Efficient land seismic data reconstruction from sparse measurements

Iga Pawelec*, Michael Wakin† & Paul Sava*

* *Center for Wave Phenomena and Dept. of Geophysics, Colorado School of Mines, Golden CO 80401*
email ipawelec@mines.edu
† *Dept. of Electrical Engineering, Colorado School of Mines, Golden CO 80401*

ABSTRACT

Acquisition of high-quality land seismic data requires (expensive) dense source and receiver geometry to avoid aliasing-related problems. Alternatively, acquisition using the concept of compressive sensing (CS) allows for similarly high quality land seismic data using fewer measurements provided that the designed geometry and sparse recovery strategy are well matched. We propose a complex wavelet-based sparsity-promoting wavefield reconstruction strategy to overcome challenges in land seismic data interpolation using the compressive sensing framework. Despite having lower angular sensitivity than curvelets, complex wavelets enable the improved reconstruction of sparsely acquired land data while being faster and requiring less storage. The complex wavelet transform operates outside of the Fourier domain, which localizes aliasing-related artifacts likely to be present in field data, and yields reconstructions with fewer artifacts and higher signal-to-noise ratios. We demonstrate that the data recovery success depends on both the number and the geometry of the missing traces as revealed by analyzing reconstructions from multiple realizations of trace geometry and data decimation ratios. Using half the number of traces required by the regular sampling rules and thus reducing the acquisition costs, we show that data are appropriately reconstructed provided that there are no big gaps in the strategic places.

Key words: compressive sensing, interpolation, sparsity, wavelets, curvelets

1 INTRODUCTION

Land seismic data are notorious for being particularly challenging to handle. The challenges can be separated into two categories: acquisition and processing. On the acquisition side, the main challenge is that poor coupling between receivers and the medium (sand, soil or ice) significantly reduces the signal-to-noise ratio of the recorded signal. There are also legal and operational access restrictions that can result in regions with no data coverage. On the processing side, challenges are associated with various types of noise (though in some applications, what we term here as noise can be used to derive information about the subsurface). The ambient noise background is typically higher on land than in a marine environment due to the proximity to different cultural sources of noise (e.g., industry, traffic). Even remote areas can present unique noise challenges such as ice breaking in the Arctic (Li et al., 2017). Furthermore, the commonly used Vibroseis source generates both the desired sweep and other signals (harmonics of the sweep, sound generated by the engine) (Polom, 1997; Denisov et al., 2019), which need to be attenuated either during acquisition or processing. Finally, one of the most important causes of land data quality degradation is the highly heterogeneous and generally poorly consolidated near surface. This portion of the subsurface, called the low-velocity layer (Sheriff, 2002), gives rise to strong and dispersive surface waves and scattering noise.

Due to these challenges, land seismic data quality tends to be low, which can adversely impact reservoir characterization. In the case of unconventional reservoirs, the value of acquiring seismic data in the first place is sometimes questioned. However, even in these scenarios, seismic data can provide a wealth of information useful for characterizing unconventional plays, such as anisotropy maps, fracture orientation maps, and pressure distribution. The land seismic acquisition goal is to increase the reliability of seismic-derived products, which depend critically on having dense and high signal-to-noise ratio seismic data as an input.

While we are on the verge of a revolution in terms of the number of geophones deployable in a cost-effective manner (Manning et al., 2018), some sites with particularly slow surface waves require ultra-fine sampling (on the order of 1m), and certain areas are inaccessible to both sources and receivers. To overcome these economic and logistical obstacles, one can explore alternative ways of acquiring data that could yield substantial savings on the acquisition efforts. Mosher et al. (2017) show that a novel approach, using ideas from compressive sensing, can significantly speed-up acquisition without jeopardizing the data quality. The underlying idea is to randomize receiver placement and shot timing according to compressive sampling methodologies and to subsequently solve a large-scale regularization problem, recovering non-aliased data from sparse measurements. Most conventional 3D acquisition geometries have regular, but coarse sampling in at least one direction (Trad, 2009) and sometimes large data gaps due to access restrictions. In contrast, compressive sensing surveys have deliberate irregular sampling and often simultaneous shooting, account for geometry restrictions, and are tuned for the data recovery strategy.

Commonly, data recovery strategies rely on the presence of known, repetitive patterns in the data and/or data sparsity in some representation. Current techniques for infilling missing data can be divided in several categories, including prediction error filters (PEFs) (Spitz, 1991; Crawley et al., 1999; Fomel and Claerbout, 2003; Naghizadeh and Sacchi, 2009; Claerbout, 2014), matrix or tensor completion (Kreimer and Sacchi, 2012; Kreimer et al., 2013; Ma, 2013; Kumar et al., 2015), rank reduction (Trickett et al., 2010; Gao et al., 2013; Chen et al., 2016) and machine learning (Jia and Ma, 2017; Pilikos and Faul, 2017; Jia et al., 2018; Wang et al., 2019). However, the most widely used interpolation techniques are transform-based approaches that can take advantage of known data characteristics in the transform domain. Such methods are well-studied in the context of data aliasing and irregular sampling, and rely on data representation in a known 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; Hollander and Yilmaz, 2019) and wavelet or seislet transforms (Yu et al., 2007; Gan et al., 2015; Liu et al., 2017), the Fourier transform remains the most popular choice because it is easy to implement and fast to compute, as long as data are regularly sampled. Liu and Sacchi (2004) develop a framework for data recovery based on L_2 norm minimization, using spectral weights bootstrapped from FK representation of data that 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 can be 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 and angles, which can be prohibitively expensive for large 3D datasets. Furthermore, the choice of number of scales and angles in the curvelet transform is non-intuitive and strongly data-dependent.

In this paper, we discuss the challenges inherent in seismic data reconstruction - such as significant temporal or spatial changes in signal amplitude and frequency content or data aliasing - and propose how to overcome them by exploiting the complex wavelet domain (Kingsbury, 2001; Selesnick et al., 2005). We demonstrate that complex wavelets outperform the curvelets in the challenging field data reconstruction examples while being faster to compute and requiring less memory. We also emphasize the key role that gap pattern plays for successful data recovery.

The paper is organized as follows. First, we explore the specific challenges inherent in seismic data reconstruction with an emphasis on land seismic data. We then discuss the conditions for sparse data recovery and review several transforms available for sparse representation of seismic data. This review is followed by a description of our data recovery approach and field data examples to compare the performance of a wavelet transform (both real and complex) to that of a curvelet transform. Finally, we share key insights from our analysis, including computational advantages of using the complex wavelet transform for seismic data reconstruction, the relevance of data decimation ratio with respect to the Nyquist wavenumber, and the impact of the gap patterns on successful data recovery.

2 CHALLENGES IN DATA RECONSTRUCTION

Every seismic survey is different, but there are some common characteristics of land seismic data which make them particularly difficult to reconstruct: the presence of data aliasing, the pattern of missing traces, the size of data gaps, and the large dynamic range. Land seismic data tend to be more problematic than marine seismic data due to the highly complex heterogeneous near surface, which traps the majority of energy released by the seismic source and produces slowly propagating surface waves (Keho

and Kelamis, 2012). In the following, we discuss the challenging features of land seismic data in more detail and explain how different transforms handle them.

2.1 Aliasing

In land seismic data, aliasing of surface waves can be especially severe due to the much slower velocities of the 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 their frequency-wavenumber spectra. Note that the energy corresponding to the surface waves is present between 10-60 Hz and the phase velocity is about 200 m/s. The spatial sampling needed to acquire non-aliased data would be

$$\Delta x \leq \frac{v_{\min}}{2f_{\max}}, \tag{1}$$

where $v_{\min}$ is the slowest surface-wave velocity and $f_{\max}$ is the maximum frequency in the data. We can conclude that the needed sampling interval for recording the data shown in Figure 1 non-aliased is $\Delta x \leq 1.66\mathrm{m}$, which is much finer than would typically be acquired for a large-scale exploration project.

High-fidelity recording of surface waves is beneficial for near-surface characterization (Foti et al., 2014). For example, one could use surface-wave analysis to build a velocity model for a low-velocity layer (Socco et al., 2010) or to compute static corrections (Papadopoulou et al., 2020), both of which can lead to an enhanced image of the subsurface. Furthermore, non-aliased surface waves are much easier to remove from shot record for subsequent reflection processing (Manning et al., 2018). Although the surface-wave velocity is site-dependent and the case presented in Figure 1 has unusually low velocities, it clearly demonstrates that to obtain non-aliased representation of surface waves one needs to either drastically increase the sampling on the receiver side or to develop a strategy to reconstruct the non-aliased wavefield from reduced measurements.

Reconstruction of an aliased wavefield poses a serious challenge. One possible approach uses prediction error filters (PEF) (Spitz, 1991) by exploiting the idea that filter coefficients derived from low, non-aliased frequencies can be used to interpolate aliased data components (Naghizadeh and Sacchi, 2009). Naghizadeh and Sacchi (2010) also use non-aliased 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 (Fomel and Liu, 2010). 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 some land seismic data can be severe, suggesting that altering the data reconstruction approach would be a better solution (Baraniuk and Steeghs, 2017).

2.2 Gap pattern

Historically, seismic data have been acquired on a regular grid or are 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., 2006b) which dictates a sampling rate of at least two points per wavelength for successful recovery of a non-aliased signal. The number of sensors needed to record good quality, non-aliased land data on a regular grid is exceedingly high for slowly propagating waves. The advent of compressive sensing (CS) (Candès et al., 2006c) opened up 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 to regularly sampled, dense grid acquisition. Mosher et al. (2017) demonstrate that compressive sensing can be successfully applied to field seismic acquisition. The success of compressive sensing depends on finding the optimal gap pattern combined with sparse signal representation in a known transform domain.

