Introduction

Coherent subsampling-related interferences are the main enemies of successful seismic acquisition and processing. These interferences can come from any number of places: traditionally they are caused by badly designed regular sub-Nyquist samplings of sources and receivers, and more recently by badly designed simultaneous sources. Conversely, well-designed *randomized* subsamplings—through jittered sampling (Hennenfent and Herrmann, 2008) of the source-receiver positions or through randomly phase encoded source signatures (Berkhout, 2008; Neelamani et al., 2008; Herrmann et al., 2009)—lead to manageable subsampling artifacts that manifest themselves as incoherent noise with a level that depends on the degree of subsampling; the more subsampled you are the higher your expected noise level becomes. Herein lies an unique opportunity: as long as we are able to separate subsampling noise from the desired signal, we are in the position to recover the fully-sampled signal. This is where transform-domain sparsity enters into the equation, because the sparser we can represent our desired signal—i.e., the more of the signal's energy we can store into the fewer largest transform-domain coefficients—the better we can separate this incoherent subsampling noise from the signal. This task of recovering fully-sampled signals from deliberate subsampling is actually far less daunting as it may seem; we all know that seismic data contains structure that can be exploited with certain multiscale and multidirectional transforms such as curvelets (Candès et al., 2006). These transforms—possibly in conjunction with *focusing* procedures that map multiple energy onto primaries (see e.g. Herrmann and Wang, 2008; van Groenestijn and Verschuur, 2009; Lin and Herrmann, 2009b, and our other contribution to the proceedings of this conference) or that collapse primary energy onto reflectors during imaging (Herrmann, 2009, also see our other contribution to the proceedings of this conference)—translate this structure into transform domain *sparsity*.

Mathematically, our recovery from deliberate subsampling with sparsity promotion can be formulated as the inversion of a flat matrix $\mathbf{A}$ via $\min_{\mathbf{x}} \|\mathbf{x}\|_1 := \sum_i |x_i| \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{b}$, with $\mathbf{A} = \mathbf{R}\mathbf{M}\mathbf{S}^*$. This optimization procedure seeks amongst all possible transform-domain vectors $\mathbf{x}$ the one that has the smallest one norm—i.e., it seeks the sparsest vector using a heuristically derived one-norm measure. The true miracle in this nonlinear recovery lies in the fact that this optimization problem is able to recover the original coefficient vector—and hence the original data—with high accuracy from *randomized* subsampled data in $\mathbf{b}$. Here, the estimate for the recovered data is given by $\widetilde{\mathbf{d}} = \mathbf{S}^*\widetilde{\mathbf{x}}$ with $\widetilde{\mathbf{x}}$ the vector that solves this optimization program. The accuracy of the recovery depends on the following design principles: (i) the length of the measurement vector $\mathbf{b}$ = hight of the *randomized* restriction matrix $\mathbf{R}$, which randomly removes rows. The longer $\mathbf{b}$ the better the recovery; (ii) properties of the measurement matrix $\mathbf{M}$ that preferably represents *randomized* physics underlying the measurements. The more *randomized* this matrix—i.e., the more its action on $\mathbf{S}^*$ resembles a matrix with random Gaussian noise—the better the recovery. (iii) the *sparsity* attained by the transformed domain spanned by $\mathbf{S}$. The sparser this domain the better the recovery. These design principles find their origin and theoretical justification in a new field of mathematics known as *compressive sensing* (Donoho, 2006; Candès et al., 2006), where sparse signals are recovered from *randomized* samples using sparsity-promoting programs. In the next two sections, we illustrate the above examples by two concrete examples that underline the importance of following the above design principles.

Example I: recovery of fully sampled data from jittered sampling and simultaneous acquisition

To illustrate the importance of selecting the appropriate *randomized* measurement matrix $\mathbf{M}$, we compare sparsity-promoting recoveries from three incomplete (50 % of shots missing) data sets that are all equal in size but that differ in acquisition strategy. First, we compare the recovery from two sequential source experiments where the shots are either periodically or randomly (jittered) sampled at half the Nyquist rate. In the second, experiment we compare the recovery from randomized sequential with randomized simultaneous acquisition. In the latter case, we study the recovery from 50 % somewhat fictitious simultaneous source experiments where phase-encoded sweeps go off simultaneously at *all* source positions (Lin and Herrmann, 2009a). Results from sparsity-promoting recovery for data collected according to these scenarios are summarized in Fig. 2. Comparison of the recovery from regular, jittered, and simultaneous shots shows a drastic improvement in the recovery quality as we move from regular subsampled, to *randomized* jittered source locations all the way to *randomized* simultaneous sources. These findings clearly underline the importance of *randomization* in the collection of seismic data. This example also nicely illustrates that *randomization* of the source locations by itself is not optimal and that a lot is to be gained by designing *randomized* incoherent simultaneous-source acquisitions such as acquisitions using phase-encoded sweeps.

