Introduction

With the recent resurgence of full-waveform inversion, the computational cost of solving forward modeling problems has become—aside from issues with non-uniqueness—one of the major impediments withstanding successful application of this technology to industry-size data volumes. To overcome this impediment, we argue that further improvements in this area will depend on a problem formulation with a computational complexity that is no longer strictly determined by the *size* of the discretization but by transform-domain *sparsity* of its solution. In this new paradigm, we bring computational costs in par with our ability to compress seismic data and images. This premise is related to two recent developments. First, there is the new 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. Second, there is in the seismic community the recent resurgence of simultaneous-source acquisition (Beasley, 2008; Krohn and Neelamani, 2008; Herrmann et al., 2009; Berkhout, 2008; Neelamani et al., 2008), and continuing efforts to reduce the cost of seismic modeling, imaging, and inversion through phase encoding of simultaneous sources (Morton and Ober, 1998; Romero et al., 2000; Krohn and Neelamani, 2008; Herrmann et al., 2009), removal of subsets of angular frequencies (Sirgue and Pratt, 2004; Mulder and Plessix, 2004; Lin et al., 2008) or plane waves (Vigh and Starr, 2008). By using CS principles, we remove sub-sampling interferences asocciated with these approaches through a combination of exploiting transform-domain sparsity, properties of certain sub-sampling schemes, and the existence of sparsity promoting solvers.

Theory

Full-waveform inversion entails solving PDE-constrained optimization problems of the following form

$$\min_{\mathbf{U}\in\boldsymbol{\mathcal{U}},\,\mathbf{m}\in\boldsymbol{\mathcal{M}}} \frac{1}{2}\|\mathbf{RM}\big(\mathbf{d}-\mathbf{DU}\big)\|_2^2 \quad \text{subject to} \quad \mathbf{H}[\mathbf{m}]\mathbf{U}=\mathbf{B}, \tag{1}$$

where $\mathbf{d}$ is the observed data volume, $\mathbf{U}$ the solution of the multi-shot (in its columns)-frequency Helmholtz equation, $\mathbf{D}$ the detection operator that extracts the simulated data from time-harmonic solutions at the receiver locations, $\mathbf{H}$ is a matrix with the discretized multi-frequency Helmholtz equation, and $\mathbf{B}$ a matrix with the frequency-transformed source distributions in its columns. In the above optimization problem (from which, after casting Eq. 1 in its unconstrained form, most quasi-Newton type full-waveform inversion schemes derive), solutions for the unknown velocity model, $\mathbf{m}$, and the solution to the wave equation, $\mathbf{U}$, are pursued that minimize the energy between the observations and simulations after detection. Solutions to Eq.1 typically require multiple solves of the implicit Helmholtz equation, which even after our preconditioning (yielding a complexity for this solver of $\mathcal{O}(n^4)$ in 2-D (Erlangga and Nabben, 2007; Erlangga and Herrmann, 2008)) proves prohibitive. We address this issue by using CS (Herrmann et al., 2009) to reduce the size of the seismic data volume through $\mathbf{y}=\mathbf{RMd}$ with

$$\mathbf{RM}=\overbrace{\begin{bmatrix}\mathbf{R}_1^{\Sigma}\otimes\mathbf{I}\otimes\mathbf{R}_1^{\Omega}\\ \vdots \\ \mathbf{R}_{n_{s'}}^{\Sigma}\otimes\mathbf{I}\otimes\mathbf{R}_{n_{s'}}^{\Omega}\end{bmatrix}}^{\text{sub sampler}}\overbrace{\left(\mathbf{F}_2^{*}\left(e^{\hat{i}\boldsymbol{\theta}}\right)\otimes\mathbf{I}\right)\mathbf{F}_3}^{\text{random phase encoder}}, \tag{2}$$

with $\mathbf{F}_{2,3}$ the 2,3-D Fourier transforms, and $\boldsymbol{\theta}=\mathrm{Uniform}([0,2\pi])$ a random phase rotation. The matrices $\mathbf{R}^{\Omega}$ and $\mathbf{R}^{\Sigma}$ represent CS-subsampling operators (see Figure 1) acting along the rows (frequency coordinate) and columns (shot coordinate), respectively. Application of this CS-sampling matrix, $\mathbf{RM}$, to the original source wavefields in $\mathbf{s}$ turns these single shots into a subset ($n_s'\ll n_s$ 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 (see Figure 1). As was shown by the authors (Herrmann et al., 2009), this sub-sampling operator commutes

with the Helmholtz system and this allows us recast Eq. 1 into the following reduced form

$$\min_{\underline{\mathbf{U}}\in\underline{\boldsymbol{\mathcal{U}}},\,\mathbf{m}\in\boldsymbol{\mathcal{M}}} \frac{1}{2}\|\mathbf{y}-\underline{\mathbf{D}}\underline{\mathbf{U}}\|_2^2 \quad \text{subject to} \quad \underline{\mathbf{H}}[\mathbf{m}]\underline{\mathbf{U}}=\underline{\mathbf{B}}, \tag{3}$$

where the underbarred quantities relate to the reduced Helmholtz system.