2.3 Dynamic range and Fourier domain representation

Although visually seismic data do not look more complex than many natural images (i.e., photographs of real objects), the key differences lie in the dynamic range and representation of seismic data in the Fourier domain. Dynamic range can be defined as the ratio between the largest and smallest values that a given signal can assume. Figure 2 shows a land shot record with absolute values of amplitudes displayed on the logarithmic scale. Note that the coherent events which could be reliably labeled as seismic signals easily span five orders of magnitude. With the exception of high dynamic range images (HDR) which are stored as floating point numbers (i.e., 32 bits per color channel), natural images tend to have low dynamic range with the fixed number of possible values

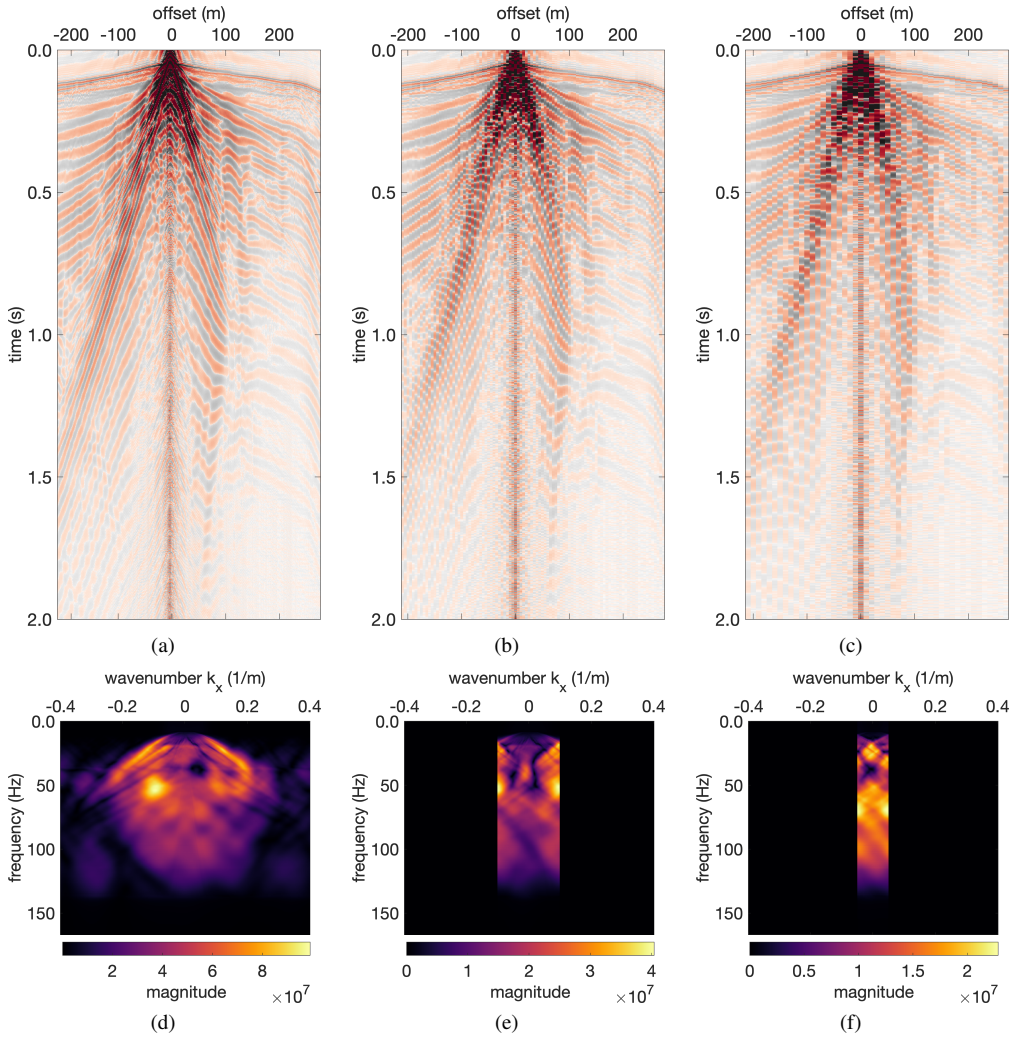


Figure 1. (a)-(c): Land data sampled at 1.25m, 5m and 10m trace interval, and (d)-(f): corresponding frequency spectra. Data amplitudes are gained for display. Note that aliasing occurs even at 5m sampling interval due to the slow surface waves. Data sampled at 10m are difficult to interpret.

per each color channel: 256 for 8-bit images and 65536 for a 16-bit images. The bigger the dynamic range, the more challenging it is to find a sparse signal representation which does not compromise low data amplitudes. Thus, the many advances in image compression and reconstruction are not immediately applicable to seismic data. One can partially bypass that problem by sorting seismic data into common offset gathers (smaller lateral amplitude variation) or by applying reversible gain functions (to preserve the relative amplitudes). Both operations may reduce the dynamic range, enhance sparsity in some transform domains, and yield a better reconstruction of missing data. However, as we later demonstrate, sparsity alone does not guarantee a successful signal recovery.

Another challenge in seismic data reconstruction is the Fourier-domain data representation. Natural images have comparable sampling rates in all image dimensions. Seismic data are typically oversampled along the time axis and undersampled along the spatial axes. The differences in sampling lead to differences in the Fourier-domain representation of these two types of data and consequently differences in how transforms relying on sampling of the Fourier space see them. For instance, wavelets and curvelets are sensitive to angular sampling (curvelets more so than wavelets). However, while downsampling the seismic data along the time axis can usually be done without the risk of information loss, the same cannot be said about the spatial dimension. As a result, the curvelet and wavelet transforms may behave in non-intuitive ways, with large energy transform coefficients concentrated at unexpected scales and angles and being absent at angles identified in data. Thus, the conventional wisdom about angles represented by the transform applies to data with similar sampling along all dimensions. Since seismic data typically do not meet this requirement,

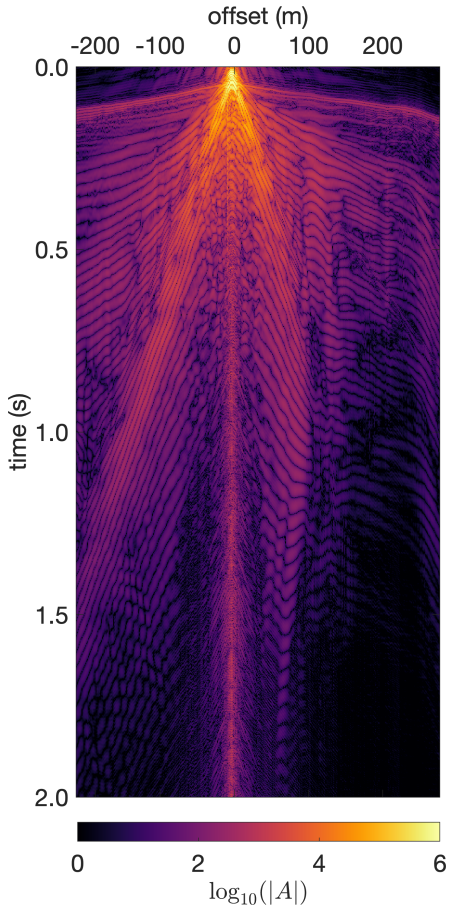


Figure 2. Dynamic range of land seismic data. Note that significant amplitudes span several orders of magnitude and that most of the shot generated energy is trapped in the near source region.

it is beneficial to adjust sampling such that spatial and temporal frequencies occupy a similar portion of the Fourier space in all dimensions. This allows one to use insights derived from the application of curvelet- and wavelet-based natural image reconstruction to improve the process of seismic data interpolation.

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 representations, at the expense of increased storage requirements, which makes curvelets a memory-expensive choice for large datasets.

The complex wavelet transform on the other hand offers a good middle ground between the Fourier and curvelet domains. In a sense, complex wavelets can be viewed as an isotropic version of curvelets (Douma and de Hoop, 2007), though wavelets have more limited directional sensitivity. Complex wavelets provide a multiresolution approximation and in 2D are sensitive to six directions: $\pm 15^\circ, \pm 45^\circ$ and $\pm 75^\circ$. Another advantage of the wavelet transform is its linear computational complexity and only 2^D redundancy for $D-$dimensional signals, thus making wavelets suitable for analysis of large datasets.

Due to their multiscale nature, wavelets are also well-suited for handling non-stationary signals. The large dynamic range of seismic data 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 infill the gaps assuming that observed trends are also present in gaps. However, given a raw land seismic record, such task becomes nearly 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.