Example II: estimation of primaries from simultaneous data

The above scheme of recovery from *randomized* data can even be carried a step further by including more information on the physics, i.e., *focusing* in the matrix $\mathbf{M}$. For instance, if we include in this matrix—aside from the randomization of the sources—an operator that generates surface related multiples, our inversion procedure will map surface-related multiples to "primaries" (that include internal multiples). This approach has two advantages. First, "primaries" are sparser then multiples. Second, multiples are mapped to primaries and thereby facilitate the decoding by sparsity promoting. To illustrate how this works we consider the following two scenarios. First, we recover the total data, including the surface-related multiples, from simultaneous acquired data followed by a prediction of the primaries. Second, we estimate the primaries directly from the simultaneously collected data. As we can see from Fig. 1, the recovery according to the second scenario is far superior because we incorporated more physics into the formulation of our problem (see our other contribution to the proceedings of this conference).

Discussion and conclusions

The above examples illustrate that we are at the cusp of very exciting developments where our design principles for acquisition and processing no longer need to be dominated by our fear of creating coherent subsampling related artifacts. Instead, we arrive at a formulation where we have control over these artifacts—by turning harmful coherent interferences into harmless incoherent noise. In this way, we facilitate the removal of subsampling related artifacts by our sparsity-promoting inversion procedure. This opens enticing new perspectives towards a new formulation of seismic data acquisition and processing. To summarize, the success of this new formulation depends on three key design principles, namely (i) *randomize*—break coherent aliases by introducing randomness, e.g. by designing randomly perturbed acquisition grids, or by designing randomized simultaneous sources and blended receivers; (ii) *sparsify*—utilize sparsifying transforms in conjunction with sparsity promoting programs to remove incoherent subsampling artifacts, e.g. by exploiting curvelet-domain sparsity; (ii) *focus*–leverage physical focusing principles that concentrate seismic energy in order to further promote sparsity in the final solution, e.g. by turning multiples into primaries or primaries into images.

We have made the case that information can be obtained from *randomized* subsamplings. This allows us to formulate rigorous and cost-effective acquisition and processing schemes based on the principles of *compressive sensing*. According to these principles, data can be reconstructed from *randomized* subsamplings commensurate with their *complexity*. We verified this behavior experimentally and this, in conjunction with the intrinsic linearity of the *randomized* sampling, opens a number of enticing new perspectives because acquisition and processing costs are decoupled from the acquisition area and grid size. Instead, these costs depend on sparsity. Because of this linearity, we envision a seamless incorporation of this new *paradigm* into seismic exploration.

Acknowledgments

FJH would like to Eric Verschuur and Tim Lin for their input. This work was financed through the SINBAD II project with support from BG, BP, Chevron, Petrobras, and WesternGeco and was partly supported by NSERC Discovery Grant (22R81254) and CRD Grant DNOISE (334810-05).

References

Berkhout, A.J. [2008] Changing the mindset in seismic data acquisition. *The Leading Edge*, **27**(7), 924–938, ISSN 7, doi: 10.1190/1.2954035.

Candès, E.J., Demanet, L., Donoho, D.L. and Ying, L. [2006] Fast discrete curvelet transforms. *Multiscale Modeling and*

Simulation, **5**, 861–899, doi:10.1137/05064182X.

Candès, E., Romberg, J. and Tao, T. [2006] Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, **59**(8), 1207–1223.

Donoho, D.L. [2006] Compressed sensing. *IEEE Transactions on Information Theory*, **52**(4), 1289–1306.

Hennenfent, G. and Herrmann, F.J. [2008] Simply denoise: wavefield reconstruction via jittered undersampling. *Geophysics*, **73**, V19–V28, doi:10.1190/1.2841038.

Herrmann, F.J. [2009] Compressive imaging by wavefield inversion with group sparsity. *SEG Technical Program Expanded Abstracts*, SEG, SEG, vol. 28, 2337–2341.

Herrmann, F.J., Erlangga, Y.A. and Lin, T.T.Y. [2009] Compressive simultaneous full-waveform simulation. *Geophysics*, **74**(4), A35–A40, doi:10.1190/1.3115122.

Herrmann, F.J. and Wang, D. [2008] Seismic wavefield inversion with curvelet-domain sparsity promotion. *SEG Technical Program Expanded Abstracts*, **27**(1), 2497–2501, doi:10.1190/1.3063862.

