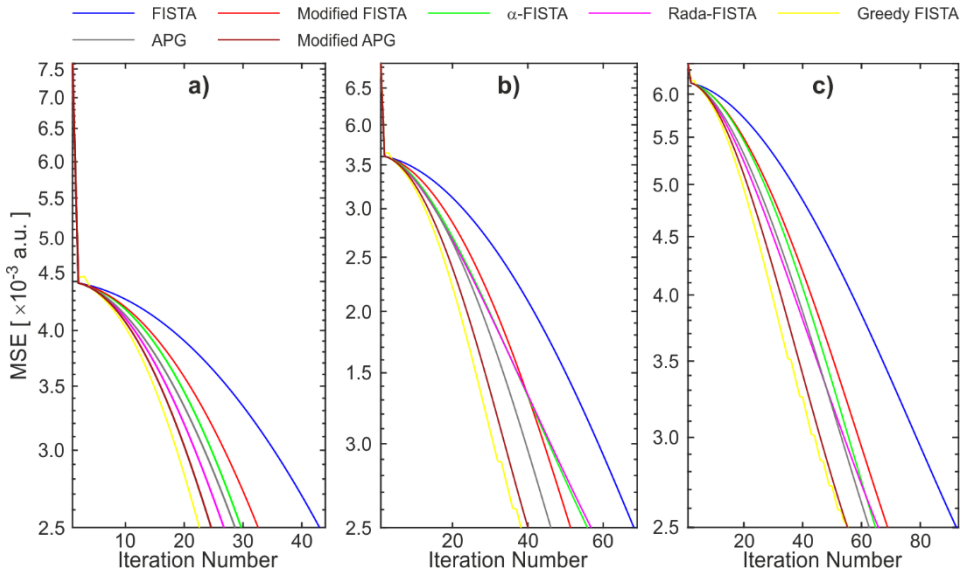

Fig. 6. The same as in Fig. 3, but for data shown in Fig. 5a–c.

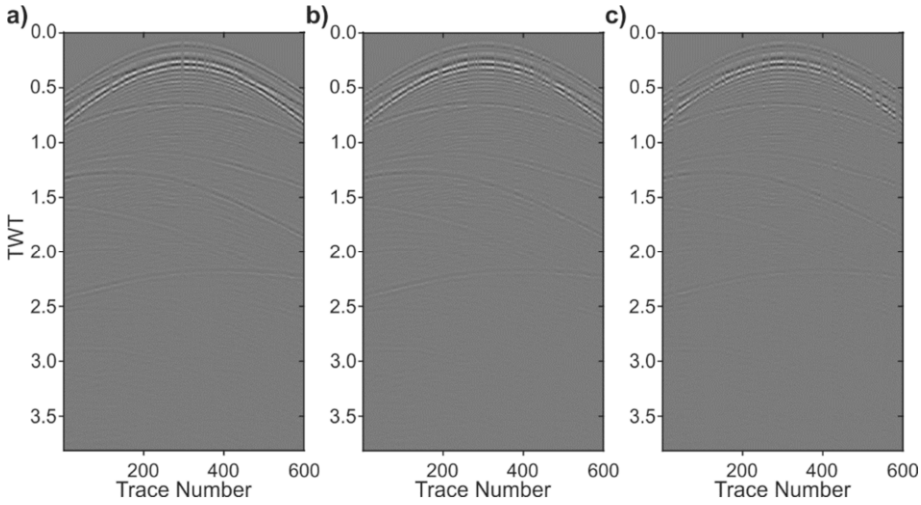

Fig. 7. **a)** Reconstruction of data in Fig. 5c (70% of traces are lacking), and results of reconstruction using **b)** NUFFT and **c)** Greedy-FISTA method; *TWT* stands for two-way travel time.

