Wavefield reconstruction using wavelet transform

Iga Pawelec[1], Paul Sava[1] & Michael Wakin[2]
[1] *Center for Wave Phenomena, Colorado School of Mines*
[2] *SINE Center for Research in Signals and Networks, Colorado School of Mines*

ABSTRACT
We propose wavelet-based data reconstruction to interpolate land data with large dynamic range without amplitude processing or windowing. We test two approaches using compressive sampling to recover full unaliased data: sparsity promoting reconstruction by ℓ_1 minimization and projection onto convex sets. Unlike the Fourier domain, the wavelet domain provides a good representation of non-stationary signals and allows to rebuild data of high dynamic range with relatively small percentage of all coefficients. We solve an ℓ_1 minimization problem to find a sparse representation of full data in the wavelet domain and compare it with results of a wavelet-domain POCS algorithm. Tests on synthetic and field data reveal that both approaches can recover missing data highly coherent with the existing data, while taking advantage of the full dynamic range of the data.

Key words: interpolation, wavelets, compressive sensing, land acquisition

1 INTRODUCTION

Modern techniques of seismic data analysis and inversion, such as AVO/AVA, full waveform inversion (FWI) and least squares reverse time migration (LSRTM), operate on big data volumes. However, multi-azimuth, long offset dense surveys are expensive to acquire and sometimes access restrictions do not permit acquisition over certain areas, resulting in data gaps. Seismic data reconstruction plays a key role in both scenarios, allowing to fill-in missing data before advanced processing and inversion are performed.

Mosher et al. (2017) show that a novel approach, utilizing ideas from compressive sensing, can significantly speed up acquisition without jeopardizing data quality. The underlying idea is to randomize receiver placement and shot timing according to compressive sampling rules and solve a large scale regularization problem, recovering a full data volume from reduced measurements. To recover full data, one must develop a strategy for dense data recovery that takes into account acquisition geometry: parameters such as source and receiver locations, maximum gap size and timing of the shots have to be considered. For example, most conventional 3D acquisition geometries have regular but poor sampling in at least one direction (Trad, 2009) and one has to contend with aliasing and sometimes big data gaps due to access restrictions. In contrast, compressive sensing surveys would have deliberate irregular sampling and often simultaneous shooting. Thus, data recovery strategy needs to be tuned with particular geometry restrictions in mind.

Current techniques for infilling missing data can be divided in several categories, including prediction error filters (PEFs) (Spitz, 1991), tensor completion (Kreimer et al., 2013), rank reduction (Chen et al., 2016) and deep learning (Wang et al., 2019). However, the most widely used interpolation techniques are transform-based approaches. Such methods are well-studied in the context of data aliasing and irregular sampling and rely on data representation in a transform domain to recover missing information. Although different transforms have been used, including Radon transform (Kabir and Verschuur, 1995; Yu et al., 2007; Wang et al., 2010) and wavelet or seislet transforms (Yu et al., 2007; Gan et al., 2015), the Fourier transform remains the most popular choice because it is easy to interpret and fast to compute. Liu and Sacchi (2004) develop a framework for data recovery based on weighted norm minimization, using spectral weights bootstrapped from FK representation of data. This framework can be extended to five dimensions (Trad, 2009). The Fourier domain is also used in the projection onto convex sets (POCS) method described by Abma and Kabir (2006). To deal with problems of non-uniform sampling and aliasing artifacts in the Fourier domain, Xu et al. (2005, 2010) propose an antileakage version of the Fourier transform. One downside of Fourier-based approaches is that data have to be windowed for non-stationarity and, as a consequence, only local information is used for interpolation. Another attractive transform for seismic data interpolation, gaining significant popularity, is the curvelet transform (Hennenfent et al., 2010; Herrmann, 2010; Naghizadeh and Sacchi, 2010). Curvelets provide an optimally sparse representation of seismic wavefields (Candès and Demanet, 2005), but their redundancy implies that for a dataset of size N, as many

as $7 \times N$ curvelet coefficients have to be computed, depending on the chosen number of scales, which can be prohibitively expensive for large 3D datasets.

In this paper, we present a seismic data reconstruction strategy exploiting wavelet domain sparsity under randomized acquisition. We discuss the features of seismic signals in the wavelet domain, review the theory highlighting the favorable recovery conditions and present data reconstruction results for synthetic and field data with two strategies: projection onto convex sets (POCS) and ℓ_1 norm minimization. Wavelet transform is fast to compute and represents well large dynamic range in data. The POCS approach iteratively restores missing data, but its success hinges on developing a good thresholding strategy. ℓ_1 optimization is less robust to acquisition geometry, but does not need any thresholding.

2 CHALLENGES IN DATA RECONSTRUCTION

Several challenges have to be addressed for successful data reconstruction: the presence of data aliasing, the pattern of missing traces and size of data gaps, and the dynamic range of seismic data. These challenges are more prominent for land seismic acquisition due to the highly complex heterogeneous shallow subsurface, which traps a large portion of energy released by the seismic source and produces slowly propagating surface waves (Keho and Kelamis, 2012). In the following, we discuss these challenges in more detail and explain how different transforms handle them.

In land seismic data, aliasing of surface waves can be especially severe due to the much slower velocities of surface waves compared with the body waves. Figure 1 shows the same land data record sampled at different trace intervals (coarse sampling results from discarding a portion of the full data) and the corresponding frequency spectra. Aliasing occurs in this example even at 2.5m sampling interval, which is 10 times finer than what would commonly be used in big land surveys. Severe aliasing makes it difficult to use surface waves for characterizing the shallow subsurface or to remove the surface waves from the seismic record entirely.

One way to overcome the aliasing problem is by data reconstruction exploiting prediction error filters (PEF) (Spitz, 1991). The underlying idea is that filter coefficients derived from low, unaliased frequencies can be used to interpolate aliased data components. Naghizadeh and Sacchi (2008) use this concept to develop adaptive PEFs, Naghizadeh and Sacchi (2010) utilize unaliased scales in the curvelet domain for reconstructing aliased data, and Gan et al. (2015) take advantage of low-pass filtered data to interpolate using seislets. Another popular data reconstruction strategy, the minimum weighted norm interpolation (Liu and Sacchi, 2004), requires adjustments to spectral weights to handle aliased data since additional energy is present for aliased components. Despite these advances, the degree of aliasing present in land seismic data may prove to be too much to handle in an elegant way, suggesting that a change in the approach to data reconstruction might be a better solution.

Historically, seismic data have been acquired on a regular grid or have been regularized after acquisition - a pragmatic choice, since many processing and imaging algorithms require regular spacing. However, such acquisition is limited by the Nyquist - Shannon sampling theorem (Candès et al., 2006a) which dictates a sampling rate of at least two points per wavelength for successful recovery of a non-aliased signal. Furthermore, the number of sensors needed to record good quality, unaliased land data on a regular grid is exceedingly high. The advent of compressive sensing (CS) (Candès et al., 2006b) opened new, exciting possibilities for signal reconstruction from incomplete information. Hennenfent and Herrmann (2008) and Herrmann (2010) examine randomized acquisition using much fewer sensors than a regular-grid survey and achieve comparable data density and quality. Mosher et al. (2017) demonstrate that compressive sensing can be successfully applied to field seismic acquisition. The main requirement for data reconstruction is sparse representation in a known transform domain. We discuss sparse recovery in more detail in the following section.

The curvelet domain is optimal for representing wave phenomena (Candès and Demanet, 2005). The curvelet transform divides the frequency plane into dydaic bands which are then split into overlapping angular wedges doubling in every other dydaic scale. The curvelet transform is highly redundant: there is no unique representation of a signal in the curvelet domain and the number of curvelet coefficients is much larger than the number of data points. This feature of the curvelet transform is favorable for denoising and finding sparse signal representation, at the expense of increased storage requirements, which makes curvelets a memory-expensive choice for large datasets.

The wavelet transform on the other hand offers a good middle ground between the frequency and curvelet domains. Wavelets provide a so called multiresolution approximation and in 2D are sensitive to three directions: horizontal, vertical and diagonal. Although the wavelet representation of wavefields is less sparse than a curvelet representation, the wavelet transform can be orthogonal, providing a unique representation of the signal and preserving its total energy. Another advantage of the wavelet transform is its computational speed surpassing even that of a Fast Fourier Transform, thus making wavelets suitable for analysis of large datasets.

Wavelets are also a good choice for handling non-stationary signals. The dynamic range of seismic data, spanning several orders of magnitude, is particularly difficult to handle by data reconstruction algorithms, so windowing or data gaining are often used to avoid dealing with the full data range. Consider the way humans would interpolate missing data. We would look at the available portion of data to find patterns and then fill-in the gaps under assumptions that observed trends are also present in gaps. However, given a raw land seismic record, such task becomes impossible, because unless gain or trace balancing is applied, only a small range of offsets and early times are visible to the eye. We would be unable to interpolate something we cannot see.

