Unified compressive sensing framework for simultaneous acquisition with primary estimation

Tim T.Y. Lin and Felix J. Herrmann, University of British Columbia, EOS*

SUMMARY

The central promise of simultaneous acquisition is a vastly improved crew efficiency during acquisition at the cost of additional post-processing to obtain conventional source-separated data volumes. Using recent theories from the field of compressive sensing, we present a way to systematically model the effects of simultaneous acquisition. Our formulation form a new framework in the study of acquisition design and naturally leads to an inversion-based approach for the separation of shot records. Furthermore, we show how other inversion-based methods, such as a recently proposed method from van Groenestijn and Verschuur (2009) for primary estimation, can be processed together with the demultiplexing problem to achieve a better result compared to a separate treatment of these problems.

INTRODUCTION

The physical limitations of seismic data acquisition have traditionally been one of the strongest driving forces behind the development of signal processing techniques. Well-known examples of this include interpolation, a post-processing step called seismic regularization, and deconvolution for the (sweep) source functions. Similarly, the high cost of acquisition work compels the adoption of simultaneous acquisition. Aimed at improving the performance of marine- and land-acquisition crews, simultaneous acquisition calls for development of a new set of design principles and post-processing tools. In this paper, we focus on new techniques to separate (demultiplex) simultaneously acquired data (Beasley, 2008; Krohn and Neelamani, 2008; Berkhout, 2008; Neelamani et al., 2008; Herrmann et al., 2009).

In our approach, we leverage developments from the field of compressive sensing (CS in short throughout the paper, Candès et al., 2006; Donoho, 2006)—where the argument is made, and rigorously proven—that compressible signals can be recovered from severely sub-Nyquist sampling by solving a sparsity promoting program. The CS approach differs from most simultaneous acquisition/recovery schemes because it combines three indispensable components namely, (i) the design of subsampling schemes that turn coherent sub-sampling interferences into harmless Gaussian-like noise (see e.g. Hennenfent and Herrmann (2008), where this principle is used to recover seismic data volumes from missing traces with curvelet-domain sparsity), (ii) the selection of a sparsifying domain (such as curvelets) in which the data can be represented parsimoniously, and (iii) the use of sparsity-promoting programs to recover the source-separated data volumes. As long as a sampling/recovery scheme adheres to this principles (this is not true for the majority of the current simultaneous acquisition strategies where typically one of the components is missing), CS guarantees recovery to high fidelity as long as the degree of sub-sampling is commensurate with the transform-domain sparsity. This means that sparser signals allow for larger degrees of subsamplings as long as these subsamplings are carried out according to CS principles.

By using CS principles, we present a rigorous framework for simultaneous acquisition and subsequent recovery by sparsity promotion. For the actual design of the subsampling scheme, we are motivated by recent work of Neelamani et al. (2008), and Herrmann et al. (2009) where phase-encoded simultaneous sources were used to reduce the computational cost of wavefield simulations. Difference, here is that we include a sweep function that makes our methodology relevant for land acquisition with vibroseis trucks. We use the principle of superposition, which allows us to work with simultaneous sources where all sources are firing.

SIMULTANEOUS SOURCES AS A CASE OF COMPRESSIVE SENSING

Compressive sensing (CS) theory proves that recovery through sparsity promotion is possible from a sample size m that is proportional to the signal's sparsity (here, the number of nonzeros, k) as opposed to the signal length N. The main contribution of this paper is to recognize simultaneous acquisition as an instance of compressive sensing and taking advantage of it in our formulation of the simultaneous acquisition problem. The benefits afforded by a well-designed subsampling scheme designed according to the compressive sensing framework are (i) an improved demultiplexing into source-separated data volumes by recovery through transform-domain sparsity promotion, and (ii) compression of imaging operators through a reduction of the number of sources (i.e., number of right-hand sides for the wavefield simulators) and number of frequencies per simultaneous source (i.e., the number of block-diagonals in the discretization of the Helmholtz equation) (Herrmann et al., 2009). Before discussing an example of practical source design for this subsampling scheme, let us first review the different components that allow us to recover from simultaneous data.

