

# Compressive simultaneous full-waveform simulation

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

## SUMMARY

The fact that the computational complexity of wavefield simulation is proportional to the size of the discretized model and acquisition geometry, and not to the complexity of the simulated wavefield, is a major impediment within seismic imaging. By turning simulation into a compressive sensing problem—where simulated data is recovered from a relatively small number of independent simultaneous sources—we remove this impediment by showing that compressively sampling a simulation is equivalent to compressively sampling the sources, followed by solving a reduced system. As in compressive sensing, this allows for a reduction in sampling rate and hence in simulation costs. We demonstrate this principle for the time-harmonic Helmholtz solver. The solution is computed by inverting the reduced system, followed by a recovery of the full wavefield with a sparsity promoting program. Depending on the wavefield’s sparsity, this approach can lead to significant cost reductions, in particular when combined with the implicit preconditioned Helmholtz solver, which is known to converge even for decreasing mesh sizes and increasing angular frequencies. These properties make our scheme a viable alternative to explicit time-domain finite-differences.

## INTRODUCTION

Currently, simulation costs are largely dictated by the mesh size of the computational domain, the size of the source-receiver acquisition grid, and the maximum-desired frequency of the source function. As opposed to solutions based on certain approximations, e.g. asymptotic (high-frequency) solutions that typically involve diagonalization of solution operators (see e.g. ten Kroode et al., 1998), efficient computation of high-fidelity large-scale full-waveform simulations remains an elusive area of research. To overcome this impediment, we argue that future improvements in wave-equation imaging and inversion will depend on a problem formulation with a computational complexity that is no longer strictly determined by the size of the discretization but by the transform-domain compressibility of its solution. In this new paradigm, we bring computational costs in par with our ability to compress seismic data and images.

The above premise is related to recent developments in theoretical signal processing—known as 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. In this approach, sub-sampling interferences are removed by exploiting transform-domain sparsity, properties of certain sub-sampling schemes, and the existence of sparsity promoting solvers. Following earlier work by Lin and Herrmann (2007) (with a formal proof recently established by Demanet and Peyré,

2008), we adapt the ideas from CS towards the problem of seismic waveform simulation. Instead of compressively sampling along the source/receiver coordinates in the modal domain (spanned by the eigenfunctions of the Helmholtz operator), we propose to compressively sample the source wavefields. Both approaches can be seen as instances of simultaneous-source acquisition, an observation also made by Neelamani et al. (2008). Our main contribution here lies in the design of a simultaneous acquisition methodology, based on fast implementations of CS-sampling matrices (Romberg, 2008), for numerical waveform simulation (albeit our ideas are extendible to field acquisition, see e.g. Krohn and Neelamani, 2008; Neelamani et al., 2008).

By designing the acquisition according to CS principles, the interference phenomenon experienced in this type of acquisition can be mitigated. CS is important because it provides a link between subsampling strategies and the quality of recovery, which is largely lacking in simultaneous acquisition (Beasley, 2008), and in cost-reduction during waveform simulation and imaging with subsets of frequencies (Sirgue and Pratt, 2004; Mulder and Plessix, 2004).

## MOTIVATION

Even though numerical solutions to wave equations vary in complexity, they generally compress in transformed domains such as curvelets (Candès et al., 2006). In principle, this finding allows for recovery through sparsity promotion, following the sampling strategies prescribed by CS (see e.g. Herrmann and Hennenfent, 2008; Hennenfent and Herrmann, 2008, where this approach was followed to regularize data with missing traces through curvelet-domain sparsity promotion). To illustrate how these guidelines translate to subsampling (= compressing) solutions of the wave equation, we first briefly outline the principles of CS and sparse recovery, followed by a brief discussion on the Helmholtz wave-equation solver.