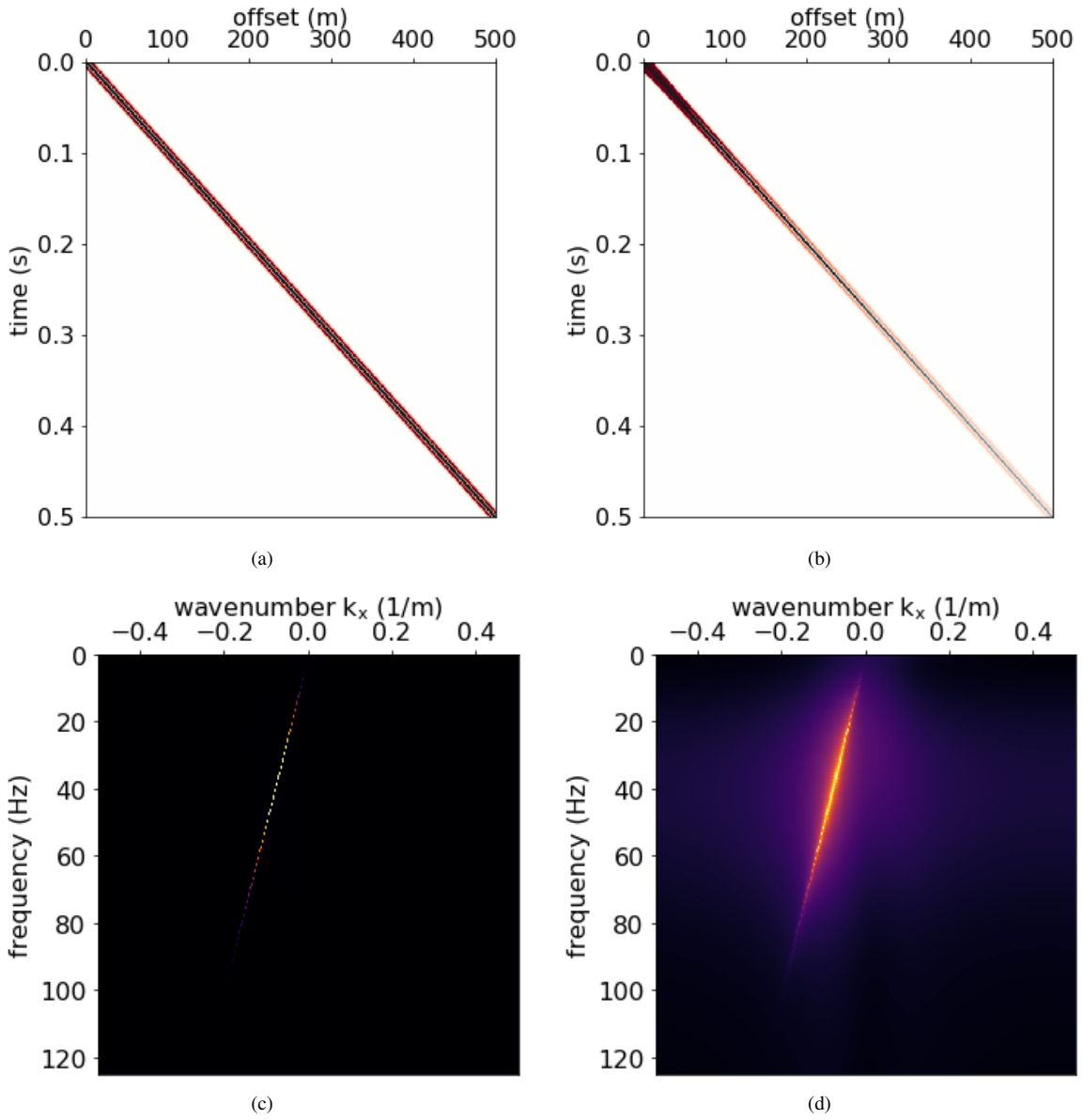


Figure 3. Plane wave: (a) of constant amplitude, (b) with amplitude decay proportional to $1/r$. (c), (d): Fourier domain representations of (a) and (b). Note that Fourier representation of (b) is much less sparse than that of (a).

Numerical interpolation algorithms struggle in the same way. Many approaches can only be applied to small data windows or to amplitude processed data because the transform-domain representation they use is strongly affected by the dynamic range. Consider for example a plane wave of constant amplitude (Figure 3(a)). The Fourier representation of this object is also a line with a few non-zero coefficients (Figure 3(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 3(b) and 3(d)). Plane waves with decaying amplitude do not have sparse frequency-domain representations, causing attempts at signal recovery to fail if the algorithm relies on sparsity. In the case of local transforms such as the wavelet or curvelet transforms, large transform coefficients correspond to strong events, enabling much better recovery of signals with decaying amplitudes. We compare wavelet and curvelet domain data recovery schemes to overcome the dynamic range problem without the necessity of amplitude pre-processing. 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) and improve reservoir characterization.

3 THEORY

3.1 Sparse signal recovery

Consider an $N-$length signal $\mathbf{m}$ that can be represented as a vector of coefficients $\boldsymbol{\alpha}$ in some basis or dictionary expansion: $\mathbf{m} = \boldsymbol{\Phi}^T\boldsymbol{\alpha}$. $\mathbf{m}$ is said to be sparse if only $K \ll P$, where $P \geq N$ is the number of the dictionary coefficients in $\boldsymbol{\alpha}$, are non-zero. $\mathbf{m}$ is compressible when sorted coefficients $\boldsymbol{\alpha}$ decay rapidly enough to zero, so that $\boldsymbol{\alpha}$ can be well-approximated as sparse using a small number of large magnitude coefficients (Baraniuk et al., 2010).

In seismic data acquisition, we acquire $\mathbf{d} = \mathbf{S}\mathbf{m}$, where $\mathbf{d}$ are recorded data, $\mathbf{S}$ is a sampling matrix and $\mathbf{m}$ 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 $\mathbf{S}$ is sparse. Using the techniques from compressive sensing, it is possible to recover the full data volume from the sparse acquisition under certain assumptions.

Successful recovery of $K-$sparse or compressible signal depends on three key components: the sampling strategy, the sparsifying transform, and the sparsity-promoting recovery algorithm. Results from compressive sensing suggest that sparse signals can be recovered without loss of information if the sampling matrix satisfies the restricted isometry property (RIP) (Baraniuk, 2007). RIP is satisfied with high probability for Gaussian (each entry is independent and follows a normal distribution) and random Bernoulli matrices (entries are ± 1 with equal probability) or when sampling non-uniformly Fourier-sparse signals. Depending on the choice of the sampling matrix, the number of measurements to recover K-sparse signal is at $M = \mathcal{O}(K\log(N/K))$, where N is signal length and M is the number of measurements. However, this result may not hold for non-uniform sampling in other domains (for example for non-uniform sampling of a wavefield in conjunction with wavelet or curvelet transform).

Let $\mathbf{m} = \boldsymbol{\Phi}^T\boldsymbol{\alpha}$ and $\mathbf{d} = \mathbf{S}\boldsymbol{\Phi}^T\boldsymbol{\alpha}$, where $\boldsymbol{\alpha}$ represents the signal in the sparse domain. The matrix that ideally satisfies the RIP in this case is $\mathbf{S}\boldsymbol{\Phi}^T$. If $\mathbf{S}$ has sample locations chosen uniformly at random with sufficient number of measurements ($M = \mathcal{O}(K\text{polylog}(N/K))$) and $\boldsymbol{\Phi}$ is a Fourier transform, the RIP is satisfied with high probability and the sparsity promoting recovery can be achieved by solving the following ℓ_1 optimization problem:

$$\tilde{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \text{ subject to } \mathbf{d} = \mathbf{S}\boldsymbol{\Phi}^T\boldsymbol{\alpha}. \tag{2}$$

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.

Finding a good sparsifying transform for seismic data is the key for applying techniques from compressive sensing to infill the missing traces. The sparsifying capabilities of a transform can be quantified by approximating a target signal with K largest transform coefficients and computing the approximation error. In the following, we review a couple of popular choices for a sparse domain and highlight their benefits and pitfalls.

3.2 Fourier transform

The classic D-dimensional Fourier transform is an excellent tool to examine spatio-temporal frequencies present in seismic data. The transform is orthogonal and its computational complexity is $\mathcal{O}(N\log N)$. One advantage of the Fourier-domain representation of seismic data is that it is straightforward to interpret which coefficients represent the majority of coherent seismic energy. However, Fourier coefficients are not localized in time and space, making it challenging to identify what features of data are represented by specific coefficients. Furthermore, the typically high number of non-zero Fourier coefficients necessary to represent the data depends on sampling in time and space. In other words, the Fourier representation of seismic data is rarely sparse. Figure 4 shows the non-linear approximation of a 2D field seismic shot using $\rho = K/P$ fraction of the largest transform coefficients (P represents the number of transform coefficients) in different domains. The approximation error is quantified by signal-to-noise ratio that reflects the power of the original signal to the power of non-linear approximation error. The Fourier domain is the least sparse out of all tested, as discussed next.

3.3 Discrete wavelet transform

The introduction of wavelets in geophysics (Morlet et al., 1982b,a) was motivated by the desire to examine the characteristics of seismic reflection signals (amplitude, shape, frequency and phase) with time. While it is possible to use the windowed Fourier transform for this purpose, the achieved time-frequency resolution is often insufficient to identify subtle signal characteristics due to the fixed window size. The wavelet transform (Mallat, 1989b) achieves improved resolution and time-frequency characteristics of non-stationary signals by varying the window size based on the scale. The continuous wavelet transform can be sampled to obtain the discrete wavelet transform (DWT) which forms an orthonormal basis for a large class of wavelets (e.g., Daubechies, symlets or

