Introduction

Estimation of Primaries by Sparse Inversion (EPSI) is a seismic free-surface removal technique recently proposed by van Groenestijn and Verschuur (2009a,b) as an alternative to the traditional adaptive-subtraction process that currently underlies the overwhelming majority of multiple elimination processes. Its theoretical pedigree stems from the familiar relation between the primary and multiple parts of the recorded wavefield that underlies Surface-Related Multiple Elimination (SRME). In vein of Amundsen (2001), EPSI attempts to obtain the primaries via a multidimensional inversion of down-going wavefields with the up-going wavefields. This process can be very unstable in typical field conditions and a regularization process is almost always needed to make the problem well-posed. The choice can be made to control the uncertainty in either the source signature or the "primary" impulse response of the Earth (that is, anything not caused by the existence of its free-surface and the source signature). Here lies the main thematic difference between Amundsen Inversion and EPSI; the former attempts to iteratively determine the source properties exactly using the direct waves, while EPSI imposes the assumption that the correct primary events should be as sparsely-populated in time as possible. This choice naturally leads to a wavefield inversion for the primary impulse response where the penalty function is the energy mismatch between the actual recorded up-going wavefield and the prediction that is produced from the 'SRME' relationship.

Assuming sparsity on the primary wavefield can work like a double-edged sword for EPSI. In general, enforcing the solution's sparsity level in any inversion problem requires a fair bit of nuance and pragmatism to avoid problematic convergence. We previously reported in Lin and Herrmann (2010) on why the original EPSI algorithm converges, and proposed and alternative framework where the minimum sparsity assumption is relaxed to a minimum ℓ_1 norm assumption. Unlike the intractable minimum sparsity problem, this formulation yields a more easily (and reliably) solved convex problem that typically leads to a very good approximation to the sparsest solution. More interestingly, we removed the myriad of physical constraints imposed on individual primary wavefield updates in the inversion, which are necessary to make the original algorithm converge. As mentioned in our work, without these constraints we are free to carry out the sparse inversion process in any representation domain or frame for our solution, as long as the general notion of "sparse events" are conserved. Here, we present the results of enforcing minimum ℓ_1 norm in a hybrid domain of wavelet and 2D curvelet using the reformulated EPSI algorithm. Since wavefronts and edge-type singularities are sparsely represented in the curvelet domain, the results show vastly improved continuity when compared to sparsity regularization in the physical domain, especially in the low-energy events at later arrival times. Such qualities may be greatly desirable for applications of the seismic data that do not inherently involve stacking.

An alternative ℓ_1-relaxed EPSI problem expressed in optimization form

The theoretical foundations of EPSI is usually expressed in a similar fashion to SRME, using the detail-hiding notation of Berkhout and Pao (1982), making use of monochromatic slices of wavefields arranged into a matrix that have columns representing common shot gathers. Here, hatted quantities represent the monochromatic property, and upper-case variables denote matrices or linear operators. The matrix multiplication of two hatted wavefield quantities synonymously imply multidimensional convolutions in the physical domain. With this notation, we introduce the quantities $\mathbf{G}$ representing the primary impulse response, $\mathbf{P}$ the total up-going wavefield, and $\mathbf{Q}$ a (possibly shot-dependent) source signature function. $\mathbf{R}$ is the reflection operator at the surface that is approximated to $-\mathbf{I}$ for the rest of this text. The main physical relationship is thus expressed as

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

Following our previous work, we express EPSI as an optimization problem in the canonical form of solving for an unknown vector quantity $\mathbf{x}$ from a vector observation $\mathbf{b}$, using the relationship $\mathbf{A}\mathbf{x} = \mathbf{b}$ where $\mathbf{A}$ is some linear operator. To write out the proper linear operator for EPSI we introduce the convention of vectorized wavefields in lower case, e.g. $\mathbf{p} = \mathrm{vec}(\mathbf{P})$. We can now express Eq. 1 in terms of an "EPSI prediction" linear operator $\mathbf{E}$ acting on vectorized primary impulse response $\mathbf{g}$:

$$\mathbf{E}\mathbf{g} := \mathscr{F}_t^* \mathrm{BlockDiag}_f[(\hat{\mathbf{Q}} - \hat{\mathbf{P}}^-)^* \otimes \mathbf{I}]\mathscr{F}_t \mathbf{g} = \mathbf{p}, \tag{2}$$