Numerical interpolation struggles in the same way. Many

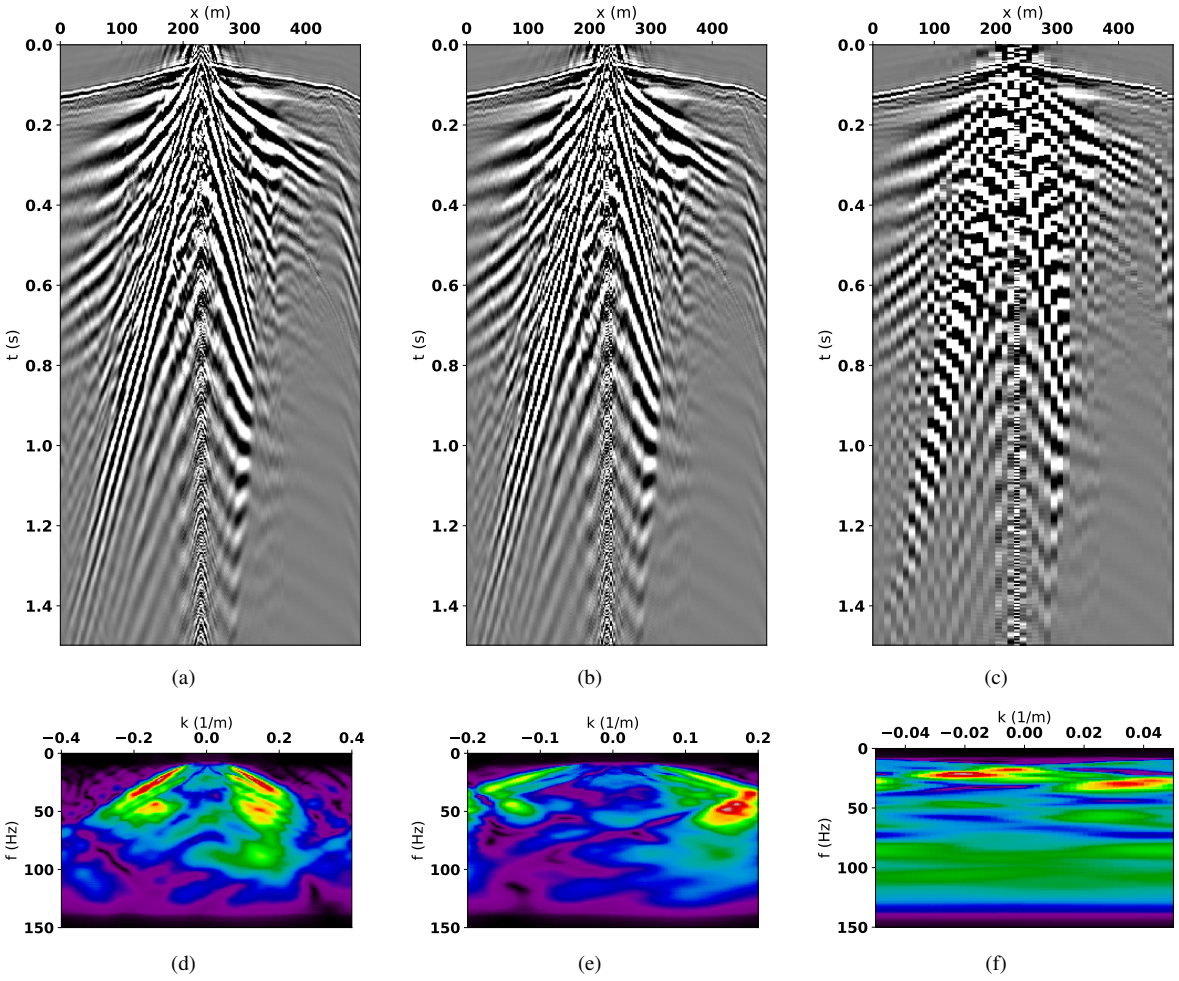


Figure 1. (a)-(c): Land data sampled at 1.25m, 2.5m and 10m trace interval, and (d)-(f): corresponding frequency spectra. Note that aliasing occurs even at 2.5m sampling interval due to the slow surface waves. Data sampled at 10m are difficult to interpret.

algorithms can only be applied to small data windows or to amplitude processed data because the transform domain representation they use is strongly affected by large dynamic range. Consider for example a plane wave of constant amplitude (Figure 2(a)). The Fourier representation of this object is also a line with just a few non-zero coefficients (Figure 2(c)). However, if one introduces an offset-dependent amplitude decay on the order of $1/r$, where r is the offset, the spectral representation changes: a large region of non-zero coefficients surrounds the previously sparse line (Figures 2(b) and 2(d)). Plane waves with decaying amplitude do not have sparse representations in the frequency domain, causing attempts at signal recovery to fail if the algorithm relies on sparsity. In the case of wavelets, large wavelet coefficients correspond to strong events, enabling much better recovery of signals with decaying amplitudes. We use wavelet-domain data recovery schemes to overcome the dynamic range problem without the necessity of amplitude pre-processing or data windowing. This approach enables interpolation of raw land seismic data and typically aliased surface waves, which in turn has the potential to solve some of the key near surface challenges (Keho and Kelamis, 2012).

3 THEORY

3.1 Wavelet transform

The continuous wavelet transform can be thought of as a generalization of a windowed Fourier transform. Wavelets are localized in time and frequency, and their support in time and frequency depends on the scale. In contrast, the time-frequency window in the local Fourier transform has a constant size. For a 1D signal $f(t)$, the continuous wavelet transform is defined by:

$$Wf(u,s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \psi^* \Big(\frac{t-u}{s} \Big) dt, \tag{1}$$

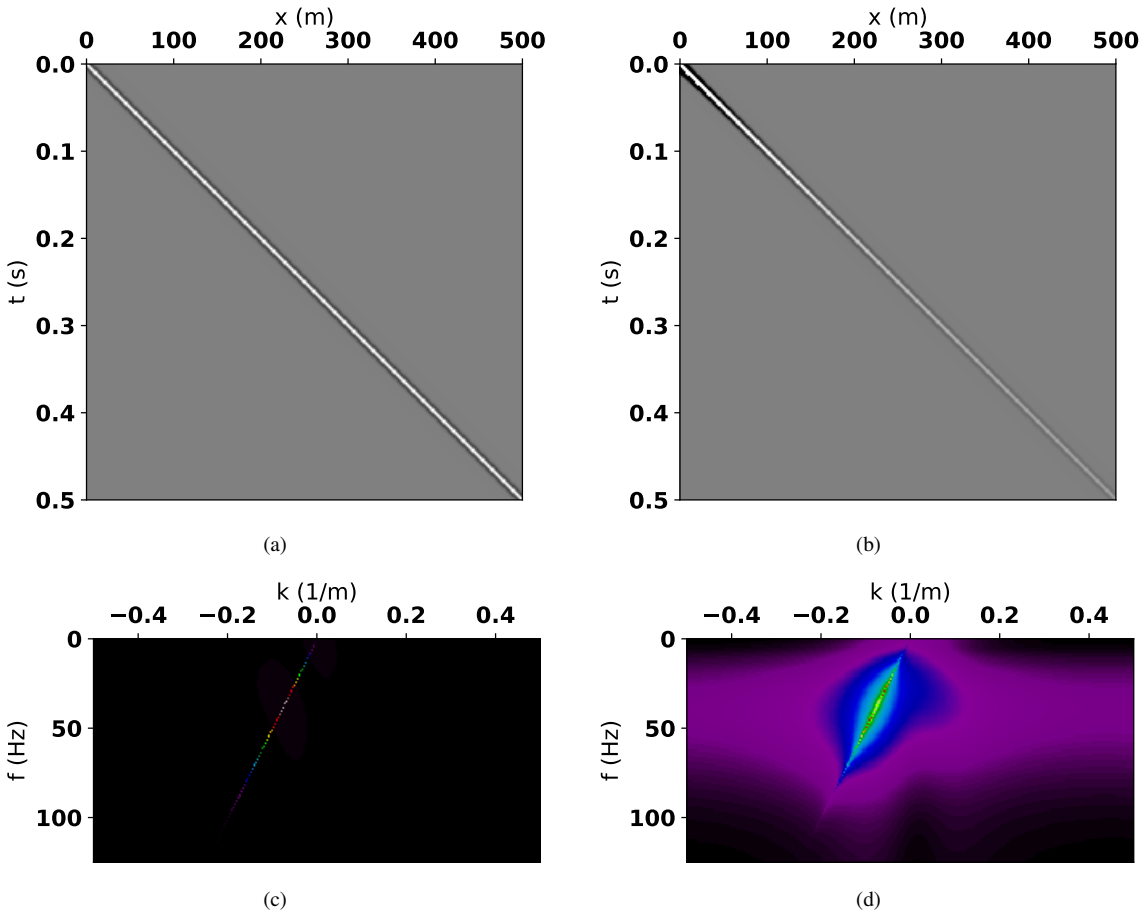


Figure 2. Band limited line: (a) of constant amplitude, (b) with amplitude decay proportional to $1/r$. (c), (d): Fourier domain representations of (a) and (b). f denotes normalized frequency and k denotes normalized wavenumber.

where u denotes translation, s is scale and ψ^* is complex conjugate of wavelet function (Mallat, 1999). The wavelet $\psi_{u,s}$ has a time support centered at u and proportional to s. The continuous wavelet transform conserves energy if the wavelet admissability condition is satisfied:

$$C_\psi = \int_0^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < +\infty, \tag{2}$$

where $\hat{\psi}(\omega)$ denotes the Fourier transform of the wavelet function and ω is angular frequency. The admissibility condition is met if the wavelet is zero mean (i.e. $\hat{\psi}(0) = 0$) and $\hat{\psi}(\omega)$ is continuously differentiable, which is ensured if $\psi(t)$ decays sufficiently fast in the time domain: $\int_{-\infty}^{+\infty}(1+|t|)|\psi(t)|dt < \infty$. When the admissability condition is satisfied, the continuous wavelet transform is invertible:

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} Wf(u,s) \frac{1}{\sqrt{s}} \psi\Big(\frac{t-u}{s}\Big) du \frac{ds}{s^2}. \tag{3}$$

The continuous wavelet transform represents 1D signals by highly redundant images in (u, s). A discrete wavelet transform reduces this redundancy by sampling the continuous wavelet transform in such a way that a frame is obtained. A set of vectors $\{\psi_n\}_{n\in\Gamma}$ forms a frame for an inner product space H if there exist two constants A and B such that $0 < A \leq B < \infty$ and

$$A\|f\|_H^2 \leq \sum_{n\in\Gamma} |\langle f, \psi_n \rangle|^2 \leq B\|f\|_H^2 \tag{4}$$

for all $f \in H$. One has to sample the time-frequency plane in such a way that the entire plane is covered. In practice, this is achieved by sampling scale s along an exponential sequence $\{a^j\}_{j\in\mathbb{Z}}, a > 1$. The sampling of time translation u is uniform with resolution proportional to $1/s$. Thus, the basis vectors of the wavelet frame are of the form:

$$\psi_{j,n}(t) = \frac{1}{\sqrt{a^j}} \psi\Big(\frac{t - nu_0 a^j}{a^j}\Big) \tag{5}$$

In the filter bank implementation of the wavelet transform, the total number of operations to compute the wavelet coefficients for an $N-$length discrete 1D signal is $O(N)$ (Mallat, 1999); less than the Fast Fourier transform, which requires $O(N \log_2 N)$ operations.

