Simply denoise: wavefield reconstruction via jittered undersampling

Gilles Hennenfent[1] and Felix J. Herrmann[1]

ABSTRACT

In this paper, we present a new discrete undersampling scheme designed to favor wavefield reconstruction by sparsity-promoting inversion with transform elements that are localized in the Fourier domain. Our work is motivated by empirical observations in the seismic community, corroborated by recent results from compressive sampling, which indicate favorable (wavefield) reconstructions from random as opposed to regular undersampling. As predicted by theory, random undersampling renders coherent aliases into harmless incoherent random noise, effectively turning the interpolation problem into a much simpler denoising problem.

A practical requirement of wavefield reconstruction with localized sparsifying transforms is the control on the maximum gap size. Unfortunately, random undersampling does not provide such a control and the main purpose of this paper is to introduce a sampling scheme, coined jittered undersampling, that shares the benefits of random sampling, while offering control on the maximum gap size. Our contribution of jittered sub-Nyquist sampling proves to be key in the formulation of a versatile wavefield sparsity-promoting recovery scheme that follows the principles of compressive sampling.

After studying the behavior of the jittered undersampling scheme in the Fourier domain, its performance is studied for curvelet recovery by sparsity-promoting inversion (CRSI). Our findings on synthetic and real seismic data indicate an improvement of several decibels over recovery from regularly-undersampled data for the same amount of data collected.

INTRODUCTION

While the argument has been made that there is no real theoretical requirement for regular spatial sampling of seismic data (Bednar, 1996), most of the commonly-used multi-trace processing algorithms, e.g., Surface-Related Multiple Elimination (SRME - Verschuur et al., 1992) and wave-equation migration (WEM - Claerbout, 1971), need a dense and regular coverage of the survey area. Field datasets, however, are typically irregularly and/or coarsely sampled along one or more spatial coordinates and need to be interpolated before being processed.

---

[1]Seismic Laboratory for Imaging and Modeling, Department of Earth and Ocean Sciences, The University of British Columbia, 6339 Stores Road, Vancouver, V6T 1Z4, BC, Canada

[1]**e-mail:** ghennenfent@eos.ubc.ca, fherrmann@eos.ubc.ca

For regularly-undersampled data along one or more spatial coordinates, i.e., data spatially sampled below Nyquist rate, there exists a wide variety of wavefield reconstruction techniques. Filter-based methods interpolate by convolution with a filter designed such that the error is white noise. The most common of these filters are the prediction error filters (PEF's) that can handle aliased events (Spitz, 1991). Wavefield-operator-based methods represent another type of interpolation approaches that explicitly include wave propagation (Canning and Gardner, 1996; Biondi et al., 1998; Stolt, 2002). Finally, transform-based methods also provide efficient algorithms for seismic data regularization (Sacchi et al., 1998; Trad et al., 2003; Zwartjes and Sacchi, 2007; Herrmann and Hennenfent, 2007). However, for irregularly-sampled data, e.g., binned data with some of the bins that are empty, or data that are continuous random undersampled, the performance of most of the aforementioned interpolation methods deteriorates.

The objective of this paper is to demonstrate that irregular/random undersampling is not a drawback for particular transform-based interpolation methods and for many other advanced processing algorithms as was already observed by other authors (Zhou and Schuster, 1995; Sun et al., 1997; Trad and Ulrych, 1999; Xu et al., 2005; Abma and Kabir, 2006; Zwartjes and Sacchi, 2007). We explain why random undersampling is an advantage and how it can be used to our benefit when designing coarse sampling schemes. To keep the discussion as clear and concise as possible, we focus on regular sampling with randomly missing data points, i.e., discrete random (under)sampling. Our conclusions extend to continuous random undersampling though. Unless otherwise specified, the term random is used in the remaining of the text in the discrete sense.

Motivation

Recent results in Information Theory and Approximation Theory established that a signal can be recovered *exactly* from (severely) undersampled data points provided that 1) the signal exhibits sparsity in a known transform domain, 2) the artifacts introduced by undersampling look like incoherent random noise in the sparsifying domain, and 3) a data-consistent sparsity-promoting procedure is used for the recovery. It is possible to build an intuitive understanding of these theoretical results, termed *Compressive Sampling* (CS - Candès et al., 2006; Donoho, 2006; Candès and Romberg, 2006), by considering a simple example.

Figure 1(a) shows the superposition of three cosine functions. This signal is sparse in the Fourier domain (condition 1) and is regularly sampled above Nyquist rate. Its amplitude spectrum is plotted in Figure 1(b). When the signal is randomly three-fold undersampled according to a discrete uniform distribution as in Figure 1(c), its amplitude spectrum, plotted in Figure 1(d), is corrupted by artifacts (condition 2) that look like additive incoherent random noise. In this case, the significant coefficients of the to-be-recovered signal remain above the "noise" level. These coefficients can be detected with a denoising technique that promotes sparsity, e.g., nonlinear

thresholding (dashed line in Figures 1(d) and 1(f)), and exactly recovered by an amplitude-matching procedure to fit the acquired data (condition 3). This experiment illustrates a favorable recovery from severely undersampled data points of a signal that is sparse in the Fourier domain.

When the original signal is regularly three-fold undersampled (Figure 1(e)), the undersampling artifacts coherently interfere, giving rise to well-known aliases that look like the original signal components (Figure 1(f)). In this case, the above sparsity-promoting recovery scheme may fail because the to-be-recovered signal components and the aliases are both sparse in the Fourier domain. This example suggests that random undersampling according to a discrete uniform distribution is more favorable than regular undersampling for reconstruction algorithms that promote sparsity in the Fourier domain. In general terms, the above observations hint at undersampling schemes that lead to more favorable recovery conditions. Within the field of CS, significant advances have been made regarding the main ingredients that go into the design of an undersampling scheme that favors sparsity-promoting recovery. In this paper, we draw on these results to design a new coarse spatial sampling scheme for seismic data.

Main contributions

We propose and analyze a coarse sampling scheme, termed *jittered undersampling* (Leneman, 1966; Dippe and Wold, 1992), which creates, under specific conditions, a favorable recovery situation for seismic wavefield reconstruction methods that impose sparsity in Fourier or Fourier-related domains (see e.g. Sacchi et al., 1998; Xu et al., 2005; Zwartjes and Sacchi, 2007; Herrmann and Hennenfent, 2007). Jittered undersampling differentiates itself from random undersampling according to a discrete uniform distribution, which also creates favorable recovery conditions (Xu et al., 2005; Abma and Kabir, 2006; Zwartjes and Sacchi, 2007), by controlling the maximum gap in the acquired data. This control makes jittered undersampling very well suited to methods that rely on transforms with localized elements, e.g., windowed Fourier or curvelet transform (Candès et al., 2005a, and references therein). These methods are known to be vulnerable to gaps in the data that are larger than the spatio-temporal extent of the transform elements (Trad et al., 2005).

Outline

After a brief overview of the CS framework and the criteria for a favorable recovery, the effects of different undersampling schemes are studied for signals that are sparse in the Fourier domain. Next, we discuss the advantages of random undersampling and design our jittered undersampling strategy that offers increased control on the acquisition grid. The performance of this new scheme for curvelet-based recovery is illustrated on synthetic and real data.
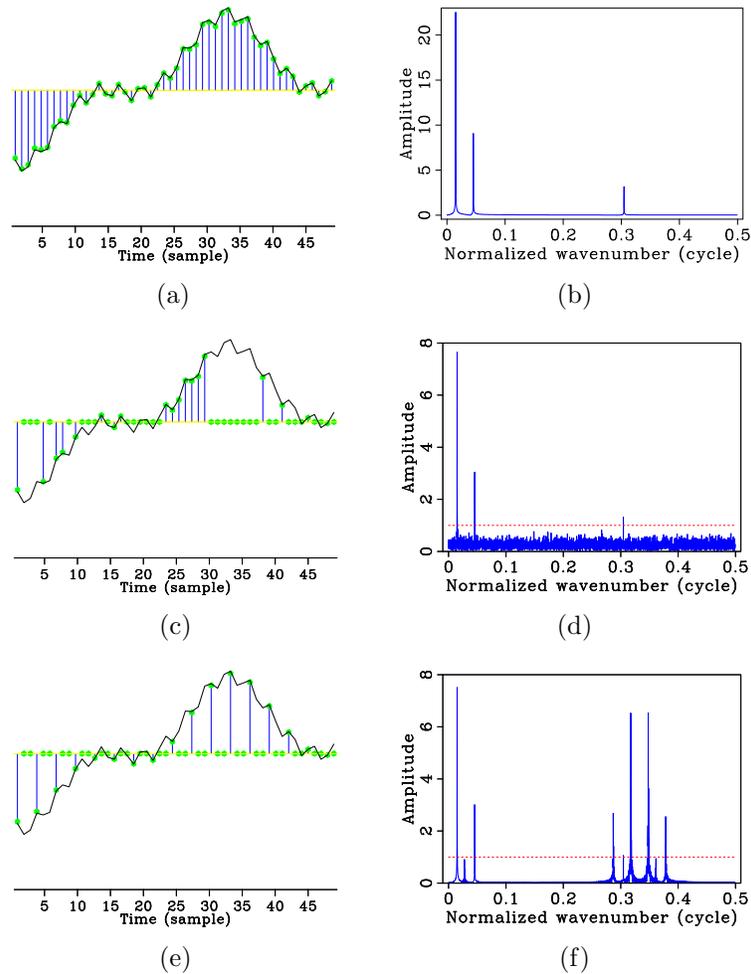
Figure 1: Different (under)sampling schemes and their imprint in the Fourier domain for a signal that is the superposition of three cosine functions. Signal (a) regularly sampled above Nyquist rate, (c) randomly three-fold undersampled according to a discrete uniform distribution, and (e) regularly three-fold undersampled. The respective amplitude spectra are plotted in (b), (d) and (f). Unlike aliases, the undersampling artifacts due to random undersampling can easily be removed using a standard denoising technique promoting sparsity, e.g., nonlinear thresholding (dashed line), effectively recovering the original signal.

