Design of two-dimensional randomized sampling schemes for curvelet-based sparsity-promoting seismic data recovery

Gang Tang[12] (gtangthu@gmail.com), Reza Shahidi[2] (rshahidi@gmail.com),

Jianwei Ma[1] (jma@tsinghua.edu.cn) and Felix J. Herrmann[2] (fherrmann@eos.ubc.ca)

[1]Insititute of Seismic Exploration, School of Aerospace, Tsinghua University, Beijing, China

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

ABSTRACT

The tasks of sampling, compression and reconstruction are very common and often necessary in seismic data processing due to the large size of seismic data. Curvelet-based Recovery by Sparsity-promoting Inversion, motivated by the newly developed theory of compressive sensing, is among the best recovery strategies for seismic data. The incomplete data input to this curvelet-based recovery is determined by randomized sampling of the original complete data. Unlike usual regular undersampling, randomized sampling can convert aliases to easy-to-eliminate noise, thus facilitating the process of reconstruction of the complete data from the incomplete data. Randomized sampling methods such as jittered sampling have been developed in the past that are suitable for curvelet-based recovery, however most have only been applied to sampling in one dimension. Considering that seismic datasets are usually higher dimensional and extremely large, in the present paper, we extend the 1D version of jittered sampling to two dimensions, both with underlying Cartesian and hexagonal grids. We also study separable and non-separable two dimensional jittered sampling, the former referring to the Kronecker product of two one-dimensional jittered samplings. These different categories of jittered sampling are compared against one another in terms of signal-to-noise ratio and visual quality, from which we find that jittered hexagonal sampling is better than jittered Cartesian sampling, while fully non-separable jittered sampling is better than separable sampling. Because in the image processing and computer graphics literature, sampling patterns with blue-noise spectra are found to be ideal to avoid aliasing, we also introduce two other randomized sampling methods, possessing sampling spectra with beneficial blue noise characteristics, Poisson Disk sampling and Farthest Point sampling. We compare these methods, and apply the introduced sampling methodologies to higher dimensional curvelet-based reconstruction. These sampling schemes are shown to lead to

better results from CRSI compared to the other more traditional sampling protocols, e.g. regular subsampling.

Keywords: curvelets, irregular sampling, seismic data, acquisition design, jittered sampling, blue noise

INTRODUCTION AND MOTIVATION

Seismic data volumes are usually extremely large, requiring immense amounts of storage space. They are also often incomplete, with traces missing due to complex acquisition conditions. Due to economical reasons (it can be costly to place many sources and receivers), we also wish to reduce the number of sources or receivers, i.e. sample only some traces or positions by design. Thus seismic data compression and recovery become very important issues.

Obviously, in general, if samples are added to an already selected sample set, we obtain higher-resolution reconstructed data, which is desirable. However on the other hand, in order to save measurement costs in the field, the fewer samples that are acquired the better. These two competing requirements lead us to design sampling methodologies that minimize the number of samples necessary while at the same time maintaining the quality of the reconstructed data volume from these samples.

For traditional uniform regular sampling, the Shannon/Nyquist sampling theorem states that it is necessary to sample at a frequency over at least twice the signal's maximum frequency magnitude. A newly developed theory called "compressed sensing" (CS) (Candès et al., 2006b; Donoho, 2006) provides new insights, and opens up the possibility of reconstructing compressible images or signals of scientific interest accurately from a number of samples far fewer than that dictated by the Nyquist rate, though now using irregular randomized sampling, and not regular sampling.

Irregular sampling of seismic images

Based on CS theory, a successful recovery method for seismic data, named Curvelet-based Recovery by Sparsity-promoting Inversion (CRSI), was developed by (Herrmann and Hennenfent, 2008). It is derived from a sparsifying transform, in this case the discrete curvelet transform (Candès et al., 2006a), in conjunction with an undersampling scheme that favours recovery. Curvelets have a wide variety of applications in many fields (Ma and Plonka, 2009), and pertinent to this paper, it has been demonstrated that curvelets are a very good choice for seismic data sets, with their localized plane-wave-like frame elements and well-documented sparsity for seismic data wavefronts (Candès et al., 2006a; Herrmann et al., 2007; Chauris and Nguyen, 2008). We therefore use the discrete curvelet transform for the sparsifying transform in CRSI.

When regular periodic undersampling is used as the underlying sampling scheme input to this reconstruction, performance is poor due to the presence of well-known periodic aliases (Zwartjes and Sacchi, 2007). On the other hand, randomized undersampling can render coherent aliases into easy-to-remove incoherent noise in the frequency domain, so that CRSI reconstruction becomes a simple denoising problem (Hennenfent and Herrmann, 2008). Expressions for the variance of this incoherent noise are given in the appendix of this paper. Unfortunately, commonly used discrete uniform random sampling cannot control gap lengths between missing traces, and this can greatly affect the quality of reconstruction, since there may be too much contiguous data that is unsampled, leading to a severely underdetermined problem which cannot be easily regularized to give a satisfactory solution. Thus, jittered sampling was introduced to mitigate this issue (Hennenfent and Herrmann, 2007, 2008), so that gap size is limited while at the same time there is enough randomness that aliases are

converted to noise.

But most other methods are only applicable to 1D sampling, i.e. they sample some traces along only one space axis of the data (the complete data is shown in Figure 1(a)), as in Figure 2(a). However, as we know, higher dimensional seismic data is more common, e.g., seismic volumes, which are always very large, necessitating a strategy to reconstruct them from as few traces as possible, in order to design seismic data acquisition with less shots and/or receivers (Tang et al., 2009; Zwartjes and Sacchi, 2007). Moreover, as mentioned in (Herrmann and Hennenfent, 2008), compared to the common shot/receiver reconstruction, a significant improvement can be achieved with shot-receiver interpolation, because the 3D geometry of seismic lines will be fully exploited by the 3D curvelet transform.

However, there exist only a few higher dimensional sampling strategies for curvelet-based seismic data recovery in the literature. (Herrmann and Hennenfent, 2008) used 2D uniform discrete random sampling while performing 3D interpolation, and though there was some improvement over 2D regular subsampling, there was still room for improvement in their obtained interpolated results. Knowing that jittered sampling is better than uniform discrete random sampling in one dimension, it is natural to extend it to two dimensions (László, 1995) in the hope that it would also be better than uniform random sampling in this case. In this paper, first we compare 2D jittered sampling on underlying Cartesian and hexagonal regular grids. In addition to this, we also introduce another two types of blue-noise sampling schemes, Poisson Disk sampling (Cook, 1986) and Farthest Point sampling (Eldar et al., 1997), in order to increase the choice of sampling schemes for CRSI, since such irregular sampling methods are commonly used in the computer graphics and image processing communities. These schemes can be used both for 1D and higher dimensional samplings.

In the context of seismic acquisition, sampling corresponds to the determination of source and receiver geophone locations. Traditionally, spacing these geophones regularly was common since this was an "easier" process to perform in the field. When a full seismic survey is done in two dimensions, in a traditional experimental setup, often source and receiver geophones are dragged along a straight line perpendicular to their common axis on cables and then shots fired from the source(s) at certain intervals as the truck (or boat in the case of marine data acquisition) travels along what is called the *inline* direction. The common axis for the source and receiver geophone is called the *crossline* direction, while as just stated the direction along which the truck travels is called the inline direction (Vermeer, 1990). The case where the inline and crossline directions are sampled regularly with the same sampling interval for each corresponds to the placement of receivers as in Figure 3(a). Typically the inline direction is sampled more densely than the crossline direction Long (2004), so in fact the sampling intervals along each axis could be different. There is also some choice as to where the cables are positioned behind the truck or boat, and this positioning can be done in an irregular fashion. As will be seen, introducing randomness leads to lesser magnitude, or at least less coherent aliasing, and thus to superior seismic data recovery from the incomplete data. In Figure 3(b) the positions of the cables are chosen according to a discrete uniform random distribution. Unfortunately, for the CRSI scheme used for reconstruction, large gaps in the incomplete data can lead to problems in the recovery. Therefore, in Figure 3(c), we show an example where the cable positions are chosen according to a jittered distribution. Also, (in this case) the vertical positions at which the shots are fired from the source can be regular or irregular. In Figure 3(d), both the cable positions and the positions where the shots are fired are from a jittered distribution with the same jittered distribution used for all receivers. It is conceivably possible for each receiver to follow independent jittered

distributions, and this case is illustrated in Figure 3(e).

Finally, at the present time, with the advent of new wireless portable Global Positioning Systems (GPSs), it is much more feasible than in the past to place sources and receivers at arbitrary locations on the area being surveyed. Figure 3(f) shows one such possible sampling configuration, where sample locations are not confined to lie along the cable paths. In this paper, we explore such fully two-dimensional sampling schemes in the context of seismic acquisition, and these schemes hold much promise for future acquisition design because of the advent of exciting new portable wireless positioning technologies.

Figure 1 shows the reference synthetic model we use in some of our experiments along with its Fourier spectrum.

In Figure 2(a), we show the incomplete data from 1D discrete uniform random sampling, where complete traces are taken at uniform random positions along the receiver and in Figure 2(b), incomplete data with fully 2D random sample positions is shown, where systems like GPS could be used to implement the sampling in practice.

Blue-noise and randomized sampling

Blue noise refers to a signal whose energy is concentrated at high frequencies with little energy concentrated at lower non-zero frequencies. Sampling patterns with blue-noise spectra, common in the field of image processing, have been proven to be able to scatter aliasing artifacts throughout the Fourier spectrum out of the signal band as high-pass noise, which is more easily filtered out (Dippé and Wold, 1985; Ignjatovic and Bocko, 2005). This feature is very effective in obtaining good reconstructions with some Fourier-based methods, and as we see in this paper for curvelet-based techniques. Jittered sampling, Poisson Disk sampling and

Farthest Point sampling are among the best schemes in the image processing and computer graphics literature yielding blue-noise spectra (Mitchell, 1987; Hennenfent and Herrmann, 2008). In the field of magnetic resonance imaging, (Lustig et al., 2009) have used Poisson Disk sampling to provide a locally uniform distribution for compressive sensing. Here, we will introduce all three of these randomized sampling schemes for seismic reconstruction by curvelet-based sparsity promotion.

In the experiments in this paper, we made a comparison between these schemes in terms of reconstruction quality, so that some comparisons of these samplings for curvelet-based seismic data reconstruction could be made. We intend to do a more complete study of Poisson Disk and Farthest Point sampling in seismic acquisition design in a future work, along with possibly other methods in a future work and adaptive versions of these methods, however the results we obtained in this paper are promising.

The main contributions of this paper are the generalization of the one-dimensional jittered sampling in (Hennenfent and Herrmann, 2008) to two dimensions. There is more than one way of doing this, depending on the underlying regular tiling of the plane, e.g. rectangular or hexagonal. We find that both give similar aliasing magnitudes, however hexagonal jittered sampling permits one to use a lower sampling density to sample data of the same bandwidth, due to its more efficient packing in the plane. We also experiment with some other two-dimensional randomized sampling methods found in the computer graphics and image processing literature, namely Poisson Disk sampling and Farthest Point sampling. We find that both these methods give similar to and slightly better results than two-dimensional hexagonal jittered sampling. All of these methods, by virtue of their aperiodicity and randomness, improve upon regular subsampling when used in conjunction with CRSI to reconstruct higher-dimensional seismic data.

CRSI RECONSTRUCTION METHOD

Restriction operator

Defining *vec* to be a vectorization operator on a 2D data matrix, and $\mathbf{A} \otimes \mathbf{B}$ to be the Kronecker product of the two matrices $\mathbf{A}$ and $\mathbf{B}$, one may formulate 2D Curvelet-based Recovery by Sparsity-Promoting Inversion (CRSI) reconstruction as an inverse problem. Though we limit the discussion here to CRSI in 2 dimensions, it is straightforward to generalize the method to higher dimensions. The reconstruction from an incomplete 2D seismic dataset follows the forward model

$$\mathbf{b} = \mathbf{R}vec(\mathbf{s}), \tag{1}$$

where $\mathbf{b} \in \mathbb{R}^p$ represents the vector of acquired incomplete 2D data with missing traces or positions, $\mathbf{s} \in \mathbb{R}^{m \times n}$ is the 2D signal, to be recovered, i.e. the adequately sampled data, and $\mathbf{R} \in \mathbb{R}^{p \times mn}$ is the restriction operator that collects the p acquired samples from $\mathbf{s}$, $mn \gg p$. Thus $\mathbf{R}$ is a sampling matrix, on which both the acquired data $\mathbf{b}$, and the recovery of model $\mathbf{m}$ depend.