It is useful to think about wavelet transform as a mul-

tiresolution approximation formed by a scaling function $\phi(t)$. Suppose that the wavelet coefficients are known only for the scales $s < s_0$. A scaling function contains information that corresponds to wavelet coefficients for scales $s > s_0$ (Mallat, 1999). Consider an approximation of a signal f in subspace V_j and a subspace V_{j-1}, corresponding to the finer scale. The approximation in V_{j-1} contains all information from approximation in V_j and some additional details. Thus, we can write that:

$$V_{j-1} = V_j \oplus W_j, \tag{6}$$

where $\oplus$ denotes union of orthogonal subspaces and W_j is the orthogonal complement of V_j in V_{j-1} containing details present in V_{j-1} but not in V_j. The scaling functions ϕ_j form the basis for subspace V_j, while wavelet functions ψ_j span the W_j space. Figure 3 illustrates wavelet decomposition on a real seismic trace. The signal is decomposed into 4 scales. The finest scale subspace, V_{-4}, can be written as:

$$V_{-4} = V_{-3} \oplus W_{-3} = V_0 \oplus W_0 \oplus W_{-1} \oplus W_{-2} \oplus W_{-3}, \tag{7}$$

meaning that the signal, which is its own approximation at scale -4, can be decomposed into the low resolution approximation at scale 0 (computed as an inner product of the signal with scaling function at scale 0) and a set of details at scales -1, -2 and -3 (computed as inner products of the signal with wavelet functions at scales -1, -2 and -3). Note that large wavelet coefficients are concentrated at the coarsest scale approximation, which is effectively a low-frequency portion of the signal. Due to the fast decay of wavelet coefficients, the seismic signal can be considered compressible in the wavelet domain, enabling the use of compressive sensing techniques for signal recovery.

Wavelet transform can be extended to higher dimensions. Let ϕ be a scaling function and ψ a wavelet which yields an orthogonal basis of $\mathbf{L^2}(\mathbb{R})$. The basis for $p-$dimensional multiresolution approximation is formed by 2^p basis vectors as an outer product between scaling and wavelet functions. In 2D case, the subspace V_j^2 is spanned by an outer product of a scaling functions along two dimensions: $\phi^2 = \phi_1 \otimes \phi_2$ and the detail subspaces W_j^2 are formed by $\psi^1 = \phi \otimes \psi$, $\psi^2 = \psi \otimes \phi$, and $\psi^2 = \psi \otimes \psi$. The ψ^1 is sensitive to horizontal features in the signal, ψ^2 to vertical features, and ψ^3 to diagonal features. In 3D, the detail subspace W_j^3 has sensitivity in seven directions, as illustrated in Figure 4.

3.2 Sparse recovery conditions

Consider an $N-$dimensional signal f which can be represented as a vector of coefficients α in some basis expansion: $f = \Phi\alpha$. f is said to be sparse if only $K \ll N$ of the coefficients α are non-zero. f is compressible when sorted coefficients α decay rapidly enough to zero, so that α can be well-approximated as sparse (Baraniuk et al., 2010).

Successful recovery of $K-$sparse or compressible signal depends on three key components:

- the sampling strategy,
- the sparsifying transform,
- the sparsity-promotion recovery.

In seismic data acquisition, we acquire $d = Tm$, where d are recorded data, T is a sampling matrix and m is the full data volume needed for processing and inversion. The only freedom that can be taken in the field is where receivers are placed and when shots are fired. Thus, for seismic applications, the matrix T is sparse.

Results from compressive sensing suggest that sparse signals can be recovered without loss of information if the sampling matrix satisfies the so called *restricted isometry property* (RIP) (Baraniuk, 2007). RIP is satisfied with high probability for Gaussian and sub-Gaussian matrices or when sampling non-uniformly Fourier-sparse signals. Then, the number of measurements to recover K-sparse signal is $M = O(K \log(N/K))$. However, this results may not hold for non-uniform sampling in other domains.

Let $m = \Phi^T\alpha$ and $d = T\Phi^T\alpha$, where α represents the signal in the sparse domain. In this scenario, the RIP is satisfied if Φ is a Fourier transform and the sparsity promoting recovery can be achieved by solving the following ℓ_1 optimization problem:

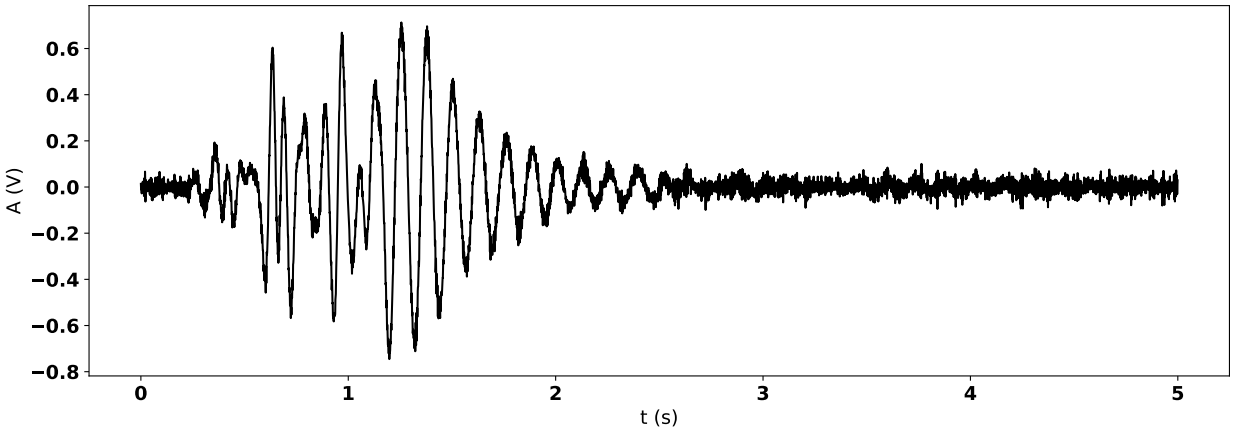
$$\tilde{\alpha} = \arg\min_{\alpha} \|\alpha\|_1 \text{ subject to } d = T\Phi^T\alpha. \tag{8}$$

However, for other transforms (such as the wavelet transform or the curvelet transform), there is no practical algorithm to compute RIP constants (Herrmann, 2010) and provide similar recovery guarantees. A way to gain an insight into recovery behavior of a given sampling matrix - transform pair, one can construct a phase transition diagram. For different realizations of a $K-$sparse signal, one computes the chance of successful recovery by looking at the difference between original and recovered signal. Figure 5 shows such a diagram for 1D wavelet domain $K-$sparse signal with a binary sampling matrix (data are a subset of true wavelet coefficients). Note that for successful recovery in this scenario, one needs a relatively high number of measurements.

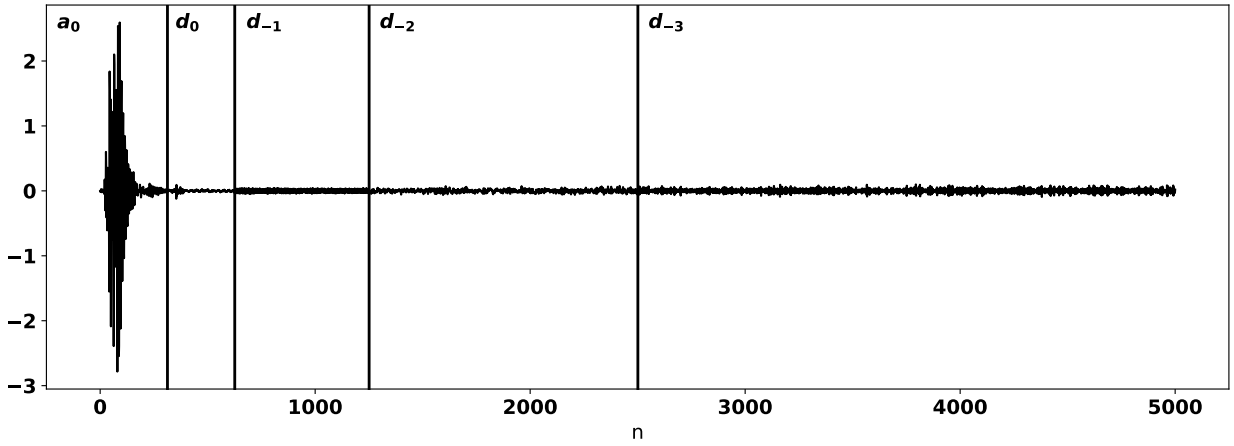
3.3 Sampling patterns in the wavelet domain

Hennenfent and Herrmann (2008) present a detailed analysis of sampling patterns in the frequency domain and show that signals can be recovered well when sampling artifacts, known as spectral leakage, do not have a sparse representation in the frequency domain. Regular undersampling results in aliases at particular frequencies. In contrast, irregular (random) downsampling results in white noise-like spectral leakage. Therefore, random downsampling yields much more favorable recovery conditions. The question is, however, whether these results hold for other Fourier-related domains?

Figure 6 shows uniform and random downsampling of the trace from Figure 3(a) and the resulting wavelet coefficients. Note that signal aliases appear at scales -2 and -3 for the uniform downsampling, whereas for random downsampling, noisy images appear at all scales. This is a consequence of the



(a)



(b)

Figure 3. (a) Field seismic trace and (b) its wavelet transform.

