Introduction

Accurate removal of surface related multiples is a key step in seismic data processing. The industry standard for removing multiples is SRME, which involves convolving the data with itself to predict the multiples, followed by an adaptive subtraction procedure to recover the primaries (Verschuur and Berkhout, 1997). Other methods involve multidimensional division of the up-going and down-going wavefields (Amundsen, 2001). However, this apporach may suffer from stability problems. With the introduction of the "estimation of primaries by sparse inversion"(EPSI), van Groenestijn and Verschuur (2009) resentely reformulated SRME to jointly estimate the surface-free impulse response and the source signature directly from the data. The advantage of EPSI is that it recovers the primary response directly, and does not require a second processing step for the subtraction of estimated multiples from the original data. However, because it estimates both the primary impulse response and source signature from the data EPSI must be regularized.

Motivated by recent successful application of the curvelet transform in seismic data processing (Herrmann et al., 2007), we formulate EPSI as a bi-convex optimization problem that seeks sparsity on the surface-free Green's function and Fourier-domain smoothness on the source wavelet. Our main contribution compared to previous work (Lin and Herrmann, 2009), and the conntribution of that author to the proceedings of this meeting(Lin and Herrmann, 2010), is that we employ the physical principle of as source-receiver reciprocity to improve the inversion.

The acoustic wave equation obeys a reciprocal relationship with respect to source and receiver location. This relationship means that the impulse response for a shot at location 1 and receiver at location 2 is equal to the response for a shot at location 2 and receiver at location 1 (Knopoff and Gangi, 1959). It has been shown that real data is reciprocal to within noise levels, and that deviations from reciprocity are due to acquisition geometry, and source directivity (Fenati and Rocca, 1984). When seismic data is organized along time, receiver, and shot location axis, reciprocity makes the data symmetrical about the shot and receiver location axis. Using this fact, we are able to decompose data volumes into their symmetric and asymmetric components, which allows us to use the asymmetric decomposition as a penalty term when estimating the surface-free impulse response. A similar approach has been used in surface consistant amplitude corrections (van Vossen et al., 2006). Source arrays are designed to have a signature that is as impulsive as possible—i.e., short time windows. This is done to make it easier to differentiate between different reflections in the recorded data. Because of this, we regularize the source signature estimation problem with a penalty function that punishes any values outside of a small window in time. The proposed penalty terms are validated on synthetic data, and it is shown that they make improvements to the SNR of the recovered primaries compared to previous work. A comparison of our method with previous work on marine data is also provided.

Theory

Using Berkhouts discrete formulation relating the total upgoing wavefield to the surface free impulse response and source signature, we write single frequency slices taken from a prestack data volume as upper case bold quantities (Berkhout, 1982). The columns and rows are made of the monochomatic entries of common-shot, and common-receiver gathers, respectively. Upper case bold quantities will be overloaded to represent linear operators. In this notation, the monochromatic upgoing wavefield measured at the surface $\hat{\mathbf{P}}$ is equal to the primary impulse response $\hat{\mathbf{G}}$, convolved with the source signature $\hat{\mathbf{Q}}$, plus the reflection at the free surface—i.e.,

$$\hat{\mathbf{P}} = \hat{\mathbf{G}}(\hat{\mathbf{Q}} + \mathbf{R}\hat{\mathbf{P}}). \tag{1}$$

We approximate $\mathbf{R}$ as $-\mathbf{I}$, with $\mathbf{I}$ being the identity operator. The source function $\hat{\mathbf{Q}}$ will be the same for all shot locations— i.e., $\hat{\mathbf{Q}} = \hat{\mathbf{q}}(\omega)\mathbf{I}$. With a slight abuse in notation, and using basic linear algebra, Eq. 1 can be written as $\hat{\mathbf{p}} = \mathrm{BlockDiag}\big[\big((\hat{\mathbf{Q}} - \hat{\mathbf{P}})^* \otimes I\big)_{1..n_f}\big]\mathcal{F}_t\mathbf{g}$. Here $\mathcal{F}_t$ is the 1-D Fourier transform along the time coordinate $\mathcal{F}_t := [\mathbf{I} \otimes \mathbf{I} \otimes \mathcal{F}]$, and $\hat{\mathbf{p}}$, $\mathbf{g}$, $\hat{\mathbf{q}}$ are vectorized versions of the total upgoing wavefield, the surface-free impulse response, and the source signature, respectively. For example $\hat{\mathbf{p}} = \mathrm{vec}\big([\mathbf{P}_{1..n_t}]\big)$. Because there are more unknowns then knowns in the above equations ($\mathbf{g}$, and $\hat{\mathbf{q}}$, versus $\hat{\mathbf{p}}$), the problem of inverting the action of $\mathrm{BlockDiag}\big[\big((\hat{\mathbf{Q}} - \hat{\mathbf{P}})^* \otimes I\big)_{1..n_f}\big]\mathcal{F}_t$ to find $\mathbf{g}$ and $\hat{\mathbf{q}}$ must be regularized.