If the restriction operator $\mathbf{R}$ can be expressed as the Kronecker product of two matrices $\mathbf{R}_1 \in \mathbb{R}^{p_1 \times m}$ and $\mathbf{R}_2 \in \mathbb{R}^{p_2 \times n}$, with $p_1 \cdot p_2 = p$, then we call the sampling associated with $\mathbf{R}$ a *separable* sampling method. Otherwise, we call the sampling method *non-separable*. If we sample only certain traces, but each trace consists of complete data, then $\mathbf{R} = \mathbf{R}_t \otimes \mathbf{I}_m$, with $\mathbf{R}_t$ the restriction to the sampled traces, and $\mathbf{I}_m$ the mxm identity matrix. The matrix $\mathbf{R}_t$ is a restriction operator corresponding to a 1D sampling of traces. This matrix can thus use any 1D sampling scheme, e.g. $\mathbf{R}_r$ for 1D regularly subsampling, $\mathbf{R}_{j1}$ for 1D jittered subsampling or $\mathbf{R}_u$ for 1D uniform random subsampling. It is also possible to do

completely 2D subsampling in which case $\mathbf{R}$ will not be expressible as a Kronecker product in a non-trivial way, e.g. $\mathbf{R}_{j2}$ for fully two-dimensional jittered sampling. Several choices of restriction matrix $\mathbf{R}$ are shown in Figure 3.

We will present more about sampling schemes that define the matrix $\mathbf{R}$ in the next section.

CRSI method

There is more than one possible solution of Equation (1) because any model that agrees with the incomplete data $\mathbf{y}$ after being restricted by the matrix $\mathbf{R}$ will satisrfy the equation – this is an underdetermined inverse problem. According to compressive sampling theory, it was suggested to reformulate the problem as follows:

$$\mathbf{b} = \mathbf{A}\mathbf{x} \quad \text{with} \quad \mathbf{A} \stackrel{\text{def}}{=} \mathbf{R}\mathbf{D}^H, \tag{2}$$

where $\mathbf{D}$ is a sparse transform and $\mathbf{D}^H$ is its adjoint – i.e., its conjugate transpose, and $\mathbf{x} \in \mathbb{R}^N$ with $N \gg n$ is the representation of $\mathbf{s}$ in the sparse domain.

In (Herrmann and Hennenfent, 2008), the discrete curvelet transform is adopted as the sparse transform, i.e., $\mathbf{D} \stackrel{\text{def}}{=} \mathbf{C}$, where $\mathbf{C}$ is the forward curvelet transform matrix, so Equation (2) becomes

$$\mathbf{b} = \mathbf{R}\mathbf{C}^H\mathbf{x} \tag{3}$$

The curvelet transform gives a compressible representation of $\mathbf{m}$, which in other words means that the vector $\mathbf{x}$ has few large and many small non-zero coefficients. These properties make it possible to successfully recover $\mathbf{m}$ according to the theory of *compressive sampling* (Candès et al., 2006b; Candès, 2006). However to solve this underdetermined problem, additional information must be provided to *regularize* the problem. The CRSI method

promotes sparsity in the curvelet domain as a regularization term and gives a solution to problem (3) by solving

$$\mathbf{P}_\sigma : \quad \begin{cases} \widetilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \\ \widetilde{\mathbf{s}} = \mathbf{C}^H \widetilde{\mathbf{x}}, \end{cases} \tag{4}$$

where $\|\mathbf{x}\|_1 \stackrel{\text{def}}{=} \sum_{i=1}^{N} |x_i|$ is the ℓ_1 norm. The recovered vector that solves $\mathbf{P}_\sigma$ is $\widetilde{\mathbf{x}}$ and $\widetilde{\mathbf{s}} \in \mathbb{R}^m$ is the estimate of the recovered data obtained by applying $\mathbf{C}^H$. Here we solve this problem with the Spectral Projected Gradient for ℓ_1 solver (SPGL1 - (Berg and Friedlander, 2008, 2007)).

RANDOMIZED SAMPLING SCHEMES

From traditional sampling theory, we know that, with regularly sampled points, aliasing will occur at frequencies higher than the Nyquist limit, due to the regular and periodic nature of the sampling. If we sample in an irregular manner to make the sizes of unsampled regions unequal, it is possible for these aliases to be converted into easy-to-remove noise. Randomized sampling is a way of achieving this, by distributing samples randomly, so that every point has a finite probability of being sampled (Dippé and Wold, 1985).

Discrete uniform random sampling, where each location on a grid has exactly the same probability of being chosen, can lead to poor reconstructions with CRSI since it cannot control the size of gaps between samples. Such uniform random sampling converts aliases into white noise in the frequency domain, from which in fact it can be very difficult to remove and distinguish the required spectrum if the noise is of too high an amplitude.

On the other hand, uniform jittered sampling first subdivides the space into n regions, with n the predetermined number of samples we wish to take, and then randomly takes one

sample in each region. Because each region is sampled, and the regions form a partition of the space (they are contiguous), the size of gaps can be controlled, and experimentally it is found that a blue-noise spectrum is obtained. (Hennenfent and Herrmann, 2008) introduced 1D jittered sampling into CRSI and proved its effectiveness, and in the current paper, we extend jitterd sampling to 2D.

Generalization of jittered sampling to 2D

For one-dimensional jittered sampling, there is only one degree of freedom, and that is the selection of the positions of the traces to be sampled. But when jittered sampling is extended to two dimensions, there are more choices for the underlying tiling of the field, e.g. Cartesian or hexagonal tiling. Cartesian tiling simply consists of the usual regularly spaced samples along each dimension, so this does not require further explanation. We now describe jittered sampling with hexagonal tiling in more detail. Both the Cartesian and hexagonal jittered samplings are non-separable sampling methods since they cannot be expressed as the Kronecker product of two one-dimensional samplings. Later in the text, we sometimes refer to such samplings as fully two-dimensional jittered samplings.

Hexagonal sampling

Regular hexagonal sampling grids have become popular due to their increased packing density (about 13.4% greater) over the usual regular Cartesian grids. Larger bandwidth signals can be sampled with hexagonal grids than with Cartesian grids if the same number of samples is used for each.

In two dimensions, Cartesian sampling of a function f can be described by the equation:

$$f_s^{Cart}(x,y) = \sum_{k=-\infty}^{\infty} \sum_{l=\infty}^{\infty} f(kT_x, lT_y)\delta(x - kT_x)\delta(y - lT_y), \tag{5}$$

with T_x and T_y the sampling intervals in x and y respectively, δ the continuous Dirac delta function, and assuming the sampling domain is all of $\mathbb{R}^2$. The resulting sample train f_s^{Cart} consists of impulses modulated by the function f's value at the corresponding impulse locations. In general, the horizontal and vertical sampling intervals, T_x and T_y may be different from each other, though for the purposes of this paper we only consider square Cartesian sampling, where both of these intervals are equal.

On the other hand, hexagonal sampling can be thought of as being the union of two Cartesian samplings staggered with respect to one another by half a sample. Such a sampling may be described by the equation:

$$f_s^{Hex}(x,y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f(2kT_x,\, lT_y)\, \delta(x - 2kT_x)\, \delta(y - lT_y) +$$
$$f\left((2k+1)T_x, \left(l + \frac{1}{2}\right) T_y\right) \delta(x - (2k+1)T_x)\, \delta\left(y - \left(l + \frac{1}{2}\right) T_y\right),$$

where now the resulting sample train is $f_s^{Hex}(x,y)$. Once again, we only consider regular hexagonal sampling, which in this case, means that $T_x = \frac{3W}{2}$ and $T_y = \sqrt{3}W$ for some positive constant W. This regular hexagonal sampling is shown in Figure 4.

Comparisons between Cartesian and hexagonal jittered sampling

Both Cartesian (rectangular) and hexagonal jittered strategies are shown in Fig. 5(a) and 5(b). First, the field is tiled with Cartesian or hexagonal grids, and then a random perturbation is made around the center of each tile while making sure the sample is kept inside its tile. To some extent, the sampling depends on the tiling itself because of the restriction that random perturbations keep a sample within a tile or a subset thereof. In

Figure 6, the theoretical formulae in the appendix of this paper are used to calculate the aliasing in the expected power spectrum of the jittered sampling process both for Cartesian jittered sampling and hexagonal jittered sampling. This aliasing is plotted as a function of the perturbation ratio, or simply the maximum allowed jitter ζ vs. the cell size γ. For Cartesian jittered sampling, the cell size is computed as twice the distance from the cell's centre to any edge of the cell's square perimeter, while for hexagonal jittered sampling, the cell size is calculated to be the diameter of the circle circumscribing the hexagonal cell. In this case, because the packing densities of Cartesian jittered and hexagonal jittered sampling are different, the cell sizes are chosen so that the same number of samples is taken for each scheme.

Jittered hexagonal undersampling ($0 < \zeta < \gamma$)

When $\zeta = 0$, we have no jitter, thus this case corresponds to regular hexagonal undersampling, and there are full aliasing replicas. As the value of ζ increases, the amplitudes of the aliases decreases. In Figure 6, we use the expressions derived in the appendix to plot the amount of aliasing for the expected spectra of hexagonal and Cartesian jittered sampling vs. the perturbation ratio $\beta = \frac{\zeta}{\gamma}$ as β ranges from 0 to 1. Note that for optimal (i.e. with $\zeta = \gamma$) jittered sampling the amount of aliasing for the expected spectrum is zero, however for any typical instance of optimal jittered sampling, the actual amount of aliasing will be non-zero due to the variance of the random jittered sampling spectrum. This variance is shown later in Figure 7.

As an example, we also select $\beta = 0.3$ for a domain of dimensions 200x200, and plot an instance of this random sampling spectrum in Figure 8(a). Observe that in this figure, aliases are present, but their amplitudes are less than that of the base signal spectrum.

Optimally-jittered hexagonal undersampling ($\zeta = \gamma$)

When the amount of jitter in each hexagonal cell is maximal, then it can be easily shown that the coefficients of all the impulses in Equation 17, as derived in the Appendix, are zero except the impulse corresponding to the DC component at the origin, with coefficient 1.

In other words, if we define the Gram matrix of the selection and sparsity operator

$$\mathbf{G} \stackrel{\text{def}}{=} \mathbf{A}^H \mathbf{A} \tag{6}$$

Except for the unit diagonal the expected value of each element of the Gram matrix is zero, or more simply, the (ignoring gridding effects) expected value of the Gram matrix is the identity matrix. However there will still be some aliasing since the variances of the entries of this matrix are non-zero. These variances (ignoring gridding effects) are given by the expression $\frac{8}{3\sqrt{3}\gamma^2}(1 - |\psi_Z(\mathbf{f})|^2)$, as derived in the appendix of this paper. The variances are not dependent on the jitter ζ, but are dependent on the spatial frequency vector $\mathbf{f} = (f_1, f_2)$. It is seen that for the zero-frequency, the variance of the spectrum is 0, meaning that the base signal replica will always be reproduced exactly without being scaled. An example spectrum for optimal jittered sampling is shown in Figure 8(b).

We also plot the sampling spectrum for an instance of optimal hexagonal jittered sampling in Fig. 8(a), where it can be seen that there are no aliases, but just wideband noise of low magnitude which may be easily filtered out.

Alternative progressive randomized sampling schemes

As stated in the recent work by (Grundland et al., 2009), jittered sampling attempts to replicate a blue-noise spectrum, which will yield good reconstruction results. There are other sampling methods that have spectra, which are more "blue" in the sense that there is

more of a clear threshold frequency below which the spectrum magnitude is insignificant (except for the zero frequency), and above which there is substantial spectral energy. Two such methods in that paper, and used in the larger computer graphics and image processing communities are Poisson Disk and Farthest Point sampling. However these methods are not tailored to CRSI reconstruction since they do not explicitly limit the distances between samples. Despite this, it is possible to find some bounds on these distances, so that there would be extremely large unsampled areas, which would hinder the performance of CRSI. Therefore, we also wish to look at these methods to determine their performance compared to the other sampling methods discussed in this paper.

Poisson Disk sampling selects n points at random iteratively, and only keeps a sample if it is at a sufficient distance away from all previously selected samples. This distance can be slowly shrunk as the number of samples, and hence the sampling density, increases. This also leads to a blue noise spectrum, as Fig. 9(a) and 9(c) show.

Farthest Point sampling is another irregular sampling scheme with excellent anti-aliasing properties, and is based on the computational geometry concept of the Voronoi diagram (Eldar et al., 1997). The main idea for that scheme, as suggested by its name, is to repeatedly place the next sample point to be the farthest point from all previously selected samples. This ensures that there are no regions of the image that are not adequately sampled, while at the same time, some randomness is maintained by initially selecting a small number of randomly selected seed samples. It can be proven that the farthest point from the previous samples is a vertex of the Voronoi diagram of those samples. So points are added one at a time to the sample set, and the Voronoi diagram updated incrementally. Please see Fig. 9(b) and 9(d).