THEORY

Basics of compressive sampling

An overview of the CS framework and criteria for favorable recovery conditions is given. As mentioned before, CS relies on a sparsifying transform for the to-be-recovered signal and uses this sparsity prior to compensate for the undersampling during the recovery process. For the reconstruction of wavefields in the Fourier (Sacchi et al., 1998; Xu et al., 2005; Zwartjes and Sacchi, 2007), Radon (Trad et al., 2003), and curvelet (Hennenfent and Herrmann, 2005; Herrmann and Hennenfent, 2007) domains, sparsity promotion is a well-established technique documented in the geophysical literature. The main contribution of CS is the new light shed on the favorable recovery conditions.

Recovery by sparsity-promoting inversion

Consider the following linear forward model for the recovery problem

$$\mathbf{y} = \mathbf{R}\mathbf{f}_0, \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^n$ represents the acquired data, $\mathbf{f}_0 \in \mathbb{R}^N$ with $N \gg n$ the unaliased signal to be recovered, i.e., the model, and $\mathbf{R} \in \mathbb{R}^{n \times N}$ the restriction operator that collects the acquired samples from the model. Assume that $\mathbf{f}_0$ has a sparse representation $\mathbf{x}_0 \in \mathbb{C}^N$ in some known transform domain $\mathbf{S}$, equation 1 can now be reformulated as

$$\mathbf{y} = \mathbf{A}\mathbf{x}_0 \quad \text{with} \quad \mathbf{A} \stackrel{\text{def}}{=} \mathbf{R}\mathbf{S}^H, \tag{2}$$

where the symbol ^{H} represents the conjugate transpose. As a result, the sparsity of $\mathbf{x}_0$ can be used to overcome the singular nature of $\mathbf{A}$ when estimating $\mathbf{f}_0$ from $\mathbf{y}$. After sparsity-promoting inversion, the recovered signal is given by $\tilde{\mathbf{f}} = \mathbf{S}^H \tilde{\mathbf{x}}$ with

$$\tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} ||\mathbf{x}||_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{3}$$

In these expressions, the symbol $\tilde{\ }$ represents estimated quantities and the ℓ_1 norm is defined as $||\mathbf{x}||_1 \stackrel{\text{def}}{=} \sum_{i=1}^{N} |\mathbf{x}[i]|$, where $\mathbf{x}[i]$ is the ith entry of the vector $\mathbf{x}$.

Minimizing the ℓ_1 norm in equation 3 promotes sparsity in $\mathbf{x}$ and the equality constraint ensures that the solution honors the acquired data. Among all possible solutions of the (severely) underdetermined system of linear equations ($n \ll N$) in equation 2, the optimization problem in equation 3 finds a sparse or, under certain conditions, the sparsest (Donoho and Huo, 2001) possible solution that explains the data.

Favorable recovery conditions

Following Verdu (1998) and Donoho et al. (2006), we define the matrix $\mathbf{L} \stackrel{\text{def}}{=} \mathbf{A}^H\mathbf{A} - \alpha\mathbf{I}$ to study the undersampling artifacts $\mathbf{z} \stackrel{\text{def}}{=} \mathbf{L}\mathbf{x}_0$. The matrix $\mathbf{I}$ is the identity matrix and the parameter α is a scaling factor such that $\text{diag}(\mathbf{L}) = \mathbf{0}$. For more general problems and in particular in the field of digital communications, these undersampling artifacts $\mathbf{z}$ are referred to as *Multiple-Access Interference* (MAI).

According to the CS theory (Candès et al., 2006; Donoho, 2006), the solution $\tilde{\mathbf{x}}$ in equation 3 and $\mathbf{x}_0$ coincide when two conditions are met, namely 1) $\mathbf{x}_0$ is sufficiently sparse, i.e., $\mathbf{x}_0$ has few nonzero entries, and 2) the undersampling artifacts are incoherent, i.e., $\mathbf{z}$ does not contain coherent energy. The first condition of sparsity requires that the energy of $\mathbf{f}_0$ is well concentrated in the sparsifying domain. The second condition of incoherent random undersampling artifacts involves the study of the sparsifying transform $\mathbf{S}$ in conjunction with the restriction operator $\mathbf{R}$. Intuitively, it requires that the artifacts $\mathbf{z}$ introduced by undersampling the original signal $\mathbf{f}_0$ are not sparse in the $\mathbf{S}$ domain. When this condition on $\mathbf{z}$ is not met, sparsity alone is no longer an effective prior to solve the recovery problem. Albeit qualitative, the second condition provides a fundamental insight in choosing undersampling schemes that favor recovery by sparsity-promoting inversion.

Fourier-domain undersampling artifacts

Undersampling artifacts in the Fourier domain are studied for two reasons. Firstly, several interpolation methods are based on the Fourier transform (Sacchi et al., 1998; Xu et al., 2005; Zwartjes and Sacchi, 2007). Secondly, the curvelet transform, a dyadic-parabolic partition of the Fourier domain, forms the basis of our recently-introduced recovery scheme (Herrmann and Hennenfent, 2007). Curvelets are in many situations to be preferred over Fourier because of their ability to sparsely represent complex seismic data. For a detailed discussion on this topic, we refer to Candès et al. (2005a) and Hennenfent and Herrmann (2006).

In the coming discussion, the sparsifying transform is defined as the Fourier transform, i.e., $\mathbf{S} \stackrel{\text{def}}{=} \mathbf{F}$. For this definition, the vector generating the Hermitian Toeplitz and circulant matrix $\mathbf{A}^H\mathbf{A}$ is the discrete Fourier transform of the (under)sampling pattern. This pattern has ones where samples are taken, zeros otherwise. Besides, the undersampling artifacts generated by the convolution operator $\mathbf{L}$ are known as *spectral leakage* (Xu et al., 2005).

Regular (under)sampling

When $\mathbf{R}$ keeps all the data points of $\mathbf{f}_0$, i.e., $\mathbf{R} = \mathbf{I}$, the matrix $\mathbf{A}^H\mathbf{A}$ is the identity matrix, as depicted in Figure 2(a), $\mathbf{L} = 0$, as plotted in Figure 2(d), and there is no

spectral leakage. This property holds for any orthonormal sparsifying transform.

When **R** corresponds to a regular undersampling scheme, the matrix $\mathbf{A}^H\mathbf{A}$ is no longer diagonal. It now also has a number of nonzero off-diagonals as depicted in Figure 2(b). These off-diagonals create aliases, i.e., undersampling artifacts that are the superposition of circular-shifted versions of the original spectrum. Since $\mathbf{x}_0$ is assumed to be sparse, these aliases are sparse as well. Therefore, they are also likely to enter in the solution $\tilde{\mathbf{x}}$ during sparsity-promoting inversion. Because the ℓ_1 norm can not efficiently discriminate the original spectrum from its aliases, regular undersampling is the most challenging case for recovery.

In the seismic community, difficulties with regularly undersampled data are acknowledged when reconstructing by promoting sparsity in the Fourier domain. For example, Xu et al. (2005) write that the anti-leakage Fourier transform for seismic data regularization "may fail to work when the input data has severe aliasing".

Random undersampling according to a discrete uniform distribution

When **R** corresponds to a random undersampling according to a discrete uniform distribution, the situation is completely different. The matrix $\mathbf{A}^H\mathbf{A}$ is dense (Figure 2(c)) and the convolution matrix **L** is a random matrix (Figure 2(f)). Consequently, we have

$$\mathbf{A}^H\mathbf{y} = \mathbf{A}^H\mathbf{A}\mathbf{x}_0 \approx \alpha\mathbf{x}_0 + \mathbf{n}, \tag{4}$$

where the spectral leakage is approximated by additive white Gaussian noise **n**. For infinitely large systems (Donoho et al., 2006), this approximation becomes an equality. Because of this property, the recovery problem turns into a much simpler denoising problem, followed by a correction for the amplitudes. Remember that the acquired data **y** are noise-free (cf. equation 2) and that the noise **n** in equation 4 only comes from the underdeterminedness of the system. In other words, random undersampling according to a discrete uniform distribution spreads the energy of the spectral leakage across the Fourier domain turning the noise-free underdetermined problem (cf. equation 2) into a noisy well-determined problem (cf. equation 4) whose solution can be recovered by solving equation 3. This observation was first reported by Donoho et al. (2006).

