Improved wavefield reconstruction from randomized sampling via weighted one-norm minimization

Hassan Mansour[*†], **Felix J. Herrmann, and Özgür Yılmaz**[*]

[*] *Mathematics Department,*

The University of British Columbia,

Vancouver, BC V6T1Z2.

[†] *Seismic Laboratory of Imaging and Modeling,*

The University of British Columbia,

Vancouver, BC V6T1Z4.

(May 20, 2013)

GEO-2012-0383.R2

Running head: **Weighted one-norm minimization**

ABSTRACT

Missing-trace interpolation aims to recover the gaps caused by physical obstacles or deliberate subsampling to control acquisition costs in otherwise regularly-sampled seismic wavefields. While transform-domain sparsity promotion has proven to be an effective tool to solve this recovery problem, current recovery techniques do not fully utilize a priori information derived from the *locations* of the transform-domain coefficients especially when curvelet domain sparsity is exploited. In this paper, we propose recovery by weighted one-norm minimization, which exploits correlations between the locations of significant curvelet coefficients of different partitions, e.g., shot records, common-offset gathers, or frequency slices of the acquired data. We use these correlations to define a sequence of 2D curvelet-

based recovery problems that exploit 3D continuity exhibited by seismic wavefields without relying on the highly redundant 3D curvelet transform. To illustrate the performance of our weighted algorithm, we compare recoveries from different data sorting and partitioning scenarios for a seismic line from the Gulf of Suez. These examples demonstrate that our method is superior to standard ℓ_1 minimization in terms of anti-aliasing capability, reconstruction quality and computational memory requirements.

INTRODUCTION

The problem of interpolating random missing traces in otherwise regularly sampled seismic data to a complete regular grid often occurs in 2D and 3D seismic settings. For example, recent developments in randomized acquisition in marine, via randomized coil sampling or randomized streamers with deliberate feathering, and on land, via randomization of the source and/or receiver positions, result in a vast array of random missing traces. Since most multi-trace processing algorithms do not handle irregularly sampled (and aliased) data, interpolation to a regular grid is necessary in order to process and subsequently interpret the data.

Irregular sampling often leads to image artifacts (referred to as acquisition footprint) due to the uneven illumination of the subsurface as well as poor levels of repeatability between time-lapse imaging. To overcome these problems, the acquired dataset is regularized and/or interpolated (Hennenfent et al., 2010). Regularization, or bin-centering, simply relocates traces from their irregular recording locations to locations on a regular grid. The traces may be altered depending on the method used and some traces may be discarded. On the other hand, interpolation generates synthesized traces that are interleaved with the recorded traces in order to increase the spatial sampling rate or fill in the gaps. By combining regularization with interpolation, it is possible to significantly impact the quality of pre-stack time migration (Symes, 2007), 3-D surface-related multiple elimination (Berkhout and Verschuur, 1997; Verschuur and Berkhout, 1997), wave-equation pre-stack depth migration (Claerbout, 1971), and time-lapse imaging where interpolation and regularization homogenize the sampling between all vintages of the 4-D dataset (Smith et al., 2012).

In this paper, we are interested in the scenario where the complete data is sampled on

a regular grid and we address the problem of interpolating missing traces in an otherwise-regular dataset. Recently, it has been shown both by simulation (Herrmann and Hennenfent, 2008; Hennenfent and Herrmann, 2006) and in practice (Mosher et al., 2012) that randomizing the acquisition of seismic data by randomly spacing the receivers on a regular grid as an acquisition design guideline produces better interpolated traces than from uniformly spaced receivers. The recovery from subsampled receivers has been addressed by several works in the seismic exploration literature. For instance, Claerbout (1992) and Spitz (1991) formulate the interpolation problem as a least-squares optimization problem where the missing traces are interpolated using a prediction filter that is estimated in the frequency domain. Ferguson (2010), applies simultaneous regularization and wave-equation statics (WEstatics) implemented using least-squares inversion to address the problem of irregular trace spacing and statics correction. Alternatively, transform-based methods have dominated the literature in recent years utilizing particular transforms e.g. the (non)uniform discrete Fourier transform (see e.g. Sacchi and Ulrych, 1996; Liu and Sacchi, 2004; Duijndam et al., 1999; Zwartjes, 2005), the antileakage Fourier transform (Xu et al., 2005), the Radon transform (see e.g. Thorson and Claerbout, 1985; Hampson, 1986; Herrmann et al., 2000), and the curvelet transform (see e.g. Hennenfent and Herrmann, 2005, 2006; Herrmann and Hennenfent, 2008; Hennenfent et al., 2010; Naghizadeh and Sacchi, 2010; Neelamani et al., 2010).

Recently, compressed sensing (Donoho, 2006; Candès and Tao, 2006) has emerged as a process of acquiring incomplete random linear measurements of a signal and then reconstructing it by utilizing the prior knowledge that the signal is sparse or compressible in some transform domain. A measurement process is considered to be linear if it can be described by a set of linear equations. In seismic exploration, data consists of wavefronts that

exhibit structure in multiple dimensions. In the curvelet transform domain, it is possible to capture this structure by a small number of significant transform coefficients, resulting in a sparse representation of the data. Several works in the literature have formulated the seismic data interpolation problem as an instance of recovery from compressive samples and this approach resulted in considerable improvement in reconstruction quality over linear reconstruction techniques such as applying the adjoint of the sampling operator (see e.g., Hennenfent and Herrmann, 2005, 2006, 2008; Herrmann et al., 2012). The interpolation problem becomes that of finding the curvelet synthesis coefficients with the smallest ℓ_1 norm that best fits the randomly subsampled data in the physical domain.

Fully utilizing transform domain sparsity sometimes requires the use of redundant transforms, i.e., transforms that can result in a significant expansion in the size of the model space. One such highly redundant transform is the directional curvelet transform where applying the 3D curvelet transform to a seismic line induces a curvelet domain representation with a 24 fold expansion in the size of the seismic line. A less optimal but memory saving approach applies the 2D curvelet transform to two dimensional partitions of the seismic line, e.g., shot records, common-offset gathers, or frequency slices of the acquired data. Consequently, the data are interpolated by solving several ℓ_1 minimization problems that capture the structure in every 2D partition of the seismic line. This approach can be parallelized since the recovery of each 2D partition is independent of the remaining partitions. However, recovering 2D partitions independently does not utilize the wavefront continuity that exists across partitions. This is due to the fact that the ℓ_1 minimization problem does not incorporate additional prior information related to the structure of the signal.

In this paper, we propose to use weighted ℓ_1 minimization to improve the performance of transform based data interpolation when information related to the locations of the non-

zero entries (also called support) of the signal is available. The continuity of wavefronts across adjacent partitions of a seismic line manifests itself as a high correlation in the support of the curvelet coefficients of the partitions. Moreover, by virtue of the transient character of seismic source functions, this correlation is also present amongst the support of different frequency slices as well as amongst different offset slices. Our approach is motivated by the results of Friedlander et al. (2012), which proved that given a support estimate that is highly correlated with the true support of the signal, the solution of the weighted ℓ_1 minimization problem has better recovery performance than that of standard ℓ_1 minimization. The idea of using weights to improve the recovery of regularized inverse problems has been previously explored in the literature. Liu and Sacchi (2004) proposed minimizing a weighted ℓ_2 norm in the wavenumber domain that constrains the solution to be spatially bandlimited. Our approach differs from that of Liu and Sacchi (2004) in that we make no assumption on the bandlimitedness of the signal. Instead, we utilize the sparsity and correlated structure of seismic data partitions in the curvelet domain and solve a sequence of weighted ℓ_1 minimization problems that is well suited for such structure. Moreover, our approach is not restricted to operating in the F-X domain but extends to the time-midpoint domain. We present numerical simulations conducted on a subsampled seismic line from the Gulf of Suez demonstrating that weighted ℓ_1 minimization significantly outperforms standard ℓ_1 minimization both in terms of the recovered signal-to-noise-ratio (SNR) and a visual inspection of the quality of reconstructed shot gathers.

Outline

We start by formulating the seismic data interpolation problem as a sparse recovery problem and then present an overview of the recovery guarantees of ℓ_1 minimization and weighted

ℓ_1 minimization. Next, we describe how the interpolation of a seismic line can be achieved by solving a sequence of weighted ℓ_1 minimization problems and we discuss different ways of partitioning seismic lines to improve the recovery. Finally, we present the results of numerical simulations that illustrate the performance of weighted ℓ_1 in interpolating real seismic data.

THEORY

Seismic data interpolation by sparse recovery

Seismic data interpolation is an underdetermined inverse problem since a high dimensional signal model is recovered from a smaller number of measurements. The signal model is assumed to be sparse (or nearly sparse) since the forward model is taken to be the inverse of the redundant curvelet transform. Therefore, the interpolation approach minimizes a data misfit between the measurements and the forward model in addition to a sparse regularization term such as the ℓ_1 norm that captures sparsity of the signal model (Hennenfent and Herrmann, 2006).

Consider a seismic line with N_s sources, N_r receivers, and N_t time samples, where all sources share the same receivers. This scenario is becoming more feasible with the advent of receivers that are deployed via autonomous underwater nodes and wireless land-based geophones (Howie et al., 2008; Savazzi and Spagnolini, 2008). Note that our approach is not restricted to fixed geometries and can also be applied to streamer data acquisition utilizing the continuity of the waveforms across adjacent shot-gathers. The unknown fully-sampled seismic line can be reshaped into an N dimensional vector $\mathbf{f}$, where $N = N_s N_r N_t$. It is well known that seismic data admit sparse representations by curvelets that capture "wavefront

sets" efficiently (see e.g., Smith, 1998; Candès and Demanet, 2005; Candès et al., 2006a; Herrmann et al., 2008). Therefore, we wish to recover an approximation $\tilde{\mathbf{f}}$ of the discretized wavefield $\mathbf{f}$ from measurements

$$\mathbf{b} = \mathbf{RMf},$$

where $\mathbf{RM}$ is a sampling operator composed of the product of a restriction matrix $\mathbf{R}$ with a measurement basis matrix $\mathbf{M}$. The restriction matrix $\mathbf{R}$ is the selection matrix that specifies the sample locations that have data in them. The measurement matrix $\mathbf{M}$ represents the basis in which the measurements are taken and corresponds to the Dirac basis (identity matrix) in the missing trace interpolation scenario. In the case of simultaneous-source acquisition, $\mathbf{M}$ represents the multiplexing or blending operator.

Let $\mathbf{S}$ be a sparsifying operator that characterizes the transform domain of $\mathbf{f}$, such that $\mathbf{S} \in \mathbb{C}^{P \times N}$ with $P \geq N$. In the case of the redundant curvelet transform (Candès et al., 2006a), $\mathbf{S}$ is a tight frame with $P > N$ and $\mathbf{S}^H\mathbf{S} = \mathbf{I}$, and the transform domain representation $\mathbf{x}$ of $\mathbf{f}$ in $\mathbf{S}$ is not unique. The curvelet transform is highly redundant and when all angles and the finest scales are incorporated it could result in the frame expansion $\frac{P}{N} \approx 8$ when a 2D curvelet transform is used, or $\frac{P}{N} \approx 24$ when a 3D curvelet transform is used. This redundancy could easily become a computational impediment in 2D seismic surveys and especially when scaling the surveys to 3D seismic.

Let $\mathbf{A} := \mathbf{RMS}^H$, the measurements $\mathbf{b}$ can then be written as $\mathbf{b} = \mathbf{Ax}$, where $\mathbf{x}$ is the $\mathbf{S}$-transform of $\mathbf{f}$. We obtain the sparse approximation $\tilde{\mathbf{f}}$ of $\mathbf{f}$ by first finding the vector $\tilde{\mathbf{x}}$ that solves the sparse recovery problem with the underdetermined linear constraints

$$\mathbf{A}\tilde{\mathbf{x}} = \mathbf{b}, \tag{1}$$

and then computing $\tilde{\mathbf{f}} = \mathbf{S}^H\tilde{\mathbf{x}}$, where the superscript H denotes the Hermitian transpose.

Next, we formulate the corresponding sparse recovery problem that recovers the estimate $\tilde{\mathbf{x}}$ of the curvelet synthesis coefficients.