For the purposes of this paper, we use a discrete approximation to Farthest Point sampling, which although takes more CPU time because of the fine grid spacing taken, is much simpler algorithmically. In our approximation algorithm, we maintain a distance array $\mathbf{D}$, each element of which contains the value of the distance squared of that discrete grid point to the closest already selected sample. At each iteration, the array element with maximum $\mathbf{D}$ value is chosen as the next sample $\mathbf{s}^{n+1} = (x^{n+1}, y^{n+1})$, and then the distance map updated using the formula:

$$\mathbf{D}^{n+1}(i,j) = \min(\mathbf{D}^n(i,j), (i - x^{n+1})^2 + (j - y^{n+1})^2),$$

for all array positions (i, j) in the image. If there is more than one maximal array element, then one is chosen at random. If the discrete grid is fine enough, this results in a good approximation to the continuous version of Farthest Point sampling, without dealing with the intricacies of computational geometry constructs such as the Voronoi diagram.

Both Poisson Disk and Farthest Point sampling fall under the category of *progressive* sampling methods, while jittered sampling does not. A sampling algorithm is called progressive if samples are selected one at a time, and if the sampling can be stopped after any desired number of samples is taken, still giving a good reconstruction result with this incomplete sampling. It can be easily seen from the descriptions of Poisson Disk and Farthest Point sampling above, that they both fit this definition. On the other hand, for jittered sampling a regular tiling is first taken, and then samples chosen within each tile. If we stop the sampling before all samples have been taken, then clearly there will be some tiles which are left unsampled, and as described earlier, this will lead to reconstructions of poor quality by CRSI.

Now we wish to translate what progressive and non-progressive samplings mean into

practical terms when sampling in the field for seismic data acquisition. In general, successive samples for a progressive sampling scheme will be relatively far apart, especially in the early stages of the sampling, since when one location has been sampled, we do not need to sample close to it. Rather, it is better to sample locations far away from all previously selected samples since at least in the case of a stationary image model, unsampled regions of images need to be sampled more than regions already sampled, since they are more likely to exhibit error when reconstructed. So if there is only one individual who is laying down the receivers using a GPS system as explained in the introduction of this paper, and we wish to maintain the progressive nature of the sampling so that it can be stopped at any time without significantly affecting reconstruction quality for that number of samples, they would have to travel a very large distance, zigzagging between successive samples as determined by one of these schemes.

A progressive sampling method is especially useful if we can very quickly compute the reconstruction from the incomplete sampled data at any given instant. In that case, one could inspect if the reconstruction was of good quality, and this would help in determining if more samples would be required. A compromise between fully progressive and non-progressive sampling methods would be a semi-progressive sampling algorithm. In this case, the sampling pattern locations could be determined for example after every K samples, e.g. for K, $2K$, $3K$, ... samples, with $K > 1$, and K sufficiently large so that the newest K samples would be of sufficient density over the acquisition domain so that receivers could e.g. be placed by sweeping through the image domain top-to-bottom and left-to-right, without requiring excessive travelling distance for the individual performing the placement of the receivers. This would be similar to what could be done for a non-progressive sampling method, except that in this semi-progressive case, each K samples is treated as a non-progressive sampling

subset. In general, unequal numbers of samples could be chosen for each subset so that instead of $K, 2K, 3K, ...$ the sampling number sequence could be $k_1, k_2, k_3, ...$ with $k_i \in \mathbb{N}$ and increasing.

An alternative solution would be to distribute a number of individuals to place these receivers, each one covering a subdomain of the entire domain being sampled. Each individual would only be responsible for placing receivers which would lie within their respective subdomain. This would require synchronization between the individuals placing the receivers since the order of placement is also important for a progressive sampling method. However the benefits in determining the reconstruction progressively could outweigh any potential drawbacks in terms of number of individuals involved and coordination required between them.

RESULTS AND DISCUSSION

In this section, we conduct experiments on the 2D synthetic reference model found in Figure 1. We first compare the reconstruction results for when we sample with a 1D discrete uniform random distribution along the receiver axis and taking full traces, vs. taking the same amount of data with a fully 2-dimensional discrete uniform random distribution over the model. We would expect that taking samples with a fully 2D uniform random distribution would lead to better results, and in Figure 10 we see that this is indeed the case. The reconstruction in Figure 10(b) is free of the many reconstruction artifacts that are clearly visible in Figure 10(a). We define the signal-to-noise ratio (SNR) as $SNR = 20 \log_{10} \frac{\|\mathbf{f}_0\|_2}{\|\mathbf{f}-\mathbf{f}_0\|_2}$, where $\mathbf{f}_0$ is the original data, and $\mathbf{f}$ is the interpolated data. The signal-to-noise ratio of the result in Figure 10(b) is also ~ 3 dB higher than that in Figure 10(a).

To compare the 2D hexagonal jittered undersampling with discrete uniform random undersampling, a synthetic time slice is chosen to do the experiments. As the sampling rate varying from 10% to 50%, we record SNR values, as shown in Fig. 11, for different rates of undersampling, it is found that 2D jittered sampling gives higher SNRs than discrete uniform random sampling for all sampling rates. For example, for a sampling rate of 50%, the uplift of jittered sampling over discrete uniform random sampling is approximatly 2 dB.

As expected, from Fig. 11, we see that as the sampling rate increases from 10% to 50%, the SNR of the reconstruction increases. What should be also noted is that because the CRSI reconstruction scheme is based on compressed sensing, the sampling rate necessary is not related to the maximum frequency of the underlying wavefield, as is the case with Nyquist sampling, but instead on the wavefield's complexity. This leads to a further reduction in the number of samples necessary to achieve a given SNR for less complicated wavefields. More information on this observation may be found in (Herrmann et al., 2009), and also later in this paper.

Figure 12 shows the spectra from the different sampling methods discussed in this paper. The original spectrum of the model is shown in Figure 1. In Figure 12(a), there are full copies of the original spectrum in Figure 1 of the same magnitude as the original spectrum. In Figure 12(b), where discrete random sampling is introduced along one axis, these aliases are less pronounced, and also it can be seen along the zero source frequency there is dimming in the spectrum, and this is due to the jittered sampling, which is taken in the crossline direction. The spectrum in Figure 12(c) is for a sampling similar to that in (b), except that instead of discrete uniform random sampling, we use jittered sampling in the crossline direction. This leads to a slight improvement in the sampling spectrum, with the incoherent noise less obvious. In Figure 12(d), each direction is sampled with a 1D jittered distribution,

each one possibly being a different instance of such a distribution. The overal 2-D sampling pattern is expressible as the Kronecker product of the two one-dimensional jittered sampling distributions. The aliases are much less visible than in Figures (b) and (c), due to the fact that regular samples are not taken along any direction. Also, there is dimming of the spectrum, however this time along both the zero source and receiver frequencies. The spectrum in Figure 12(e) corresponds to a jittered sampling pattern as in Figure 3(e), where now different jittered sampling positions are taken for each receiver position. This is again better than the previous spectra, with very little aliasing or noise present. Finally, in Figure 12(f), we see the spectrum for fully 2D (optimal) jittered sampling, and now there is no aliasing, with only some incoherent noise that can be easily filtered out from the spectrum.

In Figure 13, we show the reconstructions from the incomplete data determined by the various sampling methods discussed in this paper. As expected, the results progressively improve from Figure 13(a) to 13(f), as the sampling positions become less constrained and less regular. There is a notable (~ 1.64 dB) improvement even when samples are only taken irregularly in one dimension and regularly in the other, over completely 2D regular sampling.

Next, we employ 3D seismic lines for comparison of the sampling methods, including 2D Poisson disk sampling and Farthest Point sampling. Here, we only sample 25% of the total number of traces with different sampling methods, then do the recovery by CRSI with 3D curvelets. The interpolated results for the different schemes are shown in Figs. 14 and 15, and the residuals in Figs. 16 and 17. Again, jittered sampling is better than discrete uniform random sampling, while Poisson Disk and Farthest Point sampling are somewhat better than jittered sampling. Visually, it is a bit difficult to see the differences between the reconstructions from these sampling methods in Figures 14 and 15, and therefore we plot the residuals (the original model minus the reconstructions in Figures 16 and 17. From

this latter figure, we see that Poisson Disk and Farthest Point sampling have residuals that look quite similar to each other, and both have slightly less energy than the residual for jittered sampling. The difference between the Farthest Point/Poisson Disk vs. the jittered sampling residuals is evident at the apex of the residual in the time slice at roughly $t = 0.4$. Additionally, the jittered sampling residual can be seen visually to be of less energy than the residual from discrete uniform random sampling, with the difference between the two especially obvious at the elliptical structure in the top square of the data cube plot.

Real data

In Figure 18, the real SAGA dataset (see e.g. (Verschuur et al., 1992) for a detailed description of this dataset) is shown and the reconstruction from CRSI with hexagonal jittered sampling is shown in Figure 19. The reconstruction is faithful to the original with most important structures present. In the top square of the data cube, it can be seen that some of the fainter structures around the main diagonal where the source and receiver positions are close to each other are not very visible in the CRSI reconstruction. Otherwise, the reconstruction is very good, and some structures are smoother, which is important for visualization of the reconstructed volume.

EXTENSIONS

Non-uniform CRSI

The present work mainly used CRSI, where the curvelet transform is regular. Non-uniform CRSI (NCRSI) (Fénélon, 2008), based on the non-equispaced Fast Discrete Curvelet Transform (NFDCT), can deal with irregular sampling points or traces better without binning. The current version of NCRSI is for 1D sampling; we are currently developing NCRSI for higher dimensions, in combination with the non-equispaced curvelet transform, and results from this higher-dimensional NCRSI will be presented in a future work.

Adaptive sampling and reconstruction

If we know a priori some properties of the model for which we are acquiring data, then it is possible to sample more densely those areas of the image for which we know there is more information content. In general, this would correspond to higher frequency parts of the model since more samples are required to capture the wavefronts that are present in such regions, which are rapidly changing, and for which if a change is not captured by the samples, it will not show up in the reconstruction.

In addition to this, the reconstruction may also be made adaptive. The analytic expressions for the sampling spectra for hexagonal and Cartesian jittered sampling in the appendix, can be used to weight the unknown solution vector so that more curvelet-domain sparsity is enforced in the image at frequencies in the sampling spectrum where there is less energy. These two extensions are left to future work, along with symmetric sampling, which also allows the sampling density to be reduced.

Insights from compressive sampling theory

The profound implication of compressive sensing, of which this paper is a particular instance, is that the sampling rates are no longer determined by the Nyquist sampling criterion. Instead, the sampling rate can be chosen in accordance with the recovery quality one desires. This error decays for increasing sampling rates or for more compressive wavefields. In addition, (Herrmann et al., 2009; Lin and Herrmann, 2009) recently showed that simultaneous acquisition is another instance of compressive sensing, which leads to even better recovery. Combination of simultaneous acquisition with the approach presented in this paper will allow us to obtain an even greater reduction in the number of sources and receivers required for sampling a wavefield. Because both methods are linear, we envisage the development of a progressive sampling strategy, where the reconstruction from previously taken samples is updated as new samples are added until a reconstructed wavefield of desired quality is obtained.

CONCLUSIONS

In this paper, we explored sampling schemes with blue-noise patterns for curvelet-based interpolation. We extended jittered sampling to 2D and from our experiments, we have found that for higher dimensions, 2D jittered sampling is again significantly better than uniform random sampling in that it leads to reconstructions of higher SNR and of better visual quality than uniform random sampling when used to select the incomplete restricted data for CRSI. It was found that hexagonal jittered sampling is preferred over Cartesian jittered sampling due to its less stringent requirements on the sampling density for images of the same bandwidth. Additionally, we studied separable versus non-separable jittered

sampling, which indicates that completely non-separable jittered sampling is better. As well, we tested two other blue-noise pattern sampling schemes, i.e. Farthest Point and Poisson Disk sampling, which also gave good results under some conditions, despite not having been designed for curvelet-based reconstruction as jittered sampling is. We will explore these other methods more in a future work. We applied all these sampling techniques to 2D and 3D seismic data interpolation by CRSI, and obtained very good results, as compared to methods previously used in the literature. It would also be valuable to investigate the combination of the ideas of spatial subsampling design as we have done in this paper with the approach of using simultaneous independent sources as mentioned previously. This would lead to an even greater savings in the required sampling rate to achieve a desired SNR for the reconstructed wavefield.

