

Fixed-Point Proximity Algorithms Solving an Incomplete Fourier Transform Model for Seismic Wavefield Modeling *

Tingting Wu[†]

Lixin Shen[‡]

Yuesheng Xu[§]

Abstract

Seismic wavefield modeling is an important tool for the seismic interpretation. We consider modeling the wavefield in the frequency domain. This requires to solve a sequence of Helmholtz equations of wave numbers governed by the Nyquist sampling theorem. Inevitably, we have to solve Helmholtz equations of large wave numbers, which is a challenging task numerically. To address this issue, we develop two methods for modeling the wavefield in the frequency domain to obtain an alias-free result using lower frequencies of a number fewer than typically required by the Nyquist sampling theorem. Specifically, we introduce two ℓ^1 regularization models to deal with incomplete Fourier transforms, which arise from seismic wavefield modeling in the frequency domain, and propose a new sampling technique to avoid solving the Helmholtz equations of large wave numbers. In terms of the fixed-point equation via the proximity operator of the ℓ^1 norm, we characterize solutions of the two ℓ^1 regularization models and develop fixed-point algorithms to solve these two models. Numerical experiments are conducted on seismic data to test the approximation accuracy and the computational efficiency of the proposed methods. Numerical results show that the proposed methods are accurate, robust and efficient in modeling seismic wavefield in the frequency domain with only a few low frequencies.

Key Words. Compressed sensing, seismic wavefield modeling, incomplete Fourier transform, proximity algorithms.

1 Introduction

Seismic wavefield modeling is an important tool in geological hypothesis testing including synthetic seismograms, crosshole tomography, and seismic migration. It can be carried out in both the time-space domain and the frequency-space domain [14, 26]. The time-domain finite-difference technique for the generation of synthetic seismograms has achieved considerable success in waveform crosshole tomography and seismic reverse-time migration. Finite-difference frequency-domain modeling for the generation of synthetic seismograms and crosshole tomography has been an active field of

*T. Wu was supported in part by the Natural Science Foundation of China under grants 11971276 and 11771257, and Shandong Province Higher Educational Science and Technology Program under grant J18KA221. L. Shen was supported in part by the National Science Foundation under grant DMS-1913039. Y. Xu was supported by the National Science Foundation under grant DMS-1912958 and by the Natural Science Foundation of China under grant 11771464.

[†]School of Mathematics and Statistics, Shandong Normal University, Jinan 250014, P. R. China. tingtingwu@sdsu.edu.cn.

[‡]Department of Mathematics, Syracuse University, Syracuse, NY 13244, USA. lshen03@syr.edu. All correspondence should be sent to this author.

[§]Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA 23529, USA. y1xu@odu.edu.

research since 1980s. The frequency domain method was pioneered by Lysmei and Drake in [19]. The fundamental advantages of frequency domain methods were clearly pointed out in papers [15, 21, 26, 27, 28, 30] and the references cited therein. The key advantage of a frequency domain approach is that direct solution methods are available that allow solutions to the wave equation for additional source positions to be obtained at a minimal extra cost. Each frequency can be computed independently, which favors parallel computing. Furthermore, frequency domain modeling allows accurate modeling of attenuation effects, and allows optimal spatial discretization intervals to be chosen at each frequency.

Seismic wavefield modeling in the frequency domain can be formulated as a frequency domain boundary value problem (cf. [9, 15, 18]). Mathematically, we convert the wave equation, which models seismic wave propagation, to a collection of Helmholtz equations of different frequencies (wave numbers) by using the continuous Fourier transform. When solutions to the Helmholtz equations for all frequencies are available, we can obtain the corresponding time domain results by the inverse discrete Fourier transform (IDFT) (cf. [4, 14, 29]), in which case the Nyquist-Shannon criterion must be satisfied. Otherwise, aliasing artifacts will be visible in the resulting modeled wavefield. As a result, in order to represent a solution of the wave equation accurately for a given time interval, we are required to solve many Helmholtz equations of different wave numbers including large wave numbers.

Solving the Helmholtz equation of large wave numbers is a challenging task in the field of computational mathematics [2]. As the wave number becomes large, the solution of the Helmholtz equation becomes highly oscillatory and the numerical solution of this equation requires high computational complexity. Hence, there is a growing interest in developing numerical methods whose computational complexity increases only moderately as the wave number increases (see, [9, 12, 15, 30, 31]). However, there are still some cases for which iterative methods fail to converge for all frequencies (especially for high frequencies), or the computer storage is exhausted while direct methods are used to solve the resulting linear system for all frequencies. Thus, we shall have to invert an incomplete Fourier transform data.

The incomplete Fourier transform usually has the form of an underdetermined linear system. Hence, to reconstruct a digital signal or image from incomplete Fourier transform data is an inverse problem, and it is an ill-posed problem. There are infinitely many candidate signals or images to satisfy the related linear system. But one could perhaps imagine a way out by relying on realistic models of objects which naturally exist. Compressed sensing [11] is a method to invert incomplete Fourier transforms. It is based on the idea that a signal can be reconstructed from a very small number of measurements, provided that these measurements are obtained in a correct basis. Candès, Romberg and Tao [5, 6] studied the stable recovery of sparse signals from incomplete and inaccurate measurements, and they gave conditions for stable recovery. Goldstein and Osher [13] applied the Split-Bregman method to deal with the inverting incomplete Fourier transform problem. Lin and Herrmann demonstrated that, information redundancy exists for the seismic wavefield itself in the frequency domain [17]. Furthermore, Lebed and Herrmann [16] showed numerically that, in the problem of recovering seismic signals from missing temporal frequencies, one can expect to see better results from transforms that are more spread in the frequency domain and hence more incoherent, for example, using shift-invariant wavelets. In the paper [18], based on the shift-invariant wavelet transform and random sampling, the problem of recovering seismic signals from missing temporal frequencies is resolved by solving an ℓ^1 -regularization minimization. In addition, the resulting ℓ^1 -regularization minimization is dealt with an iterative soft thresholding method. However, as random sampling is used in [18], solving the Helmholtz equation with large wave numbers is still needed.

To avoid solving the Helmholtz equation with large wave numbers, in this paper we first propose

a new sampling method for seismic wavefield modeling in the frequency domain, that is, only the solutions of the Helmholtz equations for a band of low frequencies are required. Then we use the ℓ^1 regularization technique to reconstruct the wave solution from the solutions of the Helmholtz equations of low frequencies. A tight frame wavelet transform is used as the “sparsifying” transform that maps a signal (e.g., an image) to a sparse vector. Based on the wavelet transform, we propose two regularization models for inverting incomplete Fourier transforms, and finally solve these models by using fixed-point proximity algorithms.

This paper is organized in five sections. In Section 2, we review seismic wavefield modeling in the frequency domain and describe the mathematical formulation of the incomplete Fourier transform arising from the context of wavefield modeling. A new sampling method is also presented in this section. In Section 3, we present two regularization models based on the wavelet transform and propose two convergent algorithms to compute the models based on fixed-point equations. In Section 4, five numerical examples are presented to show that the proposed methods are robust and efficient for seismic wavefield modeling in the frequency domain. Finally, Section 5 concludes this paper.

2 An incomplete Fourier transform model for seismic wavefield modeling in the frequency domain

The seismic wave propagation is often modeled by the 2D acoustic wave equation. In the time domain the equation has the form

$$\frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} - \Delta u = g, \quad \text{on } \mathbb{R}^2, \quad (2.1)$$

where u , v and g denote, respectively, the unknown pressure of the wave field, the background velocity and the source term in the medium. Both u and g are functions in the spatial-time space $\mathbb{R}^2 \times \mathbb{R}_+$, while v is a function in the spatial space $\mathbb{R}^2$. By using the Fourier transform, one may convert the wave equation as a family of Helmholtz equations. Upon solving these Helmholtz equations and converting back to the solution of the wave equation, one can model the propagation of seismic wavefield. This is the frequency approach for modeling the seismic wave propagation.

We present below the 2D acoustic wave equation in the frequency domain. To this end, for a function φ of time $t \in \mathbb{R}$, its continuous Fourier transform at frequency $f \in \mathbb{R}$ is given as

$$\widehat{\varphi}(f) := \int_{\mathbb{R}} \varphi(t) e^{-i2\pi f t} dt.$$

The function φ can be reconstructed from $\widehat{\varphi}$ through the inverse Fourier transform. For any $(x, z) \in \mathbb{R}^2$, we denote by $\widehat{u}(x, z, f)$ and $\widehat{g}(x, z, f)$, respectively, the continuous Fourier transforms at the frequency f of the functions $u(x, z, \cdot)$ and $g(x, z, \cdot)$ which appear in (2.1). With the Fourier transform, the acoustic wave equation (2.1) is converted to the well-known Helmholtz equation

$$-\Delta \widehat{u} - k^2 \widehat{u} = \widehat{g}, \quad (2.2)$$

where k is the wave number defined by $k := \frac{2\pi f}{v}$, and f denotes the frequency in Hz. The solution $u(x, z, t)$ of the acoustic wave equation (2.1) may be obtained by the inverse Fourier transform from the solutions $\widehat{u}(x, z, f)$ of the Helmholtz equation, for all $f \in \mathbb{R}$. Therefore, the fundamental problem for solving the acoustic wave model in the frequency domain is to solve the family of the Helmholtz equations (2.2) for all $f \in \mathbb{R}$.

Take the synthetic seismogram for a fixed point (x_r, z_r) in the computational domain as an example. To obtain the synthetic seismogram for this fixed point, we need to evaluate the solution of the wave equation (2.1) at the point (x_r, z_r) . In the context of seismic modeling, two commonly used source functions, the first order derivative of the Gaussian function and the Ricker wavelet, approximately have limited spectrums. We take the first order derivative of the Gaussian function as an example, which is expressed as

$$G(t, t_0, \tilde{\alpha}) = -2\tilde{\alpha}(t - t_0) \exp(-\tilde{\alpha}(t - t_0)^2). \quad (2.3)$$

Clearly, its Fourier transform has the form

$$\widehat{G}(f, t_0, \tilde{\alpha}) = 2\sqrt{\frac{\pi}{\tilde{\alpha}}} \pi f \exp\left(-\frac{\pi^2 f^2}{\tilde{\alpha}}\right) [\sin(2\pi f t_0) + i \cos(2\pi f t_0)]. \quad (2.4)$$

In Figure 1, we plot the first order derivative of the Gaussian function (2.3) and its frequency spectrum for the case when $t_0 = 0.3$ and $\tilde{\alpha} = 200$. From Figure 1 (b), we find that the function (2.3) is approximately band limited, and its highest frequency can be approximately seen as $15H_z$. In general, we assume that there exists $T > 0$ such that the solution u of the wave equation (2.1)

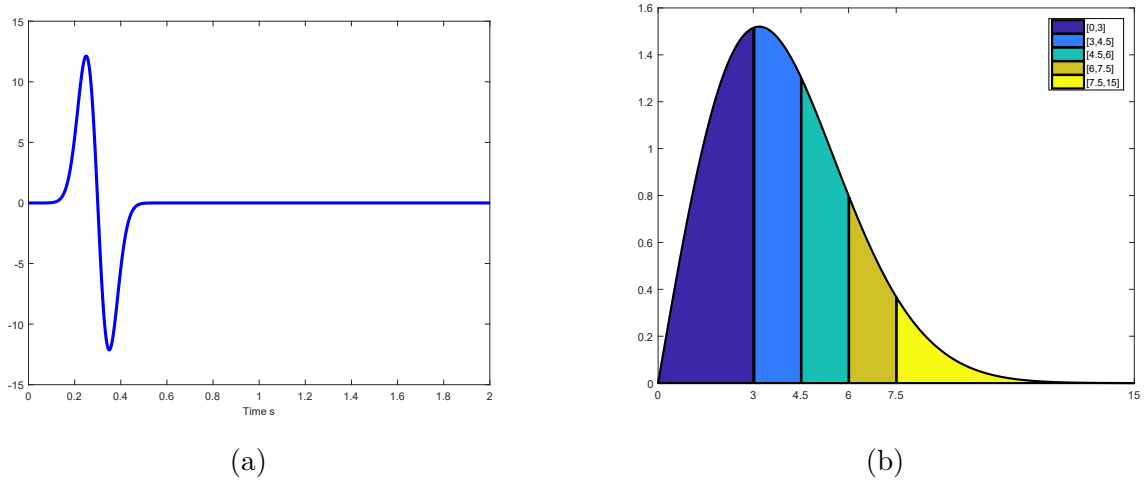


Figure 1: (a) the first derivative of the Gaussian function; (b) the frequency spectrum for the first derivative of the Gaussian function in the frequency domain.

at the point (x_r, z_r) satisfies the condition $u(x_r, z_r, t) = 0$ for all $t \notin [0, T]$. The Fourier transform of $u(x_r, z_r, \cdot)$ is written as

$$\widehat{u}(x_r, z_r, f) = \int_0^T u(x_r, z_r, t) e^{-i2\pi f t} dt. \quad (2.5)$$

For convenient exposition, we shall use the following notation $u_r(t) := u(x_r, z_r, t)$ and $\widehat{u}_r(f) := \widehat{u}(x_r, z_r, f)$.