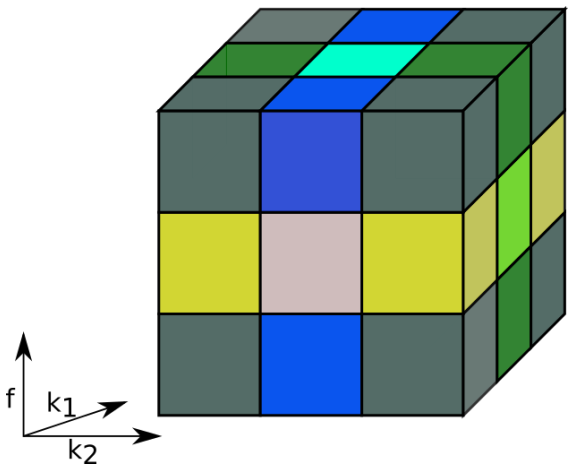


Figure 4. Partitioning of frequency cubes in 3D wavelet transform. The axes origin is at the center of the cube. Different colors correspond to detail subspaces. The low frequency approximation, i.e. the cube at the center, is not visible.

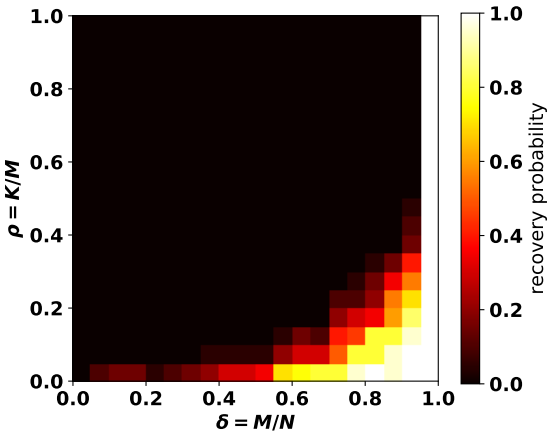


Figure 5. Phase transition diagram for sparse sampling matrix and wavelet domain sparse data. N is the signal length, K is the sparsity, and M is the number of measurements.

implementation of the digital wavelet transform with a filter bank, as described in Mallat (1999).

3.4 Projection onto convex sets

Another method of data reconstruction, which does not rely on sparsity explicitly, is projection onto convex sets (POCS). It this method, the data to be recovered are assumed to have some known properties (such as being band-limited). We also assume that:

- all signals are members of a Hilbert space $\mathcal{H}$,
- signals with desired properties lie in subsets of $\mathcal{H}$,
- subsets describing properties are *closed linear manifolds*.

Let $\mathcal{P}$ be a closed linear manifold in a Hilbert space that represents the desired property and let P be orthogonal projection operator onto $\mathcal{P}$. Suppose the recorded singnal $f \in \mathcal{H}$. Then the signal with desired properties g is the unique projection $Pf = g \in \mathcal{P}$ that is closest to f in some appropriate norm (Moon and Stirling, 2000). We solve an optimization problem of the form

$$\min_{g \in \mathcal{P}} \mathcal{J} = \|f - g\|, \tag{9}$$

by the method of alternating projections (Moon and Stirling, 2000).

Abma and Kabir (2006) introduce the application of POCS to seismic data reconstruction by alternating projections in the $F-K$ domain. Their idea is to transform data with gaps to the $F-K$ domain, project back with only a few biggest coefficients, re-inject original data and repeat the process, while progressively lowering the threshold applied to the coeficients in the transform domain. The challenge lies in deciding on a thresholding strategy. Starting from just a few highest coefficients may lead to undesirable results for signals with large dynamic range (recall Figure 2). Therefore, the choice of transform and knowledge about signal properties in the transform domain is of key importance for the success of this strategy.

We show that POCS can successfuly recover missing data with large dynamic range if the wavelet transform is used instead of the $F-K$ transform. Wavelets reconstruct large-dynamic range signals starting from the high-energy portion of signals; progressive lowering of the threshold allows to accurately reconstruct weaker portions of signals. We find that logarithmic thresholding, starting from 0.1% of largest coefficients and relaxing the threshold values to 20% of the highest coefficients yields satisfactory results. To achieve similar results in the $F-K$ domain, windowing or amplitude processing would be required.

4 DATA RECONSTRUCTION EXAMPLES

We demonstrate data reconstruction for data with large dynamic range using two wavelet domain approaches: sparsity promoting ℓ_1 optimization problem described in equation 8 and the POCS approach described in equation 9. In the first case, we use the Spectral Projected Gradient solver (Van Den Berg and Friedlander, 2008). For POCS, we follow the scheme proposed by Abma and Kabir (2006), while replacing the Fourier transform with a wavelet transform. We use logarithmic thresholding applied to all scales except the coarse scale, which we leave unchanged. For both synthetic and field data examples, we use the same wavelet transform parameters (coiflet10 wavelet and two levels of decomposition) and percent thresholds.

We design our experiments for data reconstruction from compressive seismic acquisition, following the jittered undersampling strategy described by Hennenfent and Herrmann (2008). Both synthetic and field data example showcase data with large dynamic range, strong and slow surface waves, and weaker seismic signals. The input for the reconstruction algorithms are unwindowed, unprocessed data.

To quantify the quality of the reconstruction, we use

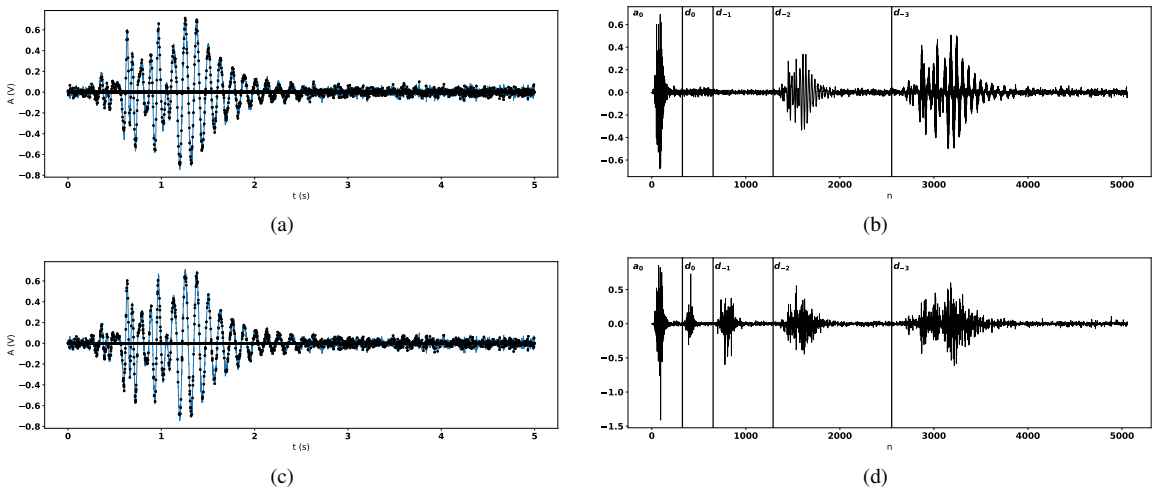


Figure 6. Different undersampling strategies and their imprint in the wavelet domain. (a) Regular downsampling by a factor of 4, (c) random downsampling by the same factor. (b), (d): Wavelet coefficients for strategies shown in (a), (c), respectively.

signal-to-noise ratio, defined as:

$$SNR = 20 \log_{10} \Big(\frac{\|x\|}{\|x - \hat{x}\|} \Big), \tag{10}$$

where x represents original full data and $\hat{x}$ represents reconstructed data. We also look at the coherence between full and recovered data, computed separately for every trace as a measure of similarity between original and reconstructed data, and the difference image.

4.1 Synthetic data example

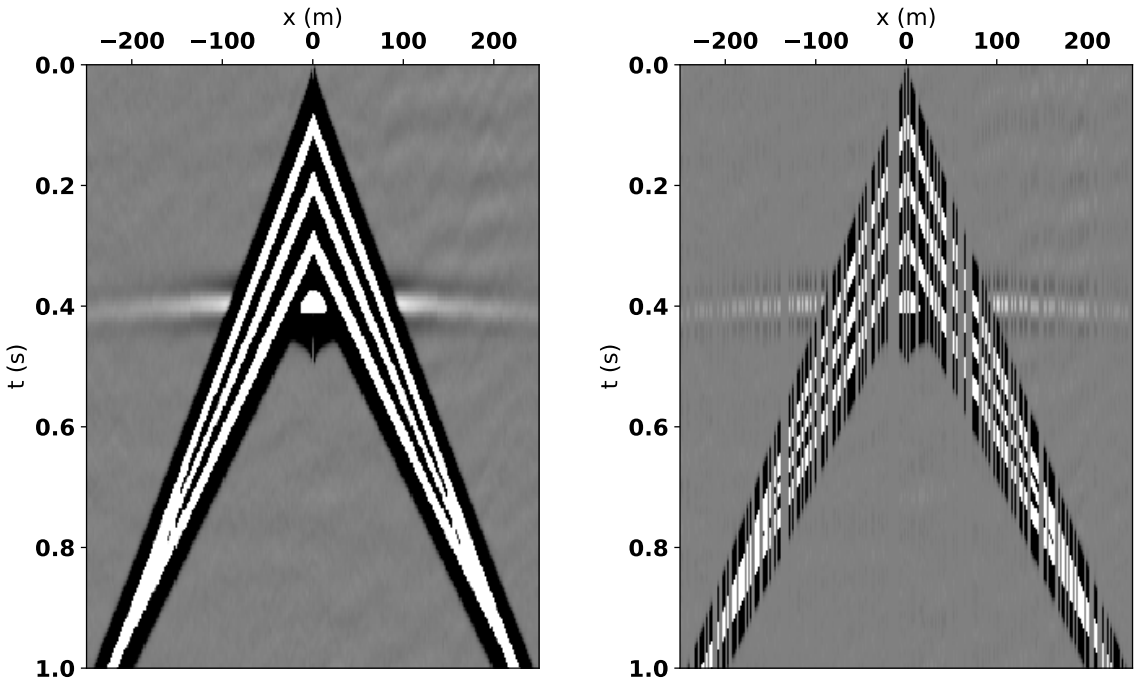
Figure 7(a) shows a synthetic shot containing strong surface waves, a reflection and background noise. Slow surface waves require fine receiver sampling. The ambient seismic noise at fine sampling is correlated and coherent. We use portions of passive records sampled finely to increase the realism of this example. The peak amplitude of the surface wave train is 9 times stronger than the peak reflection amplitude. The ambient noise is 100 times weaker than the peak amplitude reflection, but does not decay with offset.

