Randomized marine acquisition "without" simultaneous sourcing

Hassan Mansour[1], *Haneet Wason*[2], *Tim T.Y. Lin*[2], *and Felix J. Herrmann*[2]

ABSTRACT

Seismic data acquisition in marine environments is a costly process that calls for the adoption of simultaneous-source or randomized acquisition - an emerging technology that is stimulating both geophysical research and commercial efforts. Simultaneous marine acquisition calls for the development of a new set of design principles and post-processing tools. In this paper, we discuss the properties of a specific class of randomized simultaneous acquisition matrices and demonstrate that sparsity-promoting recovery improves the quality of the reconstructed seismic data volumes. We propose a practical randomized marine acquisition scheme where the sequential sources fire airguns at only randomly time-dithered instances. We demonstrate that the recovery using sparse approximation from random time-dithering with a single source approaches the recovery from simultaneous-source acquisition with multiple sources. Established findings from the field of compressive sensing indicate that the choice of the sparsifying transform that is incoherent with the compressive sampling matrix can significantly impact the reconstruction quality. Leveraging these findings, we then demonstrate that the compressive sampling matrix resulting from our proposed sampling scheme is incoherent with the curvelet transform. The combined measurement matrix exhibits better isometry properties than other transform bases such as a non-localized multi-dimensional Fourier transform. We illustrate our results with simulations of "ideal" simultaneous-source marine acquisition, which dithers both in time and space, compared with periodic and randomized time-dithering.

INTRODUCTION

Data acquisition in seismic exploration forms one of the bottlenecks in seismic imaging and inversion. It involves the collection and processing of massive data volumes, which can be up to 5-dimensional in nature (2D for the source positions $\times$ 2D for the receiver positions $\times$ 1D for the time dimension). Constrained by the Nyquist sampling rate, the increasing sizes of these data volumes pose a fundamental shortcoming in the traditional sampling paradigm as the size and desired resolution of the survey areas continue to grow.

Conventional marine acquisition is carried out as single-source experiments of the subsurface response. This means that the size of the field recording is a product of the number of source locations, the number of receiver locations active per source experiment, and the number of discretized time samples proportional to the length of the reflection series. To reduce the threat of aliasing, the source and receiver locations are preferably bound to a grid spacing of less than 50 meters, therefore the number of receivers and sources are in turn direct functions of the size of the survey area.

Geological considerations, such as the presence of salt bodies in the Gulf of Mexico, create a need for wide azimuth coverage (i.e., receivers at large offsets), which pushes survey areas to cover thousands of square kilometres, with transverse lengths of 50 kilometres or more. However, adhering to a conventional single-source recording scheme requires recording vessels to move at a pace no more than 10 kilometres per hour to maintain the desired source spacing. This represents a direct conflict to the productivity of surveys, and makes large area acquisition particularly expensive. On the other hand, it is entirely possible and even desirable in terms of streamer stability for survey vessels to move faster (up to operational limits such as streamer drag), but doing so will not usually allow sufficient time for the single-source seismic responses to completely decay before the vessels have reached the next source position.

Several works in the seismic imaging literature have explored the concept of simultaneous or blended source activation to account for this situation Beasley et al. (1998); de Kok and Gillespie (2002); Beasley (2008); Berkhout (2008); Hampson et al. (2008). When sources are fired simultaneously, the main issue is the resulting interference between the responses of the different sources that makes it difficult to estimate interference-free shot gathers. Therefore, the challenge is to recover subtle late reflections that can be overlaid by interfering seismic response from other shots. We will show that this challenge can be effectively addressed through recovery by sparsity promotion.

Recently, "compressed sensing" (Donoho, 2006; Candès and Tao, 2006) has emerged as an alternate sampling paradigm in which randomized sub-Nyquist sampling is used to capture the structure of the data with the assumption that it is sparse or compressible in some transform domain. A signal is said to admit a sparse (or compressible) representation in a transform domain if only a small number k of the transform coefficients are nonzero (or if the signal can be well approximated by the k largest-in-magnitude transform coefficients). In seismic exploration, data consists of wavefronts that exhibit structure in multiple dimensions. With the appropriate data transformation, we capture this structure by a small number of significant transform coefficients resulting in a sparse representation of data.

In this paper, we rely on the compressed sensing literature to analyze a physically realizable simultaneous-source marine acquisition technology where acquisition related costs are no longer determined by the Nyquist sampling criteria. We also propose a random time-dithered acquisition scheme whose performance using a single source approaches that of simultaneous-source acquisition. Under this paradigm, data are no longer collected as separate shot records with single-source experiments. Instead, we continuously record over the whole acquisition process, collecting a single long "supershot" record that is acquired over a time interval shorter than the cumulative time of conventional marine acquisition (excluding downtime and overhead such as vessel turning). We then recover the canonical sequential single-source shot record by solving a sparsity promoting problem. The contributions of this work can be summarized as follows:

- We develop the relation between simultaneous-source sampling that is physically realizable in a marine setting and curvelet-based sparse recovery.
- We propose a random time-dithering marine acquisition scheme which can deliver even in the case of a single source a sparse recovery performance that approaches that of simultaneous-source acquisition with multiple sources with random time and space dithering.

- Through Monte-Carlo simulations, we estimate quantitative measures introduced in Compressive Sensing that predict the performance of sparse recovery for a particular acquisition design.
- Through simulated experiments, we demonstrate the performance of the proposed sampling matrices in recovering real prestack seismic data.

RELATED WORK

The earliest works on simultaneous acquisition were formulated with land acquisition in mind, where the prevalent use of vibroseis sweeps have allowed the freedom of employing sophisticated codes in source signatures as a way to differentiate between the responses due to different simultaneous sources (Allen et al., 1998; Beasley et al., 1998; Bagaini, 2006; Lin and Herrmann, 2009). Marine acquisition, on the other hand, rarely employs the marine vibrators that are the analog of the vibroseis due to poor signal-to-noise ratio and low-frequency content. Consequently, the marine case of simultaneous acquisition was less well-explored, as the more commonly used impulsive airgun sources are considerably more rigid when it comes to manipulating its signature. Literature directly discussing the marine case topic did not appear until Beasley et al. (1998), where the source "encoding" are limited to dithered activation times and locations. Near-simultaneous marine cases were also discussed in de Kok and Gillespie (2002) and Hampson et al. (2008). The main messages in these works seem to be that interferences due to simultaneous firing are often ignorable after stacking and simple filtering, and thus do not seriously impact the imaging step.

However, many subsurface attribute inversion schemes still rely on single-source prestack data. Recovering these volumes from simultaneous marine recordings did not truly become feasible until the recognition that, as long as the shot timings are suitably randomly delayed, the resulting interferences will appear noise-like in specific gather domains such as common-offset and common-receiver. This property differentiates these events from responses due to single-source experiments that remain coherent in these gather domains. This observation was reported in Stefani et al. (2007); Moore et al. (2008) and Akerberg et al. (2008) with application to land acquisition in Bagaini and Ji (2010). Subsequent processing techniques, which aim to remove noise-like source crosstalk, vary from simple filters (Huo et al., 2009) to more sophisticated inversion-type algorithms (Moore, 2010; Abma et al., 2010; Mahdad et al., 2011). The latter are designed to take advantage of sparse representations for coherent seismic signals.

The aforementioned works did not investigate the link between sparsity-based recovery and the specific properties of the acquisition system, but theoretical results from compressive sensing *do* suggest a direct relationship between acquisition design and the expected fidelity of the achievable recovery. An analysis for practical acquisition systems exist in terms of incoherency arguments (Blacquière et al., 2009,which interestingly also considers receiver-side blending), but analysis in terms of compressive sensing arguments remain challenging, as most existing mathematical results in compressive sensing deal with rather abstract acquisition systems. Evidently, existing works relating to seismic acquisition and compressive sensing only seem to suggest schemes suitable for forward-modelling in the computer. Neelamani et al. (2008, 2010) suggested an acquisition that uses noise-like signals as sweeps for land-based acquisition. Herrmann et al. (2009) on the other hand uses

impulsive sources but requires modulation of each source by a randomly-determined scaling in amplitude. Both of these papers also suggested recovery using sparse inversion of the curvelet representation of the data. In this paper we will keep the sparse inversion technique but consider more practical acquisition systems suitable for field marine acquisition.

COMPRESSED SENSING OVERVIEW

Compressive sensing (abbreviated as CS throughout the paper) is a process of acquiring random linear measurements of a signal and then reconstructing it by utilizing the prior knowledge that the signal is sparse or compressible in some transform domain. One of the main advantages of CS is that it combines sampling and compression in a single linear step, thus reducing the cost of traditional Nyquist sampling followed by dimensionality reduction through data encoding. A direct application which can benefit from this feature of CS is seismic acquisition where the acquisition costs are now quantified by the transform-domain sparsity of seismic data instead of by the grid size.

The sparse recovery problem

Suppose that $\mathbf{x}_0$ is an P dimensional vector with at most $k \ll P$ nonzero entries. The sparse recovery problem involves solving an underdetermined system of equations

$$\mathbf{b} = \mathbf{A}\mathbf{x_0}, \tag{1}$$

where $\mathbf{b} \in \mathbb{C}^n$, $n < P$ represents the compressively sampled data of n measurements, and $\mathbf{A} \in \mathbb{C}^{n \times P}$ represents the measurement matrix. When $\mathbf{x_0}$ is sparse—i.e., when there are only $k < n$ nonzero entries in $\mathbf{x_0}$— sparsity-promoting recovery can be achieved by solving the ℓ_0 minimization problem

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathbb{C}^P} \|\mathbf{x}\|_0 \quad \text{subject to } \mathbf{b} = \mathbf{A}\mathbf{x}, \tag{2}$$

where $\tilde{\mathbf{x}}$ represents the sparse approximation of $\mathbf{x_0}$, and the ℓ_0 norm $\|\mathbf{x}\|_0$ is the number of non-zero entries in $\mathbf{x_0}$. Note that if the ℓ_0 minimization problem were solvable in practice and every $n \times n$ submatrix of $\mathbf{A}$ is invertible, then $\tilde{\mathbf{x}} = \mathbf{x_0}$ when $k < n/2$ (Donoho and Elad, 2003).