Now, we illustrate how we obtain the values of $u_r(t)$ at N equally spaced points taken in the interval $[0, T]$ in the frequency modeling. Applying the rectangle method to equation (2.5) yields

$$\widehat{u}_r(f) \approx \lambda \sum_{n=0}^{N-1} u_r(\lambda n) e^{-i2\pi f \lambda n}, \quad \text{for any } f \in \mathbb{R}, \quad (2.6)$$

where $\lambda := \frac{T}{N}$. As the source function in the seismic case approximately has a limited spectrum (see, Figure 1 (b)), we denote by f_{max} the approximate highest frequency of the source. Therefore, we assume that equation (2.6) holds approximately for $f \in [0, f_{max}]$. Then, following the Nyquist sampling theorem, the frequency step size Δf satisfies

$$\Delta f \leq \frac{1}{T}. \quad (2.7)$$

Hence, we choose $f = \frac{m}{T}$, $m = 0, 1, \dots, N-1$, and the corresponding frequency step size $\Delta f = \frac{1}{T}$. Also, let

$$\mathbf{u} := \begin{bmatrix} u_r(0) \\ u_r\left(\frac{T}{N}\right) \\ \vdots \\ u_r\left(\frac{(N-1)T}{N}\right) \end{bmatrix} \quad \text{and} \quad \hat{\mathbf{u}} := \begin{bmatrix} \hat{u}_r(0) \\ \hat{u}_r\left(\frac{1}{T}\right) \\ \vdots \\ \hat{u}_r\left(\frac{N-1}{T}\right) \end{bmatrix}.$$

Hence, from equation (2.6), we have that

$$\hat{\mathbf{u}} \approx \lambda \sqrt{N} F \mathbf{u},$$

where F is the $N \times N$ discrete Fourier transform (DFT) matrix with $\frac{1}{\sqrt{N}} \exp\left(-i \frac{2\pi(m-1)(n-1)}{N}\right)$ as its (m, n) -th entry. We can reconstruct $\mathbf{u}$ from $\hat{\mathbf{u}}$ with the inverse discrete Fourier transform (IDFT), that is,

$$\mathbf{u} \approx \frac{1}{\lambda} \frac{1}{\sqrt{N}} F^\top \hat{\mathbf{u}}.$$

Here, the notation $A^\top$ denotes the conjugate transpose of A . Note that each component of $\hat{\mathbf{u}}$ is obtained by solving the Helmholtz equation.

In computation, the Helmholtz equation is usually discretized by finite difference [9, 15] or finite element methods [2, 12]. For a large scale problem, solving the Helmholtz equation with large wave numbers is a difficult task (see, [2, 14]). We even have difficulty in obtaining $\hat{u}_r(f)$ for frequencies f satisfying $f \leq f_{max}$. Therefore, only part of components of $\hat{\mathbf{u}}$ are available. We use the vector $\mathbf{r}_{obs}$ to record the available components of $\hat{\mathbf{u}}$. The relation between $\hat{\mathbf{u}}$ and $\mathbf{r}_{obs}$ is linked by a “row selector” matrix R that comprises a subset of the rows of an identity matrix such that

$$R \hat{\mathbf{u}} \approx \mathbf{r}_{obs}. \quad (2.8)$$

Let $\mathbf{r} := \frac{1}{\sqrt{N}} \mathbf{r}_{obs}$. With the DFT, the relation of $\mathbf{u}$ and $\mathbf{r}$ can be presented as

$$\lambda R F \mathbf{u} \approx \mathbf{r}, \quad (2.9)$$

which is an incomplete Fourier transform system.

Different row selector matrices correspond to different ways of sampling. Since solving the Helmholtz equation with a large wave number is a challenging task, we propose to sample only lower frequencies in recovering the seismic wavefield. In this case we choose to use the following “row selector” matrix

$$R := \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & \dots & 0 \end{bmatrix},$$

where the number of its rows is less than $\frac{N}{2}$ with N being the number of its columns. The choice of R depends on the value of the highest frequency we compute.

3 Regularization Models for Inverting Incomplete Fourier Transforms

Reconstructing a digital signal or an image from incomplete Fourier data is an inverse and ill-posed problem. In practice, noise also exists in the observed data. To solve these problems, regularization methods are usually used. In this section, we propose two regularization models for inverting incomplete Fourier transforms and develop algorithms for solving these models.

We assume that the underlying signal $\mathbf{u} \in \mathbb{R}^N$ in (2.9) has a sparse representation with a redundant tight framelet system in $\mathbb{R}^N$ [16]. Suppose that $W \in \mathbb{R}^{\tilde{N} \times N}$ (with $\tilde{N} := LN$) satisfies $W^\top W = I$, where L denotes a positive integer and I is the $N \times N$ identity matrix. Then W is a tight frame in $\mathbb{R}^N$. The components of the vector $W\mathbf{u}$ are called framelet coefficients representing $\mathbf{u}$. Below we present a concrete tight framelet matrix W . To this end, for a given filter with length $2J+1$

$$h := [h(-J), h(-J+1), \dots, h(-1), h(0), h(1), \dots, h(J-1), h(J)],$$

we define an $N \times N$ matrix $S(h)$ by

$$(S(h))_{i,j} = \begin{cases} h(i-j) + h(i+j-1), & \text{if } i+j \leq J+1; \\ h(i-j) + h(-1-(2N-i-j)), & \text{if } i+j \geq 2N-J+1; \\ h(i-j), & \text{otherwise.} \end{cases}$$

We shall use the matrix $S(h)$ for specific filter h to construct W . Next we choose to use specific piecewise linear tight frame filters. This frame has $h_0^{(0)} := [\frac{1}{4}, \frac{1}{2}, \frac{1}{4}]$ as its low-pass filter and $h_1^{(0)} := [\frac{\sqrt{2}}{4}, 0, -\frac{\sqrt{2}}{4}]$ and $h_2^{(0)} := [-\frac{1}{4}, \frac{1}{2}, -\frac{1}{4}]$ as its high-pass filters. Performing a multilevel wavelet decomposition without downsampling using this tight frame demands $h_0^{(\ell)}$, the filter $h_0^{(0)}$ at level ℓ :

$$h_0^{(\ell)} = \left[\frac{1}{4}, \underbrace{0, \dots, 0}_{2^{\ell-1}-1}, \frac{1}{2}, \underbrace{0, \dots, 0}_{2^{\ell-1}-1}, \frac{1}{4} \right].$$

The $h_1^{(\ell)}$ and $h_2^{(\ell)}$, the filters corresponding to $h_1^{(0)}$ and $h_2^{(0)}$ at level ℓ , respectively, can be constructed similarly. For $i = 0, 1, 2$, we let $H_i^{(\ell)} := S(h_i^{(\ell)})$. The tight framelet matrix W is then defined by

$$W := \begin{bmatrix} \prod_{\ell=0}^{L-1} H_0^{(L-\ell)} \\ H_1^{(L)} \prod_{\ell=1}^{L-1} H_0^{(L-\ell)} \\ H_2^{(L)} \prod_{\ell=1}^{L-1} H_0^{(L-\ell)} \\ \vdots \\ H_1^{(2)} H_0^{(1)} \\ H_2^{(2)} H_0^{(1)} \\ H_1^{(1)} \\ H_2^{(1)} \end{bmatrix}. \quad (3.1)$$

It can be verified that $W^\top W = I$.

With the help of a tight framelet matrix W of size $\tilde{N} \times N$, we rewrite (2.9). To this end, we define

$$\mathbf{y} := \lambda W \mathbf{u} \quad \text{and} \quad K := RFW^\top.$$

Then model (2.9) becomes

$$K\mathbf{y} \approx \mathbf{r}. \quad (3.2)$$

We shall develop regularization methods to “invert” equation (3.2) to obtain the vector $\mathbf{y}$.

The following ℓ^1 -norm based optimization model is proposed for the recovery of the sparse vector $\mathbf{y}$ from (3.2)

$$\tilde{\mathbf{y}} = \operatorname{argmin} \left\{ C(\mathbf{y}) : \mathbf{y} \in \mathbb{R}^{\tilde{N}} \right\}, \quad (3.3)$$

where

$$C(\mathbf{y}) := \|\Gamma\mathbf{y}\|_1 + \varphi(K\mathbf{y} - \mathbf{r}) \quad (3.4)$$

with Γ being a diagonal matrix with non-negative diagonal entries and φ a function measuring the distance between $K\mathbf{y}$ and $\mathbf{r}$. Recent results show that if the matrix K obeys the uniform uncertainty principle, and $\mathbf{y}$ is sufficiently sparse, then the solution $\tilde{\mathbf{y}}$ to equation (3.3) is exactly $\mathbf{y}$ (cf. [7, 11]). As the tight framelet system, which is a kind of shift-invariant wavelet transforms, is used in this paper, the two conditions above are satisfied [18].

Two different choices for the function φ are considered. When the observation $\mathbf{r}$ is contaminated with the Gaussian noise, we choose $\varphi := \frac{1}{2}\|\cdot\|_2^2$. When the observation $\mathbf{r}$ is contaminated with non-Gaussian noise, we choose φ as the Moreau envelope of the ℓ^1 -norm. For a positive number τ and a convex function ψ in $\mathbb{R}^d$, the Moreau envelope of ψ with index τ at $\mathbf{z} \in \mathbb{R}^d$ is defined as

$$\operatorname{env}_{\tau\psi}(\mathbf{z}) := \min \left\{ \frac{1}{2\tau}\|\mathbf{x} - \mathbf{z}\|_2^2 + \psi(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^d \right\}, \quad (3.5)$$

which is again a convex function in $\mathbb{R}^d$ (see, e.g., [24]). Moreover, there exists a unique vector, denoted by $\operatorname{prox}_{\tau\psi}(\mathbf{z})$, in $\mathbb{R}^d$ such that the minimum value $\operatorname{env}_{\tau\psi}(\mathbf{z})$ can be achieved, that is,

$$\operatorname{env}_{\tau\psi}(\mathbf{z}) = \frac{1}{2\tau}\|\operatorname{prox}_{\tau\psi}(\mathbf{z}) - \mathbf{z}\|_2^2 + \psi(\operatorname{prox}_{\tau\psi}(\mathbf{z})).$$

We call $\operatorname{prox}_{\tau\psi} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ the proximity operator of ψ with index τ [22, 23]. The second choice of φ in (3.3) is then $\operatorname{env}_{\tau\|\cdot\|_1}$, whose i th entry at $z \in \mathbb{R}^{\tilde{N}}$ is, via the definition of the Moreau envelope,

$$(\varphi(\mathbf{z}))_i = \begin{cases} \frac{1}{2\tau}|z_i|^2, & |z_i| \leq \tau; \\ |z_i| - \frac{\tau}{2}, & \text{otherwise.} \end{cases}$$

We can easily see that this function φ approaches to the ℓ^1 -norm $\|\cdot\|_1$ as the parameter τ goes to zero. It has been well demonstrated in [1, 25] that the ℓ^1 -norm fidelity term is particularly effective for handling non-Gaussian additive noise such as impulsive noise and the Laplace noise.