ACKNOWLEDGMENTS

This work was in part financially supported by the NSERC Discovery Grant 22R81254 and CRD Grant DNOISE 334810-05 of F.J. Herrmann. The authors would also like to thank the China Scholarship Council, National Basic Research Program of China (973 program, 2007CB209505), and NSFC 40704019. The authors are grateful to SAGA Petroleum A.S. for providing the field data set.

REFERENCES

Bartlett, M. S., 1963, The spectral analysis of point processes: J. Roy. Statist. Soc. Ser. B, **25**, 264–296.

Berg, E. v., and M. P. Friedlander, 2007, SPGL1: A solver for large-scale sparse reconstruction. (http://www.cs.ubc.ca/labs/scl/spgl1).

———, 2008, Probing the pareto frontier for basis pursuit solutions: SIAM Journal on Scientific Computing, **31**, 890–912.

Candès, E. J., 2006, Compressive sampling: Presented at the International Congress of Mathematicians, Madrid, Spain.

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

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

Chauris, H., and T. Nguyen, 2008, Seismic demigration/migration in the curvelet domain: Geophysics, **73**, S35–S46.

Cook, R. L., 1986, Stochastic sampling in computer graphics: ACM Transactions on Graphics.

Dippé, M. A. Z., and E. H. Wold, 1985, Antialiasing through stochastic sampling: Proceedings of SIGGRAPH 85, 69–78.

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

Eldar, Y., M. Lindenbaum, M. Porat, and Y. Y. Zeevi, 1997, The farthest point strategy for progressive image sampling: IEEE Transactions on Image Processing.

Fénélon, L., 2008, Nonequispaced discrete curvelet transform for seismic data reconstruction: Master's thesis.

Grundland, M., J. Patera, M. Zuzana, and A. Neil, 2009, Image Sampling with Quasicrystals:

Arxiv preprint arXiv:0907.3604.

Hennenfent, G., and F. J. Herrmann, 2007, Irregular (sub-)sampling: from aliasing to noise: Presented at the 69^{th} EAGE Conference and Exhibition.

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

Herrmann, F. J., U. Boeniger, and D. J. Verschuur, 2007, Nonlinear primary-multiple separation with directional curvelet frames: Geophysical Journal International, **170**, 781–799.

Herrmann, F. J., Y. A. Erlangga, and T. Lin, 2009, Compressive simultaneous full-waveform simulation: Geophysics, **74**, A35.

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

Ignjatovic, Z., and M. F. Bocko, 2005, A method for efficient interpolation of discrete-time signals by using a blue-noise mapping method: Proceedings of the IEEE ICASSP 2005, 213– 216.

László, S.-K., 1995, Stochastic sampling of two-dimensional images: COMPUGRAPHICS 95, 292–299.

Lin, T., and F. J. Herrmann, 2009, Unified compressive sensing framework for simultaneous acquisition with primary estimation: Presented at the .

Long, A. S., 2004, The revolution in seismic resolution: High density 3d spatial sampling developments and results: Presented at the ASEG Extended Abstracts.

Lustig, M., M. Alley, S. Vasanawala, D. Donoho, and J. Pauly, 2009, Autocalibrating parallel imaging compressed sensing using l1 spir-it with poisson-disc sampling and joint sparsity constraints: Presented at the ISMRM Workshop on Data Sampling and Image

Reconstruction, Sedona '09.

Ma, J., and G. Plonka, 2009, A review of curvelets and recent applications: IEEE Signal Processing Magazine, to appear.

Mitchell, D., 1987, Generating antialiased images at low sampling densities: Proc. of the 14th annual conf. on computer graphics and interactive techniques, 65–72.

P. Brémaud, L. M., and A. Ridolfi, 2002, Power spectra of random spike fields and related processes: Technical report, École Polytechnique Fédérale de Lausanne, School of Computer and Information Sciences. (Technical Report 200280).

Tang, G., R. Shahidi, F. J. Herrmann, and J. Ma, 2009, Higher dimensional blue-noise sampling schemes for curvelet-based seismic data recovery: Presented at the 79^{th} SEG Annual Meeting, Houston. Expanded Abstracts.

Vermeer, G., 1990, Seismic wavefield sampling: Soc. Expl. Geophys.

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

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

APPENDIX - ANALYSIS OF 2D JITTERED SAMPLING

From the technical report of Brémaud et al. (P. Brémaud and Ridolfi, 2002), if there is a wide sense stationary point process S jittered by a point process Z, so that the jitter point process J is given by $J = S + Z$, then the *Bartlett spectrum* of J, $\mu_J(d\mathbf{f})$ is given by

$$\mu_J(d\mathbf{f}) = |\psi_Z(\mathbf{f})|^2 \mu_S(d\mathbf{f}) + \lambda(1 - |\psi_Z(\mathbf{f})|^2)d\mathbf{f}, \tag{7}$$

where $\mu_S(d\mathbf{f})$ is the Bartlett spectrum of the process S, $\psi_Z(\mathbf{f})$ is the characteristic function of Z, and λ is the *intensity* of the process S. The arguments of many of the functions in Equation 7 are $d\mathbf{f}$ because the power spectral density measures the expected power in small intervals around each point in the frequency domain. The Bartlett power spectrum is a generalization of the usual power spectral density for point processes which are not wide-sense stationary (Bartlett, 1963). Since S is indeed a deterministic process, the Bartlett spectrum of S will just be the usual power spectrum. Additionally, the intensity of S will simply be the sampling density. Then Equation 7 becomes:

$$S_J(\mathbf{f}) = |\psi_Z(\mathbf{f})|^2 S_S(\mathbf{f}) + \lambda(1 - |\psi_Z(\mathbf{f})|^2) = |E\{e^{-2\pi i(\mathbf{f}\cdot\bar{\epsilon})}\}|^2 |\hat{S}(\mathbf{f})|^2 + \lambda(1 - |E\{e^{-2\pi i(\mathbf{f}\cdot\bar{\epsilon})}\}|^2) \tag{8}$$

The above can be shown to be equal to

$$S_J(\mathbf{f}) = (|E\{\hat{J}(\mathbf{f})\}|)^2 + \lambda(1 - |E\{e^{-2\pi i(\mathbf{f}\cdot\bar{\epsilon})}\}|^2), \tag{9}$$

since the jitter point process Z is assumed to be independent of the regular sampling process S.

Below, we derive the expected value and standard deviation of the Fourier spectrum for regular hexagonal jittered sampling. This can then be shown to be equivalent to finding the expectation and standard deviation of the entries of the Gram matrix $\mathbf{A}^H\mathbf{A}$ for the sampling, which can then be used to analyze the amount of aliasing present for this type of

jittered sampling. We could also derive similar expressions for regular rectangular jittered sampling, however elect not to, since using the same techniques as for hexagonal jittered sampling, these can be easily derived, but not easily vice-versa.

Analysis of aliasing for 2D regular hexagonal jittered sampling

As in (Hennenfent and Herrmann, 2008), we derive expressions to analyze the aliasing for 2D regular hexagonal jittered sampling. In order to do this, we need to find the spectrum of this sampling process. Since this process is random, the spectrum will also be a random process. Hence we calculate the expectation and variance of this random spectrum. Let $\frac{\gamma}{2}$ be the radius of any regular hexagon in the hexagonal sampling grid, so that $\gamma = 2W$ as in Figure 4. In this case, the jittered sampling locations $\mathbf{r}_n$ are determined by the formula:

$$\mathbf{r}(k_1, k_2) = \mathbf{r}_1(k_1, k_2) \bigcup \mathbf{r}_2(k_1, k_2), \tag{10}$$

where

$$\mathbf{r}_1(k_1, k_2) = \left(\frac{3}{2}\gamma k_1, \frac{\sqrt{3}}{2}\gamma k_2 \right) + \epsilon_{\mathbf{1}k_1,k_2} \tag{11}$$

and

$$\mathbf{r}_2(k_1, k_2) = \left(\frac{3}{2}\gamma k_1 + \frac{3}{4}\gamma, \frac{\sqrt{3}}{2}\gamma k_2 + \frac{\sqrt{3}}{4}\gamma \right) + \epsilon_{\mathbf{2}k_1,k_2}, \tag{12}$$

and the $\epsilon_{\mathbf{i}k_1,k_2}$ are a set of identically uniformly and independently distributed random variables on the hexagon with circumscribing circle of radius $\frac{\zeta}{2}$, centered at the origin, for all indices k_1 and k_2 and $i = 1$ or 2. Call this distribution of the $\epsilon_{\mathbf{i}k_1,k_2}$, $p_{\frac{\zeta}{2}}(\epsilon)$.

Then similar to (Hennenfent and Herrmann, 2008), the sampling operator s is given by the expression:

$$s(\mathbf{r}) = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \delta(\mathbf{r} - \mathbf{r}_1(k_1, k_2)) + \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \delta(\mathbf{r} - \mathbf{r}_2(k_1, k_2)). \tag{13}$$

The Fourier transform of the sampling operator s is

$$\hat{S}(\mathbf{f}) = \frac{4}{3\sqrt{3}\gamma^2} \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \delta\left(f_1 - k_1 \frac{2}{3\gamma}\right) \delta\left(f_2 - k_2 \frac{2}{\sqrt{3}\gamma}\right) \left(1 + (-1)^{k_1+k_2}\right) e^{-2\pi i \mathbf{f} \cdot \epsilon_{\mathbf{1}k_1,k_2}}, \tag{14}$$

where $\mathbf{f} = (f_1, f_2)$ is the frequency variable vector. Therefore, the expected value of the Fourier transform of the sampling operator s is:

$$E\{\hat{S}(\mathbf{f})\} =$$
$$E\{e^{-2\pi i(f_1\bar{\epsilon}_{\mathbf{x}} + f_2\bar{\epsilon}_{\mathbf{y}})}\} \cdot \frac{4}{3\sqrt{3}\gamma^2} \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \delta\left(f_1 - k_1 \frac{2}{3\gamma}\right) \delta\left(f_2 - k_2 \frac{2}{\sqrt{3}\gamma}\right) \left(1 + (-1)^{k_1+k_2}\right). \tag{15}$$

Here, because all the $\epsilon_{\mathbf{1}k_1,k_2}$'s are identically distributed, they can be replaced by a generic $\epsilon = (\bar{\epsilon}_{\mathbf{x}}, \bar{\epsilon}_{\mathbf{y}})$ in the expected value of the sampling operator's Fourier transform. So we are left to compute $E\{e^{-2\pi i(f_1\bar{\epsilon}_{\mathbf{x}} + f_2\bar{\epsilon}_{\mathbf{y}})}\}$, where $(\bar{\epsilon}_{\mathbf{x}}, \bar{\epsilon}_{\mathbf{y}})$ is distributed uniformly on the hexagon $\mathbf{H}_\zeta$ with circumscribing circle of radius $\frac{\zeta}{2}$ centered at the origin. This is in fact the *characteristic* function $\psi_Z(\mathbf{f})$ as in Equation 7, of the distribution $p_{\frac{\zeta}{2}}(\epsilon)$. After some calculations, the following expression is obtained for this characteristic function:

$$\psi_Z(\mathbf{f}) = \frac{4\sqrt{3}}{9\pi^2 f_1 f_2 (f_1^2 - 3f_2^2)\zeta^2} \left(2\sin\left(\frac{\pi\zeta}{2} f_1\right) \sin\left(\frac{\pi\zeta}{2}\sqrt{3} f_2\right) (f_1^2 - 3f_2^2) - 2f_1 f_2 \sqrt{3} \cos(\pi\zeta f_1) + \right. \tag{16}$$
$$\left. \cos\left(\frac{\pi\zeta}{2}(f_1 + \sqrt{3}f_2)\right) f_1(f_1 + \sqrt{3}f_2) - \cos\left(\frac{\pi\zeta}{2}(f_1 - \sqrt{3}f_2)\right) f_1(f_1 - \sqrt{3}f_2)\right)$$

By the sifting property of the δ function, the expected value of $\hat{S}(\mathbf{f})$ may be simplified

to:

$$E\{\hat{S}(\mathbf{f})\} = \sum_{\substack{k_1,k_2=-\infty \\ k_1+k_2\ even}}^{\infty} \frac{8\sqrt{3}}{9\pi^2 k_1 k_2 (k_1^2 - 9k_2^2)\zeta^2} \left(\cos\left(\frac{\pi\zeta}{3\gamma}(k_1+3k_2)\right) k_1(k_1+3k_2) - \right.$$
$$\cos\left(\frac{\pi\zeta}{3\gamma}(k_1-3k_2)\right) k_1(k_1-3k_2) + 2\sin\left(\frac{\pi\zeta}{3\gamma}k_1\right)\sin\left(\frac{\pi\zeta}{\gamma}k_1\right)(k_1^2-9k_2^2) \qquad (17)$$
$$\left. -6k_1k_2\cos\left(\frac{2\pi\zeta}{3\gamma}k_1\right)\right) \delta\left(f_1 - \frac{k_1}{\frac{3}{2}\gamma}\right)\delta\left(f_2 - \frac{k_2}{\frac{\sqrt{3}}{2}\gamma}\right).$$

The intensity of the regular hexagonal sampling process, λ is computed as the sampling density, or the expected number of samples per unit area, which in this case is simply $\lambda = \frac{8}{3\sqrt{3}\gamma^2}$.