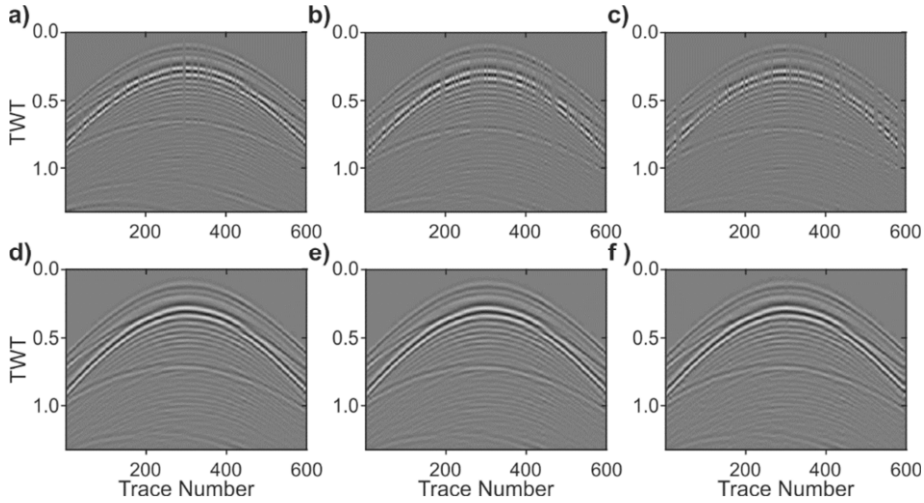

Fig. 8. **a)–c)** Zoomed sections of the respective upper parts of Fig. 5 (incomplete input data), and **d)–f)** zoomed sections of the respective upper parts of Fig. 7 (reconstructed data); *TWT* stands for two-way travel time.

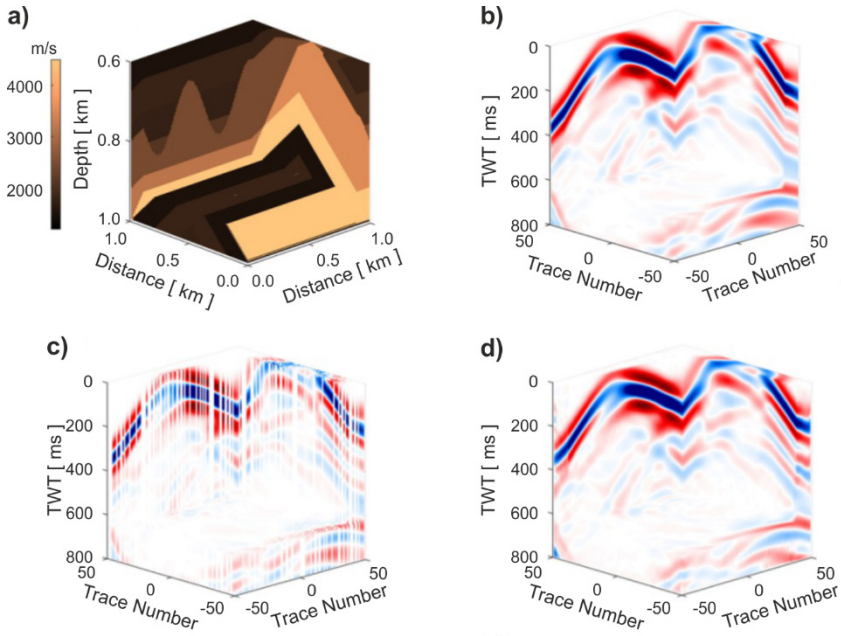

Fig. 9. **a)** 3D P-velocity synthetics model, **b)** true data, **c)** data with 50% missing traces, and **d)** data reconstructed with the NUFFT and Greedy-FISTA method; *TWT* stands for two-way travel time.

