Comparison of sparsity-constrained regularization methods for denoising and interpolation

Lucas Almeida[1], Michael Wakin[2] & Paul Sava[1]
[1]*Center for Wave Phenomena, Colorado School of Mines*
[2]*Department of Electrical Engineering and Computer Science, Colorado School of Mines*

ABSTRACT

Missing trace reconstruction is an ongoing challenge in seismic processing due to incomplete acquisition schemes and irregular grids. Noise is also a concern because it is naturally present in acquired seismic data through several mechanisms such as natural noise and equipment noise. Both problems need to be adequately addressed, especially because they negatively affect several important processing steps such as migration. While there are many approaches to solve these problems, most of the recent research on the subject focuses on transform domain approaches. These approaches commonly use sparsity-constrained inversion, i.e., one assumes that the signal to be estimated is sparse in a transform domain, in order to obtain a reasonable solution. Several formulations for denoising and missing trace reconstruction have been proposed based on the compressive sensing (CS) framework, which states that sparse signals can be recovered from a highly incomplete set of measurements. In particular, a specific kind of constraint, called synthesis approach, has been widely used in geophysical problems. The analysis approach, which can be considered as the synthesis' dual problem, is an alternative for sparsity-constrained inversion. Although less popular than the synthesis problem, the analysis approach is more effective in several problems, such as denoising of natural images. In this paper, we compare the analysis and synthesis approaches as sparsity constraints for the denoising of seismic images and missing trace reconstruction. Our experiments show that the analysis approach can yield more accurate results than the synthesis approach for both problems, which makes it a viable approach for sparsity-constrained inversion for geophysical problems and should be considered along with the synthesis approach.

Key words: inverse problem, sparsity, synthesis, analysis, curvelet

1 INTRODUCTION

Ideally, seismic acquisition seeks to sample densely and regularly in every spatial direction, with the intent of obtaining signals that adequately represent data observed at the surface. Adequate sampling has important consequences for many applications, such as reverse-time migration and multiple removal. However, acquisition costs and field obstacles can easily make the acquisition both irregular and sparse. The purpose of missing trace reconstruction is to fill the data gaps or to resample the data as accurately as possible.

The missing trace reconstruction problem has been approached in several ways. One technique consists of designing a filter that predicts the dip of the data and use this information to reconstruct the missing traces (Crawley et al., 1999). Porsani (1999); Spitz (1991) propose interpolating seismic events in Fourier domain windows by solving an overdetermined least-squares problem. Although their theory assumes seismic data without lateral amplitude variation, the method is robust enough to perform reasonably inside this restriction. A different technique is to explicitly incorporate a wave equation in the interpolation procedure in order to enhance its accuracy, as proposed by Stolt (2002) and others. More recently, the approach of using transform domains to estimate the missing traces has been proposed by several authors, such as Hennenfent and Herrmann (2008) and Naghizadeh and Sacchi (2010). This approach assumes that the analyzed wavefield is sparse in some domain and involves the minimization of a convex function with an L_1 norm, which is known to promote sparsity. Examples of transforms that are used in this setting are the Fourier, curvelet (Candes et al., 2006), seislet (Fomel and Liu, 2010) and data-driven transforms (Yu et al., 2015).

Transform domain approaches are also common when tackling the denoise problem. For coherent noise, a multiscale transform (such as wavelets or curvelets) is used in order to separate signal from noise at various scales, thus enabling efficient removal of the noise (Oliveira et al., 2012; Almeida et al., 2015). For incoherent noise, a transform domain where the signal is sparse is used under the assumption that the noise is not sparse in that domain (Hennenfent and Herrmann, 2006). This procedure, along with soft or hard thresholding, can elimitate incoherent noise efficiently. L_1 minimization is also used to denoise complex data corrupted by incoherent noise, especially when data-driven transforms are applied (Herrmann et al., 2007; Zhu et al., 2015).

The transform domain approach is usually related to the compressive sensing problem (Candès et al., 2006), which provides theoretical guarantees of exact recovery for underdetermined inverse problems where the solution to be estimated is sparse. In the common compressive sensing formulation, one tries to find the sparsest representation that can describe the signal to be estimated in a transform domain. This formulation is known as the **synthesis approach** to the L_1 minimization problem and it has been widely used for several geophysical problems, such as full-waveform inversion (Li et al., 2012), deblending (Wason et al., 2011), salt body detection (Ramirez et al., 2016), microseismic (Rodriguez and Sacchi, 2014) and missing trace reconstruction (Hennenfent and Herrmann, 2008).

The **analysis approach** provides a different way of solving the L_1 minimization problem by estimating the signal in the ambient domain using a set of forward transforms. Although less popular than the synthesis approach, the analysis approach and its geometry has been studied theoretically from the perspective of compressive sensing (Candes et al., 2011) and the cosparse model (Nam et al., 2013). Empirically, the analysis approach has been found to be superior than the synthesis approach when used to recover specific types of signals (Elad et al., 2007). Improvements over the synthesis approach were observed especially for the denoising problem on several types of images (Selesnick and Figueiredo, 2009; Elad et al., 2007). However, the differences between the analysis and synthesis approaches are highly dependent on the application, dataset and transform domain, and a definitive conclusion regarding which approach is better remains elusive. In geophysics, the analysis approach has been used sparingly to solve problems such as denoising with data-driven frames (Chen et al., 2016), multiple removal and interpolation (Yang and Fomel, 2015).

In this paper, we propose to compare the analysis and synthesis approaches when applied to two geophysical problems, namely missing trace reconstruction and random noise attenuation. We provide an overview of the underdetermined inverse problem and the synthesis approach. We detail the implementation of the analysis approach and explain the main differences between both approaches. Following Hennenfent and Herrmann (2008), we provide an overview of how different undersampling techniques affect the estimation of the signal in the missing trace reconstruction problem. We illustrate both approaches by denoising and interpolating missing traces in a complex shot gather from the Sigsbee 2A model under different spatial sampling schemes.

2 THE UNDERDETERMINED INVERSE PROBLEM

The underconstrained inverse problem can be formulated mathematically as

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x}, \tag{1}$$

where $\mathbf{\Phi}$ is a linear operator that maps a high dimensional vector $\mathbf{x} \in \Re^n$ to a low dimensional vector $\mathbf{y} \in \Re^m$. In a general setting, this problem is related to the embedding of a signal, which belongs to a high dimensional subspace, into a lower dimensional subspace. Several problems currently pursued by the signal processing community, such as denoising and edge detection (Selesnick and Figueiredo, 2009), can be formulated in a way that resembles equation 1.

2.1 The synthesis approach

In general, the problem of obtaining a high dimensional signal from its low dimensional representation has infinite solutions; solving the problem described in equation 1 using a least-squares approach yields only an approximate solution, which may differ significantly from the original signal. An interesting way to constrain the problem is to assume that the signal to be recovered is sparse, which means that even though the signal is high dimensional, it contains a small number of non-zero elements. Formally, sparsity is defined as

$$\{\mathbf{x} : \|\mathbf{x}\|_0 = k\}. \tag{2}$$

In other words, the sparsity is defined by the number of non-zero elements in a given signal: a signal that has k non-zero elements is called k-sparse. In their seminal paper, Candès et al. (2006) observe that under specific conditions such as the type of the matrix $\mathbf{\Phi}$, sparsity of the signal to be estimated and number of measurements available, the underdetermined problem has a unique solution. This solution can be found by solving the following convex minimization problem:

$$\min_{\mathbf{x}} \quad \|\mathbf{x}\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2. \tag{3}$$

The problem described by equation 3 has become increasingly popular in the signal processing field and is commonly known as the **compressive sensing problem**. While signals are generally not sparse in the ambient domain, most signals have a concise representation in a transform domain. Denoting the transform matrix as $\mathbf{\Theta}$, the equation 3 can be modified to

$$\min_{\mathbf{u}} \quad \|\mathbf{u}\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{\Phi}\mathbf{\Theta}^*\mathbf{u}\|_2, \tag{4}$$

where $\mathbf{u}$ is the representation of the original signal $\mathbf{x}$ in the

transform domain, which is supposed to be sparse. Knowing $\mathbf{u}$, one can obtain the original signal $\mathbf{x}$ using

$$\mathbf{x} = \mathbf{\Theta}^* \mathbf{u}. \tag{5}$$

Because this formulation involves synthesizing the signal from its representation in the transformed domain, it is commonly referred to as the synthesis approach.

The following example illustrates the capability of this formulation when the signal to be estimated is sparse in a transform domain. Consider the sinusoid with length $n = 512$, shown in Figure 1(a), and suppose $\mathbf{\Phi}_{m\times n}$ is a matrix whose values are taken from a Gaussian distribution with zero mean and variance $1/m$. We set $m = 20$, which means that the aforementioned sinusoid is embedded in a low dimensional signal $\mathbf{y} \in \Re^{20}$ through the action of $\mathbf{\Phi}$. In general, it is impossible to recover the original signal $\mathbf{x}$ from $\mathbf{y}$, as shown in Figure 1(b).

However, a sinusoid is sparse in the Fourier domain. Note that the real part of the frequency spectrum of said sinusoid, shown in Figure 1(c), contains only two non-zero elements. Setting $\mathbf{\Theta}_{n\times n}$ equal to the discrete Fourier transform matrix and solving the inverse problem in equation 4, we obtain a perfect reconstruction from $\mathbf{y}$, as observed in Figure 1(d).

2.2 The analysis approach

Although less popular than the formulation in equation 4, the following formulation also solves the underdetermined problem:

$$\min_{\mathbf{x}} \quad \|\mathbf{\Theta x}\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{\Phi x}\|_2. \tag{6}$$

Instead of estimating a sparse representation of the signal in the transformed domain and subsequently synthesizing the signal using the inverse transform as in the problem defined in equation 5, equation 6 estimates a signal in the ambient domain with the constraint of a sparse forward transform. For this reason, it is commonly referred to as the analysis approach (Elad et al., 2007).

2.3 Sparsity in seismic wavefields

As previously observed, sparsity is defined by the number of non-zero elements in a given signal. With a few exceptions, no signal is strictly sparse in the ambient domain. However, several transforms have the ablility to act on these signals in such a way that their representation in the transform domain is more concise.

As expected, a non-adaptive transform cannot make every signal sparse. Indeed, certain classes of signals are sparse in different bases. For instance, sufficiently smooth signals are sparse in the Fourier domain (Mallat, 2008), while piecewise smooth signals (such as natural images) are better represented in a wavelet basis (Daubechies et al., 1992). When performing sparsity-constrained inversion, it is necessary to carefully choose the transform domain in order to obtain a representation that is as sparse as possible in order to improve the quality of the signal reconstruction (Donoho, 2006).

Seismic wavefields are usually composed of curves, i.e., seismic events. Candès and Donoho (2004) show that curvelets optimally represent piecewise smooth images with C^2 curves as discontinuities, which can be taken as the seismic events. Hence, curvelets are an attractive non-adaptive transform to represent seismic events and also can be efficiently applied to large problems (Candes et al., 2006). The curvelet transform is a tight frame (see Appendix A) that can be roughly seen as a dyadic-parabolic division of the Fourier spectrum of an image, which essentially means that curvelets are anisotropic and efficiently capture the directionality of events in the ambient domain.