where the block-diagonal elements vary over frequency. The symbol $\otimes$ defines a Kronecker product that, in this case, simply helps reformulate the matrix-matrix operations into matrix-vector operations. $\mathscr{F}_{\mathrm{t}}$ is a Fourier transform in the time axis that also organized the different monochromatic wavefields in a vectorized manner, such that $\mathscr{F}_{\mathrm{t}}\mathbf{g} = \hat{\mathbf{g}} := [\hat{\mathbf{g}}_{f1}, \hat{\mathbf{g}}_{f2}, \ldots, \hat{\mathbf{g}}_{fn}]^T$, while the adjoint operation $\mathscr{F}_{\mathrm{t}}^*$ on the left brings the wavefield back to the time domain. Note that $\mathbf{E}$ is a simple linear operator that depends both on a source signature estimate $\mathbf{Q}$ and the (recorded) up going wavefield $\mathbf{P}$. $\mathbf{Q}$ is an unknown quantity that must be explicitly inverted for at the same time as $\mathbf{g}$.

In this language, the goal of EPSI is to solve an instance of the following *non-convex* optimization problem:

$$\tilde{\mathbf{g}} = \underset{\mathbf{g}}{\operatorname{argmin}} \|\mathbf{p} - \mathbf{E}\mathbf{g}\|_2 \quad \text{subject to} \quad \|\mathbf{g}\|_0 \leq k\rho, \tag{3}$$

where k is the number of iterations taken in the EPSI process. The ℓ_0 pseudonorm $\|\mathbf{g}\|_0$ measures the number of non-zero elements, or in other words the *sparsity*, of $\mathbf{g}$. The parameter ρ is an inversion parameter that control the overall sparsity on the data.

EPSI states that a reasonable estimate of the primary impulse response can be obtained by a steepest-descent inversion process, but it is known that dealing with an ℓ_0 measure in optimization is famously difficult, since uniqueness and existence of a solution cannot be proved in general and no method outside of combinatorial searches are guaranteed to yield the optimal solution. EPSI works around this problem by severely limiting the size of the feasible set at every iteration, such as imposing constraint $\|\delta\tilde{\mathbf{g}}\|_0 = \rho$ for every single gradient update to the solution, so that the update making the largest possible progress to the minimization objective in Eq. 3 can be found by simple searching over the whole set, i.e. zeroing everything in the gradient except the τ-th largest elements.