Let the ratio of the perturbation range to the edge length be $\beta = \frac{\zeta}{\gamma}$. Then the above expression becomes:

$$E\{\hat{S}(\mathbf{f})\} =$$
$$\sum_{\substack{k_1,k_2=-\infty \\ k_1+k_2\ even}}^{\infty} \frac{8\sqrt{3}}{9\pi^2 k_1 k_2 (k_1^2 - 9k_2^2)\zeta^2} \left(\cos\left(\frac{\pi\beta}{3}(k_1+3k_2)\right) k_1(k_1+3k_2) - \cos\left(\frac{\pi\beta}{3}(k_1-3k_2)\right) k_1(k_1-3k_2) \right.$$
$$\left. +2\sin\left(\frac{\pi\beta}{3}k_1\right)\sin(\pi\beta k_1)(k_1^2-9k_2^2) - 6k_1k_2\cos\left(\frac{2\pi\beta}{3}k_1\right)\right) \delta\left(f_1 - \frac{k_1}{\frac{3}{2}\gamma}\right)\delta\left(f_2 - \frac{k_2}{\frac{\sqrt{3}}{2}\gamma}\right). \qquad (18)$$

Observe that there is still a dependence on the maximum jitter magnitude ζ, but that this is expected, since for a given perturbation ratio β, as γ, and thus ζ increases, the hexagonal jittered sampling grid spreads out and becomes bigger, which implies that the sampling mask becomes lower in frequency. As $(k_1, k_2) \to (0, 0)$, the limit of the above is 1, as expected, since this corresponds to the base replica of the sampling spectrum, and should have magnitude of unity. Other values of the spectrum can be calculated by substituting various values of k_1 and k_2, and these values can only be non-zero when the condition that $k_1 + k_2$ is even holds.

The expression in Equation 18 is the expected value of the regular hexagonal jittered sampling spectrum. We now wish to find the variance of this random spectrum. This can be done by computing the power spectral density for the regular hexagonal jittered sampling process, and then using the fact that

$$\sigma_J^2 = S_J(\mathbf{f}) - [E\{\hat{S}(\mathbf{f})\}]^2 = \lambda(1 - |\psi_Z(\mathbf{f})|^2) \tag{19}$$

Therefore, the variance of the hexagonal jittered sampling process is given by

$$\sigma_J^2 = \frac{8}{3\sqrt{3}\gamma^2}\left(1 - \left(\frac{4\sqrt{3}}{9\pi^2 f_1 f_2 (f_1^2 - 3f_2^2)\zeta^2}\left(2\sin\left(\frac{\pi\zeta}{2}f_1\right)\sin\left(\frac{\pi\zeta}{2}\sqrt{3}f_2\right)(f_1^2 - 3f_2^2) - \right.\right.\right.$$
$$\left.\left.\left. 2f_1 f_2\sqrt{3}\cos(\pi\zeta f_1) + \cos\left(\frac{\pi\zeta}{2}(f_1 + \sqrt{3}f_2)\right) f_1(f_1 + \sqrt{3}f_2) - \cos\left(\frac{\pi\zeta}{2}(f_1 - \sqrt{3}f_2)\right) f_1(f_1 - \sqrt{3}f_2)\right)\right)^2\right) \tag{20}$$

Ignoring discretization effects, the elements of the Gram matrix $\mathbf{L} = \mathbf{A^H A}$ are each random with mean given by Equation 18 and variance given by Equation 20, but where only the elements of this matrix corresponding to grid points closest to the impulse locations in the frequency response for the regular hexagonal jittered samples are equal to the corresponding magnitudes of the impulse functions and can be non-zero.

Analysis of aliasing for 2D regular Cartesian jittered sampling

We now briefly find the expectation and standard deviation of the sampling spectrum for Cartesian jittered sampling. Since we have already gone through the derivation of these functions for the more complicated case of hexagonal jittered sampling, we skip through many of the steps in the present derivation.

For regular Cartesian jittered sampling, we find that the characteristic function is:

$$\psi_Z(\mathbf{f}) = \frac{\sin(\pi f_1\zeta)\sin(\pi f_2\zeta)}{\pi^2\zeta^2 f_1 f_2} \tag{21}$$

Hence the expected value of the spectrum for this type of sampling can be found to be:

$$E\{\hat{S}(\mathbf{f})\} = \frac{\sin(\pi k_1 \frac{\zeta}{\gamma}) \sin(\pi k_2 \frac{\zeta}{\gamma})}{\pi^2 \zeta^2 k_1 k_2} \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \delta\left(f_1 - \frac{k_1}{\gamma}\right) \delta\left(f_2 - \frac{k_2}{\gamma}\right) \tag{22}$$

The variance, once again wideband noise, is computed to be:

$$\sigma_J^2 = \lambda(1 - |\psi_Z(\mathbf{f})|^2) = \frac{1}{\gamma^2}\left(1 - \frac{\sin^2(\pi f_1 \zeta) \sin^2(\pi f_2 \zeta)}{\pi^4 \zeta^4 f_1^2 f_2^2}\right) \tag{23}$$

LIST OF FIGURES

same jittered positions for each receiver (separable), (e) 2D jittered sampling, jittered along receiver and source axes, different jittered positions for each receiver (non-separable), (f) fully 2D jittered sampling (non-separable)

13 Reconstructions from 75% missing traces: (a) 2D regular sampling, SNR=3.91 dB, (b) regular along source axis, discrete uniform random along receiver axis, SNR=7.30 dB, (c) regular along source axis, jittered along receiver axis, SNR=8.94 dB, (d) 2D jittered sampling, jittered sampling along receiver and source axes, same source pos. for all receivers, SNR=9.65 dB, (e) 2D jittered sampling along receiver and source axes, different source pos. for each receiver, SNR=10.03 dB, (f) Fully 2D jittered sampling, SNR=10.86 dB

14 CRSI reconstruction from 75% missing traces sampled by (a) discrete uniform random sampling, SNR=8.13 dB and (b) jittered sampling, SNR=8.43 dB

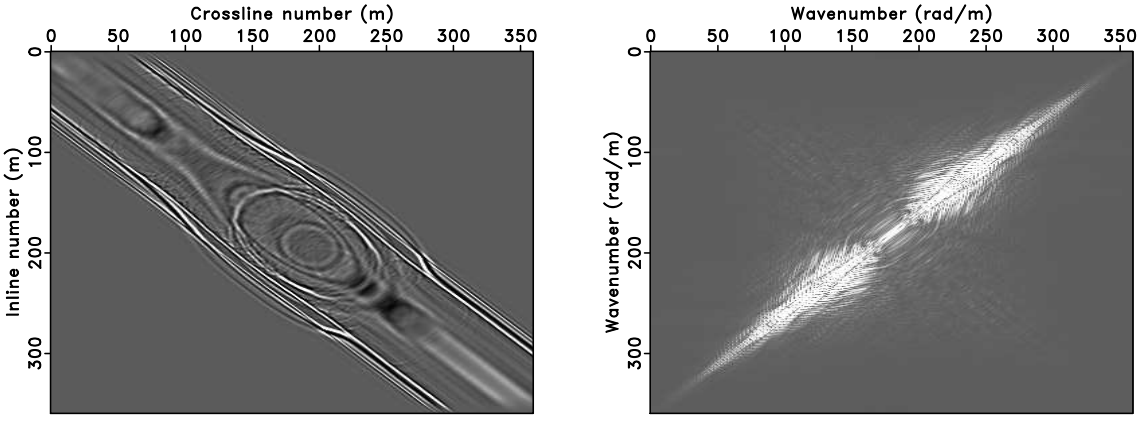
15 CRSI reconstruction from 75% missing traces sampled by (a) Farthest Point sampling, SNR=8.50 dB and (b) Poisson Disk sampling, SNR=8.48 dB

16 CRSI reconstruction residuals from (a) random sampling and (b) jittered sampling

17 CRSI reconstruction residuals from (a) Farthest Point sampling and (b) Poisson Disk sampling

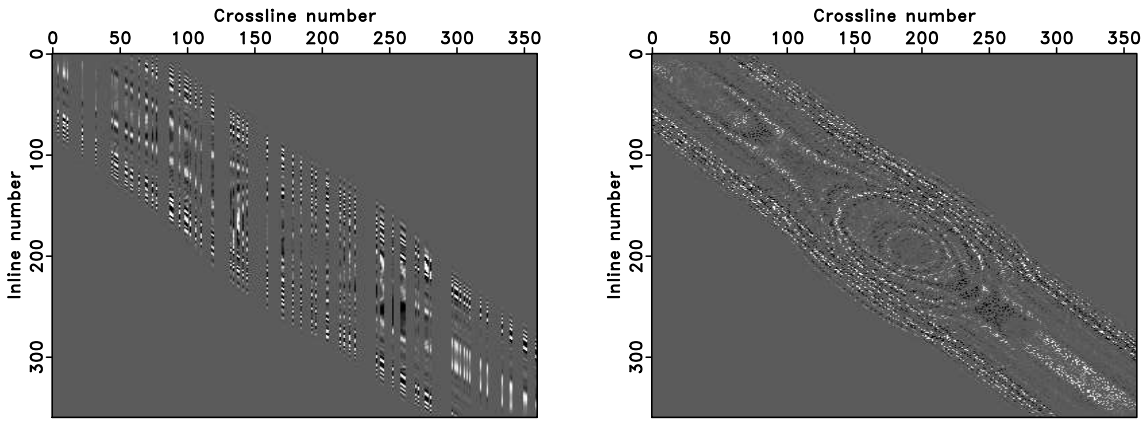
18 Saga example: model

19 Saga example: CRSI reconstruction from 75% missing traces sampled by jittered hexagonal sampling



(a) (b)

Figure 1: (a) Reference synthetic seismic data and (b) Fourier spectrum of (a) –



(a) (b)

Figure 2: 25% random samples on time slice via: (a) 1D and (b) 2D sampling –

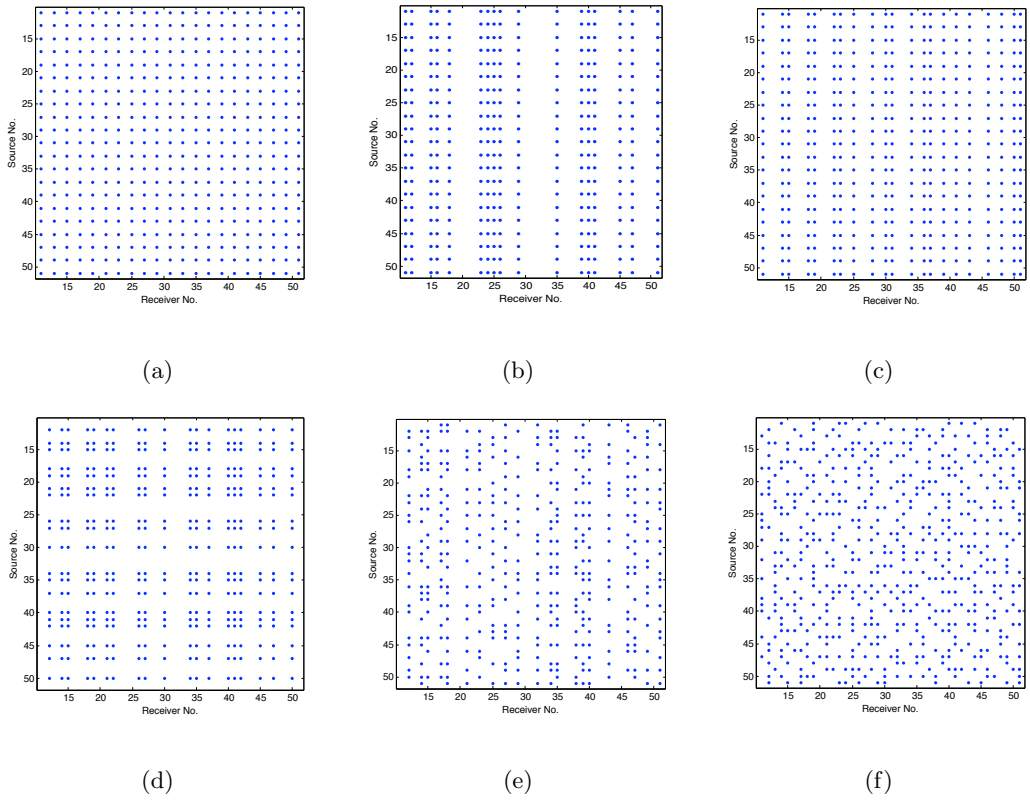


Figure 3: A zoom of different types of sampling: (a) regular ($\mathbf{R} = \mathbf{R}_r \otimes \mathbf{R}_r$) (b) uniform random inline and regularly sampled crossline($\mathbf{R} = \mathbf{R}_u \otimes \mathbf{R}_r$), (c) jittered sampling crossline and regularly sampled inline ($\mathbf{R} = \mathbf{R}_{j1} \otimes \mathbf{R}_r$), (d) 2D jittered sampling, with jittered sample positions crossline, and shots fired at positions determined by another jittered sampling pattern, (e) 2D jittered sampling, jittered positions crossline, and different jittered positions along each cable inline (non-separable), (f) Fully 2D jittered sampling ($\mathbf{R} = \mathbf{R}_{j2}$).