The functions φ for both choices are differentiable. Using their differentiability together with a result from [8, 10, 20, 22], we can characterize the solutions of model (3.3) in terms of a fixed-point equation. Precisely, if $\mathbf{y} \in \mathbb{R}^{\tilde{N}}$ is a solution to model (3.3), then for any positive number δ

$$\mathbf{y} = \operatorname{prox}_{\frac{1}{\delta}\|\cdot\|_1 \circ \Gamma} \left(\mathbf{y} - \frac{1}{\delta} K^\top \nabla \varphi(K\mathbf{y} - \mathbf{r}) \right). \quad (3.6)$$

Conversely, if $\mathbf{y} \in \mathbb{R}^{\tilde{N}}$ satisfies the equation (3.6) for some $\delta > 0$, then $\mathbf{y}$ is a solution of model (3.3).

The fixed-point equation (3.6) together with an idea of [3] leads to efficient algorithms for finding a solution of model (3.2) via the optimization model (3.3). We first consider the case $\varphi := \frac{1}{2}\|\cdot\|_2^2$.

Algorithm 1

Input: the matrix K , the vector $\mathbf{r}$, and the diagonal matrix $\Gamma = \gamma I$

Initialization: $\mathbf{y}^0 = \mathbf{v}^1 = 0$, $t_1 = 1$.

repeat ($k \geq 0$)

$$\begin{aligned}\mathbf{y}^k &= \text{prox}_{\|\cdot\|_1 \circ \Gamma}(\mathbf{v}^k - K^\top(K\mathbf{v}^k - \mathbf{r})) \\ t_{k+1} &= \frac{1 + \sqrt{1 + 4t_k^2}}{2} \\ \mathbf{v}^{k+1} &= \mathbf{y}^k + \left(\frac{t_k - 1}{t_{k+1}}\right)(\mathbf{y}^k - \mathbf{y}^{k-1})\end{aligned}$$

until $\|\mathbf{y}^k - \mathbf{y}^{k-1}\|_2 / \|\mathbf{y}^{k-1}\|_2 > \text{tol}$

Return: $\mathbf{u} = \frac{1}{\lambda} W^\top \mathbf{y}^\infty$

Implementing Algorithm 1 requires computing the proximity operator $\text{prox}_{\frac{1}{\delta}\|\cdot\|_1 \circ \Gamma}$. This operator at a vector $\mathbf{z} \in \mathbb{R}^{\tilde{N}}$ can be explicitly evaluated by

$$\text{prox}_{\frac{1}{\delta}\|\cdot\|_1 \circ \Gamma}(\mathbf{z}) = \left[\text{prox}_{\frac{\gamma_1}{\delta}|\cdot|}(z_1), \text{prox}_{\frac{\gamma_2}{\delta}|\cdot|}(z_2), \dots, \text{prox}_{\frac{\gamma_{\tilde{N}}}{\delta}|\cdot|}(z_{\tilde{N}}) \right]^\top, \quad (3.7)$$

where γ_i is the i -th diagonal entry of Γ and

$$\text{prox}_{\frac{\gamma_i}{\delta}|\cdot|}(z_i) = \text{sign}(z_i) \cdot \max\left\{|z_i| - \frac{\gamma_i}{\delta}, 0\right\}. \quad (3.8)$$

Making use of formulas (3.7) and (3.8), we can implement Algorithm 1 efficiently.

Algorithm 1 can be seen as a special case of FISTA proposed in [3]. We next make a similar argument to Theorem 4.4 in [3] to analyze the convergence property of Algorithm 1. For the completion of this paper, we include this theorem here.

Theorem 3.1 (Theorem 4.4 in [3]) *Consider the optimization problem*

$$\min\{C(\mathbf{y}) \equiv p(\mathbf{y}) + q(\mathbf{y}) : \mathbf{y} \in \mathbb{R}^n\}$$

where both p and q are convex on $\mathbb{R}^n$, and p is continuously differentiable with Lipschitz continuous gradient $L(p)$. Let $\{t_k\}_{k=1}^\infty$ be a sequence with $t_1 = 1$ and $t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2}$ for all $k \geq 1$. Beginning with $\mathbf{y}^0 = \mathbf{v}^1 \in \mathbb{R}^n$, we generate two sequences $\{\mathbf{y}^k\}_{k=1}^\infty$, $\{\mathbf{v}^k\}_{k=1}^\infty$ of $\mathbb{R}^n$ as follows:

$$\begin{aligned}\mathbf{v}^{k+1} &= \text{prox}_{\frac{1}{L(p)}q}\left(\mathbf{y}^k - \frac{1}{L(p)}\nabla p(\mathbf{y}^k)\right) \\ \mathbf{v}^{k+1} &= \mathbf{y}^k + \left(\frac{t_k - 1}{t_{k+1}}\right)(\mathbf{y}^k - \mathbf{y}^{k-1}).\end{aligned}$$

Then for any $k \geq 1$

$$C(\mathbf{y}^k) - C(\mathbf{y}^*) \leq \frac{2L(p)\|\mathbf{y}^k - \mathbf{y}^*\|^2}{(k+1)^2},$$

where $\mathbf{y}^*$ is any solution of the optimization problem.

A convergence result for Algorithm 1 is given as follows.

Theorem 3.2 *Let $\{\mathbf{y}^k\}$ and $\{\mathbf{v}^k\}$ be generated by Algorithm 1. If the set of solutions to model (3.3) with $\varphi = \frac{1}{2}\|\cdot\|_2^2$ is nonempty, then for any $k \geq 1$ and any solution $\mathbf{y}^*$ of model (3.3)*

$$C(\mathbf{y}^k) - C(\mathbf{y}^*) \leq \frac{2\|\mathbf{y}^k - \mathbf{y}^*\|^2}{(k+1)^2}.$$

Proof: This result follows from Theorem 3.1. Specifically, for $\mathbf{y} \in \mathbb{R}^{\tilde{N}}$, we let

$$p(\mathbf{y}) := \frac{1}{2}\|K\mathbf{y} - \mathbf{r}\|_2^2.$$

Clearly, we have that

$$\nabla p(\mathbf{y}) = K^\top(K\mathbf{y} - \mathbf{r}), \quad \mathbf{y} \in \mathbb{R}^{\tilde{N}}$$

and the gradient ∇p of p is Lipschitz continuous with the Lipschitz constant $L(p) := \|K\|^2$. Moreover, due to the relation

$$KK^\top = RFW^\top WF^\top R^\top = I,$$

we have that $\|K\| = 1$. Thus, $L(p) = 1$.

We further define

$$q := \|\cdot\|_1 \circ \Gamma.$$

In this notation, Algorithm 1 is a special case of FISTA. By directly applying Theorem 3.1 (Theorem 4.4 of [3]) to the sequence generated by Algorithm 1, we complete the proof of this theorem. $\square$

Next, we consider the case $\varphi := \text{env}_{\tau\|\cdot\|_1}$. In this case, the flow of finding a solution of model (3.2) is described in Algorithm 2.

Algorithm 2

Input: the matrix K , the vector $\mathbf{r}$, the diagonal matrix $\Gamma = \gamma I$, and the parameter τ .

Initialization: $\mathbf{y}^0 = \mathbf{v}^1 = 0$, $t_1 = 1$.

repeat ($k \geq 0$)

$$\begin{aligned} \mathbf{y}^k &= \text{prox}_{\tau\gamma\|\cdot\|_1} \left(\mathbf{v}^k - K^\top(I - \text{prox}_{\tau\|\cdot\|_1})(K\mathbf{v}^k - \mathbf{r}) \right) \\ t_{k+1} &= \frac{1 + \sqrt{1 + 4t_k^2}}{2} \\ \mathbf{v}^{k+1} &= \mathbf{y}^k + \left(\frac{t_k - 1}{t_{k+1}} \right) (\mathbf{y}^k - \mathbf{y}^{k-1}) \end{aligned}$$

until $\|\mathbf{y}^k - \mathbf{y}^{k-1}\|_2 / \|\mathbf{y}^{k-1}\|_2 > \text{tol}$

Return: $\mathbf{u} = \frac{1}{\lambda} W^\top \mathbf{y}^\infty$

Implementing Algorithm 2 requires computing the proximity operator $\text{prox}_{\alpha\|\cdot\|_1}$. It is well-known [22] that for $\alpha > 0$ and $\mathbf{x} \in \mathbb{R}^{\tilde{N}}$, the i -th component of $\text{prox}_{\alpha\|\cdot\|_1} \mathbf{x}$ is $\text{prox}_{\alpha|\cdot|} x_i$, which is $\text{prox}_{\alpha|\cdot|} x_i = \max\{|x_i| - \alpha, 0\} \text{sign}(x_i)$. With these formulas, we can implement Algorithm 2 efficiently.

A convergence result of Algorithm 2 is given in the following theorem.

Theorem 3.3 *Let $\{\mathbf{y}^k\}$ and $\{\mathbf{v}^k\}$ be generated by Algorithm 2. If the set of solutions to model (3.3) with $\varphi = \text{env}_{\tau\|\cdot\|_1}$ is nonempty, then for all $k \geq 1$ and all solutions $\mathbf{y}^*$ of model (3.3)*

$$C(\mathbf{y}^k) - C(\mathbf{y}^*) \leq \frac{2\|\mathbf{y}^k - \mathbf{y}^*\|^2}{\tau(k+1)^2}.$$

Proof: Again, this result follows from Theorem 3.1. Define

$$p(\mathbf{y}) := \text{env}_{\tau\|\cdot\|_1}(K\mathbf{y} - \mathbf{r}), \quad \text{for } \mathbf{y} \in \mathbb{R}^{\tilde{N}}.$$

Recalling a well-known result about the Moreau envelope [24]

$$\nabla(\text{env}_{\tau\|\cdot\|_1}) = \frac{1}{\tau}(I - \text{prox}_{\tau\|\cdot\|_1}),$$

we observe that

$$\nabla p(\mathbf{y}) = \frac{1}{\tau}K^\top(I - \text{prox}_{\tau\|\cdot\|_1})(K\mathbf{y} - \mathbf{r}).$$

Note that the operator $I - \text{prox}_{\tau\|\cdot\|_1}$ is nonexpansive and $\|K\| = 1$. Hence, the gradient of ψ is Lipschitz continuous with the Lipschitz constant $L(p) = \frac{1}{\tau}$. We further define $q := \gamma\|\cdot\|_1$. Then the result of this theorem follows from Theorem 4.4 of [3] with $L = L(p)$. $\square$

4 Numerical experiments

In this section, five numerical experiments which include one for noise-free data, one for noisy data, one homogenous velocity model with two different sources and a three layered velocity model, are presented to illustrate the efficiency of the proposed algorithms in this paper. All the experiments are performed with Matlab 7v on an Intel Xeon (4-core) with 3.60 GHz and 16 Gb RAM.

We begin with setting up model (3.2). The matrices R , F and W involved in the model are chosen in the same way as that in previous sections. In the following computation, by Δf we denote the step size of the frequency, $f_{\min}$ and $f_{\max}$ represent, respectively, the lowest frequency and the highest frequency that we compute. The signal-to-noise ratio (SNR) defined by

$$\text{SNR} := 10 \log_{10} \left(\frac{\|data_{orig}\|_2^2}{\|data_{orig} - data_{reco}\|_2^2} \right)$$

is adopted, where $data_{orig}$ and $data_{reco}$ represent the original data and the recovered data, respectively. Furthermore, we set the tolerance as $tol = 10^{-6}$, and obtain the SNR-values by implementing Algorithms 1 and 2 with large enough iteration numbers to guarantee convergence.

4.1 Recovering time-domain data with exact data of low frequencies

In this subsection, we concentrate on recovering the first derivative of the Gaussian function (see equation (2.3)) with insufficient *exact* data in the frequency domain (see equation (2.4)) for the case of $t_0 = 1$ and $\tilde{\alpha} = 200$. In our computation, let $T := 2s$ and $N := 129$. Comparisons between Algorithm 1 and Algorithm 2 are made to demonstrate the efficiency of the proposed methods.