The practical requirement of maximum gap control

As shown in the previous section, random undersampling according to a discrete uniform distribution creates favorable recovery conditions for a reconstruction procedure that promotes sparsity in the Fourier domain. However, a global transform such as the Fourier transform does not typically permit a sparse representation for complex seismic wavefields (Hennenfent and Herrmann, 2006). It requires a more local transform, e.g., windowed Fourier (Zwartjes and Sacchi, 2007) or curvelet (Herrmann and Hennenfent, 2007) transform. In this case, problems arise with gaps in
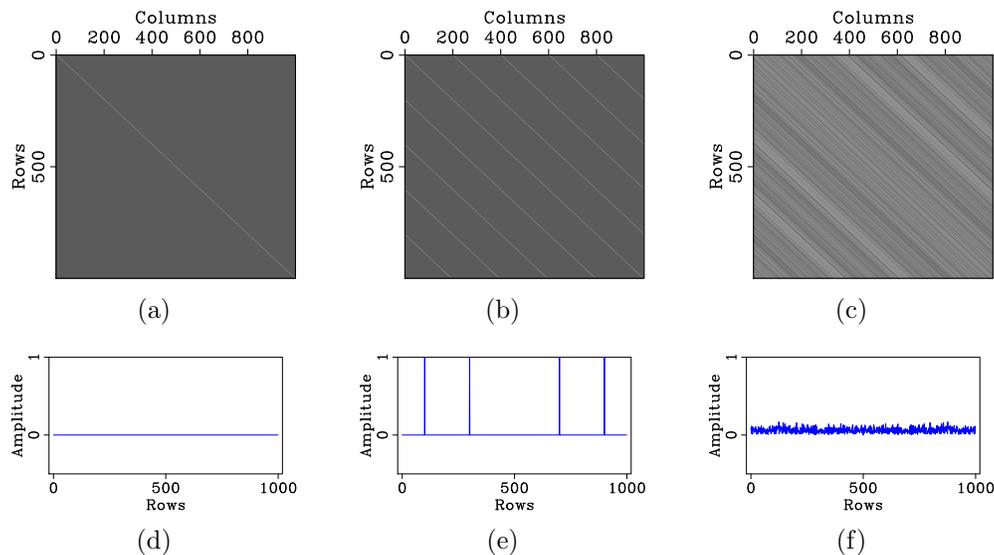
Figure 2: Convolution matrix (in amplitude) for (a) regular sampling above Nyquist rate, (b) regular five-fold undersampling, and (c) random five-fold undersampling according to a discrete uniform distribution. The respective convolution kernels (in amplitude) that generate spectral leakage are plotted in (d), (e) and (f). Despite the same undersampling factor, regular and random undersamplings produce very different spectral leakage.

the data that are larger than the spatio-temporal extent of the transform elements (Trad et al., 2005). Consequently, undersampling schemes with no control on the size of the maximum gap, e.g., random undersampling according to a discrete uniform distribution, become less attractive. The term gap refers here to the interval between two adjacent acquired traces minus the interval associated with the fine interpolation grid, such that adequate sampling has gaps of zero. We present an undersampling scheme that has, under some specific conditions, an anti-aliasing effect, yet offering control on the size of the maximum gap.

Uniform jittered undersampling on a grid

First, the undersampling grid is defined for a discrete uniform jitter. Next, the spectral leakage caused by this scheme is studied.

Definition of the jittered grid

The basic idea of jittered undersampling is to regularly decimate the interpolation grid and subsequently perturb the coarse-grid sample points on the fine grid. As for random undersampling according to a discrete uniform distribution, where each

location is equally likely to be sampled, a discrete uniform distribution for the perturbation around the coarse-grid points is considered (see Appendix A and Leneman (1966) for more details).

To keep the derivation of our jittered undersampling scheme succinct, the undersampling factor, γ, is taken to be odd, i.e., $\gamma = 1, 3, 5, \ldots$ We also assume that the size N of the interpolation grid is a multiple of γ so that the number of acquired data points $n = N/\gamma$ is an integer. For these choices, the jittered-sampled data points are given by

$$\mathbf{y}[i] = \mathbf{f}_0[j] \quad \text{for} \quad i = 1, \ldots, n \quad \text{and} \quad j = \underbrace{\frac{1-\gamma}{2} + \gamma \cdot i}_{\text{deterministic}} + \underbrace{\epsilon_i}_{\text{random}}, \tag{5}$$

where the discrete random variables ϵ_i are integers independently and identically distributed (iid) according to a uniform distribution on the interval between $-\lfloor (\xi - 1)/2 \rfloor$ and $\lfloor (\xi - 1)/2 \rfloor$. The jitter parameter $0 \leq \xi \leq \gamma$ relates to the size of the perturbation around the coarse regular grid. The floor function of a real number q, denoted $\lfloor q \rfloor$, is a function that returns the highest integer less than or equal to q. The above sampling can be adapted for the case γ is even.

In Figure 3, schematic illustrations are included for samplings with increasing randomness. The fine grid of open circles denotes the interpolation grid on which the model $\mathbf{f}_0$ is defined. The solid circles correspond to the coarse sampling locations. These illustrations show that for jittered undersampling, the maximum gap size can not exceed $(\gamma - 1) + 2 \cdot \lfloor (\xi - 1)/2 \rfloor$ data points. For regular undersampling, all the gaps are of size $\gamma - 1$ and for random undersampling according to a discrete uniform distribution, the maximum gap size is $N - n$. Remember that the number of samples is the same for each of these undersampling schemes.

As mentioned earlier, recovery with localized transforms depends on both the maximum gap size and a sufficient sampling randomness to break the coherent aliases. In the next section, we show how the value of the jitter parameter controls these two aspects in our undersampling scheme.

Fourier-domain artifacts of the jittered grid

When $\mathbf{R}$ describes a jittered undersampling scheme according to a discrete uniform distribution, the stochastic expectation $\mathrm{E}\{\cdot\}$ of the first column $\mathbf{a}$ of the circulant matrix $\mathbf{A}^H \mathbf{A}$ is given by

$$\mathrm{E}\{\mathbf{a}[k]\} \approx \begin{cases} n \cdot \mathrm{sinc}\left(\frac{\xi}{N}(k-1)\right), & \text{if} \quad k = 1 + l \cdot n \quad \text{for} \quad l = 0, \ldots, \frac{\gamma-1}{2} \\ n \cdot \mathrm{sinc}\left(\frac{\xi}{N}(k-1-N)\right), & \text{if} \quad k = 1 + l \cdot n \quad \text{for} \quad l = \frac{\gamma+1}{2}, \ldots, \gamma - 1 \\ 0 & \text{otherwise,} \end{cases}$$
$$\tag{6}$$

where $\mathrm{sinc}(\cdot)$ is the normalized sinc function defined as $\mathrm{sinc}(x) \overset{\text{def}}{=} \sin(\pi x)/\pi x$.
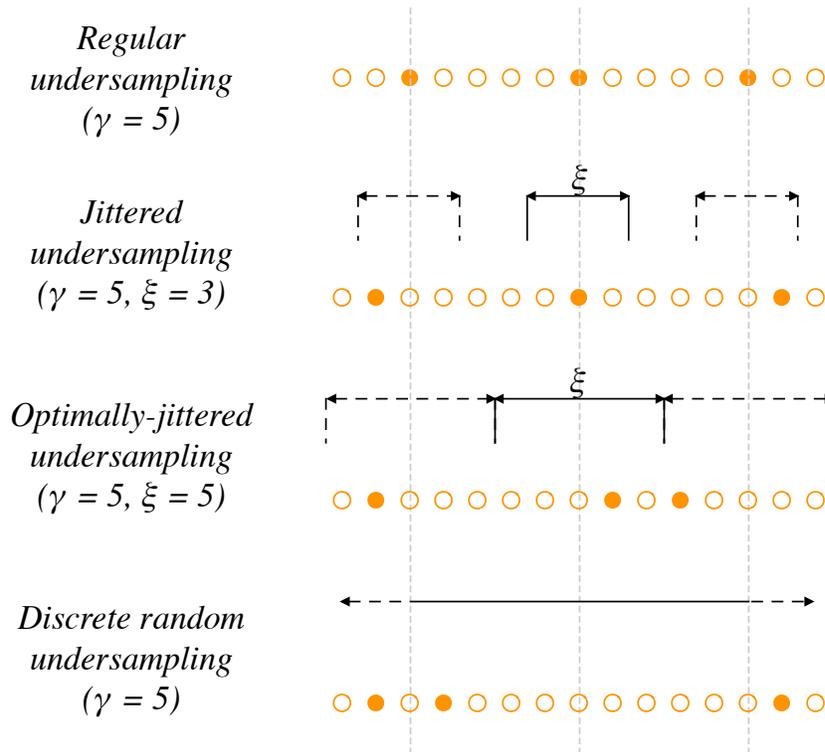
Figure 3: Schematic comparison between different undersampling schemes. The circles define the fine grid on which the original signal is alias-free. The solid circles represent the actual sampling points for the different undersampling schemes. The jitter parameter ξ relates to how far the actual jittered sampling point can be from the regular coarse grid, effectively controlling the size of the maximum acquisition gap.
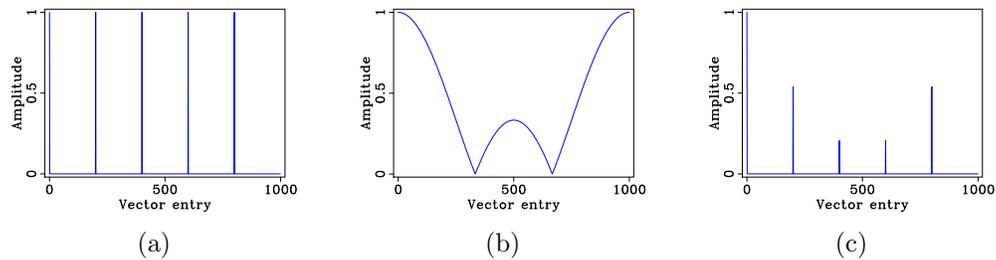

Figure 4: Amplitude spectrum of (a) a five-fold ($\gamma = 5$) regular undersampling vector, (b) a three-sample wide uniform distribution ($\xi = 3$), and (c) the resulting jittered undersampling vector. The first half of the vectors contains the positive frequencies starting with zero, the second half contains the negative frequencies in decreasing order.