Adaptive transforms have also been proposed to represent seismic data (e.g. Zhu et al. (2015); Yu et al. (2015)), and often obtaining good results. The power of an adaptive transform comes from the fact that it is trained using the available data, so that it produces a specific transform that provides a sparse representation of that specific data. More often than not, this representation is sparser than that of a non-adaptive transform. However, a non-adaptive transform lacks an efficient way of applying it to the data. This makes its application and training expensive when compared to non-adaptive transforms, specially on large datasets.

We restrict our discussion in this paper to non-adaptive transforms for both cost efficiency and generality. We adopt the curvelet frame as the transform domain $\mathbf{\Theta}$ in our examples and results. We use the convention $\mathbf{C}$ and $\mathbf{C}^*$ to represent the curvelet frame and the curvelet dual frame, respectively.

2.4 Analysis vs. synthesis

The analysis and synthesis approaches are two different ways of solving a sparsity-constrained inverse problem and, although both approaches seem to be equivalent, subtle differences arise between the formulations. Consider the following unconstrained minimization problem (Cai et al., 2012):

$$\min_{\mathbf{u}} \quad \|\mathbf{\Phi\Theta}^*\mathbf{u} - \mathbf{y}\|_2 + \kappa\|(\mathbf{I} - \mathbf{\Theta\Theta}^*)\mathbf{u}\|_2 + \gamma\|\mathbf{u}\|_1. \tag{7}$$

where κ and γ are regularization parameters that control the importance of the second and third terms in the minimization problem, respectively. The first term of equation 7 represents the data misfit. We define

$$\mathbf{u}_* = \mathbf{\Theta x}, \tag{8}$$

as the canonical coefficients of a given $\mathbf{x}$ in the transform domain, which means that among all possible $\mathbf{u}$ than can represent $\mathbf{x}$ in the transform domain, the canonical coefficients are the ones obtained by the forward transform of $\mathbf{x}$. Thus, the second term relates the distance between the estimated solution $\mathbf{u}$ and $\mathbf{u}_*$. The third term controls the sparsity of $\mathbf{u}$ using the L_1 norm.

The general formulation in equation 7 changes with re-

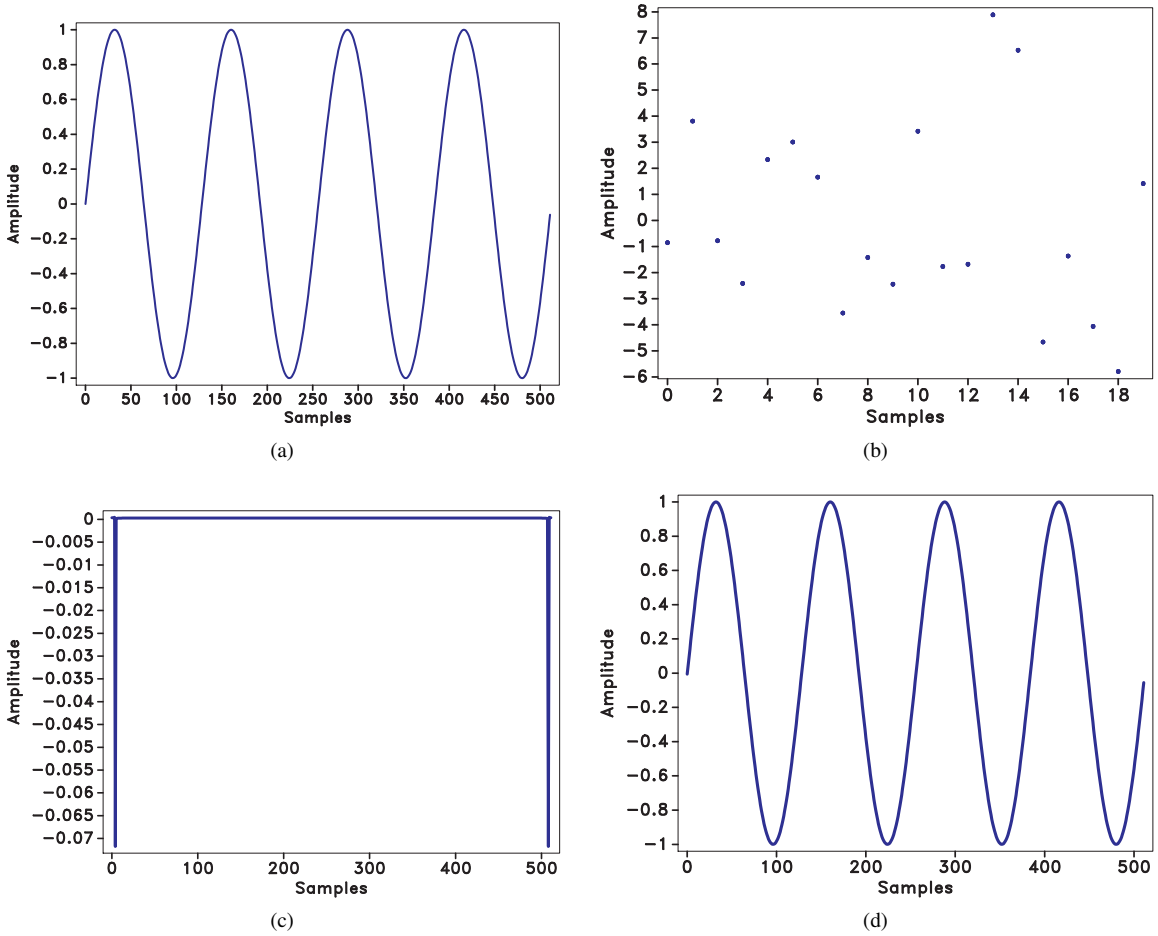


Figure 1. Compressive sensing toy problem: (a) - Input signal, which is denoted as $\mathbf{x}$. (b) - The measured data $\mathbf{y} = \mathbf{\Phi x}$. (c) - Real part of the frequency spectrum of the sine in (a). (d) - Reconstruction result.

spect to how the second term affects the minimization problem. In fact, two special cases are of particular relevance. When $\kappa = 0$, then equation 7 becomes the synthesis approach analogous to equation 4, which finds the sparsest set of coefficients $\mathbf{u}$ that agrees with a certain data misfit. When $\kappa \longrightarrow \infty$, then

$$\min_{\mathbf{u} \in Range(\mathbf{\Theta})} \quad \|\mathbf{\Phi \Theta}^* \mathbf{u} - \mathbf{y}\|_2 + \gamma \|\mathbf{u}\|_1. \tag{9}$$

Note that, since $\mathbf{u} \in Range(\mathbf{\Theta})$, then the estimated vector can be described as $\mathbf{\Theta x}$. Therefore, equation 9 can be written as

$$\min_{\mathbf{x}} \quad \|\mathbf{\Phi x} - \mathbf{y}\|_2 + \gamma \|\mathbf{\Theta x}\|_1, \tag{10}$$

which is an unconstrained version of the analysis approach in equation 6.

When $\mathbf{\Theta}$ is an orthonormal basis, the equation 7 reduces to a unique formulation independent of the value of κ because the second term is zero. For example, the sine problem described in Figure 1 leads to the same solution regardless of which approach is used, because the Fourier transform is an orthonormal basis. However, if $\mathbf{\Theta}$ is a frame such as the curvelet transform $\mathbf{C}$, the synthesis and analysis approaches lead to different solutions because $\mathbf{CC}^* \neq \mathbf{I}$. One of the main differences between the approaches comes from the fact that the solutions are estimated in different domains: while the synthesis approach seeks the sparsest sequence $\mathbf{u}$ such that it complies with equation 5, the analysis approach seeks to find a dataset $\mathbf{x}$ which has a forward transform similar to $\mathbf{u}_*$. In particular, the solution of the synthesis approach is closely related to the soft thresholding operator (Daubechies et al., 2004)

$$T_\gamma(u[k]) = \begin{cases} u[k] - \gamma/2 & \text{if } u[k] \geq \gamma/2 \\ 0 & \text{if } |u[k]| < \gamma/2 \\ u[k] + \gamma/2 & \text{if } u[k] \leq -\gamma/2, \end{cases} \tag{11}$$

where $u[k]$ is an element of the vector $\mathbf{u}$. The above operator is, in fact, present in almost every algorithm that solves

equation 4. Moreover, if a frame is used, there are several sequences $\mathbf{u}$ that represent a given dataset $\mathbf{x}$, and only the sparsest of them is a solution of equation 4 (see Appendix A). On the other hand, the analysis approach is dominated by the second term of equation 7, which means that its solution is closely related to the orthogonal projection operator

$$\mathbf{\Theta}_{proj} = \mathbf{I} - \mathbf{\Theta}\mathbf{\Theta}^*. \tag{12}$$

Because $\mathbf{\Theta}\mathbf{\Theta}^*$ spans the column space of $\mathbf{\Theta}$, $\mathbf{\Theta}_{proj}$ projects a vector to the subspace spanned by the orthogonal complement of said column space, which is the null space of $\mathbf{\Theta}^*$. Setting $\kappa \longrightarrow \infty$ in equation 7 implies $\mathbf{\Theta}_{proj}\mathbf{u} = \mathbf{0}$, which means that $\mathbf{u} = \mathbf{u}_*$ is the only non-trivial vector in the null space of $\mathbf{\Theta}_{proj}$ for a given $\mathbf{x}$. This acts as another constraint in addition to the sparsity constraint given by the third term of equation 7, so that the analysis approach looks for the sparsest $\mathbf{u}$ that belongs to the null space of $\mathbf{\Theta}_{proj}$.

If the frame used in both approaches is formed by rows vectors in general position, then the solution $\mathbf{u}$ of the synthesis approach is much sparser than $\mathbf{u}^*$. This is an advantage for the synthesis approach in terms of descriptive power, i.e., it needs a smaller number of coefficients to represent a given data when compared to the analysis approach. However, it can also be detrimental in cases where the estimation of the coefficients is not accurate, i.e., in the presence of noise. This is because a smaller number of coefficients makes every coefficient carry much more significance. If, for instance, the support or magnitude of (possibly some of) the coefficients of the synthesis solution are erroneously estimated, then the final signal $\mathbf{x} = \mathbf{\Theta}^*\mathbf{u}$ might be very different from the desired signal. Since the solution of the analysis approach is not as sparse in the transform domain, erroneous estimation does not affect the estimated signal as much. This motivates us to study the performance of both approaches on noise related geophysical problems.

In general, the exact distinction between the synthesis and analysis approaches when it comes to specific applications and data is difficult to pinpoint. Elad et al. (2007) observe that there is indeed a gap between the geometries of both problems and that each approach has its own favorable set of signals for which the estimation is optimal. Furthermore, it is highly likely that the differences between the approaches also depend on aspects of the chosen frame, such as redundancy (Becker et al., 2011). That said, it is seldom possible to assure which approach is better without prior knowledge on the application, dataset and transform domain.