**Compressive sampling:** Recovering a sparse (‘spiky’) signal from sub samplings is the modern equivalent of finding ‘needles’ in a haystack. For instance, consider the problem of locating  $k$  arbitrary non-zero entries from a sparse spike train of length  $N \gg k$ . According to CS theory,  $m$  incoherent measurements of a data vector  $\mathbf{d} \in \mathbb{R}^N$  suffice to recover these non-zero entries. Here, the CS-measurement vector is given by  $\mathbf{y} = \mathbf{R}\mathbf{M}\mathbf{d} \in \mathbb{C}^m$  with  $\mathbf{M}$  the  $(N \times N)$  measurement matrix (e.g. the Fourier basis) and  $\mathbf{R}$  the  $(m \times N)$  restriction matrix that randomly selects  $m$  rows with  $m \sim k$  ( $\sim$  means proportional to within  $\log N$  factors). Two basis are *incoherent*, when the inner products of the columns are small. For instance, Dirac and Fourier are incoherent and so are any orthonormal basis and a orthonormalized Gaussian random matrix. 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 non-zeros,  $k$ ) as opposed to the signal length  $N$ . We use this important finding in our formulation of the compressive or blended simulation problem.

**Feasible Helmholtz solver for full-waveform simulation:** Since

their inception, iterative implicit matrix-free solutions to the time-harmonic Helmholtz equation have been plagued by lack of numerical convergence for decreasing mesh sizes and increasing angular frequencies (Riyanti et al., 2006). By including deflation, a way to handle small eigenvalues that lead to slow convergence, Erlangga and Nabben (2007); Erlangga and Herrmann (2008) successfully removed this impediment, bringing 2- and 3-D solvers for the time-harmonic Helmholtz into reach. For a given shot (right-hand side  $\mathbf{b}$ ) and angular frequency  $\omega$  ( $:= 2\pi f$ , with  $f$  the temporal frequency in Hz), the frequency-domain wavefield  $\mathbf{u}$  is computed with a Krylov method that involves the following system of equations

$$\mathcal{H}[\omega]\mathcal{M}^{-1}\mathcal{Q}\hat{\mathbf{u}} = \mathbf{b}, \quad \mathbf{u} = \mathcal{M}^{-1}\mathcal{Q}\hat{\mathbf{u}}, \quad (1)$$

where  $\mathcal{H}[\omega]$ ,  $\mathcal{M}$ , and  $\mathcal{Q}$  represent the discretized monochromatic Helmholtz equation, the preconditioner, and projection matrices, respectively. As shown by Erlangga et al. (2004, 2006), convergence is guaranteed by defining the preconditioning matrix  $\mathcal{M}$  in terms of the discretized shifted or damped Helmholtz operator  $\mathcal{M} := -\nabla \cdot \nabla - \frac{\omega^2}{c(x)^2}(1 - \beta\hat{i})$ ,  $\hat{i} = \sqrt{-1}$ , with  $\beta > 0$ . With this preconditioning, the eigenvalues of  $\mathcal{H}\mathcal{M}^{-1}$  are clustered into a circle in the complex plane. By the action of the projector matrix  $\mathcal{Q}$ , these eigenvalues move towards unity on the real axis as shown in Figure 1. These two operations lead to an improved condition number, which explains the superior performance of this solver.

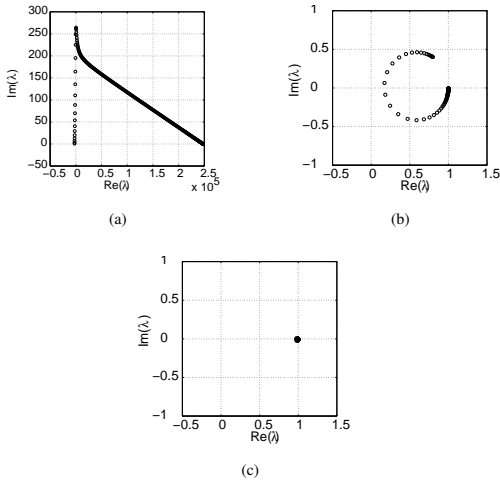


Figure 1: Eigenvalues ( $\lambda$ ) of the 1-D Helmholtz operator before and after preconditioning and spectral shaping. (a) Original eigenvalue spectrum of  $\mathcal{H}$ . (b) Spectrum after preconditioning with  $\mathcal{M}^{-1}$ . (c) Spectrum after deflation with  $\mathcal{Q}$ .