The above expression corresponds to an elementwise multiplication of the periodic Fourier spectrum of the discrete regular sampling vector with a sinc function. This sinc function follows from the Fourier transform of the probability density function for the perturbation with respect to a point of the regularly decimated grid.

In Figure 4 the amplitudes for this Fourier-domain multiplication are plotted for jittered undersampling with $\gamma = 5$ and $\xi = 3$, i.e., on average four-out-of-five samples are missing for a jitter that includes the decimated grid point, one sample on the right and one sample on the left (cf. Figure 3, second row).

Equation 6 is a special case of the result for jittered undersampling according to an arbitrary distribution introduced by Leneman (1966) and further detailed in Appendix A. Because these results were originally derived for the continuous case, the above expression is approximate. In practice, however, this formula proves to be accurate, an observation corroborated by numerical results presented below. Consider the following cases for a fixed undersampling factor γ.

Regular undersampling ($\xi = 0$): As observed from the first row of Figure 3, there is no jitter in this case and equation 6 becomes

$$\mathbf{a}[k] = \begin{cases} n, & \text{for} \quad k = 1 + l \cdot n \quad \text{with} \quad l = 0, \cdots, \gamma - 1 \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

The undersampling artifacts $\mathbf{z}$ consist of aliased energy.

Optimally-jittered undersampling ($\xi = \gamma$): Now the sampling points are perturbed within contiguous windows, as depicted in the third row of Figure 3, and equation 6 reduces to

$$\mathrm{E}\{\mathbf{a}[k]\} \approx \begin{cases} n, & \text{for} \quad k = 1 \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

In this special case, the cause of the aliases is removed by the zeros of the sinc function. As with random undersampling according to a discrete uniform distribution, the off-diagonals of the matrix $\mathbf{A}^H\mathbf{A}$ (cf. Figure 5(b) and 2(c)) are random, turning aliases into noise. Again, the kernel of $\mathbf{L}$ does not contain coherent energy, as observed in Figure 5(d), for a five-fold undersampling ($\gamma = 5$) and a jitter parameter of $\xi = 5$. In that sense, this specific relation between the jitter parameter and the undersampling factor is optimal because it creates the most favorable conditions for recovery with a localized transform.

Jittered undersampling ($0 < \xi < \gamma$): In this regime, both coherent aliases and incoherent random undersampling noise are present. Depending on the choice for the jitter parameter, the energy either localizes or randomly spreads across the spectrum. Again, the reduction of the aliases is related to the locations of the zero crossings of the sinc function that move as a function of ξ. As ξ increases, the zeros move closer to

the aliases. As expected, the matrix $\mathbf{A}^H\mathbf{A}$, plotted in Figure 5(a), still contains the imprint of coherent off-diagonals, resulting in a kernel of $\mathbf{L}$, included in Figure 5(c), that is a superposition of coherent aliases and incoherent random noise. Although this regime reduces the aliases, coherent energy remains in the undersampling artifacts. This residue creates a situation that is less favorable for recovery. Depending on the relative strength of the aliases compared to the magnitude n of the diagonal of $\mathbf{A}^H\mathbf{A}$, recovery becomes increasingly more difficult, an observation that can be established experimentally.

In the next section, a series of controlled experiments is conducted to compare the recovery from regularly, randomly according to a discrete uniform distribution and optimally-jittered undersamplings.
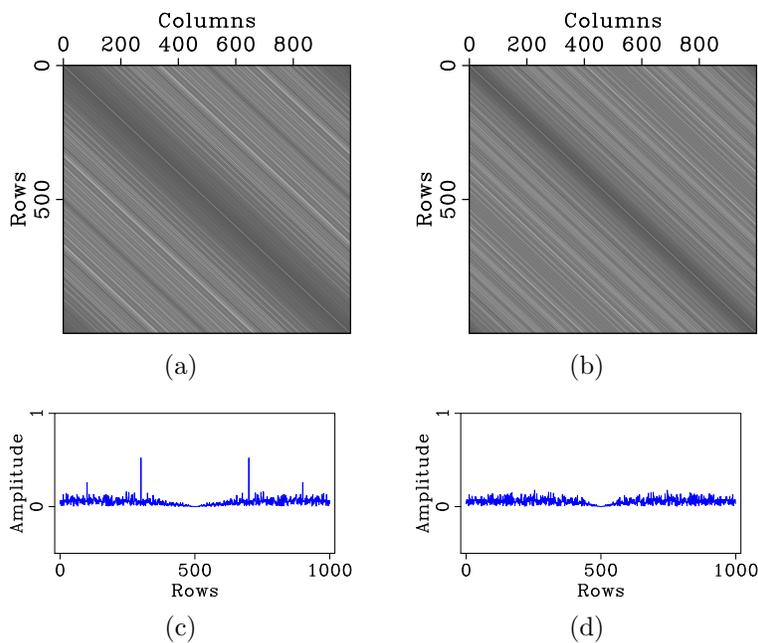
Figure 5: Jittered undersampling according to a discrete uniform distribution. (a) Suboptimal and (b) optimal jittered five-fold undersampling convolution matrices (in amplitude). The respective convolution kernels (in amplitude) that generate spectral leakage are plotted in (c) and (d). If the regular undersampling points are not shuffled enough, only part of the undersampling artifacts energy is spread, the rest of the energy remaining in weighted aliases. When there is just enough shuffling, all the undersampling artifacts energy is spread making jittered undersampling like random undersampling, yet controlling the size of the largest gap between two data points.

Controlled recovery experiments for different sampling schemes

With the favorable sampling schemes identified, it remains to be shown that these samplings lead to an improved recovery compared to the unfavorable regular under-

sampling. In particular, we want to experimentally confirm that jittered undersampling behaves similarly as random undersampling according to a discrete uniform distribution.

For this purpose, we define the sparsifying transform $\mathbf{S}$ as the Fourier transform $\mathbf{F}$, i.e., $\mathbf{S} \stackrel{\text{def}}{=} \mathbf{F}$, and generate a vector $\mathbf{x}_0$ with k nonzero entries and of length $N = 600$. The nonzero entries of $\mathbf{x}_0$ are distributed at random with random signs and amplitudes. The to-be-recovered signal $\mathbf{f}_0$ is given by $\mathbf{f}_0 = \mathbf{S}^H \mathbf{x}_0$ and the observations $\mathbf{y}$ are obtained by undersampling $\mathbf{f}_0$ regularly, randomly according to a discrete uniform distribution, or optimally-jittered, i.e., $\xi = \gamma$. Finally, the estimated spectrum $\tilde{\mathbf{x}}$ of $\mathbf{f}_0$ is obtained by solving equation 3 with the Spectral Projected Gradient for ℓ_1 solver (SPGL1 - van den Berg and Friedlander, 2007). Keep in mind that the number k of nonzero entries of $\mathbf{x}_0$ is not known a priori. Each experiment is repeated 100 times for the different undersampling schemes and for varying undersampling factors γ, ranging from 2 to 6. The reconstruction error is the number of entries at which the estimated representation $\tilde{\mathbf{x}}$ and the true representation $\mathbf{x}_0$ of $\mathbf{f}_0$ in the Fourier domain disagree by more than 10^{-4}. This error accounts for both false positives and false negatives. The averaged results for the different experiments are summarized in Figures 6(a), 6(b), and 6(c), which correspond to regular, random, and optimally-jittered undersampling, respectively. The horizontal axes in these plots represent the relative underdeterminedness of the system, i.e., the ratio of the number k of nonzero entries in $\mathbf{x}_0$ to the number n of acquired data points. The vertical axes represent the average reconstruction error. The different curves represents the different undersampling factors. In each plot, the curves from top to bottom correspond to $\gamma = 2, \ldots, 6$.

Figure 6(a) shows that, regardless of the undersampling factor, there is no range of relative underdeterminedness for which $\mathbf{x}_0$ can accurately be recovered from a regular undersampling of $\mathbf{f}_0$. Sparsity is not enough to discriminate the signal components from the spectral leakage. The situation is completely different in Figures 6(b) and 6(c) for the random and optimally-jittered sampling. In this case, one can observed that exact recovery is possible for $0 < k/n \lesssim 1/4$. The main purpose of these plots is to qualitatively show the transition from successful to failed recovery. The quantitative interpretation for these diagrams to the right of the transition is less well understood but also observed in phase diagrams published in the literature (Donoho et al., 2006). A possible explanation for the observed behavior of the error lies in the nonlinear behavior of the solvers and on an error not measured in the ℓ_2 sense.