3 DENOISING OF SEISMIC DATA AS AN INVERSE PROBLEM

The denoising problem can be formulated using equation 1. In this case, the signal $\mathbf{x}$ represents the original seismic data one wants to estimate and the signal $\mathbf{y}$ represents the seismic data corrupted with random noise. Because the measured data $\mathbf{y}$ are complete, the problem can be formulated using $\mathbf{\Phi} = \mathbf{I}$. Using these modifications, then equation 4 and equation 6 can be modified to

$$\|\mathbf{u}\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{I}\mathbf{C}^*\mathbf{u}\|_2, \tag{13}$$

and

$$\|\mathbf{C}\mathbf{x}\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{I}\mathbf{x}\|_2. \tag{14}$$

Because the curvelet transform is a Fourier-related transform, this problem amounts to denoising the f-k spectrum of the noisy image, which is contaminated by random noise.

4 THE MISSING TRACE RECONSTRUCTION AS AN INVERSE PROBLEM

The missing trace reconstruction problem can also be formulated using equation 1. In this case, the signal $\mathbf{x}$ represents the original seismic data one wants to estimate and the signal $\mathbf{y}$ represents the seismic data with missing traces. The missing trace representation can be obtained using equation 1 if $\mathbf{\Phi}$ is, for example, a restriction matrix with 1's on the positions where samples are taken in the missing trace representation and 0's elsewhere. Denoting the restriction matrix as $\mathbf{R}$, the equation 4 and equation 6 can be modified to

$$\|\mathbf{u}\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{R}\mathbf{C}^*\mathbf{u}\|_2, \tag{15}$$

and

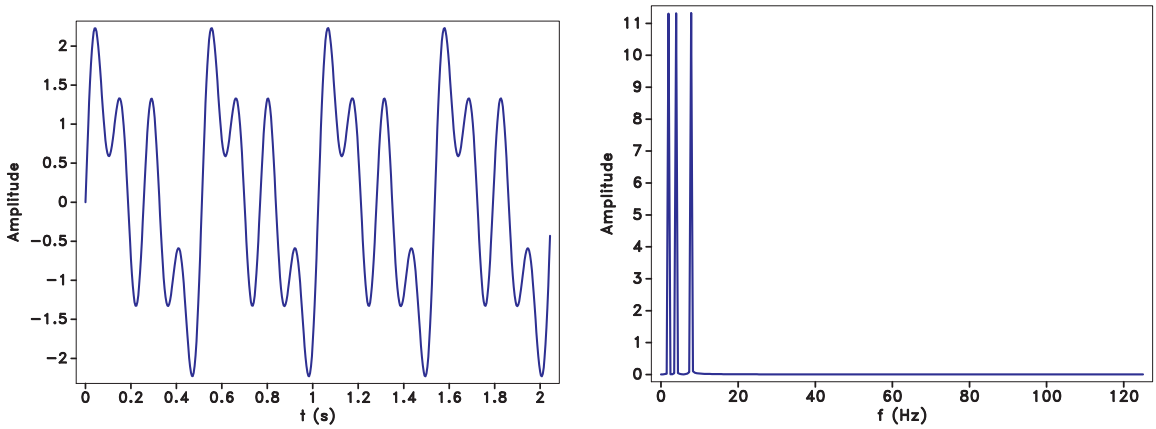
$$\|\mathbf{C}\mathbf{x}\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{R}\mathbf{x}\|_2. \tag{16}$$

The restriction matrix $\mathbf{R}$ can be defined in a variety of ways, which correspond to the way the data are undersampled during the acquisition. Hennenfent and Herrmann (2008) and others observe that the undersampling scheme plays an important role in determining the overall success of the aforementioned minimization problems using Fourier-related transforms, because it affects how the spectral leakage (i.e., aliases) is present in the Fourier spectrum of the missing trace data.

4.1 The restriction matrix

One of the components of the minimization problems described in equations 15 and 16 is the restriction matrix $\mathbf{R}$. This matrix is constructed based on which traces are sampled during the acquisition. Because the acquired data are incomplete, their Fourier spectrum varies depending on how the data are undersampled. This variation is important for our purposes because we assume that the seismic wavefield is sparse in the curvelet domain, which is a Fourier-related transform. In this section, we illustrate some undersampling schemes and their effects on the frequency spectrum of the simple signal in Figure 2(a).

A common undersampling scheme selects traces regularly as a function of space. This pattern is known as the uni-



(a) (b)

Figure 2. (a) Signal containing three different frequencies. (b) Corresponding frequency spectrum.

form undersampling scheme, illustrated in Figure 3(a), for the case where one out of every two samples removed. Figure 3(b) shows the frequency spectrum of the uniformly undersampled data. Note that the usual aliasing effects are present due to insufficient sampling in the spatial direction.

This type of undersampling renders minimization problems based on sparsity priors more challenging because the aliases and the original spectrum of the data have similar structure. Therefore, the aliases are sparse and likely to be a part of the solution of the minimization problem. Authors such as Hennenfent and Herrmann (2008) and Xu et al. (2005) encounter the same issue as they try to obtain a soluton for the missing trace reconstruction problem using uniform undersampling.

A more general undersampling scheme is obtained when the traces are taken out following a uniform probability distribution. This is known as the random undersampling scheme and is illustrated in Figure 3(c). The frequency spectrum of the randomly undersampled data is shown in Figure 3(d). Note that instead of the aliasing that usually appears in the f-k spectrum of uniformly undersampled data, the spectral leakage turns into somewhat incoherent artifacts (Donoho, 2006).

The randomly undersampling scheme is quite favorable for minimization problems that involve sparsity priors because the incoherent artifacts produced by the scheme are commonly not sparse in any reasonable transform domain. Consequently, the inversion should be able to "denoise" the f-k spectrum effectively, thus providing a good solution. This undersampling schemes also provides a more general way to undersample the data because it allows for varying gap sizes between traces in the decimated data.

A third type of undersampling scheme, called jittered undersampling, is analyzed in Hennenfent and Herrmann (2008). Although similar to the random undersampling scheme, the jittered scheme does not remove traces according to an uniform probability distribution. Instead, it perturbs a uniformly undersampling scheme around a given gap size. Figure 3(e) shows an example of the jittered undersampling scheme. Notice that, in the jittered scheme, one can control the gap sizes and prevent them from becoming large.

The frequency spectrum of the jittered undersampled data is shown in Figure 3(f) and it is very similar to the spectrum in Figure 3(d), which indicates that the jittered and randomly undersampled schemes are similarly favorable to inversion using sparsity priors. However, as seen in Hennenfent and Herrmann (2008), the usage of the randomly undersampling scheme with Fourier-related localized transforms (such as curvelets) might decrease the quality of the estimated data when compared to the jittered undersampling. This is because arbitrarily big gaps can be larger than the transform element (i.e., larger than a curvelet), which cannot promote accurate reconstruction inside these gaps.

5 RESULTS

In this section, we evaluate the reconstruction results of the missing trace reconstruction and denoising problems using equations 15 and 16. We propose to evaluate the quality of the results through the signal-to-noise equation

$$SNR = 20log(\frac{\|\mathbf{x}\|_2}{\|\mathbf{x} - \tilde{\mathbf{x}}\|_2}), \tag{17}$$

where $\mathbf{x}$ is the original data and $\tilde{\mathbf{x}}$ is the estimated data using either equation 15 or equation 16. We also provide difference plots, which feature the elementwise difference $\mathbf{x} - \tilde{\mathbf{x}}$, as a secondary evaluation tool.

For both problems, we use a shot gather from the Sigsbee 2A model, shown in Figure **??**. Figure 2(b) shows the f-k spectrum of the shot gather. This model is appropriate for testing both approaches because it features complex geology, which generates shots with complex events with a wide range of amplitudes and conflicting dips due to the salt body. Note that the spectrum is not symmetric and features varying amplitudes.

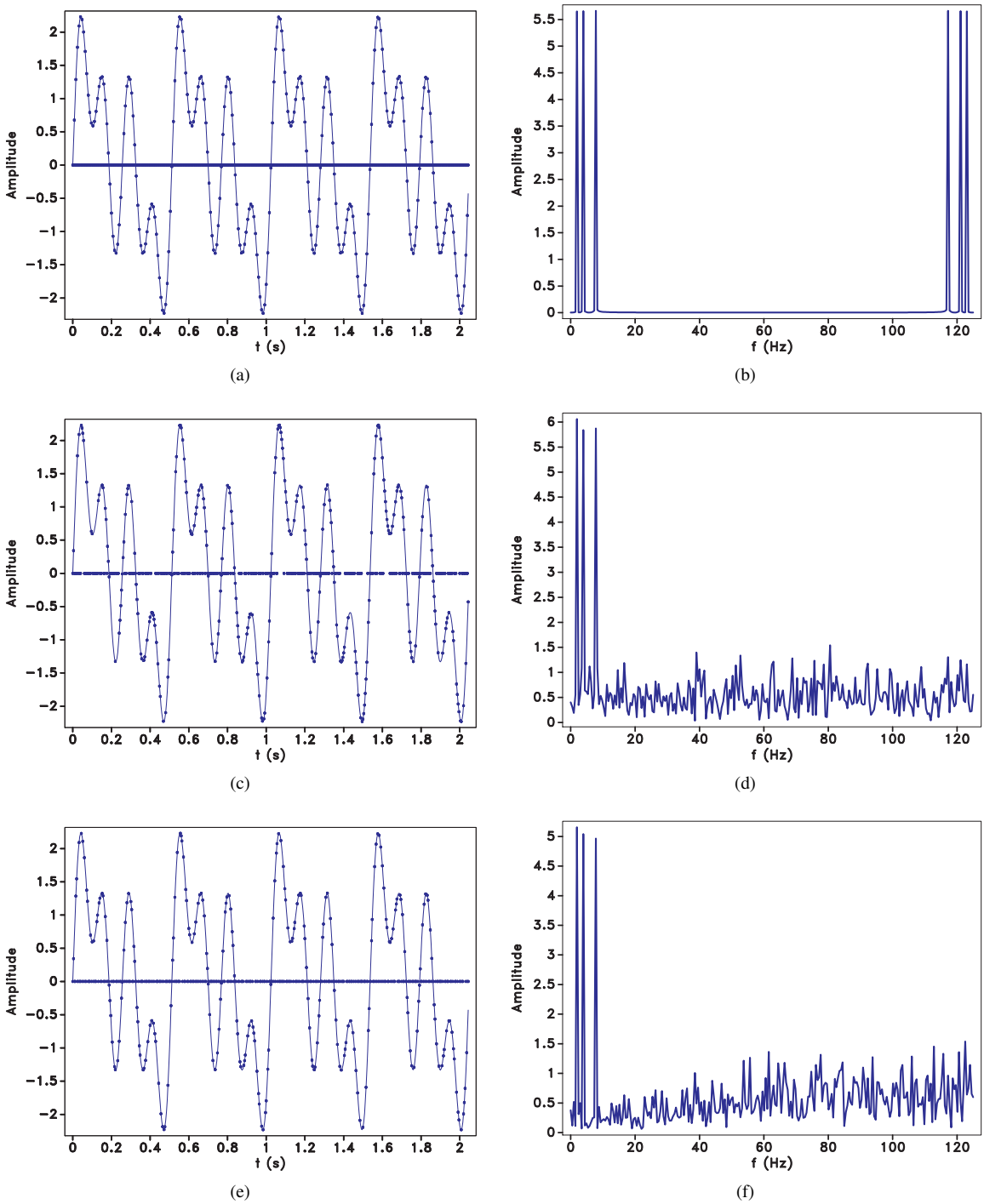
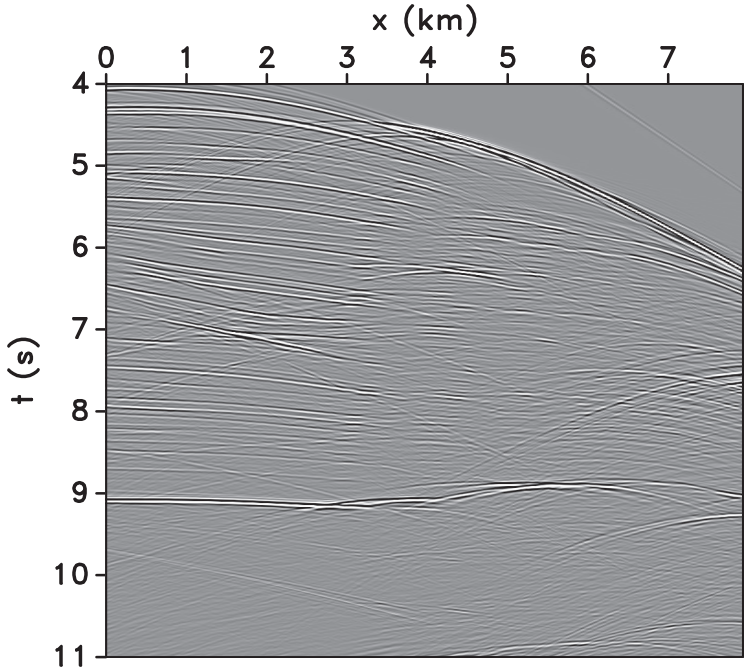
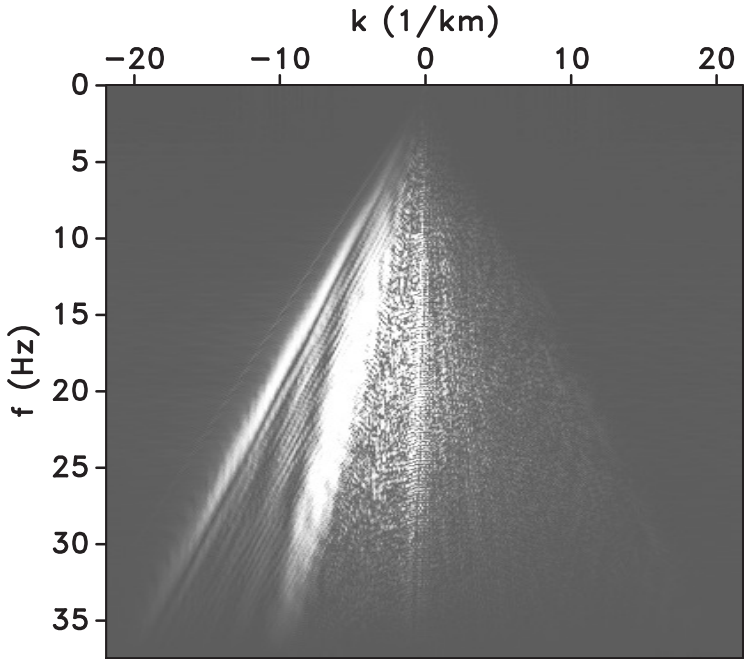