Lin, T. and Herrmann, F.J. [2009a] Designing simultaneous acquisitions with compressive sensing. EAGE, EAGE.

Lin, T. and Herrmann, F.J. [2009b] Unified compressive sensing framework for simultaneous acquisition with primary estimation. *SEG Technical Program Expanded Abstracts*, SEG, vol. 28, 3113–3117.

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

van Groenestijn, G.J.A. and Verschuur, D.J. [2009] Estimating primaries by sparse inversion and application to near-offset data reconstruction. *Geophysics*, **74**(3), A23–A28, doi:10.1190/1.3111115.

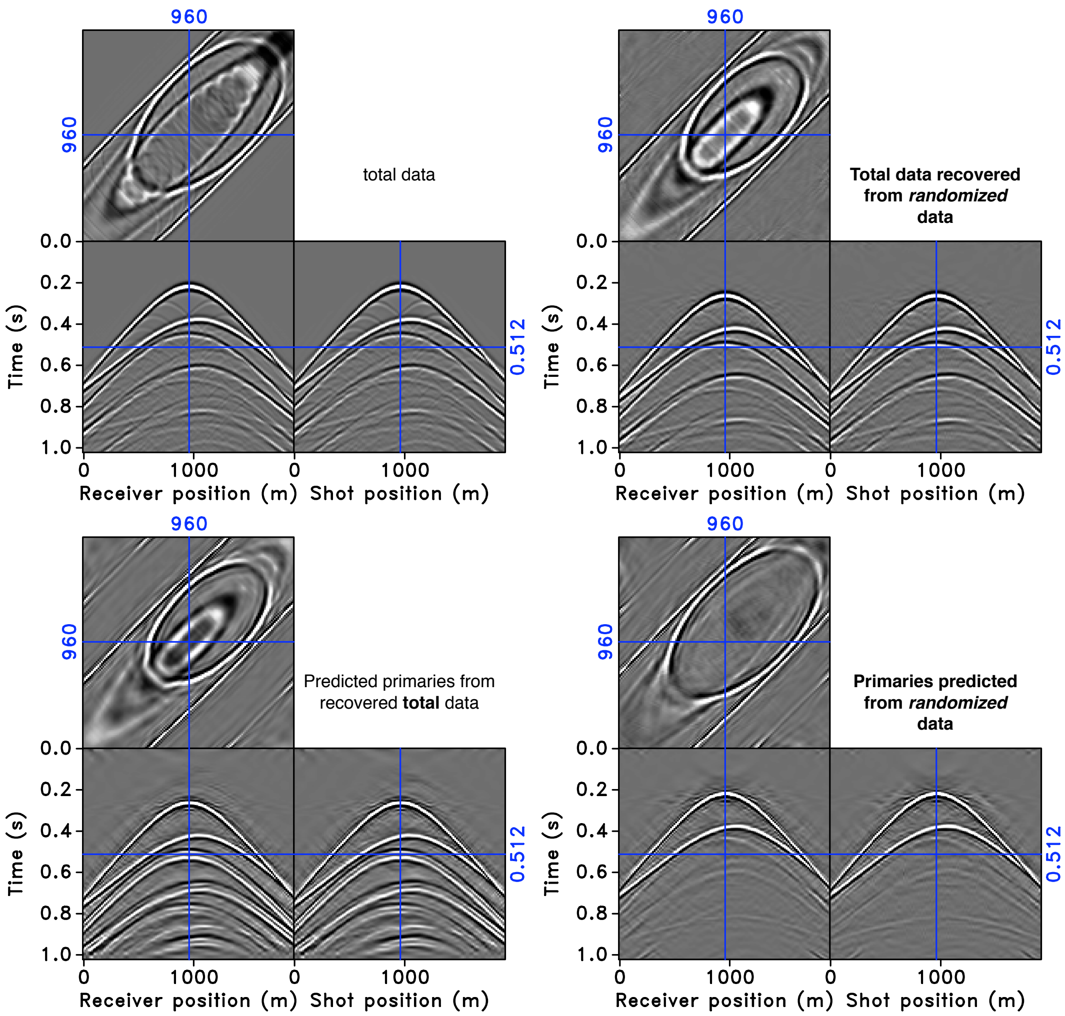

Figure 1 *Sparsity-promoting recovery from simultaneously sampled data with 50 % of the shots missing.* ***(a)*** *Original data.* ***(b)*** *Estimation of the total data by sparsity promotion.* ***(c)*** *Estimation of primaries from recovered total data under* ***(b)****.* ***(d)*** *Estimation of primaries directly from the simultaneously acquired data. Notice the remarkable improvement in the estimation of the primaries* directly *from the simultaneously acquired data.*