The compressive-sampling matrix: The success of compressive simulation depends on devising a subsampling of the physically distinct source and frequency axes where coherent interferences are turned into random noise (Hennenfent and Herrmann, 2008). We follow recent work by Romberg (2008) and implement the CS matrix through a random phase encoder in Fourier space. To maximize independence amongst the sources, we apply different restrictions for each of the n_s' simultaneous shots—i.e., we have

$$\mathbf{RM} = \overbrace{\begin{bmatrix} \mathbf{R}_1^{\Sigma} \otimes \mathbf{I} \otimes \mathbf{R}_1^{\Omega} \\ \vdots \\ \mathbf{R}_{n_{s'}}^{\Sigma} \otimes \mathbf{I} \otimes \mathbf{R}_{n_{s'}}^{\Omega} \end{bmatrix}}^{\text{sub sampler}} \overbrace{\left(\mathbf{F}_{\mathrm{sr}}^{*} \mathrm{diag}\left(e^{\hat{i}\boldsymbol{\theta}}\right) \mathbf{F}_{\mathrm{sr}} \otimes \mathbf{I}\right) \mathbf{F}_{\mathrm{t}}}^{\text{random phase encoder}}, \qquad (1)$$

with $\mathbf{F}_{\mathrm{t}}$ the Fourier transform in the time axis, and $\mathbf{F}_{\mathrm{sr}}$ the 2D-Fourier transform in the shot-receiver plane. $\boldsymbol{\theta}$ is a randomly determined phase chosen from Uniform$([0,2\pi])$. The matrices $\mathbf{R}^{\Omega}$ and $\mathbf{R}^{\Sigma}$ represent CS-subsampling operators acting along the rows (frequency coordinate) and columns (shot coordinate) of the data volume, respectively. As shown by Herrmann et al. (2009) application of this CS-sampling matrix, $\mathbf{RM}$, to the data is equivalent to applying it to the source wavefields directly turning single-source shots into a subset ($n_s' \ll n_s$ with n_s the number of separated single-source shots) of time-harmonic simultaneous sources that are randomly phase encoded and that have, for each simultaneous shot, a different set of angular frequencies missing—i.e., there are $n_f' \ll n_f$ (with n_f the number of frequencies of fully sampled data) frequencies non-zero.

The sparsfying transform: Aside from proper CS sampling the recovery from simultaneous simulations depends on a sparsifying transform that compresses seismic data, is fast, and reasonably incoherent with the CS sampling matrix. We accomplish this by defining the sparsity transform as the Kronecker product between the 2-D discrete curvelet transform (Candès et al., 2006) along the source-receiver coordinates, and the discrete wavelet transform along the time coordinate—i.e., $\mathbf{S} := \mathbf{C} \otimes \mathbf{W}$ with $\mathbf{C}$, $\mathbf{W}$ the curvelet- and wavelet-transform matrices, respectively. The choice of using a wavelet representation in the time domain is justified by the high incoherence between the wavelet dictionary and the discrete Fourier transform (Herrmann et al., 2009). This makes wavelets a more ideal sparsity basis for reconstruction from missing frequencies according to CS theory.

Recovery by sparsity promotion: As prescribed by CS theory, we reconstruct the seismic wavefield by solving the following nonlinear but convex optimization problem

$$\widetilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 < \rho, \tag{2}$$

with $\widetilde{\mathbf{d}} = \mathbf{S}^*\widetilde{\mathbf{x}}$ recovered data, $\mathbf{A} := \mathbf{RMS}^*$ the CS matrix, $\mathbf{y}$ the simultaneously acquired data, and ρ a tolerance level on noise. This is solved with SPGℓ_1, a projected-gradient algorithm with root finding (Berg and Friedlander, 2008).

Comparison of interpolation and demultiplexing The optimization problem in Eq. 2 can also be used to describe a type of interpolation process called Curvelet Reconstruction by Sparsity-promoting Inversion (CRSI) discussed at length in Hennenfent and Herrmann (2008). In summary, the method prescribes solving Eq. 2 with $\mathbf{A} := \mathbf{RS}^*$ with $\mathbf{S}$ the 3D Curvelet transform and $\mathbf{R}$ acting only on either the receiver location (missing traces) or the shot location (missing shots), chosen in a special way as to minimize the effects of coherent interference from downsampling. CRSI can be interpreted as a method to deal with undersampled shots by interpolating missing data, while the approach presented in this abstract can be seen as dealing with undersampled shots by separating them out of records that also contain signals from all other shot locations. It is interesting to compare the results possible from both methods using the same inversion algorithm. Fig. 1 shows the recovery from 50% undersampled shots on a synthetic dataset using both CRSI and the approach presented in this abstract. The much higher SNR achieved from the demultiplexing approach strongly suggests that this is a better posed problem compared to an interpolation based approach.