We first test the effect of the parameters L , γ , and τ of Algorithms 1 and 2. In this test, we choose $\Delta f := 0.5Hz$, $f_{\min} := 0.5Hz$ and $f_{\max} := 7.5Hz$. For Algorithm 1, Figures 2 (a)-(b) show that decreasing γ yields a higher SNR-value, while the number of iterations increases quickly. Therefore, a proper γ should be chosen to balance the quality of a reconstructed signal and the computational cost for obtaining the signal. Figure 2 (c) shows that the reconstructed signal cannot have a good SNR-value if L is too small. However, if L is too large, the computational cost increases quickly. Figure 2 (d) shows that $L = 3$ or 4 may be a proper choice. For Algorithm 2, Figures 3 (a)-(b) illustrate that a balance between SNR-values and iterations is also needed. Figures 3

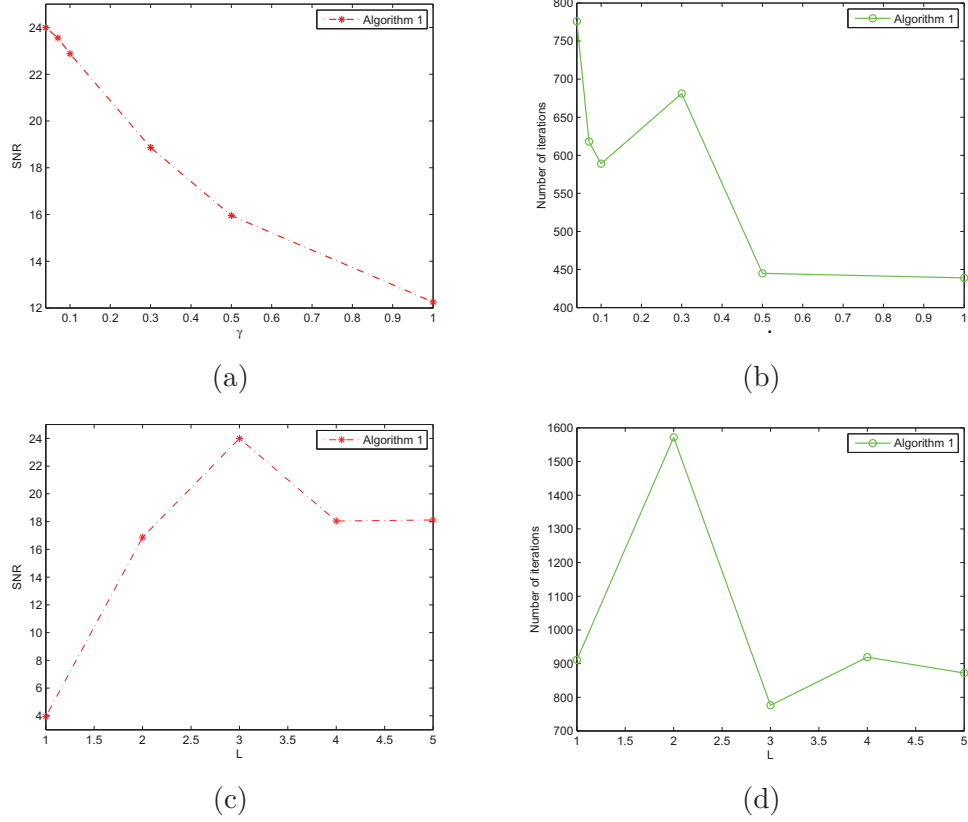


Figure 2: Parameters' effect in Algorithm 1 : (a) γ 's effect on SNR, (b) γ 's effect on iterations, (c) L 's effect on SNR, (d) L 's effect on iterations.

(c)-(d) show that when the parameter τ decreases the SNR-value usually increases, but the number of iterations increases rapidly. Therefore, we should choose τ properly.

Next, we test the restoration ability of Algorithms 1 and 2 for *exact* data sampled from different intervals $[0.5Hz, f_{max}]$ with Δf being $0.5Hz$, where f_{max} are 7.5Hz, 6Hz, 4.5Hz and 3Hz. Note that f_{max} in each test is much smaller than $15Hz$ deemed by the Nyquist sampling theorem. In Table 1, we compare SNR-values that are obtained by Algorithms 1 and 2. From Table 1, we find that Algorithm 1 seems comparable with Algorithm 2. In Figure 4 and 5, the figures obtained by Algorithm 1 and Algorithm 2 are presented, respectively. These figures show that both Algorithms 1 and 2 can well recover the first derivative of the Gaussian function (see equation (2.3)) with exact data of lower frequencies.

Finally, we test the anti-aliasing ability of Algorithms 1 and 2 for different intervals $[0.5Hz, f_{max}]$

Table 1: The SNR results of Algorithm 1 and Algorithm 2 for recovering time-domain data from *exact* data sampled from intervals $[0.5Hz, f_{max}]$ with $\Delta f = 0.5Hz$.

f_{max}	7.5Hz	6Hz	4.5Hz	3Hz
Algo.				
Algorithm 1	24.5150	16.0542	13.6275	13.4741
Algorithm 2	24.5149	16.1785	13.6355	13.5449

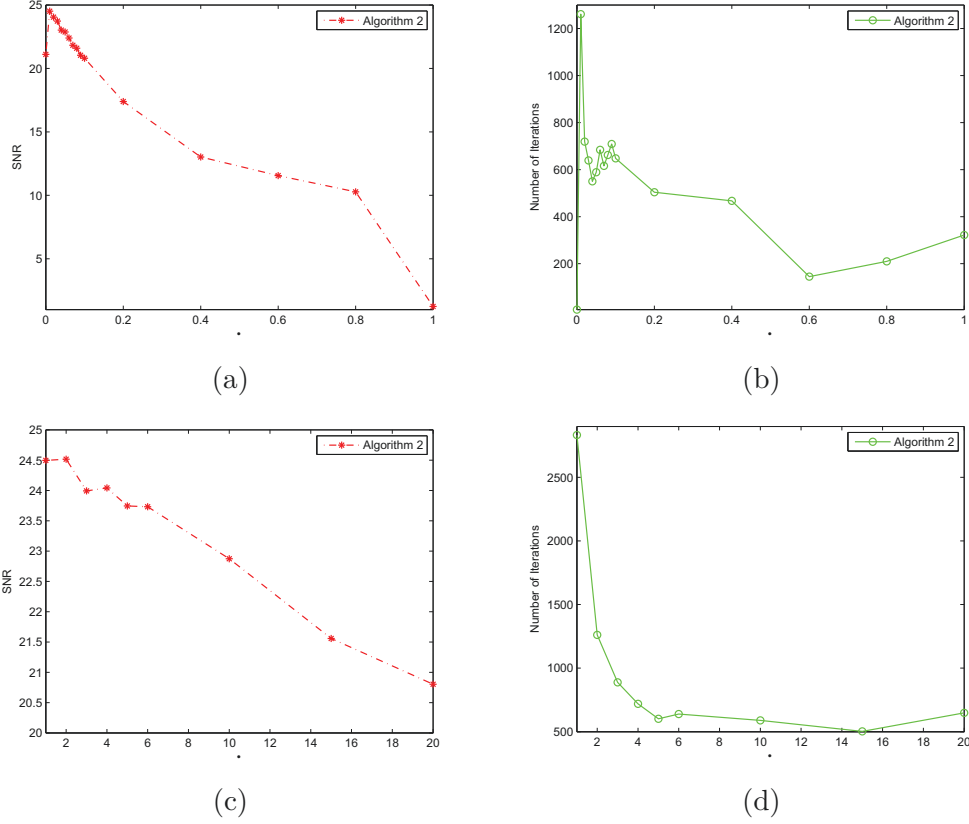


Figure 3: Parameters' effect on Algorithm 2 : (a) γ 's effect on SNR, (b) γ 's effect on iterations, (c) τ 's effect on SNR, (d) τ 's effect on iterations.

with Δf being $1.5Hz$, where f_{max} are $8Hz$, $6.5Hz$, $5Hz$ and $3.5Hz$. The selector matrix R is chosen according to the values of f_{max} . We remark that the frequency step size $\Delta f = 1.5Hz$ is bigger than $0.5Hz$ deemed by the Nyquist sampling theorem. Also, f_{max} in each test is much smaller than $15Hz$ deemed by the Nyquist sampling theorem. In Table 2, we show the SNR-value for Algorithms 1 and 2. In Figure 6 and Figure 7, we show the corresponding figures obtained by Algorithms 1 and 2. It is easy to see that, in the case $\Delta f = 1.5Hz$, both Algorithms 1 and 2 also work well.

From this example, we know that by our proposed algorithms, the first derivative of the Gaussian function can be recovered well with *exact* data of low frequency components.

Table 2: The SNR results of Algorithm 1 and Algorithm 2 for recovering time-domain data from *exact* data sampled from intervals $[0.5Hz, f_{max}]$ with $\Delta f = 1.5Hz$.

f_{max}	8 Hz	6.5Hz	5 Hz	3.5Hz
Algo.				
Algorithm 1	19.7405	13.9947	13.2246	13.4987
Algorithm 2	24.1218	14.5552	13.4151	13.4987

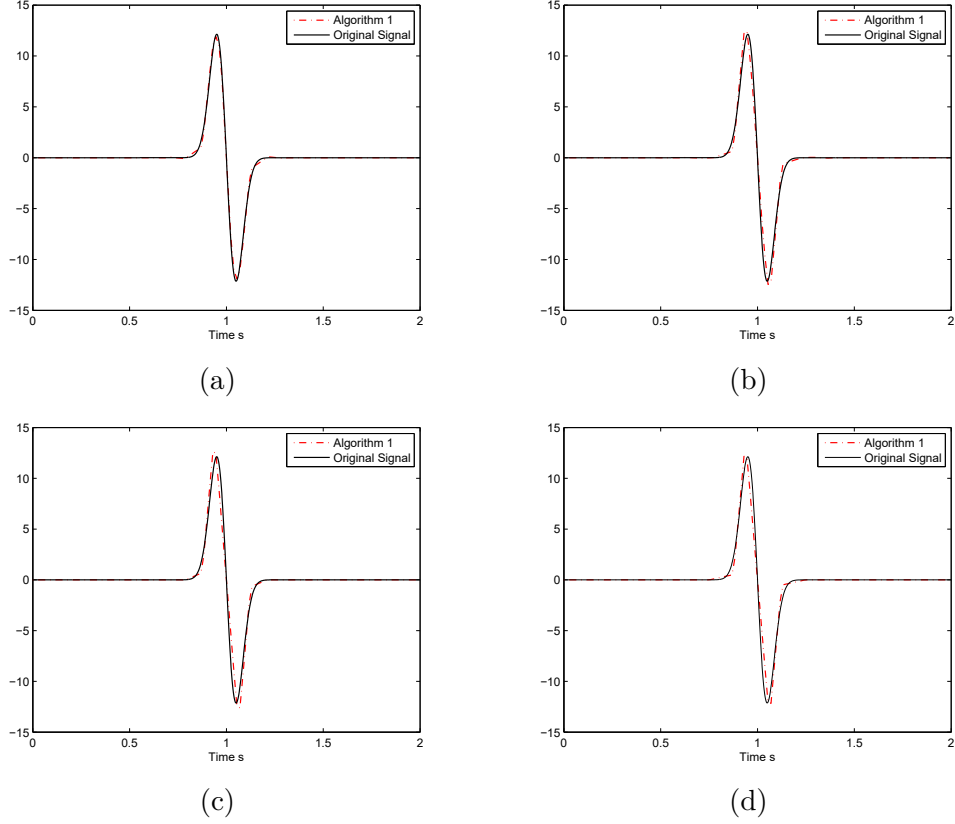


Figure 4: Algorithm 1, $\Delta f = 0.5Hz$: (a) $f_{max} = 7.5Hz$, $\gamma = 0.02$, (b) $f_{max} = 6Hz$, $\gamma = 0.01$, (c) $f_{max} = 4.5Hz$, $\gamma = 0.005$, (d) $f_{max} = 3Hz$, $\gamma = 0.15$.

4.2 Recovering time-domain data with noisy data of low frequencies

We consider an example of noisy data in this subsection. All conditions that we impose in this example are the same as those in the last one except the *exact* data is contaminated with noise (of standard deviation $\sigma = 0.1, 0.3, 0.5$). Each SNR-value reported in all tables for this example is the average over five runs.

We first test the restoration ability of Algorithms 1 and 2 for noisy data. The data are first sampled from the intervals $[0.5Hz, f_{max}]$ with Δf being $0.5Hz$ and then contaminated by noise with $\sigma = 0.1, 0.3, 0.5$, where f_{max} are chosen to be $7.5Hz, 6Hz, 4.5Hz$ and $3Hz$. Again, the selector matrix R for this example is chosen according to the values of f_{max} . f_{max} in each test is much smaller than $15Hz$, which is deemed by the Nyquist sampling theorem. Table 3 reports the SNR-values of the restored signals by Algorithms 1 and 2. From Table 3, we find that both Algorithms 1 and 2 can restore the signal well with noisy data of low frequency components.