The sparsity promoting approach (Figure 8(a)) preserves the general trends of the data: the visual similarity between the original and the recovered data is high, except at the center and at the edges, where wavelet approximation is suboptimal due to the limited directionality of wavelets. The coherence between original and recovered data is close to 1 for all traces and all frequencies. However, the difference image shows that the biggest discrepancies occur at sharp contrast boundaries. This is explained by analyzing the wavelet coefficient spectrum of original and recovered data (Figure 9). Some detail information is missing in the recovered data, especially at the finest scale: the resulting wavelet spectrum is sparser than the spectrum of the original data. The POCS approach yields fewer reconstruction artifacts, but we note that there are some visible edge effects, similar to the ℓ_1 reconstruction. The difference image in Figure 8(d) reveals that more details are preserved by the POCS reconstruction than by sparse reconstruction. This is also reflected in the signal-to-noise ratios for the two reconstructions: SNR is 7.1dB for POCS, whereas it is 4.8dB for the sparsity promoting approach. The comparison of wavelet spectra in Figure 9 reveals that the wavelet spectrum of POCS reconstructed data is closer to the spectrum of the original full data than the wavelet spectrum of sparsity promoting reconstruction. The differences in wavelet spectra translate into the missing details in reconstructed data, visible in the difference images.

We repeat the experiment with the same synthetic shot for downsampling by a factor of 4. Figure 10 shows the reconstructed volumes, revealing that 75% of missing data is too much for the sparsity promoting approach, as one could expect from the phase transition plot in Figure 5. However, the POCS approach is able to reconstruct the original data shape, though the quality of that reconstruction is poorer than in the previous example, as revealed by the difference image. A glance at wavelet spectra in Figure 11 demonstrates that the ℓ_1 minimization does not recover the coarsest scale coefficients correctly. As those coefficients correspond to the low frequency information and are of key importance, the low quality reconstruction is not surprising. In contrast, in our implementation of POCS, we do not change the coarsest scale coefficients from the downsampled data. This choice preserves the low frequency character of the data and allows to rebuild some of the higher frequencies, though not as well as in the example with only 50% missing data.

4.2 Field data example

The field data example uses an unprocessed shot from a dense 2D land survey conducted by Colorado School of Mines in Pagosa Springs, Colorado. Regular downsampling by a factor of 2 would introduce aliasing, so instead we follow the jittered



(a) (b)

Figure 7. Synthetic data used for (a) reconstruction tests and (b) the data with 50% missing traces.

undersampling approach of Hennenfent and Herrmann (2008). Figure 12(a) shows the selected shot; the data are clipped for display but reconstruction is run on the raw shot, without windowing or amplitude balancing.

The reconstructed data are shown in Figure 13. The most noticeable differences between full and reconstructed data are in the immediate vicinity of the source. The signal-to-noise ratio is higher for ℓ_1 recovery at 7.0dB compared to 4.3dB for POCS, which is opposite to the synthetic data example. Examining the wavelet spectra in Figure 14 shows that ℓ_1 reconstruction yields a sparser spectrum than that of the original data, while POCS reconstruction increases the strength of some coefficients compared to the original data. This observation can explain the higher SNR for sparse reconstruction. The trace coherence in both cases is close to 1. Looking at the difference images reveals that reconstruction mostly distorts the high amplitude jumps. Similarly to the synthetic data example, this phenomenon is related to the wavelet spectra of reconstructed data and to the limited directional sensitivity of the 2D wavelet transform. Furthermore, the field shot record contains many more seismic events than the synthetic example, and thus it is less sparse in the wavelet domain, which also affects the quality.

5 DISCUSSION

The performance of wavelet-based data reconstruction depends on the wavelet choice. There are many wavelet families to choose from. Some of the popular wavelet families include Daubechies, symlets, biorthogonal wavelets and coiflets. We consider Daubechies (Daubechies, 1988) and coiflets (Daubechies, 1993). Daubechies wavelets are designed to have the most compact support for a given number of vanishing moments: more vanishing moments enable an approximation of wider class of piecewise-smooth signals. Coiflets are designed so that the scaling function has a prescribed number of vanishing moments. In general, a good choice is a wavelet that resembles the source signature, if known. Another consideration is the support length of the wavelet. If the chosen wavelet has support comparable to the size of the data gaps, the reconstruction performance deteriorates.

One has two additional choices to make regarding the wavelet transform. First is the number of decomposition levels, whose maximum depends on the size of data along the transform axis and on the support of the wavelet. Choosing the maximum possible number of decomposition levels allows to better distinguish features in data and is a good rule of thumb. A second choice is the signal extension mode for computing the transform. To obtain N coefficients from an $N-$length signal, N has to divisible by 2^L, where L is number of decomposition levels and the whole signal needs periodic padding while computing the wavelet transform.

We show that the sparsity promoting approach requires more data for successful recovery than the POCS approach. As seen in the phase transition diagram (Figure 5), the number of data needed for good recovery is quite high even for very sparse data. Note that we created the phase transition dia-

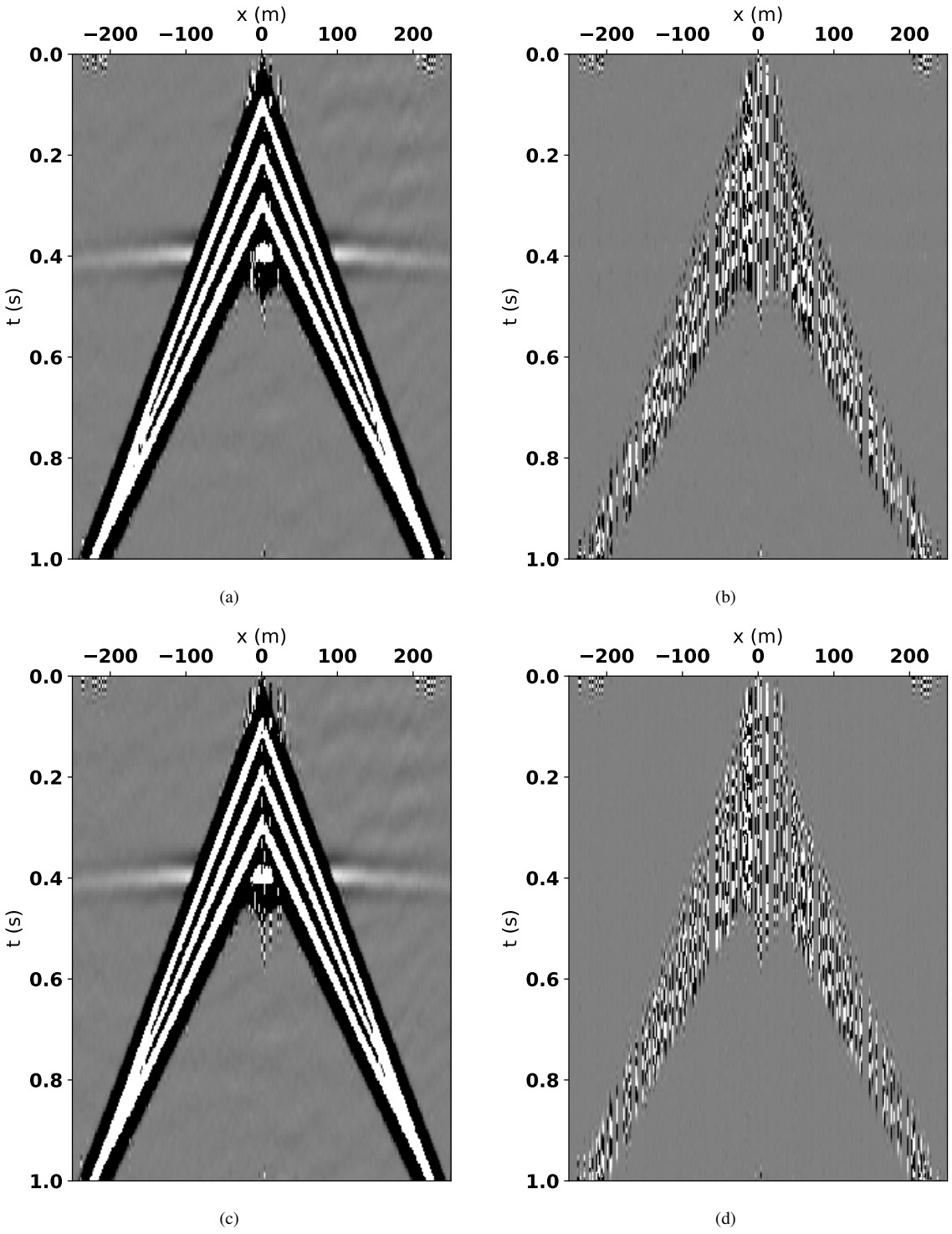
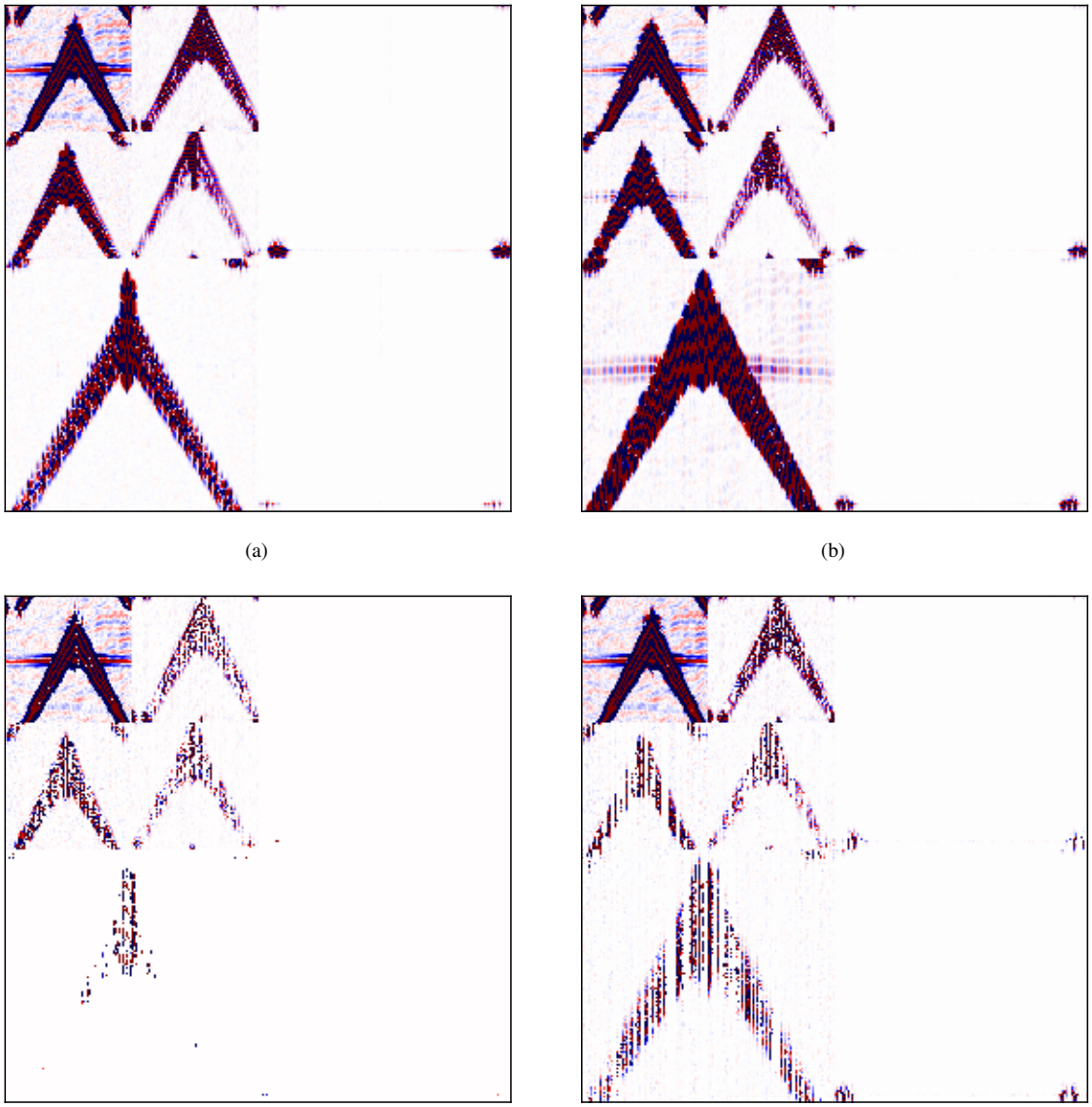