**Source-receiver CS-sampling equivalence:** Aside from the required number of frequencies, the computational complexity of full-wavefield simulation is determined by the number of shots—i.e., the number of right-hand sides. In the current simulation paradigm, these shots determine the number of single-source simulations. As prescribed by CS, these costs can be reduced by designing a survey that consists of a relatively *small number* of simultaneous ‘shots’ with simultaneous sources that contain *subsets* of angular frequencies. Mathematically, we can accomplish this by applying the CS-sampling matrix,  $\mathbf{RM}$ , to the individual sources collected in the vector  $\mathbf{s}$ . Now, if we can show that the output of this procedure to generate simultaneous sources,  $\underline{\mathbf{s}} = \mathbf{RM}\mathbf{s}$ , yields the same results as modeling the *complete* system followed by compressive sampling then we are in the position to speed up our computations. This speed up is the result of a decreased number of shots and angular frequencies (quantities related to the compressed system are underlined) that are present in the simultaneous source vector. For this to work, the solution  $\underline{\mathbf{y}}$  must be equivalent to the solution  $\mathbf{y}$  obtained by compressively sampling the full solution. More specifically, we need to demonstrate that the solutions for the full and compressed systems are equivalent—i.e.,  $\mathbf{y} = \underline{\mathbf{y}}$  in

$$\mathbb{P}_1 : \begin{cases} \mathbf{B} = \mathbf{D}^* \mathbf{s} \\ \mathbf{H}\mathbf{U} = \mathbf{B} \\ \mathbf{y} = \mathbf{R}\mathbf{M}\mathbf{D}\mathbf{U} := \mathbf{R}\mathbf{M}\mathbf{d} \end{cases} \iff \mathbb{P}_2 : \begin{cases} \underline{\mathbf{B}} = \underline{\mathbf{D}}^* \underline{\mathbf{s}} = \underline{\mathbf{D}}^* \mathbf{R}\mathbf{M}\mathbf{s} \\ \underline{\mathbf{H}}\underline{\mathbf{U}} = \underline{\mathbf{B}} \\ \underline{\mathbf{y}} = \underline{\mathbf{D}}\mathbf{U}. \end{cases} \quad (2)$$