Recovery by sparsity promotion

Blind application of the above sub-sampling formalism has led —i.e., without changing the optimization problems in Eqs. 1, 3—to successful reductions in computations but generally at the expense of a (minor) decrease in signal-to-noise and estimated amplitudes. The explanation for this success lies in the fact that subsamplings are typically carried out in such a way that the artifacts become incoherent (read Gaussian noise like) and these are subsequently "stacked out" by virtue of data redundancy (e.g. in 2-D seismic data volumes are 3-D while the image is only 2-D). By keeping the sub-sampling ratio under control, these artifacts can be managed. However, according to the principles of CS this procedure can be pushed much further and we therefore propose to supplement these sub-sampling schemes by a nonlinear recovery based on transform-domain sparsity promotion (see e.g. Hennenfent and Herrmann (2008), where we used this principle to recover seismic data volumes from missing traces using curvelet-domain sparsity). Following earlier work (Neelamani et al., 2008; Lin et al., 2008; Herrmann et al., 2009), we follow a similar principle to recover source-separated data volumes from solutions of the reduced Helmholtz system for a small subset of simultaneous sources with subsets of angular frequencies removed. Because seismic data is typically larger and more complex than realistic velocity distributions (see e.g. Herrmann et al. (2008) where a real-data seismic image was recovered from a small fraction of its curvelet coefficients), successful recovery of complete data volumes will be a strong indication that CS principles will mitigate the adverse affects of high degrees of subsampling. As shown by Neelamani et al. (2008); Herrmann et al. (2009), recovery from simultaneous simulations depends on a sparsifying transform that compresses seismic data. 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. We reconstruct the source-separated wavefield at the surface by solving the following nonlinear optimization problem

$$\widetilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{x}=\mathbf{y}, \tag{4}$$

with $\widetilde{\mathbf{d}} = \mathbf{S}^*\widetilde{\mathbf{x}}$ recovered data, $\mathbf{A} := \mathbf{R}\mathbf{M}\mathbf{S}^*$ the CS matrix, and $\mathbf{y}$ the compressive simulation. This is solved with SPGℓ_1, a projected-gradient algorithm with root finding (Berg and Friedlander, 2008).

Example

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 for a spread with 128 col-located shots and receivers sampled at a 30 m interval. The time sample interval is $0.004\,\mathrm{s}$ 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 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, SNR $= -20\log_{10}(\frac{\|\mathbf{d}-\widetilde{\mathbf{d}}\|_2}{\|\mathbf{d}\|_2})$ for reconstructions with the curvelet-wavelet sparsity transform for different subsample and frequency-to-shot ratios.

Discussion and conclusions

Compressive sampling—where data with structure (read transform-domain sparsity) can be reconstructed from degrees of subsamping commensurate their sparsity—is considered a paradigm shift. We show that this new paradigm can also be applied to reduce the computational complexity of solving PDEs or better of PDE-constrained optimization problems. When we have control over the sources—such as on land—CS extends to acquisition itself. This is exciting because CS decouples simulation- and acquisition-related costs from the model size. Instead, these costs depend on sparsity. Finally, by using the sparsity argument further improvements can be expeced when recovering the model itself rather the data.

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: Technical Report 2, Department of Computer Science, University of British Columbia.

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.

Erlangga, Y. A. and F. J. Herrmann, 2008, An iterative multilevel method for computing wavefields in frequency-domain seismic inversion: SEG Technical Program Expanded Abstracts, 1957–1960, SEG.

Erlangga, Y. A. and R. Nabben, 2007, On multilevel projection Krylov method for the preconditioned Helmholtz system. Submitted for publication.

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. TR-2008-09. to appear in geophysics.

Herrmann, F. J., P. P. Moghaddam, and C. C. Stolk, 2008, Sparsity- and continuity-promoting seismic image recovery with curvelet frames: Applied and Computational Harmonic Analysis, **24**, no. 2, 150–173.

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

Lin, T., E. Lebed, Y. A. Erlangga, and F. J. Herrmann, 2008, Interpolating solutions of the helmholtz equation with compressed sensing: SEG Technical Program Expanded Abstracts, 2122–2126, SEG.

Morton, S. A. and C. C. Ober, 1998, Faster shot-record depth migrations using phase encoding: SEG Technical Program Expanded Abstracts, 1131–1134, SEG.

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: SEG International Exposition and 78th Annual Meeting, 2107–2110.

Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, no. 2, 426–436.

Sirgue, L. and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248.

Vigh, D. and E. W. Starr, 2008, 3d prestack plane-wave, full-waveform inversion: Geophysics, **73**, VE135–VE144.