Figure 8. Synthetic data reconstruction results for (a) ℓ_1 optimization and (c) POCS. (b) and (d) are the corresponding data differences with respect to the full data in Figure 7(a). Note that the outline of surface waves is more visible on the difference image for the sparsity promoting approach, hence its lower SNR.



(a) (b)

(c) (d)

Figure 9. Wavelet spectra of (a) full synthetic data, (b) downsampled synthetic data, synthetic data reconstructed with (c) ℓ_1 optimization and (d) with POCS. Note that random gaps in data introduce additional wavelet coefficients of alternating sign (e.g., the coefficients corresponding to reflection at bottom left of (b)). Sparsity promoting reconstruction recovers much fewer coefficients compared to the full data, especially at the finest scale. POCS is able to recover more fine detail coefficients, but the visible discontinuities result in missed details in reconstruction.

gram for a 1D example and a Daubechies 8 wavelet. A similar diagram for a different wavelet and in 2D might look differently, especially if one uses a randomized sampling with gap size control. Increasing the support length of a wavelet and introducing gap size control increases the chance of sampling non-zero wavelet coefficients. If one randomly selects a subset of coefficients from signal of length N, only K of which are non-zero, there is a risk of only picking zeros, which does not bring the needed new information about the reconstructed signal. Looking at signals in higher dimensions can result in a sparser representation, depending on how fast the signal character changes in a given direction. For seismic signals, temporal variations are more rapid than spatial variations, hence 2D reconstruction has a better chance of success than the Figure 5 would imply.

Acquisition of dense unaliased data on land is an expensive undertaking. Our data wavelet-based reconstruction results are a first step towards designing acquisition geometry which uses fewer sensors and allows recovery of full data without the need for amplitude processing or window-

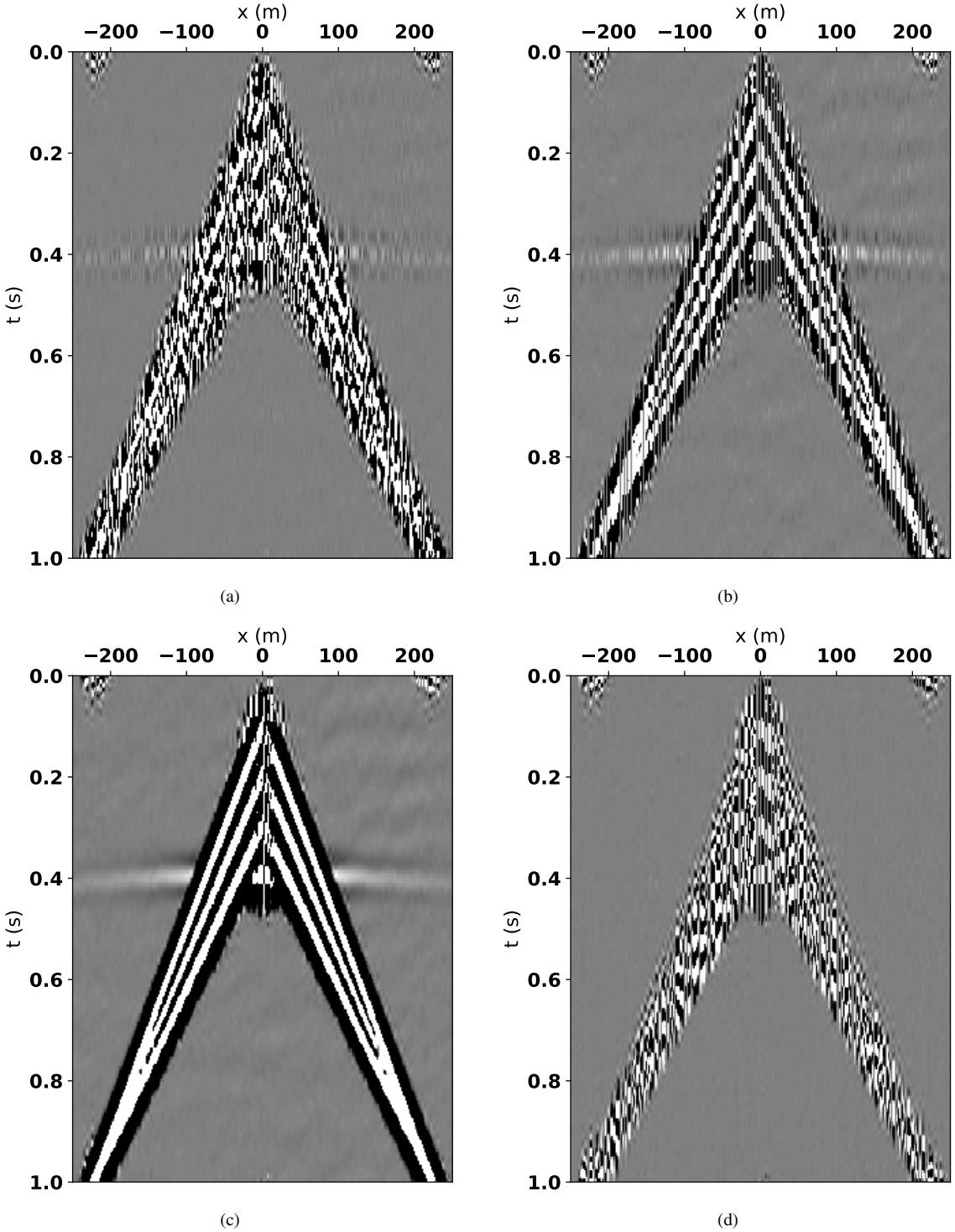


Figure 10. Synthetic data reconstruction results after undersampling by a factor of 4 for (a) ℓ_1 optimization and (c) POCS. (b) and (d) are the corresponding data differences with respect to the full data in Figure 7(a). Note that POCS is able to restore the shape of original full data, while sparse reconstruction fails at that task.