INCORPORATION OF THE SOURCE SIGNATURE

Even though the above sampling scheme is viable as-is for forward modeling simulations, straightforward application of this scheme to seismic acquisition is impractical. The main hurdle is the fact that vibroseis trucks have physical constraints that limit the type of source functions they can emit. To incorporate this limitation into the model, we slightly refine the CS-sampling matrix defined in Eq. 1 by replacing the identity matrix $\mathbf{I}$ in the phase encoder with $\Psi = \operatorname{diag}\{\psi\}$, where ψ is a discretization of a pre-defined causal signature. This is mathematically equivalent to globally convolving each shot impulse in the source with the signature. A seismic surveyor may set this signature to be identical to the frequency sweeps programmed into the veibroseis, or may instead use a modulated sweep to reflect some knowledge of the coupling between the veibroseis and the earth surface. Note that any arbitrary source signature may be injected into the model in this way. Consequently if provided with a source signature which can be estimated through some external means, our refined model will be immediately extendable to all source types without modification to the underlying computational structure.

ESTIMATION OF PRIMARIES BY SPARSE INVERSION

Recent works on surface-related multiple removal include a method proposed by van Groenestijn and Verschuur (2009), where the primary impulse response is directly inverted from a linear operator which maps it to the up-going data wavefield. This operator is written as the sum of two terms; the first being a straightforward time-convolution with the source signature, while the second term is a non-stationary convolution with the up-going wavefield data. Expressed mathematically, the action of this operator $\mathbf{P}$ acting on the unknown primary impulse response $\mathbf{x}_0$ can be defined as

$$\mathbf{P}\mathbf{x}_0 = \mathbf{F}_{\mathrm{t}}^* \left[\psi \mathbf{X}_0 - \mathbf{X}_0 \mathbf{D}_{\mathrm{u}}\right]_{\omega} \mathbf{F}_{\mathrm{t}} = \mathbf{d}_{\mathrm{u}}, \tag{3}$$

where $\mathbf{d}_{\mathrm{u}}$ is the up-going data wavefield, $\mathbf{X}_0$ represent a single frequency slice in the shot-receiver plane of the primary impulse response and $\mathbf{D}_{\mathrm{u}}$ likewise for the up-going data wavefield, respectively. ψ is overloaded here to mean the individual coefficients of the source signature estimate in the frequency domain. The subscript ω on the inner bracket expression implies that the contained expression is carried out separately for each frequency.

According to the authors, a reasonable estimate of the primary impulse response can be obtained by an iterative steepest-descent inversion process on $\mathbf{P}$. The gradients of the objective function $f(\mathbf{x}) = \|\mathbf{d}_\mathrm{u} - \mathbf{P}\mathbf{x}\|_2^2$, is evaluated at $\widetilde{\mathbf{x}}_0$ (a current estimate on $\mathbf{x}_0$) according to

$$\nabla f(\widetilde{\mathbf{x}}_0) = 2\,(\mathbf{d}_\mathrm{u} - \mathbf{P}\widetilde{\mathbf{x}}_0)\,\mathbf{P}^*. \tag{4}$$

An update to $\delta\widetilde{\mathbf{x}}_0$ on $\widetilde{\mathbf{x}}_0$ is then obtained by picking the τ-th largest elements of the gradient and setting the rest to zero, followed by a scaling factor of $1/2$. The next update will then be calculated on the gradient of $(\widetilde{\mathbf{x}}_0 + \delta\widetilde{\mathbf{x}}_0)$. The process is repeatedly carried out until a desired image of the primary is formed, resulting in a method coined Estimation of Primaries by Sparse Inversion (EPSI) by its authors, who have reported that for the case of synthetic marine data a reasonable estimate can typically be obtained within 100 such steps.

It is illuminating to recognize that EPSI belongs to a class of so-called projected gradient methods in optimization. The implied goal of EPSI is to solve an instance of the following *non-convex* and NP-hard optimization problem:

$$\widetilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \quad \text{subject to} \quad \|\mathbf{x}\|_0 < k\tau, \tag{5}$$