However, ℓ_0 minimization is a combinatorial problem and quickly becomes intractable as the dimensions increase. Instead, the basis pursuit (BP) convex optimization problem shown below can be used to recover an estimate $\tilde{\mathbf{x}}$ at the cost of decreasing the level of recoverable sparsity k — e.g. $k \lesssim n/\log(N/n) < n/2$ when $\mathbf{A}$ is a Gaussian matrix with independent identically distributed (i.i.d.) entries (Candès et al., 2006; Donoho, 2006). The BP problem is given by

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathbb{C}^P} \|\mathbf{x}\|_1 \quad \text{subject to } \mathbf{b} = \mathbf{A}\mathbf{x}, \tag{3}$$

where $\tilde{\mathbf{x}}$ represents the sparse (or compressible) approximation of $\mathbf{x_0}$, and the ℓ_1 norm $\|\mathbf{x}\|_1$ is the sum of absolute values of the elements of a vector $\mathbf{x}$. The BP problem typically finds a sparse or (under some conditions) the sparsest solution that explains the data exactly.

Finally, we note that $\mathbf{x_0}$ can be the sparse expansion of a physical domain signal $\mathbf{f}_0 \in \mathbb{C}^N$ in some transform domain characterized by the operator $\mathbf{S} \in \mathbb{C}^{P\times N}$ with $P \geq N$. In this case, $\mathbf{A}$ can be composed of the product of a sampling operator $\mathbf{RM}$ with the sparsifying operator $\mathbf{S}$ such that $\mathbf{A} = \mathbf{RMS}^H$, where ^{H} denotes the Hermitian transpose. Consequently, the acquired measurements $\mathbf{b}$ are given by

$$\mathbf{b} = \mathbf{Ax_0} = \mathbf{RMf_0}.$$

We will elaborate more on this concept of sparse recovery in the randomized marine acquisition section.

Recovery conditions

Next, we discuss some conditions that make unique recovery possible despite the fact that the linear system we are solving is underdetermined, meaning that we have fewer equations than unknowns. We present two sets of conditions, those that guarantee recovery of any arbitrary signal $\mathbf{x_0}$, and those that are specialized for a particular class of signals $\mathbf{x_0}$.

Suppose that the vector $\mathbf{x_0}$ is an arbitrary signal that can be well approximated by the vector $\mathbf{x_k}$ which contains only the largest k coefficients of $\mathbf{x_0}$, i.e., the largest $k \ll P$ nonzero entries of $\mathbf{x_k}$ contain most of the energy of $\mathbf{x_0}$. Let a be a number larger than 1. As long as any subset Λ of ak columns of the $n \times P$ matrix $\mathbf{A}$ are linearly independent and constitute a submatrix $\mathbf{A}_\Lambda$ which is invertible and has a small condition number (close to 1), there exists some algorithm that recovers $\mathbf{x_0}$ exactly.

To quantify this property of $\mathbf{A}$, Candès and Tao (2005) define the *restricted isometry property* (RIP) which states that there exists a constant $0 < \delta_{ak} < 1$ for which

$$(1 - \delta_{ak})\|\mathbf{u}\|_2^2 \leq \|\mathbf{A}_\Lambda \mathbf{u}\|_2^2 \leq (1 + \delta_{ak})\|\mathbf{u}\|_2^2, \tag{4}$$

where Λ is any subset of $\{1 \dots P\}$ of size $|\Lambda| \leq ak$, $\mathbf{A}_\Lambda$ is the submatrix of $\mathbf{A}$ whose columns are indexed by Λ, and $\mathbf{u}$ is an arbitrary k-dimensional vector. The definition above indicates that if every submatrix $\mathbf{A}_\Lambda$ has an RIP constant that is close to zero, then its condition number approaches 1. More precisely, let σ_{min} and σ_{max} be the smallest and largest singular values of $\mathbf{A}_\Lambda$, respectively, the RIP constant

$$\delta_{ak} = \sup_{\Lambda \in \{1,\dots P\}} \max\{1 - \sigma_{\min}, \sigma_{\max} - 1\}. \tag{5}$$

That is, the RIP constant is the smallest upper bound on the maximum of $\{1-\sigma_{\min}, \sigma_{\max}-1\}$ for all subsets $\Lambda \in \{1, \dots P\}$ of size ak.

The RIP constant is difficult to compute since it requires evaluating δ_{ak} for every subset of ak columns of $\mathbf{A}$ and there are $\binom{P}{ak}$ of such subsets. However, it is possible to find theoretical upper bounds on the RIP constant for matrices whose entries are drawn i.i.d from sub-Gaussian distributions. Otherwise, Monte Carlo simulations are used to approximate the value of δ_{ak}. It was shown in (Candès et al., 2006) that if $\delta_{(a+1)k} < \frac{a-1}{a+1}$ (e.g. $\delta_{3k} < \frac{1}{3}$), then the BP problem can recover an approximation $\tilde{\mathbf{x}}$ to $\mathbf{x_0}$ with an error bounded by

$$\|\tilde{\mathbf{x}} - \mathbf{x_0}\|_2 \leq \frac{C(\delta_{ak})}{\sqrt{k}} \|\mathbf{x_k} - \mathbf{x_0}\|_1, \tag{6}$$

where $C(\delta_{ak})$ is a well-behaved constant. This error bound indicates that if the matrix $\mathbf{A}$ has the RIP for a specific sparsity level k, then the recovery error is bounded by the best k-term approximation error of the signal. Finally, we note that the condition on the RIP constant was later improved by Candès (2008) to $\delta_{2k} < \sqrt{2} - 1$.

An easier to compute but less informative characterization of $\mathbf{A}$ is the *mutual coherence* $\mu(\mathbf{A})$. The mutual coherence, which measures correlations between the columns of $\mathbf{A}$, provides an upper bound on $\delta_k < (k-1)\mu(\mathbf{A})$ and is given by

$$\mu(\mathbf{A}) = \max_{1 \le i \ne j \le P} \frac{|\mathbf{a}_i^H \mathbf{a}_j|}{(\|\mathbf{a}_i\|_2 \cdot \|\mathbf{a}_j\|_2)}, \tag{7}$$

where $\mathbf{a}_i$ is the i^{th} column of $\mathbf{A}$. It is evident from equation (7) that the mutual coherence is much easier to compute than the RIP constant. If we normalize the columns of $\mathbf{A}$ and form the Gram matrix $\mathbf{G} = \mathbf{A}^H\mathbf{A}$, the mutual coherence is then simply equal to the maximum off-diagonal element of $\mathbf{G}$. Therefore, the mutual coherence is the largest absolute normalized inner product between different columns of $\mathbf{A}$ (Bruckstein et al., 2009). Because near orthogonal matrices have small correlations amongst their columns, matrices with small mutual coherence favor recovery.

COMPRESSED SENSING AND RANDOMIZED MARINE ACQUISITION

Our focus in this section is on the design of source subsampling schemes that favor recovery in combination with the selection of the appropriate sparsifying transform. To illustrate the importance of transform-domain sparsity and mutual coherence, we include sparse recovery by the curvelet transform and the 3D Fourier transform in our simulations. This Fourier transform should not be confused with windowed Fourier transforms that are typically used for sparse recovery.

Randomized marine acquisition as a CS problem

Consider marine data organized in a seismic line with N_s sources, N_r receivers, and N_t time samples. For simplicity, we assume that all sources see the same receivers, which makes our method applicable to marine acquisition with ocean-bottom cables. The seismic line can be reshaped into an N dimensional vector $\mathbf{f}$, where $N = N_s N_r N_t$. It is well known that seismic data admit sparse representations by curvelets that capture "wavefront sets" efficiently (see e.g. Smith, 1998; Candès and Demanet, 2005; Candès et al., 2006; Herrmann et al., 2008,and the references therein). Therefore, we wish to recover a sparse approximation $\tilde{\mathbf{f}}$ of the discretized wavefield $\mathbf{f}$ from measurements $\mathbf{b} = \mathbf{RMf}$.

Let $\mathbf{S}$ be a sparsifying operator that characterizes a transform domain of $\mathbf{f}$, such that $\mathbf{S} \in \mathbb{C}^{P \times N}$ with $P \ge N$. When $\mathbf{S}$ is an orthonormal basis, i.e. $P = N$ and $\mathbf{SS}^H = \mathbf{S}^H\mathbf{S} = \mathbf{I}$ where $\mathbf{I}$ is the identity matrix, the signal $\mathbf{f}$ admits a unique transform domain representation $\mathbf{x} = \mathbf{Sf}$. On the other hand, if $\mathbf{S}$ is a tight frame with $P \ge N$ and $\mathbf{S}^H\mathbf{S} = \mathbf{I}$, as in the case of the redundant curvelet transform (Candès et al., 2006), then the expansion of $\mathbf{f}$ in $\mathbf{S}$ is not unique. The sparse approximation $\tilde{\mathbf{f}}$ is obtained by solving the inverse problem

$$\mathbf{A} := \mathbf{RMS}^H \tag{8}$$

with the basis pursuit sparsity-promoting program

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \tag{9}$$

yielding $\tilde{\mathbf{f}} = \mathbf{S}^H\tilde{\mathbf{x}}$. To solve this one-norm optimization problem, we use the SPGℓ_1 solver (Berg and Friedlander, 2008).

By solving a sparsity-promoting problem (Candès and Tao, 2006; Donoho, 2006; Herrmann et al., 2007; Mallat, 2009), it is possible to reconstruct high-resolution data volumes from the randomized samples at moderate sampling overhead compared to data volumes obtained after conventional compression (see e.g., Donoho et al. (1999) for wavelet-based compression, and Herrmann (2010) for empirical oversampling rates for seismic problems). As in conventional compression, the recovery error is controllable, but in the case of CS this recovery error depends on the sampling ratio $\gamma = \frac{n}{N}$. This ratio is given by the number of compressive samples and the number of conventionally acquired samples. From a simultaneous-source marine seismic perspective, this is the ratio between the size of the continuous and simultaneous recordings and the size of the conventional sequential data. From the perspective of the proposed single-source randomly time-dithered marine acquisition scheme, this is the ratio between the size of the randomly overlapping sequential source recordings and the size of the conventional non-overlapping sequential data. Consequently, in both the simultaneous-source and the randomized acquisition scenarios, the survey time is reduced for a fixed number of shots. Conversely, the number of shots recorded can also be increased given the same amount of recording time as a conventional survey, which is useful for projects afflicted with poor shot coverage.