The key observations from these experiments are threefold. First, it is possible, under specific conditions, to *exactly* recover by sparsity-promoting inversion the original spectrum $\mathbf{x}_0$ of $\mathbf{f}_0$ from (very) few data points. Secondly, optimally-jittered undersampling behaves like random undersampling according to a discrete uniform distribution. For practical purposes, the former can thus be seen as equivalent to the latter. Thirdly, not all undersampling schemes for a given undersampling factor are comparable from a CS perspective. Regular undersampling is the most challenging. Random and optimally-jittered undersamplings according to a discrete uniform distribution are among the most favorable. In particular, if the signal is sufficiently

sparse, these schemes lead to a reconstruction *as good as* dense regular sampling. Translated to the reconstruction of seismic wavefields, these results hint at a new nonlinear sampling theory based on a sparsifying transform for complex seismic data, e.g., the curvelet transform, and a coarse random sampling scheme that creates favorable recovery conditions for that transform, e.g., optimally-jittered undersampling.

APPLICATION TO SEISMIC DATA

Following recent work on Curvelet Reconstruction with Sparsity-promoting Inversion (CRSI - Herrmann and Hennenfent, 2007), seismic wavefields are reconstructed via $\tilde{\mathbf{f}} = \mathbf{C}^H \tilde{\mathbf{x}}$ where

$$\tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} ||\mathbf{x}||_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{RC}^H \mathbf{x}. \tag{9}$$

In this formulation, $\mathbf{C}$ is the discrete wrapping-based curvelet transform (Candès et al., 2005a). Similarly to any other data-independent transforms, curvelets do not provide a sparse representation of seismic data in the strict sense. Instead, the curvelet transform provides a compressible, arguably the most compressible (Hennenfent and Herrmann, 2006), representation. Compressibility means that most of the wavefield energy is captured by a few significant coefficients in the sparsifying domain. Since CS guarantees, for sparse-enough signal representations, the recovery of a fixed number of largest coefficients for a given undersampling factor (Candès et al., 2005b), a more compressible representation yields a better reconstruction, which explains the success of CRSI.

Synthetic data example

Figure 7(a) shows a synthetic dataset sampled above Nyquist rate along both the time and receiver axes. The corresponding amplitude spectrum is plotted in Figure 7(b). These two figures serve as references. Comparisons are made between the interpolation results of three-fold spatially undersampled data, collected either regularly or optimally-jittered. As expected, the amplitude spectrum (Figure 8(c)) of the regularly undersampled data (Figure 8(a)) is severely aliased. Unfortunately, these coherent f-k undersampling artifacts remain coherent in the curvelet domain and hence create a challenge for the reconstruction. To the contrary, there is no observable coherent spectral leakage in the amplitude spectrum (Figure 8(d)) for the optimally-jittered undersampled data (Figure 8(b)). Instead, the amplitude spectrum looks noisy in the temporal frequency band of the seismic signal.

Figure 9 shows the CRSI results for these two experiments. Figures 9(a) and 9(b) depict the reconstructions given data regularly (Figure 8(a)) and optimally-jittered (Figure 8(b)) sampled, respectively. Figures 9(c) and 9(d) represent the corresponding amplitude spectra. Unlike Figure 9(d) that is only slightly corrupted
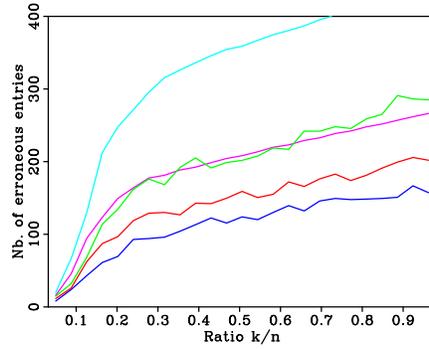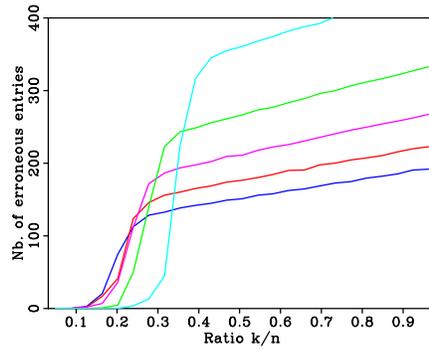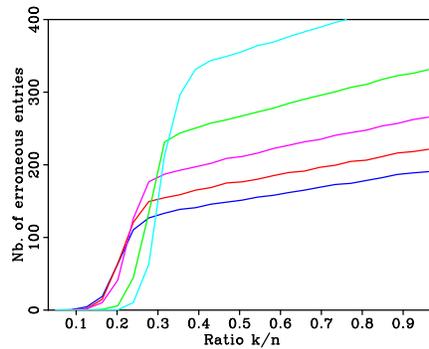
Figure 6: Averaged recovery errors for a *k*-sparse Fourier vector reconstructed from *n* time samples taken (a) regularly, (b) randomly, and (c) optimally jittered from the model. In each plot, the curves from top to bottom correspond to an undersampling factor ranging from two to six, i.e., $\gamma = 2, \ldots, 6$.

by incoherent errors, Figure 9(c) still contains substantial energy from the coherent undersampling artifacts. This observation is corroborated by the respective signal-to-reconstruction-error-ratios of 6.91 dB and 10.42 dB. The signal-to-reconstruction-error-ratio, defined as $20 \cdot \log_{10}(\|\mathbf{f}_0\|_2/\|\mathbf{f}_0 - \tilde{\mathbf{f}}\|_2)$, accounts for the energy of the error but not its type. It is important to keep in mind that the difference in reconstruction quality is solely due to the difference in spatial sampling, the undersampling factor and the recovery procedure were kept the same. This behavior leads us to conclude that, for a given undersampling factor, spatial optimally-jittered undersampling is (much) more favorable for CRSI than regular undersampling.

In addition, Figure 10 shows a recovery experiment given randomly three-fold spatially undersampled data. Figure 10(a) depicts the simulated acquired data and Figure 10(b) the CRSI result. The signal-to-reconstruction-error-ratio is 9.72 dB. Figures 10(c) and 10(d) contain the corresponding amplitude spectra. As can be observed by comparing Figure 8(d) with Figure 10(c), both random and optimally-jittered samplings create favorable recovery conditions. However, the larger size of the acquisition gaps in randomly undersampled data deteriorates the overall performance of CRSI. This result corroborates the importance of controlling the size of the maximum gap in optimally-jittered undersampling for reconstruction with curvelets.

Field data example

The far-offsets of a regularly-sampled shot taken from a real marine dataset are considered. Our model consists of 255 traces separated by 6.25 m. The simulated data are obtained by three-fold undersampling this model either regularly (Figure 11(a)) or optimally-jittered (Figures 11(d)). In both cases, the nominal spatial sampling is 18.75 m. Again, the CRSI algorithm is applied (cf. equation 9). No assumption is made regarding the maximum dip in the data. Figures 11(b) and 11(e) show the CRSI results for the data plotted in Figures 11(a) and 11(d), respectively. Figures 11(c) and 11(f) show the differences scaled by a factor of four between the model and the CRSI results. The signal-to-reconstruction-errors are respectively 12.98 dB and 15.22 dB, which corroborates our observations from the synthetic data example. The performance of wavefield reconstruction by CRSI improves when the input data is optimally-jittered sampled.

DISCUSSION

Undersampled data contaminated by noise

Although we focused on a noise-free (severely) underdetermined system of linear equations, the CS theory, and hence our work, both extend to the recovery from under-sampled data contaminated by noise (Candès et al., 2005b). In this case, the noise $\mathbf{e}$
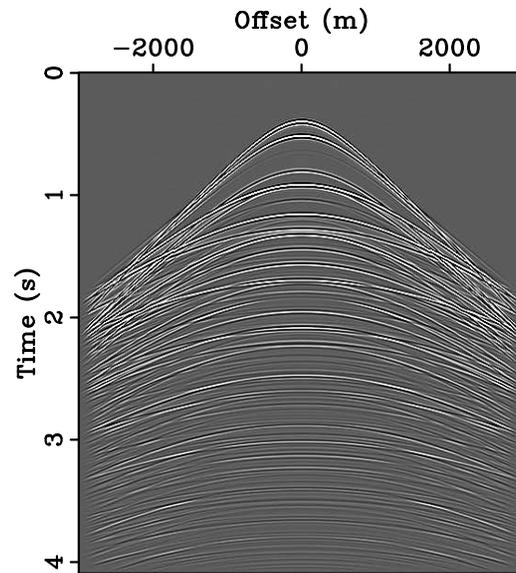
(a)
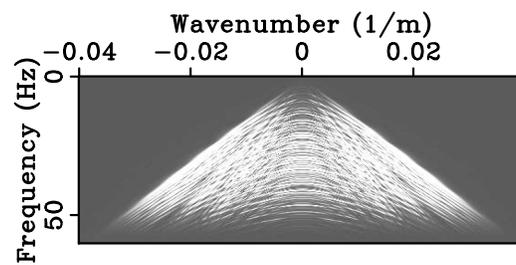
(b)