We next test the anti-aliasing ability of Algorithms 1 and 2 for noisy data. The data are first sampled from the interval $[0.5Hz, f_{max}]$ with $\Delta f = 1.5Hz$ and then contaminated by noise with $\sigma = 0.1, 0.3$, where f_{max} are $8Hz, 6.5Hz, 5Hz$ and $3.5Hz$. We remark that the frequency step size $\Delta f = 1.5Hz$ is bigger than $0.5Hz$ which is deemed by the Nyquist sampling theorem. f_{max} in each test is much smaller than $15Hz$ which is again deemed by the Nyquist sampling theorem. The restored results by Algorithms 1 and 2 in terms of SNR-values are reported in Table 4. Comparing the results in Table 4 with those in Table 3, we find that the results of Table 3 are much better, which indicates that the step size of the frequency Δf should abide by the Nyquist sampling

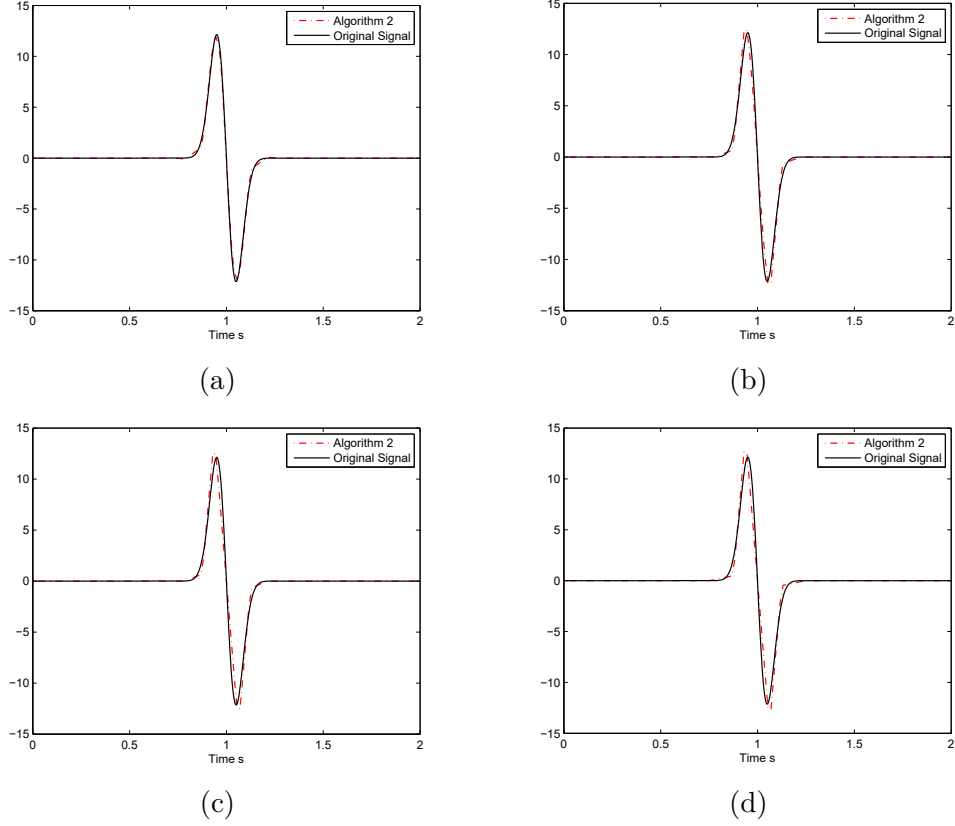


Figure 5: Algorithm 2, $\Delta f = 0.5Hz$: (a) $f_{max} = 7.5Hz$, $\gamma = 0.01$, $\tau = 2$, (b) $f_{max} = 6Hz$, $\gamma = 0.002$, $\tau = 7$, (c) $f_{max} = 4.5Hz$, $\gamma = 0.01$, $\tau = 3$, (d) $f_{max} = 3Hz$, $\gamma = 0.01$, $\tau = 5$.

theorem when the observed data are contaminated with noise.

4.3 One homogenous velocity model with the Ricker wavelet as the source

We consider generating the synthetic seismogram for a given point based on solving the 2D Helmholtz equation in a homogeneous medium. The related seismic model is shown in Figure 8 (a). The velocity of the medium is $1500m/s$. The receiver is located at the point $(x_r, z_r) = (1500m, 1000m)$. The point source $\hat{g} := \delta(x - x_s, z - z_s)\hat{R}(f, f_0)$ is located at the point $(x_s, z_s) = (500m, 1000m)$, where $\hat{R}(f, f_0)$ is the Ricker wavelet in the frequency domain defined by

$$\hat{R}(f, f_0) := \int_{-\infty}^{+\infty} (1 - 2\pi^2 f_0^2 t^2) \exp(-\pi^2 f_0^2 t^2) e^{-i2\pi f t} dt,$$

whose dominant frequency is $f_0 = 25Hz$. The maximum frequency of the above Ricker wavelet is approximately equal to $60Hz$. In our computation, we let $h := 10m$ and $T := 1.3440s$, and choose $N := 168$, $\Delta f := \frac{1}{T}$, $L := 4$ and $tol := 1e - 6$. In addition, in this experiment all results by the IDFT is obtained with frequency sampled from $[1Hz, 60Hz]$.

For the wave equation (2.1) with $v := 1500m/s$ and $g := \delta(x - x_s, z - z_s)R(t, f_0)$, its exact solution can be obtained by the D'Alembert formula:

$$u(x, z, t) = \frac{1}{4\pi r} R(t - \frac{r}{v}, f_0),$$

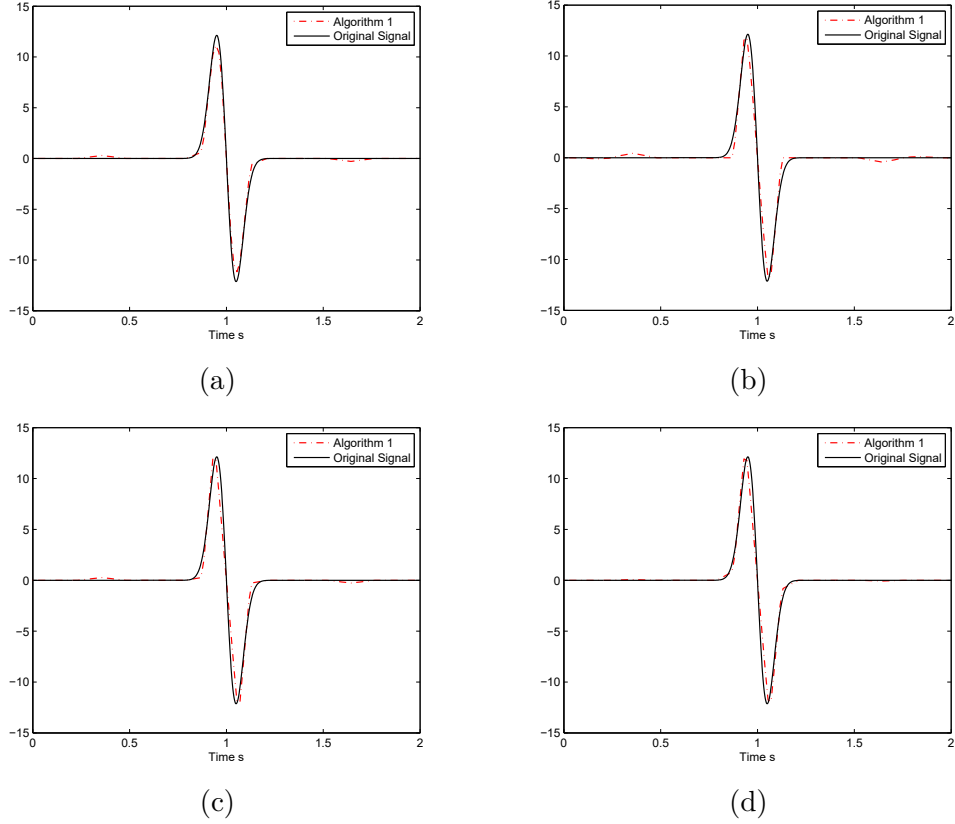


Figure 6: Algorithm 1, $\Delta f = 1.5Hz$: (a) $f_{max} = 8Hz$, $\gamma = 0.1$, (b) $f_{max} = 6.5Hz$, $\gamma = 0.07$, (c) $f_{max} = 5Hz$, $\gamma = 0.07$, (d) $f_{max} = 3.5Hz$, $\gamma = 0.05$.

where $R(t, f_0) := (1 - 2\pi^2 f_0^2 t^2) \exp(-\pi^2 f_0^2 t^2)$ and $r := \sqrt{(x - x_s)^2 + (z - z_s)^2}$. In the following, we take the signal $u(x_r, z_r, t)$, $t \in [0, T]$ obtained by the D'Alembert formula as the original signal.

In Figure 9 (a)-(b), we show the original signal and synthetic seismograms obtained by the IDFT and the proposed methods. Figure 9 (a) compares results obtained by Algorithm 1 and the IDFT, and Figure 9 (b) compares results obtained by Algorithm 2 and the IDFT. From these figures, we see that the signals recovered by both Algorithms 1 and 2 are much better than that by the IDFT, as the IDFT creates many spurious oscillations in the recovered signals.

To test our proposed methods, we show in Figure 10 (a)-(d) the original signal and the synthetic seismograms by Algorithm 1 with frequency sampled from intervals $[1Hz, f_{max}]$, where f_{max} are 54Hz, 48Hz, 42Hz and 36Hz, respectively. In Figure 11 we show the corresponding results for Algorithm 2. According to these figures, we find that even though only low frequencies are used, a better synthetic seismogram is obtained by Algorithms 1 and 2.

In Table 5, we list the parameter values used in the IDFT and Algorithm 1, which generate Figure 10 (a)-(d). Let N_H denote the number of times needed to solve the Helmholtz equation and k_{max} represent the largest dimensionless wave number. The dimensionless wave number $\tilde{k}$ on a nondimensional $[0, 1] \times [0, 1]$ domain is defined by $\tilde{k} := \frac{2\pi f M}{v}$, when the domain of interest is $[0, M] \times [0, M]$. As shown in Table 5, when the IDFT is applied, we need to solve the Helmholtz equation 81 times with the largest dimensionless wave numbers $k = 502.65$. On the other hand, when Algorithm 1 is used for Figure 10 (c), we only solve the equation 56 times with the largest dimensionless wave numbers $k = 351.86$. Similar conclusions can be drawn for Algorithm 2 which

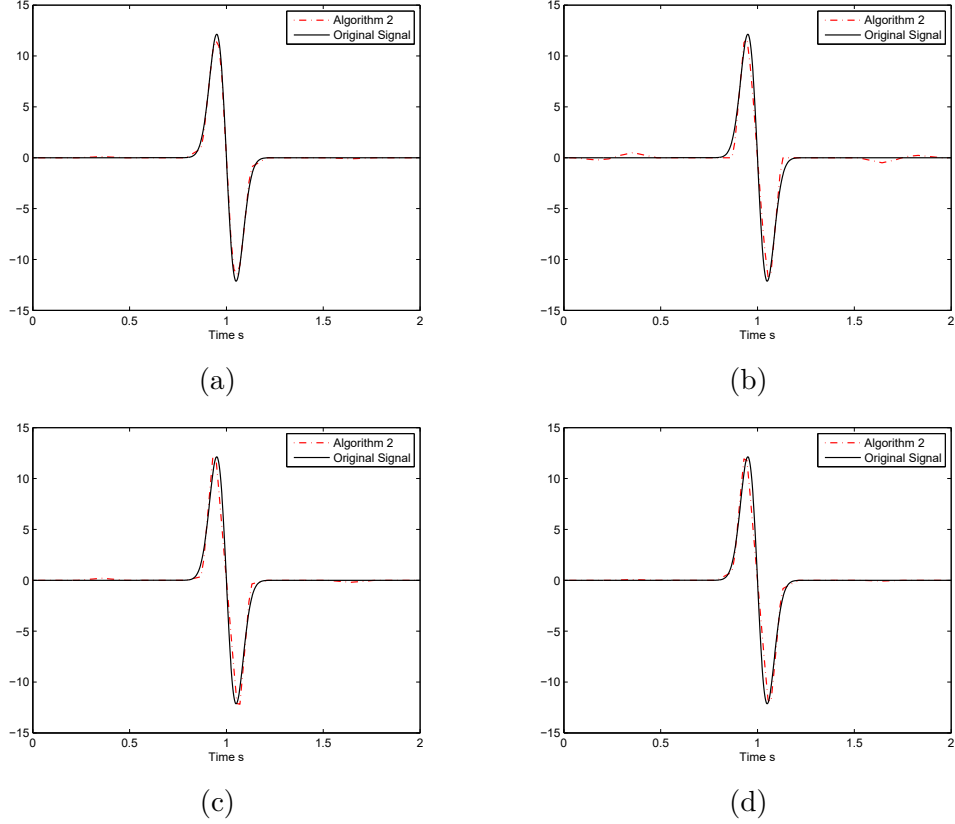


Figure 7: Algorithm 2, $\Delta f = 1.5Hz$, $\gamma = 0.1$, $\tau = 0.5$: (a) $f_{max} = 8Hz$, (b) $f_{max} = 6.5Hz$, (c) $f_{max} = 5Hz$, (d) $f_{max} = 3.5Hz$.

generates the results shown in Figure 11 (a)-(d). From Figures 10-11 and Table 5, we find that with the proposed method low frequencies of a signal contain enough information to recover the signal. It avoids solving the Helmholtz equation with large wave numbers.