Designing the randomized operator

The design of the linear sampling operator $\mathbf{RM}$ is critical to the success of the recovery algorithm. $\mathbf{RM}$ may in some cases be separable and composed of an $n \times N$ restriction matrix $\mathbf{R}$ multiplying an $N \times N$ mixing matrix $\mathbf{M}$. This is not true in the case of simultaneous-source or the single source random time-dithered marine acquisition where, as we will illustrate, the sampling operator $\mathbf{RM}$ is nonseparable. In the simultaneous marine acquisition scenario, the classic sequential acquisition with a single airgun is replaced with continuous acquisition with multiple airguns firing at random times and at random locations that span the entire survey area. This "ideal" simultaneous-source sampling scheme is illustrated in Figure 1(a) where the circles indicate the firing times and locations of the multiple sources. Such simultaneous acquisition schemes require an airgun to be located at each source location throughout the survey, which is infeasible. Alternatively, it may be possible yet costly to send out several vessels with airguns that swarm over an ocean-bottom array.

In this paper we present a new alternative which requires a very small number of vessels (possibly one) that map the entire survey area while firing sequential shots at randomly time-dithered instances. In the random time-dithered acquisition scheme a single airgun or multiple airguns are fired sequentially with random lag intervals between shots. This random time-dithered marine acquisition scheme is illustrated in Figure1(b) where, similar to the simultaneous source scheme, the firing times are still random but the source positions are sorted with respect to survey time. Remember that the ordered acquisition is still random by virtue of the random time shifts as opposed to the case of a periodic time-dithering

scheme where we simply decrease the intershot time delays as depicted in Figure 1(c). In the remainder of this paper we use the CS criteria of the previous section to analyze the efficacy of this random time-dithered scheme combined with the appropriate transform $\mathbf{S}$ and sparse approximation algorithm in recovering the discretized wavefield $\mathbf{f}$ from measurements $\mathbf{b} = \mathbf{RMf}$.

First we develop the structure of the sampling operator $\mathbf{RM}$. Suppose as before we have N_s shots, N_r receivers, and every shot decays after N_t time samples. We first map the seismic line into a series of sequential shots $\mathbf{f}$ of total length $N = N_s N_t N_r$ and apply the sampling operator $\mathbf{RM}$ to reduce $\mathbf{f}$ to a single long "supershot" of length $n \ll N$ that consists of a superposition of N_s impulsive shots. Again, to make the analysis more tractable, we ignore varying detector coverage by assuming a fixed receiver spread.

Since the subsampling/mixing is performed in the source-time domain, the resulting sampling operator is defined as follows

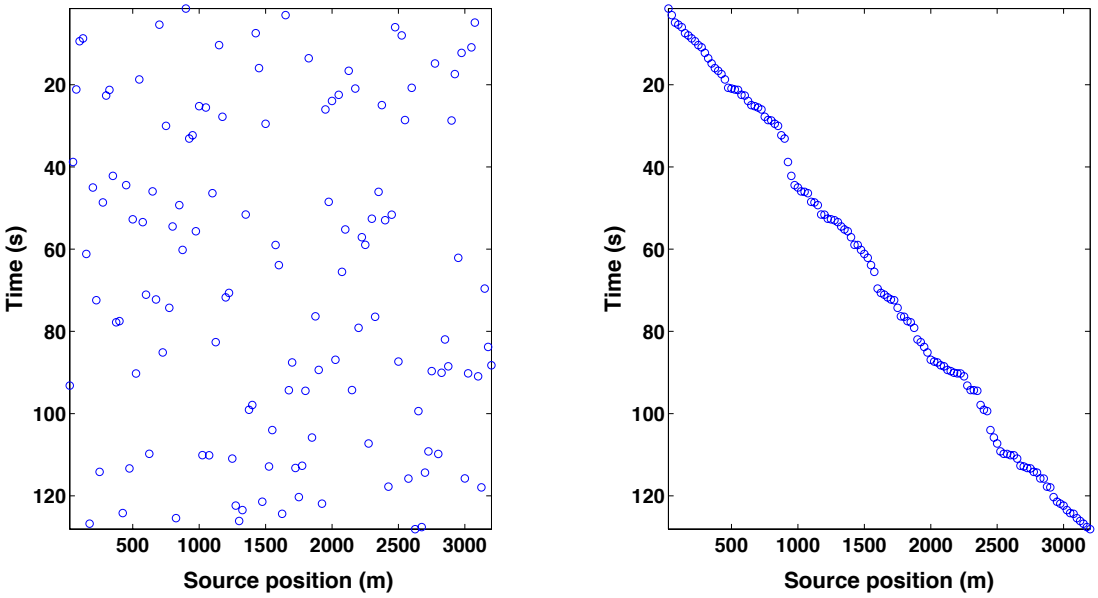
$$\mathbf{RM} := [\mathbf{I} \otimes \mathbf{T}], \tag{10}$$

where $\otimes$ is the Kronecker product, $\mathbf{I}$ is an $N_r \times N_r$ identity matrix, and $\mathbf{T}$ is a combined random shot selector and time shifting operator. This structure decouples the receiver axis from the source-time axis in the sampling operator allowing $\mathbf{T}$ to operate on the vectorized common receiver gathers. Taking the Kronecker product of $\mathbf{T}$ with $\mathbf{I}$ simply repeats the operation of $\mathbf{T}$ on every available receiver. The operator $\mathbf{T}$ turns the sequential-source recordings into continuous recordings with N_s impulsive sources, and firing at time instances selected uniformly at random from $\{1, \dots (N_s - 1)N_t\}$ discrete times[1]. Consequently, the operator $\mathbf{T}$ subsamples the $N_s N_t$ samples recorded at each receiver to $m \ll N_s N_t$ samples resulting in a total number $n = mN_r$ compressive samples[2].

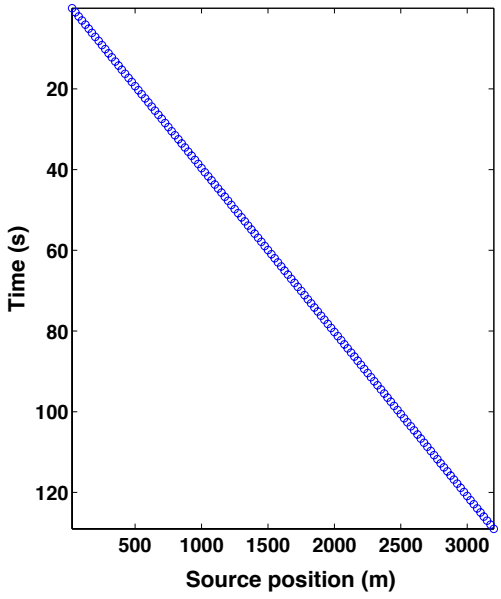
In the marine case with airgun sources, we can only work with binary $(0, 1)$-matrices because we have virtually no control over the source signature and energy output of airgun arrays. In the conventional sequential acquisition scheme where no overlap exists between the source responses, the operator $\mathbf{T}$ would be a block diagonal matrix of N_s blocks, each block being an $N_t \times N_t$ identity matrix. Each $N_t \times N_t$ identity matrix corresponds to the time taken for a source response to decay. This results in a large identity matrix of size $N_s N_t \times N_s N_t$. The simultaneous-source acquisition scheme destroys the block diagonal structure by placing the $N_t \times N_t$ identity matrices at random positions inside the matrix $\mathbf{T}$. An example of the corresponding operator is shown in Figure 2(a). In the case of random time-dithering, the $N_t \times N_t$ identity matrices are situated in an overlapped block diagonal structure as illustrated in Figure 2(b). The effect of random time-dithering on the structure of the operator $\mathbf{T}$ can be seen when we look at the operator in Figure 2(c) which corresponds to the periodic time-dithering scheme where we simply decrease the intershot time delays.

[1]It is possible to subsample the number of sources such that $n_s < N_s$ shots are selected uniformly at random from $\{1, \dots N_s\}$ source locations. Such a configuration requires a modified operator $\mathbf{T}$ in which the number of $N_t \times N_t$ identity submatrices is equal to n_s.

[2]It is also possible to subsample the receiver axis or equivalently to randomize the locations of the available receivers in order to produce a higher resolution receiver grid. This is achieved by replacing the $N_r \times N_r$ identity matrix in $\mathbf{RM}$ by an $n_r \times N_r$ restriction matrix that selects the $n_r < N_r$ physical receiver coordinates from the high resolution grid. Consequently, the number of collected CS measurements would be $n = mn_r$.

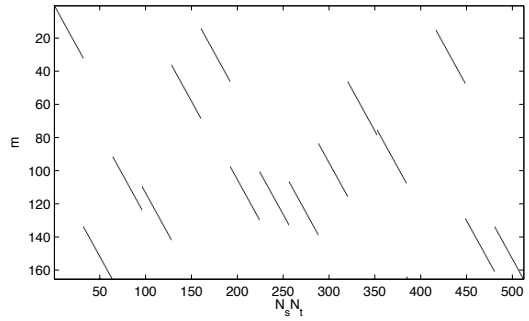


(a) (b)

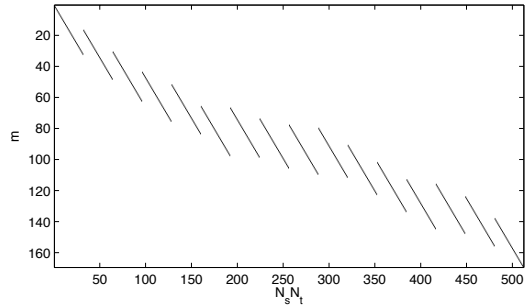


(c)

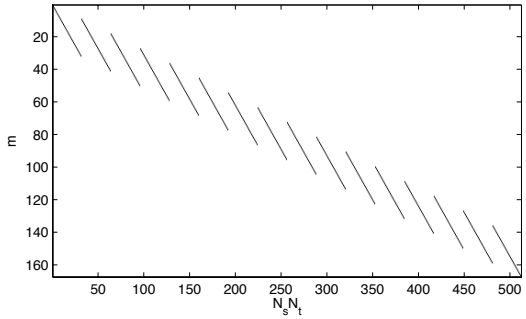
Figure 1: Examples of (a) random dithering in source location and trigger times, (b) sequential locations and random time-triggers, and (c) periodic source firing triggers.



