Random sampling: new insights into the reconstruction of coarsely-sampled wavefields

Gilles Hennenfent, EOS-UBC and Felix J. Herrmann, EOS-UBC

ABSTRACT

In this paper, we turn the interpolation problem of coarsely-sampled data into a denoising problem. From this point of view, we illustrate the benefit of random sampling at sub-Nyquist rate over regular sampling at the same rate. We show that, using nonlinear sparsity-promoting optimization, coarse random sampling may actually lead to significantly better wavefield reconstruction than equivalent regularly sampled data.

INTRODUCTION

Dense sampling of seismic data is traditionally understood as evenly-distributed time and space measurements of the reflected wavefield. Moreover, the sampling rate along each axis must be equal to or above twice the highest frequency/wavenumber of the continuous signal being discretized (Shannon/Nyquist sampling theorem). In practice, however, seismic data is often randomly and/or sparsely sampled along spatial coordinates, which is generally considered as a challenge since it breaks one or both previously-stated conditions of dense sampling. It turns out that these acquisition geometries are not necessarily a source of adversity to accurately reconstruct densely-sampled data when using nonlinear optimization promoting sparsity. This new insight, developed in the information theory community, is referred to in the literature by the terms "compressed sensing" or "compressive sampling" (see e.g. Donoho, 2006; Candes et al., 2005, and references therein).

THEORY

Sampling below Nyquist rate

Discretizing a continuous function bandlimited to the frequency interval $[-B, B]$ corresponds to a multiplication with a Dirac comb in the time domain and thus to a convolution with another Dirac comb in the frequency domain. As a consequence, the original spectrum becomes periodic after discretization. Problems occur when regularly sampling below Nyquist rate, i.e. $f_s < 2B$. Replicas of the original spectrum overlap, which creates an indetermination in the reconstruction process. This is the well-known phenomenon of aliasing. In contrast, random sampling at the same sub-Nyquist rate is less likely to create strong aliases but rather weak broadband noise. Consider for example a random sampling operator $\mathbf{s}$ over $[0, N]$ where N is the size of the sampling region. Suppose that $\mathbf{s}$ samples $n < N$ points uniformly distributed in the interval. Then, the expectation of the power spectrum of $\mathbf{s}$ over $[0, N]$ is given by (Leneman, 1966; Dippe and Wold, 1985)

$$\mathrm{E}\left[|\hat{\mathbf{s}}_N(u)|^2\right] = n^2\delta_{u0} + (1 - \delta_{u0})\, n, \tag{1}$$

where the symbol ˆ denotes Fourier coefficients, $\mathrm{E}[\cdot]$ the mathematical expectation, and u is the frequency variable. The first term in this expression is only nonzero at the origin and gives through convolution a scaled version of the original spectrum. The second term is only zero at the origin and can be assimilated to broadband noise.
Fig. 1 summarizes these observations. Fig. 1(a) shows the amplitude spectrum of a densely-sampled signal consisting of the superposition of three cosine functions. Figs. 1(b) and 1(c) show the spectra of the same signal being regularly- and randomly-sampled below Nyquist rate, respectively.

Reconstruction by denoising

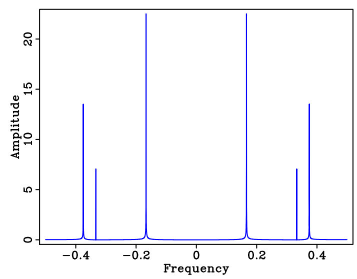
Consider the following linear forward model for the interpolation problem

$$\mathbf{y} = \mathbf{R}\mathbf{f}_0, \tag{2}$$

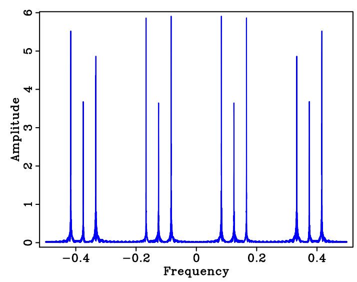
where $\mathbf{y} \in \mathbb{R}^n$ represents the data acquired at sub-Nyquist rate, $\mathbf{f}_0 \in \mathbb{R}^N$ the densely-sampled data to be recovered, and $\mathbf{R} \in \mathbb{R}^{n \times N}$ the restriction operator that selects the acquired samples among the desired samples. Assume $\mathbf{f}_0$ has a sparse representation $\mathbf{x}_0$ in some transform domain $\mathbf{S}$ where random sampling creates incoherent noise (e.g. Fourier domain, $\mathbf{S} := \mathbf{F}$), then interpolation to a dense grid becomes a denoising problem in the $\mathbf{S}$ domain (Donoho et al., 2007). This problem is solved by the following nonlinear sparsity-promoting optimization (Candes et al., 2005; Hennenfent and Herrmann, 2006)