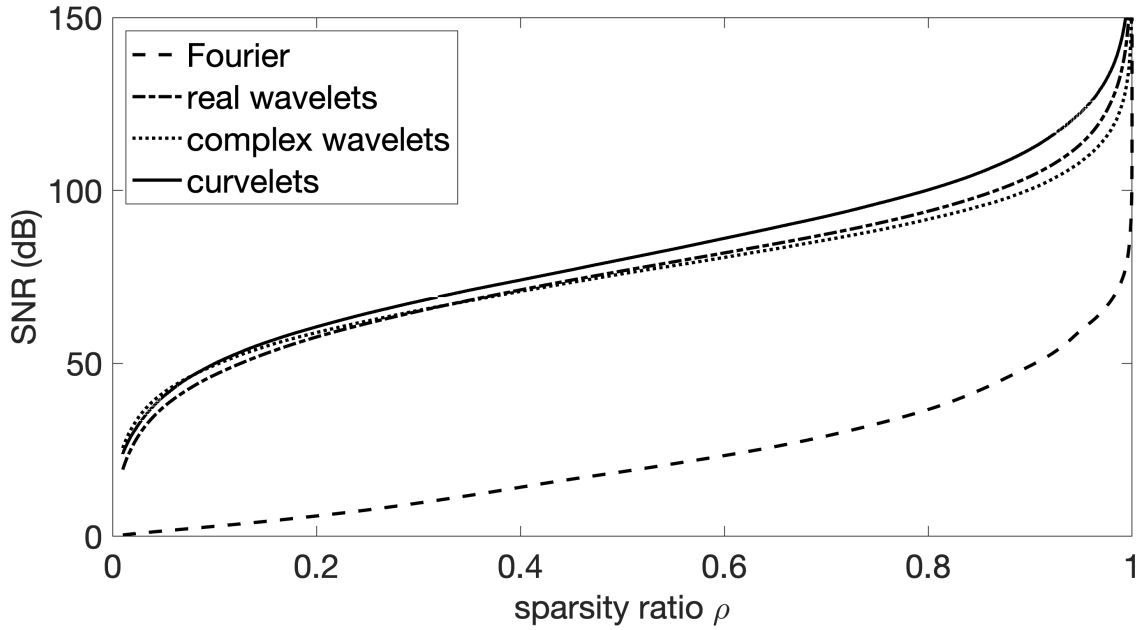


Figure 4. The non-linear approximation error of a seismic field record as a function of sparsity ratio $\rho = K/N$, where N represents the total number of coefficients in the transform domain and K is the number of top magnitude coefficients used in non-linear approximation. The Fourier transform performs the worst while curvelets and wavelets are closely matched.

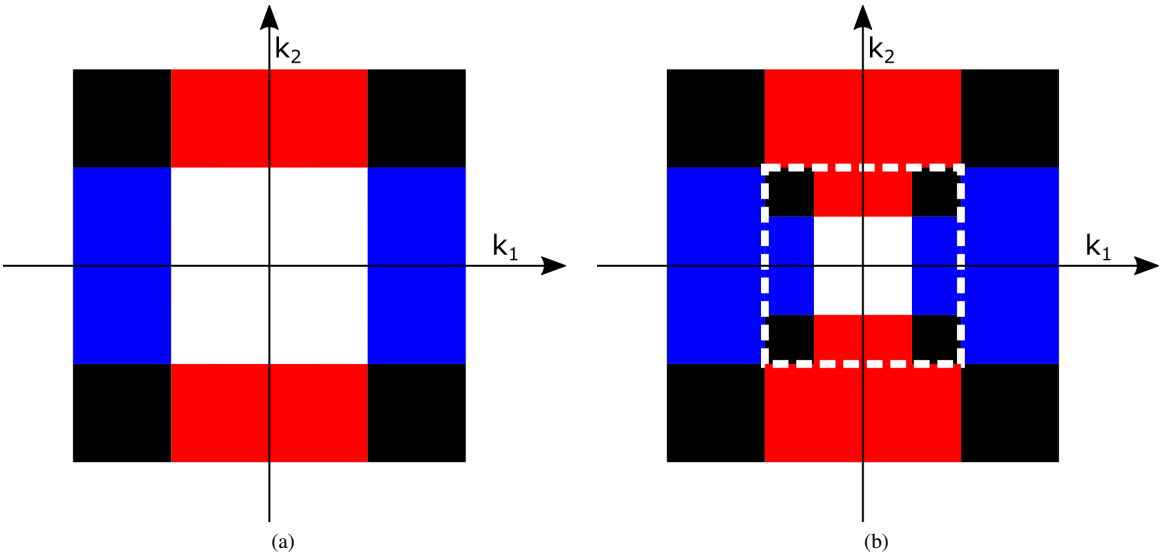


Figure 5. Idealized Fourier domain support for 2D real wavelet transform. (a) is one level decomposition with white box in the middle representing low frequency signal approximation. Blue, red and black represent vertical, horizontal and diagonal details, respectively. Level two wavelet decomposition in (b) results from further decomposing the white box from (a).

coiflets). This feature makes wavelets an attractive choice for multiresolution approximations of signals (Mallat, 1989a), allowing to analyze signal characteristics in different frequency bands with high accuracy and to localize them in space-time domain at the same time.

The multiresolution character of the wavelet transform can be understood by examining the idealized Fourier domain support of wavelet coefficients. Figure 5 shows that support for a 2D signal, but in higher dimensions similar reasoning applies. A first-level wavelet decomposition splits a signal into two parts along each dimension: the low- part and the high-frequency parts. The 2^D

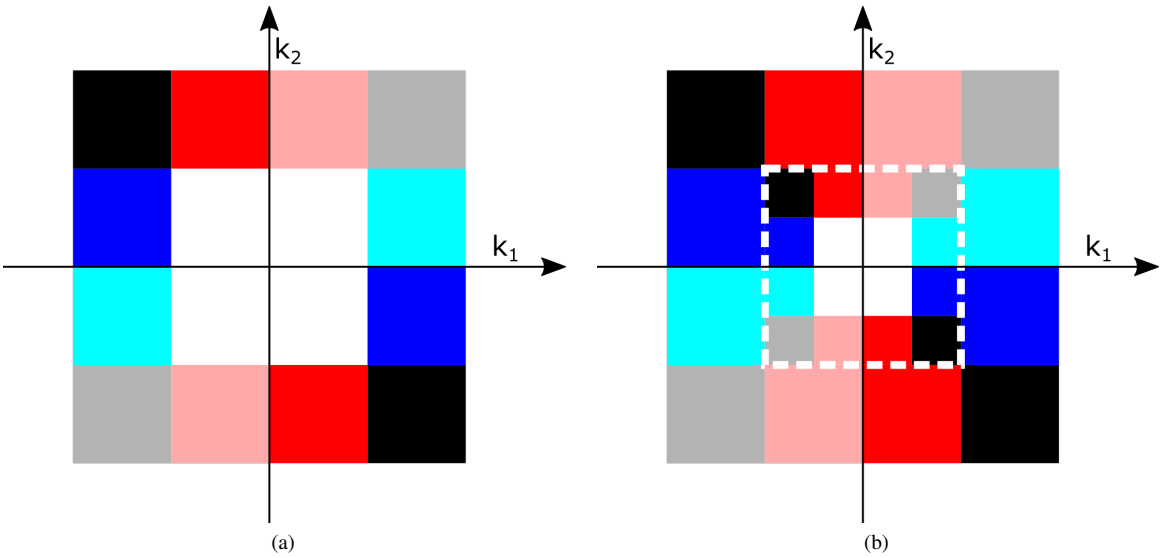


Figure 6. Idealized Fourier domain support for 2D dual-tree complex wavelet transform. (a) is one level decomposition with white box in the middle representing low frequency signal approximation. The colors denote support of distinct detail subspaces. Level two wavelet decomposition in (b) results from further decomposing the white box from (a).

colors in Figure 5(a) represent different bandwidths along signal dimensions. Thus, in 2D the white box corresponds to the low frequency along both axes, the black boxes correspond to high frequency along both axes, while red and blue represent low-high and high-low frequency portions of the signal. The black, blue and red boxes can also be linked to directional sensitivity of the wavelet transform. Since real-valued signals have symmetric amplitude spectra, the real DWT cannot distinguish between events dipping to the left or the right. Thus, for a 2D case only three directions can be distinguished: $0^\circ, \pm 45^\circ$ and 90°. More generally, in D dimensions the sensitivity is along $2^D - 1$ directions. Figure 5(b) depicts the second level wavelet decomposition of a 2D signal. Note that by decomposing further, we increase the number of bands in which directional details can be observed. Thus, the real DWT allows to examine the signals in multiple bands, which results in more sparse representation of high dynamic range data such as wavefields when compared to the Fourier transform (see Figure 4).

3.4 Complex wavelet transform

The complex wavelet transform is an enhancement to the real discrete wavelet transform offering attractive additional properties: near shift invariance and directional selectivity in 2D and higher dimensions (Selesnick et al., 2005).

In a classical DWT, a small shift of signal greatly perturbs the wavelet coefficients around signal singularities such as zero traces - an undesirable property while using overlapping spatial windows for seismic data reconstruction. Furthermore, the limited directional selectivity of DWT does distinguish betweenleft- and right-dipping events on a seismic record. The complex wavelet transform ($\mathbb{C}$WT) is able to overcome the shortcomings of DWT by replacing the classical real wavelet with a complex, approximately analytic wavelet:

$$\psi_C(t) = \psi_r(t) + \mathrm{i}\psi_i(t), \tag{3}$$

where real functions $\psi_r(t)$ and $\psi_i(t)$ are even and odd, respectively. Similarly, $\mathbb{C}$WT utilizes a complex scaling function:

$$\phi_C(t) = \phi_r(t) + \mathrm{i}\phi_i(t), \tag{4}$$

with similar properties to the complex wavelet. The scaling function acts as a lowpass filter, while the wavelet function is a bandpass filter. Since the analytic signals have one-sided amplitude support in the Fourier domain, the $\mathbb{C}$WT is able to distinguish between events of opposing dip. Figure 6 shows the idealized Fourier domain support of complex wavelet coefficients in 2D (similar principles apply in higher dimensions). Note that the number of distinct detail subspaces increased from three for DWT to six for $\mathbb{C}$WT.

One way of achieving an approximately analytic wavelet is by forming a slightly redundant frame such that both $\psi_r(t)$ and $\psi_i(t)$ form orthonormal or biorthogonal bases (Selesnick et al., 2005). An implementation following this approach is the dual-

tree $\mathbb{C}$WT (Kingsbury, 2001) based on filter banks. In fact, the dual-tree $\mathbb{C}$WT can be computed using the infrastructure used for classical DWT, though with specially designed filters (Selesnick et al., 2005). This makes the complex wavelet transform fast to compute ($\mathcal{O}(N)$ complexity) and 2^D times redundant tight frame (independent of the decomposition level) for $D-$dimensional signals.