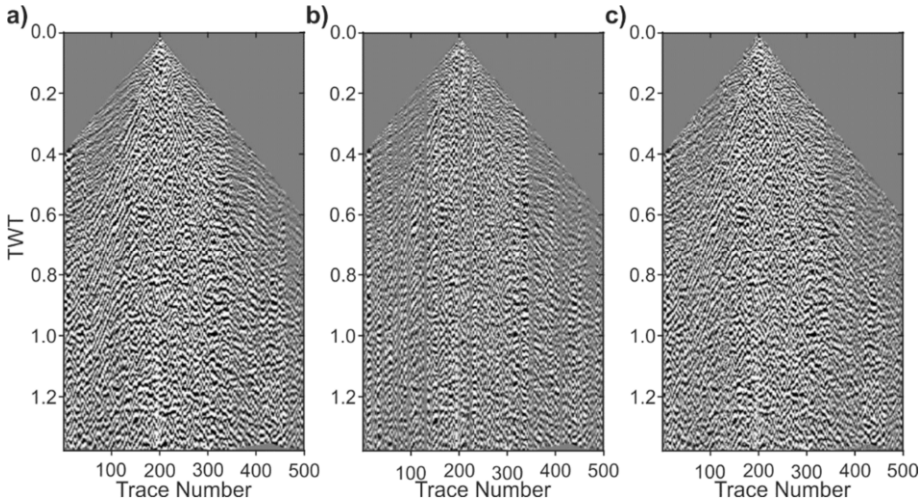

Fig. 10. **a)** Shot gather field data before randomly removing traces, **b)** the shot gather with 30% of randomly removed traces, **c)** reconstruction of the removed traces from shot-gather using the Greedy-FISTA approach; *TWT* stands for two-way travel time.

5. CONCLUSIONS

We show that the NUFFT method can be used for reconstructing missing traces in shot records when employing an appropriate optimization strategy. We use NUFFT for more generality. Although for regularly sampled data with missing traces simple FFT can be employed, in real land data with irregular trace intervals NUFFT should be used. Among the various enhanced and modified versions of the FISTA method used for trace reconstruction with NUFFT, the Greedy-FISTA approach stands out as the most promising. It exhibits superior convergence rates compared to other methods and demonstrates its effectiveness in reconstructing both simple and complex synthetic as well as real field data. One of its significant advantages is its higher convergence rate which in our examples was near half the number of iterations required for convergence of other methods, especially when dealing with large 3D datasets. This method addresses real-world data challenges and provides promising results.

References

Abedi M.M. and Pardo D., 2022. A multidirectional deep neural network for self-supervised reconstruction of seismic data. *IEEE Trans. Geosci. Remote Sensing*, **60**, https://doi.org/10.1109/TGRS.2022.3227212

Abma R. and Kabir N., 2005. Comparisons of interpolation methods. *The Leading Edge*, **24**, 984–989

Beck A. and Teboulle M., 2009. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, **2**, 183–202

Ben-Tal A. and Nemirovski A., 2001. *Lectures on Modern Convex Optimization: Analysis, Algorithms, Engineering Applications*. MPS-SIAM Series on Optimization. SIAM, Philadelphia, PA

Bredies K. and Lorenz D., 2008. Linear convergence of iterative soft-thresholding. *J. Fourier Anal. Appl.*, **14**, 813–837, https://doi.org/10.1007/s00041-008-9041-1

Chambolle A., De Vore R.A., Lee N.Y. and Lucier B.J., 1998. Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Trans. Image Process.*, **7**, 319–335

Ciabarri F., Mazzotti A., Stucchi E. and Bienati N., 2015. Appraisal problem in the 3D least squares Fourier seismic data reconstruction. *Geophys. Prospect.*, **63**, 296–314

Daubechies I., Defrise M. and De Mol C., 2004. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.*, **57**, 1413–1457

Duijndam A.J.W. and Schonewille M.A., 1999. Nonuniform fast Fourier transform. *Geophysics*, **64**, 539–551

Eslami R. and Radha H., 2006. Translation-invariant contourlet transform and its application to image denoising. *IEEE Trans. Image Process.*, **15**, 3362–3374