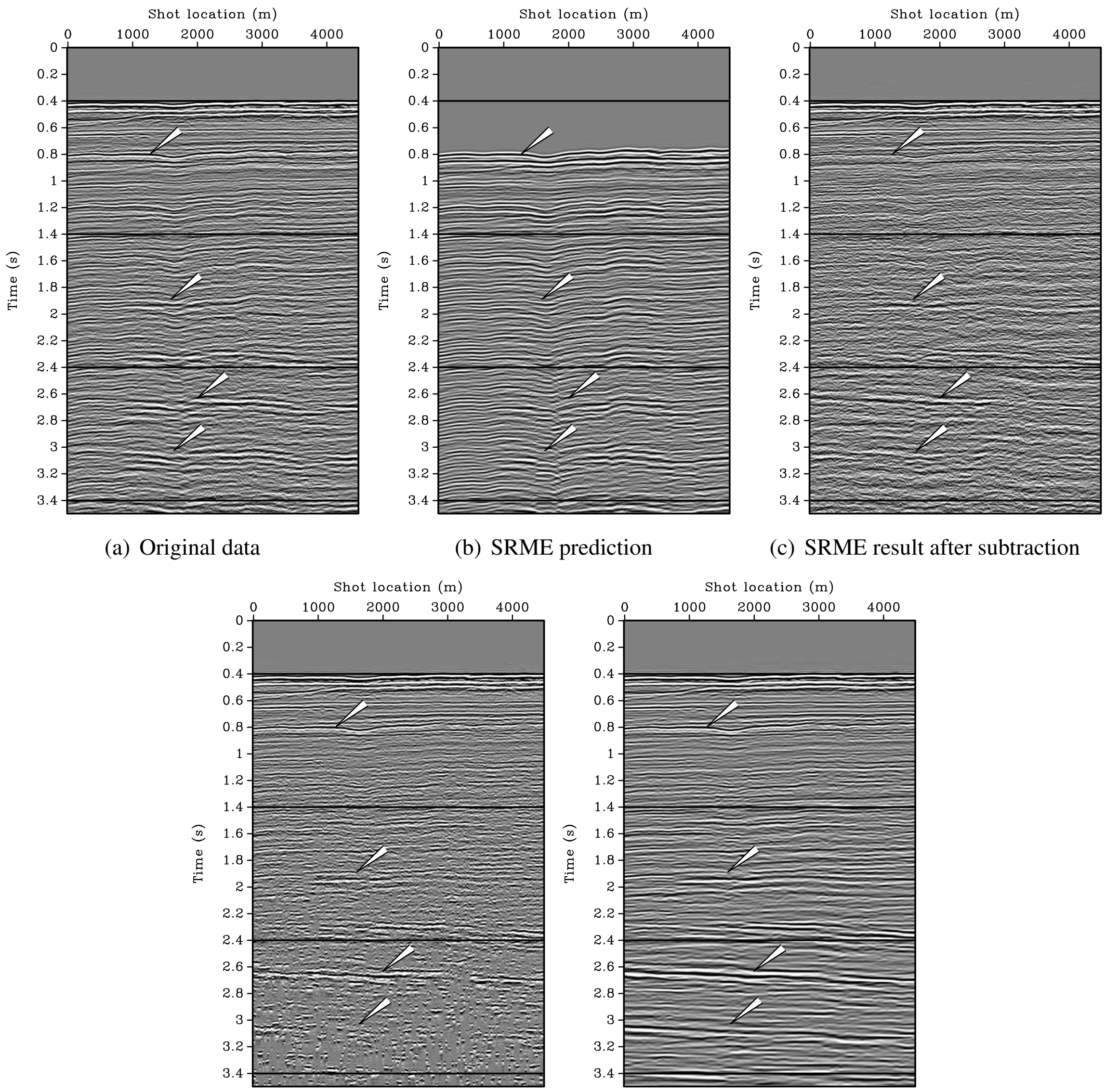
Note that although the previous argument for enforcing the sparsity stems from a purely algorithmic point of view, this particular regularization also has a very physically intuitive meaning. In practice, the allowed ρ events are evenly distributed across every single trace, and is applied in conjunction with an ad-hoc time window that hopes to limit the update to a region without multiples. The number of spikes allowed per trace per update is on the order of 5 or so, translating to the idea that *wavefield updates should be the result of a few discontinuities in the subsurface*. Should conditions be ideal, the allowed events on individual traces should all belong to the same event, with a nice continuous appearance when viewed in a panel. This assumption is, however, easily violated by factors such as complex geometry and noise. This often causes discontinuous wavefronts in the final estimate, especially when they are sufficiently low in energy compared to other events in the data or the overall noise floor. Also, the physical significance behind this regularization breaks down if the updates are not formed in the time-space domain.

We postulated a more tractable version of EPSI by performing a convexification of the above optimization problem. A very well-known strategy when tackling cardinality-constrained problems is to replace the ℓ_0 term with the ℓ_1 norm $\|\mathbf{g}\|_1$, the sum over the element-wise absolute value of $\mathbf{g}$ (Donoho, 2006). Minimizing the ℓ_1 norm typically results in the sparsest solution, a result which is widely known and exploited in the field of compressive sensing. Since taking an ℓ_1-norm is a convex function, this formulation of EPSI is also convex. Convexity is a desirable property, because convex problems in general are stable with no local minima in the objective function. Furthermore, we are allowed to exploit a well-known duality result (van den Berg and Friedlander, 2008) to form a related problem that explicitly asks for the *sparsest* result, rather than guessing at the level of the final sparsity as in Eq. 3. We write it as

$$\tilde{\mathbf{g}} = \underset{\mathbf{g}}{\operatorname{argmin}} \|\mathbf{g}\|_1 \quad \text{subject to} \quad \|\mathbf{p} - \mathbf{E}\mathbf{g}\|_2 \leq \sigma. \tag{4}$$

Here σ is seen as the residual between the recorded data and the total up-going wavefield predicted by the estimated primaries, which is linked to the noise level of the shot record. Due to its physical significance, a good estimate for σ should be more easily determined compared to ρ.