To understand how the dual-tree $\mathbb{C}$WT works, consider the complex wavelet decomposition of a binary image of a circle shown in Figure 7. A circle is a simple object with all angles equally represented, thus it is optimal for assessing the directional selectivity of the $\mathbb{C}$WT. Since the transform yields complex-valued coefficients, we can interpret the decompositions in terms of magnitude (Figure 7(a) and 7(b)) and phase (Figure 7(c) and 7(d)). Note that despite the uniform distribution of angles, the energy of complex wavelet coefficients is not distributed evenly between detail subspaces. This implies that the $\mathbb{C}$WT is not equally sensitive to all angles in the data. The study of phase plots suggests that there is a phase shift of about 90° between the two trees except for the low frequency approximation. The shift ensures the enhanced directional selectivity of the $\mathbb{C}$WT.

The straightforward interpretation of complex wavelet coefficients combined with enhanced directional selectivity and limited redundancy of the $\mathbb{C}$WT makes it an attractive domain for sparsely representing seismic data. Note that in Figure 4, $\mathbb{C}$WT outperforms DWT when $\rho < 0.35$, making $\mathbb{C}$WT a good transform for sparsity-promoting recovery of seismic data.

3.5 Curvelet transform

The curvelet transform can be thought of as a localized oriented Fourier transform or an anisotropic generalization of the complex wavelet transform. Since curvelets are specified by scale, angle, and position, the transform contains location, orientation, and frequency information. These features come at an expense of highly redundant representation. Redundancy depends on the number of scales and angles, but is much higher than that of $\mathbb{C}$WT. For the example presented in Figure 4, there are 7.2 times more curvelet coefficients than data samples.

Although the curvelet transform is said to provide an optimally sparse representation of wave propagators (Candès and Demanet, 2005), the gains may not be as significant as expected when applied to the field seismic data. One pitfall of field data is the often prominent issue of aliased energy. Since the Fourier transform is at the core of the digital implementation of a curvelet transform (with a similar computational complexity of $\mathcal{O}(N \log N)$) (Candès et al., 2006a), it may spread aliasing artifacts across the entire domain instead of keeping them localized (Yu et al., 2017). Note that in Figure 4, the curvelets outperform real and complex wavelets for $\rho > 0.1$, but not as much as one would expect from a similar experiment with synthetic data.

4 DATA RECONSTRUCTION EXAMPLES

We demonstrate data reconstruction with field data using real wavelets, complex wavelets and curvelets. We design our experiments for data reconstruction using compressive seismic acquisition, which follows the uniform random downsampling strategy, similar to Herrmann (2010). Let $\delta = \frac{n}{N}$ denote data decimation ratio, where N is the number of traces in the undecimated data and n is the number of traces remaining after random data decimation. For every value of δ, we generate 100 realizations of uniform random sampling, with different geometries of missing traces. We then run the reconstruction for each realization to evaluate the impact of gap pattern on the reconstructed data.

We formulate the data reconstruction as sparsity promoting ℓ_1 optimization problem in the tested domains. Since field data inevitably contain noise, instead of solving equation 2, we change the constraints such that $\|\mathbf{d} - \mathbf{S}\mathbf{\Phi}^T\boldsymbol{\alpha}\|_2^2 \leq \epsilon$, where ϵ is a noise level inferred from log-amplitude plot similar to Figure 2. This is to ensure that there is no data over-fitting which could introduce high frequency artifacts to the reconstructed wavefields. The optimization is solved using the spectral projected gradient solver (Van Den Berg and Friedlander, 2008). To quantify the quality of the reconstruction, we use the signal-to-noise ratio, defined as:

$$SNR = 20 \log_{10} \Big(\frac{\|x\|}{\|x - \hat{x}\|} \Big), \tag{5}$$

where x represents the original full data and $\hat{x}$ represents the reconstructed data. We also examine the individual reconstructions, data difference and the FK spectrum of the reconstructed data to understand the domain-specific reconstruction artifacts and the effect of geometry on the reconstructed wavefields.

The undecimated shot is shown in Figure 8(a) and the corresponding FK spectrum in Figure 8(b). The data come from a mountainous region. The dominant lithologies are shale and sandstone with occasional intrusions of igneous rocks. Note that this field record showcases the described challenges with land data: slow and dispersive surface waves, source-generated noise, and large dynamic range. Furthermore, field data are critically sampled, i.e. they occupy the entire FK space. The real benefit of compressive

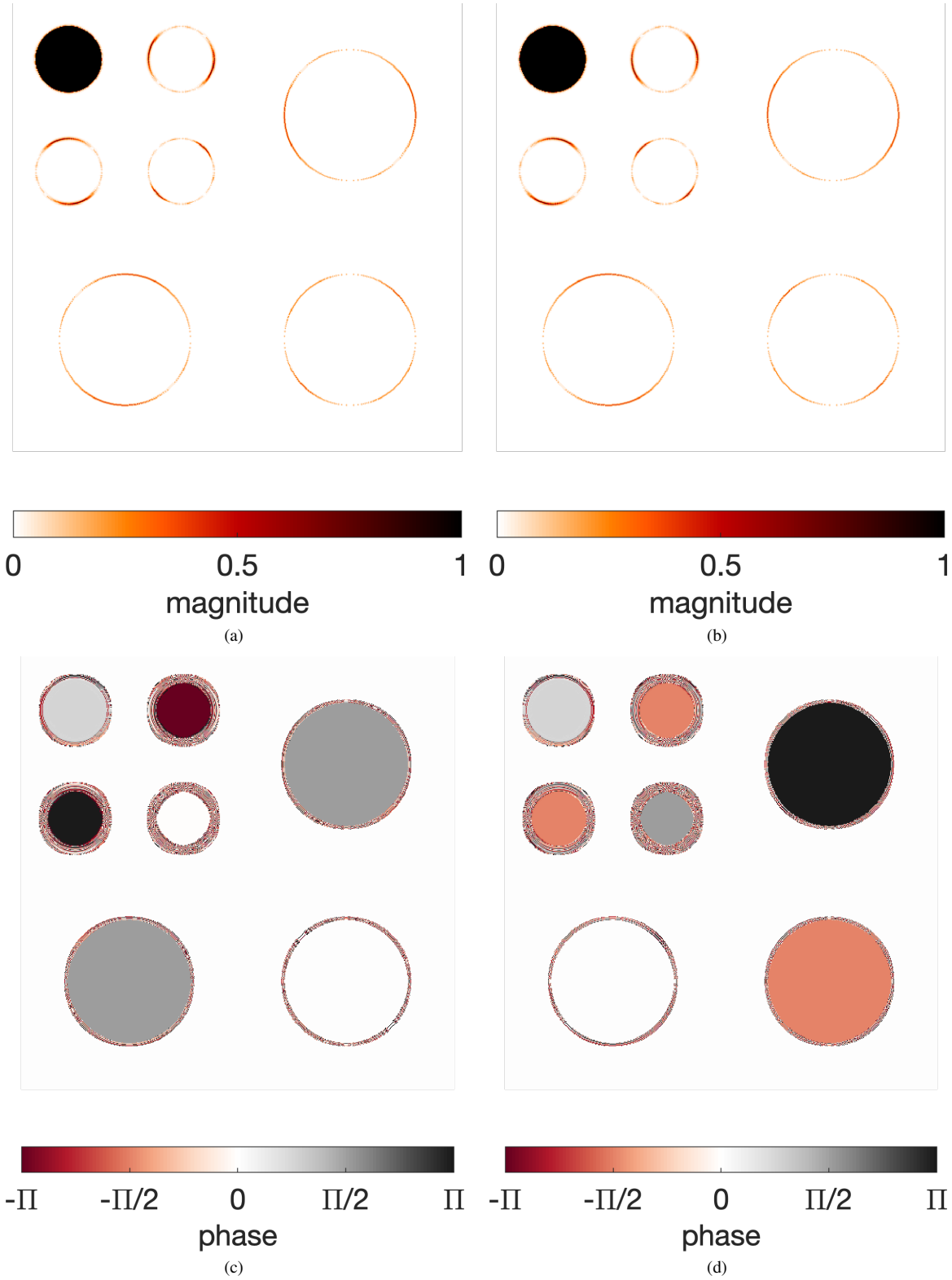


Figure 7. Magnitudes and phases of the two-level complex wavelet decomposition of a circle. Note the selective directional sensitivity and phase shift between the two trees.