Figure 3. Figures displaying different undersampling schemes. The solid line indicates the original signal and the dots indicate the selected samples. Signals in time and corresponding spectra for (a), (b) - Uniform sampling, (c), (d) Random sampling, (e), (f) Jittered sampling.



(a)



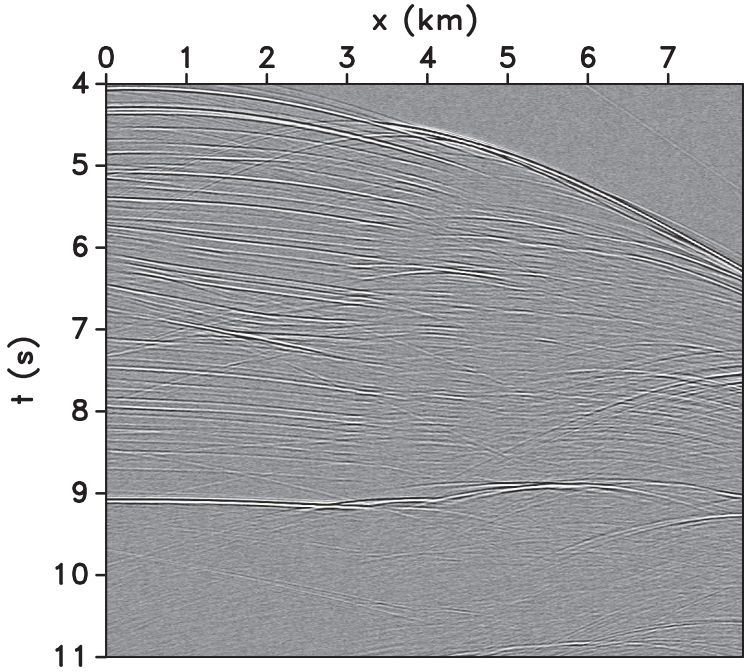
(b)

Figure 4. (a) Shot gather from the Sigsbee 2A model. (b) Corresponding f-k spectrum.

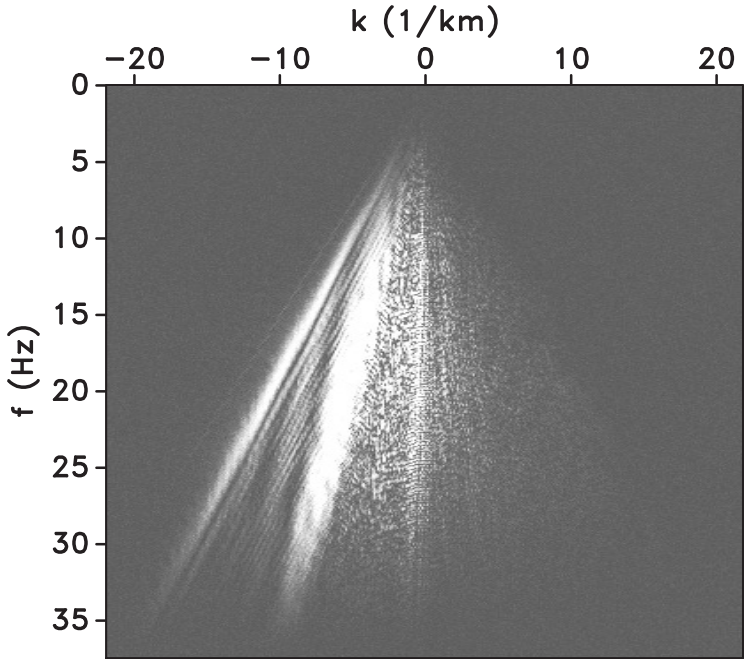
5.1 Denoising

We evaluate the results of the denoising problem on the shot gather in Figure **??**. We contaminate the shot gather with Gaussian noise, which is shown in Figure **??**. We show the f-k spectrum of the noisy shot gather in Figure 5(b). Note that the noise has a completely random pattern, as expected.

We apply the formulations present in equations 13 and 14 to denoise the shot gather. Figure 6 shows the estimated solutions by the synthesis and analysis approaches, as well as the



(a)



(b)

Figure 5. (a) Noisy shot gather. (b) Corresponding f-k spectrum.

corresponding f-k spectra and difference plots. One can observe in the difference plots that both approaches do well in removing noise, but also damage seismic events that should stay untouched. However, the difference plots and the SNR indicate that the analysis approach does a better job at preserving the signal than the synthesis approach. This indicates that the analysis approach is more effective at denoising incoherent noise than the synthesis approach in this experiment.

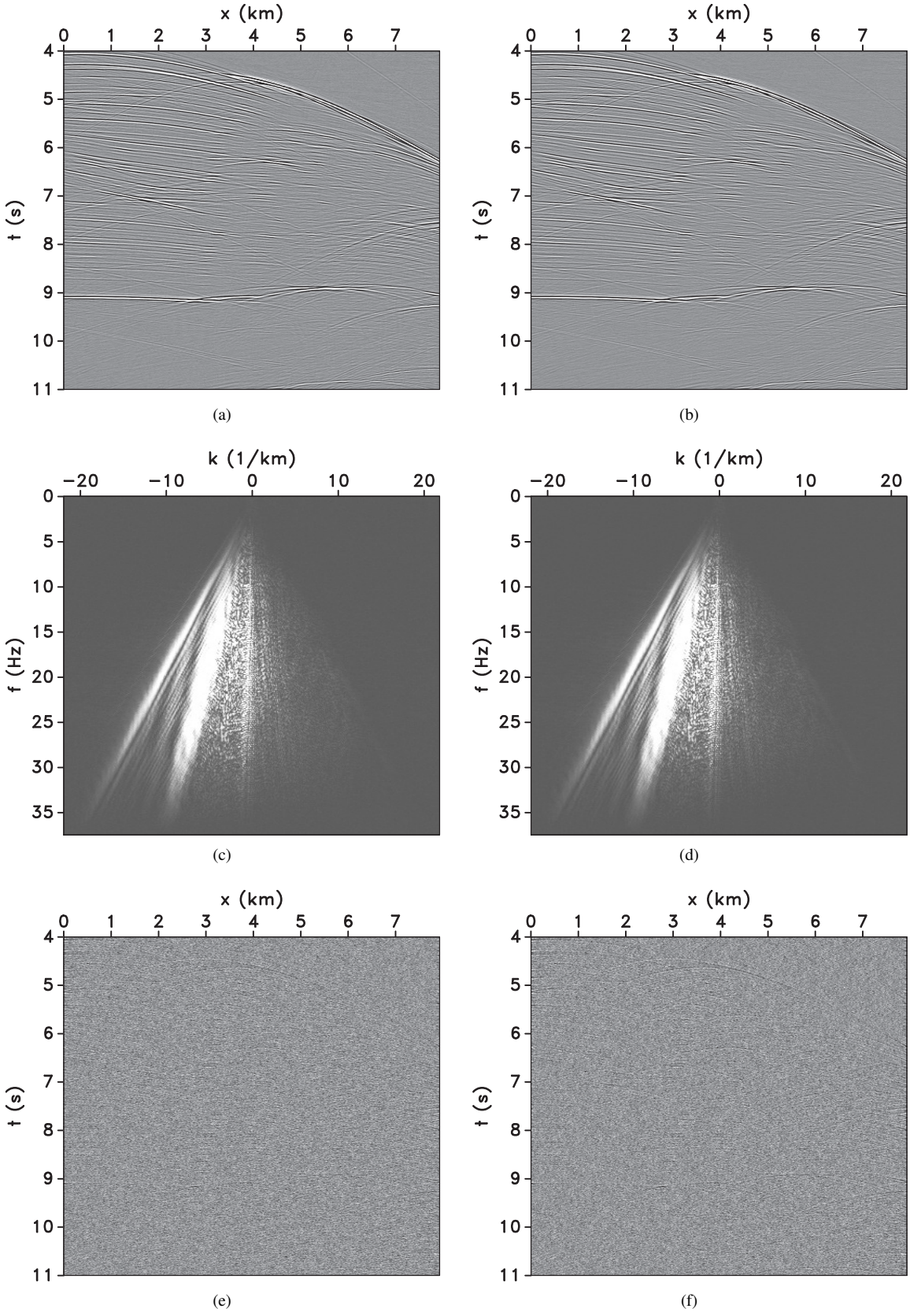


Figure 6. Denoised shot gathers: (a), (b) Synthesis and analysis solutions with SNR of 17.46 dB and 18.22 dB, respectively. (c), (d) Corresponding f-k spectrum. (e), (f) Corresponding difference plots.