To end this subsection, we give some illustrations on the phase displacement between the the original signal and the synthetic seismograms obtained by the IDFT or our proposed algorithms. This is due to the difference between the numerical phase velocity and the exact velocity (see, [15]).

4.4 The homogenous velocity model with the first order derivative of the Gaussian function as the source

We continue to generate synthetic seismograms based on the 2D Helmholtz equation in a homogeneous medium, and its seismic model is shown in Figure 8 (a). All the conditions are the same as those in the last subsection, except the point source function is now chosen as $\delta(x - x_s, y - y_s)\hat{G}(f, t_0, \tilde{\alpha})$ for $t_0 = 0.3$, $\tilde{\alpha} = 200$ (see, the equation (2.4)). The original signal of this problem is also obtained by the D'Alembert formula. The maximum frequency for the first order derivative of the Gaussian function is approximately $15Hz$. In addition, we set $h := 10m$ and $T := 2s$, and choose $N := 129$, $f_{min} := 0.5Hz$, $\Delta f := 0.5Hz$ and $L := 4$ in this experiment. Moreover, all results by the IDFT in this experiment is obtained with frequency sampled from $[0.5Hz, 15Hz]$.

Figure 12 shows the original signal and synthetic seismograms obtained by the IDFT, Algorithms

Table 3: A summary of the SNR results of Algorithm 1 and Algorithm 2 for recovering time-domain data from *noisy* data sampled from intervals $[0.5Hz, f_{max}]$ with $\Delta f = 0.5Hz$ and $\sigma = 0.1, 0.3, 0.5$.

Algorithm \ f_{max}	7.5Hz	6Hz	4.5Hz	3Hz
$\sigma = 0.1$				
Algorithm 1	18.8278	13.4524	13.3601	13.0725
Algorithm 2	19.6462	13.4857	13.3537	13.2617
$\sigma = 0.3$				
Algorithm 1	15.6276	12.5826	12.1345	10.7584
Algorithm 2	15.2982	12.6605	11.8586	11.7015
$\sigma = 0.5$				
Algorithm 1	11.5818	11.4273	10.3744	6.7984
Algorithm 2	12.1898	11.1381	11.0302	9.1298

Table 4: A summary of the SNR results of Algorithm 1 and Algorithm 2 for recovering time-domain data from *noisy* data sampled from intervals $[0.5Hz, f_{max}]$ with $\Delta f = 1.5Hz$ and $\sigma = 0.1, 0.3$.

Algorithm \ f_{max}	8Hz	6.5Hz	5Hz	3.5Hz
$\sigma = 0.1$				
Algorithm 1	15.1852	12.2777	11.6868	9.3177
Algorithm 2	14.4361	12.8019	11.8733	10.8679
$\sigma = 0.3$				
Algorithm 1	9.5592	10.7386	10.9311	9.2843
Algorithm 2	12.2044	10.0740	9.1445	7.6307

1 and 2, with the frequency sampled from $[0.5Hz, 15Hz]$. From these pictures, we find that the results obtained by both Algorithms 1 and 2 are much better than that by the IDFT, as the IDFT creates many spurious oscillations in the recovered signals.

Figures 13 (a)-(d) show the original signal and the synthetic seismograms by Algorithm 1 with frequency sampled from $[0.5Hz, f_{max}]$, where f_{max} are 9Hz, 7.5Hz, 6Hz, and 4.5Hz, respectively. The corresponding results for Algorithm 2 are presented in Figure 14. From Figures 13 and 14, we find that both Algorithms 1 and 2, even though only a few low frequencies are used, yield better synthetic seismograms than that by the IDFT.

We summarize in Table 6 the parameters used in the IDFT and Algorithm 1 to generate the numerical results demonstrated in Figures 13 (a)-(d). The notations listed in Table 6 have the same meaning as that in Table 5. As shown in Table 6, corresponding to Figures 13 (a)-(d), the largest dimensionless wave number k we need to compute by Algorithm 1 are 75.398, 62.832, 50.266 and 37.699, respectively, while the largest dimensionless wave number associated with the IDFT is 125.66. Moreover, the IDFT needs to solve the Helmholtz equation 30 times to obtain the synthetic seismogram, while Algorithm 1 only needs to solve this equation with low frequencies for as low as

Table 5: The parameter values used in the IDFT and Algorithm 1 that generate Figure 10 (a)-(d)

	IDFT	Algo.1 in (a)	Algo.1 in (b)	Algo.1 in (c)	Algo.1 in (d)
f_{max}	60 Hz	54 Hz	48 Hz	42 Hz	36 Hz
k_{max}	502.65	452.39	402.12	351.86	301.59
N_H	81	73	65	56	48

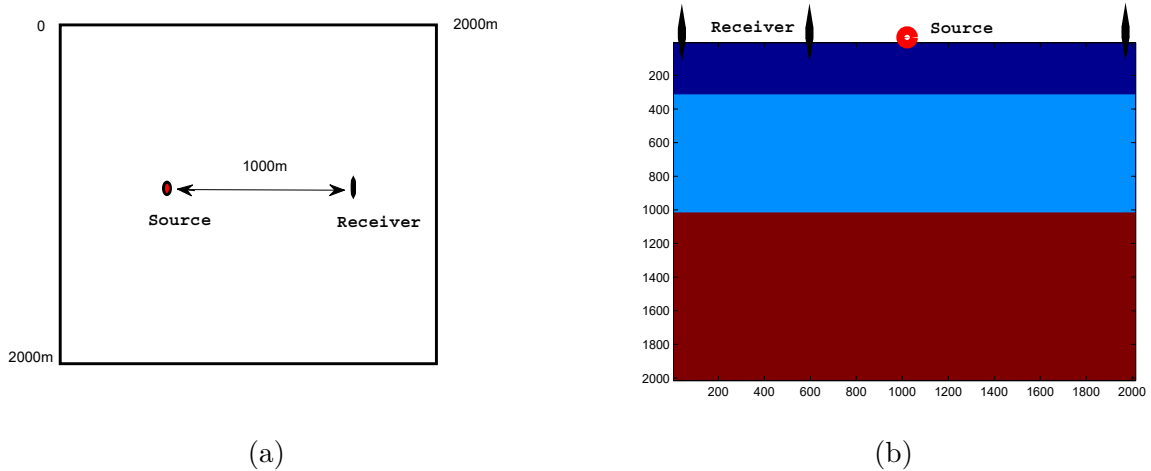


Figure 8: Velocity model: (a) The homogenous model; (b) The layered model.

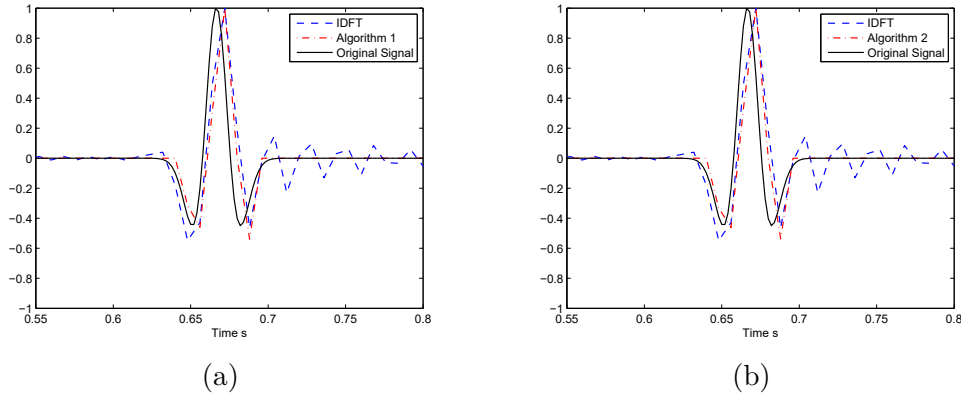


Figure 9: $f_{max} = 60Hz$ (a) Algorithm 1: $\gamma = 0.5$, (b) Algorithm 2: $\gamma = 0.8$, $\tau = 0.6$.

9 times. Similar conclusion can be drawn for Algorithm 2 based on Figure 14 (a)-(d). According to Figures 13-14 and Table 6, to obtain a reliable synthetic seismogram, we need not to compute the wave number which is larger than 50.266. Hence, solving the Helmholtz equation with large wave numbers is avoided.

4.5 A three layered velocity model

We consider generating common-shot-point records (shot profiles) based on the 2D Helmholtz equation in a heterogenous medium, and its seismic model is shown in Figure 8 (b). Our interested domain is $[0, 2000m] \times [0, 2000m]$. There are three velocities (from the top to the bottom) in this velocity model: $v = 2000m/s$, $2500m/s$, $4000m/s$. The source is located at the point $(0m, 1000m)$. As in Subsection 4.3, Ricker wavelet with domain frequency 25Hz is used as the source. In addition, the receivers are located on the top ground, that is, they are in the line of $z = 0$. The grid size is taken as $\Delta x = \Delta z := 10m$ and the interval of time increment $\Delta t := 8ms$. Here we only present the results obtained by Algorithm 2, as Algorithm 1 performs similar to Algorithm 2 in efficiency.

Figure 15 (a)-(b) show the common-shot-point records obtained by the IDFT, in which frequency is sampled from intervals $[1Hz, 60Hz]$ and $[1Hz, 42Hz]$, respectively. The step size of the frequency

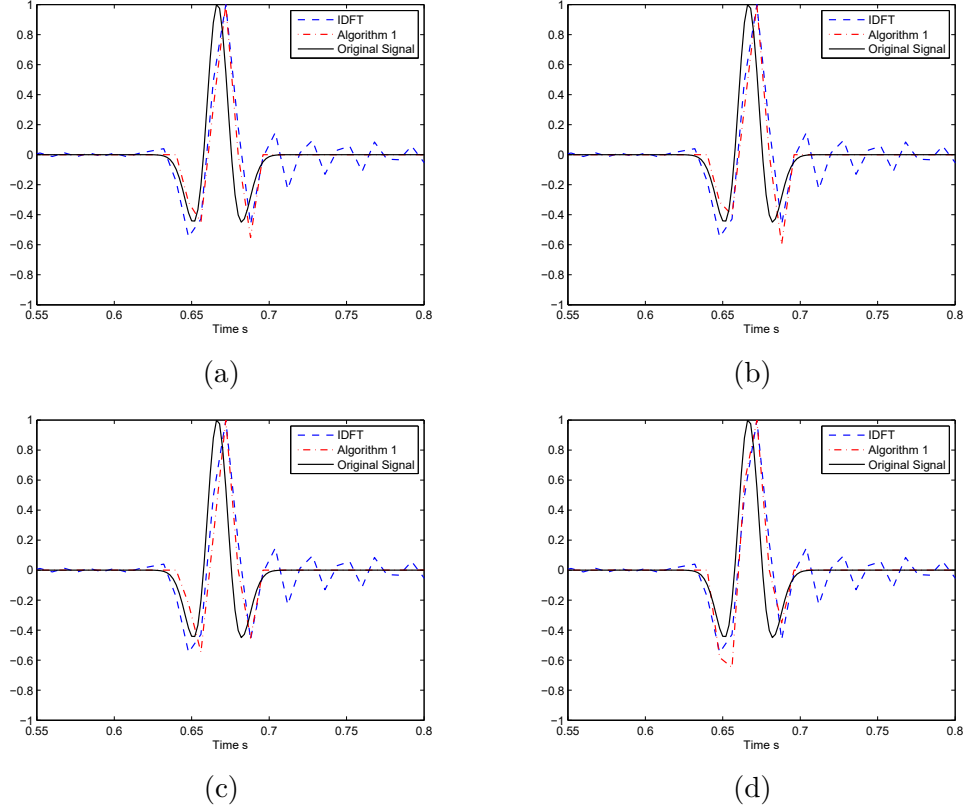


Figure 10: Algorithm 1: (a) $f_{max} = 54Hz$, $\gamma = 0.5$, (b) $f_{max} = 48Hz$, $\gamma = 0.51$, (c) $f_{max} = 42Hz$, $\gamma = 0.51$, (d) $f_{max} = 36Hz$, $\gamma = 0.5$.