The sparse recovery problem

Let $\mathbf{x}$ be a P dimensional vector in $\mathbb{C}^P$ and let $\mathbf{b} \in \mathbb{C}^n$, $n \ll P$ represent the compressively sampled data of n measurements. The sparse recovery problem involves finding the solution $\mathbf{x}$ with the smallest number of non-zero entries that solves an underdetermined system of equations

$$\mathbf{b} = \mathbf{A}\mathbf{x}, \tag{2}$$

where $\mathbf{A} \in \mathbb{C}^{n \times P}$ represents the measurement matrix. When $\mathbf{x}$ is k-sparse—i.e., when there are only $k < n$ nonzero entries in $\mathbf{x}$— sparsity-promoting recovery can be achieved by solving the ℓ_0 minimization problem

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{u} \in \mathbb{C}^P} \|\mathbf{u}\|_0 \quad \text{subject to } \mathbf{b} = \mathbf{A}\mathbf{u}, \tag{3}$$

where $\tilde{\mathbf{x}}$ represents the sparse approximation of $\mathbf{x}$, and the ℓ_0 norm $\|\mathbf{x}\|_0$ is the number of non-zero entries in $\mathbf{x}$. The case that is more commonly observed in practice is when $\mathbf{x}$ is not sparse, but can be well approximated by a sparse vector, e.g., when the magnitudes of the sorted coefficients of $\mathbf{x}$ have a fast decay. More precisely, an N dimensional signal is said to be compressible if it can be well approximated by its largest $k \ll N$ coefficients. For example, a signal x is considered to be compressible if the sorted magnitudes of its coefficients decay according to a power law, i.e. the i^{th} largest coefficient will have a magnitude at most $|x[i]| \leq ci^{-p}$ with $p > 1$, where c is some scaling constant. Here, $x[i]$ is the i^{th} entry of the sorted coefficients of x. In this case, a multitude of techniques can be used to recover a stable and robust approximation of $\mathbf{x}$. We will focus on the performance on one

such approach in this paper, namely, recovery by ℓ_1 minimization. Note that ℓ_0 minimization is a combinatorial optimization problem, which quickly becomes computationally intractable as the size P of the problem increases. However, if it were possible to solve large scale ℓ_0 minimization problems in practice and if every $n \times n$ submatrix of $\mathbf{A}$ is invertible, then $\tilde{\mathbf{x}}$ would be equal to $\mathbf{x}$ when $k < n/2$ (Donoho and Elad, 2003).

Several greedy algorithms, such as orthogonal matching pursuit (OMP) (Pati et al., 1993; Tropp et al., 2007), and the antileakage Fourier transform (ALFT) (Xu et al., 2005) are also capable of finding the sparsest solution for systems of equations while still being computationally feasible. However, these greedy algorithms can only recover sparse signals with far fewer non-zero entries than what ℓ_0 minimization can handle. In fact, Tropp et al. (2007) has shown that given a signal $\mathbf{x}$ with sparsity level $k \lesssim n/\log(P/\rho) < n/2$ for some constant $\rho \in (0, 0.36)$ and if $\mathbf{A}$ is a Gaussian random matrix, then OMP succeeds in recovering sparse signals with probability $1-2\rho$. Moreover, these greedy algorithms typically only allow a few nonzero components to enter into the solution set for every matrix-vector multiply. While this may not be a problem for Fourier-based methods on small cubes, it quickly becomes a computational problem for curvelets that work on large data volumes.

Alternatively, the basis pursuit (BP) (Chen et al., 2001) convex optimization problem shown below, has demonstrated to enjoy better sparse recovery capabilities than greedy algorithms while still remaining computationally tractable. The BP problem is guaranteed to recover an estimate $\tilde{\mathbf{x}}$ for all signals $\mathbf{x} \in \mathbb{C}^P$ with sparsity $k \lesssim n/\log(P/n) < n/2$ when $\mathbf{A}$ is a Gaussian matrix with independent identically distributed (i.i.d.) entries and normalized columns (Candès et al., 2006b; Donoho, 2006). Moreover, the recovery is exact, i.e. $\tilde{\mathbf{x}} = \mathbf{x}$, when the measurement process is free of noise. The BP problem, also known as the ℓ_1

minimization problem, is given by

$$\tilde{\mathbf{x}} = \arg \min_{\mathbf{u} \in \mathbb{C}^P} \|\mathbf{u}\|_1 \quad \text{subject to } \mathbf{b} = \mathbf{A}\mathbf{u}, \tag{4}$$

where $\tilde{\mathbf{x}}$ represents an approximation of $\mathbf{x}$, and the ℓ_1 norm $\|\mathbf{u}\|_1 = \sum_{i=1}^{P} |\mathbf{u}(i)|$ is the sum of absolute values of the elements of a vector $\mathbf{u}$.

The BP problem typically finds a compressible or (under some conditions) the sparsest solution that explains the data exactly. So, if the initial data is sparse, or compressible, it is plausible that the solution of (BP) coincides with or approximates well the transform coefficients of the data. Indeed, this was mathematically proven to be the case if $\mathbf{A}$ obeys certain "incoherence" properties, formally the restricted isometry property (RIP) (Candès et al., 2006b), and if the original $\mathbf{x}$ is sufficiently sparse or compressible. Moreover, the recovery is stable and robust to measurement noise when $\mathbf{A}$ and $\mathbf{x}$ obey these properties. In fact, the BP problem and its robust extension

$$\tilde{\mathbf{x}} = \arg \min_{\mathbf{u} \in \mathbb{C}^P} \|\mathbf{u}\|_1 \quad \text{subject to } \|\mathbf{b} - \mathbf{A}\mathbf{u}\|_2 \leq \varepsilon, \tag{5}$$

are guaranteed to recover an approximation $\tilde{\mathbf{x}}$ to $\mathbf{x}$ for a wider range of measurement matrices $\mathbf{A}$ and under less strict conditions than the greedy algorithms mentioned above.

Remark: Gaussian matrices are not practical measurement matrices. However, Gaussian matrices have the measure-concentration property that allow for the proof of theoretical recovery guarantees. Therefore, when practical measurement matrices have the measure-concentration property (as in the case of uniform random subsampling of Fourier matrices) then the results derived for Gaussian matrices extend to the practical measurement matrices.

In particular, for general (not necessarily sparse) signals $\mathbf{x} \in \mathbb{C}^P$, Donoho (2006) and

Candès et al. (2006b) proved that the reconstruction error is bounded by

$$\|\tilde{\mathbf{x}} - \mathbf{x}\|_2 \leq C_0\varepsilon + \frac{C_1}{\sqrt{k}}\|\mathbf{x}_{T_0^c}\|_1, \tag{6}$$

where C_0 and C_1 are constants that depend on the condition number of submatrices of $\mathbf{A}$ of size $n \times k$, ε is an upper bound on the measurement noise variance, $T_0 = \text{supp}(\mathbf{x}|_k)$ is the index set (also called *support*) of the k largest in magnitude entries of $\mathbf{x}$, and T_0^c is the set complement of T_0, i.e. T_0^c is the set of entries with small magnitudes. The one-norm $\|\mathbf{x}_{T_0^c}\|_1$ therefore represents the error in the best k-term approximation of $\mathbf{x}$. For orthonormal bases, this best k-term approximation corresponds to selecting the k-largest magnitude transform coefficients. This approximation can be extended to redundant transforms by taking the k largest coefficients that solve the synthesis problem, i.e., that are obtained by solving a sparsifying program. Note that when T_0 is unique and $\mathbf{x}$ is k-sparse with all the non-zero entries lying in T_0, then the term $\|\mathbf{x}_{T_0^c}\|_1 = 0$ and the bound given by Equation 9 in the next section implies that the recovery by the BP problem should be exact, i.e., $\tilde{\mathbf{x}} = \mathbf{x}$, when the measurements $\mathbf{b}$ are not contaminated with noise. In other words, the result indicates that if $\mathbf{x}$ is k-sparse, then the reconstruction error is bounded only by the measurement noise. On the other hand, if $\mathbf{x}$ is a general (non-sparse) signal in $\mathbb{C}^P$, then the reconstruction error is bounded by the measurement noise plus the one-norm of the tail $\mathbf{x}_{T_0^c}$. Note that when the transform domain coefficients $\mathbf{x}$ are compressible, i.e. the magnitudes of the sorted coefficients have a fast decay, as in the case of the curvelet transform coefficients of seismic data, where the one-norm of the tail $\mathbf{x}_{T_0^c}$ is small.

Remark: Successful sparse recovery from the compressive measurements $\mathbf{b} = \mathbf{A}\mathbf{x}$, where $\mathbf{A} = \mathbf{R}\mathbf{M}\mathbf{S}^H$ requires that the sampling operator $\mathbf{RM}$ be incoherent with the sparsifying transform $\mathbf{S}$ (Candès and Tao, 2006). This incoherence can be achieved when the

samples are collected on a random subset of a regular grid, i.e. when the restriction operator $\mathbf{R}$ corresponds to a random restriction of the identity matrix. The random locations can generally be drawn from a uniform distribution. Otherwise, Hennenfent and Herrmann (2008) demonstrated that jittered sampling can improve the sparse recovery capabilities even further in applications where the signal to be recovered is bandlimited and reconstruction is performed in the curvelet domain. More precisely, the authors use jitter sampling to control the largest gap between subsequent samples especially since curvelets are localized in the physical domain. See Herrmann et al. (2012) for further discussion.

Interpolation with prior support information

In many situations, the data $\mathbf{f}$ exhibit continuity along one or more of its physical dimensions. In such situations, it is desirable to employ transforms that capitalize on this continuity in order to improve the sparsity of the coefficients $\mathbf{x}$. However, it is often the case that the dimensionality of the problem is too large and recovery is performed by first partitioning (or windowing) the seismic data volumes into frequency slices, or into common offset-azimuth gathers and then solving a sequence of individual subproblems. Moreover, the correlation of the support of the data partitions is better realized among frequency slices than among time-slices. When $\mathbf{f}$ is windowed across the dimensions of continuity, this continuity manifests itself as a high correlation in the support sets of the transform coefficients of the windowed data since the wavefields are repeated approximately in adjacent partitions. Therefore, sequentially recovering the windowed sections provides additional support information that could be incorporated in the recovery algorithm. However, the ℓ_1 minimization problem defined in Equation 4 does not incorporate prior information about the support of $\mathbf{x}$.

One approach that utilizes prior information in the recovery algorithm is to replace ℓ_1 minimization in Equation 4 with *weighted* ℓ_1 *minimization*

$$\min_{\mathbf{u}} \ \|\mathbf{u}\|_{1,\mathbf{w}} \text{ subject to } \mathbf{A}\mathbf{u} = \mathbf{b}, \tag{7}$$

where $\mathbf{w} \in [0,1]^P$ and $\|\mathbf{u}\|_{1,\mathbf{w}} := \sum_i \mathrm{w}_i |u_i|$ is the weighted ℓ_1 norm.

Consider the case where we are given a support estimate $\widetilde{T} \subset \{1, \ldots, P\}$ for $\mathbf{x}$ with a certain accuracy relative to the true support T_0 of the k largest in magnitude entries of $\mathbf{x}$. Friedlander et al. (2012) investigated the performance of weighted ℓ_1 minimization, as described in Equation 7, where the weights are assigned such that

$$\mathrm{w}_i = \begin{cases} 1, & i \in \widetilde{T}^c, \\ \gamma, & i \in \widetilde{T}. \end{cases} \tag{8}$$

Figure 1 illustrates the allocation of the weights given a support estimate set $\widetilde{T}$. Small weights $\gamma, 0 \leq \gamma \leq 1$, are applied to the set $\widetilde{T}$, while elements in the set $\widetilde{T}^c$ are assigned weights equal to 1. Such a selection of weights decreases the cost of having large nonzero entries on the set $\widetilde{T}$ since the weighted one norm of the vector $\mathbf{u}$ will be equal to

$$\|\mathbf{u}\|_{1,\mathbf{w}} = \gamma \|\mathbf{u}_{\widetilde{T}}\|_1 + \|\mathbf{u}_{\widetilde{T}^c}\|_1.$$

Consequently, solutions $\tilde{\mathbf{x}}$ with small entries outside of the set $\widetilde{T}$ are more likely to have smaller weighted one-norms than solutions with larger entries outside of $\widetilde{T}$. Therefore, using the weighted ℓ_1 norm in Equation 7 favours solutions that have small entries outside of $\widetilde{T}$.