where in the case of EPSI, $\widetilde{\mathbf{x}}$ corresponds to $\widetilde{\mathbf{x}}_0$ the estimated primary, operator $\mathbf{A}$ to the linear mapping $\mathbf{P}$, $\mathbf{y}$ to the data wavefield $\mathbf{d}_\mathrm{u}$, and k to the number of iterations taken in the EPSI process. The cardinality constraint $\|\mathbf{x}\|_0 < k\tau$ limiting the number of non-zeros in the solution makes the problem in Eq. 5 non-convex. As a result, such a method is not theoretically guaranteed to converge to a global solution. Indeed, it is reported by van Groenestijn and Verschuur (2009) that if the cardinality constraint for fixed k is imposed on the entire current estimate $\widetilde{\mathbf{x}}_0$ at each step, then ESPI ceases to converge in a reasonable amount of time. A workaround is to severely limit the size of the feasible set at every iteration, such as imposing constraint $\|\delta\widetilde{\mathbf{x}}_0\|_0 = \tau$, so that the update making the largest possible progress to the minimization objective in Eq. 5 can be found by simple searching over the whole set. It is easy to show that for $\|\delta\widetilde{\mathbf{x}}_0\|_0 = \tau$ this is done by zeroing everything in the gradient except the τ-th largest elements, exactly as prescribed EPSI. It is very possible to end up choosing overlapping element positions between successive updates, but this not a problem since the resulting solution the end will remain feasible for the original constraint $\|\mathbf{x}\|_0 < k\tau$.

Another method inspired by convex optimization theory is replacing the cardinality constraint $\|\mathbf{x}\|_0 < k\tau$ with a convex relaxation, namely a 1-norm constraint $\|\mathbf{x}\|_1 < \sigma$. This approach has been both theoretically and experimentally justified as being a very effective heuristic. In fact, showing that these two terms are substitutable under the assumption of sparsity in $\mathbf{x}$ is one of the main thesis of CS theory. More importantly, with this substitution the problem in Eq. 5 becomes convex, meaning that a global solution exists that can be solved in polynomial time using any number of widely studied algorithms. Serendipitously, the SPGℓ_1 algorithm with which we originally proposed to solve Eq. 2 is specifically designed to efficiently solve just such a problem, with the addition that it solves it for a number of different values of σ in an attempt to find the smallest σ that will satisfy Eq. 2. We therefore propose to combine the demultiplexing of seismic data and the estimation of primaries into the same optimization problem, by defining $\mathbf{A} := \mathbf{RMPS}^*$ and solving Eq. 2. Theoretically, this should yield three benefits:

(i) the computational overhead of the optimization algorithm and any associated costs such as computing the sparsity basis analysis can be amortized by solving the two problems together instead of separately

(ii) an implicit resistance to the effect of reconstruction errors in the separation on EPSI, due to the nature of the inverse problem

(iii) either an improvement in SNR for a given number of simultaneous shots, or the ability to further subsample the shots while retaining similar SNR, following the results shown in (Herrmann et al., 2009) where the sparsity of the underlying solution is associated with the quality of the recovery

The last argumens can also be intuited from the physical argument that the combined optimization problem treats the multiple energy as a redundant sampling of the primaries.

Simulation experiment

To illustrate our claim, we devise a experiment on a combined source separation and multiple removal problem. A synthetic marine dataset shown as a crossing-planes plot in Fig. 2(a) is subject to the CS sampling matrix $\mathbf{RM}$ with a 50% restriction in $\mathbf{R}$ on both the shot and frequency direction, for a combined 75% missing data on the whole volume. We model a typical seismic "sweep" type signature of veibroseis trucks by choosing $\psi(t) = \cos\left(2\pi(f_b t + \frac{f_e}{2}t^2)\right)$ with $f_b = 5\,\mathrm{Hz}$ and $f_e = 100\,\mathrm{Hz}$. The resulting transform on the source volume and the data wavefield can be seen in Fig. 2(b) and 2(c) respectively. Solution of Eq. 2 using $\mathbf{A} := \mathbf{RMS}^*$ is seen in Fig. 2(d) while a combined recovery using $\mathbf{A} := \mathbf{RMPS}^*$ which directly gives the primary impulse response is shown in Fig. 2(e). Note that the results in Fig. 2(d) and 2(e) are obtained using the same algorithm SPGℓ_1 ran for the same number of iterations. EPSI is then carried out on the solution wavefield in Fig. 2(d) by solving Eq. 2 with $\mathbf{A} := \mathbf{PS}^*$ setting ψ as a Ricker wavelet centered at 10Hz to estimate natural earth wave dispersion. The solution, seen in Fig. 2(f), shows how non-uniform errors in the wavefield separation severely weakens the effectiveness of EPSI since the primaries can no longer be matched perfectly to the surface-related multiples, even though the separated wavefield have a relatively high SNR of 16.3.