Δf is $0.4464Hz$. Figure 16 (a)-(f) show some results obtained by Algorithm 2. From Figure 15 (b), we find that some nonphysical oscillations appear in the seismic wavefields obtained by the IDFT, as the Nyquist-Shannon criterion may not be satisfied. However, as shown in Figure 16 (a)-(e), the direct waves of the source, the reflected waves of the top side of the second layer and the reflected waves of the bottom side of the second layer are displayed clearly in the seismic wavefields obtained by Algorithm 2. It is easy to find that, Figure 16 (b)-(e) are almost as clear as Figure 15 (a). Nonphysical oscillations appear only in Figure 16 (f). Therefore, frequencies sampled from $[1Hz, 24Hz]$ with $\Delta f = 0.4464Hz$ are enough to restore the seismic wavefield by Algorithm 2, which confirms the efficiency of our proposed algorithms.

Table 6: The parameter values used in the IDFT and Algorithm 1 that generate Figure 13(a)-(d)

	IDFT	Algo.1 in (a)	Algo.1 in (b)	Algo.1 in (c)	Algo.1 in (d)
f_{max}	15 Hz	9 Hz	7.5 Hz	6 Hz	4.5 Hz
k_{max}	125.66	75.398	62.832	50.266	37.699
N_H	30	18	15	12	9

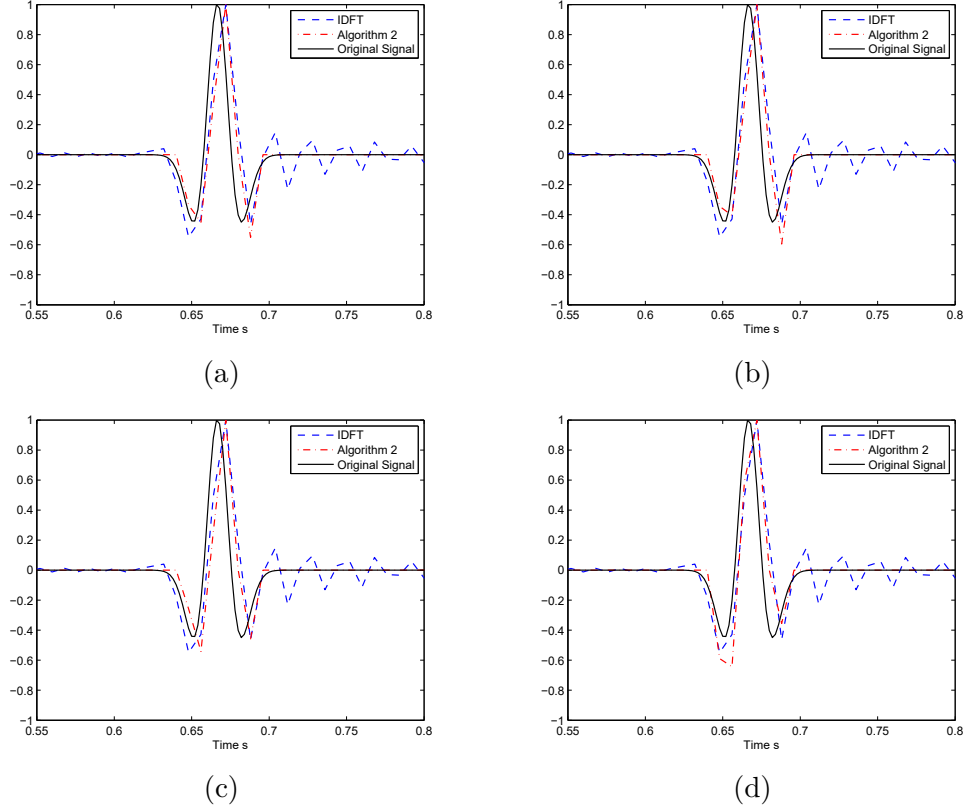


Figure 11: Algorithm 2, $\gamma = 0.8$, $\tau = 0.6$: (a) $f_{max} = 54Hz$, (b) $f_{max} = 48Hz$, (c) $f_{max} = 42Hz$, (d) $f_{max} = 36Hz$.

5 Conclusion

In this article, we use fixed-point proximity algorithms to solve the incomplete Fourier transform arising in seismic wavefield modeling in the frequency domain. The piecewise linear tight framelet is used as the sparse transform, and a new sampling is introduced to avoid solving the Helmholtz equation with large wave numbers. Based on the proximity operator, two algorithms are proposed. Numerical results show that this sampling method is practical, and our proposed algorithms are efficient and robust.

References

- [1] S. Alliney, A property of the minimum vectors of a regularizing functional defined by means of the absolute norm, *IEEE Transactions on Signal Processing*, 45 (1997), 913-917.
- [2] I. Babuška and S. A. Sauter, Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers? *SIAM Review*, 42 (2000), 451-484.
- [3] A. Beck and M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM Journal on Imaging Sciences*, 2 (2009), 183-202.
- [4] E. Brigham, *The fast Fourier transform and its application*, Prentice-Hall Inc., NJ, 1988.

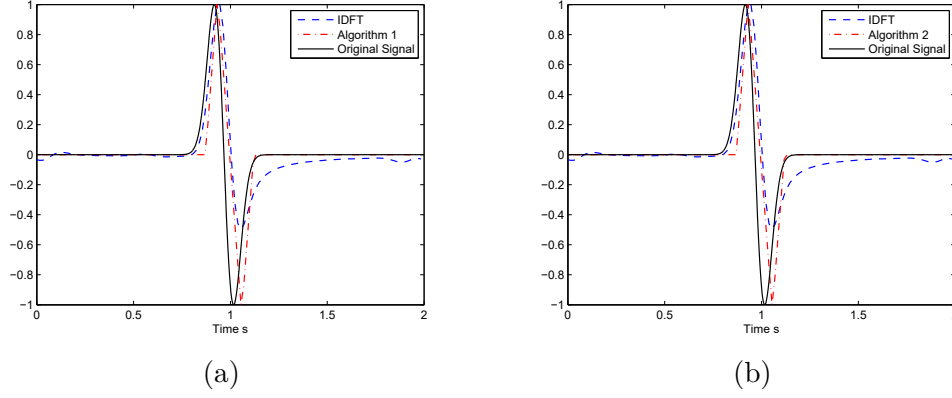


Figure 12: $f_{max} = 15\text{Hz}$ (a) Algorithm 1: $\gamma = 29$, (b) Algorithm 2: $\gamma = 0.7$, $\tau = 41$.

- [5] E. J. Candès, J. Romberg and T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Transactions on Information Theory*, 52 (2006), 489-509.
- [6] E. J. Candès, J. Romberg and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, *Communications on Pure and Applied Mathematics*, 59 (2006), 1207-1223.
- [7] E. J. Candès, Compressive sampling, *Proceedings of the International Congress of Mathematicians*, 3 (2006), 1433-1452.
- [8] F. Chen, L. Shen, Y. Xu, and X. Zeng, The Moreau Envelope Approach for the L1/TV Image Denoising Model, *Inverse Probl. Imaging*, 8 (2014), 53-77.
- [9] Z. Chen, D. Cheng, W. Feng and T. Wu, An optimal 9-point finite difference scheme for the Helmholtz equation with PML, *International Journal of Numerical Analysis and Modeling*, 10 (2013), 389-410.
- [10] P. Combettes and V. Wajs, Signal recovery by proximal forward-backward splitting, *Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal*, 4 (2005), 1168-1200.
- [11] D. L. Donoho, Compressed sensing, *IEEE Transactions on Information Theory*, 52 (2006), 1289-1306.
- [12] X. Feng and H. Wu, Discontinuous Galerkin methods for the Helmholtz equation with large wave number, *SIAM J. Numer. Anal.*, 47 (2009), 2872-2896.
- [13] T. Goldstein and S. Osher, The split Bregman method for l^1 regularization problems, *SIAM Journal on Imaging Sciences*, 2 (2009), 323-343.
- [14] B. Hustedt, S. Operto and J. Virieux, Mixed-grid and staggered-grid finite-difference methods for frequency domain acoustic wave modelling, *Geophys. J. Int.*, 157 (2004), 1269-1296.
- [15] C. -H. Jo, C. Shin, and J. H. Suh, An optimal 9-point, finite-difference, frequency-space, 2-D scalar wave extrapolator, *Geophysics*, 61 (1996), 529-537.
- [16] E. Lebed and F. J. Herrmann, A hitchhiker's guide to the galaxy of transform-domain sparsification, *In SEG Technical Program Expanded Abstracts*, SEG, 27 (2008).

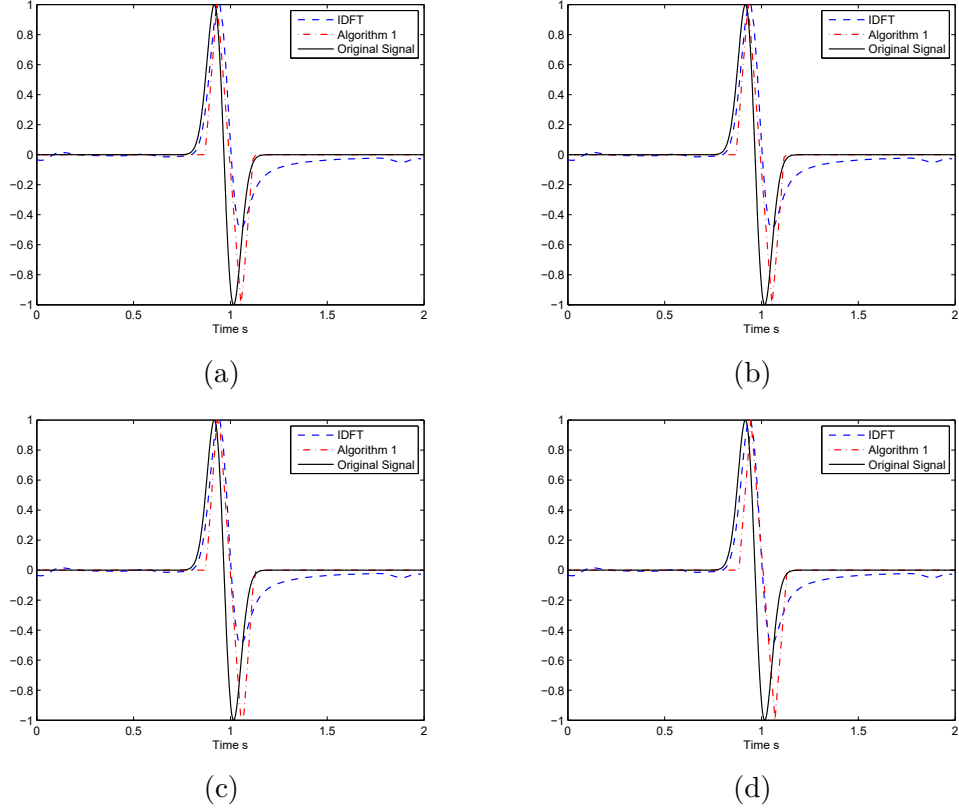


Figure 13: Algorithm 1: (a) $f_{max} = 9\text{Hz}$, $\gamma = 29$, (b) $f_{max} = 7.5\text{Hz}$, $\gamma = 29$, (c) $f_{max} = 6\text{Hz}$, $\gamma = 29$, (d) $f_{max} = 4.5\text{Hz}$, $\gamma = 26$.