The accuracy of the support estimate $\widetilde{T}$ plays a critical role in determining how well the solution to Equation 7 approximates the true signal $\mathbf{x}$ since the weighted ℓ_1 problem favours solutions that are supported on $\widetilde{T}$. It was proved by Friedlander et al. (2012) that the reconstruction error from weighted ℓ_1 minimization is bounded by

$$\|\tilde{\mathbf{x}} - \mathbf{x}\|_2 \leq C_0(\alpha, \gamma)\varepsilon + \frac{C_1(\alpha, \gamma)}{\sqrt{k}} \|\mathbf{x}_{T_0^c}\|_1, \tag{9}$$

where ε is an upper bound on the measurement noise variance, and $\alpha = \frac{|\widetilde{T} \cap T_0|}{|\widetilde{T}|}$ is a parameter that quantifies the accuracy of the support estimate and is defined by the proportion of entries in $\widetilde{T}$ that also lie in T_0. Here $|T|$ denotes the number of entries (cardinality) in a set T. Moreover, the constants $C_0(\alpha, \gamma)$ and $C_1(\alpha, \gamma)$ were shown to be smaller than the corresponding constants for standard ℓ_1 minimization when $\alpha > 0.5$ and $\gamma < 1$. The result indicates that when the support estimate is accurate enough, weighted ℓ_1 minimization outperforms standard ℓ_1 minimization (BP) in terms of accuracy, stability, and robustness.

WEIGHTED ℓ_1 MINIMIZATION FOR SEISMIC DATA INTERPOLATION

Seismic data is a discretization of the Green's function, which is a solution to the wave equation restricted to the surface where the receivers are located. As a result, seismic data organized in a seismic line exhibit continuity in the time/frequency dimension as well as continuity across the offset/azimuth directions. Several works in the seismic trace interpolation literature have taken advantage of waveform continuity by incorporating prior information in the recovery algorithms (Herrmann et al., 2000; Liu and Sacchi, 2004). For instance, Herrmann et al. (2000) used the un-aliased parabolic decomposition at low frequencies to guide the parabolic decomposition at higher frequencies. In the curvelet domain, the waveform continuity translates into a high correlation between the support sets of adjacent time/frequency slices and of adjacent common-offset/ common-azimuth slices. The high correlation is visible in the size of the intersection between the support sets of adjacent slices. In this section, we set up the weighted ℓ_1 minimization problem as a means to exploit this correlation among the supports and improve the performance of seismic data interpolation.

A general algorithm

Recall that our objective is to recover a high dimensional seismic data volume $\mathbf{f}$ by interpolating between a small number of measurements $\mathbf{b} = \mathbf{RMf}$ collected on an irregular grid defined by the restriction operator $\mathbf{R}$. The measurements therefore represent irregular, in particular, random samples from the high dimensional data volume. In order to cope with the large dimensionality of the problem, we tackle the problem by windowing the data along some dimension and then sequentially recovering the windowed partitions.

Let $\mathbf{b}^{(j)}$ be the subsampled measurements of the data $\mathbf{f}^{(j)}$ in partition j. The corresponding compressed sensing matrix used for sparse recovery is given by

$$\mathbf{A}^{(j)} = \mathbf{R}^{(j)}\mathbf{S}^H, \tag{10}$$

where $\mathbf{R}^{(j)}$ is the subsampling operator restricted to the j^{th} partition, and $\mathbf{S}$ is the 2D curvelet transform. Here we assumed that $\mathbf{M}^{(j)}$ is the identity matrix and removed it from the expression of $\mathbf{A}^{(j)}$. The classical approach of sparse recovery finds a sparse approximation of each windowed partition by solving the ℓ_1 minimization problem

$$\tilde{\mathbf{x}}^{(j)} = \arg\min_{\mathbf{u}} \|\mathbf{u}\|_1 \text{ s.t. } \mathbf{A}^{(j)}\mathbf{u} = \mathbf{b}^{(j)} \tag{11}$$

for each partition j independent of the remaining partitions. Note that although we restrict the partitions in this paper to 2D slices of the seismic line, our approach is general and can be trivially extended to higher dimensional partitions.

Alternatively, we propose to use the support information of the *previously* recovered partition in order to apply weights to a support estimate $\widetilde{T}$ of the adjacent partition and then solve a weighted ℓ_1 minimization problem to recover the subsampled partition. We present below a general algorithm that can be applied to arbitrary partitioning of the

same subsampled data. Again, the only requirement for improved recovery over standard ℓ_1 minimization is that adjacent partitions indexed by j should have sufficiently correlated support sets in the transform domain. Details of the algorithm are presented in the following subsections.

Algorithm 1 Weighted ℓ_1 recovery of seismic data.

1: **Input** $\mathbf{b}^{(j)} = \mathbf{A}^{(j)}\mathbf{x}^{(j)}$ for all j,

where $\mathbf{A}^{(j)} = \mathbf{R}^{(j)}\mathbf{S}^H$, and choose γ, k

2: **Output** $\tilde{\mathbf{x}}^{(j)}$

3: **Initialize** $j = 1$

$$\tilde{\mathbf{x}}^{(j)} = \arg\min_{\mathbf{u}} \|\mathbf{u}\|_1 \text{ s.t. } \mathbf{A}^{(j)}\mathbf{u} = \mathbf{b}^{(j)}$$

4: **loop**

5: $j = j + 1$

6: $\widetilde{T} = \text{supp}(\mathbf{S}\mathbf{S}^H\tilde{\mathbf{x}}^{(j-1)}|_k)$

7: Set

$$\mathrm{w}_i = \begin{cases} 1, & i \in \widetilde{T}^c \\ \gamma, & i \in \widetilde{T} \end{cases} \tag{12}$$

8: $\tilde{\mathbf{x}}^{(j)} = \arg\min_{\mathbf{u}} \|\mathbf{u}\|_{1,\mathbf{w}} \text{ s.t. } \mathbf{A}^{(j)}\mathbf{u} = \mathbf{b}^{(j)}$

9: **end loop**

Remark: Notice that in step 6 of Algorithm 1, the support estimate is chosen as the locations of the largest k analysis coefficients $\mathbf{S}\mathbf{S}^H\tilde{\mathbf{x}}^{(j-1)} = \mathbf{S}\tilde{\mathbf{f}}^{(j-1)}$ of the recovered partitions instead of the synthesis coefficients $\tilde{\mathbf{x}}^{(j-1)}$. This is due to the fact that the curvelet transform $\mathbf{S}$ is highly redundant, causing adjacent partitions to have synthesis coefficients with near but non-overlapping supports. The analysis coefficients, on the other hand, smear the nonzero entries in the synthesis coefficients, allowing the synthesis coefficients

of adjacent partitions to fall within the support of the analysis coefficients of previously recovered partitions. A similar approach was used by Saab et al. (2007) to define the weights in their curvelet-based Bayesian primary-multiple separation problem. Finally, we note that the parameter k can be set equal to the number of the largest in magnitude curvelet coefficients that contribute to a large percentage (e.g. 90%) of the signal energy.

Partitioning of irregularly subsampled seismic lines

There are several ways in which seismic data can be partitioned while still exploiting the continuity of the waveforms between partitions. We present in this section two partitioning scenarios. In the first scenario, we propose to partition seismic lines in the frequency domain and recover a sequence of frequency slices by sparse recovery scanning from the low frequencies to the high frequencies since the low frequencies are not aliased. In the second scenario, we propose partitioning the data into offset slices (or azimuth slices) and recovering a sequence of common-offset gathers (common-azimuth gathers) by sparse recovery scanning from the near-offsets to the far-offsets. The choice of the partitioning scheme can be important algorithmically. For example, the first scenario is beneficial when combined with wave-equation based inversion, whereas the second scenario is useful when combined with common-offset/azimuth migration.

Partitioning in the time/frequency domain

Consider the fully sampled seismic line from the Gulf of Suez illustrated in Figure 2a by a time slice in the source-receiver domain. Random subsampling of the seismic line is simulated by applying the mask shown in Figure 2b to the seismic line. This results in

the subsampled data shown in Figure 2c, corresponding to measurements collected from irregularly spaced receivers whose locations constitute a randomly chosen subset of the complete regular grid.

Sparse recovery in this setup takes advantage of the sparsity of the curvelet synthesis coefficients of a frequency slice in the source-receiver domain. Therefore, we first transform the data from the time domain to the frequency domain by applying the N_t DFT to the time axis of every source-receiver coordinate. Let $\mathbf{f}^{(0)}$ correspond to the lowest frequency slice of the resulting Fourier transformed seismic line. A sparse approximation $\tilde{\mathbf{f}}^{(0)}$ is obtained by solving the sparse recovery problem defined in Equation 4. The sparse recovery problem is set up by defining $\mathbf{b}^{(0)} = \mathbf{R}^{(0)}\mathbf{f}^{(0)} = \mathbf{R}^{(0)}\mathbf{S}^H\mathbf{x}^{(0)}$ as the subsampled lowest frequency slice, $\mathbf{S}$ as the two dimensional curvelet transform in the source-receiver (or midpoint-offset) domain, and $\mathbf{x}^{(0)}$ as the curvelet synthesis coefficients of $\mathbf{f}^{(0)}$. The operator $\mathbf{R}^{(j)}$ is the restriction matrix which maps a fully sampled frequency slice $\mathbf{f}^{(j)}$ to the measurements $\mathbf{b}^{(j)}$ for all frequency slices indexed by j.

The next step is to identify the support $\widetilde{T}^{(0)}$ of $\mathbf{S}\tilde{\mathbf{f}}^{(0)}$ and use it as a support estimate for the adjacent frequency slice $\mathbf{f}^{(1)}$ to solve the weighted ℓ_1 minimization problem in Equation 7 with weights γ applied to the set $\widetilde{T}^{(0)}$. Subsequent frequency slices are then recovered by repeating the process of using the support of the analysis coefficients of previous frequency slices to estimate the support of the subsequent frequency slice. Again, the full recovery of the seismic line is performed by scanning from the low frequencies to the high frequencies.

Recovery using weighted ℓ_1 minimization benefits from this setup since the support sets of the curvelet analysis coefficients of adjacent frequency slices are highly correlated. We

demonstrate this fact by plotting the parameter

$$\alpha_j = \frac{|T_j \cap T_{j+1}|}{|T_j|}, \tag{13}$$

where $T_j = \text{supp}(\mathbf{x}^{(j)})$ is the support of the synthesis coefficients of the frequency slice indexed by j, T_{j+1} is the support of $\mathbf{x}^{(j+1)}$. The accuracy parameter α_j is plotted in blue in Figure 3 for a seismic line from the Gulf of Suez. Notice that the curve is mostly above the 0.5 threshold for which weighted ℓ_1 minimization guarantees improved recovery. The analysis coefficients as they are calculated from the reconstruction of the synthesis coefficients will have peaks centered around the synthesis coefficients with a time-frequency localized spread. The synthesis coefficients of the next slice might not coincide fully with the analysis region of the previous slice. However, due to the high redundancy of the curvelet transform, weighting the analysis regions makes it possible to find synthesis coefficients with a slightly shifted support that overlaps with the analysis support that give a similar physical domain recovery.

As mentioned earlier, the structure of the subsampling matrix as well as the relation of the alignment of the curvelet transform with respect to the wavefields in the data play an important role in determining the quality of the reconstruction from ℓ_1 minimization. To investigate this avenue, we resort the data and perform the recovery along frequency slices in the midpoint-offset domain. Figures 4a, 4b, and 4c show the same time slice of Figure 2a with the corresponding mask and subsampled data transformed into the midpoint-offset domain.

Similar to the source-receiver domain, the support sets of the analysis coefficients of adjacent frequency slices are also highly correlated in the midpoint-offset domain. This can be observed in the value of α_j illustrated by the green curve in Figure 3. Since $\alpha_j > 0.5$ for

most values of j, the figure suggests that solving weighted ℓ_1 problems sequentially along frequency slices could improve the reconstruction quality over standard ℓ_1.