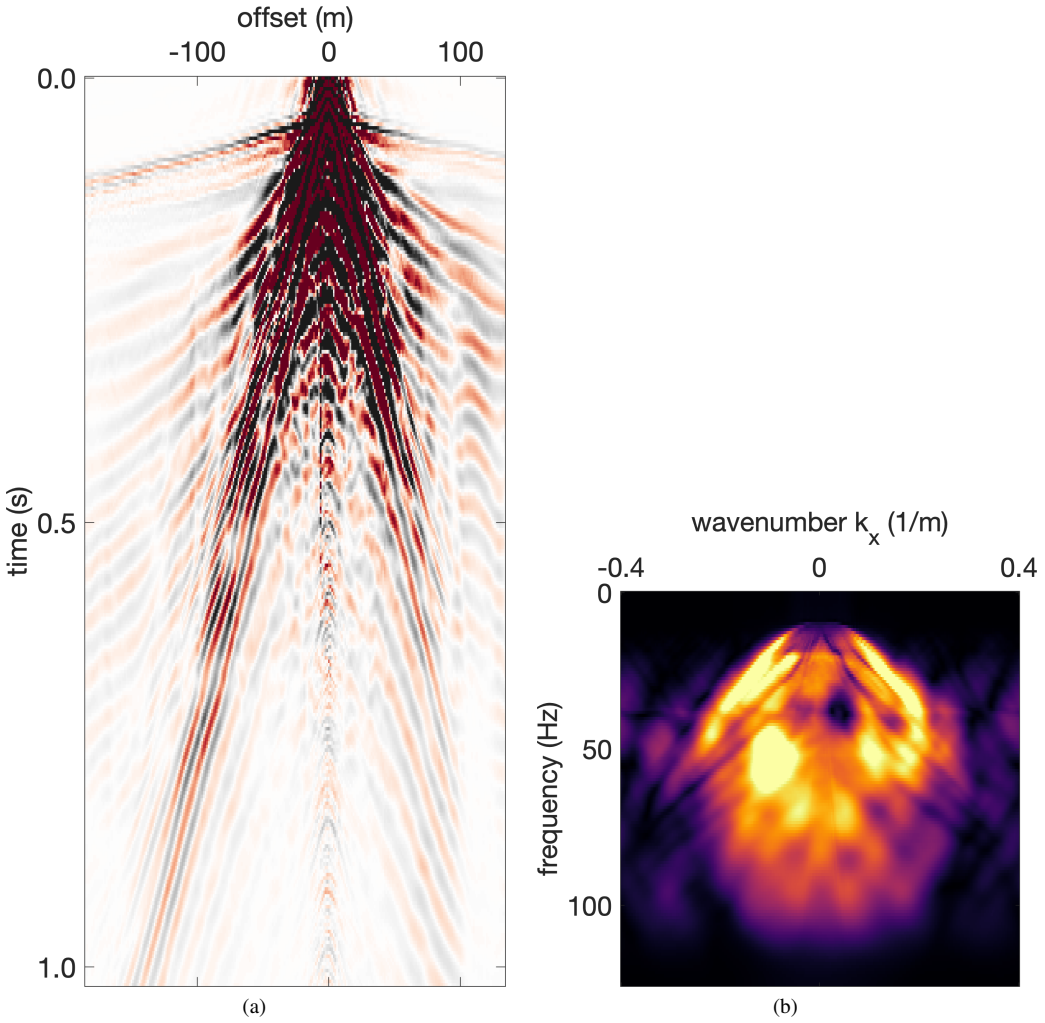


Figure 8. Field data (a) and their spectrum (b) used in reconstruction experiments. Note that the field data are critically sampled.

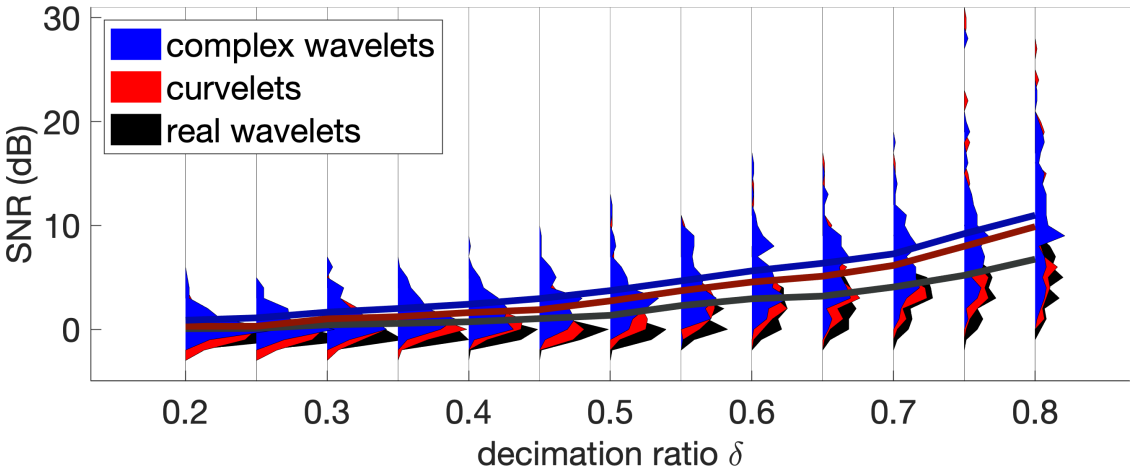


Figure 9. Summary of the reconstruction results for field data in Figure 8. Solid lines indicate the average performance of each domain.

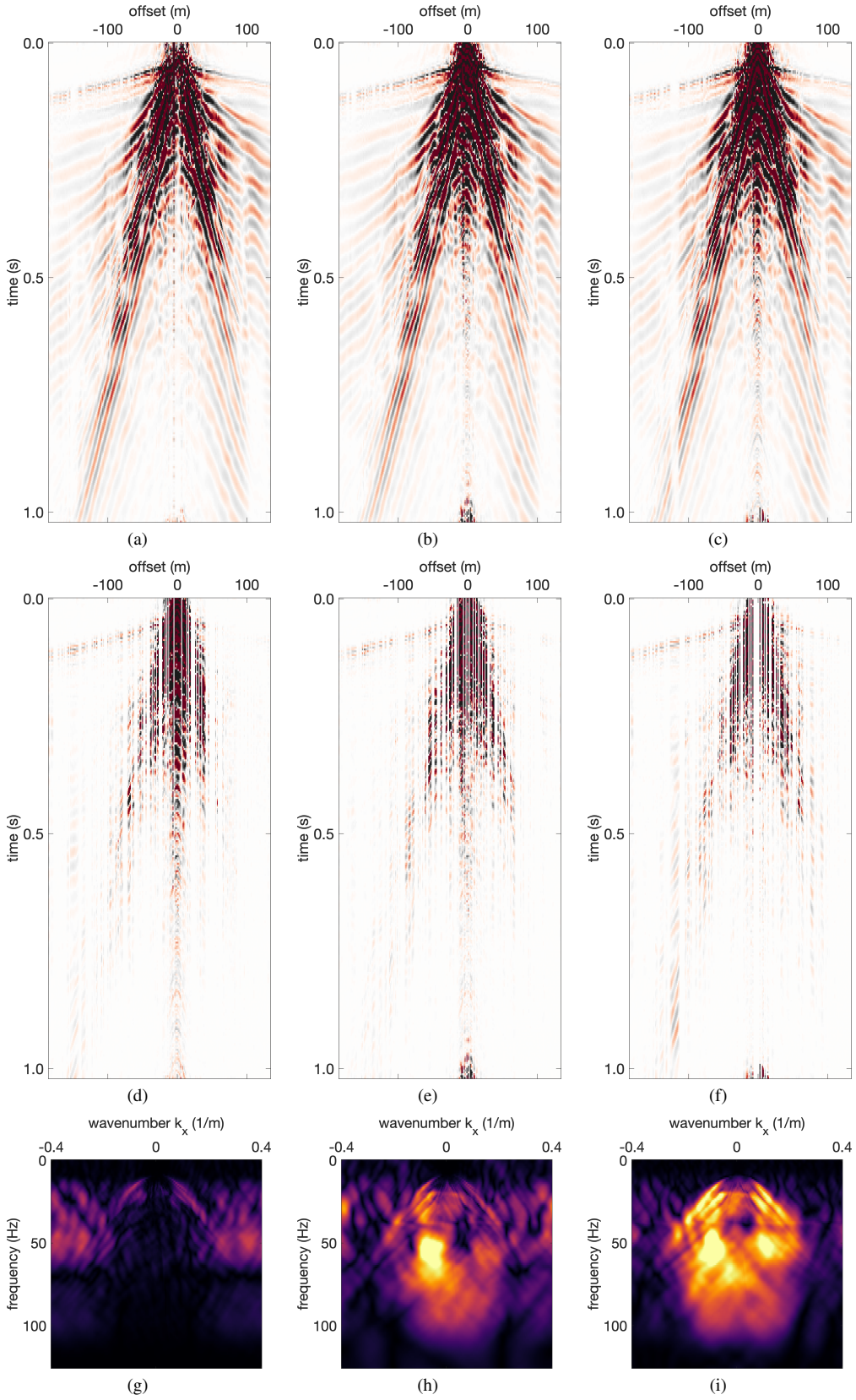


Figure 10. From left to right: the worst, average and best case **complex wavelet**-based reconstruction for $\delta = 0.5$. The middle row shows the data difference and the bottom row the FK spectrum of the reconstructed data.