5.2 Missing trace reconstruction

We evaluate the results of the missing trace reconstruction for the shot gather in Figure **??** using the three different undersampling schemes proposed earlier. For every experiment, we undersample the data by 50%. In our first experiment, we uniformly undersample the original shot gather by removing one out of every two traces. Figure **??** shows the shot gather after the undersampling, while Figure **??** shows the corresponding f-k spectrum. Note that the usual aliasing is present, as expected. In Figure **??**, we show the result, difference plot and f-k spectrum for both approaches. We can see that the estimated shot gathers have some artifacts, which is expected because the uniform undersampling scheme does not favor the recovery through sparsity constraints. The f-k spectra also shows that the f-k spectrum of the synthesis solution still features less residual alias than that of the analysis approach. This fact, along with the difference plots and the SNR difference, points that the analysis estimate is closer to the original shot gather than that of the synthesis.

In the second experiment, we randomly undersample the original shot gather as a function of receiver position. Figure **??** shows the shot gather after random undersampling, and Figure **??** shows the corresponding f-k spectrum. Note that the undersampled shot gather features gaps of varying size. Also, the f-k spectrum shows the spectral leakage now appears as somewhat incoherent artifacts, as expected from our previous example in Figure 3. Although more incoherent than the aliases present in the f-k spectra of Figure **??**, these artifacts are not as incoherent as the ones in the noisy case. This might be because the undersampling is only performed in one dimension of the two dimensional image, in contrast with our one dimensional examples. Figure **??** shows the result, difference plot and f-k spectrum for both approaches. These solutions are closer to the original shotgather when compared to the one obtained using the uniform undersampling scheme, as evidenced by the SNR and difference plots. Also, the artifacts that are present in the solutions using the undersampling scheme are mostly gone. This is accurate because the random undersampling produces incoherent aliasing, which favors sparsity-constrained inversion. However, notice that in the areas where the gaps between traces are large, the quality of the estimated solution is lower. The f-k spectra of the solutions show that most of the incoherent alias is attenuated for both approaches, but the spectra of the analysis solution show less spectral leakage. This assures that the solution estimated by the analysis approach is more accurate than that of the synthesis approach for this undersampling scheme.

Finally, Figure **??** illustrates the shot gather after the jittered undersampling and Figure **??** the corresponding f-k spectrum. The undersampled shot gather has controlled gap sizes between traces. Notice that the f-k spectrum is very similar to the one obtained after random undersampling. Figure **??** shows the result, difference plot and f-k spectrum for both approaches. In terms of SNR and difference plots, one can observe that this undersampling scheme yields the best solutions for sparsity-constrained inversion among the three schemes used. However, the SNR difference between these solutions and those using the random undersampling schemes are small, indicating that the benefit of the jittered undersampling scheme is marginal when compared to the random undersampling scheme. Also, notice that, because the gaps between traces are smaller, no artifacts or low quality traces are visible in the estimated shot gathers. The f-k spectra of both solutions show that the analysis approach can also attenuate more spectral leakage than the synthesis approach with the jittered undersampling scheme.

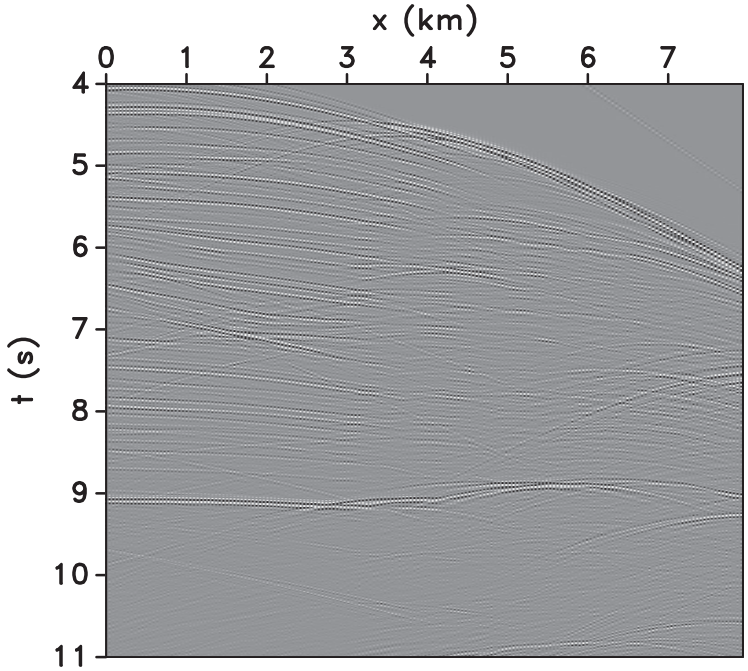
6 DISCUSSION

We define two approaches for sparsity-constrained inversion, called the synthesis and analysis approaches. We show that they are derived from the same problem and, although similar, search for different solutions. In particular, the synthesis approach attemps to find the sparsest solution for the problem in the transform domain using a soft thresholding operator, while the analysis approach uses an orthogonal projection approach to find a dataset with a forward transform given by the canonical coefficients.

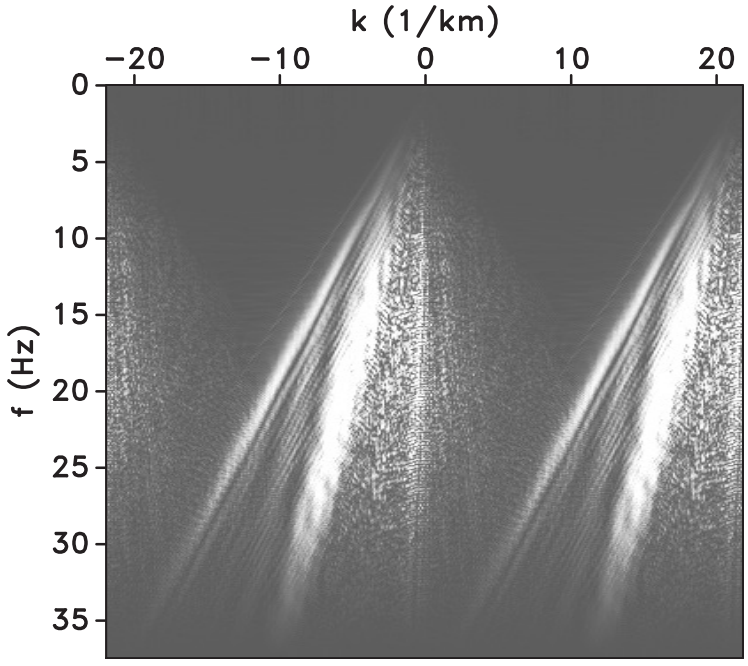
An analysis of the SNR results and difference plots demonstrates that the solution of the analysis approach is superior (in the sense that it is closer to the original data) compared to that of the synthesis approach for both problems. Because the missing trace reconstruction problem can also be interpreted as the denoising of the f-k spectrum of the missing data, this corroborates that the analysis approach is better than the synthesis approach for denoising purposes. The results are coherent with our prediction that the analysis approach is more robust to noise due to the fact that the solution is denser than that of the synthesis approach, which means that erroneous estimation of the solution coefficients is less important. However, we remark that the success of both approaches is very application dependent, which means that the synthesis approach might yield superior or equivalent results with a different type of data or application.

For the missing trace reconstruction problem, when considering the different undersampling schemes, recovery using uniform undersampling is the least effective due to the presence of coherent aliases in the f-k spectrum of the measured data. This is noticed both in terms of SNR and visually, as some artifacts are present in the estimated solutions of both approaches. However, even in this adverse scenario, the analysis solution features less residual spectral leakage in its f-k spectrum, when compared to that of the synthesis solution. This indicates that the analysis approach may be more robust to regular undersampling and coherent noise than the synthesis approach, at least when the curvelet frame is the transform domain.

The random and jittered undersampling schemes yield better solutions in terms of SNR when compared to the uniform undersampling case for the analysis and synthesis approaches. Also, the SNR values of the estimated solutions are very similar for both schemes, which indicates that the benefit of jittered undersampling is marginal in terms of SNR compared to random undersampling. However, one can observe



(a)



(b)

Figure 7. (a) Shot gather after uniform undersampling. (b) Corresponding f-k spectrum.

that the quality of the solution using the random undersampling scheme is lower than that of the jittered one, especially in areas where gaps between traces are larger. As explained before, this is due to the fact of the gaps are larger than the transform elements employed in the reconstruction scheme, which makes the reconstruction less accurate.

In this report, we use the algorithm proposed by Becker et al. (2011) (see Appendix B) to perform the synthesis and analysis minimizations. As outlined in the theory section of