A 2D curvelet transform ($\mathbf{C}$) along the spatial dimension Kroneckered with a wavelet transform ($\mathbf{W}$) for the temporal dimension has been shown to do an excellent job of sparsely representing seismic wavefields such as $\mathbf{g}$ (Herrmann et al., 2009). We can use sparsity in this transform domain as a constraint, to regularize the inversion problem (Lin and Herrmann, 2009). To do so the adjoint of the combined transform $\mathbf{S}^* = (\mathbf{C} \otimes \mathbf{W})^*$, which takes data from the sparse domain to the physical domain, is inserted into the above equation for the upgoing wavefield—i.e., $\hat{\mathbf{p}} = \mathrm{BlockDiag}\big[\big((\hat{\mathbf{Q}} - \hat{\mathbf{P}})^* \otimes I\big)_{1..n_f}\big]\mathcal{F}_t\mathbf{S}^*\mathbf{x}$ or $\mathbf{p} = \mathbf{A}\left[\hat{\mathbf{Q}}\right]\mathbf{x}$. Here $\mathbf{x}$ is the transform-domain representation of the primary impulse response, for all shots, receivers and time. This equation is used as part of an bi-convex optimization program to recover the unknowns $\mathbf{g} = \mathbf{S}^*\mathbf{x}$ and $\mathbf{q}$ (Lin and Herrmann, 2010).

Because there are two distinct unkown quantities that most be found we break the problem down into two separate optimization programs. Initially $\mathbf{g}$ and $\mathbf{q}$ are set to zero. First we estimate the impulse response $\mathbf{g}$ by solving

$$OP_x: \quad \tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \quad \mathbf{A}\mathbf{x} = \mathbf{p}. \tag{2}$$

We run the sparsity promoting solver (SPGL1 Berg and Friedlander, 2007) for a number of iterations to give a first estimate for the surface-free data. Next, we use this approximation to compute the source function $\tilde{\hat{\mathbf{q}}}$ by solving the second optimization problem:

$$OP_q: \quad \tilde{\mathbf{q}} = \arg\min_{\mathbf{q}} \quad \big\|\big[\hat{\mathbf{P}} + \hat{\mathbf{G}}\hat{\mathbf{P}}\big]_{1..n_f} - \big[\hat{\mathbf{G}}\big]_{1..n_f}\mathcal{F}_t\mathbf{q}\big\|_2. \tag{3}$$

Using the new estimate of $\tilde{\hat{\mathbf{q}}}$ the program OP_x is solved again. This alternating optimization procedure is repeated until a satisfactory solution is found. By splitting the formula up over different frequencies both programs can be implemented in an embarrassingly parallel fashion.

Reciprocity makes $\mathrm{vec}^{-1}(\mathbf{g})$ symmetrical about the shot and receiver axis. Defining a operator $\mathbf{T}\mathbf{g} := \mathrm{vec}(\mathrm{vec}^{-1}(\mathbf{g})^T)$ that transposes data along the shot and receiver axis, allows us to decompose $\mathbf{g}$ into its symmetric $\frac{\mathbf{g}+\mathbf{T}\mathbf{g}}{2}$ and asymmetric components $\frac{\mathbf{g}-\mathbf{T}\mathbf{g}}{2}$. The asymmetric decomposition of $\mathbf{g} = \mathbf{S}^*\mathbf{x}$ is used to redefine

$$OP_x: \quad \tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \quad \|\mathbf{A}\mathbf{x} - \mathbf{p}\|_2 + \alpha_x \Big\|\frac{\mathbf{S}^*\mathbf{x} - \mathbf{T}\mathbf{S}^*\mathbf{x}}{2}\Big\|_2 \le \sigma_x, \tag{4}$$

where α_x and σ_x are the regularization parameter, and the allowable misfit. As stated above, source arrays are constructed to make the source signature as short as possible in time. Because of this OP_q can be redefined with a regularization term that uses a function ($\mathbf{M}$) that penalizes any values outside a small time window.