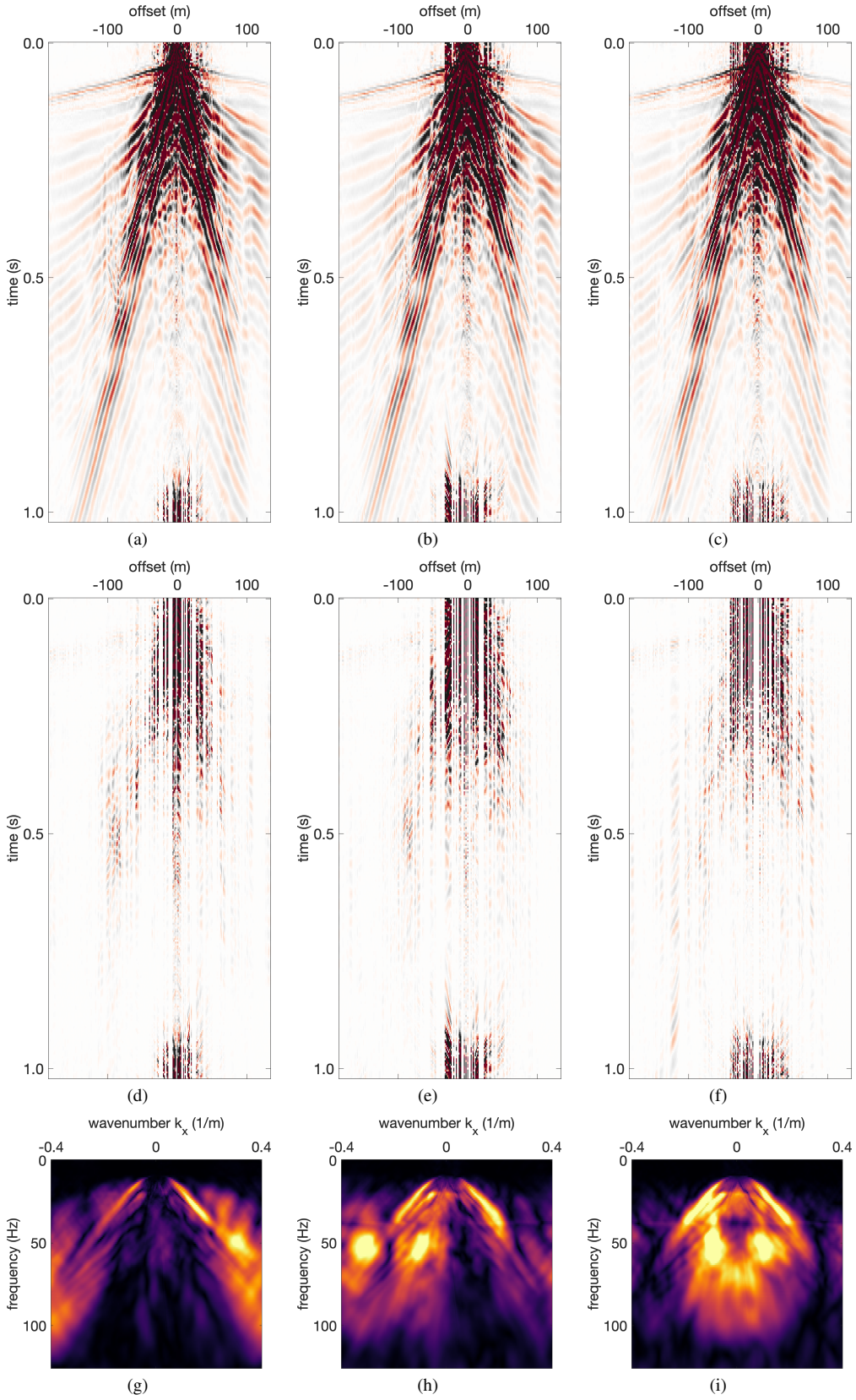


Figure 11. From left to right: the worst, average and best case **curvelet**-based reconstruction for $\delta = 0.5$. The middle row shows the data difference and the bottom row the FK spectrum of the reconstructed data.

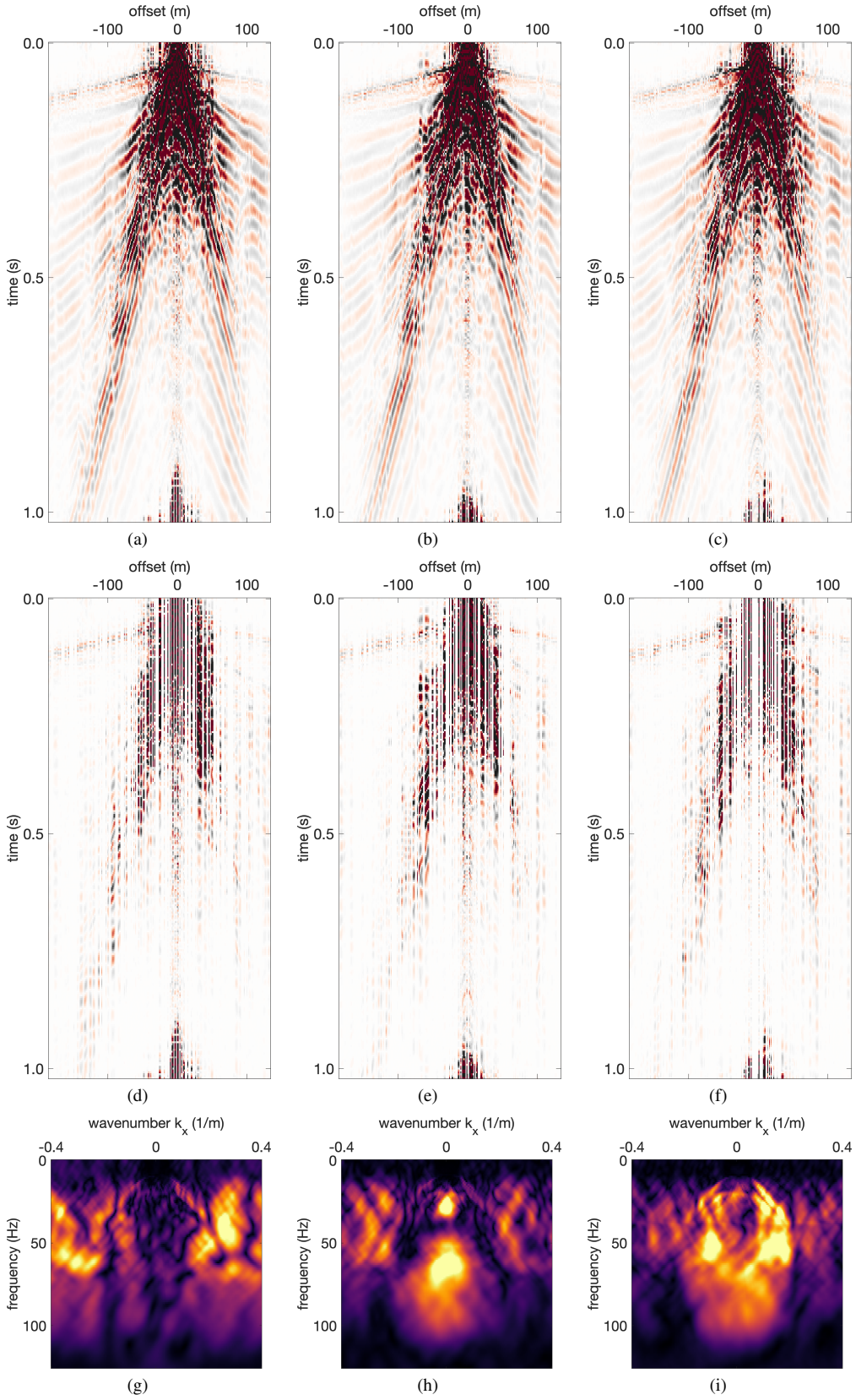


Figure 12. From left to right: the worst, average and best case **real wavelet**-based reconstruction for $\delta = 0.5$. The middle row shows the data difference and the bottom row the FK spectrum of the reconstructed data.

sensing is when one can decimate such data (which would introduce aliasing if done regularly) and still recover the non-aliased wavefield without the loss of information.

The reconstructions for real wavelets, curvelets and complex wavelets are summarized in Figure 9. On average, the complex wavelets perform the best in terms of SNR, followed by curvelets and real wavelets. The reconstruction performance is consistently poor for large decimation ratios ($\delta < 0.3$). For small decimation ratios ($\delta > 0.7$), there is a big spread in the reconstruction quality in all domains, with the best and the worst reconstruction differing by as much as 30dB. This suggests that the geometry of missing traces has a big impact on the results.

Since we use the uniform random sampling, there is no control over the size of the gaps. Scenarios in which gaps are bigger and contain unique information not present in the remaining traces yield poor reconstructions compared to the scenarios with more but smaller gaps.

Figures 10 - 12 show the worst, average, and the best reconstructions in terms of SNR, the data difference and the FK spectrum of the reconstructed data for $\delta = 0.5$ using complex wavelets, curvelets and real wavelets, respectively. Note that although the reconstructed data may look similar within the same reconstruction domain, the SNR is related to the FK spectrum of the reconstructed data: high values of SNR imply higher fidelity of spectral representation and vice versa. Accurate frequency content, especially the low temporal frequencies, is essential in many applications for reservoir characterization, for example the acoustic impedance inversion or full waveform inversion. Poorly reconstructed data may hinder instead of help the subsequent processing and inversions.

Comparing the reconstructions from different domains we note that complex wavelets and curvelets perform similarly, while the real wavelets struggle the most to reconstruct the data with the accurate spectrum. The reconstructions utilizing DWT suffer from aliased coefficients - a known shortcoming of a critically sampled wavelet transform (Selesnick et al., 2005). Thus, the DWT is not a good candidate for sparse recovery with field data. The main difference between the complex wavelets and curvelets is the reconstruction artifacts. Complex wavelets are unable to reconstruct the first arrivals, resulting in significantly lower amplitudes where traces were missing. This behavior stems from the fact that first arrivals contain high frequencies which tend to be represented by smaller transform coefficients than the low frequencies. Furthermore, as shown in Figure 7, the $\mathbb{C}$WT is not equally sensitive to all angles in the data so without providing additional structural information in the recovery process, the angles represented by small transform coefficients may be lost. Curvelets have better angular coverage and do not struggle with the first arrival, but instead exhibit strong "wrap-around" artifacts along the time axis, likely due to their implementation using wrapped Fourier transforms. Given the consistently better reconstruction quality and reduced artifacts in the reconstructed wavefields, complex wavelets are the best choice for field data reconstruction out of all tested options.