- [17] T. T. Y. Lin and F. J. Herrmann, Compressed wavefield extrapolation, *Geophysics*, 72 (2007), SM77-SM93.
- [18] T. T. Y. Lin, E. Lebed, Y. Erlangga and F. J. Herrmann, Interpolating solutions of the Helmholtz equation with compressed sensing, *In SEG Technical Program Expanded Abstracts, SEG*, 27 (2008), 2122-2126.
- [19] J. Lysmer and L. A. Drake, A finite element method for seismology, *Methods in Computational Physics*, 11 (1972), 181-216.
- [20] Q. Li and L. Shen and Y. Xu and N. Zhang, Multi-step fixed-point proximity algorithms for solving a class of optimization problems arising from image processing, *Advances in Computational Mathematics*, 41(2)(2014), 387-422.
- [21] K. J. Marfurt, Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations, *Geophysics*, 49 (1984), 533-549.
- [22] C. A. Micchelli, L. Shen and Y. Xu, Proximity Algorithms for Image Models: Denoising, *Inverse Problems*, 27 (2011), 45009-45038.
- [23] C. A. Micchelli, L. Shen, Y. Xu, and X. Zeng, Proximity algorithms for the l_1/tv image denosing models, *Advances in Computational Mathematics*, 38 (2013), 401-426.

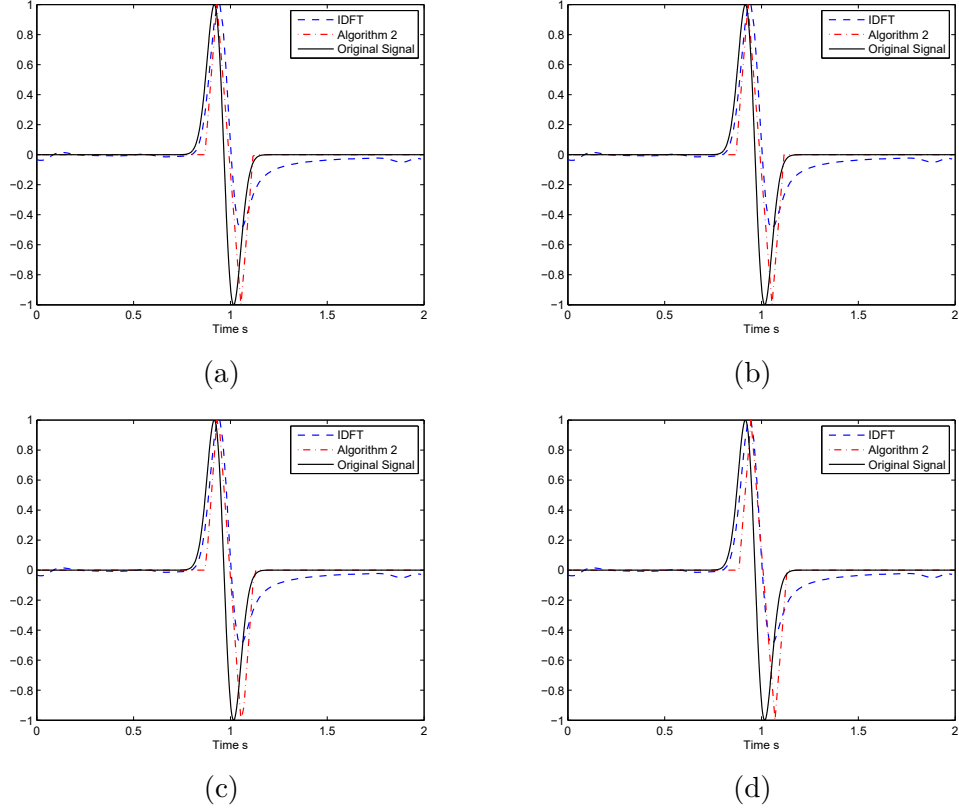


Figure 14: Algorithm 2: (a) $f_{max} = 9Hz$, $\gamma = 0.7$, $\tau = 40$, (b) $f_{max} = 7.5Hz$, $\gamma = 0.7$, $\tau = 40$, (c) $f_{max} = 6Hz$, $\gamma = 0.69$, $\tau = 40$, (d) $f_{max} = 4.5Hz$, $\gamma = 0.62$, $\tau = 40$.

- [24] J.-J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien. *C.R. Acad. Sci. Paris Sér. A Math.*, 255 (1962), 1897-2899.
- [25] M. Nikolova, A variational approach to remove outliers and impulse noise, *Journal of Mathematical Imaging and Vision*, 20 (2004), 99-120.
- [26] R. G. Pratt, C. Shin and G. J. Hicks, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion, *Geophys. J. Int.*, 133 (1998), 341-362.
- [27] R. G. Pratt and M. H. Worthington, Inverse theory applied to multi-source cross-hole tomography. Part I: Acoustic wave-equation method, *Geophysical Prospecting*, 38 (1990), 287-310.
- [28] H. Ren, H. Wang and T. Gong, Seismic modeling of scalar seismic wave propagation with finite difference scheme in frequency-space domain, *Geophysical Prospecting for Petroleum*, 48 (2009), 20-27.
- [29] C. D. Riyanti, A. Kononov, Y. A. Erlangga, C. Vuik, C. W. Oosterlee, R. E. Plessix and W. A. Mulder, A parallel multigrid-based preconditioner for the 3D heterogeneous high-frequency Helmholtz equation, *Journal of Computational Physics*, 224 (2007), 431-448.
- [30] C. Shin and H. Sohn, A frequency-space 2-D scalar wave extrapolator using extended 25-point finite-difference operator, *Geophysics*, 63 (1998), 289-296.

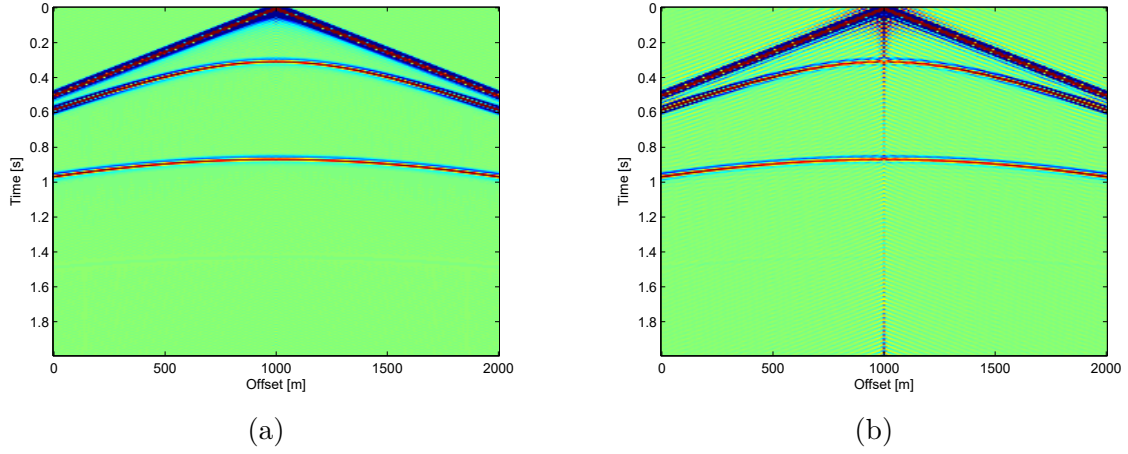
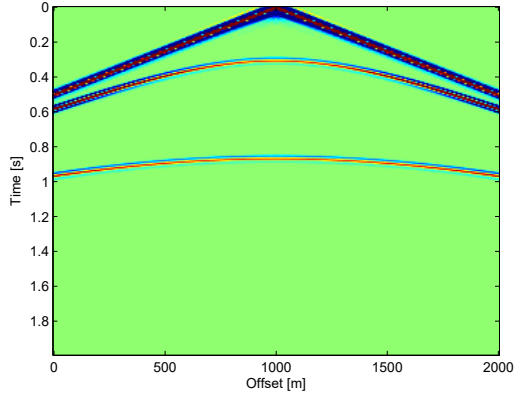
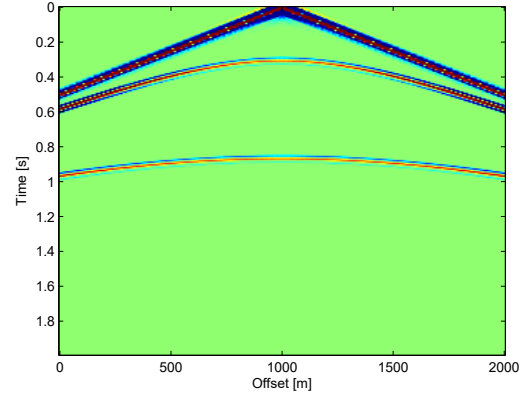


Figure 15: The common-shot-point records via IDFT with the frequency samples taken from : (a) [1Hz,60Hz]; (b) [1Hz,42Hz].

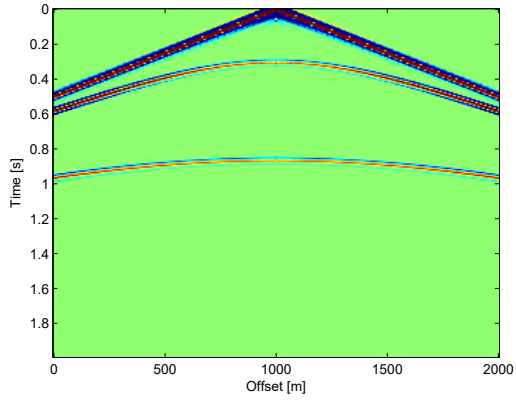
- [31] T. Wu, A dispersion minimizing compact finite difference scheme for the 2D Helmholtz equation, *Journal of Computational and Applied Mathematics*, 311 (2017), 497-512.



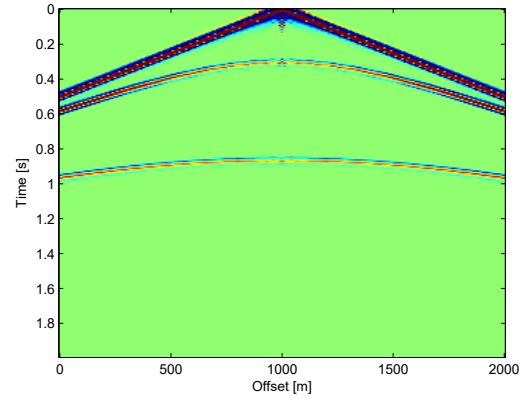
(a)



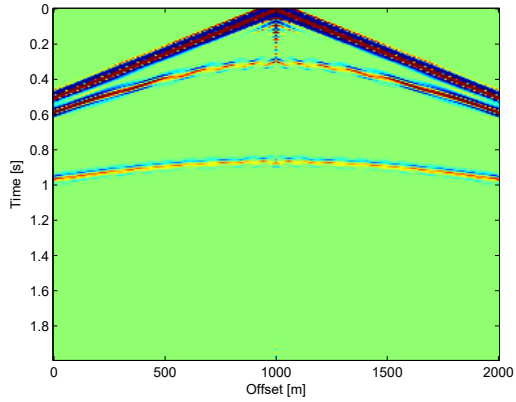
(b)



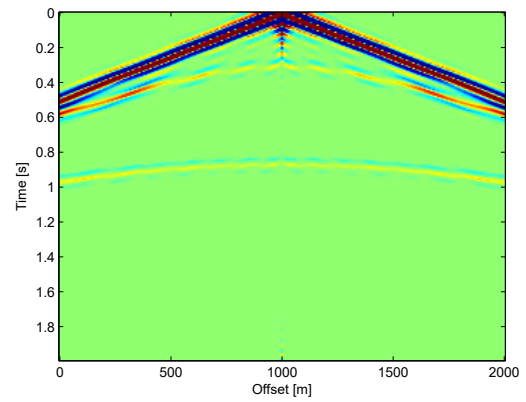
(c)



(d)



(e)



(f)

Figure 16: The common-shot-point records via Algorithm 2 with the frequency samples taken from : (a) [1Hz,60Hz]; (b) [1Hz,42Hz]; (c) [1Hz,36Hz]; (d) [1Hz,30Hz]; (e) [1Hz,24Hz]; (f) [1Hz,18Hz].