(a)



(b)



(c)

Figure 2: Example of (a) "ideal" simultaneous-source operator defined by a Bernouilli matrix, (b) operator that corresponds to the more realizable Marine acquisition by the random time-dithering, and (c) sampling operator with periodic time-dithering.

Assessment of the sampling operators

In this section, we limit the assessment to the 2D case due to the large dimensionality of the data in the 3D case. Consequently, the sampling operator $\mathbf{A}$ that we analyze here is constructed as $\mathbf{A} = \mathbf{T}\mathbf{S}^H$, where $\mathbf{S}$ is a 2D curvelet or Fourier transform. The operator $\mathbf{A}$ is then the lowest dimensional nonseparable component of the 3D Kronecker structure. These results constitute worst case bounds for the 3D case since the performance of the Kronecker structure is bounded by the worst case performance of its components (Duarte and Baraniuk, 2011).

Mutual-coherence based assessment

The randomized time-dithering operator results in a measurement matrix $\mathbf{A}$ that exhibits a smaller mutual coherence compared to the periodic time-dithering sampling operator (cf. Figures 2(b) and Figure 2(c)). Notice that in both cases we fixed the number of source experiments and the number of collected samples. Aside from the randomization of the sampling operator, the choice of sparsifying transform also determines the mutual coherence. To study this combined effect, let us consider the Gram matrix $\mathbf{G} = \mathbf{A}^H\mathbf{A}$ of deterministic versus random sampling matrices $\mathbf{A}$ using either curvelets or Fourier as sparsifying transforms.

Recall that the mutual coherence $\mu(\mathbf{A})$ is given by the largest off-diagonal element of $\mathbf{G}$. As Figures 3 and 4 indicate, there is big difference between the coherences for the deterministic versus the randomized acquisitions. In fact, the mutual coherence of the curvelet based operator in this example is 0.695 for random time-dithering compared with 0.835 for periodic time-dithering. Similarly, the mutual coherence of the Fourier based operator is 0.738 fo random time-dithering compared with 0.768 for periodic time-dithering. For the Fourier and curvelet-based CS-sampling there is only a slight difference between the coherences despite significant differences in appearances of the Gram matrices. Therefore, the mutual coherence is a crude measure, which is confirmed in our experimental section.

RIP-based assessment

While the calculation of the mutual coherence is straight forward, it can only give us a pessimistic upper bound on the recoverable sparsity k of a signal (Bruckstein et al., 2009). A better bound (i.e. a guarantee for larger k) can be achieved by evaluating the restricted isometry property (RIP) of $\mathbf{A}$. Bounding this RIP constant allows us to guarantee recovery of less sparse (larger k) signals than what is guaranteed by the mutual coherence. The RIP constant δ_k of $\mathbf{A}$ is evaluated for all submatrices of $\mathbf{A}$ of size $n \times k$. Let Λ be a set of column indices of $\mathbf{A}$ of size k. For any matrix $\mathbf{A}_\Lambda$ the following property holds

$$\sigma_{\min}^2 \|\mathbf{u}\|_2^2 \leq \|\mathbf{A}_\Lambda \mathbf{u}\|_2^2 \leq \sigma_{\max}^2 \|\mathbf{u}\|_2^2,$$

where $\sigma_{\min}$ and $\sigma_{\max}$ are the smallest and largest singular values of the matrix $\mathbf{A}_\Lambda$, respectively. The RIP constant δ_k is the smallest upper bound on $\hat{\delta}_\Lambda := \max\{1 - \sigma_{\min}, \sigma_{\max} - 1\}$, i.e. $\delta_k = \max_\Lambda \hat{\delta}_\Lambda$. Consequently, if we can show that for all sets $\Lambda \in \{1, \dots P\}$, $\hat{\delta}_\Lambda < 1$ then there exists some solver which can recover any k-sparse signal. Moreover, if $\hat{\delta}_\Lambda < \sqrt{2} - 1$ or

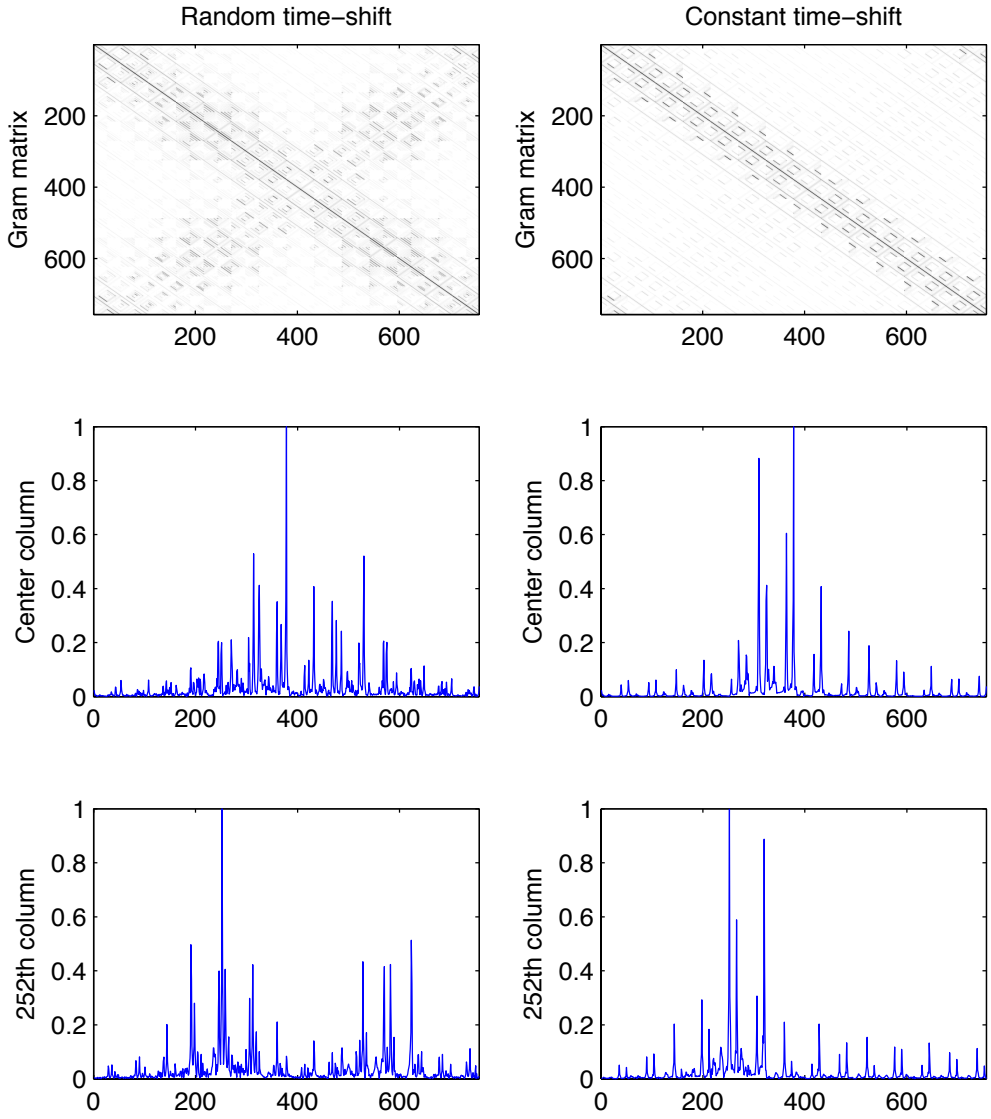


Figure 3: Gram matrices of example random time-dithering and constant time-dithering operators, top row, with $N_s = 10$ and $N_t = 40$ coupled with a **curvelet** transform. The resulting mutual coherence is 0.695 for random time-dithering compared with 0.835 for periodic time-dithering. The center plots show column the center column of the Gram matrices. The bottom row shows column 252 (one third) of the Gram matrices.