Partitioning in the offset/azimuth domain

In this scenario, we set up the weighted ℓ_1 problem in order to take advantage of the continuity of seismic data across common-offset gathers. The evidence for this continuity can be seen in the high correlation between the supports of analysis curvelet coefficients of adjacent offset gathers in the time-midpoint domain. To illustrate the high correlation in support sets of adjacent offsets, let $\mathbf{f}^{(j)}$ and $\mathbf{f}^{(j+1)}$ be two adjacent offset gathers in a seismic line indexed by j and $j+1$. As before, we illustrate the level of overlap between the synthesis coefficients of the two adjacent offset gathers by plotting the quantity α_j in Figure 5. Note that the synthesis coefficients are now defined for a 2D curvelet transform in the time-midpoint domain.

Next we set up the weighted ℓ_1 problem for partitions in the offset domain. Let $\mathbf{f}^{(0)}$ denote the zero-offset or near-offset gather in a seismic line. The sparse recovery problem can now be setup by defining $\mathbf{b}^{(0)} = \mathbf{R}^{(0)}\mathbf{f}^{(0)} = \mathbf{R}^{(0)}\mathbf{S}^H\mathbf{x}^{(0)}$ as the subsampled zero-offset gather, $\mathbf{S}$ as the two dimensional curvelet transform in the time-midpoint domain, and $\mathbf{x}^{(0)}$ as the curvelet synthesis coefficients of $\mathbf{f}^{(0)}$. The operator $\mathbf{R}^{(j)}$ is the restriction matrix which maps a fully sampled offset gather $\mathbf{f}^{(j)}$ to the measurements $\mathbf{b}^{(j)}$ for all offset gathers indexed by j.

To illustrate this setup, let $\mathbf{d}^{(0)}$ be the center column of the subsampling mask in Figure 4b, then the restriction operator $\mathbf{R}^{(0)}$ of the zero-offset gather is given by

$$\mathbf{R}^{(0)} = \mathbf{d}^{(0)} \otimes \mathbf{I}_{N_t}, \tag{14}$$

where $\otimes$ is the Kronecker (or Tensor) product, $\mathbf{I}_{N_t}$ is the identity matrix of size $N_t \times N_t$, and N_t is the number of time samples in the seismic line. Note that the operator $\mathbf{R}^{(0)}$ acts on the vectorized offset gather, where the time samples for each midpoint are stacked above each other. Hence, the formulation above ensures that every sample in $\mathbf{d}^{(0)}$ is multiplied by all the time samples of an individual midpoint. This results in an offset gather with missing columns as illustrated in Figure 6a and 6b. The solution of the sparse recovery problem in Equation 4 yields a sparse approximation $\tilde{\mathbf{x}}^{(0)}$ of the curvelet synthesis coefficients $\mathbf{x}^{(0)}$.

Adjacent offsets are then recovered by solving a weighted ℓ_1 minimization problem using the support of the near-offset gather as support estimate. Similar to the frequency partitioning scenario, let $\widetilde{T}_0$ be the support of $\mathbf{S}\tilde{\mathbf{f}}^{(0)}$, the synthesis coefficients $\mathbf{x}^{(\pm\mathbf{1})}$ of the adjacent offset gathers are then found by solving Equation 7 with support estimate $\widetilde{T}_0$ and weight γ. As before, this process is repeated over the remaining offsets moving out towards the far offsets while using the recovered support of previous adjacent offsets.

RESULTS

We tested the performance of weighted ℓ_1 minimization on recovering a seismic line from the Gulf of Suez with 50% randomly subsampled receivers using the mask shown in Figure 2b. The data are acquired from a real marine streamer from the Gulf of Suez region, 12.5m shot spacing, and 12.5m receiver spacing. The near-offset traces have been interpolated by Radon domain interpolation and converted to fixed-spread geometry using reciprocity. Preprocessing only involved a direct-wave mute and denoising by low-pass filtering. The seismic line at full resolution has $N_s = 178$ sources, $N_r = 178$ receivers, and $N_t = 500$ time samples acquired with a sampling interval of 4 ms. Figures 7a and 7b show a fully sampled shot gather extracted from our data set and its corresponding F-K spectrum. The

figures show that the seismic line has relatively simple structure. The same shot gather subsampled using our random subsampling mask can be seen in Figure 8a along with its F-K spectrum in Figure 8b. Notice how the spectrum is flooded with noise artifacts due to the random subsampling which results in spectral leakage. We note that these noise artifacts are primarily due to repetitions of the strong resonating frequencies in the fully sampled data set and to a lesser extent due to aliasing. The seismic line contains samples collected in a 2s temporal window with a maximum frequency of 125 Hz. To access frequency slices data are Fourier transformed along the time axis. We solve the ℓ_1 and weighted ℓ_1 minimization problems using the SPGL1 solver (van den Berg and Friedlander, 2008) and limit the number of iterations to 500 for each problem.

In each of the weighted ℓ_1 problems, the support estimate set $\widetilde{T}$ of a partition j is derived from the analysis coefficients of the previously recovered partition. More precisely, the set $\widetilde{T}$ of partition j is the index set of the largest k entries of the analysis coefficients of the partition $j-1$, where k is chosen to be the number of largest analysis coefficients that contribute 90% of the signal energy. The 90% figure was chosen based on empirical evidence that showed that it leads to the best recovery. Moreover, in all the weighted ℓ_1 problems the weight γ is set to 0.3, also based on experimental observations.

Recovery in the source-receiver domain

In this first setup, the data are organized in the frequency-source-receiver domain as shown in Figure 2a. We start by recovering the lowest frequency slice using ℓ_1 minimization and follow that with weighted ℓ_1 for the subsequent frequency slices using the recovered support of the previous frequency as a support estimate.

Figures 9a and 9b show a fully sampled and the corresponding subsampled shot gather, respectively. The shot gather corresponds to shot number 84 of the seismic line from the Gulf of Suez. Figures 10a and 10c show the reconstructed shot gathers using ℓ_1 minimization and weighted ℓ_1 minimization, respectively. The error plots of both reconstructions are shown in Figures 10b and 10d, and the F-K spectra of the reconstructed shot gather number 89 are illustrated for the two reconstruction techniques in Figures 11a and 11b.

The figures show that weighted ℓ_1 minimization across frequencies in the source-receiver domain preserves reflections at later times better than ℓ_1 minimization. The error plots also show that the magnitude of the reconstruction error of weighted ℓ_1 minimization is smaller than that of standard ℓ_1. Moreover, Figures 11a and 11b demonstrate the superior denoising and anti-aliasing capabilities of the weighted ℓ_1 approach compared to standard ℓ_1 minimization as witnessed by the reduced artifacts visible in the lower and mid-range frequencies. However, one can still identify subsampling artifacts in the shot gather recovered by weighted ℓ_1, especially at later arrivals and higher frequencies. Finally, we note that these results can be further improved by using the minimal velocity constraint as described by Hennenfent and Herrmann (2005). Such constraints remove the occurrence of vertical artifacts from the recovered data which correspond to highly coherent curvelets.

Recovery in the midpoint-offset domain

In the second setup, we first transform the seismic line into the frequency-midpoint-offset (MH) domain as shown in Figure 4a. Following the set-up described in the previous section, Figures 12a and 12c show the reconstructed shot gathers using ℓ_1 minimization and weighted ℓ_1 minimization, respectively. The error plots of both reconstructions are shown

in Figures 12b and 12d.

Compared to recovery in the source-receiver domain (Figures 10a and 10c), solving the sparse recovery problems in the midpoint-offset domain significantly improves the reconstruction quality for both ℓ_1 minimization and weighted ℓ_1 minimization. It can be seen that the subsampling artifacts are reduced in both cases, while the reconstructed shot gather from weighted ℓ_1 minimization has fully preserved the reflected wavefronts even at later time samples. This improvement in the recovery in the midpoint-offset domain compared to the source-receiver domain has the potential to arise from two factors: 1) a sparser representation of the curvelet coefficients in the midpoint-offset domain, and 2) a more "incoherent" interaction between the subsampling operator in the midpoint-offset domain and the corresponding curvelet transform (Trad, 2009). We inspected the decay rate of the curvelet synthesis coefficients of frequency slices in both the source-receiver and midpoint-offset domains and found that the decay rate is faster in the midpoint-offset domain compared to the source-receiver domain. We would not expect the sampling operator in the midpoint-offset domain to be more incoherent with the curvelet transform matrix than the sampling operator in the source-receiver domain. This is due to the nearly isotropic nature of the curvelet transform and the fact that the subsampling operator in the midpoint-offset domain is simply a 45 degree rotation of the same operator in the source-receiver domain. However, the operator in the midpoint-offset domain also has a large null-space since the sampling operator also removes the midpoint-offset combinations (corners) where there are no data. Both of these factors may have resulted in the improved recovery performance in the midpoint-offset domain.

Recovery in the time-midpoint domain

In the third setup, we consider offset gathers in the time-midpoint domain as shown in Figure 6a. Following the reconstruction set-up described in the previous section, we start by recovering the near-offset gather using standard ℓ_1 minimization and follow that with weighted ℓ_1 for the adjacent offset gathers. Figures 13a and 13b show the reconstruction from weighted ℓ_1 minimization of the same shot gather illustrated above. It can be seen from the figures that the recovery is comparable to that of weighted ℓ_1 applied to frequency slices in the midpoint-offset domain. One interesting observation can be seen when we consider the error artifacts shown in Figure 13b. The recovered shot gather does not contain the artifacts before the first arrivals that we see at farther offsets in the frequency based recovery. However, the error is slightly more significant in the reflections at later times than that seen in the weighted ℓ_1 recovery in the midpoint-offset domain.

Note that this behaviour is not necessarily seen across all shot gathers. Figures 14a and 14b show the reconstruction of shot gather number 20 using weighted ℓ_1 minimization in the midpoint-offset domain and in the time-midpoint domain, respectively. It can be seen from these figures that for this specific shot gather, the offset based recovery is a better reconstruction than the frequency based recovery.

Discussion of the results

The results shown above demonstrate that weighted ℓ_1 minimization consistently improves the recovery of standard ℓ_1 minimization applied in the same physical domain. However, it would be truly significant if the recovery of weighted ℓ_1 is comparable to that of ℓ_1 minimization when a 3D transform is used. To make this comparison, we recovered the

seismic line from the same subsampled data using two 3D transforms: the first is the Kronecker product of a 2D curvelet transform applied in the source-receiver domain with a 1D wavelet transform in the time domain, and the second is a 3D curvelet transform applied in the time-source-receiver domain. Figures 15a and 15c show the reconstruction quality of shot gather number 84 using the combined curvelet-wavelet transform and the quality using the 3D curvelet transform, respectively, as well as the corresponding error plots. The figures show that the recovery using the 3D transforms does not improve much on the sequential recovery using the 2D curvelet transform applied in the source-receiver domain.

Figure 16 compares the SNRs of recovered shot gathers of all the ℓ_1 minimization setups listed above. The figure shows that the best recovery is achieved by both the sequential recovery across frequencies in the midpoint-offset domain and by the 3D curvelet based recovery. In particular, the 3D curvelet based recovery is better than the recovery of the sequential 2D curvelet setup applied in the source-receiver domain. Note that we applied the 3D curvelet transform in the time-source-receiver domain on a slightly smaller data set (168 source, 168 receivers, 500 time samples) because the memory requirements of running the full 3D curvelet sparse recovery problem (178 sources, 178 receivers, 500 time samples) exceeded the capabilities of the hardware infrastructure we have available. For the same reason, we did not test the performance of the sparse recovery problem using a 3D curvelet transform applied to the time-midpoint-offset domain since the dimensionality of this problem is twice size of that in the time-source-receiver domain. However, we would expect the 3D curvelet recovery in the time-midpoint-offset domain to have a similar improvement in performance over the sequential 2D curvelet recovery in the midpoint-offset as that seen in the time-source-receiver domain.

Next, we compare the recovery of the standard ℓ_1 schemes with the weighted ℓ_1 schemes. Figure 17 shows the SNRs of all shot gathers recovered using the different weighted ℓ_1 set-ups and the best of the standard ℓ_1 set-ups. The plots demonstrate that the recovery in the midpoint-offset domain is consistently better than in the source-receiver domain. We attribute this improvement to the modified structure of the subsampling operator in the midpoint-offset domain and its better incoherence with the corresponding 2D curvelet transform compared to the source-receiver domain. However, the surprising observation is that weighted ℓ_1 minimization in the midpoint-offset domain and in the time-midpoint domain consistently outperform the 2D and 3D transform-based ℓ_1 recovery setups. Based on the plots in Figure 17 we can rank the different sparse recovery techniques in terms of their respective recovery performance in Table 1 below.