$$OP_q: \quad \tilde{\mathbf{q}} = \arg\min_{\mathbf{q}} \quad \big\|\big[\hat{\mathbf{P}} + \hat{\mathbf{G}}\hat{\mathbf{P}}\big]_{1..n_f} - \big[\hat{\mathbf{G}}\big]_{1..n_f}\mathcal{F}_t\mathbf{q}\big\|_2 + \alpha_q\|\mathbf{M}\mathbf{q}\|^2 \le \sigma_s, \tag{5}$$

where α_q and σ_q are a weight on the penalty function, and the allowable misfit. Before they are passed to OP_x the estimates of $\mathbf{q}$ are strictly forced to be short in time by applying a tapered windowing operator. This prevents any errors in the estimate of $\mathbf{q}$ at long times from affecting the solution.

Results

Our proposed formulation of EPSI was run on a synthetic data set, modeled with a 15 m shot and receiver spacing, 0.004 ms time sampling and a 60 hz Ricker wavelet source. The true primaries for this synthetic model were obtained by modeling with an absorbing boundary condition at the free surface. For further information on the synthetic data see (van Groenestijn and Verschuur, 2009). The alternating optimization program was run for a total of 150 iterations in SPGL1, with 5 source matchings by LSQR. Figure 1 shows zero offset sections of the total data, true primaries, and the estimated primaries $\hat{\mathbf{P}}_o = \hat{\mathbf{P}}\hat{\mathbf{G}}$ found with and without regularization. Figure 2 shows a comparison of the estimated source functions after one matching, and after the fifth and final matching. On this synthetic data, our method recovers the primaries with a SNR of 13.66 db, and estimates the source wavelet with a SNR of 11.00 db. This yields improvements of 2.52 and 3.86 db over EPSI without reciprocity.

References

Amundsen, L. [2001] Elimination of free-surface related multiples without need of the source wavelet. *Geophysics*, **66**(1), 327–341.

Berg, E.v. and Friedlander, M.P. [2007] SPGL1: A solver for large-scale sparse reconstruction. Http://www.cs.ubc.ca/labs/scl/spgl1.

Berkhout, A.J. [1982] *Seismic migration: Imaging of acoustic energy by wavefield extrapolation A: Theoretical aspects*. Elsevier.

Fenati, D. and Rocca, F. [1984] Seismic reciprocity field tests from the italian peninsula. *Geophysics*, **49**(10), 1690–1700.

Herrmann, F.J., Lin, T. and Erlangga, Y.A. [2009] Compressive simultaneous full-waveform simulation. *SEG Technical Program Expanded Abstracts*, SEG, SEG, vol. 28, 2577.

Herrmann, F., Wang, D., Hennenfent, G. and Moghaddam, P. [2007] Curvelet-based seismic data processing: a multiscale and nonlinear approach. *Geophysics*, **73**(1), A1–A5.

Knopoff, L. and Gangi, A.F. [1959] Seismic reciprocity. *Geophysics*, **24**(4), 681–691.

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

Lin, T. and Herrmann, F.J. [2010] Stablized estimation of primaries by sparse inversion. *EAGE Barcelona 2010*, EAGE.

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.

van Vossen, R., Curtis, A. and Trampert, J. [2006] Surface-consistent amplitude corrections for single of multi-component sources and receivers using reciprocity and waveform inversion. *Geophysics Journal International*, **165**(1), 311–322.

Verschuur, D.J. and Berkhout, A.J. [1997] Estimation of multiple scattering by iterative inversion, part ii: Practical aspects and examples. *Geophysics*, **62**(5), 1596–1611.