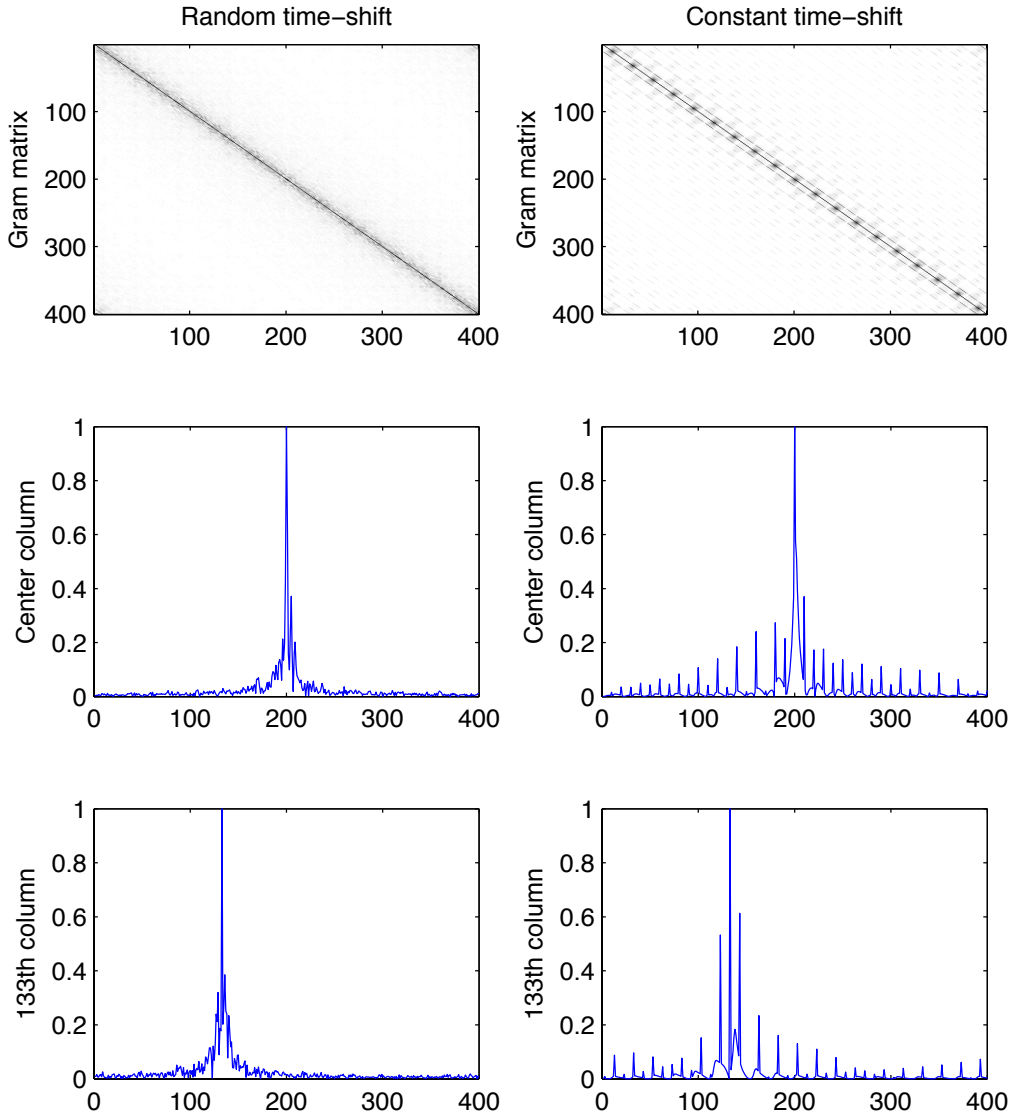


Figure 4: Gram matrices of example random time-dithering and constant time-dithering operators, top row, with $N_s = 10$ and $N_t = 40$ coupled with a **Fourier** transform. The resulting mutual coherence is 0.738 for random time-dithering compared with 0.768 for periodic time-dithering. The center plots show column the center column of the Gram matrices. The bottom row shows column 133 (one third) of the Gram matrices.

$\hat{\delta}_\Lambda < \frac{a-1}{a+1}$ for some integer $a > 1$, then with high probability we can guarantee that the BP program (3) can recover any sparse signal with sparsity less than or equal to $k/2$ or k/a, respectively.

Unfortunately, there are $\binom{P}{k}$ combinations of the $\mathbf{A}_\Lambda$ submatrices in $\mathbf{A}$, which makes evaluating $\hat{\delta}_\Lambda$ computationally infeasible in realistic settings. To overcome this difficulty, and since the transform coefficients of seismic images are often not strictly sparse, we first identify the appropriate k as the smallest number of transform coefficients that capture say 90% of the signal energy. This allows us to bound the recovery error in terms of the best k-term approximation of the signal. We are unaware of theoretical results that lead to a bound on $\hat{\delta}_\Lambda$ for our particular choice of $\mathbf{A}$. To overcome this, we estimate this quantity by Monte-Carlo sampling over different realizations of the sampling matrix and different realizations of the support. In our simulations, we generate 1000 realizations of the random time-dithering sampling matrix. For each of these matrices, we evaluate $\hat{\delta}_\Lambda$ for 100 random realizations of the set Λ. The RIP constant δ_k is estimated as the maximum of the computed values of $\hat{\delta}_\Lambda$.

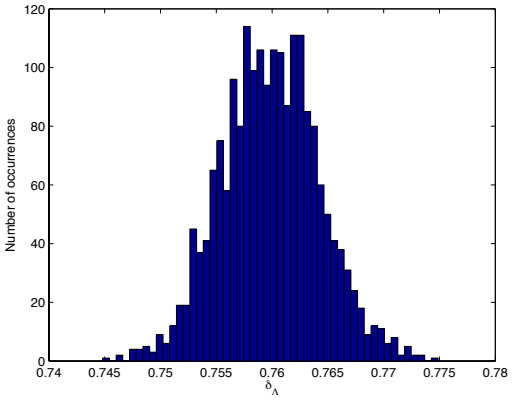
We plot the results of these simulations in Figure 5, which shows the histograms of $\hat{\delta}_\Lambda$ for the curvelet and Fourier transforms. These simulations show that the $\hat{\delta}_\Lambda$ for the curvelet transform are less than one, which means that this matrix has RIP, while the matrix based on Fourier may not have RIP for certain realizations of Λ. As a consequence, we can expect higher fidelity for curvelet-based recovery. Notice that in the simulations for the curvelet case, the estimate for $\hat{\delta}_\Lambda \approx 0.76$ is larger than the theoretical bound of $\sqrt{2}-1$ which guarantees stable recovery using BP with respect to the best $k/2$-term approximation $\mathbf{x}$. This is mainly due to the choice of k we used in our simulations. By choosing a smaller value for k, it would be possible to achieve the $\sqrt{2}-1$ mark at the expense of increasing the best k-term approximation error. For example, the condition $\delta_{9s} < 7/9$ guarantees stable recovery with respect to the best s-term approximation of the signal, where $8s = k$ and $k = |\Lambda|$ is the same as in the simulations above.

On the other hand, the RIP constant of the Fourier-based operator may exceed one. Therefore, it is not possible to find an s for which the RIP-based recovery conditions hold. We believe that this observation reflects the poorer recovery results of the Fourier-based operator compared to the curvelet-based operator as will be shown in the experimental results section.

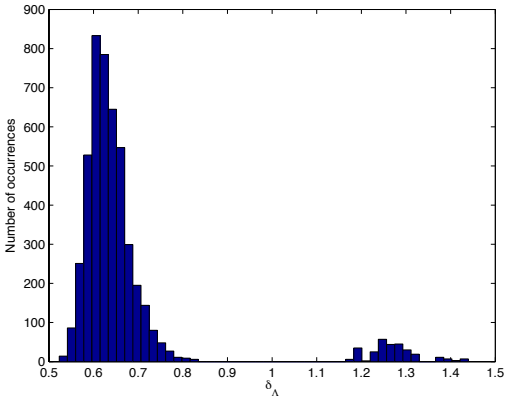
Economic considerations

Aside from mathematical factors, such as the mutual coherence that determines the recovery quality, there are also economical factors to consider. For this purpose, Berkhout (2008) proposed two performance indicators, which quantify the cost savings associated with simultaneous and continuous acquisition. The first measure compares the number of sources involved in conventional and simultaneous acquisition and is expressed in terms of the source-density ratio

$$\text{SDR} = \frac{\text{number of sources in the simultaneous survey}}{\text{number of sources in the conventional sequential survey}}. \tag{11}$$



(a)



(b)

Figure 5: Comparison between the histograms of $\hat{\delta}_\Lambda$ from 1000 realizations of $\mathbf{A}_\Lambda$, the random time-dithering sampling matrices $\mathbf{A} = \mathbf{RMS}^H$ restricted to a set Λ of size k, the size support of the largest transform coefficients of a real (Gulf of Suez) seismic image. The transform $\mathbf{S}$ is (a) the curvelet transform and (b) the nonlocalized 2D Fourier transform. The histograms show that randomized time-shifting coupled with the curvelet transform has better behaved RIP constant ($\hat{\delta}_\Lambda = \max\{1 - \sigma_{\min}, \sigma_{\max} - 1\} < 1$) and therefore promotes better recovery.

In the marine data acquisition setting, the SDR $= \frac{n_s}{N_s}$. Aside from the number of sources, the cost of acquisition is also determined by survey-time ratio

$$\text{STR} = \frac{\text{time of the conventional sequential survey}}{\text{time of the continuous and simultaneous recording}}. \tag{12}$$

The survey time ratio is therefore given by STR $= \frac{N_s N_t}{m}$ which is equal to the aspect ratio of the operator $\mathbf{T}$. The overall economic performance is measured by the product of these two ratios.

EXPERIMENTAL RESULTS

We illustrate the effectiveness of our simultaneous source acquisition approach by studying the performance of the three sampling schemes; simultaneous-source acquisition, random time-dithering, and periodic time-dithering, on a seismic line from the Gulf of Suez (Figure 6 shows a common-shot gather). The fully sampled sequential data is composed of $N_s = 128$ shots, $N_r = 128$ receivers and $N_t = 512$ time samples with 12.5 m source-receiver sampling interval. Prestack data from sequential sources is recovered using ℓ_1 minimization with 3D curvelets as the sparsifying transform. For comparison, we perform sparse recovery with a 3D Fourier transform and with the more rudimentary median filtering, which can also be used to suppress the crosstalk. We also perform linear recovery using the adjoint of the sampling operator followed by 2D median filtering in the offset domain for additional comparison.

We evaluate the recovery performance in terms of the *signal-to-noise ratio* (SNR) which is computed as follows for a signal $\mathbf{x}$ and its estimate $\tilde{\mathbf{x}}$:

$$\text{SNR}(\mathbf{x}, \tilde{\mathbf{x}}) = -20 \log_{10} \frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2}. \tag{13}$$

Simultaneous-source acquisition

We simulate simultaneous-source marine acquisition by randomly selecting 128 shots from the total survey time $t = N_s \times N_t$ with a subsampling ratio $\gamma = \frac{m}{N_s \times N_t} = \frac{1}{\text{STR}} = 0.5$. The subsampling is performed only in time so that the length of the "supershot" is half the length of the conventional survey time of sequential-source data. Figure 7(a) represents the "supershot" plotted as conventional survey by applying the sampling operator RM to sequential-source data. Notice that this type of "marine" acquisition is physically realizable only with a limited number of simultaneous sources, although truly random positioning of sources may still prove impractical depending on the manoeuvrability of source vessels.

To exploit continuity of the wavefield along all three coordinate axes, we use the 3D curvelet transform Ying et al. (2005). Figures 7(b) and 7(c) show the recovery and residual results, respectively. The recovered data volume has an SNR of 10.5 dB and was obtained with 200 iterations of solving the BP problem using SPGℓ_1.

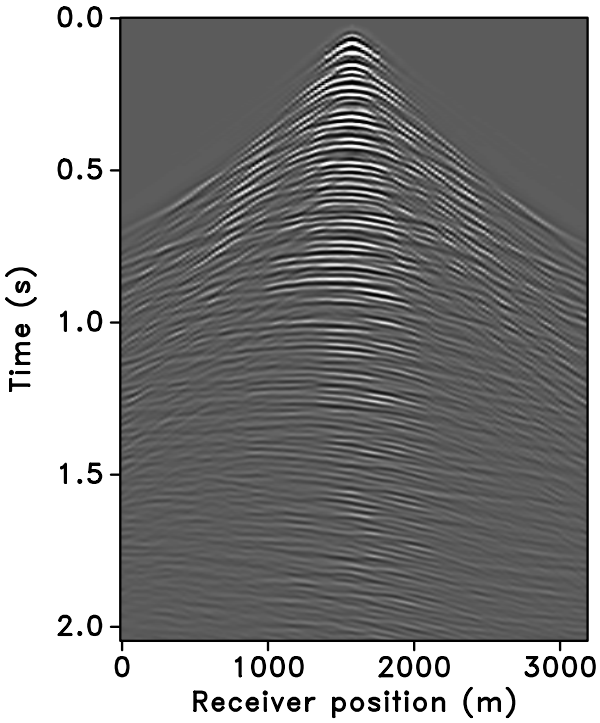


Figure 6: Common-shot gather from Gulf of Suez data set.