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} ||\mathbf{x}||_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{R}\mathbf{S}^H\mathbf{x}, \tag{3}$$

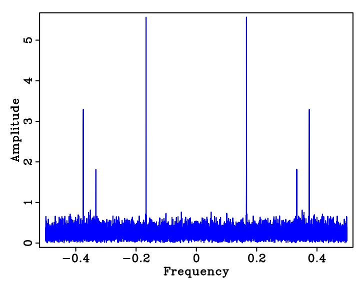
where the symbol ˜ represents an estimated quantity, and ^{H} the conjugate tranpose. The interpolated result is given by $\tilde{\mathbf{f}} = \mathbf{S}^H\tilde{\mathbf{x}}$. Note that, if the sampling at sub-Nyquist rate creates coherent noise in the $\mathbf{S}$ domain (see e.g. Fig. 1(b) when $\mathbf{S} := \mathbf{F}$), the separation between signal and noise generated by the acquisition in the $\mathbf{S}$ domain is much more delicate (if not impossible) just by imposing a sparsity constraint. We illustrate this comment by a simple 1-D experiment.
We define $\mathbf{S}$ as the discrete cosine transform. We generate a vector $\mathbf{x}_0$ of length $N = 600$ containing k nonzero entries with random amplitudes, random signs, and random positions and construct $\mathbf{f}_0 = \mathbf{S}^H\mathbf{x}_0$. The observations $\mathbf{y}$ are obtained by down-sampling $\mathbf{f}_0$ either regularly or randomly by a factor of $2, \ldots, 6$. Finally, we solve Eq. 3 and compare the estimated representation $\tilde{\mathbf{x}}$ of $\mathbf{f}_0$ to its true representation $\mathbf{x}_0$. The reconstruction error is measured as the number of false detections in the discrete cosine transform domain. The results presented in Fig. 2 are averaged over



(a)



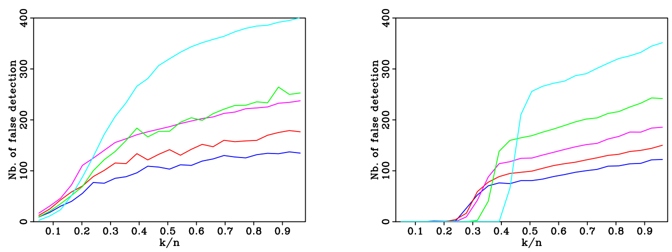
(b)



(c)

Figure 1: Spectra of a signal sampled above and below Nyquist rate. The signal consists of the superposition of three cosine functions. Amplitude spectrum of the densely-sampled signal (a), coarse regularly-sampled signal (b), and coarse randomly-sampled signal (c).

50 independent experiments. Figs. 2(a) and 2(b) show the recovery curves for regular and random downsampling, respectively. Each curve represents the results obtained for a given subsampling factor. Fig. 2(a) shows that, regardless of the subsampling factor and the sparsity of $\mathbf{f}_0$ in the discrete cosine transform domain, the solution of Eq. 3 is corrupted by some noise. However, Fig. 2(b) shows that there is a sparsity zone for each subsampling factor such that the reconstruction is perfect, i.e. no false detection. The smaller the subsampling factor, the wider the perfect reconstruction zone.



(a) (b)

Figure 2: Average recovery curves from sub-Nyquist rate samplings using nonlinear sparsity-promoting optimization. Average recovery curves (a) from regular sub-samplings by $2, \ldots, 6$ curves and (b) from random sub-samplings by the same factors as in (a)

NUMERICAL RESULTS

In the seismic context, the effect of coarse sampling in the f-k domain is illustrated in Fig. 3. Figs. 3(a) and 3(d) show densely-sampled data and the corresponding amplitude spectrum. Figs. 3(b) and 3(e) show regularly sub-sampled data and the corresponding amplitude spectrum. Finally, Figs. 3(c) and 3(f) show randomly sub-sampled data and the corresponding amplitude spectrum. Note how random sampling creates incoherent noise across the spectrum.