Here,  $\mathbf{H} = \text{diag}(\mathcal{H}[\omega_i])$  is the block-diagonal discretized Helmholtz equation for each  $\omega_i := 2\pi i \cdot \Delta f$ ,  $i = 1 \dots n_f$ , with  $n_f$  the number of frequencies and  $\Delta f$  its sample interval. The adjoint (denoted by  $*$ ) of the detection matrix  $\mathbf{D}$  injects the individual sources into the multiple right-hand sides,  $\mathbf{B} = [\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_{n_s}]$ , with  $n_s$  the number of shots. This detection matrix extracts data at the receiver positions. Its adjoint inserts data at the co-located source positions. Each column of  $\mathbf{U}$  contains the wavefields for all frequencies induced by the shots located in the columns of  $\mathbf{B}$ . Consequently, the full simulation requires the inversion of the block-diagonal system (for all shots), followed by a detection—i.e., we have  $\mathbf{d} = \mathbf{D}\mathbf{H}^{-1}\mathbf{B}$ , with  $\mathbf{H}^{-1} = \text{diag}(\mathcal{H}^{-1}[\omega_i])$ ,  $i = 1 \dots n_s$ . After CS sampling, this volume is reduced to  $\mathbf{y} = \mathbf{R}\mathbf{M}\mathbf{d}$  by applying the flat rectangular CS-sampling matrix  $\mathbf{RM}$  (defined explicitly in the next section) to the full simulation. Applying  $\mathbf{RM}$  directly to the sources  $\mathbf{s}$  in  $\mathbb{P}_2$  leads to a compressed system  $\underline{\mathbf{H}}$ , which after inversion gives  $\underline{\mathbf{y}}$ . To illustrate why  $\underline{\mathbf{y}}$  is equivalent to  $\mathbf{y}$ , consider a compressive sampling of the solution over frequency by the subsampling matrix  $\mathbf{R}^\Omega$  (for clarity, we removed the orthonormal measurement matrix). This restriction matrix removes arbitrary rows from the right-hand side. By virtue of the block-diagonal structure of our system, we have  $\mathbf{R}^\Omega \mathbf{H}^{-1} = \underline{\mathbf{H}}^{-1} \mathbf{R}^\Omega$  with  $\underline{\mathbf{H}}^{-1} = \text{diag}(\mathcal{H}^{-1}[\omega_i])$ ,  $i \in \{1 \dots n_f\}$ , yielding  $\mathbf{R}^\Omega \mathbf{U} = \underline{\mathbf{H}}^{-1} \underline{\mathbf{B}} = \underline{\mathbf{U}}$ , where  $\underline{\mathbf{B}} := \mathbf{R}^\Omega \mathbf{B}$ . This means that frequency subsampling the right-hand side, followed by solving the system for the corresponding frequencies, is the same as solving the full system, followed by frequency subsampling. A similar argument holds when subsampling the shots (removing

arbitrary columns of  $\mathbf{B}$ ). Now, we have the reduced system  $\mathbf{R}^\Omega \mathbf{U} (\mathbf{R}^\Sigma)^* = \mathbf{H}^{-1} \mathbf{B} = \mathbf{U}$ , with  $\mathbf{B} := \mathbf{R}^\Omega \mathbf{B} (\mathbf{R}^\Sigma)^*$ . Using Kronecker products, these relations can be written succinctly as  $(\mathbf{R}^\Sigma \otimes \mathbf{R}^\Omega) \text{vec}(\mathbf{U}) = \text{vec}(\mathbf{U})$  and  $(\mathbf{R}^\Sigma \otimes \mathbf{R}^\Omega) \text{vec}(\mathbf{B}) = \text{vec}(\mathbf{B})$  with  $\text{vec}(\cdot)$  being a linear operator that maps a matrix into a lexicographically-sorted array.

The inversion of  $\mathbf{H}\mathbf{U} = \mathbf{B}$  is easier because it involves only a subset of angular frequencies and simultaneous shots—i.e.,  $\{\mathbf{U}, \mathbf{B}\}$  contain only  $n'_s$  columns with  $n'_f$  frequency components each. Finally, the matrix  $\mathbf{D}$  extracts the compressed data from the solution.

## CS PROBLEM FORMULATION

With the above identification, we can now recover the full data from compressive simulations. Successful solutions of this recovery problem hinge on a delicate interplay between three key components that need to be implemented at minimal costs to avoid offsetting our computational gain through system-size reduction.

**The compressive-sampling matrix:** The success of compressive simulation depends on a subsampling of the physically distinct source and frequency axes where coherent interferences are turned into random noise (Hennenfent and Herrmann, 2008). Since speed is of the essence for the recovery, 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

$$\mathbf{RM} = \underbrace{\begin{bmatrix} \mathbf{R}_1^\Sigma \otimes \mathbf{I} \otimes \mathbf{R}_1^\Omega \\ \vdots \\ \mathbf{R}_{n'_s}^\Sigma \otimes \mathbf{I} \otimes \mathbf{R}_{n'_s}^\Omega \end{bmatrix}}_{\text{sub sampler}} \underbrace{\left( \mathbf{F}_2^* \text{diag}(e^{i\boldsymbol{\theta}}) \otimes \mathbf{I} \right) \mathbf{F}_3}_{\text{random phase encoder}}, \quad (3)$$

where  $\mathbf{F}_{2,3}$  are the 2,3-D Fourier transforms, and where  $\boldsymbol{\theta} = \text{Uniform}([0, 2\pi])$  is a random phase rotation. Notice that the  $\mathbf{F}_2$  and phase rotations act along the source/receiver coordinates.