Random time-dithering

To overcome the limitation in the number of simultaneous sources required by the "ideal" simultaneous-source approach, we propose the random time-dithering scheme. Under this scheme, we allow all 128 shots to be fired sequentially with adjacent shots firing before the previous shot fully decays. We impose a random overlap between the shot records created by a random time lag between the firing of each shot. Therefore, the start time of each shot is chosen uniformly at random between the starting time of the previous shot and the time by which the previous shot record decays.

In our simulations, we apply a sampling operator with subsampling ratio $\gamma = 0.5$. A section of the "supershot" obtained by random time-dithering is shown in Figure 8(a). Using the 3D curvelet transform, a recovery of 8.06 dB is achieved (Figure 8(b)). Figure 8(c) shows the corresponding residual plot. Figure 9 summarizes the results for recovery based on the 3D Fourier transform. The recovered data volume has an SNR = 6.83 dB, which agrees with our predictions for RIP constants estimated in the previous section.

We also demonstrate the effectiveness of sparse recovery compared with linear recovery using the adjoint of the sampling operator $\mathbf{RM}$ followed by 2D median filtering in the midpoint-offset domain. The recovery results are shown in Figure 10. The resulting SNR, 3.92 dB, is considerably lower than the SNRs achieved by sparse recovery.

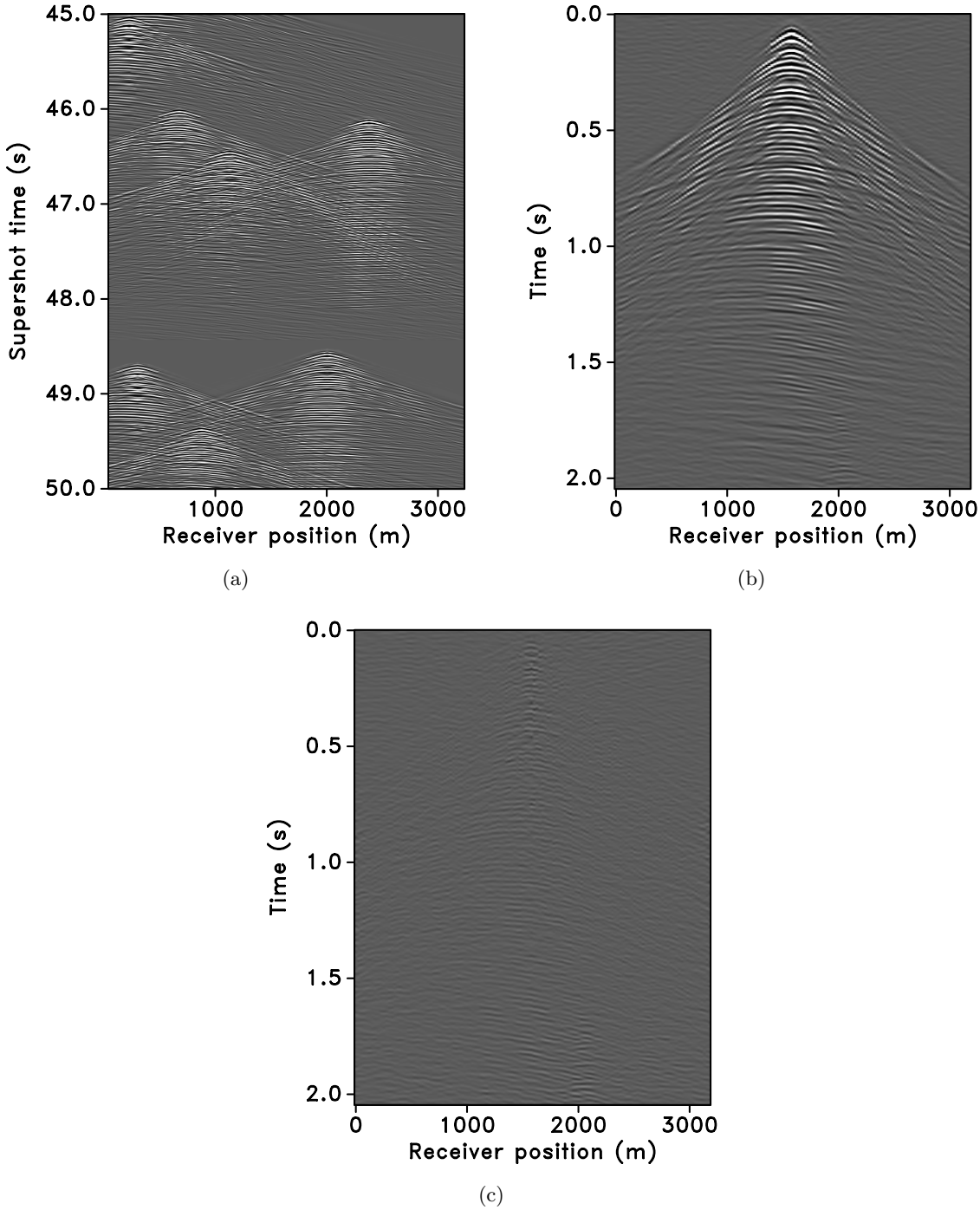


Figure 7: (a) Simultaneous-source marine data ($\gamma = 0.5$) shown as a section between 45 to 50 seconds of the "supershot". (b) Recovery from simultaneous 'marine' data (SNR = 10.5 dB). (c) Corresponding residual plot.

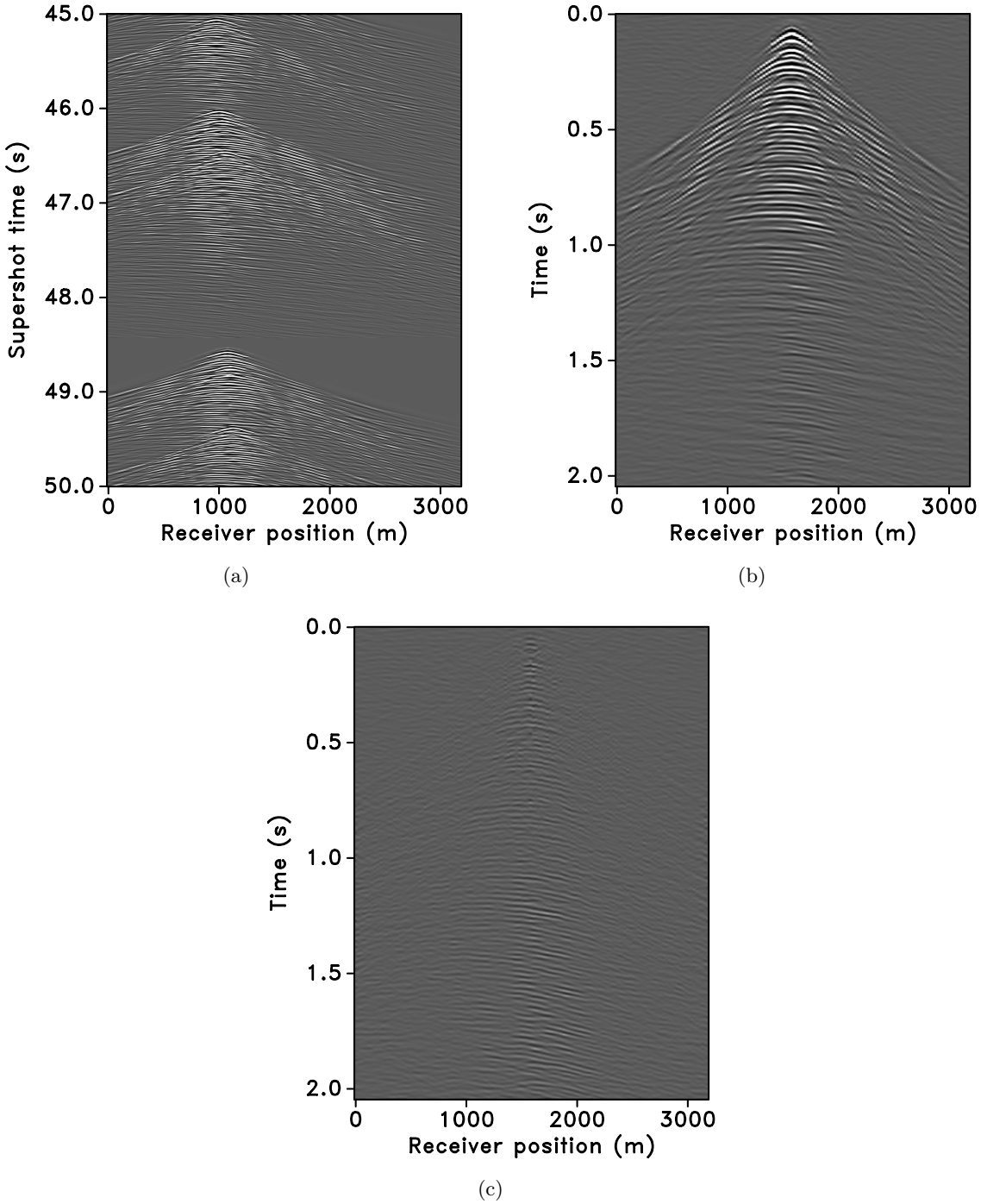
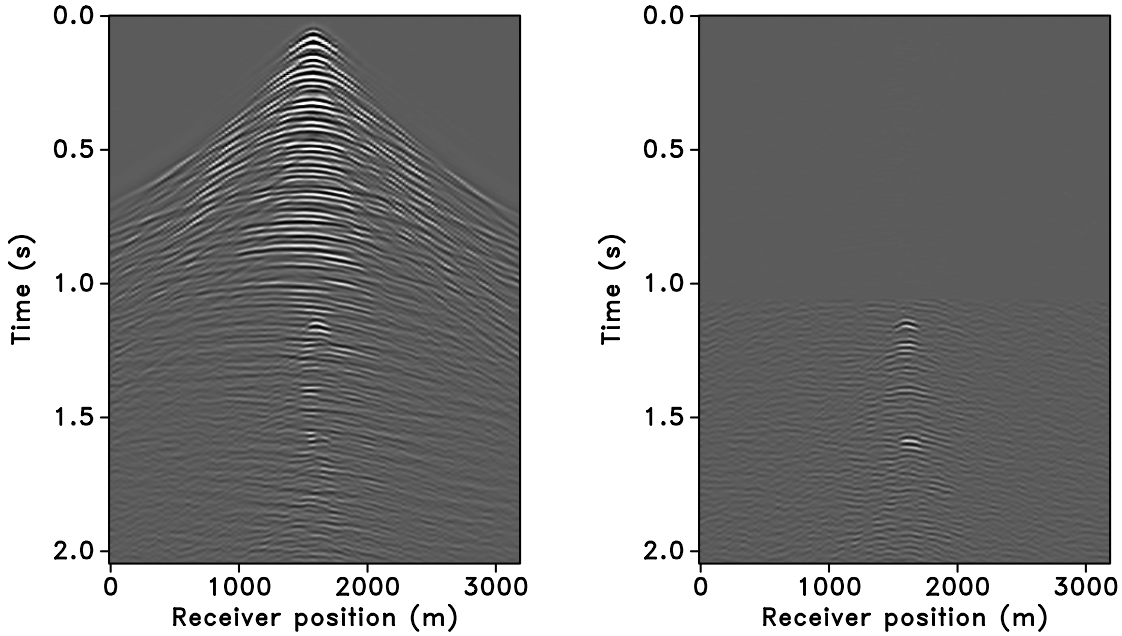
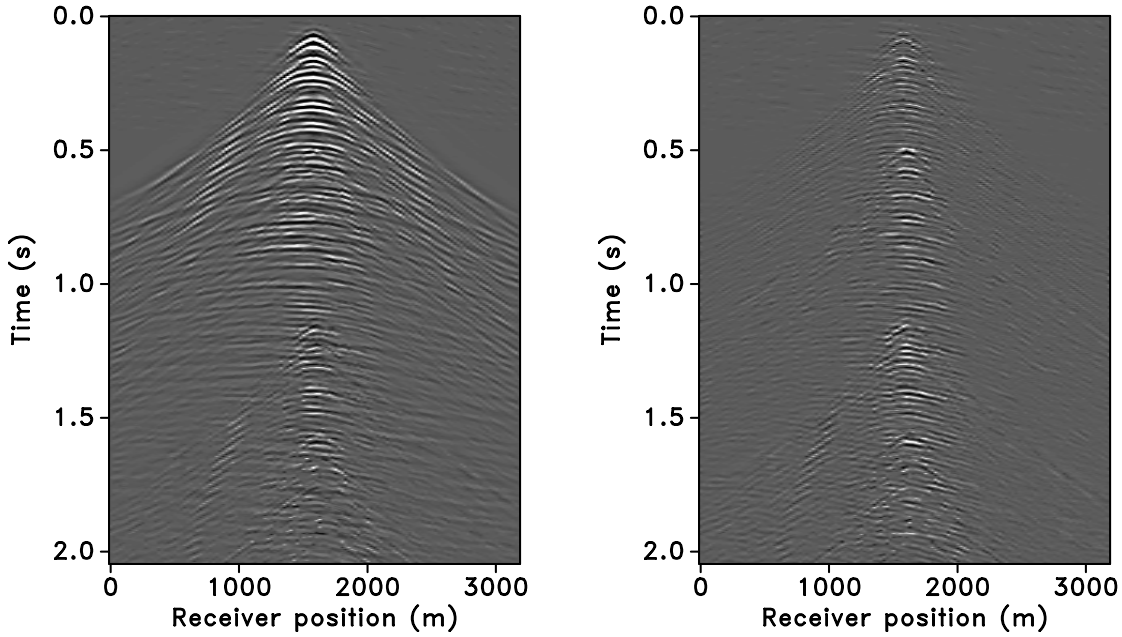