DISCUSSION AND CONCLUSIONS

Compressive sampling is considered a paradigm shift, and we have shown that simultaneous acquisition is a natural candidate for the application of its principles. Savings in acquisition time can be achieved through a deliberate reduction in the number of sweeps recorded, the extent of which we know through CS is commiserate with the complexity of the wavefield rather than the number of shot positions. Furthermore we can directly incorporate EPSI to improve the quality of the resulting signal while amortizing the cost of both separation and primary estimation. In our opinion, our proof of principle is encouraging and invites further investigation into the design and implementation of new acquisition schemes based on the principles from compressive sensing.

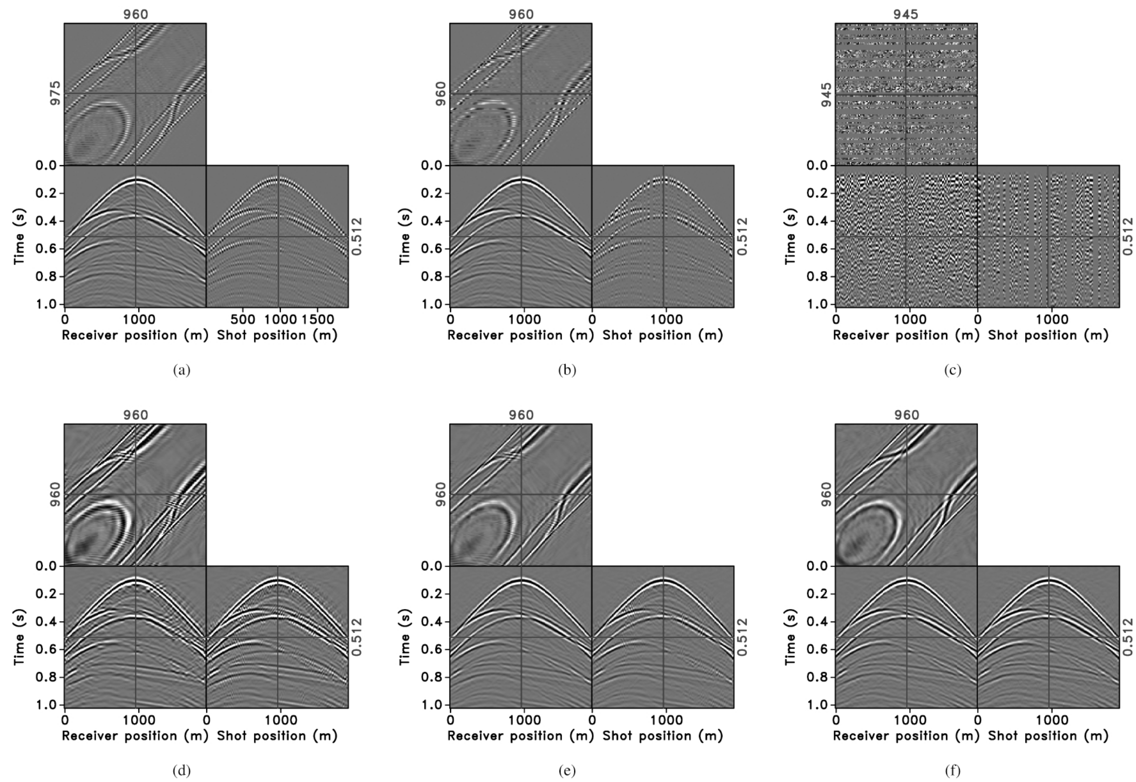

Figure 1: Comparison of interpolation and demultiplexing. Synthetic dataset created for 128 collocated shots, dt =0.004s, $ds = dg$ =15m **(a)-(c)** Simulated data for regularly missing shots, jittered missing shots ((Hennenfent and Herrmann, 2008)), and undersampled blended shots. **(d)** interpolation from regularly missing shots, SNR 8.9dB **(e)** interpolation from jittered missing shots, SNR 10.8dB, **(f)** demultiplex shots, SNR 16.1dB