We consider the first two scenarios in the list above to be almost tied for first place. In fact, the artifacts that appear before the first arrivals in the first scenario could be removed by post-processing, which should improve the recovered SNR. Also, the last two scenarios can also be considered to be tied for fifth place. The astonishing observation is that the 2D weighted ℓ_1 schemes achieve a superior reconstruction compared to the 3D standard ℓ_1 schemes at a fraction of the memory requirements of the 3D schemes. In fact, the improvement in the recovered SNR is quite remarkable and for some shot gathers it approaches 8dB (see e.g., shot number 98).

Finally, we note that the weighted ℓ_1 scenarios described in this paper are only examples of a multitude of partitioning and weighting approaches that could be used to recover seismic data. For example, the restriction to 2D partitions only illustrate the effectiveness of our approach which is in fact more general and can be trivially extended to higher dimensional partitions. Moreover, the first iteration of Algorithm 1 does not have to be performed on

the lowest frequency slice (or the near-offset gather), instead the algorithm can be initialized at the partition with the highest energy content or the partition with least noise artifacts. The ordering of the partitions may in some cases, such as the offset domain or the shot domain, affect the recovery performance of the early partitions. This effect can be mitigated by performing a forward and backward pass over the partitions. In fact, performing a backward pass over the partitions can be perceived as an iterative re-weighted application of this algorithm, which is guaranteed to improve the recovery performance over a single pass. Moreover, applying normal move-out (NMO) correction on the offset slices before recovery using weighted ℓ_1 minimization can only improve the recovery performance of our approach, since NMO correction would flatten the primaries and increase the correlation between the supports of adjacent offset gathers.

CONCLUSIONS

We proposed the use of weighted one-norm minimization in a framework for the recovery of missing-traces caused by physical obstacles or deliberate subsampling in otherwise regularly-sampled seismic wavefields. Weighted one-norm minimization exploits correlations between the locations of significant coefficients of different partitions, e.g., shot records, common-offset gathers, common-azimuth gather, or frequency slices of the acquired data. We demonstrated through simulated experiments performed on frequency slices and offset gathers of a Gulf of Suez seismic line that our method is superior to standard ℓ_1 minimization in terms of reconstruction quality and computational memory requirements. While we did not run an experiment on azimuth partitions, the extension of the method to this case is trivial. Our framework can also be extended to the seismic regularization problem by simply replacing the equispaced curvelet transform with a nonequispaced transform as

in the work of Hennenfent et al. (2010) and maintaining the exact same structure of our algorithm.

ACKNOWLEDGMENTS

We would like to thank the authors of CurveLab (curvelet.org), a toolbox implementing the Fast Discrete Curvelet Transform, Madagascar (rsf.sf.net), a package for reproducible computational experiments, SPGℓ_1 (cs.ubc.ca/labs/scl/spgl1), SPOT (http://www.cs.ubc.ca/labs/scl/spot/), a suite of linear operators and problems for testing algorithms for sparse signal reconstruction, and pSPOT, SLIM's parallel extension of SPOT. The Gulf of Suez dataset was generously provided by Eric Verschuur. This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BP, Chevron, ConocoPhillips, Petrobras, Total SA, and WesternGeco.

REFERENCES

Berkhout, A. J., and D. J. Verschuur, 1997, Estimation of multiple scattering by iterative inversion, part I: theoretical considerations: Geophysics, **62**, 1586–1595.

Candès, E. J., and L. Demanet, 2005, The curvelet representation of wave propagators is optimally sparse: Comm. Pure Appl. Math, **58**, 1472–1528.

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

Candès, E. J., J. Romberg, and T. Tao, 2006b, Stable signal recovery from incomplete and inaccurate measurements: Comm. Pure Appl. Math., **59**, 1207–1223.

Candès, E. J., and T. Tao, 2006, Near-optimal signal recovery from random projections: Universal encoding strategies: Information Theory, IEEE Transactions on, **52**, 5406 – 5425.

Chen, S. S., D. L. Donoho, and M. A. Saunders, 2001, Atomic decomposition by basis pursuit: SIAM Journal on Scientific Computing, **43**, 129–159.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 467–481.

———, 1992, Earth soundings analysis: Processing versus inversion: Blackwell Scientific publishing.

Donoho, D., and M. Elad, 2003, Optimally sparse representation in general (nonorthogonal)dictionaries via ℓ^1 minimization: PNAS, **100**.

Donoho, D. L., 2006, Compressed sensing: IEEE Trans. Inform. Theory, **52**, 1289–1306.

Duijndam, A. J. W., M. A. Schonewille, and C. O. H. Hindriks, 1999, Reconstruction of band-limited signals, irregularly sampled along one spatial direction: Geophysics, **64**, 524–538.

Ferguson, R., 2010, Efficient simultaneous wave-equation statics and trace regularization by series approximation: Geophysics, **75**, U19–28.

Friedlander, M., H. Mansour, R. Saab, and O. Yilmaz, 2012, Recovering compressively sampled signals using partial support information: IEEE Transactions on Information Theory, **58**, 1122 – 1134.

Hampson, D., 1986, Inverse velocity stacking for multiple elimination: J. Can. Soc. Expl. Geophys., **22**, 44–55.

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

Hennenfent, G., and F. J. Herrmann, 2005, Sparseness-constrained data continuation with frames: applications to missing traces and aliased signals in 2/3-d: SEG Technical Program Expanded Abstracts, SEG, SEG, 2162–2165.

———, 2006, Application of stable signal recovery to seismic data interpolation: SEG Technical Program Expanded Abstracts, SEG, SEG, 2797–2801.

———, 2008, Simply denoise: wavefield reconstruction via jittered undersampling: Geophysics, **73**, V19–V28.

Herrmann, F. J., M. P. Friedlander, and O. Yilmaz, 2012, Fighting the curse of dimensionality: Compressive sensing in exploration seismology: Signal Processing Magazine, IEEE, **29**, 88–100.

Herrmann, F. J., and G. Hennenfent, 2008, Non-parametric seismic data recovery with curvelet frames: Geophysical Journal International, **173**, 233–248.

Herrmann, F. J., P. P. Moghaddam, and C. C. Stolk, 2008, Sparsity- and continuity-promoting seismic imaging with curvelet frames: Journal of Applied and Computational

Harmonic Analysis, **24**, 150–173. (doi:10.1016/j.acha.2007.06.007).

Herrmann, P., T. Mojesky, M. Magasan, and P. Hugonnet, 2000, De-aliased, high-resolution Radon transforms: 70th Ann. Internat. Mtg., SEG, Expanded Abstracts, Soc. Expl. Geophys., Expanded abstracts, 1953–1956.

Howie, J., P. Mahob, D. Shepherd, and G. Beaudoin, 2008, Unlocking the full potential of atlantis with obs nodes: Presented at the SEG Expanded Abstracts, SEG.

Liu, B., and M. D. Sacchi, 2004, Minimum weighted norm interpolation of seismic records: Geophysics, **69**, 1560–1568.

Mosher, C., S. Kaplan, and F. Janiszewski, 2012, Non-uniform optimal sampling for seismic survey design: Presented at the , EAGE, 74th EAGE Conference & Exhibition.

Naghizadeh, M., and M. Sacchi, 2010, Hierarchical scale curvelet interpolation of aliased seismic data: SEG Expanded Abstracts, SEG, 3656 – 3661.

Neelamani, R. N., C. E. Krohn, J. R. Krebs, J. K. Romberg, M. Deffenbaugh, and J. E. Anderson, 2010, Efficient seismic forward modeling using simultaneous random sources and sparsity: Geophysics, **75**, WB15–WB27.

Pati, Y. C., R. Rezaiifar, and P. S. Krishnaprasad, 1993, Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition: Proceedings of the 27 th Annual Asilomar Conference on Signals, Systems, and Computers, 40–44.

Saab, R., D. Wang, O. Yilmaz, and F. Herrmann, 2007, Curvelet-based primary-multiple separation from a bayesian perspective: Presented at the SEG International Exposition and 77th Annual Meeting.

Sacchi, M., and T. Ulrych, 1996, Estimation of the discrete fourier transform, a linear inversion approach: Geophysics, **61**, 1128–1136.

Savazzi, S., and U. Spagnolini, 2008, Wireless geophone networks for high-density land

acquisition: Technologies and future potential: The Leading Edge, **27**, 882–886.

Smith, H. F., 1998, A Hardy space for Fourier integral operators.: J. Geom. Anal., **8**, 629–653.

Smith, P., I. Scott, and T. Traylen, 2012, Simultaneous time-lapse binning and regularization for 4d data: Presented at the 74th Ann. Internat. Mtg., EAGE, Eur. Ass. of Geosc. and Eng., Expanded abstracts.

Spitz, S., 1991, Seismic trace interpolation in the f-x domain: Geophysics, **56**, 785–794.

Symes, W. W., 2007, Reverse time migration with optimal checkpointing: Geophysics, **72**, SM213–SM221. (10.1190/1.2742686).

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

Trad, D., 2009, Five-dimensional interpolation: Recovering from acquisition constraints: Geophysics, **74**, V123–132.

Tropp, J. A., Anna, and C. Gilbert, 2007, Signal recovery from random measurements via orthogonal matching pursuit: IEEE Trans. Inform. Theory, **53**, 4655–4666.

van den Berg, E., and M. P. Friedlander, 2008, Probing the Pareto frontier for basis pursuit solutions: SIAM Journal on Scientific Computing, **31**, 890–912.

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

Xu, S., Y. Zhang, D. Pham, and G. Lambaré, 2005, Antileakage fourier transform for seismic data regularization: Geophysics, **70**, V87V95.

Zwartjes, P., 2005, Fourier reconstruction with sparse inversion: PhD thesis, Delft university of technology.

LIST OF TABLES

1 Ranking the different partitioning and reconstruction schemes according to the reconstructed SNR.

LIST OF FIGURES

1 Illustration of a signal x and the weight vector $\mathbf{w}$ showing the true support set T_0 and the support estimate $\widetilde{T}$. Small weights γ are applied to the entries in the set $\widetilde{T}$ in order to decrease the cost of assigning high valued coefficients on this set during the recovery.

2 (a) Example of a high resolution time slice at $t = 1.4$ s in the source-receiver (SR) domain, (b) the random subsampling mask where the black lines correspond to the locations of inactive receivers, and (c) the subsampled time slice. The subsampling ratio is 50%.

3 Example of the accuracy achieved from using adjacent frequencies to predict the support of the curvelet coefficients in a seismic line.

4 (a) Example of a high resolution time slice at $t = 1.4$ s in the midpoint-offset (MH) domain, (b) the resulting random subsampling mask where the black lines correspond to the locations of inactive receivers, and (c) the subsampled time slice in MH. The subsampling ratio is 50%.

5 Example of the accuracy achieved from using adjacent offsets to predict the support of the curvelet coefficients in a seismic line.

6 Example of a fully sampled (top) and subsampled (bottom) zero offset gather with missing columns arising from the diagonal entries of the source-receiver subsampling mask.

7 (a) Original shot gather number 89 from a seismic line from the Gulf of Suez. (b) Corresponding F-K-spectrum of shot gather number 89.

8 (a) Subsampled shot gather number 89 from a seismic line from the Gulf of Suez. (b) Corresponding F-K spectrum of the subsampled shot gather showing the full spectrum noise and aliasing resulting from the subsampling.

9 (a) Shot gather number 84 from a seismic line from the Gulf of Suez. (b) Subsampled shot gather using column 84 from the mask in Figure 2b.

10 (a) Recovered shot gather using ℓ_1 minimization and (c) using weighted ℓ_1 minimization in the SR domain. Error plots showing the difference between the original shot gather and the reconstruction from (b) ℓ_1 minimization in the SR domain, and (d) weighted ℓ_1 minimization in the SR domain using the support derived from the previous frequency.

11 F-K spectra of shot gather 89 after reconstruction using (a) ℓ_1 minimization in the source-receiver domain, and (b) weighted ℓ_1 minimization in the source-receiver domain utilizing curvelet support correlations across frequency slices.

12 (a) Recovered shot gather using ℓ_1 minimization and (c) using weighted ℓ_1 minimization in the MH domain. Error plots showing the difference between the original shot gather and the reconstruction from (b) ℓ_1 minimization in the MH domain, and (d) weighted ℓ_1 minimization in the MH domain using the support derived from the previous frequency.