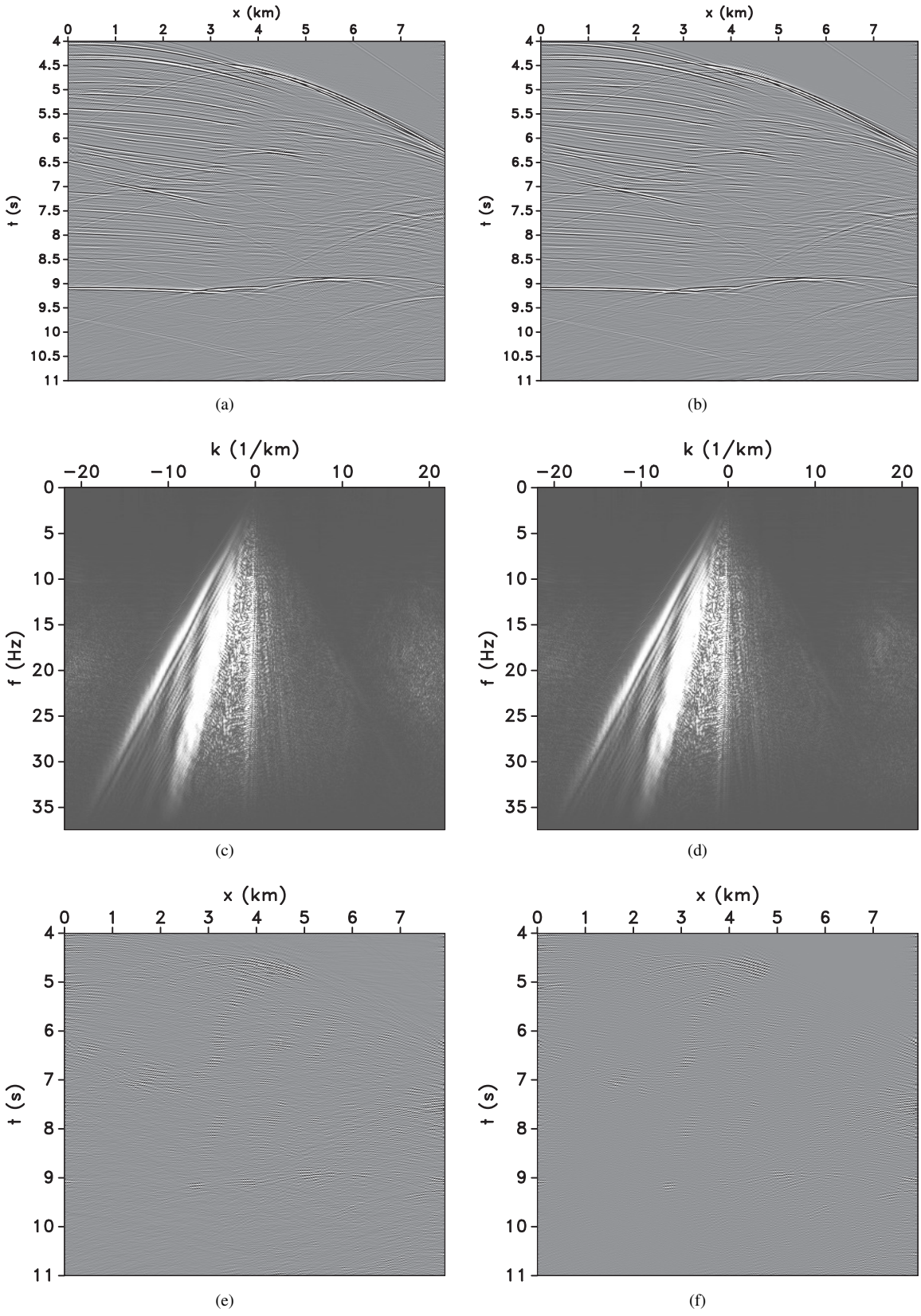
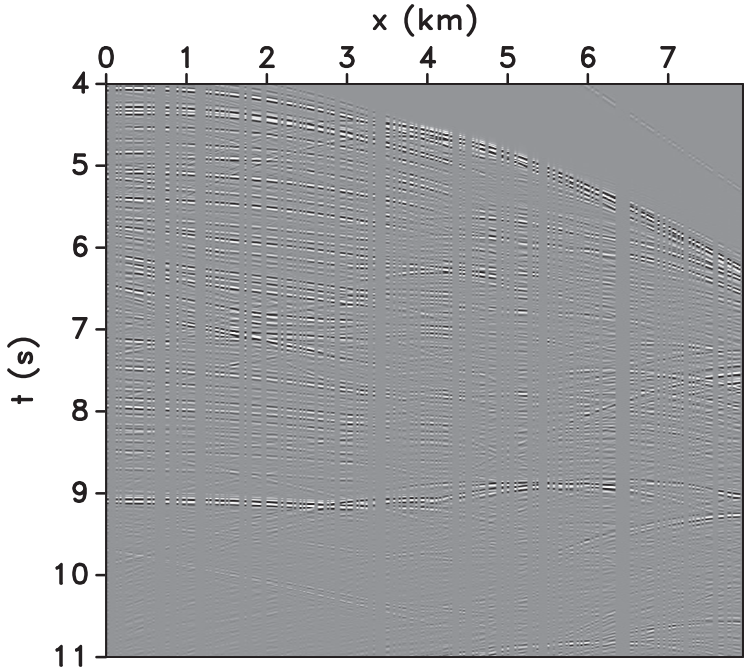
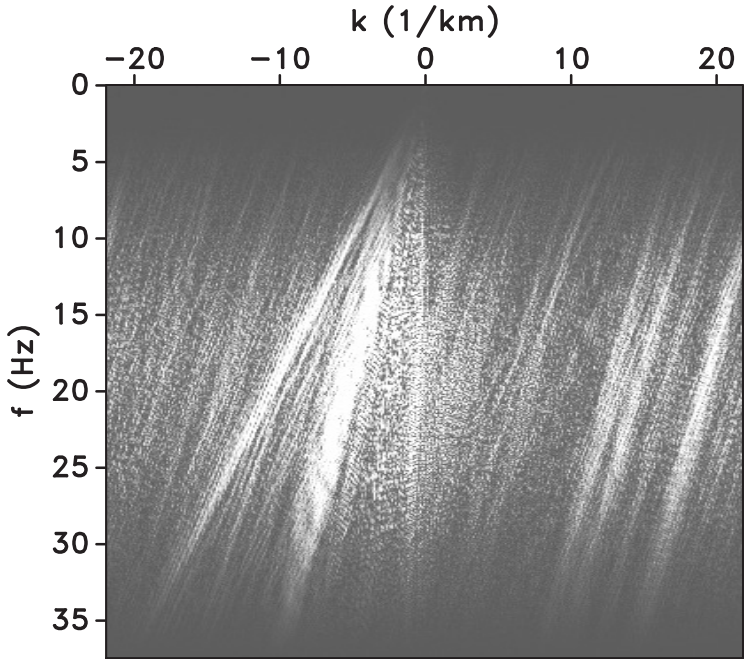


Figure 8. Uniform undersampling: (a), (b) Synthesis and analysis solutions. SNR of 13.29 dB and 14.83 dB, respectively. (c), (d) Corresponding f-k spectra. (e), (f) Corresponding difference plots scaled by ten.



(a)



(b)

Figure 9. (a) Shot gather after random undersampling. (b) Corresponding f-k spectrum.

this paper, this algorithm solves the L_1 minimization formulation of the sparsity-constrained inversion. We are aware that recent developements, such as Nam et al. (2013), point to algorithms that solve for the L_0 minimization problem which, in some cases, yield better results than the L_1 formulation. We choose to use the L_1 minimization procedure due to the convexity of the formulation and overall computational cost, since L_0 minimization approaches are usually greedy-like and thus have a higher computational cost than the L_1 formulation. Other L_1 formulation options, such as reweighted L_1 formu-

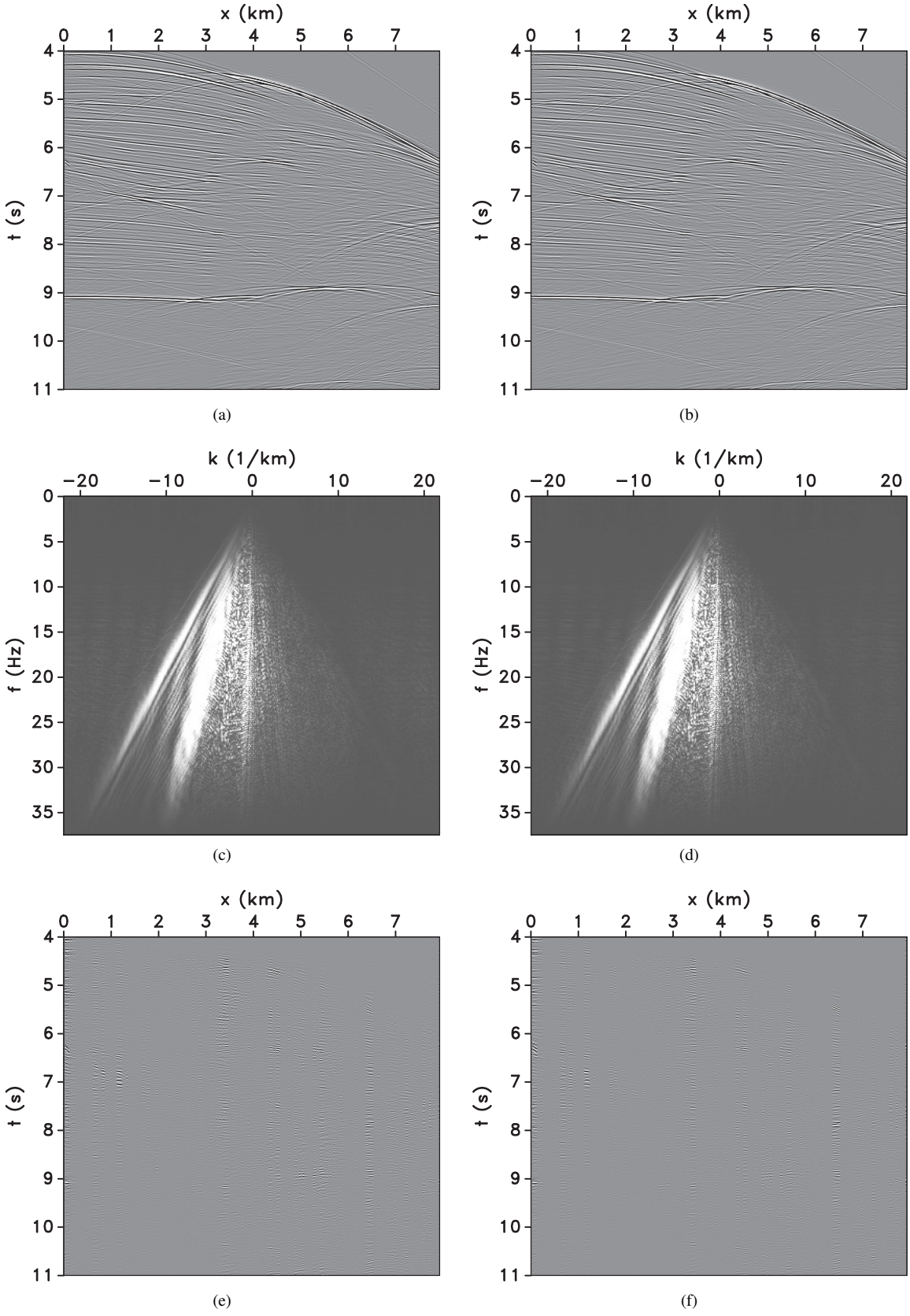
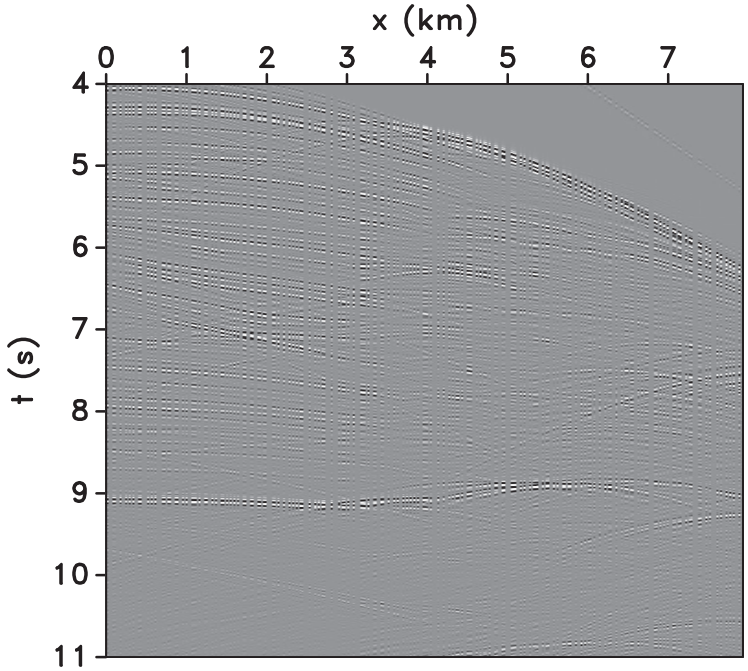
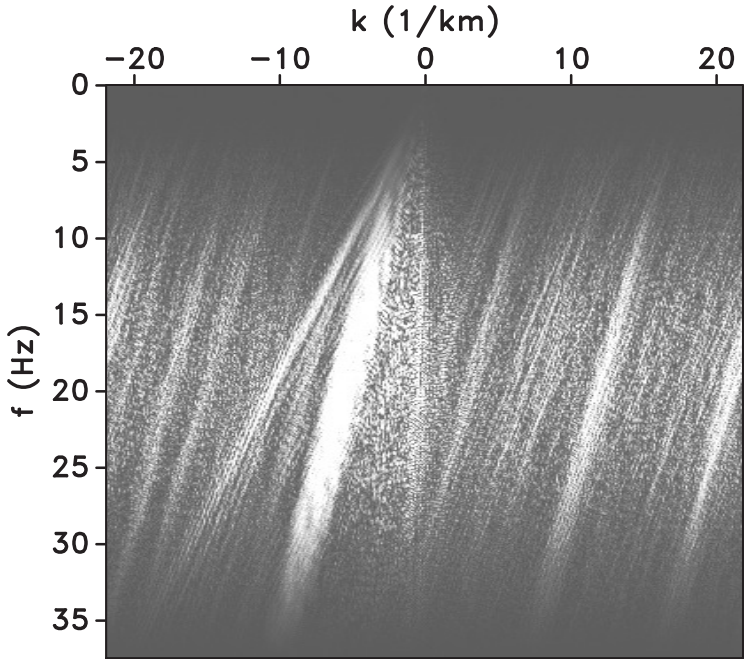