Figueiredo M.A. and Nowak R.D., 2003. An EM algorithm for wavelet-based image restoration. *IEEE Trans. Image Process.*, **12**, 906–916

Figueiredo M.A., Nowak R.D. and Wright S.J., 2007. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE J. Sel. Top. Signal Process.*, **1**, 586–597

Galvis L., Ramírez J.M., Vargas E., Villarreal O., Agudelo W. and Arguello H., 2020. Reconstruction of 2D seismic wavefields from nonuniformly sampled sources. *IS&T International Symposium on Electronic Imaging 2020*, https://doi.org/10.2352/ISSN.2470-1173.2020.14.COIMG-307

Gilles H., Fenelon L. and Herrmann F.J., 2010. Nonequispaced curvelet transform for seismic data reconstruction: A sparsity-promoting approach. *Geophysics*, **75**, WB203–WB210

Greengard L. and Lee J.Y., 2004. Accelerating the nonuniform fast Fourier transform. *SIAM Rev.*, **46**, 443–454

Lan N.Y., Zhang F.C., and Yin X.Y., 2022. Seismic data reconstruction based on low dimensional manifold model. *Pet. Sci.*, **19**, 518–533

Liang J., Fadili J. and Peyré G., 2017. Activity identification and local linear convergence of forward--backward-type methods. *SIAM J. Optim.*, **27**, 408–437

Liang J., Luo T. and Schonlieb C.B., 2022. Improving "Fast Iterative Shrinkage-Thresholding Algorithm": faster, smarter, and greedier. *SIAM J. Sci. Comput.*, **44**, A1069–A1091

Liu W., Cao S., Gan S., Chen Y., Zu S. and Jin Z., 2016. One-step slope estimation for dealiased seismic data reconstruction via iterative seislet thresholding. *IEEE Geosci. Remote Sens. Lett.*, **13**, 1462–1466

Liu W., Cao S., Li G. and He Y., 2015. Reconstruction of seismic data with missing traces based on local random sampling and curvelet transform. *J. Appl. Geophys.*, **115**, 129–139

Mallat S., 1999. *A Wavelet Tour of Signal Processing*. Second Edition. Elsevier/Academic Press, Amsterdam, The Netherlands

Mazzucchelli P., Rocca F., Spagnolini U. and Spitz S., 1998. Wavefield interpolation - continuation or prediction filter techniques? Expanded Abstract. *60th EAGE Conference and Exhibition*, https://doi.org/10.3997/2214-4609.201408278

Meng X.H., Guo L.H., Zhang Z.F., Li S.L. and Zhou J.J., 2008. Reconstruction of seismic data with least squares inversion based on nonuniform fast Fourier transform. *Chinese J. Geophys.*, **51**, 235–241 (in Chinese)

Nesterov Y., 2004. *Introductory Lectures on Convex Optimization: A Basic Course*. Springer Science+Business Media, New York, https://doi.org/10.1007/978-1-4419-8853-9

O'Donoghue B. and Candes E., 2015. Adaptive restart for accelerated gradient schemes. *Found. Comput. Math.*, **15**, 715–732

Rebollo-Neira L. and Fernandez-Rubio J., 1999. The continuous wavelet transform as a maximum entropy solution of the corresponding inverse problem. *IEEE Trans. Signal Process.*, **47**, 2046–2050. https://doi.org/10.1109/78.771053

Tian L. and Qin S., 2022. Seismic data interpolation by the projected iterative soft-threshold algorithm for tight frame. *IPMV'22: Proceedings of the 4th International Conference on Image Processing and Machine Vision*, 84–87, Association for Computing Machinery,New York, NY, https://doi.org/10.1145/3529446.3529460

Yang P.L. and Gao J.H., 2015. Enhanced irregular seismic interpolation using approximate shrinkage operator and Fourier redundancy. *J. Appl. Geophys.*, **116**, 43–50