13 (a) Recovered shot gather using weighted ℓ_1 minimization across offset-gathers in the time-midpoint domain, and (b) the corresponding error plot of the recovered shot gather.

14 Recovered shot gather number 20 using weighted ℓ_1 minimization (a) across frequency slices in the midpoint-offset domain, and (b) across offset-gathers in the time-midpoint domain.

15 (a) Recovered shot gather using ℓ_1 minimization with a 3D transform composed of a 2D curvelet transform applied in the source-receiver domain and a wavelet transform applied in the time domain, and (c) using a 3D curvelet transform in the source-receiver-time domain. The right column figures (b) and (d) illustrate the corresponding error plots of the recovered shot gathers in (a) and (b).

16 Comparison of the SNRs achieved by ℓ_1 minimization in recovering shot gathers from the Suez seismic line applied to both the source-receiver and midpoint-offset domains.

The plots also include recovery results from 3D volume reconstruction using first a 2D curvelet transform in the source-receiver domain combined by a Kronecker product with a wavelet transform in the time domain, and second a 3D curvelet transform.

17 Comparison of the SNRs achieved by ℓ_1 and weighted ℓ_1 minimization in recovering shot gathers from the Suez seismic line applied to both the source-receiver and midpoint-offset domains.

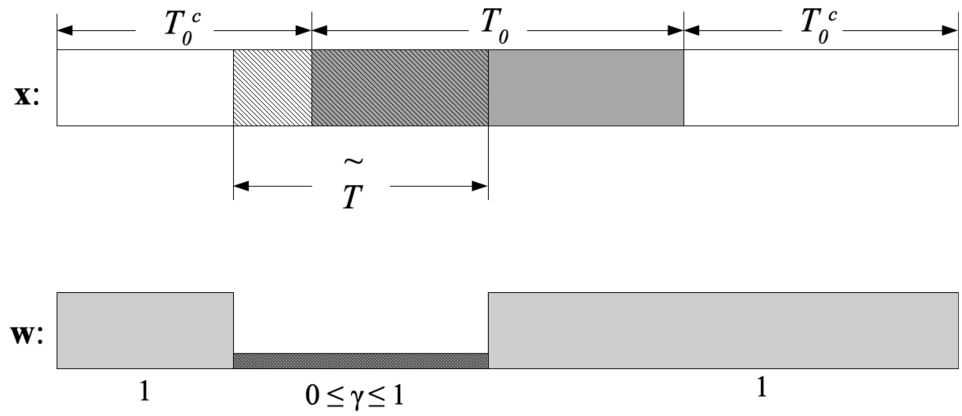
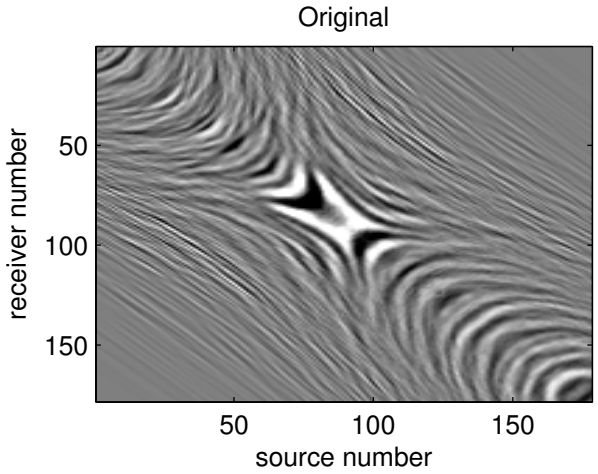
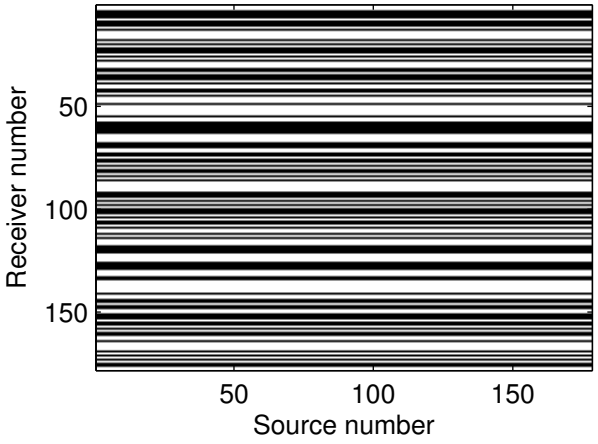


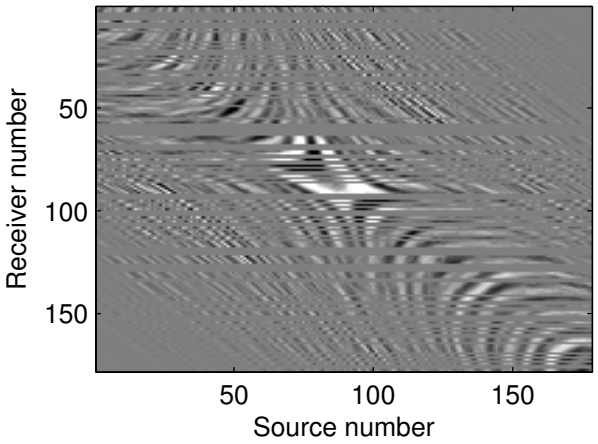
Figure 1: Illustration of a signal x and the weight vector $\mathbf{w}$ showing the true support set T_0 and the support estimate $\widetilde{T}$. Small weights γ are applied to the entries in the set $\widetilde{T}$ in order to decrease the cost of assigning high valued coefficients on this set during the recovery.



a



b



c

Figure 2: (a) Example of a high resolution time slice at $t = 1.4$ s in the source-receiver

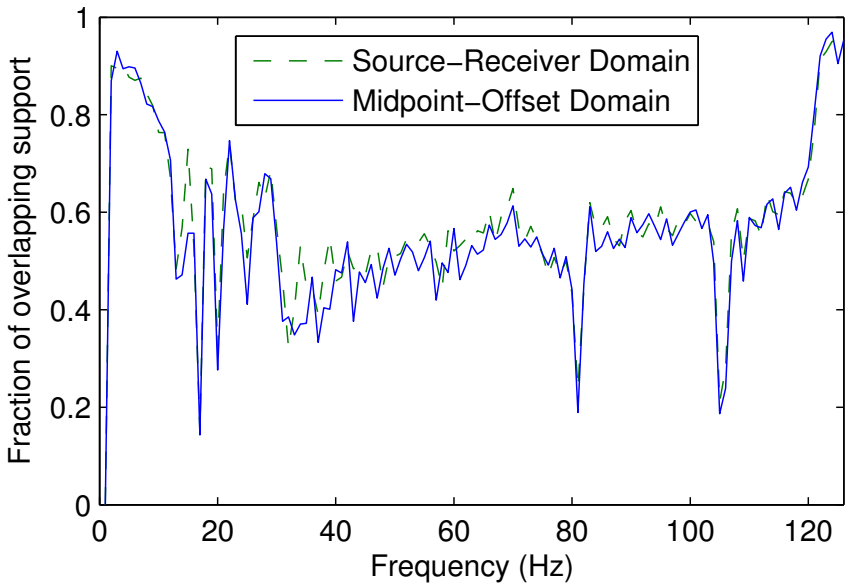
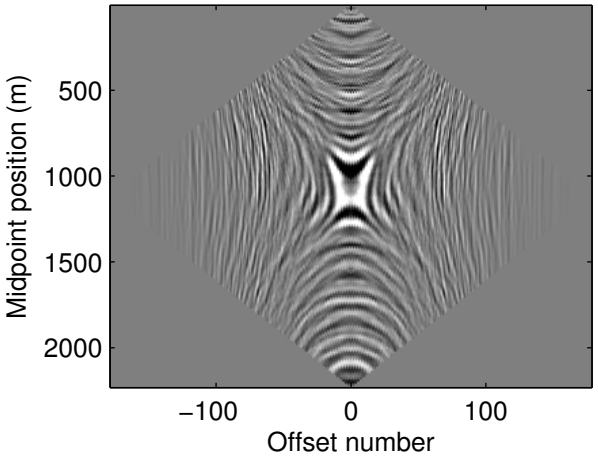
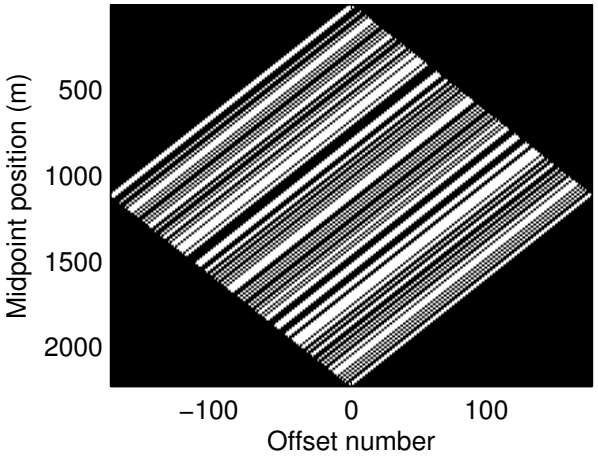


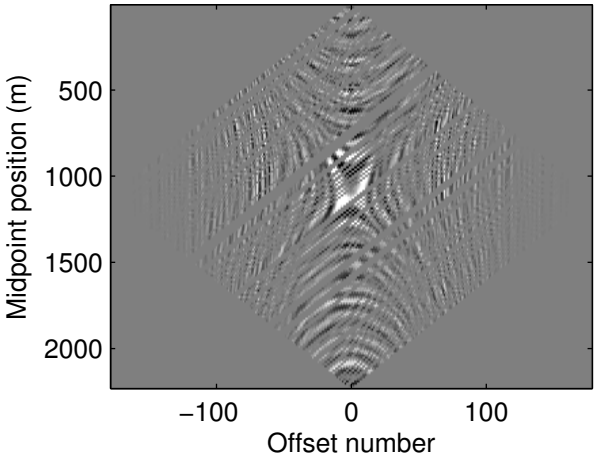
Figure 3: Example of the accuracy achieved from using adjacent frequencies to predict the support of the curvelet coefficients in a seismic line.



a



b



c

Figure 4: (a) Example of a high resolution time slice at $t = 1.4$ s in the midpoint-offset (MH)

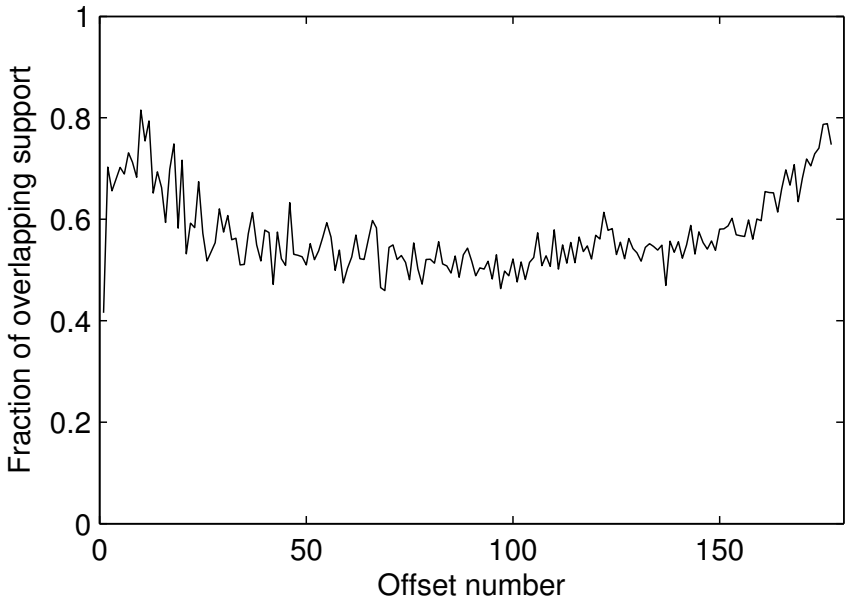
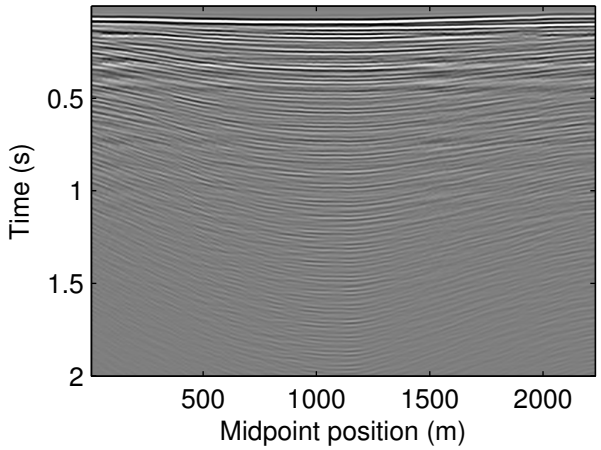
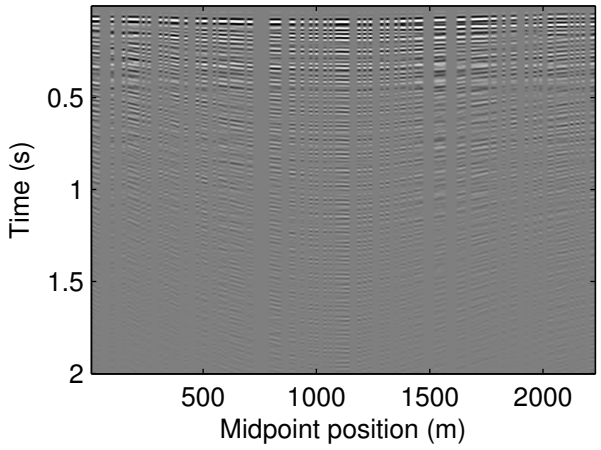


Figure 5: Example of the accuracy achieved from using adjacent offsets to predict the support of the curvelet coefficients in a seismic line.