Figure 10. Random undersampling: (a), (b) Synthesis and analysis solutions. SNR of 16.24 dB and 17.87 dB, respectively. (c), (d) Corresponding f-k spectra. (e), (f) Corresponding difference plots scaled by ten.



(a)



(b)

Figure 11. a) Shot gather after jittered undersampling. (b) Corresponding f-k spectrum.

lations (Candes et al., 2008) can also be considered to solve the problem with high accuracy.

Finally, all the experiments in this paper are carried out using either the synthesis approach or the analysis approach. However, we showed that these approaches are special cases of equation 7. If one were to solve for $\mathbf{u}$ using a value of κ that is not 0 or ∞, equation 7 would yield a solution that is neither the sparsest one possible nor the one in $Range(\mathbf{\Theta})$. Although this is a viable option and the existing algorithms that solve for the synthesis and analysis approaches can be adapted to solve

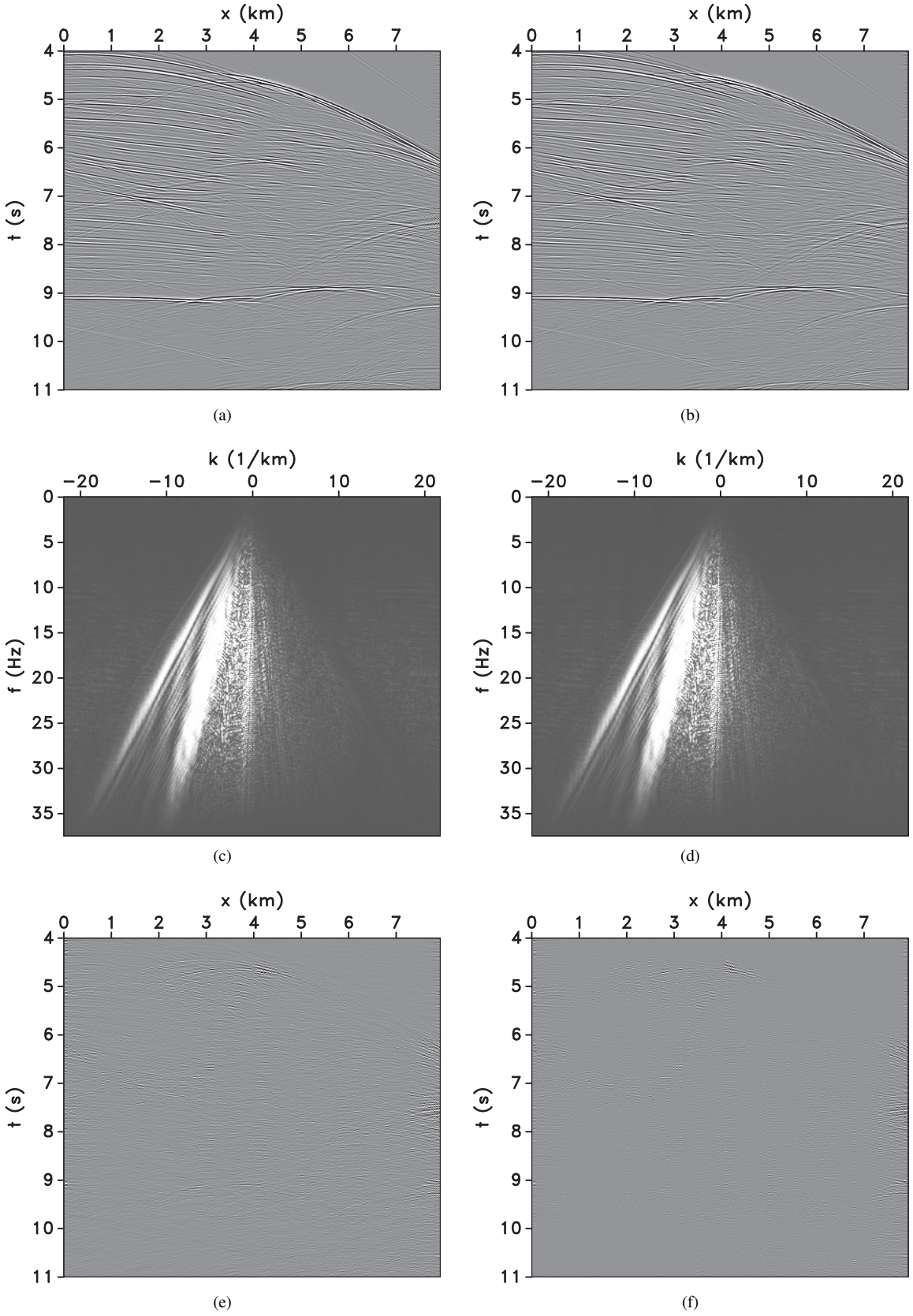


Figure 12. Jittered undersampling: (a), (b) Synthesis and analysis solutions. SNR of 16.26 dB and 17.93 dB, respectively. (c), (d) Corresponding f-k spectra. (e), (f) Corresponding difference plots scaled by ten.

for equation 7, we leave the evaluation of this approach for future work.

7 CONCLUSIONS

The state-of-the-art solutions for the denoising and missing trace reconstruciton problems are based on transform domain approaches. In the context of sparsity-promoting inversion, two key approaches are identified: the analysis approach and the synthesis approach. In this report, we compared both approaches these two types of problems. For the denoising case, our examples show that the analysis approach is superior to the synthesis approach, obtaining a clean image with less damage to the signal and thus attaining a higher SNR than that of the synthesis approach. For the missing trace reconstruction problem, our examples show that the analysis approach consistently finds better solutions than the synthesis counterpart for any of the undersampling schemes used. We attribute the success of the analysis approach in these settings to its number of constraints and amenability to noise, which makes it better in finding signals inside noisy data. We showed that the analysis approach is a strong candidate for solving inverse problems in geophysical applications when a sparsity-promoting constraint using the curvelet frame is efficient, and its use should also be considered alongside the synthesis approach. In particular, since the success of both approaches highly depends on the dataset, application and transform domain, both approaches should be tested before making a definitive conclusion for a given problem. Future work includes further evaluation of the synthesis and analysis approaches when applied to other inverse problems in the seismic field, such as deblending and full-waveform inversion.

8 ACKNOWLEDGMENTS

We would like to thank sponsor companies of the Center for Wave Phenomena, whose support made this research possible. We would also like to thank the authors of CurveLab (http://www.curvelet.org/) and NESTA (http://statweb.stanford.edu/∼candes/nesta/) for making their codes available. We would like to acknowledge the iTeam for the valuable discussions and ideas. The reproducible numeric examples in this paper use the Madagascar open-source software package (Fomel et al., 2013) freely available from http://www.ahay.org.

REFERENCES

Almeida, L. J., R. R. Manenti, M. J. Porsani, et al., 2015, Coherent noise attenuation using the wavelet transform on radial basis: Presented at the 2015 SEG Annual Meeting, Society of Exploration Geophysicists.

Becker, S., J. Bobin, and E. J. Candès, 2011, Nesta: a fast and accurate first-order method for sparse recovery: SIAM Journal on Imaging Sciences, **4**, 1–39.

Cai, J.-F., B. Dong, S. Osher, and Z. Shen, 2012, Image restoration: Total variation, wavelet frames, and beyond: Journal of the American Mathematical Society, **25**, 1033–1089.

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

Candès, E. J., and D. L. Donoho, 2004, New tight frames of curvelets and optimal representations of objects with piecewise c2 singularities: Communications on pure and applied mathematics, **57**, 219–266.

Candes, E. J., Y. C. Eldar, D. Needell, and P. Randall, 2011, Compressed sensing with coherent and redundant dictionaries: Applied and Computational Harmonic Analysis, **31**, 59–73.

Candès, E. J., J. Romberg, and T. Tao, 2006, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information: Information Theory, IEEE Transactions on, **52**, 489–509.

Candes, E. J., M. B. Wakin, and S. P. Boyd, 2008, Enhancing sparsity by reweighted 1 minimization: Journal of Fourier analysis and applications, **14**, 877–905.

Chen, Y., J. Ma, and S. Fomel, 2016, Double sparsity dictionary for seismic noise attenuation: Geophysics, **81**, V103–V116.

Crawley, S., R. Clapp, J. Claerbout, et al., 1999, Interpolation with smoothly nonstationary prediction-error filters: Presented at the 1999 SEG Annual Meeting, Society of Exploration Geophysicists.

Daubechies, I., M. Defrise, and C. De Mol, 2004, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint: Communications on pure and applied mathematics, **57**, 1413–1457.

Daubechies, I., et al., 1992, Ten lectures on wavelets: SIAM, **61**.

Donoho, D. L., 2006, Compressed sensing: Information Theory, IEEE Transactions on, **52**, 1289–1306.

Elad, M., P. Milanfar, and R. Rubinstein, 2007, Analysis versus synthesis in signal priors: Inverse problems, **23**, 947.

Fomel, S., and Y. Liu, 2010, Seislet transform and seislet frame: Geophysics, **75**, V25–V38.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013, Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, **1**.

Hennenfent, G., and F. J. Herrmann, 2006, Seismic denoising with nonuniformly sampled curvelets: Computing in Science & Engineering, **8**, 16–25.

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

Herrmann, F. J., D. Wang, G. Hennenfent, and P. P. Moghaddam, 2007, Curvelet-based seismic data processing: A multiscale and nonlinear approach: Geophysics, **73**, A1–A5.

Li, X., A. Y. Aravkin, T. van Leeuwen, and F. J. Herrmann, 2012, Fast randomized full-waveform inversion with compressive sensing: Geophysics, **77**, A13–A17.