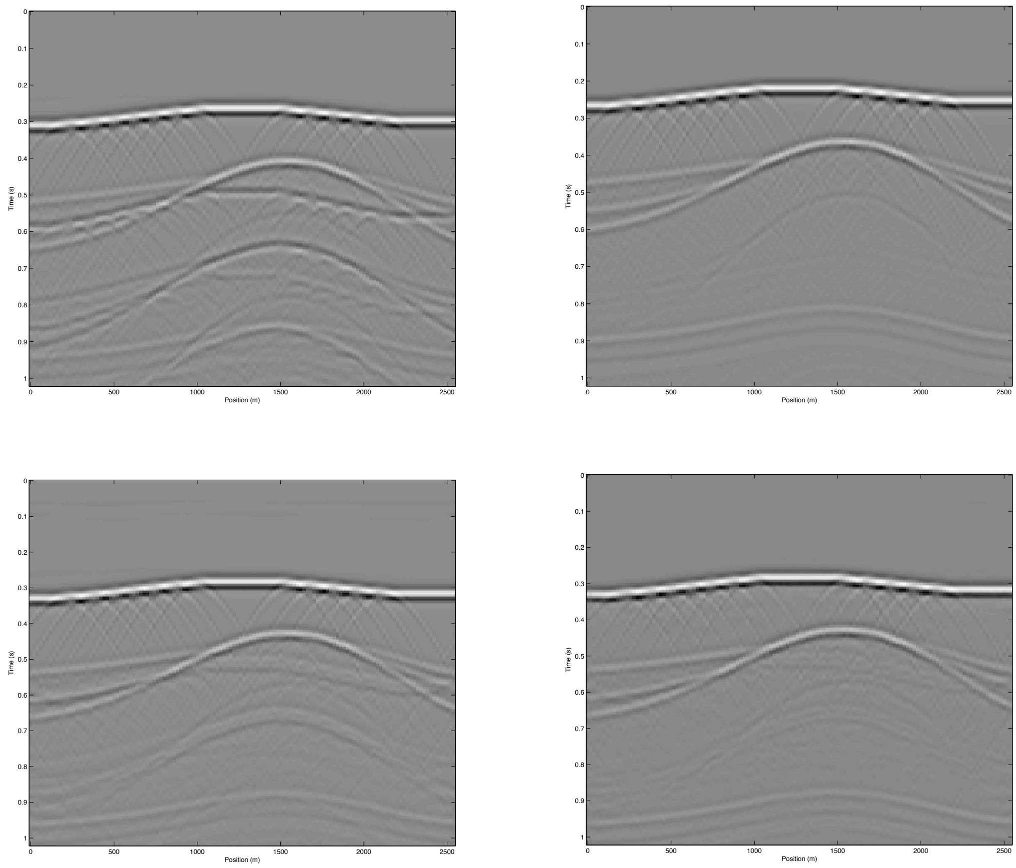

Figure 1 *Comparison of zero offset sections, top-left total input data, top-right "true primaries" (this surface-free data still contains internal multiples), bottom-left primaries estimated with out reciprocity, bottom-right primaries estimated with reciprocity. The estimated primaries are constructed by convolving the source estimate and impulse response. Notice the improvement in removal of multiple energy between with and without reciprocity.*

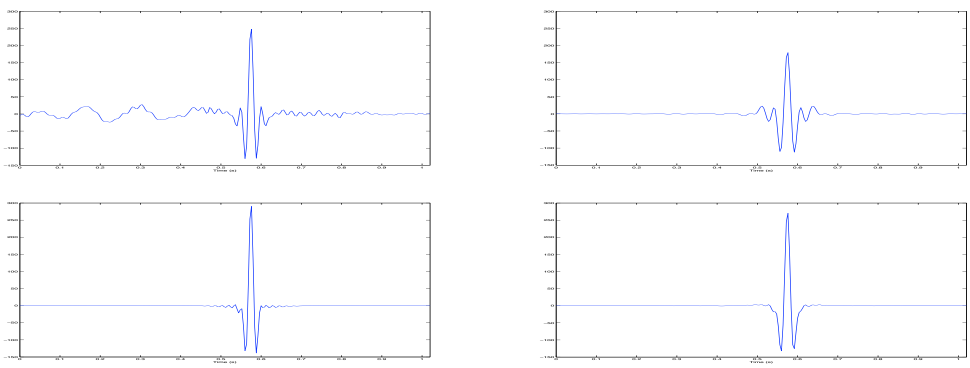

Figure 2 *Comparison of estimated source signature, top left wihtout reciprocity after one mathcing, top right with reciprocity after 1 mathcing, bottom left after 5 matchings without reciprocity, bottom right after 5 matchings with reciprocity.*