Figure 7: Reference model. (a) Synthetic data sampled above Nyquist rate and (b) corresponding amplitude spectrum.
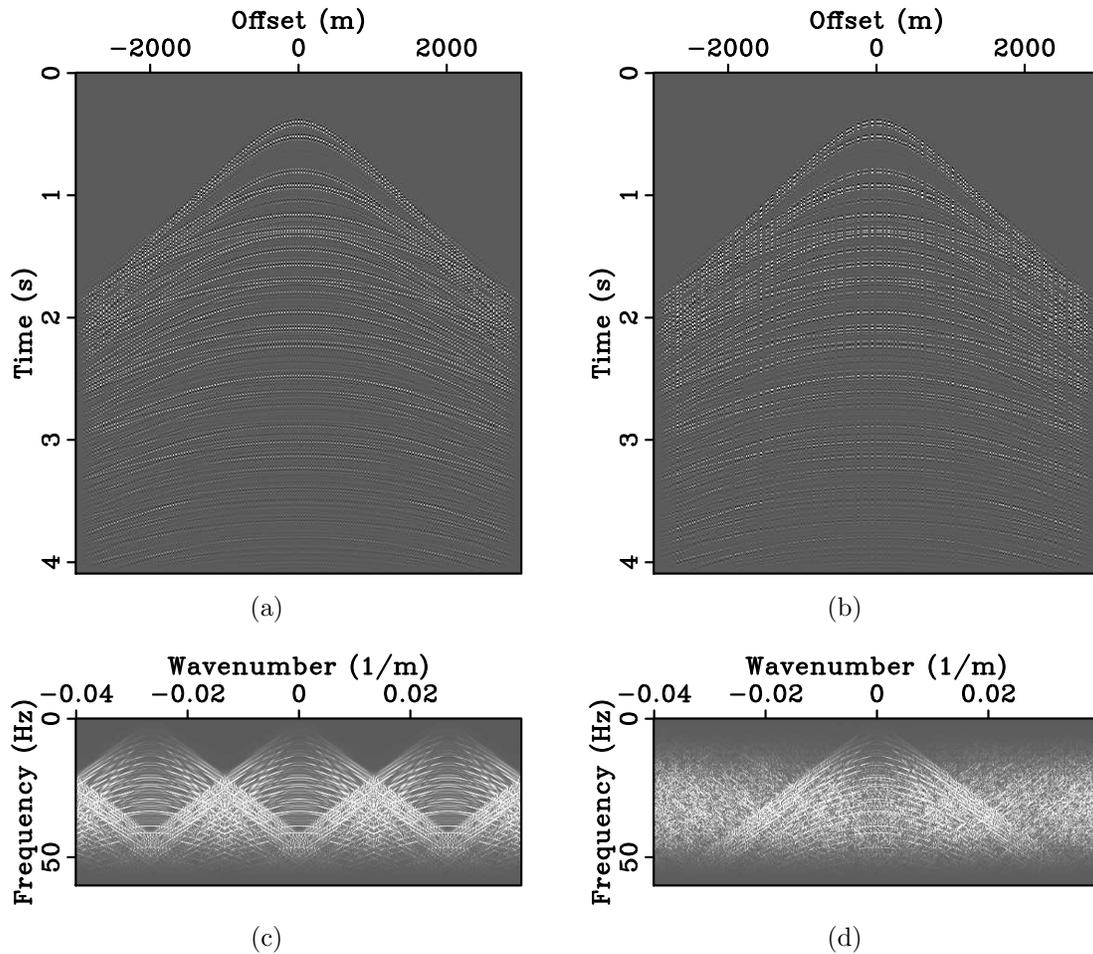
Figure 8: Synthetic data of Figure 7 (a) regularly and (b) optimally-jittered three-fold undersampled along the spatial axis. Their respective amplitude spectra are plotted in (c) and (d). For the same amount of acquired data, optimally-jittered undersampling turns the harmful coherent undersampling artifacts of regular undersampling, i.e., aliases, into incoherent random noise.

(a) (b)

(c) (d)
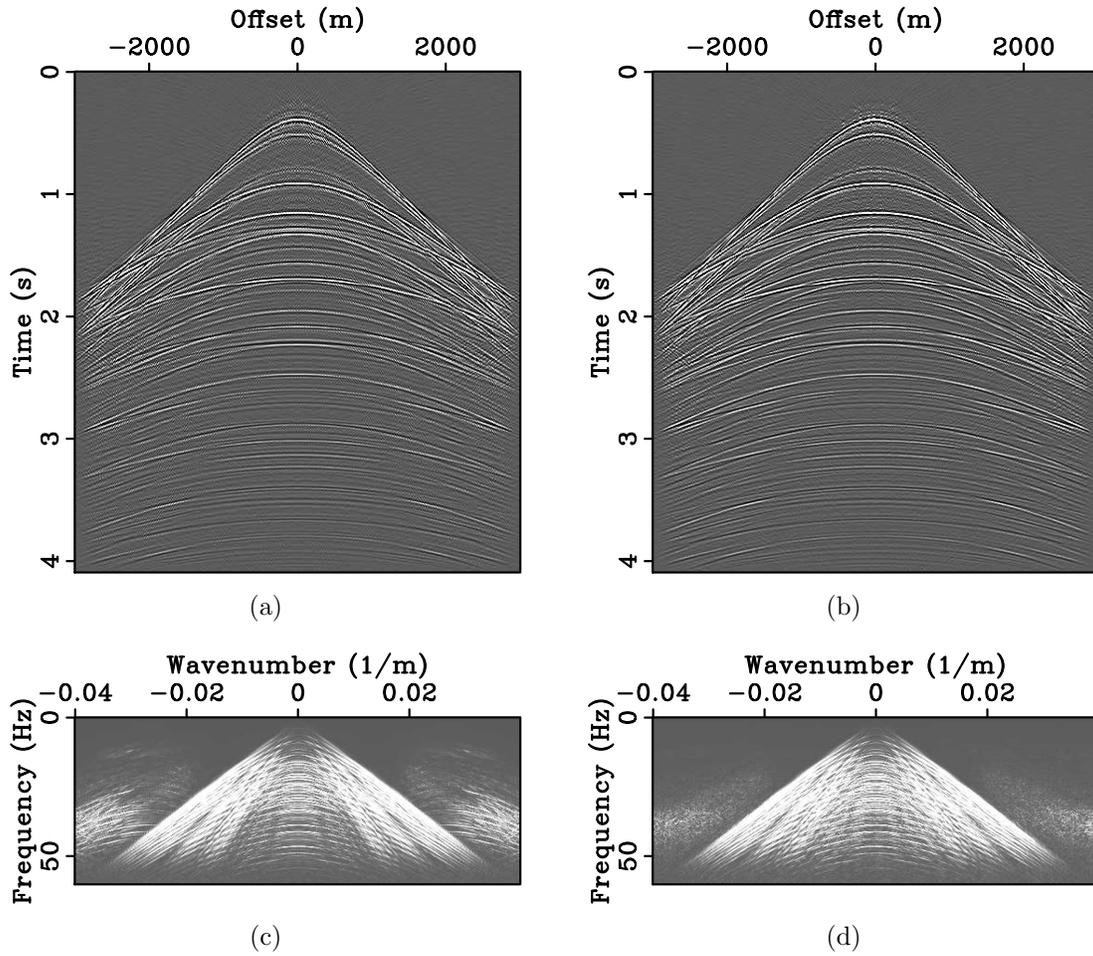
Figure 9: Curvelet reconstructions with sparsity-promoting inversion. Results given (a) data of Figure 8(a) and (b) data of Figure 8(b). The respective signal-to-reconstruction-error-ratios are 6.91 dB and 10.42 dB. For the same amount of data collected in the field, the reconstruction from optimally-jittered undersampled data is much more accurate than the reconstruction from regularly undersampled data.
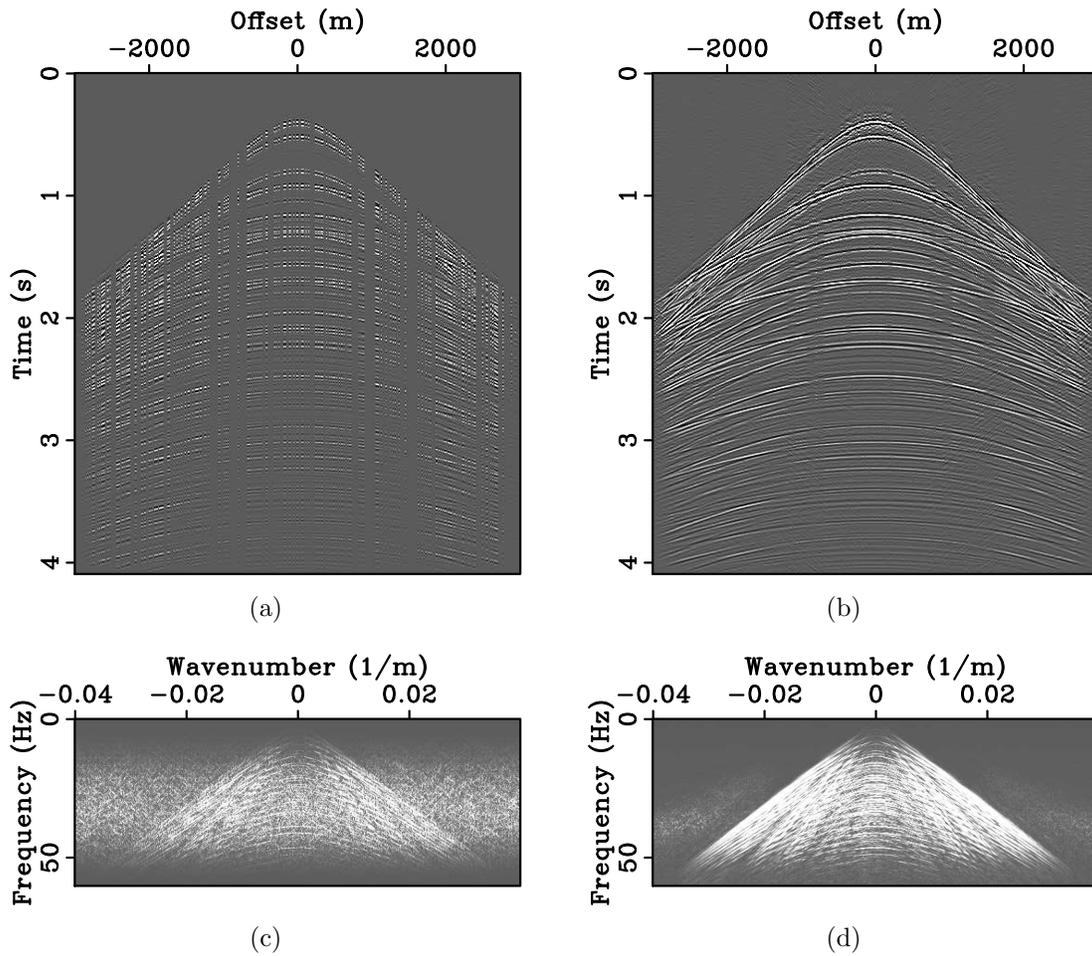
Figure 10: Randomly undersampled data and curvelet reconstruction with sparsity-promoting inversion. (a) Synthetic data randomly three-fold undersampled along the spatial axis and (b) curvelet reconstruction with sparsity-promoting inversion. Their respective amplitude spectra are plotted in (c) and (d). The signal-to-reconstruction-error-ratio is 9.72 dB. Although random and optimally-jittered undersamplings create similar favorable recovery conditions (compare (c) with Figure 8(d)), the larger size of the acquisition gaps in the randomly undersampled data deteriorates the overall performance of CRSI.
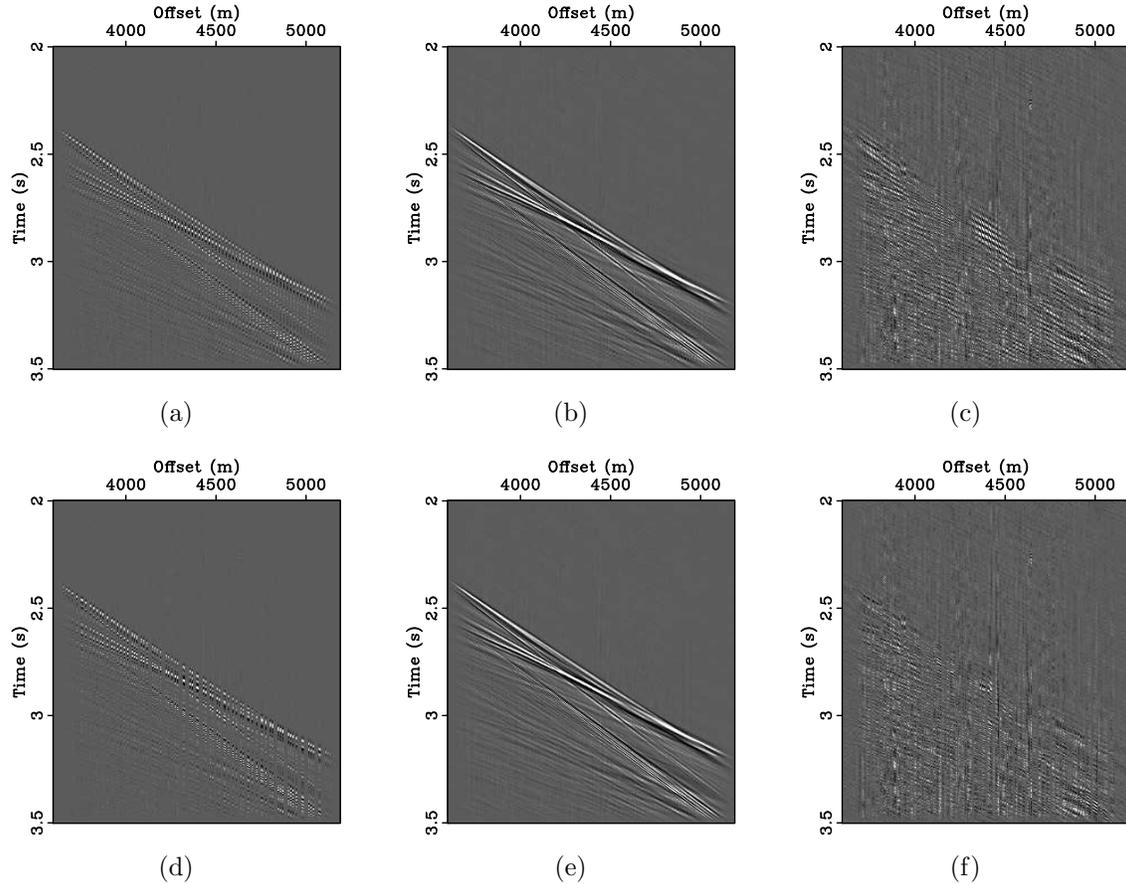
Figure 11: Field data example. The original data (not shown) is either (a) regularly or (d) optimally-jittered three-fold undersampled along the spatial coordinate. (b) and (e) are the curvelet reconstructions with sparsity-promoting inversion given data depicted in (a) and (d), respectively. (c) and (f) are differences scaled by a factor of four between the original data and the CRSI results (b) and (e), respectively. The corresponding signal-to-reconstruction-error-ratios are 12.98 dB and 15.22 dB, which corroborates that optimally-jittered undersampling is more favorable than regular undersampling.