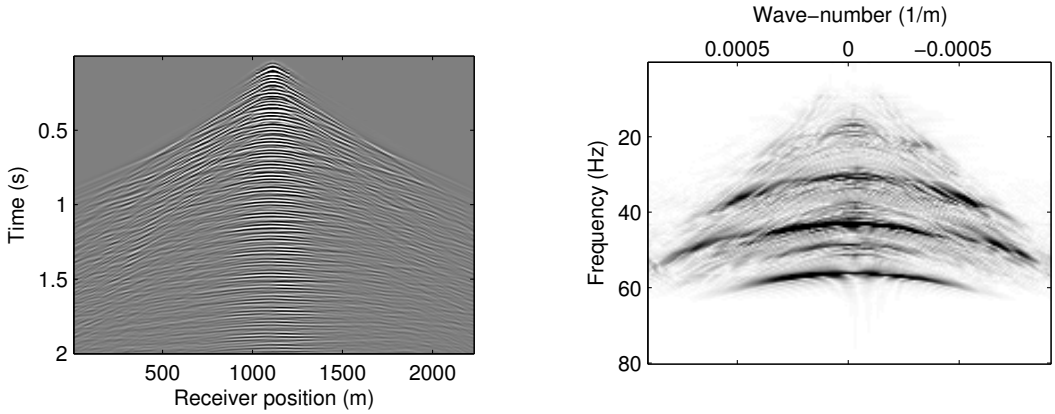


a



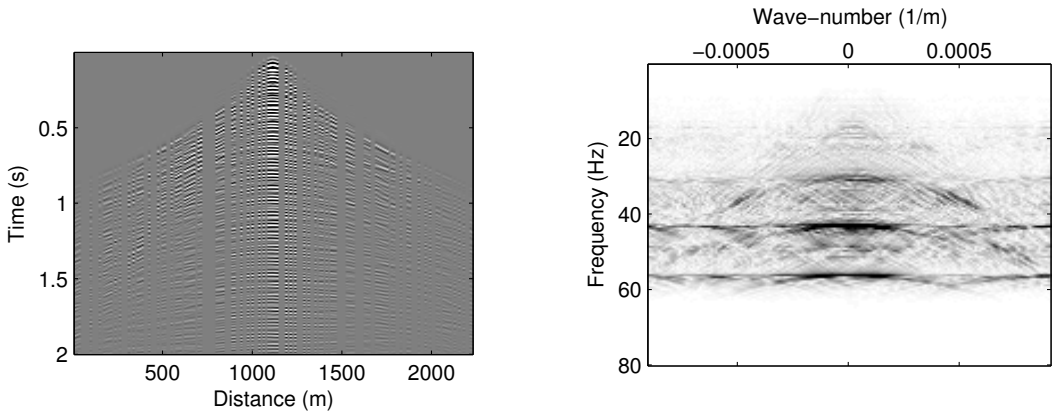
b

Figure 6: Example of a fully sampled (top) and subsampled (bottom) zero offset gather with missing columns arising from the diagonal entries of the source-receiver subsampling mask.



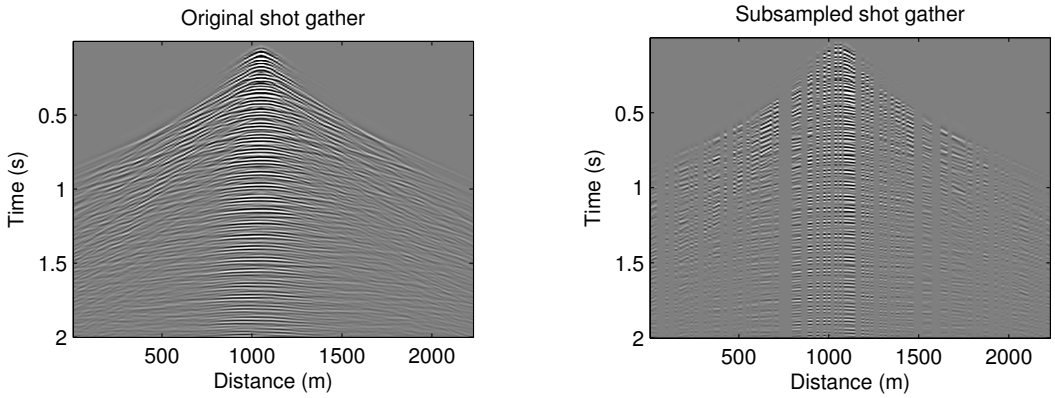
a b

Figure 7: (a) Original shot gather number 89 from a seismic line from the Gulf of Suez. (b) Corresponding F-K-spectrum of shot gather number 89.



a b

Figure 8: (a) Subsampled shot gather number 89 from a seismic line from the Gulf of Suez. (b) Corresponding F-K spectrum of the subsampled shot gather showing the full spectrum noise and aliasing resulting from the subsampling.



a b

Figure 9: (a) Shot gather number 84 from a seismic line from the Gulf of Suez. (b) Subsampled shot gather using column 84 from the mask in Figure 2b.

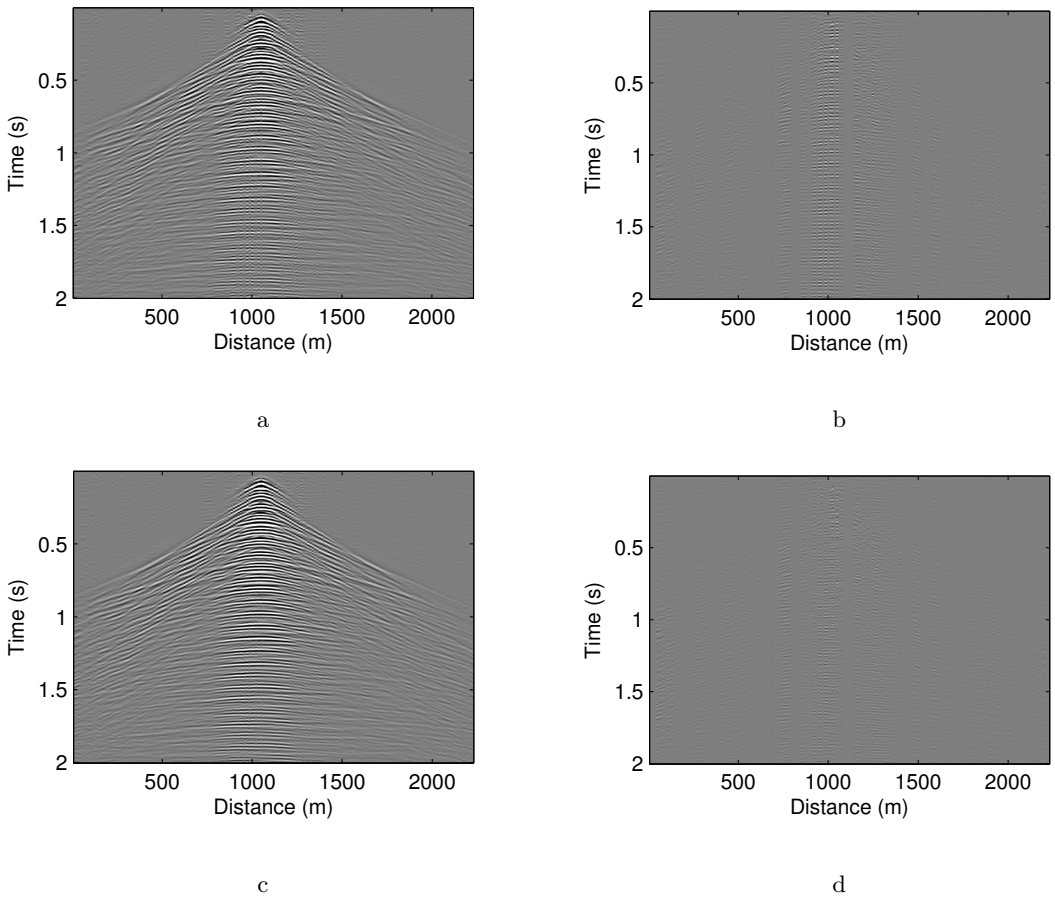
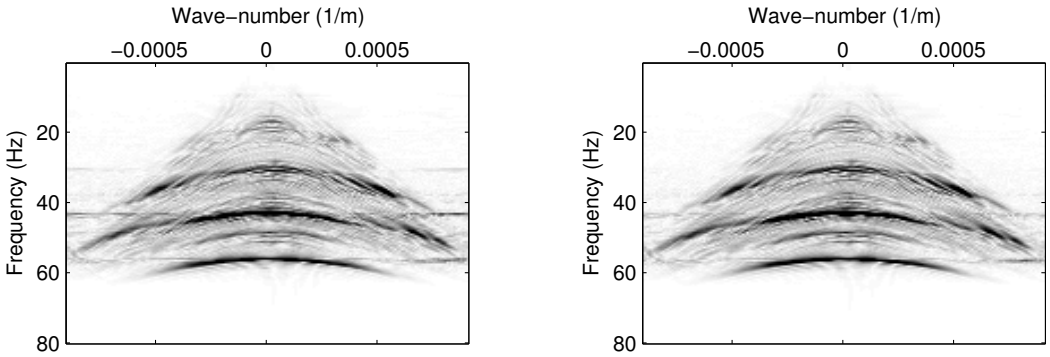


Figure 10: (a) Recovered shot gather using ℓ_1 minimization and (c) using weighted ℓ_1 minimization in the SR domain. Error plots showing the difference between the original shot gather and the reconstruction from (b) ℓ_1 minimization in the SR domain, and (d) weighted ℓ_1 minimization in the SR domain using the support derived from the previous frequency.



a b

Figure 11: F-K spectra of shot gather 89 after reconstruction using (a) ℓ_1 minimization in the source-receiver domain, and (b) weighted ℓ_1 minimization in the source-receiver domain utilizing curvelet support correlations across frequency slices.