Mallat, S., 2008, A wavelet tour of signal processing: the

sparse way: Academic press.
Naghizadeh, M., and M. D. Sacchi, 2010, Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data: Geophysics, **75**, WB189–WB202.
Nam, S., M. E. Davies, M. Elad, and R. Gribonval, 2013, The cosparse analysis model and algorithms: Applied and Computational Harmonic Analysis, **34**, 30–56.
Nesterov, Y., 2005, Smooth minimization of non-smooth functions: Mathematical programming, **103**, 127–152.
Oliveira, M., M. Henriques, F. Leite, G. Corso, and L. Lucena, 2012, Seismic denoising using curvelet analysis: Physica A: Statistical Mechanics and its Applications, **391**, 2106–2110.
Porsani, M. J., 1999, Seismic trace interpolation using half-step prediction filters: Geophysics, **64**, 1461–1467.
Ramirez, C., G. Larrazabal, and G. Gonzalez, 2016, Salt body detection from seismic data via sparse representation: Geophysical Prospecting, **64**, 335–347.
Rodriguez, I. V., and M. D. Sacchi, 2014, Microseismic source imaging in a compressed domain: Geophysical Journal International, **198**, 1186–1198.
Selesnick, I. W., and M. A. Figueiredo, 2009, Signal restoration with overcomplete wavelet transforms: comparison of analysis and synthesis priors: SPIE Optical Engineering+ Applications, International Society for Optics and Photonics, 74460D–74460D.
Spitz, S., 1991, Seismic trace interpolation in the fx domain: Geophysics, **56**, 785–794.
Stolt, R. H., 2002, Seismic data mapping and reconstruction: Geophysics, **67**, 890–908.
Wason, H., F. J. Herrmann, T. T. Lin, et al., 2011, Sparsity-promoting recovery from simultaneous data: A compressive sensing approach: Presented at the 2011 SEG Annual Meeting, Society of Exploration Geophysicists.
Xu, S., Y. Zhang, D. Pham, and G. Lambaré, 2005, Antileakage fourier transform for seismic data regularization: Geophysics, **70**, V87–V95.
Yang, P., and S. Fomel, 2015, Seislet-based morphological component analysis using scale-dependent exponential shrinkage: Journal of Applied Geophysics, **118**, 66–74.
Yu, S., J. Ma, X. Zhang, and M. D. Sacchi, 2015, Interpolation and denoising of high-dimensional seismic data by learning a tight frame: Geophysics, **80**, V119–V132.
Zhu, L., E. Liu, and J. H. McClellan, 2015, Seismic data denoising through multiscale and sparsity-promoting dictionary learning: Geophysics, **80**, WD45–WD57.

Appendix A

Orthonormal Bases and Frames

Let $\{\phi_n\}_{n:\Gamma\in\mathbb{Z}}$ be a sequence of vectors indexed by a (possibly infinite) set Γ. An orthonormal basis for a inner product space H is the set of vectors $\{\phi_n\}_{n:\Gamma\in\mathbb{Z}}$ that meet the following conditions:

- Each vector ϕ_n is an element of H,
- $\|\phi_n\|_H = 1$ for all the elements in the set,
- $\langle\phi_n, \phi_l\rangle = 0$ for $n \neq l$,
- The set of vectors $\{\phi_n\}_{n:\Gamma\in\mathbb{Z}}$ spans H.

The above conditions imply several properties of an orthonormal basis. For instance, elements of an orthonormal basis are always orthogonal to each other. Using an orthonormal basis, it is possible to represent any element of H with a linear transformation using this orthonormal basis. In particular, for a given element $f \in H$, we have

$$u_n = \langle\phi_n, f\rangle. \tag{A.1}$$

We denote the sequence of coefficients $\{u_n\}_{n:\Gamma\in\mathbb{Z}} < \infty$ as the representation of f in the transform domain. Because an orthonormal basis is orthogonal, it is not a singular matrix. We also remark that an orthobasis conforms with Parseval's theorem, i.e.,

$$\|f\|_H^2 = \|u_n\|_{L^2(\Re)}^2, \tag{A.2}$$

which means that the energy of the original element of H is preserved in the transform domain. In general, orthonormal basis provide an one-to-one mapping from H to the transform domain and provide stable recovery of the signal using the inverse matrix. Examples of orthonormal bases are the Fourier transform and wavelet transform.

A generalization of this concepts are frame constructions. We define that $\{\phi_n\}_{n:\Gamma\in\mathbb{Z}}$ is a frame if there exists constants $0 < A \leq B < \infty$ such that

$$A\|f\|_H^2 \leq \sum_{i\in\Gamma}\langle\phi_i, f\rangle^2 \leq B\|f\|_H^2. \tag{A.3}$$

Frames with constants $A = B$ are called tight frames, and every orthonormal basis is a tight frame with $A = B = 1$. Frames that are not orthonormal might have a much bigger number of vector elements than its dimension. In other words, a frame is defined as a matrix $\mathbf{\Phi}_{P\times N}$, where $P > N$. Because of its intrinsic linear dependency, frames are redundant - several representations u_n are possible for every signal $f \in H$. Stable recovery is possible using the pseudo-inverse of the frame. Common examples of frames are the undecimated wavelet, Gabor, and curvelet transforms.

Appendix B

NESTA - Nesterov Algorithm

In this report, we have used our own implementation of the NESTA solver (Becker et al., 2011) to perform the analysis and synthesis minimization problems present in equation 15 and equation 16. For completeness, we include here the main steps of this solver. For benchmark comparisons with other state-of-the-art solvers and theoretical background, we refer the reader to Becker et al. (2011) and Nesterov (2005), respectively.

The NESTA solver proposes to minimize the following function

$$f(\mathbf{x}) = \|\mathbf{x}\|_1 = \max_{\mathbf{u}\in Q_d} \langle \mathbf{u}, \mathbf{x}\rangle, \tag{B.1}$$

on the feasible set given by

$$Q_d = \{\mathbf{u} : \|\mathbf{u}\|_\infty \leq 1\}. \tag{B.2}$$

A smooth approximation to equation B.1 is given by

$$f(\mathbf{x})_\mu = \|\mathbf{x}\|_1 = \max_{\mathbf{u}\in Q_d} \langle \mathbf{u}, \mathbf{x}\rangle - \mu p_d(\mathbf{u}), \tag{B.3}$$

where p_d is a prox-function. Note that, as $\mu \longrightarrow 0$, the better the minimization of the L_1 norm of the function. The function present in equation B.3 can be minimized using the following iterative procedure (Nesterov, 2005):

Algorithm 1 General NESTA Solver

for each iteration i **do**
  $\alpha_i = 1/2(i+1)$, $\tau_i = 2/(i+3)$
  Calculate $\nabla f(\mathbf{x}_i)$
  $\mathbf{y}_i = \min_{x\in Q_p} (L/2)\|\mathbf{x} - \mathbf{x}_i\| + \langle \nabla f(\mathbf{x}_i), \mathbf{x} - \mathbf{x}_i\rangle$
  $\mathbf{z}_i = \min_{x\in Q_p} (L/\sigma_p) p_d(\mathbf{x}) + \sum_{j=0}^{i} \alpha_j \langle \nabla f(\mathbf{x}_j), \mathbf{x} - \mathbf{x}_j\rangle$
  $\mathbf{x}_{i+1} = \tau_i \mathbf{z}_i + (1-\tau_i)\mathbf{y}_i$
end for
return $\mathbf{x}_i$

In the above algorithm, σ_p is the convexity parameter of p_p and L is the Lipschitz constant of the gradient. The NESTA solver considers $p_d(\mathbf{u}) = 1/2\|\mathbf{u}\|_2^2$, which makes equation B.3 the Huber function. This function has convexity parameter equal to one and a gradient that is Lipschitz with constant $1/\mu$. The gradient for this function can be calculated as

$$\nabla f_\mu(\mathbf{x})[k] = \begin{cases} x[k]/\mu & \text{if } |x[k]| < \mu \\ sgn(x[k]) & \text{otherwise,} \end{cases} \tag{B.4}$$

where $x[k]$ is an element of the vector $\mathbf{x}$. We want to minimize equation B.3 in the feasible set $Q_d = \{\mathbf{x} : \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon\}$, which is an $L2$ ball enclosing the portion of the ambient domain where the data misfit is up to ϵ. In this appendix, $\mathbf{b}$ refers to the measured data. Assuming that $\mathbf{A}$ is an orthogonal projector, i.e., $\mathbf{A}^*\mathbf{A} = \mathbf{I}$ or $\mathbf{A}\mathbf{A}^* = \mathbf{I}$, the iterative process described above can be efficiently carried out using Algorithm 2.

Although the above algorithm works only for orthogonal projections, it is importante to notice that this assumption is widely present in compressive sensing approaches. Notably, Algorithm 2 does not require any matrix inversion, which allows for a fast solver. The aforementioned algorithm solves both the analysis and synthesis minimization problems by making the following change of variables: suppose $\mathbf{A} = \mathbf{R}\mathbf{U}$,

Algorithm 2 NESTA Solver for Orthogonal Projections

Set $\mathbf{x}_0$, $\mathbf{b}$, μ, $L = 1/\mu$, ϵ
for each iteration i **do**
  $\alpha_i = 1/2(i+1)$, $\tau_i = 2/(i+3)$
  Calculate $\nabla f(\mathbf{x}_i)$
  $\mathbf{q}_{y,i} = \mathbf{x}_i - (1/L)\nabla f(\mathbf{x}_i)$
  $\lambda_{y,i} = \max(0, \epsilon^{-1}\|\mathbf{b} - \mathbf{A}\mathbf{q}_{y,i}\|_2 - L)$
  $\mathbf{y}_i = (I - \frac{\lambda_{y,i}}{(\lambda_{y,i}+L)}\mathbf{A}^*\mathbf{A})((\lambda_{y,i}/L)\mathbf{A}^*\mathbf{b} + \mathbf{q}_{y,i})$
  $\mathbf{q}_{z,i} = \mathbf{x}_0 - (1/L)\sum_{j=0}^{i} \alpha_j \nabla f(\mathbf{x}_j)$
  $\lambda_{z,i} = \max(0, \epsilon^{-1}\|\mathbf{b} - \mathbf{A}\mathbf{q}_{z,i}\|_2 - L)$
  $\mathbf{z}_i = (I - \frac{\lambda_{z,i}}{(\lambda_{z,i}+L)}\mathbf{A}^*\mathbf{A})((\lambda_{z,i}/L)\mathbf{A}^*\mathbf{b} + \mathbf{q}_{z,i})$
  $\mathbf{x}_{i+1} = \tau_i \mathbf{z}_i + (1-\tau_i)\mathbf{y}_i$
end for
return $\mathbf{x}_i$

where $\mathbf{R}$ is a restriction matrix and $\mathbf{U}$ is a frame. Substituting $\mathbf{x} \leftarrow \mathbf{U}\mathbf{x}$, one can apply the algorithm to solve the analysis problem (Becker et al., 2011).

Regarding parameters, we use the ones recommended in (Becker et al., 2011). We set $\mu = 0.02$ for every experiment. For the noise level ϵ, we set $\epsilon = (\sqrt{m + 2\sqrt{2m}})\sigma$, where m is the number of elements in $\mathbf{b}$ and σ is the standard deviation of the noise.