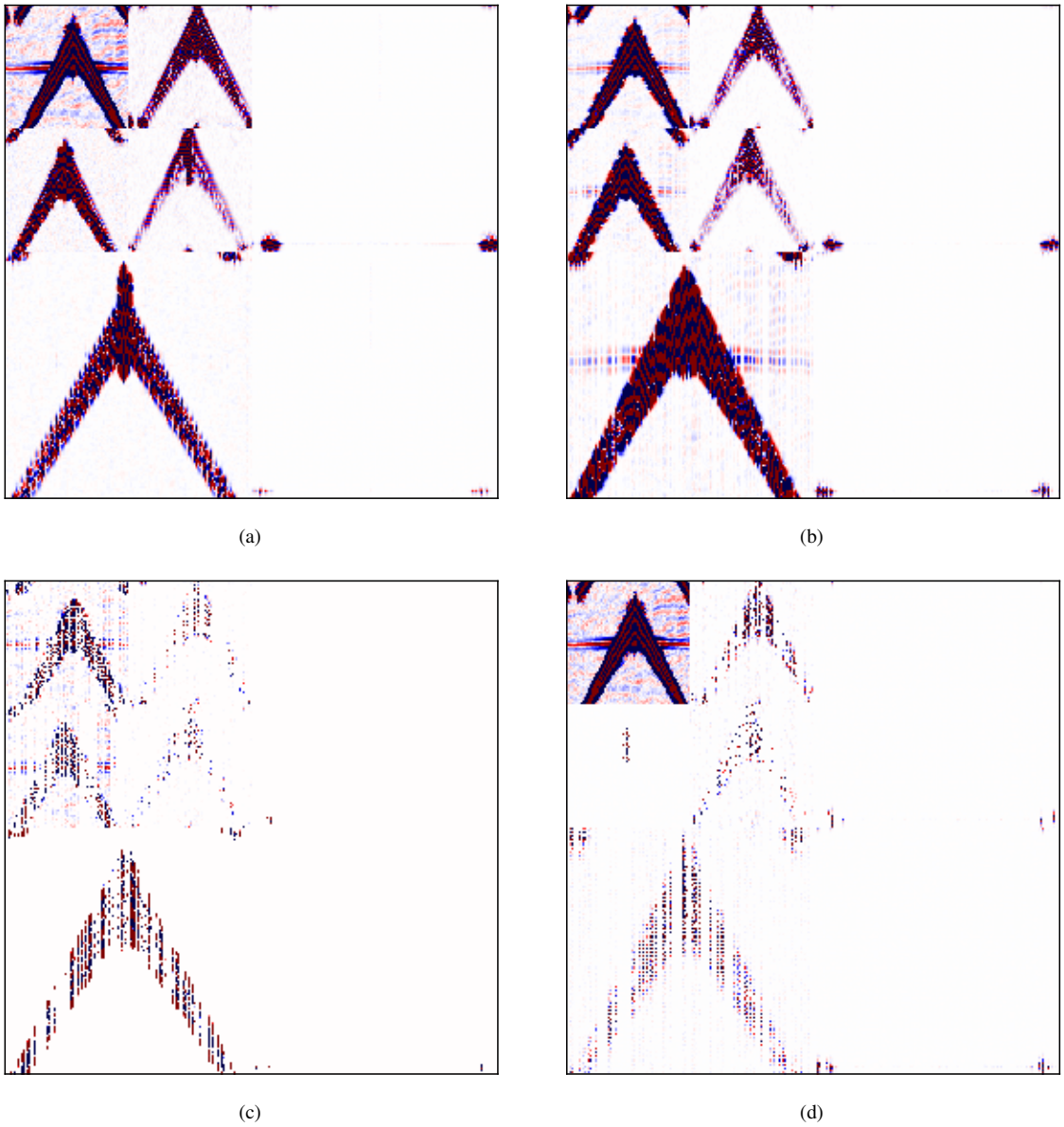


Figure 11. Wavelet spectra of (a) full synthetic data, (b) synthetic data downsampled by factor of 4, synthetic data reconstructed with (c) ℓ_1 optimization and (d) with POCS. Sparsity promoting optimization fails to reconstruct wavelet coefficients at the coarsest scale, while POCS preserves low frequency information.

ing. Non-aliased surface waves can provide a wealth of information about the shallow subsurface, potentially helping to solve near-surface challenges, such as building shallow velocity models and correctly accounting for strong shallow heterogeneity (Keho and Kelamis, 2012). However, further refinements in data reconstruction are needed to achieve even higher reduction in channel count: the average receiver spacing in our scenario is 2.5m, while practical average spacing would be on the order of 10m or more.

One promising future direction for wavelet-based data reconstruction is an extension to 5D. As explained in previous paragraphs, seismic data do not vary in space as fast as in time, resulting in sparser representations requiring fewer measurements for good reconstruction. 5D interpolation is demonstrated to work well with windows of data in the Fourier domain. Using wavelets would facilitate the analysis of larger subsets, limited only by the available computer memory.

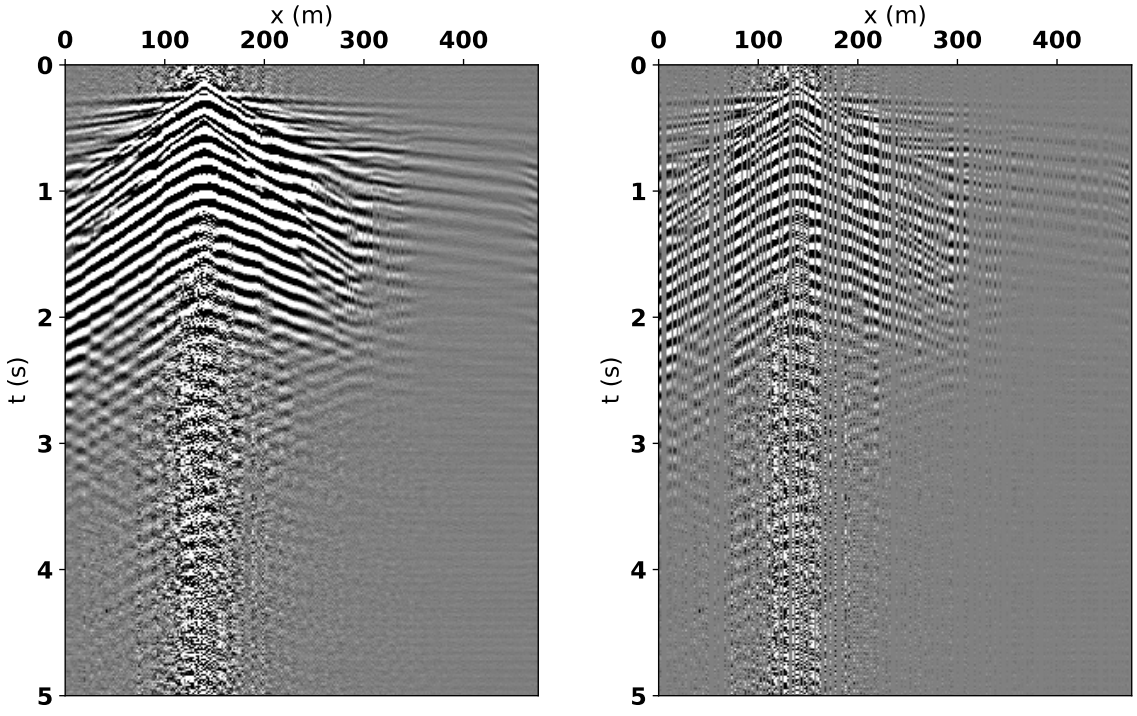


Figure 12. (a) Field data used in reconstruction tests, and (b) the data with 50% missing traces.

6 CONCLUSIONS

We demonstrate that both ℓ_1 sparsity promoting optimization and POCS can be used to recover seismic data with large dynamic range in the wavelet domain. The main advantage of using the wavelet domain is that it allows to avoid data windowing and the computations are faster than for the Fast Fourier transform. ℓ_1 recovery is more sensitive to the number of missing data. As we show with a synthetic example, POCS can handle downsampling by a factor of 4 quite well, but sparsity promotion fails.

Our reconstruction strategies are kinematically correct for strong and weak events alike, including ambient noise, as seen in the synthetic example. The relative amplitude ratio is also preserved. Observable differences between true and recovered data are primarily confined to the high amplitude events. This allows to use reconstructed data to analyze and remove surface wave, without jeopardizing much weaker reflection signals.

The wavelet transform in 2D is limited to three orientations (horizontal, vertical and diagonal), and therefore, representation of edges at different angles is suboptimal, and thus the sparsity promoting approach misses detail. In that regard curvelets have a distinct advantage if one can afford to store all coefficients of the transformation.

7 ACKNOWLEDGEMENTS

We would like to acknowledge Dawson and Sercel for providing equipment and technical support at the CSM Geophysical Field Camp. We extend our thanks to GTI for allowing us to test wireless nodes and conduct dense sampling experiments. SEG provided partial financial support of the Field Camp. We also acknowledge the CSM undergraduate students who conducted the field data acquisition.

REFERENCES

Abma, R., and N. Kabir, 2006, 3D interpolation of irregular data with a POCS algorithm: Geophysics, **71**, E91–E97.

Baraniuk, R. G., 2007, Compressive sensing [lecture notes]: IEEE Signal Processing Magazine, **24**, 118–121.

Baraniuk, R. G., V. Cevher, M. F. Duarte, and C. Hegde, 2010, Model-based compressive sensing: IEEE Transactions on Information Theory, **56**, 1982–2001.

Candès, E. J., and L. Demanet, 2005, The curvelet representation of wave propagators is optimally sparse: Communications on Pure and Applied Mathematics, **58**, 1472–1528.

Candès, E. J., et al., 2006a, Compressive sampling: Proceedings of the international congress of mathematicians, Madrid, Spain, 1433–1452.

Candès, E. J., J. Romberg, and T. Tao, 2006b, Robust uncertainty principles: exact signal reconstruction from highly in-