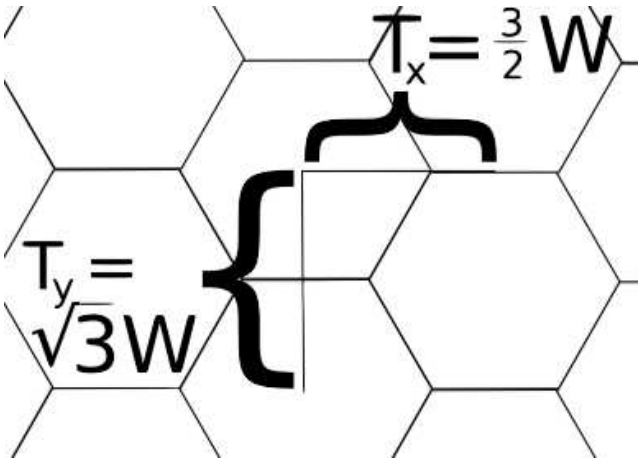
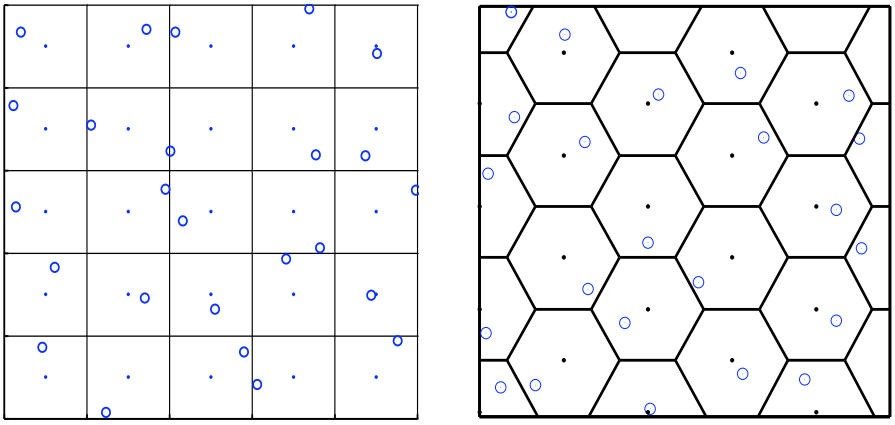


Figure 4: Regular hexagonal sampling grid –



(a) (b)

Figure 5: (a) Cartesian jittered sampling scheme, (b) hexagonal jittered sampling scheme

–

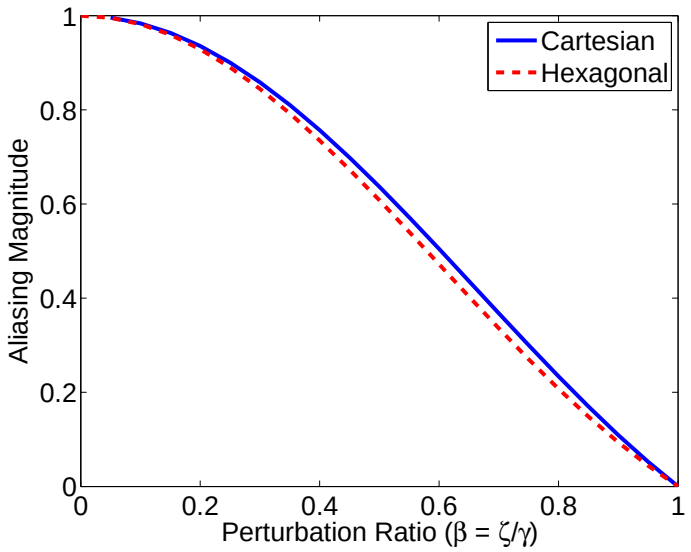


Figure 6: Comparison of the theoretical aliasing for hexagonal versus Cartesian jittered sampling.

–

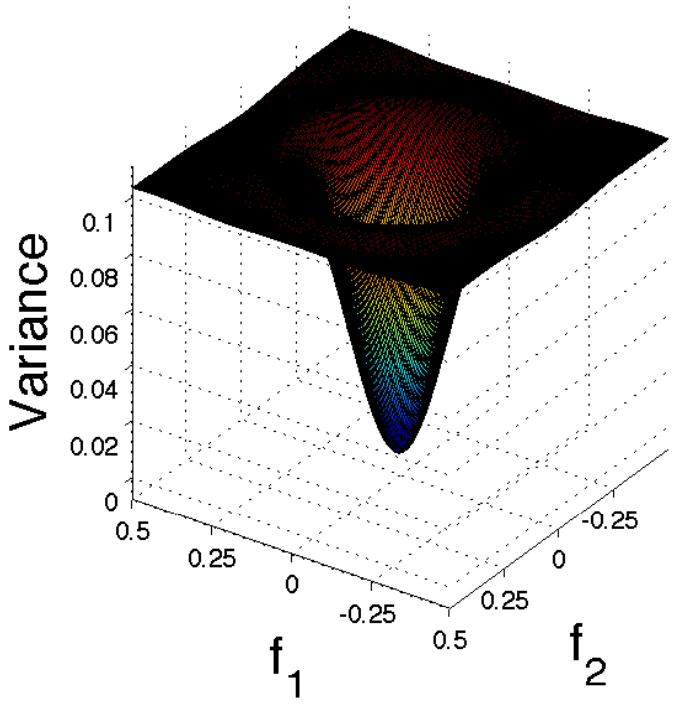
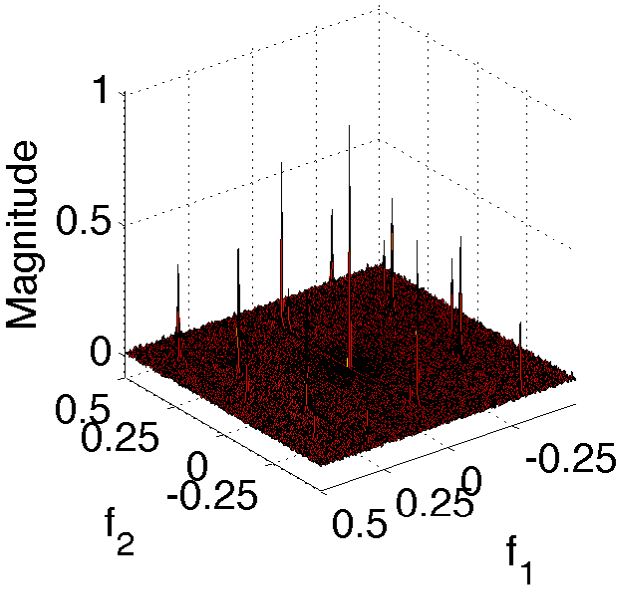
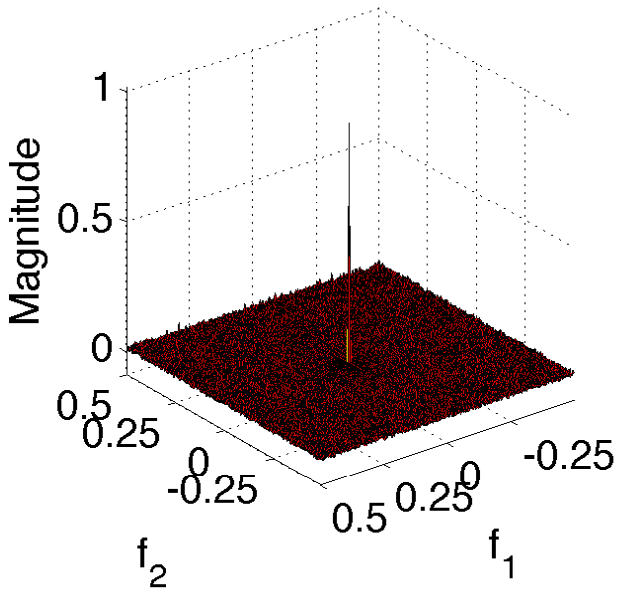


Figure 7: Theoretical sampling variance for optimally-jittered hexagonal undersampling.

–

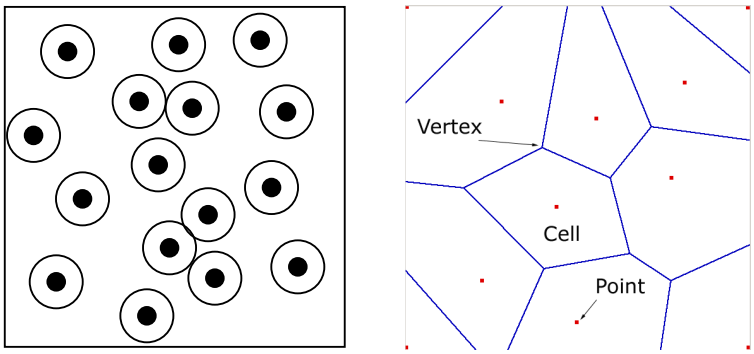


(a)

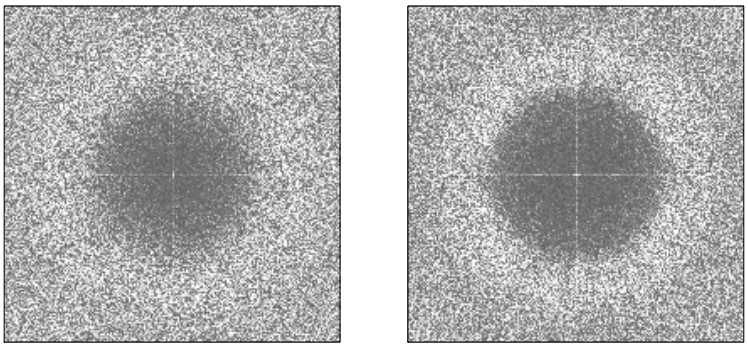


(b)

Figure 8: Example spectrum for hexagonal jittered sampling with: (a) $\zeta < \gamma$ and (b) $\zeta = \gamma$



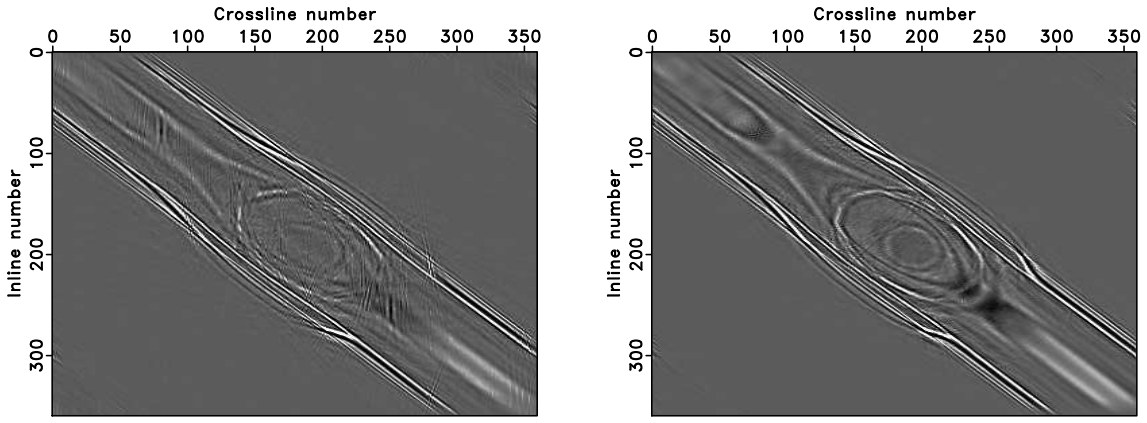
(a) (b)