5 DISCUSSION

The geometry of the missing traces plays a key role in the successful recovery of non-aliased wavefield. All local transforms struggle with gaps comparable to the transform support. Furthermore, the reconstruction algorithm cannot create new information. Suppose that a diffraction is present in the region of missing receivers. Unless we use data from additional shots which register that diffraction from a different angle, we cannot recover the diffraction inside a gap as it is missing from the registered data. Thus, the fidelity of the reconstructed data improves as more information is included (e.g., multiple shots). Furthermore, if the seismic data are viewed as a multidimensional cube, one can resort the data in a way which makes gaps appear smaller. Changing the sorting may also have a benefit of decreasing the lateral variability of data (e.g., in common offset gathers), which in turn enhances the sparsity in many transform domains and has the potential to improve the reconstruction quality.

We show in Figure 4 that curvelets and wavelets can sparsely represent seismic data, but as evidenced by the presented examples, sparsity is not sufficient to ensure high-quality reconstruction. Additional considerations include the implementation details of the transform as described before and the coherence between sampling scheme and data representation in the transform domain. Ideally, the sampling scheme and the data sparsifying transform should be incoherent, such as with uniform random sampling and the Fourier domain. However, the Fourier transform cannot sufficiently sparsify the challenging land seismic data and other transforms have some degree of coherence with randomly missing traces. Thus, randomizing along more than one dimension (e.g. by randomizing shot timing during simultaneous acquisition) is beneficial, as shown by Herrmann (2010).

The benefits of compressive sensing come from acquiring less data than required by Nyquist theorem for regularly sampled data. When discussing the data decimation ratio, it is important to keep the Nyquist wavenumbers in mind. Suppose we decimate oversampled data uniformly by half. We would be able to infill the decimated traces with little effort using one of the many techniques for interpolation of non-aliased, regular data. With half of the traces missing at random, the challenge is driven by the gap size, but reconstruction tends to be successful. However, when sampling below Nyquist requirements, the reconstructions can

range from poor to good (see Figure 9), strongly depending on the geometry of the missing traces. This has important implications for acquisition design. To acquire data at a lower cost with fewer receivers than needed by the Nyquist theorem, one needs to develop a reconstruction strategy which provides recovery guarantees with high probability. Within the framework presented here, ensuring good quality reconstruction still requires a high channel count due to the very slow surface waves.

One promising future direction for complex wavelet-based data reconstruction is an extension to 5D. Since seismic data do not vary in space as rapidly as in time, 5D $\mathbb{C}$WT of seismic data is likely to be sparser than a 2D $\mathbb{C}$WT, 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. An additional benefit of extending the framework to higher dimensions is enhanced directional selectivity: the number of directions is $2^{2D-1} - 2^{D-1}$, which yields 28 directions in 3D and 496 directions in 5D. This would further help to sparsely represent the seismic data.

6 CONCLUSIONS

We demonstrate that ℓ_1 sparsity-promoting optimization utilizing the complex wavelet transform is an attractive alternative to curvelet domain reconstruction for challenging land seismic data. The complex wavelet transform is faster and less redundant than the curvelet transform and in addition, it is straightforward to extend to four or more dimensions. Furthermore, when applied to critically sampled or aliased field data, $\mathbb{C}$WT keeps the aliasing artifacts local instead of spreading them throughout the entire domain, thus limiting artifacts in the reconstructed data. Despite superior angular sensitivity, curvelets are not as successful on field data due to the Fourier-domain based implementation spreading aliasing effects and resulting in strong artifacts in the reconstructed data.

Random sampling in conjunction with the complex wavelet domain sparsity promoting algorithm allows to reduce the number of channels needed to acquire non-aliased wavefields below the limit imposed by the Nyquist theorem, but the exact decimation ratio strongly depends on the specific receiver geometry. The successful recovery is characterized by high SNR values, which are good indicators of Fourier domain fidelity of the reconstructed data. The sparsity promoting data pre-processing or extension to higher dimensions are needed to increase the geometry robustness of the data reconstruction.

7 ACKNOWLEDGEMENTS

We thank the sponsor companies of the Center for Wave Phenomena, whose support made this research possible.

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.

Baraniuk, R. G., and P. Steeghs, 2017, Compressive sensing: A new approach to seismic data acquisition: The Leading Edge, **36**, 642–645.

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

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., 2006b, Compressive sampling: Proceedings of the International Congress of Mathematicians, Madrid, Spain, 1433–1452.

Candès, E. J., J. Romberg, and T. Tao, 2006c, Robust uncertainty principles: exact signal reconstruction from highly incomplete 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.

Claerbout, J., 2014, Geophysical image estimation by example: Lulu.com.

Crawley, S., R. Clapp, and J. Claerbout, 1999, Interpolation with smoothly nonstationary prediction-error filters, *in* SEG Technical Program Expanded Abstracts 1999: Society of Exploration Geophysicists, 1154–1157.

Denisov, M., A. Egorov, E. Kurin, and M. Shneerson, 2019, Vibroseis harmonic noise elimination based on optimized recursive filtering: 81st EAGE Conference and Exhibition 2019, European Association of Geoscientists & Engineers, 1–5.

Douma, H., and M. V. de Hoop, 2007, Leading-order seismic imaging using curvelets: Geophysics, **72**, S231–S248.

Fomel, S., and J. F. Claerbout, 2003, Multidimensional recursive filter preconditioning in geophysical estimation problems: Geophysics, **68**, 577–588.

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

Foti, S., C. Lai, G. J. Rix, and C. Strobbia, 2014, Surface wave methods for near-surface site characterization: CRC Press.

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.

Gao, J., M. D. Sacchi, and X. Chen, 2013, A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions: Geophysics, **78**, V21–V30.

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.

Hollander, Y., and O. Yilmaz, 2019, An acceleration method for the anti-leakage parabolic radon transform for seismic data interpolation, *in* SEG Technical Program Expanded Abstracts 2019: Society of Exploration Geophysicists, 4480–4484.

Jia, Y., and J. Ma, 2017, What can machine learning do for seismic data processing? An interpolation application: Geophysics, **82**, V163–V177.

Jia, Y., S. Yu, and J. Ma, 2018, Intelligent interpolation by Monte Carlo machine learning: Geophysics, **83**, V83–V97.

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.

Kingsbury, N., 2001, Complex wavelets for shift invariant analysis and filtering of signals: Applied and computational harmonic analysis, **10**, 234–253.

Kreimer, N., and M. D. Sacchi, 2012, A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation: Geophysics, **77**, V113–V122.

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.

Kumar, R., C. Da Silva, O. Akalin, A. Y. Aravkin, H. Mansour, B. Recht, and F. J. Herrmann, 2015, Efficient matrix completion for seismic data reconstruction: Geophysics, **80**, V97–V114.

Li, C., C. Mosher, Y. Tao, and Y. Zhang, 2017, A low-rank approximation method for ice-break noise attenuation of simultaneous vibroseis data, *in* SEG Technical Program Expanded Abstracts 2017: Society of Exploration Geophysicists, 4930–4934.

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

Liu, Y., P. Zhang, and C. Liu, 2017, Seismic data interpolation using generalised velocity-dependent seislet transform: Geophysical Prospecting, **65**, 82–93.

Ma, J., 2013, Three-dimensional irregular seismic data reconstruction via low-rank matrix completion: Geophysics, **78**, V181–V192.

Mallat, S. G., 1989a, Multiresolution approximations and wavelet orthonormal bases of $l^2(r)$: Transactions of the American Mathematical Society, **315**, 69–87.

———, 1989b, A theory for multiresolution signal decomposition: the wavelet representation: IEEE Transactions on Pattern Analysis and Machine Intelligence, **11**, 674–693.

Manning, T., C. Brooks, A. Ourabah, M. Popham, D. Ablyazina, V. Zhuzhel, E. Holst, and N. Goujon, 2018, The case for a nimble node, towards a new land seismic receiver system with unlimited channels, *in* SEG Technical Program Expanded Abstracts 2018: Society of Exploration Geophysicists, 21–25.

Morlet, J., G. Arens, E. Fourgeau, and D. Giard, 1982a, Wave propagation and sampling theory—Part II: Sampling theory and complex waves: Geophysics, **47**, 222–236.

Morlet, J., G. Arens, E. Fourgeau, and D. Glard, 1982b, Wave propagation and sampling theory—Part I: Complex signal and scattering in multilayered media: Geophysics, **47**, 203–221.

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, 2009, 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.

Papadopoulou, M., F. Da Col, B. Mi, E. Bäckström, P. Marsden, B. Brodic, A. Malehmir, and L. V. Socco, 2020, Surface-wave analysis for static corrections in mineral exploration: a case study from central Sweden: Geophysical Prospecting, **68**, 214–231.

Pilikos, G., and A. Faul, 2017, Bayesian feature learning for seismic compressive sensing and denoising: Geophysics, **82**, O91–O104.

Polom, U., 1997, Elimination of source-generated noise from correlated vibroseis data (the 'ghost-sweep' problem): Geophysical Prospecting, **45**, 571–591.

Selesnick, I. W., R. G. Baraniuk, and N. G. Kingsbury, 2005, The dual-tree complex wavelet transform: IEEE Signal Processing Magazine, **22**, 123–151.

Sheriff, R. E., 2002, Encyclopedic dictionary of applied geophysics: Society of Exploration Geophysicists.

Socco, L. V., S. Foti, and D. Boiero, 2010, Surface-wave analysis for building near-surface velocity models—established approaches and new perspectives: Geophysics, **75**, 75A83–75A102.

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.

Trickett, S., L. Burroughs, A. Milton, L. Walton, and R. Dack, 2010, Rank-reduction-based trace interpolation, *in* SEG Technical Program Expanded Abstracts 2010: Society of Exploration Geophysicists, 3829–3833.

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., R. Abma, J. Etgen, and C. Sullivan, 2017, Attenuation of noise and simultaneous source interference using wavelet denoising: Geophysics, **82**, V179–V190.

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