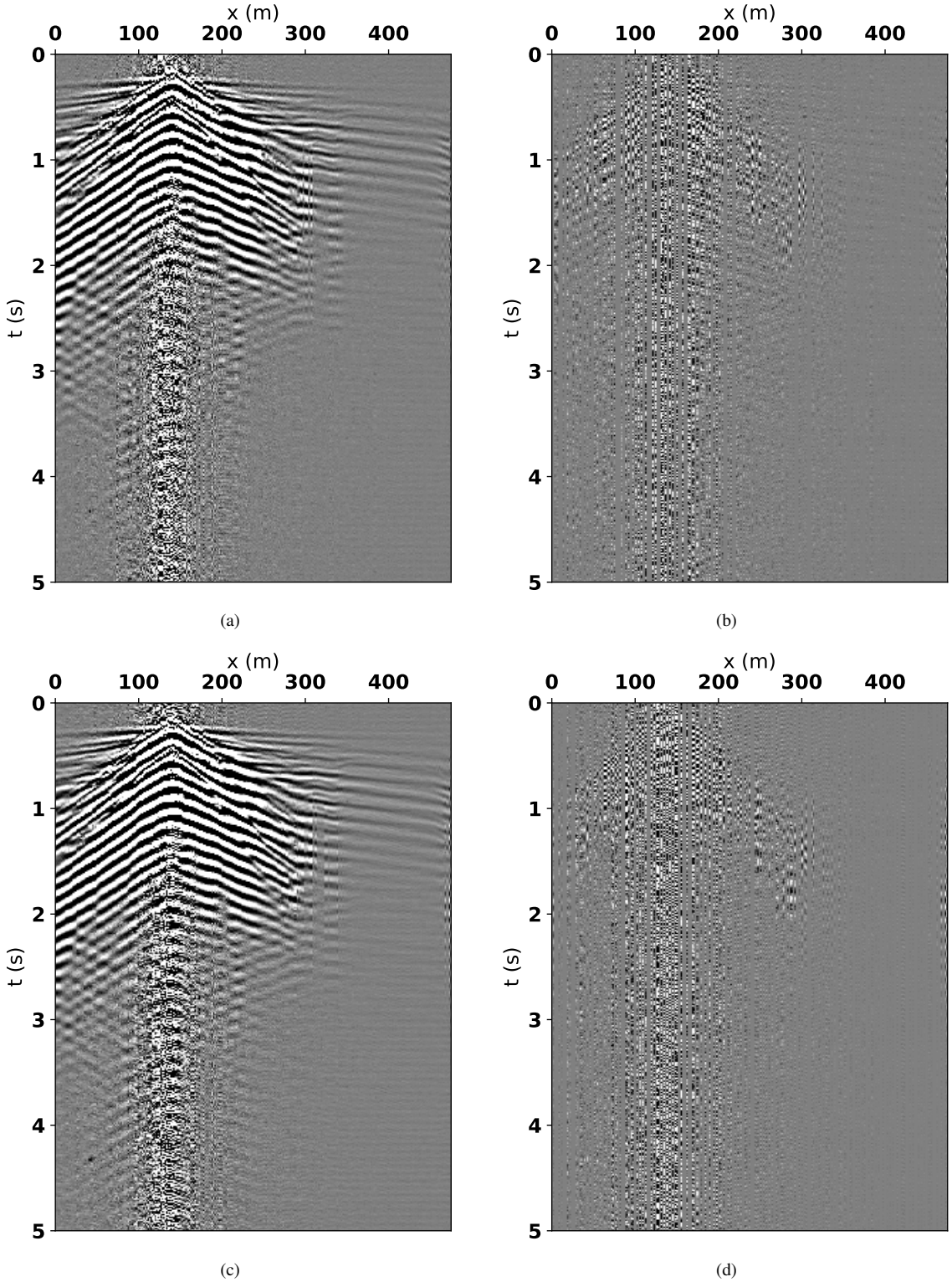
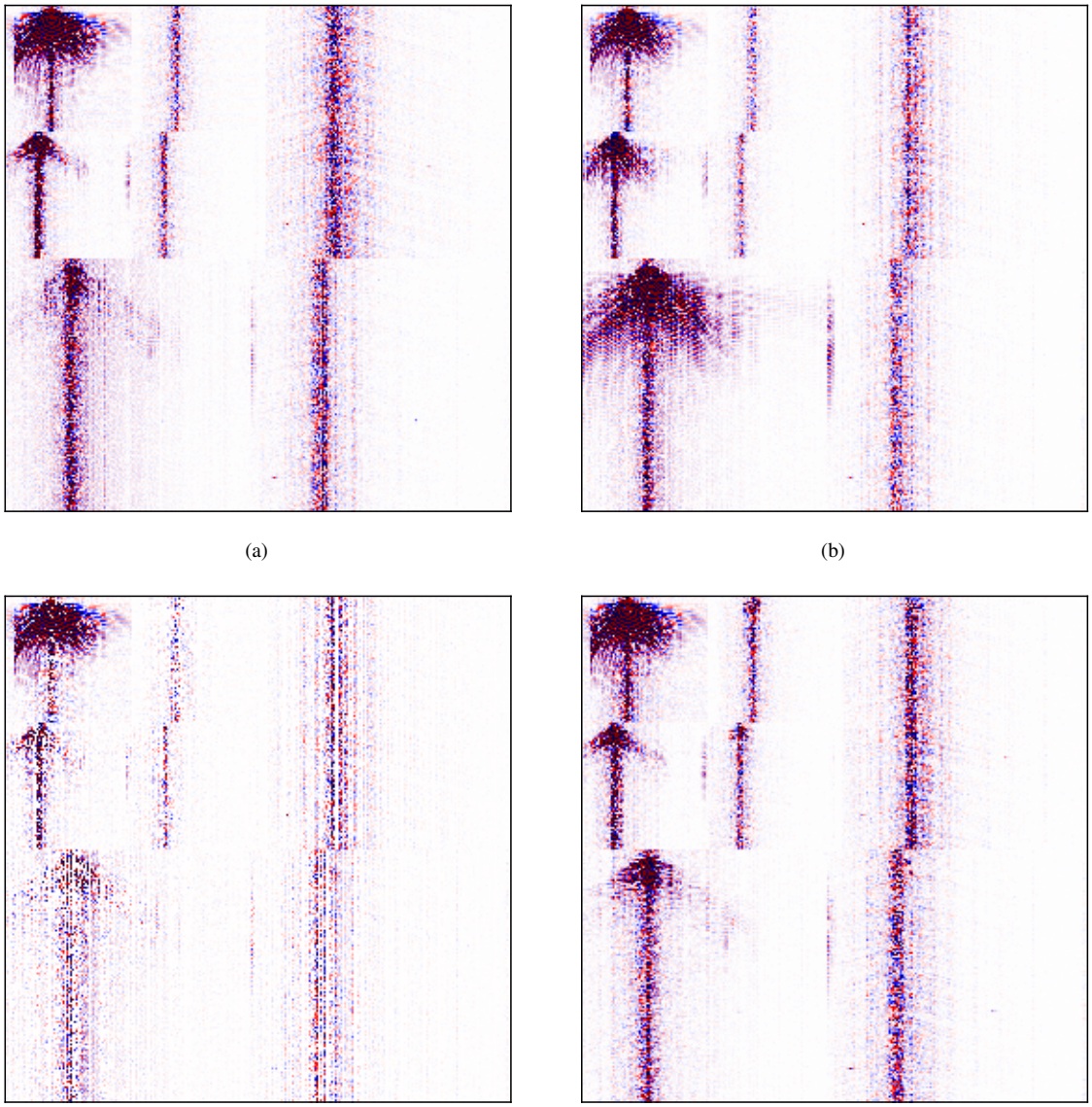


Figure 13. Field data reconstruction results for (a) ℓ_1 optimization and (c) POCS. (b) and (d) are the corresponding data differences with respect to the full data in Figure 12(a). Both reconstructions struggle in the similar way with high frequency details, but the sparsity promoting approach has higher SNR.



(a) (b)

(c) (d)

Figure 14. Wavelet spectra of (a) the full field data, (b) downsampled data, data reconstructed with (c) ℓ_1 optimization and (d) with POCS. Note that ℓ_1 reconstruction yields a sparser wavelet spectrum than POCS.

complete frequency information: IEEE Transactions on Information Theory, **52**, 489–509.

Chen, Y., D. Zhang, S. Zu, S. Gan, W. Huang, X. Chen, and Z. Jin, 2016, Simultaneous denoising and reconstruction of 5-D seismic data via damped rank-reduction method: Geophysical Journal International, **206**, 1695–1717.

Daubechies, I., 1988, Orthonormal bases of compactly supported wavelets: Communications on pure and applied mathematics, **41**, 909–996.

———, 1993, Orthonormal bases of compactly supported wavelets ii. variations on a theme: SIAM Journal on Mathematical Analysis, **24**, 499–519.

Gan, S., S. Wang, Y. Chen, Y. Zhang, and Z. Jin, 2015, Dealiased seismic data interpolation using seislet transform with low-frequency constraint: IEEE Geoscience and remote sensing letters, **12**, 2150–2154.

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

Herrmann, F. J., 2010, Randomized sampling and sparsity:

Getting more information from fewer samples: Geophysics, **75**, WB173–WB187.

Kabir, M. N., and D. Verschuur, 1995, Restoration of missing offsets by parabolic Radon transform: Geophysical Prospecting, **43**, 347–368.

Keho, T. H., and P. G. Kelamis, 2012, Focus on land seismic technology: The near-surface challenge: The Leading Edge, **31**, 62–68.

Kreimer, N., A. Stanton, and M. D. Sacchi, 2013, Tensor completion based on nuclear norm minimization for 5D seismic data reconstruction: Geophysics, **78**, V273–V284.

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

Mallat, S., 1999, A wavelet tour of signal processing: Elsevier.

Moon, T. K., and W. C. Stirling, 2000, Mathematical methods and algorithms for signal processing: Prentice Hall Upper Saddle River, NJ, **1**.

Mosher, C. C., C. Li, F. D. Janiszewski, L. S. Williams, T. C. Carey, and Y. Ji, 2017, Operational deployment of compressive sensing systems for seismic data acquisition: The Leading Edge, **36**, 661–669.

Naghizadeh, M., and M. D. Sacchi, 2008, F-X adaptive seismic-trace interpolation: Geophysics, **74**, V9–V16.

———, 2010, Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data: Geophysics, **75**, WB189–WB202.

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

Trad, D., 2009, Five-dimensional interpolation: Recovering from acquisition constraints: Geophysics, **74**, V123–V132.

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.

Wang, B., N. Zhang, W. Lu, and J. Wang, 2019, Deep-learning-based seismic data interpolation: a preliminary result: Geophysics, **84**, V11–V20.

Wang, J., M. Ng, and M. Perz, 2010, Seismic data interpolation by greedy local Radon transform: Geophysics, **75**, WB225–WB234.

Xu, S., Y. Zhang, and G. Lambaré, 2010, Antileakage Fourier transform for seismic data regularization in higher dimensions: Geophysics, **75**, WB113–WB120.

Xu, S., Y. Zhang, D. Pham, and G. Lambaré, 2005, Antileakage Fourier transform for seismic data regularization: Geophysics, **70**, V87–V95.

Yu, Z., J. Ferguson, G. McMechan, and P. Anno, 2007, Wavelet-Radon domain dealiasing and interpolation of seismic data: Geophysics, **72**, V41–V49.