(c) (d)

Figure 9: Sampling schemes of (a) Poisson Disk (b) Farthest Point, (c) is the sampling spectrum of (a), and (d) is that of (b)

–



(a) (b)

Figure 10: (a) Reconstruction from 1D samples in Figure 2(a) (SNR=6.77 dB) and (b) Reconstruction from 2D samples in Figure 2(b) (SNR=9.75 dB)

–

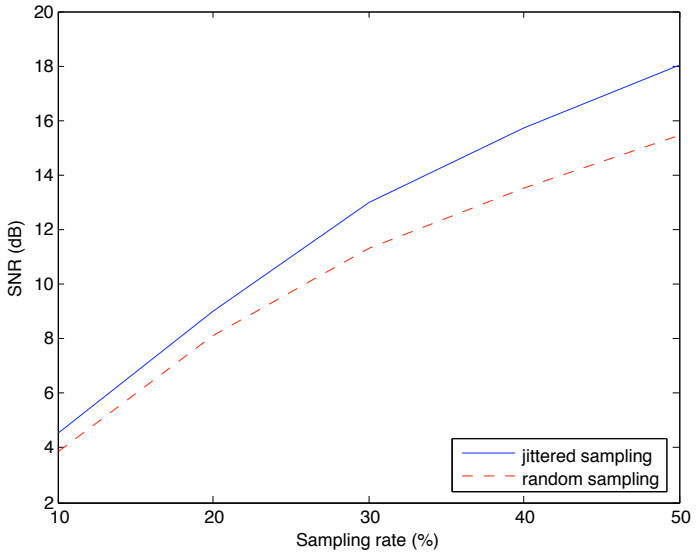


Figure 11: SNR of reconstruction from 2D jittered undersampling and discrete uniform random undersampling, for different sampling rates

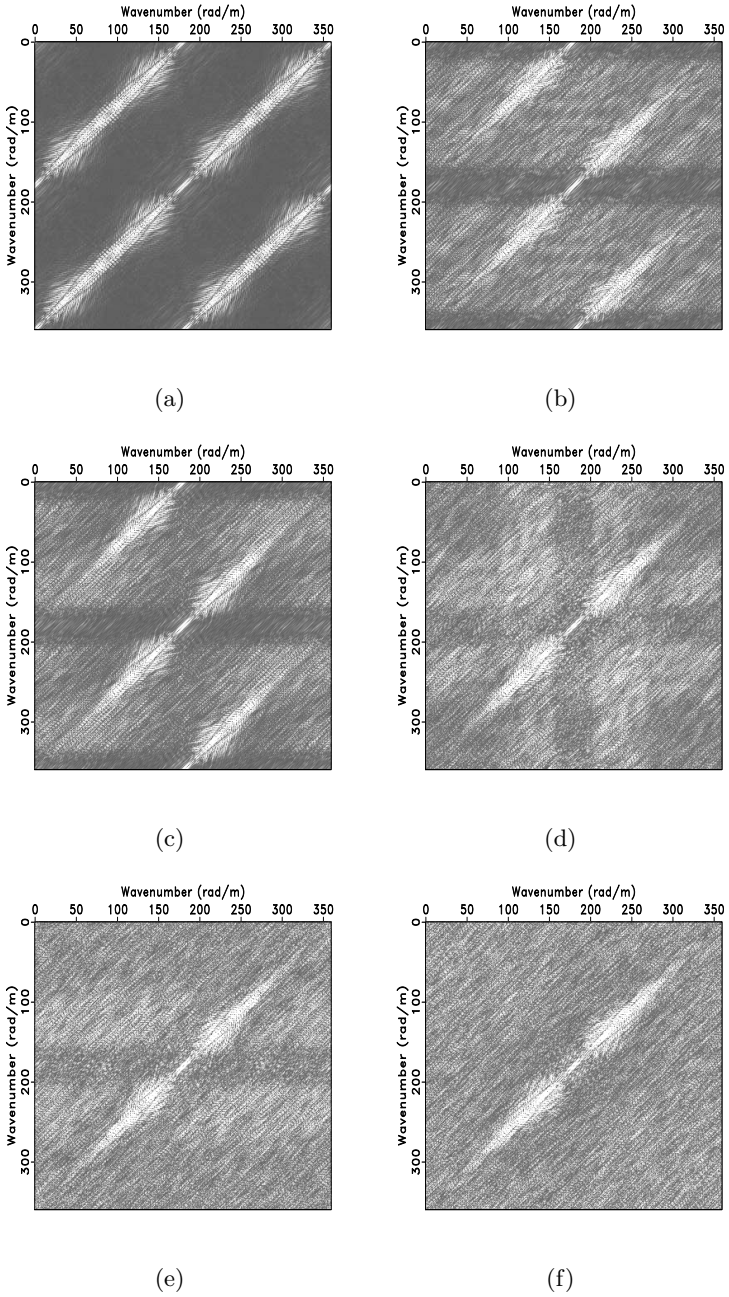


Figure 12: Spectra of different 2D sampling methods: (a) regular sampling, (b) regular along source axis, discrete uniform random along receiver axis, (c) regular along source axis, jittered along receiver axis, (d) 2D jittered sampling, jittered along receiver and source axes, same jittered positions for each receiver (separable), (e) 2D jittered sampling, jittered along receiver and source axes, different jittered positions for each receiver (non-separable), (f) fully 2D jittered sampling (non-separable)

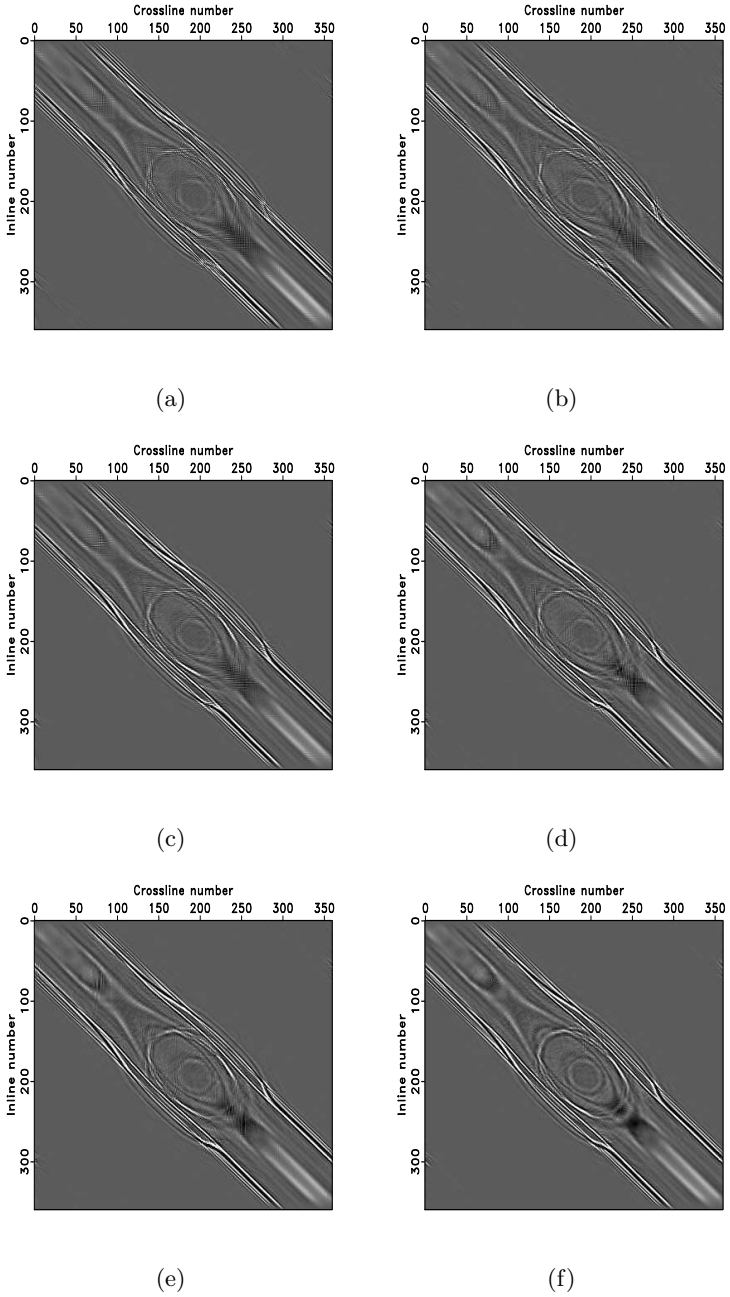
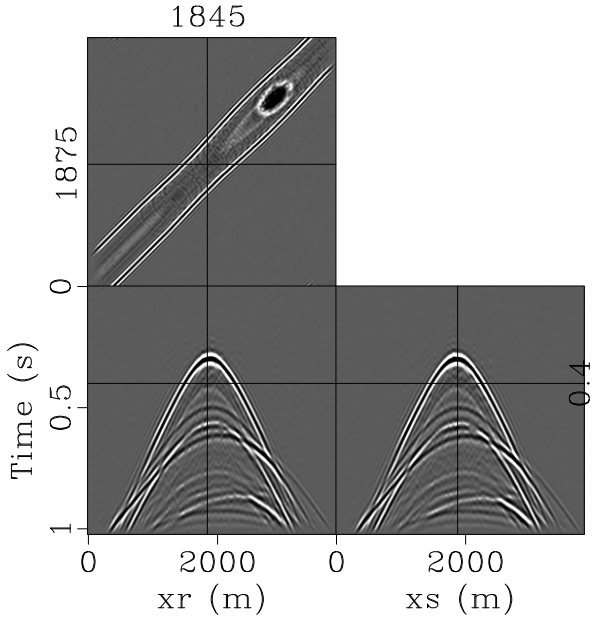
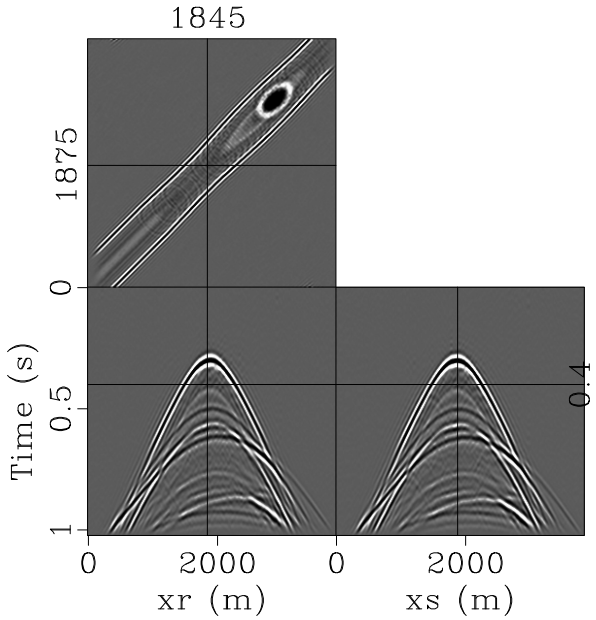


Figure 13: Reconstructions from 75% missing traces: (a) 2D regular sampling, SNR=3.91 dB, (b) regular along source axis, discrete uniform random along receiver axis, SNR=7.30 dB, (c) regular along source axis, jittered along receiver axis, SNR=8.94 dB, (d) 2D jittered sampling, jittered sampling along receiver and source axes, same source pos. for all receivers, SNR=9.65 dB, (e) 2D jittered sampling along receiver and source axes, different source pos. for each receiver, SNR=10.03 dB, (f) Fully 2D jittered sampling, SNR=10.86 dB

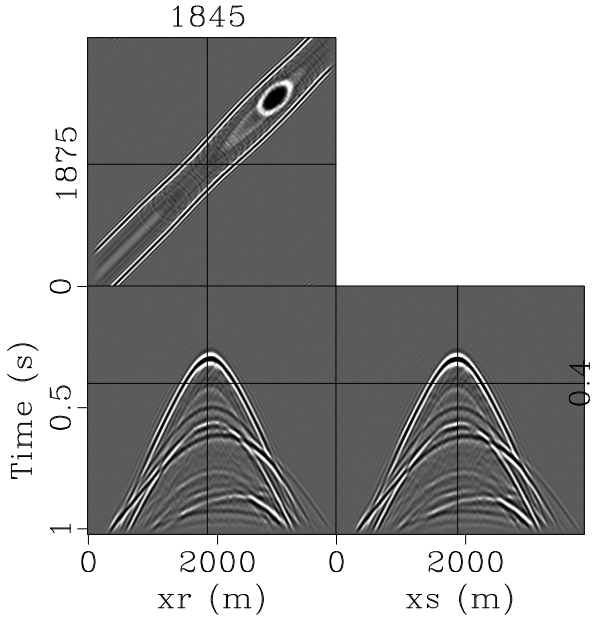


(a)