Figure 8: (a) Random time-dithered "marine" data ($\gamma = 0.5$) shown as a section between 45 and 50 seconds of the "supershot". (b) Sparse recovery with curvelet transform and SNR = 8.06dB. (c) Corresponding residual plot.



(a) (b)

Figure 9: (a) Sparse recovery with 3D Fourier transform from the same data shown in Figure 8(a), SNR = 6.83dB. (b) Corresponding residual plot.



(a) (b)

Figure 10: (a) Data recovered by applying adjoint of the sampling operator **RM** and 2D median filtering, from the same data shown in Figure 8(a), with SNR = 3.92dB. (b) Corresponding residual plot.

Periodic time-dithering

The importance of the "randomness" in time-dithering becomes evident when we compare the recovery of this sampling operator with that of a periodic time-dithering operator. Under the same subsampling conditions and sparsifying transform, the periodic time dither operator can only achieve an SNR = 4.80 dB. Figure 11 shows the periodic time-dithered "supershot", the recovered data volume and the corresponding residual. This poor performance is consistent with predictions of CS that require randomness in the design of sampling matrices. Furthermore, comparing sparse recovery with linear recovery of 1.26 dB (Figure 12), demonstrates the effectiveness of the former.

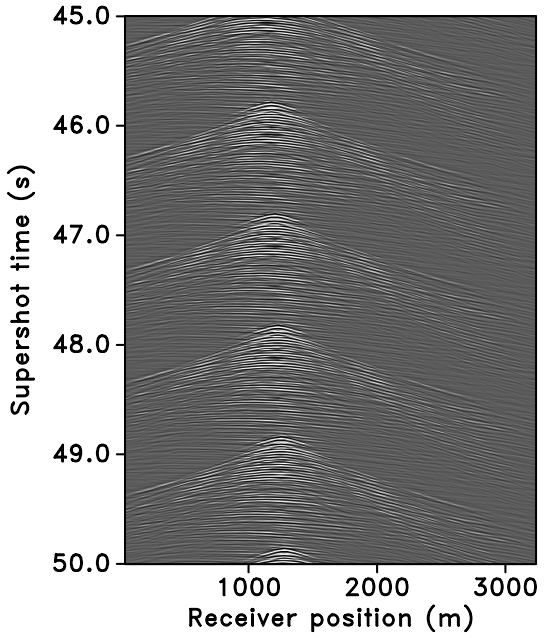
Finally, we illustrate the recovery for five subsampling ratios ($\gamma = 0.75, 0.50, 0.33, 0.25, 0.10$) for each of the schemes described above. Table 1 summarizes the SNRs for the three sampling schemes based on the 3D curvelet and 3D Fourier transforms.

Subsample ratio	**Simultaneous acquisition**		**Random time-dithering**		**Periodic time-dithering**	
1/STR	Curvelet	Fourier	Curvelet	Fourier	Curvelet	Fourier
0.75	13.0	10.2	11.2	9.44	6.93	4.93
0.50	10.5	7.06	8.06	6.83	4.80	2.42
0.33	8.31	4.50	5.33	4.10	7.32	1.37
0.25	6.55	2.93	4.35	2.88	2.85	0.89
0.10	2.82	0.27	1.14	0.20	1.60	0.19

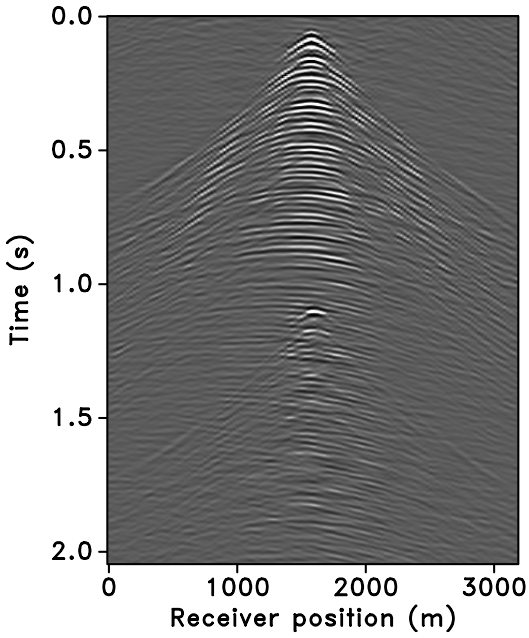
Table 1: Summary of recovery results (SNR in dB) based on the 3D curvelet and the 3D Fourier transforms for the three sampling schemes.

CONCLUSIONS

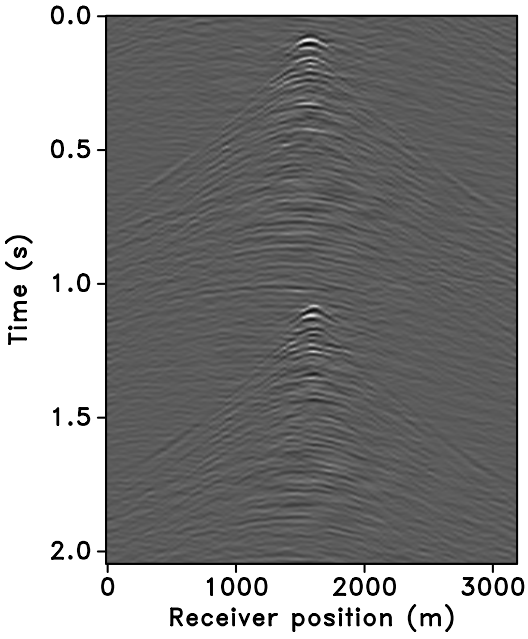
Recovering single-source prestack data volumes from simultaneously acquired marine data essentially involves removing noise-like crosstalk from coherent seismic responses. Many authors have noticed the important role of sparsity-based recovery for this problem, but few have thoroughly investigated the underlying interaction between acquisition design and reconstruction fidelity, especially in the marine setting. In contrast, we identify simultaneous marine acquisition as a linear subsampling system, which we subsequently analyze by using metrics from Compressive Sensing. We also propose a randomized time-dithering scheme which can match with a single source the performance of simultaneous-source acquisition. With the introduction of methods to calculate the mutual coherence and restricted isometry constants we are able to assert the importance of randomness in the acquisition system in combination with the appropriate choice for the sparsifying transform in the reconstruction. By comparing reconstructions on a real seismic marine line with different sparsifying transforms and sampled with different synthetic acquisitions, we quantitatively verified that more randomness in the acquisition system and more compressible transforms improve the mutual coherence and restricted isometry constants, which predict a higher reconstruction quality. As such this work represents a first step towards a comprehensive theory that predicts the reconstruction quality as a function of the type of acquisition.