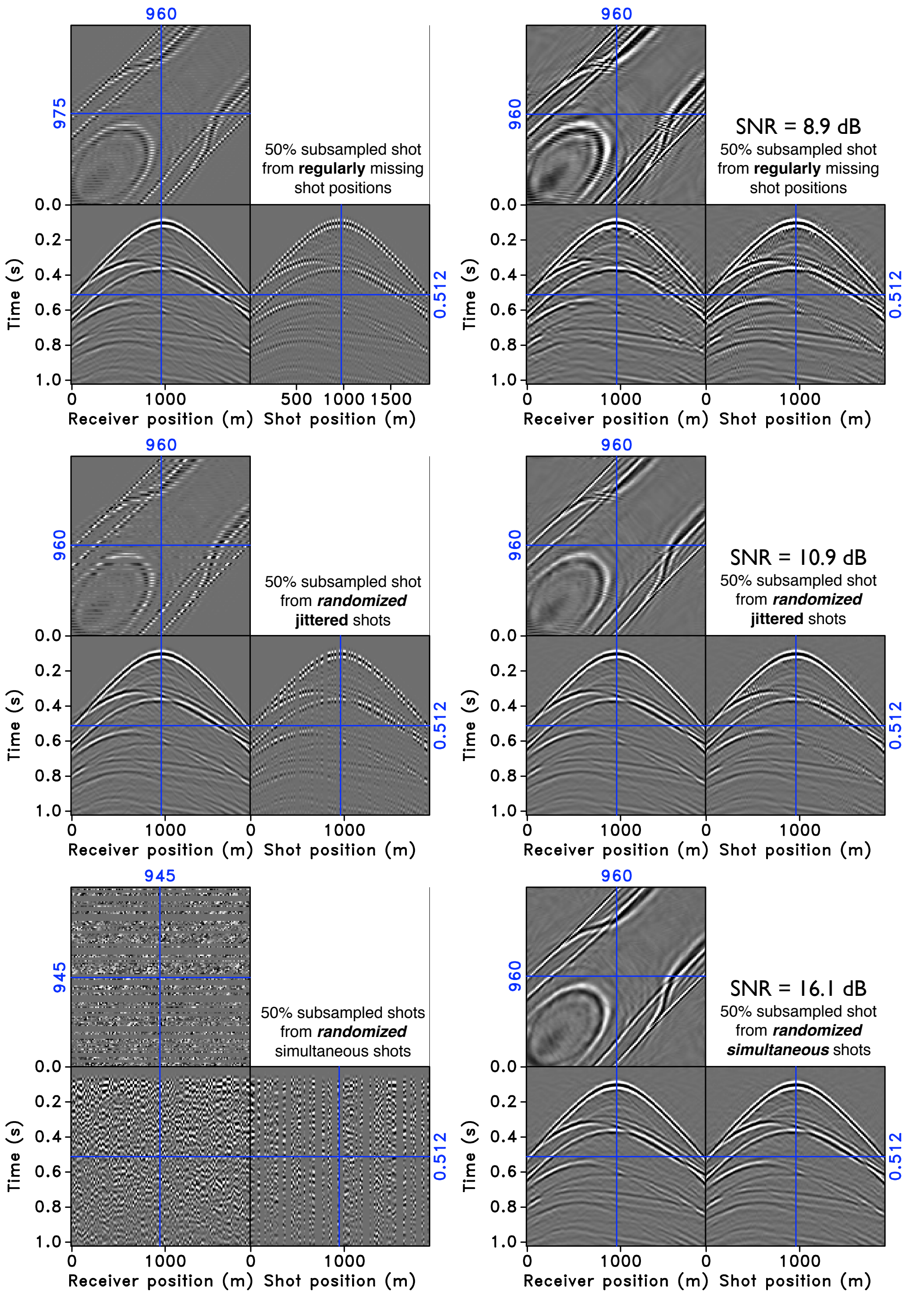

Figure 2 *Sparsity-promoting recovery from 50 % of the shots missing.* ***(a)*** *Regularly subsampled shots.* ***(b)*** *Recovery from regularly subsampled shots.* ***(c)*** *Jittered subsampled shots* ***(d)*** *Recovery from jittered subsampled shots.* ***(e)*** *Subsampled* randomized *simultaneous shots.* ***(f)*** *Recovery from* randomized *simultaneous shots. Notice the remarkable improvement in recovery from the simultaneously acquired data.*