that corrupts the data adds to the undersampling artifacts in the sparsifying domain. The quantity that relates to the recoverability is now given by $\mathbf{A}^H(\mathbf{A}\mathbf{x}_0 + \mathbf{e}) - \alpha\mathbf{x}_0$ as opposed to $\mathbf{A}^H\mathbf{A}\mathbf{x}_0 - \alpha\mathbf{x}_0$ in the noise-free case. Consequently, the undersampling artifacts $\mathbf{z}$ and the imprint of the contaminating noise in the sparsifying domain, i.e., $\mathbf{A}^H\mathbf{e}$, have to be studied jointly.

From discrete to continuous spatial undersampling

So far, undersampling schemes based on an underlying fine interpolation grid were considered. This situation typically occurs when binning continuous randomly-sampled seismic data into small bins that define the fine grid used for interpolation. Despite the error introduced in the data, binning presents some computational advantages since it allows for the use of fast implementations of Fourier or Fourier-related transforms, e.g., FFTW (Frigo and Johnson, 1998) or FDCT (Candès et al., 2005a). However, binning can lead at the same time to an unfavorable undersampling scheme, e.g., regular or poorly-jittered. In this case, one should consider working on the original data with, e.g., an extension to the curvelet transform for irregular grids (Hennenfent and Herrmann, 2006). Despite the extra computational cost for the interpolation, continuous random sampling typically leads to improved interpolation results because it does not create coherent undersampling artifacts (Xu et al., 2005).

Sparsity-promoting solvers and jittered undersampling

The applicability of CS to the large-scale problems of exploration geophysics heavily relies on the implementation of an efficient ℓ_1 solver. Despite several recent attempts to overcome this bottleneck (Tibshirani, 1996; Figueiredo et al., 2007; van den Berg and Friedlander, 2007), a wide range of large-scale applications still uses approximate ℓ_1 solvers such as iterated re-weighted least-squares (IRLS - Gersztenkorn et al., 1986), stage-wise orthogonal matching pursuit (StOMP - Donoho et al., 2006), and iterative soft-thresholding with cooling (Hennenfent and Herrmann, 2005; Herrmann and Hennenfent, 2007) derived from Daubechies et al. (2004). The success and/or efficiency of these approximate solvers depends upon the implicit-or-explicit assumption that the MAI is incoherent. Because optimally-jittered undersampling creates such a MAI, these solvers can be used for the sparsity-promoting reconstruction with curvelets or other localized Fourier-based transforms. More importantly, jittered undersampling can be useful to evaluate the efficiency/robustness of (approximate) ℓ_1 solvers since the jitter parameter controls the amount of coherent energy that enters the MAI.

Generalization of the concept of undersampling artifacts

Undersampling artifacts are only one particular case of MAI that specifically occurs in the interpolation problem, i.e., $\mathbf{A} \overset{\text{def}}{=} \mathbf{RS}^H$. The study we have done on these artifacts as a function of the restriction operator $\mathbf{R}$ can be extended to more general cases (see e.g. Lustig et al., 2007, in magnetic resonance imaging). For example, when $\mathbf{A}$ is defined as $\mathbf{A} \overset{\text{def}}{=} \mathbf{RMS}^H$ with $\mathbf{M}$ a modeling/demigration-like operator (Herrmann et al., 2007; Wang and Sacchi, 2007). In this case, $\mathbf{x}_0$ is the sparse representation of the Earth model in the $\mathbf{S}$ domain and $\mathbf{y}$ incomplete seismic data. The study of the MAI now determines which coarse spatial sampling schemes are more favorable than others in the context of sparsity-promoting migration/inversion. Based on observations in Zhou and Schuster (1995) and Sun et al. (1997), we believe that discrete random, optimally-jittered, and continuous random undersamplings will also play a key role.