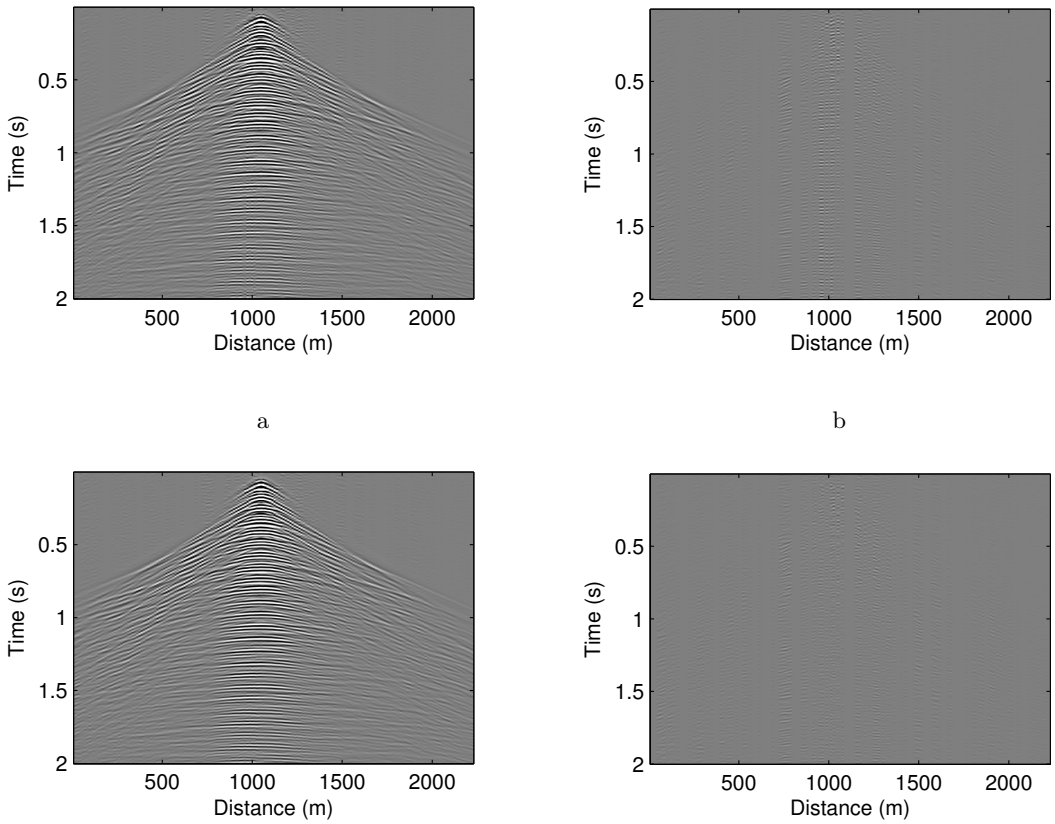
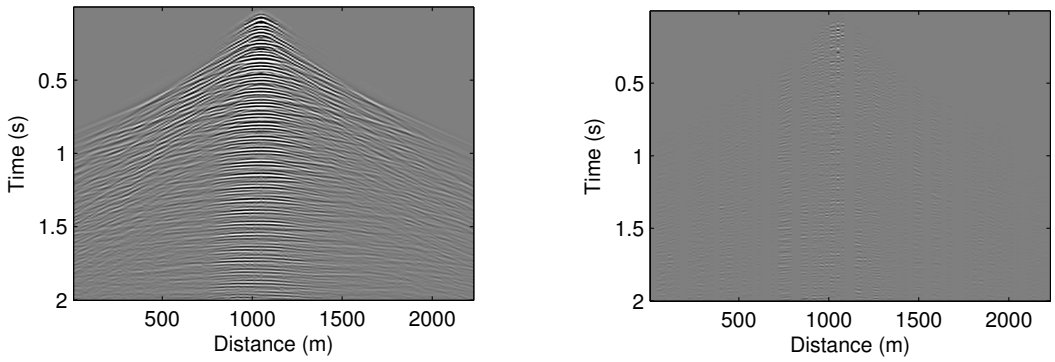
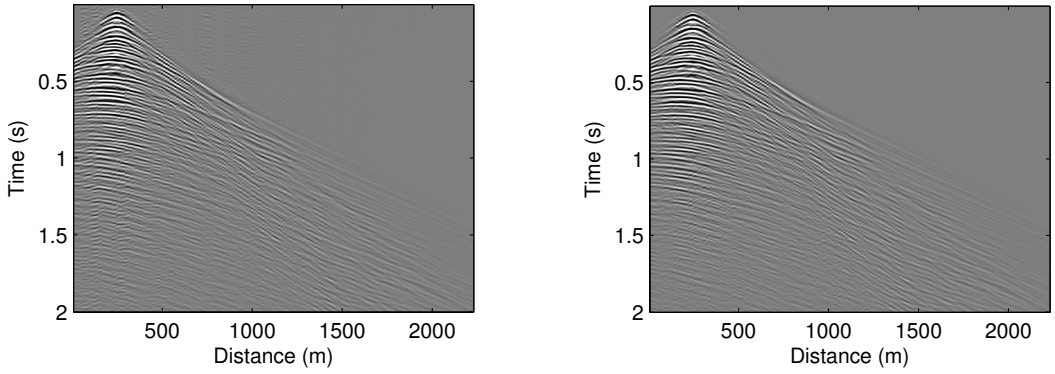


Figure 12: (a) Recovered shot gather using ℓ_1 minimization and (c) using weighted ℓ_1 minimization in the MH domain. Error plots showing the difference between the original shot gather and the reconstruction from (b) ℓ_1 minimization in the MH domain, and (d) weighted ℓ_1 minimization in the MH domain using the support derived from the previous frequency.



a b

Figure 13: (a) Recovered shot gather using weighted ℓ_1 minimization across offset-gathers in the time-midpoint domain, and (b) the corresponding error plot of the recovered shot gather.



a b

Figure 14: Recovered shot gather number 20 using weighted ℓ_1 minimization (a) across frequency slices in the midpoint-offset domain, and (b) across offset-gathers in the time-midpoint domain.

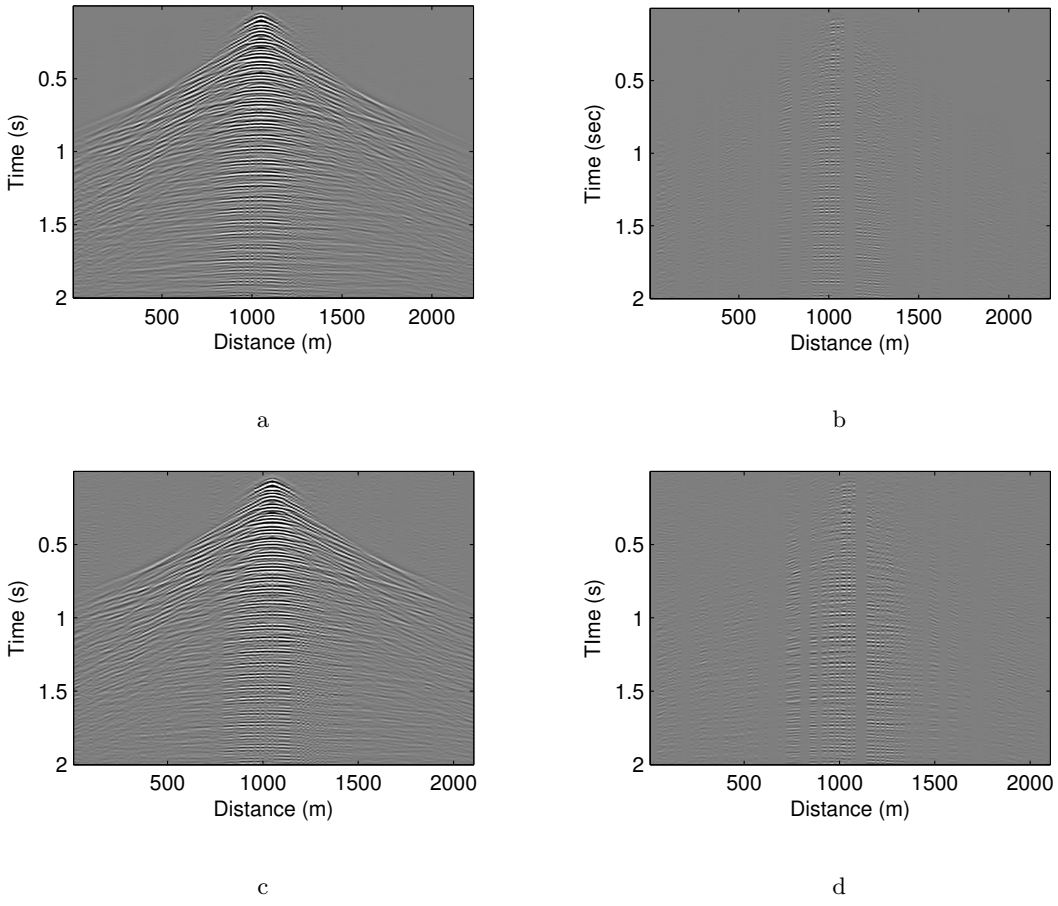


Figure 15: (a) Recovered shot gather using ℓ_1 minimization with a 3D transform composed of a 2D curvelet transform applied in the source-receiver domain and a wavelet transform applied in the time domain, and (c) using a 3D curvelet transform in the source-receiver-time domain. The right column figures (b) and (d) illustrate the corresponding error plots of the recovered shot gathers in (a) and (b).

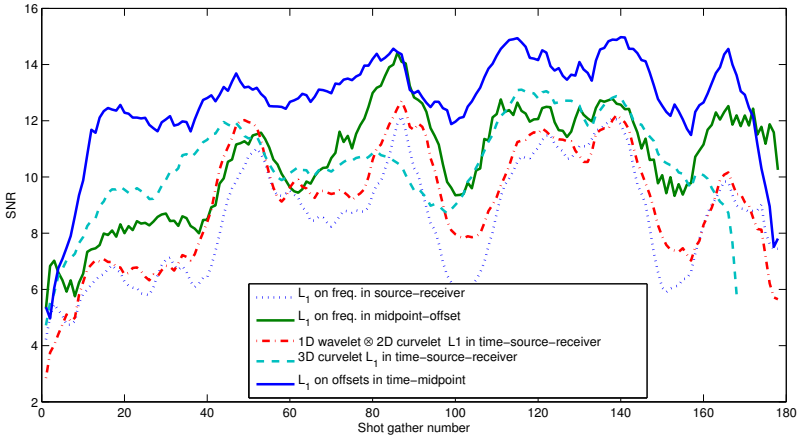


Figure 16: Comparison of the SNRs achieved by ℓ_1 minimization in recovering shot gathers from the Suez seismic line applied to both the source-receiver and midpoint-offset domains. The plots also include recovery results from 3D volume reconstruction using first a 2D curvelet transform in the source-receiver domain combined by a Kronecker product with a wavelet transform in the time domain, and second a 3D curvelet transform.

Rank	**Recovery algorithm**	**Partition axis**	**Transform domain**	**SNR**
1	weighted ℓ_1 min.	offsets	2D curvelet in time-midpoint	14.2 dB
2	weighted ℓ_1 min.	frequency	2D curvelet in midpoint-offset	13.9 dB
3	ℓ_1 min.	offsets	2D curvelet in time-midpoint	12.7 dB
4	weighted ℓ_1 min.	frequency	2D curvelet in source-receiver	11.8 dB
5	ℓ_1 min.	frequency	2D curvelet in midpoint-offset	10.58 dB
6	ℓ_1 min.	none	3D curvelet in time-source-receiver	10.45 dB

Table 1: Ranking the different partitioning and reconstruction schemes according to the reconstructed SNR.

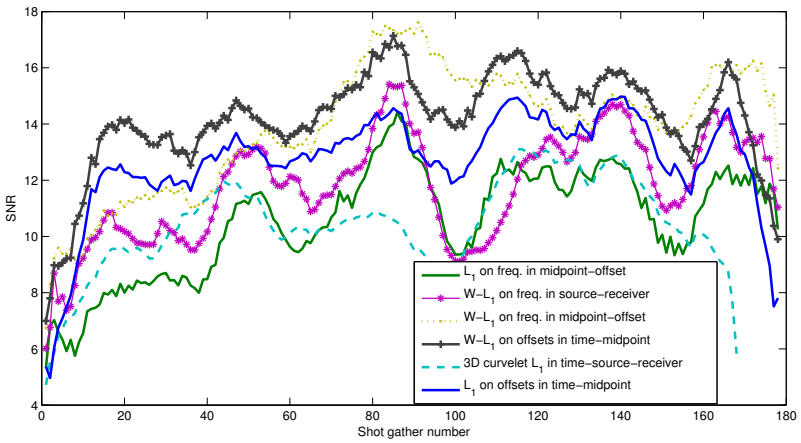


Figure 17: Comparison of the SNRs achieved by ℓ_1 and weighted ℓ_1 minimization in recovering shot gathers from the Suez seismic line applied to both the source-receiver and midpoint-offset domains.