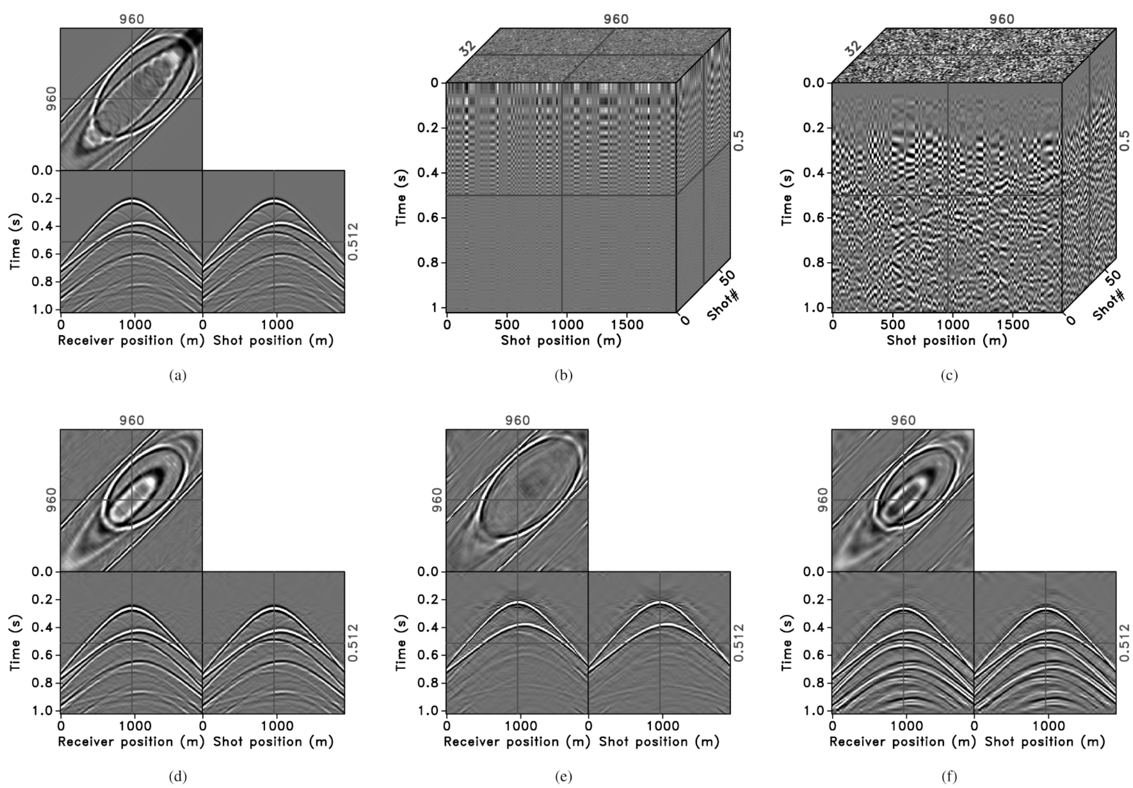

Figure 2: Simulation experiment on marine dataset. For more information on this data, including the velocity model and the true primaries see van Groenestijn and Verschuur (2009) **(a)** crossing-planes view of the model data wavefield **(b)-(c)** blended shot volume and the resulting data volume, SNR 16.3 **(d)** recovered separated shots from the blended data volume **(e)** recovered primary impulse response from combined separation and EPSI approach **(e)** EPSI on the recovered separated shots in (d).

REFERENCES

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

Berg, E. v. and M. P. Friedlander, 2008, Probing the Pareto frontier for basis pursuit solutions: University of British Columbia Technical Report TR-2008-01.

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

Candès, E., J. Romberg, and T. Tao, 2006, Stable signal recovery from incomplete and inaccurate measurements: Communications on Pure and Applied Mathematics, **59**, 1207–1223.

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

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

Hennenfent, G. and F. J. Herrmann, 2008, Simply denoise: wavefield reconstruction via jittered undersampling: Geophysics, **73**, V19–V28.

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

Herrmann, F. J., Y. A. Erlangga, and T. T. Lin, 2009, Compressive simultaneous full-waveform simulation: : Geophysics, accepted for publication.

Krohn, C. and R. Neelamani, 2008, Simultaneous sourcing without compromise: Rome 2008, 70th EAGE Conference & Exhibition, B008.

Neelamani, N., C. Krohn, J. Krebs, M. Deffenbaugh, and J. Romberg, 2008, Efficient seismic forward modeling using simultaneous random sources and sparsity: SEG International Exposition and 78th Annual Meeting, 2107–2110.

Romberg, J., 2008, Personal Communication.

van Groenestijn, G. J. A., and Verschuur, D. J., 2009, Primary estimation by sparse inversion and its application to near offset reconstruction: Geophysics, accepted for publication.