Application of this CS-sampling matrix,  $\mathbf{RM}$ , to the original source wavefields in  $\mathbf{s}$  turns these single shots into a subset ( $n'_s \ll n_s$ ) 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  $n'_f \ll n_f$  frequencies non-zero (see Figure 2(a)). Because seismic data is bandwidth limited, we sample with a probability that is weighted by the power spectrum of the source wavelet. The advantage of this implementation is that it is matrix-free, fast, and it turns interferences into harmless noise (see Figure 2(b)).

**The sparsifying 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.

**Recovery by sparsity promotion:** We reconstruct the seismic wavefield by solving the following nonlinear optimization problem

$$\tilde{\mathbf{x}} = \underset{\mathbf{x}}{\text{arg min}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y}, \quad (4)$$

with  $\tilde{\mathbf{d}} = \mathbf{S}^* \tilde{\mathbf{x}}$  the reconstruction,  $\mathbf{A} := \mathbf{RMS}^*$  the CS matrix, and  $\mathbf{y} (= \mathbf{y})$  the compressively simulated data (cf. Equation 2-right). Equation 4 is solved by SPGL<sub>1</sub> (van den Berg and Friedlander, 2008), a projected-gradient algorithm with root finding.

## COMPUTATIONAL COMPLEXITY ANALYSIS

According to Riyanti et al. (2006), the cost of the iterative Helmholtz solver equals  $n_f n_s n_{it} \mathcal{O}(n^d)$ , typically with  $n_{it} = \mathcal{O}(n)$  the number of iterations. For  $d = 2$  and assuming  $n_s = n_f = \mathcal{O}(n)$ , this cost becomes  $\mathcal{O}(n^5)$ . Under the same assumption, the cost of a time-domain solver is  $\mathcal{O}(n^4)$ . The iterative Helmholtz solver can only become competitive if  $n_{it} = \mathcal{O}(1)$ , yielding an  $\mathcal{O}(n^4)$  computational complexity. Erlangga and Nabben (2007); Erlangga and Herrmann (2008) achieve this by the method explained earlier. Despite this improvement, this figure is still overly pessimistic for simulations that permit sparse representations. As long as the simulation cost exceeds the  $\ell_1$ -recovery cost (cf. Equation 4), CS will improve on this result. This reduction depends on the cost of  $\mathbf{A}$ , which is dominated by the CS-matrix. For naive choices, such as Gaussian projections, these sampling matrices cost  $\mathcal{O}(n^3)$  for each frequency, which offers no gain. However, with our choice of fast  $\mathcal{O}(n \log n)$  projections with random convolutions (Romberg, 2008), we are able to reduce this cost to  $\mathcal{O}(n^2 \log n)$ . Remark that these costs are of the same order as those of calculating the sparsifying transforms. Now, the leading order cost of the  $\ell_1$  recovery is reduced to  $\mathcal{O}(n^3 \log n)$ , which is significantly less than the cost of solving the full Helmholtz system, especially for large problems ( $n \rightarrow \infty$ ) and for extensions to  $d = 3$ .