CONCLUSIONS

Successful wavefield recovery depends on three key factors, namely, the existence of a sparsifying transform, a favorable sampling scheme and a sparsity-promoting recovery method. In this paper, we focused on an undersampling scheme that is designed for localized Fourier-like signal representations such as the curvelet transform. Our scheme builds on the fundamental observation that irregularities in sub-Nyquist sampling are good for nonlinear sparsity-promoting wavefield reconstruction algorithms because they turn harmful coherent aliases into relatively harmless incoherent random noise. The interpolation problem effectively becomes a much simpler denoising problem in this case.

Undersampling with a discrete random uniform distribution lacks control on the maximum gap size in the acquisition, which causes problems for transforms that consist of localized elements. Our jittered undersampling schemes remedy this lack of control, while preserving the beneficial properties of randomness in the acquisition grid. Our numerical findings on a stylized series of experiments confirm these theoretically-predicted benefits.

Curvelet-based wavefield reconstruction results from jittered undersampled synthetic and field datasets are better than results obtained from regularly decimated data. In addition, our findings indicate an improved performance compared to traces taken randomly according to an uniform distribution. This is a major result, with wide ranging applications, since it entails an increased probability for successful recovery with localized transform elements. In practice, this translates into more robust wavefield reconstruction.

ACKNOWLEDGMENTS

We also appreciate the valuable comments and suggestions from the two reviewers and two associate editors.

APPENDIX A

JITTERED UNDERSAMPLING

For the convenience of the reader, we re-derive the result originally introduced by Leneman (1966) that leads to equation 6.

Jittered sampling locations r_n are given by

$$r_n = n\gamma + \varepsilon_n \quad \text{for} \quad n = -\infty, \dots, \infty \tag{A-1}$$

The continuous random variables ε_n are independent and identically distributed (iid) according to a probability density function (pdf) p on $[-\zeta/2, \zeta/2]$. The corresponding sampling operator s is given by

$$s(r) = \sum_{n=-\infty}^{\infty} \delta(r - r_n). \tag{A-2}$$

Computing the Fourier transform of the previous expression yields

$$\hat{s}(f) = \frac{1}{\gamma} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{\gamma}\right) e^{-i2\pi f \varepsilon_n} \tag{A-3}$$

which implies that

$$\mathrm{E}\{\hat{s}(f)\} = \mathrm{E}\left\{e^{-i2\pi f \varepsilon_0}\right\} \cdot \frac{1}{\gamma} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{\gamma}\right) \tag{A-4}$$

since the variables ε_n are iid. By definition, the expected value of $e^{-i2\pi f \varepsilon_0}$ is given by

$$\mathrm{E}\left\{e^{-i2\pi f \varepsilon_0}\right\} = \int_{-\zeta/2}^{\zeta/2} p(t) \cdot e^{-i2\pi f t} \mathrm{d}t \tag{A-5}$$

which is the Fourier transform of the pdf of ε_0. Hence,

$$\mathrm{E}\{\hat{s}(f)\} = \hat{p}(f) \cdot \frac{1}{\gamma} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{\gamma}\right). \tag{A-6}$$

Finally, for a pdf that is continuous uniform on $[-\zeta/2, \zeta/2]$, the expected spectrum of the sampling operator is

$$\mathrm{E}\{\hat{s}(f)\} = \mathrm{sinc}(f\zeta) \cdot \frac{\zeta}{\gamma} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{\gamma}\right). \tag{A-7}$$

This result leads us to equation 6 since the columns of $\mathbf{A}^H \mathbf{A}$ are circular-shifted versions of the Fourier transform of the discrete jittered sampling vector, i.e., $\mathrm{diag}(\mathbf{R}^H \mathbf{R})$.

REFERENCES

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

Bednar, J. B., 1996, Coarse is coarse of course unless...: The Leading Edge, **15**, 763 – 764.

Biondi, B., S. Fomel, and N. Chemingui, 1998, Azimuth moveout for 3D prestack imaging: Geophysics, **63**, 1177 – 1183.

Candès, E. J., L. Demanet, D. L. Donoho, and L. Ying, 2005a, Fast discrete curvelet transforms: Multiscale Modeling and Simulation, **5**, 861–899.

Candès, E. J. and J. Romberg, 2006, Quantitative robust uncertainty principles and optimally sparse decompositions: Foundations of Computational Mathematics, **6**, 227 – 254.

Candès, E. J., J. Romberg, and T. Tao, 2005b, Stable signal recovery from incomplete and inaccurate measurements: Communications on Pure and Applied Math, **99**, 1207 – 1223.

———, 2006, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information: IEEE Transactions on Information Theory, **52**, 489 – 509.

Canning, A. and G. H. Gardner, 1996, Regularizing 3D data-sets with DMO: Geophysics, **61**, 1103 – 1114.

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

Daubechies, I., M. Defrise, and C. De Mol, 2004, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint: Communications on Pure and Applied Mathematics, **LVII**, 1413 – 1457.

Dippe, M. and E. Wold, 1992, Stochastic sampling: theory and application: Progress in Computer Graphics, **1**, 1 – 54.

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

Donoho, D. L. and X. Huo, 2001, Uncertainty principles and ideal atomic decomposition: IEEE Transactions on Information Theory, **47**, 2845–2862.

Donoho, D. L., Y. Tsaig, I. Drori, and J.-L. Starck, 2006, Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit: Technical report, Stanford Statistics Department. (TR-2006-2. `http://stat.stanford.edu/~idrori/StOMP.pdf`).

Figueiredo, M. A. T., R. D. Nowak, and S. J. Wright, 2007, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems: Technical report, Instituto de Telecomunicacoes. (`http://www.lx.it.pt/~mtf/GPSR/Figueiredo_Nowak_Wright_twocolumn.pdf`).

Frigo, M. and S. G. Johnson, 1998, FFTW: An adaptive software architecture for the FFT: International Conference on Acoustics, Speech and Signal Processing, 1381 – 1384, IEEE.

Gersztenkorn, A., J. B. Bednar, and L. Lines, 1986, Robust iterative inversion for the one-dimensional acoustic wave equation: Geophysics, **51**, 357 – 369.

Hennenfent, G. and F. J. Herrmann, 2005, Sparseness-constrained data continuation with frames: Applications to missing traces and aliased signals in 2/3-D: SEG International Exposition and 75th Annual Meeting, 2162 – 2165.

———, 2006, Seismic denoising with non-uniformly sampled curvelets: Computing in Science and Engineering, **8**, 16 – 25.

Herrmann, F. J. and G. Hennenfent, 2007, Non-parametric seismic data recovery with curvelet frames: Technical report, UBC Earth & Ocean Sciences Department. (TR-2007-1. `http://slim.eos.ubc.ca/Publications/Public/Journals/CRSI.pdf`).

Herrmann, F. J., D. Wang, G. Hennenfent, and P. P. Moghaddam, 2007, Curvelet-based seismic data processing: a multiscale and nonlinear approach. (Accepted for publication in Geophysics. `http://slim.eos.ubc.ca/Publications/Public/Journals/curveletter.pdf`).

Leneman, O., 1966, Random sampling of random processes: Impulse response: Information and Control, **9**, 347 – 363.

Lustig, M., D. L. Donoho, and J. M. Pauly, 2007, Sparse MRI: The application of compressed sensing for rapid MR imaging: Magnetic Resonance in Medicine. (In press. `http://www.stanford.edu/~mlustig/SparseMRI.pdf`).

Sacchi, M. D., T. J. Ulrych, and C. J. Walker, 1998, Interpolation and extrapolation using a high-resolution discrete Fourier transform: IEEE Transactions on Signal Processing, **46**, 31 – 38.

Spitz, S., 1991, Seismic trace interpolation in the F-X domain: Geophysics, **67**, 890 – 794.

Stolt, R. H., 2002, Seismic data mapping and reconstruction: Geophysics, **67**, 890 – 908.

Sun, Y., G. T. Schuster, and K. Sikorski, 1997, A Quasi-Monte Carlo approach to 3-D migration: Theory: Geophysics, **62**, 918 – 928.

Tibshirani, R., 1996, Regression shrinkage and selection via the Lasso: Journal of the Royal Statistical Society, **58**, 267 – 288.

Trad, D. O., J. Deere, and S. Cheadle, 2005, Challenges for land data interpolation: Presented at the CSEG National Convention.

Trad, D. O. and T. J. Ulrych, 1999, Radon transform: beyond aliasing with irregular sampling: Presented at the Sixth International Congress of the Brazilian Geophysical Society.

Trad, D. O., T. J. Ulrych, and M. D. Sacchi, 2003, Latest view of sparse Radon transforms: Geophysics, **68**, 386–399.

van den Berg, E. and M. P. Friedlander, 2007, In pursuit of a root: Technical report, UBC Computer Science Department. (TR-2007-16. `http://www.optimization-online.org/DB_FILE/2007/06/1708.pdf`).

Verdu, S., 1998, Multiuser detection: Cambridge University Press.

Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, 1166 – 1177.

Wang, J. and M. D. Sacchi, 2007, High-resolution wave-equation amplitude-variation-with-ray-parameter (AVP) imaging with sparseness constraints: Geophysics, **72**, S11 – S18.

Xu, S., Y. Zhang, D. Pham, and G. Lambare, 2005, Antileakage Fourier transform

for seismic data regularization: Geophysics, **70**, V87 – V95.

Zhou, C. and G. T. Schuster, 1995, Quasi-random migration of 3-D field data: SEG Technical Program Expanded Abstracts, 1145 – 1148.

Zwartjes, P. M. and M. D. Sacchi, 2007, Fourier reconstruction of nonuniformly sampled, aliased data: Geophysics, **72**, V21–V32.