Although Fourier does not provide the sparsest representation for seismic data, there exists successful interpolation algorithms that solve Eq. 3 with $\mathbf{S} := \mathbf{F}$ (see e.g. Zwartjes and Hindriks, 2001; Xu et al., 2005). We use the algorithm called curvelet reconstruction with sparsity-promoting inversion (Herrmann, 2005; Hennenfent and Herrmann, 2005, 2006; Herrmann and Hennenfent, 2007) since curvelets provide a sparser representation for seismic data than Fourier (see e.g. Candes et al., 2006; Hennenfent and Herrmannn, 2006). In this case, $\mathbf{S}$ is defined as the curvelet transform (Candes et al., 2006, and references therein). The incoherent noise generated by random sampling remains incoherent in the curvelet domain since curvelets are strictly localized in the f-k domain. Fig. 4(a) and 4(b) show the interpolation results for the data of Figs. 3(b) and 3(c), respectively. The signal-to-reconstruction-error ratios are 6.92 dB for regular sub-sampling and 13.78 dB for random sub-sampling. For the same number of receivers, coarse random sampling leads to a much better reconstruction than coarse regular sampling. When a minimum velocity constraint is imposed during the reconstruction process, the same conclusion holds although the difference is reduced.

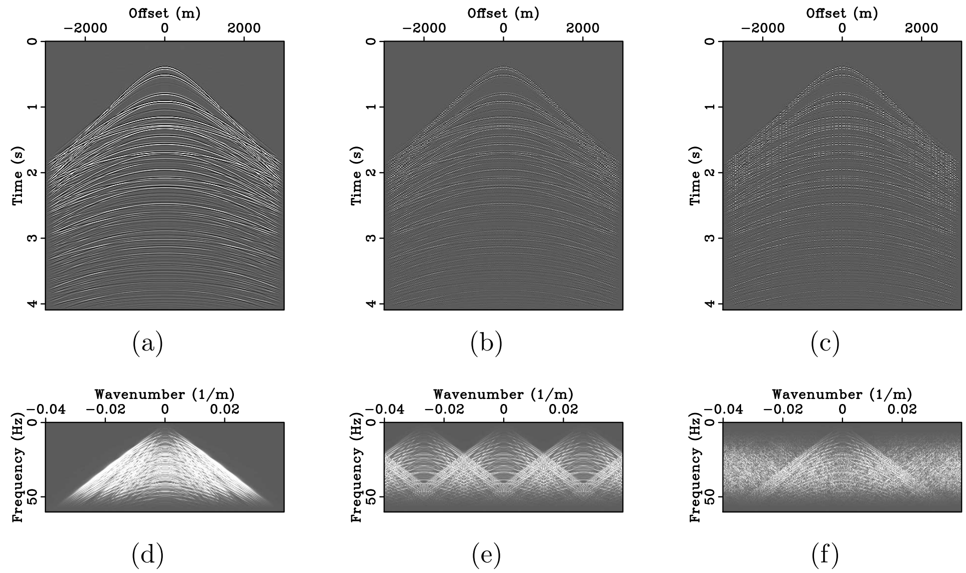


Figure 3: Seismic data and their corresponding spectrum. Densely-sampled data (a) and corresponding amplitude spectrum (d). Data regularly sampled below Nyquist rate (b) and corresponding amplitude spectrum (e) with strong aliasing beyond 25 Hz. Data randomly sampled at the same sub-Nyquist rate as (b) and corresponding amplitude spectrum (f) corrupted by broadband noise.

CONCLUSIONS

We proposed to look at the seismic data interpolation problem from a denoising perspective. From this standpoint, we showed that, for the same amount of data collected, regular subsampling geometries generate coherent acquisition noise more difficult to remove than the incoherent noise created by random subsampling geometries. Hence, random subsampling leads to a more accurate reconstruction of the seismic wavefield than equivalent regular subsampling or any subsampling that generates structured acquisition noise. We believe this new insight may lead to new acquisition strategies. On land, for example, a regular sampling may lead to (severely) aliased ground-roll that needs to be interpolated to a finer grid in order to be removed. Our observations suggest one should randomly sample on the finer grid instead. This leads to a better interpolation and hence ground-roll removal.

ACKNOWLEDGMENTS

This work was in part financially supported by NSERC Discovery Grant 22R81254 and CRD Grant DNOISE 334810-05 of F.J. Herrmann and was carried out as part