## COMPRESSIVE SIMULATION EXPERIMENT

To illustrate CS-recovery quality, we conduct a series of experiments for two velocity models, namely the complex model used in Herrmann et al. (2007), and a simple single-layer model. These models generate seismic lines that differ in complexity. During these experiments, we vary the subsampling ratio and

the frequency-to-shot subsampling ratio. All simulations are carried out with a fully parallel Helmholtz solver with  $\beta = 0.5$ , for a spread with 128 col-located shots and receivers sampled at a 30 m interval. The time sample interval is 0.004s and the source function is a Ricker wavelet with a central frequency of 10 Hz. By solving Equation 4, we recover the full simulation for the two datasets. Comparison between the full and compressive simulations in Figure 3 shows remarkable high-fidelity results even for increasing subsampling ratios. As expected, the SNR for the simple model is better because of the reduced complexity, whereas the numbers in Table 1 for the complex model confirm increasing recovery errors for increasing subsampling ratios. Moreover, the bandwidth limitation of seismic data explains improved recovery with decreasing frequency-to-shot ratio for a fixed subsampling ratio. Because the speedup of the solution is roughly proportional to the subsampling ratio, we can conclude that speedups of four to six times are possible at the expense of a minor drop in SNR.

Subsample ratio	0.25	0.15	0.07
$n'_f/n'_s$	recovery error (dB)		
2	14.3	12.1	8.6
1	18.2	14.5	10.2
0.5	22.2	16.5	10.7
Speed up (%)	400	670	1420

Table 1: Signal-to-noise ratios based on the complex model,  $\text{SNR} = -20 \log_{10}(\frac{\|\mathbf{d}-\hat{\mathbf{d}}\|_2}{\|\mathbf{d}\|_2})$  for reconstructions with the curvelet-wavelet sparsity transform with different subsample and frequency-to-shot ratios.

## CONCLUSIONS

We have identified simultaneous acquisition as compressive sensing (CS), which allowed us to derive a rigorous and cost-effective simulation scheme based on its principles. According to these principles, wavefields can be reconstructed from subsamplings commensurate with their complexity. We arrive at this result by source-receiver CS-sampling equivalence, which states that CS sampling seismic simulations is the same as CS sampling the source wavefield, followed by simulation with a reduced system. CS predicts improved recovery for compressible signals for increasing number of samples. We verified this behavior experimentally and this, in conjunction with the intrinsic linearity of the CS sampling, opens a number of enticing new perspectives. First, CS applied to simulations (and possibly during acquisition) decouples simulation- and acquisition-related costs from the model size. Instead, these costs depend on sparsity. Second, CS predicts improved recovery for increasing sample sizes, which opens the possibility to improve recovery by adding samples. Third, because of this linearity, we envision a seamless incorporation of this paradigm into seismic exploration.

## ACKNOWLEDGMENTS

The authors would like to thank Scott Morton for bringing this topic to our attention. We would like to thank the authors of SPARCO, SPGL<sub>1</sub>, and the Rice Wavelet Toolbox. This paper was prepared with Madagascar and was in part financially supported by the NSERC Discovery Grant (22R81254) of F.J.H. and CRD Grant DNOISE (334810-05).

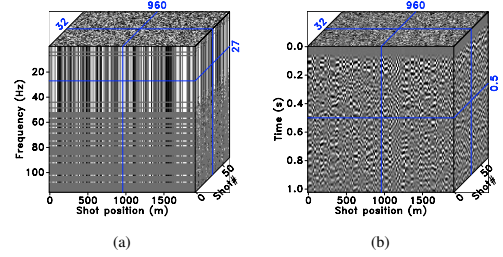


Figure 2: Compressive sampling with simultaneous sources. **(a)** Amplitude spectrum for the source signatures emitted by each source as part of the simultaneous-source experiments. These signatures appear noisy in the shot-receiver coordinates because of the phase encoding (cf. Equation 3). **(b)** CS-data after applying the inverse Fourier transform.