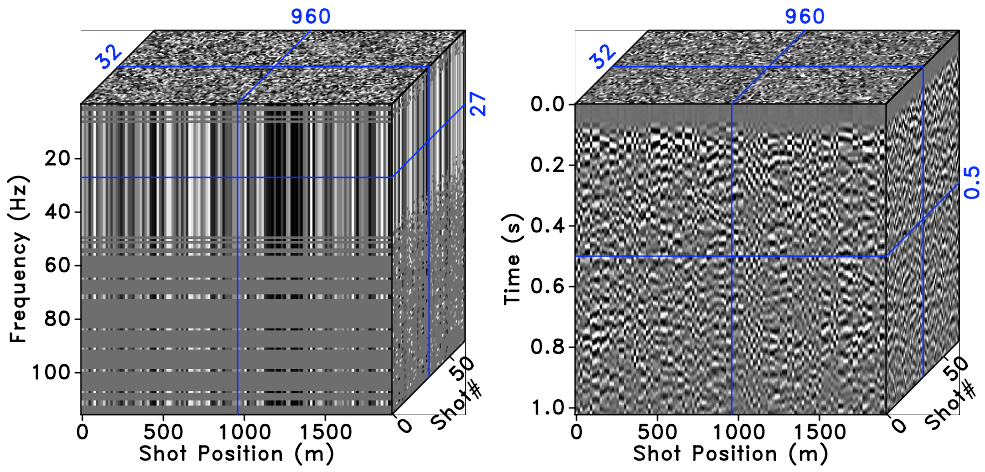

Figure 1: Compressive sampling with simultaneous sources. **(a)** Real part of the compressively-sampled source in the frequency domain. **(a)** Amplitude spectrum for the source signatures emitted by each source as part of the simultaneous-source experiments. These signatures appear noisy in the shot-receiever coordinates because of the phase encoding (cf. Equation 1). Observe that the frequency restrictions are different for each simultaneous source experiment. **(b)** CS-data after applying the inverse Fourier transform.in the time domain. Notice the noisy character of the simultaneous-shot interferences.

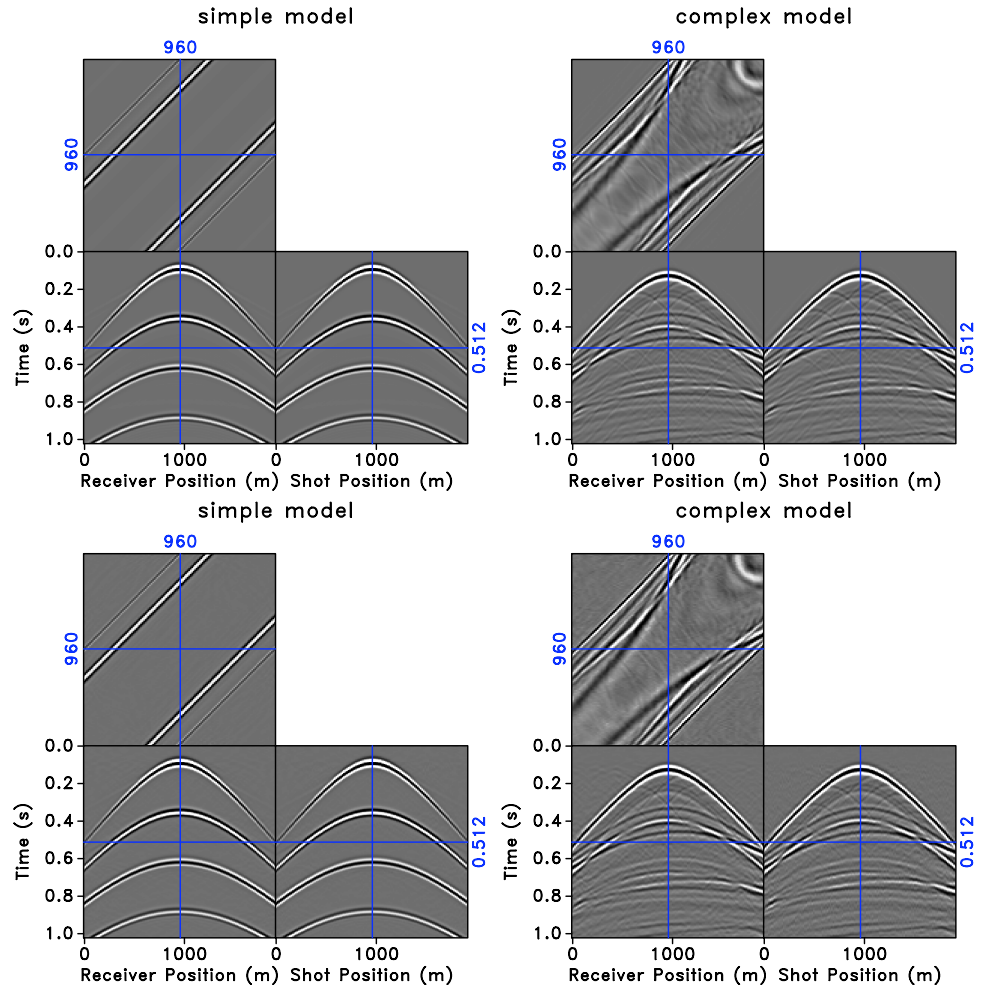

Figure 2: Comparison between conventional and compressive simulations for simple and complex velocity models. **(a)** Seismic line for the simple model. **(b)** The same for the complex model. **(c)**. Recovered simulation (with a SNR of 28.1 dB) for the simple model from 25 % of the samples with the ℓ_1-solver running to convergence. **(d)** The same but for the complex model now with a SNR of 18.2 dB.