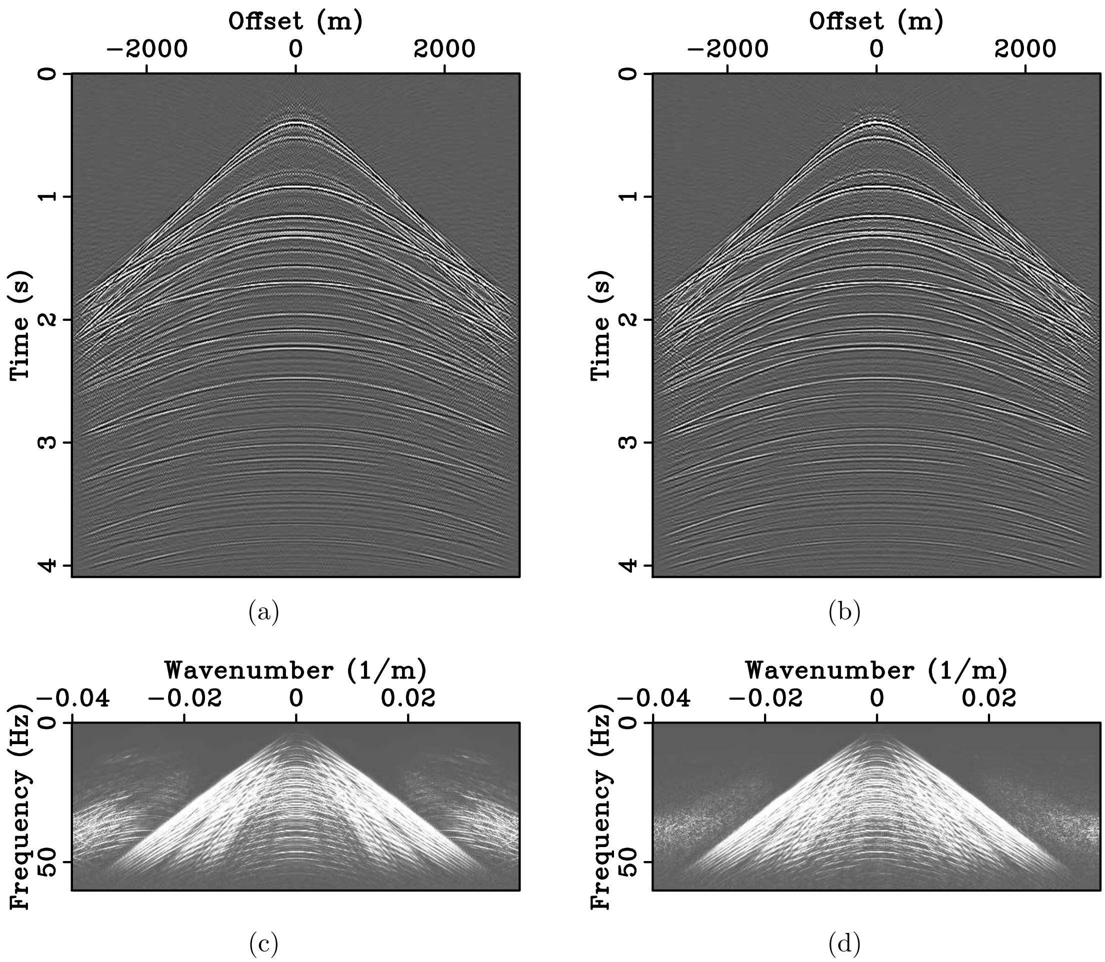


Figure 4: Synthetic seismic data reconstruction using 2-D curvelet reconstruction with sparsity-promoting inversion. Interpolation result – SNR = 6.9 dB – (a) and corresponding amplitude spectrum (c) given data of Fig. 3(b). Interpolation result – SNR = 13.78 dB – (b) and corresponding amplitude spectrum given data of Fig. 3(c). For the same number of receivers, coarse random sampling leads to a much better reconstruction than coarse regular sampling.

of the SINBAD project with support, secured through ITF, from the following organizations: BG Group, BP, Chevron, ExxonMobil and Shell.

REFERENCES

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

Candes, E., J. Romberg, and T. Tao, 2005, Stable signal recovery from incomplete and inaccurate measurements: Communications on Pure and Applied Mathematics, **59**, 1207–1223.

Dippe, M. and E. Wold, 1985, Antialiasing through stochastic sampling: Presented at the SIGGRAPH'85.

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

Donoho, D., Y. Tsaig, I. Drori, and J.-L. Starck, 2007, Sparse solution for underdetermined linear equations by stagewise orthogonal matching pursuit. (submitted).

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

———, 2006, Application of stable signal recovery to seismic interpolation: Presented at the SEG International Exposition and 76th Annual Meeting.

Hennenfent, G. and F. Herrmannn, 2006, Seismic denoising with non-uniformly sampled curvelets: Computing in Science and Engineering, **8**.

Herrmann, F., 2005, Robust curvelet-domain data continuation with sparseness constraints: Presented at the 67rd EAGE Annual Conference and Exhibition.

Herrmann, F. and G. Hennenfent, 2007, Non-parametric seismic data recovery with curvelet frames. (submitted).

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

Xu, S., Y. Zhang, and G. Lambare, 2005, Antileakage fourier transform for seismic data regularization: Geophysics, **70**, V87–V95.

Zwartjes, P. and C. Hindriks, 2001, Regularizing 3D data using Fourier reconstruction and sparse inversion: Presented at the 63rd EAGE Annual Conference and Exhibition.