(a)

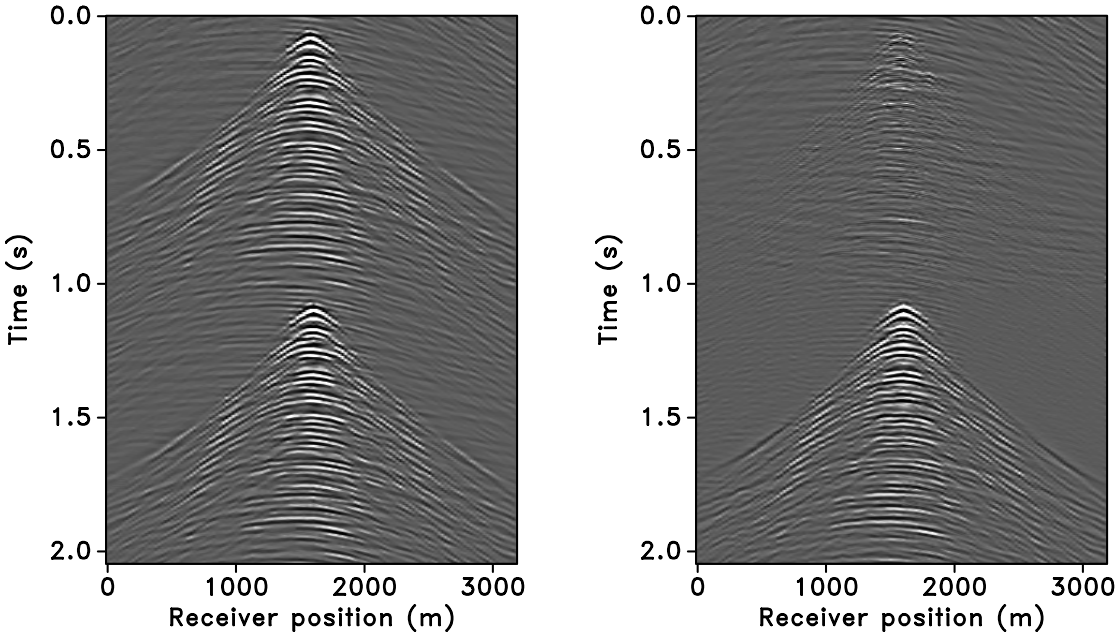


(b)



(c)

Figure 11: (a) Periodic time-dithered "marine" data ($\gamma = 0.5$) shown as a section between 45 and 50 seconds of the "supershot". (b) Sparse recovery with curvelet transform and SNR = 4.80dB. (c) Corresponding residual plot.



(a) (b)

Figure 12: (a) Data recovered by applying adjoint of the sampling operator **RM** and 2D median filtering, from the same data shown in Figure 11(a), with SNR = 1.26dB. (b) Corresponding residual plot.

ACKNOWLEDGMENTS

We would like to thank the authors of CurveLab (`curvelet.org`), a toolbox implementing the Fast Discrete Curvelet Transform, Madagascar (`rsf.sf.net`), a package for reproducible computational experiments, SPGℓ_1 (`cs.ubc.ca/labs/scl/spgl1`), SPOT (`http://www.cs.ubc.ca/labs/scl/spot/`), a suite of linear operators and problems for testing algorithms for sparse signal reconstruction, and pSPOT, SLIM's parallel extension of SPOT. The Gulf of Suez dataset was generously provided by Eric Verschuur. This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BP, Chevron, ConocoPhillips, Petrobras, Total SA, and WesternGeco.

REFERENCES

Abma, R., T. Manning, M. Tanis, J. Yu, and M. Foster, 2010, High quality separation of simultaneous sources by sparse inversion: Presented at the 72nd Ann. Internat. Mtg., EAGE, Eur. Ass. of Geosc. and Eng., Expanded abstracts.

Akerberg, P., G. Hampson, J. Rickett, H. Martin, and J. Cole, 2008, Simultaneous source separation by sparse radon transform: SEG Technical Program Expanded Abstracts, **27**, 2801–2805.

Allen, K. P., M. L. Johnson, and J. S. May, 1998, High fidelity vibratory seismic (hfvs) method for acquiring seismic data: SEG Technical Program Expanded Abstracts, **17**, 140–143.

Bagaini, C., 2006, Overview of simultaneous vibroseis acquisition methods: SEG Technical Program Expanded Abstracts, **25**, 70–74.

Bagaini, C., and Y. Ji, 2010, Dithered slip-sweep acquisition: SEG Technical Program Expanded Abstracts, **29**, 91–95.

Beasley, C. J., 2008, A new look at marine simultaneous sources: The Leading Edge, **27**, 914–917.

Beasley, C. J., R. E. Chambers, and Z. Jiang, 1998, A new look at simultaneous sources: SEG Technical Program Expanded Abstracts, **17**, 133–135.

Berg, E. v., and M. P. Friedlander, 2008, Probing the Pareto frontier for basis pursuit solutions: SIAM Journal on Scientific Computing, **31**, 890–912.

Berkhout, A. J., 2008, Changing the mindset in seismic data acquisition: The Leading Edge, **27**, 924–938.

Blacquière, G., G. Berkhout, and E. Verschuur, 2009, Survey design for blended acquisition: SEG Technical Program Expanded Abstracts, **28**, 56–60.

Bruckstein, A. M., D. L. Donoho, and M. Elad, 2009, From sparse solutions of systems of equations to sparse modeling of signals and images: SIAM Review, **51**, 34–81.

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

Candès, E., and T. Tao, 2006, Near-optimal signal recovery from random projections: Universal encoding strategies: Information Theory, IEEE Transactions on, **52**, 5406 –5425.

Candès, E. J., 2008, The restricted isometry property and its implications for compressed sensing: Comptes rendus-Mathématique, **346**, 589–592.

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

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

Candès, E. J., and T. Tao, 2005, Decoding by linear programming: IEEE Transactions on Information Theory, **51**, 4203–4215.

de Kok, R., and D. Gillespie, 2002, A universal simultaneous shooting technique: Presented at the 64th Ann. Internat. Mtg., EAGE, Eur. Ass. of Geosc. and Eng., Expanded abstracts.

Donoho, D., and M. Elad, 2003, Optimally sparse representation in general (nonorthogonal)dictionaries via ell^1 minimization: PNAS, **100**.

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

Donoho, P. L., R. A. Ergas, and R. S. Polzer, 1999, Development of seismic data compression methods for reliable, low-noise, performance: SEG Technical Program Expanded Abstracts, **18**, 1903–1906.

Duarte, M., and R. Baraniuk, 2011, Kronecker compressive sensing: to appear in IEEE Transactions on Image Processing.

Hampson, G., J. Stefani, and F. Herkenhoff, 2008, Acquisition using simultaneous sources: The Leading Edge, **27**, 918–923.

Herrmann, F. J., 2010, Randomized sampling and sparsity: Getting more information from fewer samples: Geophysics, **75**, WB173–187.

Herrmann, F. J., U. Boeniger, and D. J. Verschuur, 2007, Nonlinear primary-multiple separation with directional curvelet frames: Geoph. J. Int., **170**, 781–799. (doi:

10.1111/j.1365-246X.2007.03360.x).

Herrmann, F. J., Y. A. Erlangga, and T. T. Y. Lin, 2009, Compressive simultaneous full-waveform simulation: Geophysics, **74**, A35–40.

Herrmann, F. J., P. P. Moghaddam, and C. C. Stolk, 2008, Sparsity- and continuity-promoting seismic imaging with curvelet frames: Journal of Applied and Computational Harmonic Analysis, **24**, 150–173. (doi:10.1016/j.acha.2007.06.007).

Huo, S., Y. Luo, and P. Kelamis, 2009, Simultaneous sources separation via multi-directional vector-median filter: SEG Technical Program Expanded Abstracts, **28**, 31–35.

Lin, T., and F. J. Herrmann, 2009, Designing simultaneous acquisitions with compressive sensing: EAGE, Expanded Abstracts.

Mahdad, A., P. Doulgeris, and G. Blacquiere, 2011, Separation of blended data by iterative estimation and subtraction of blending interference noise: Geophysics, **76**, Q9–Q17.

Mallat, S. G., 2009, A Wavelet Tour of Signal Processing: the Sparse Way: Academic Press.

Moore, I., 2010, Simultaneous sources - processing and applications: Presented at the 72nd Ann. Internat. Mtg., EAGE, Eur. Ass. of Geosc. and Eng., Expanded abstracts.

Moore, I., B. Dragoset, T. Ommundsen, D. Wilson, C. Ward, and D. Eke, 2008, Simultaneous source separation using dithered sources: SEG Technical Program Expanded Abstracts, **27**, 2806–2810.

Neelamani, R., C. E. Krohn, J. R. Krebs, M. Deffenbaugh, J. E. Anderson, and J. K. Romberg, 2008, Efficient seismic forward modeling using simultaneous random sources and sparsity: SEG Technical Program Expanded Abstracts, **27**, 2107–2111.

Neelamani, R. N., C. E. Krohn, J. R. Krebs, J. K. Romberg, M. Deffenbaugh, and J. E. Anderson, 2010, Efficient seismic forward modeling using simultaneous random sources and sparsity: Geophysics, **75**, WB15–WB27.

Smith, H. F., 1998, A Hardy space for Fourier integral operators.: J. Geom. Anal., **8**, 629–653.

Stefani, J., G. Hampson, and E. Herkenhoff, 2007, Acquisition using simultaneous sources: Presented at the 69th Ann. Internat. Mtg., EAGE, Eur. Ass. of Geosc. and Eng., Expanded abstracts.

Ying, L., L. Demanet, and E. J. Candés, 2005, 3D discrete curvelet transform: Wavelets XI, Expanded Abstracts, SPIE, 591413.