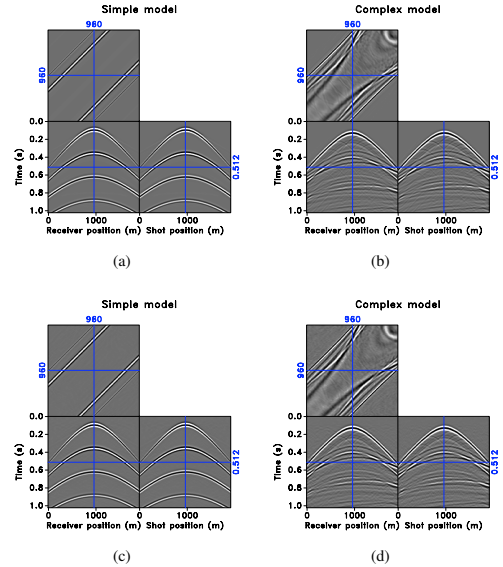


Figure 3: Comparison between conventional and compressive simulations in for simple and complex velocity models. **(a)** Crossing-planes view of the seismic line for the simple model. **(b)** The same for the complex model. **(c)** Recovered simulation (with a SNR of 28.1 dB) for the simple model from 25% of the samples with the  $\ell_1$ -solver running to convergence. **(d)** The same but for the complex model now with a SNR of 18.2dB.

## REFERENCES

- Beasley, C. J., 2008, A new look at marine simultaneous sources: *The Leading Edge*, **27**, 914–917.
- 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.
- Demanet, L. and G. Peyré, 2008, Compressive wavefield computation: submitted.
- Donoho, D. L., 2006, Compressed sensing: *IEEE Transactions on Information Theory*, **52**, 1289–1306.
- Erlangga, Y. and F. J. Herrmann, 2008, An iterative multilevel method for computing wavefields in frequency-domain seismic inversion: Presented at the SEG Technical Program Expanded Abstracts, SEG.
- Erlangga, Y. A. and R. Nabben, 2007, On multilevel projection Krylov method for the preconditioned Helmholtz system: submitted.
- Erlangga, Y. A., C. Vuik, and C. W. Oosterlee, 2004, On a class of preconditioners for solving the Helmholtz equation: *Applied Numerical Mathematics*, **50**, 409–425.
- , 2006, Comparison of multigrid and incomplete LU shifted-Laplace preconditioners for the inhomogeneous Helmholtz equation: *Applied Numerical Mathematics*, **56**, 648–666.
- 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. and G. Hennenfent, 2008, Non-parametric seismic data recovery with curvelet frames: *Geophysical Journal International*, **173**, 233–248.
- Krohn, C. and R. Neelamani, 2008, Simultaneous sourcing without compromise: Rome 2008, 70th EAGE Conference & Exhibition, B008.
- Lin, T. T. Y. and F. J. Herrmann, 2007, Compressed wavefield extrapolation: *Geophysics*, **72**, SM77–SM93.
- Mulder, W. and R. Plessix, 2004, How to choose a subset of frequencies in frequency-domain finite-difference migration: *Geophysical Journal International*, **158**, 801–812.
- Neelamani, N., C. Krohn, J. Krebs, M. Deffenbaugh, and J. Romberg, 2008, Efficient seismic forward modeling using simultaneous random sources and sparsity: Presented at the SEG International Exposition and 78th Annual Meeting.
- Riyanti, C. D., Y. A. Erlangga, R.-E. Plessix, W. A. Mulder, C. Vuik, and C. Oosterlee, 2006, A new iterative solver for the time-harmonic wave equation: *Geophysics*, **71**, E57–E63.
- Romberg, J., 2008, Compressive sensing by random convolution: submitted.
- Sirgue, L. and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: *Geophysics*, **69**, 231–248.
- ten Kroode, A., D.-J. Smit, and A. Verdel, 1998, A microlocal analysis of migration: *Wave Motion*, **28**, 149–172.
- van den Berg, E. and M. P. Friedlander, 2008, Probing the Pareto frontier for basis pursuit solutions: to appear in *SIAM Journal on Scientific Computing*.