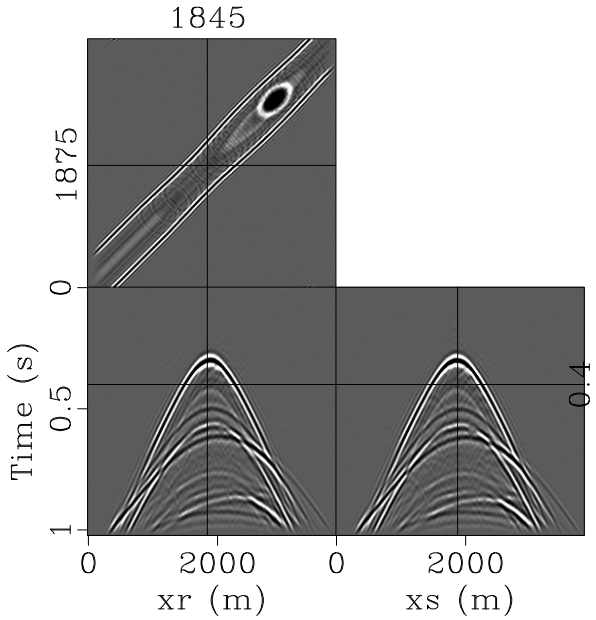


(b)

Figure 14: CRSI reconstruction from 75% missing traces sampled by (a) discrete uniform random sampling, SNR=8.13 dB and (b) jittered sampling, SNR=8.43 dB

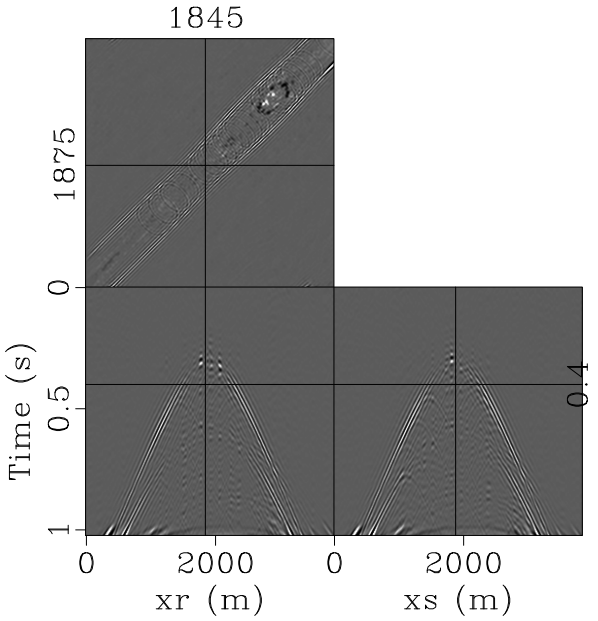


(a)

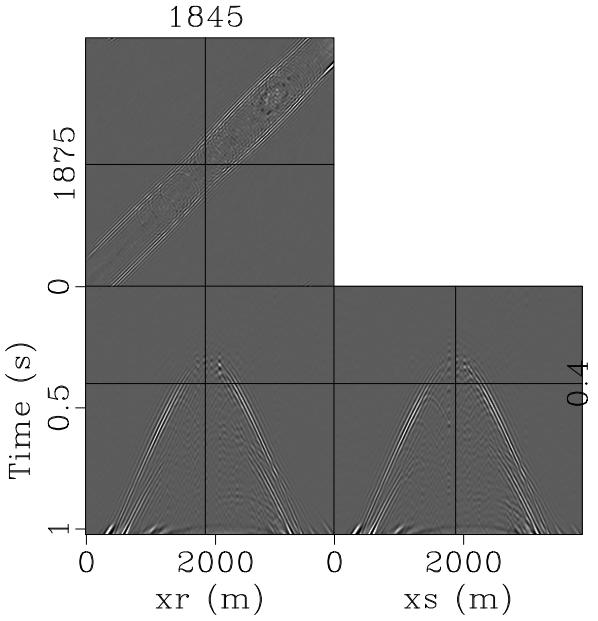


(b)

Figure 15: CRSI reconstruction from 75% missing traces sampled by (a) Farthest Point sampling, SNR=8.50 dB and (b) Poisson Disk sampling, SNR=8.48 dB



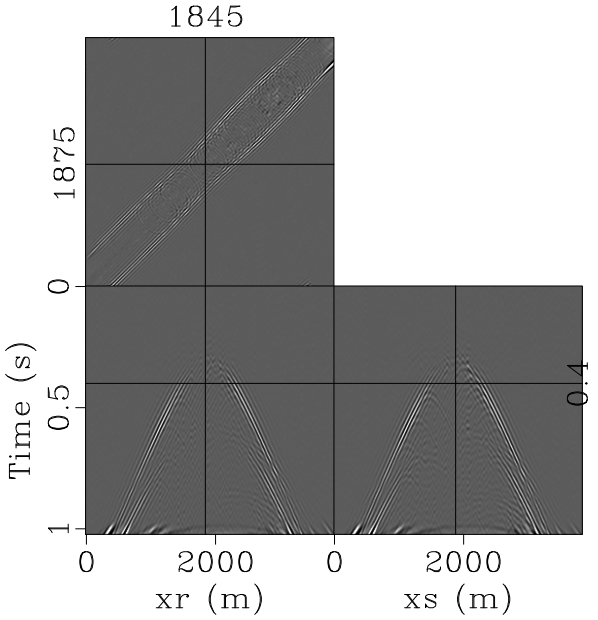
(a)



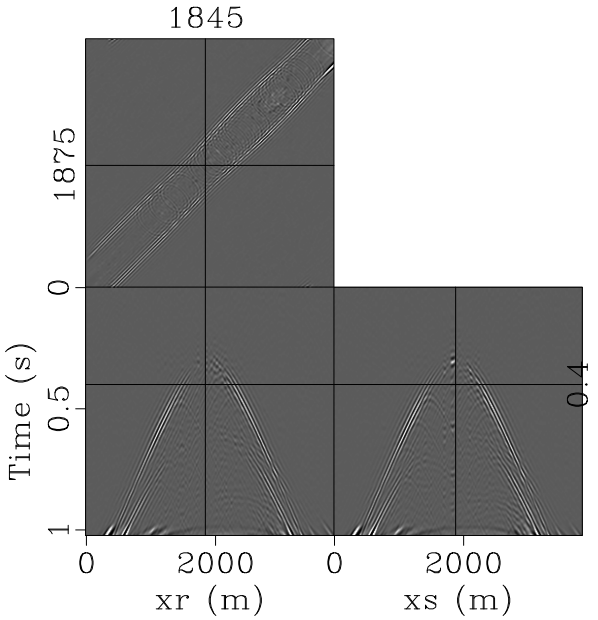
(b)

Figure 16: CRSI reconstruction residuals from (a) random sampling and (b) jittered sampling

–



(a)



(b)

Figure 17: CRSI reconstruction residuals from (a) Farthest Point sampling and (b) Poisson Disk sampling

–

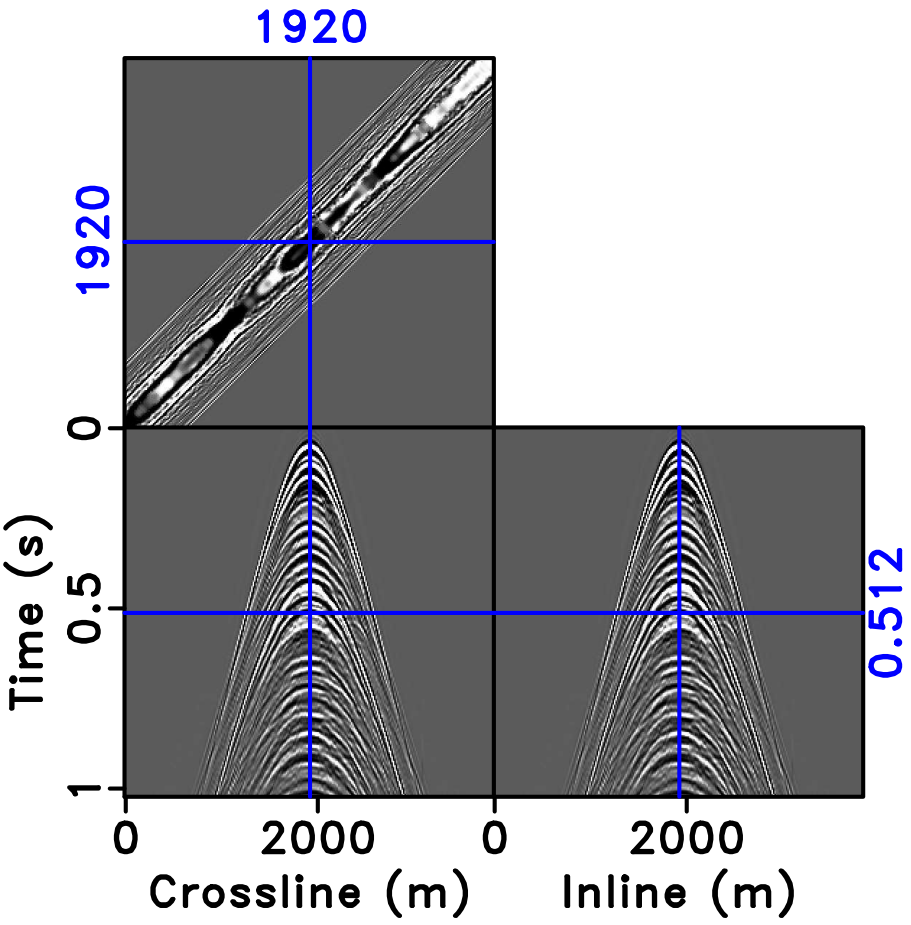


Figure 18: Saga example: model –

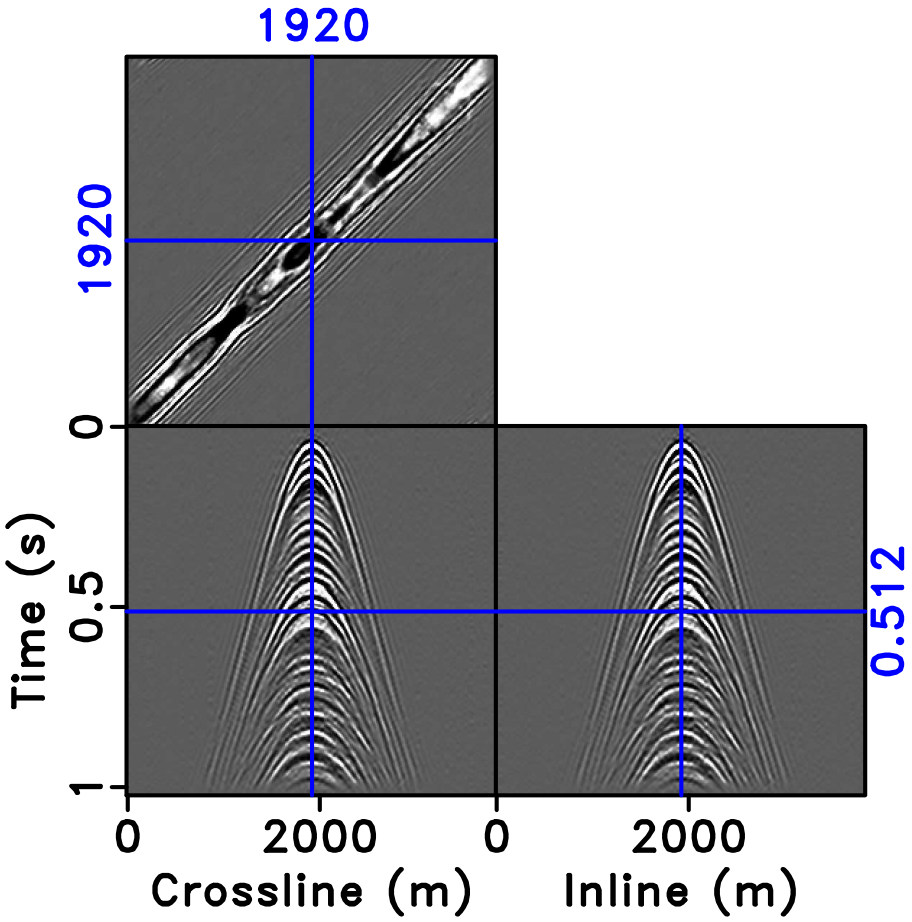


Figure 19: Saga example: CRSI reconstruction from 75% missing traces sampled by jittered hexagonal sampling