Seeking the solution under a sparsifying transform: Demonstrated in Fig. 1 is the aforementioned wavefront continuity problem that affects both the original and the ℓ_1 version of EPSI. It shows the



(a) Original data (b) SRME prediction (c) SRME result after subtraction

(d) ℓ_1 EPSI result in physical domain (e) ℓ_1 EPSI result in alternative domain

Figure 1 *Common-offset panels (200m) of a North Sea dataset with automatic gain control. Note that EPSI reconstructs late primary events more prominently that SRME, especially after 3s. Also note that the improved fidelity of the primary events in the wavelet-curvelet domain* ***(e)****, along with a smooth appearance that is bolstered by the expected denoising behaviour.*

various attempts to remove surface-related multiples on a North Sea field dataset. Fig. 1(c) shows results from a conventional 2D SRME process. We note that this dataset is known to be problematic for 2D SRME due to strong out-of-plane scattering of the reflected events. Fig. 1(d) shows the solution to Eq. 4 after 85 iterations, with σ set to 0.1 times the 2-norm energy of the recorded wavefield. The lower arrows indicate the primary events that show up more clearly due to the constructive nature — rather than the subtractive nature — of EPSI. However, note that while the strong early events are reconstructed smoothly, the much weaker late events (after 2s) exhibit severe wavefront discontinuities as a by-product of allowing a non-zero residual in a sparsity-regulated inversion.

Although we can certainly try to incorporate additional continuity constraints into the inversion process, we instead exploit the unique property that the ℓ_1 formulation requires no physical regularization on the updates to converge to a sparse solution. Therefore, we are free to consider that a very natural way of solving this problem is to find our sparse solution in an alternative representation domain. This domain should be composed of basis functions of finite energy and local support, so that the notion of a "sparse" solution still correlates with a minimum number of events in the physical space. Additionally, it should represent edge-like singularities with as few atoms as possible, so that seeking a sparse solution does not conflict with the desire for continuous wavefronts. A prime example of such a representation is the curvelet frame (Candès et al., 2006). Fig. 1(e) shows the results of the same ℓ_1 EPSI algorithm used in 1c but seeking the sparse solution partially in terms of curvelet coefficients. Due to computational constraints a hybrid wavelet-curvelet representation, explained below, is used. Notice that the lower events, especially the ones below 3s, are much better represented in this. The multiple at 0.8s remains due to amplitude errors in the data, which may be mitigated by incorporating an additional matching for the surface operator. We discuss this in a separate contribution to this proceeding.

To obtain the results in Fig. 1(e), an analysis (forward transform) operator $\mathbf{S} := \mathbf{C} \otimes \mathbf{W}$ is defined with $\mathbf{C}$ the analysis operator of 2D curvelet frame acting on the shot-receiver axes and $\mathbf{W}$ the analysis operator of a spline wavelet acting in the time direction. The adjoint of this operator (in our case the synthesis) is then compounded with $\mathbf{E}$ in Eq. 4 to form the new problem

$$\tilde{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{p} - \mathbf{E}\mathbf{S}^*\mathbf{x}\|_2 \leq \sigma, \tag{5}$$

which is solved using exactly the same algorithm as the one used for Eq. 4. We use the SPGℓ_1 algorithm (van den Berg and Friedlander, 2008) that is based on a projected gradient approach formulated specifically for large problems such as seismic wavefield reconstruction. Once the coefficients $\tilde{\mathbf{x}}$ are obtained, the solution primary impulse response can be recovered by the synthesis operation $\tilde{\mathbf{g}} = \mathbf{S}^*\tilde{\mathbf{x}}$.

Implications for non-zero-order discontinuities: Because it uses the number of non-zero spikes to represent the number of reflected events in a trace, the original EPSI formulation inherently makes the assumption that reflections originate from zero-order (jump) discontinuities. However it is known that subsurface properties often exhibit a mixture of differentiable as fractional-order discontinuities (Herrmann and Bernabé, 2004). In this case the reflected seismic impulse event cannot be easily represented by a spike in the time domain, potentially inhibiting the original EPSI's ability to reconstruct event of this type. This issue can be avoided when looking for solutions in the representation domain, as wavelet-type atoms (including curvelets) will have sufficient number of vanishing moments to represent reflections from these discontinuities in time in a sparse fashion.

Summary

Flexibility introduced by the ℓ_1 reformulation of EPSI we previously introduced in Lin and Herrmann (2010) makes it possible to construct the solution in alternative representations of a wavefield. By choosing a hybrid wavelet-curvelet domain it's possible to obtain much better wavefront continuity in the solution without changing the actual computational algorithm. We demonstrated here that this leads to a much improved fidelity in field data that typically poses problems for conventional SRME.

References

Amundsen, L. [2001] Elimination of free-surface related multiples without need of the source wavelet. *Geophysics*, **66**(1), 327–341, doi:10.1190/1.1444912.

Berkhout, A.J. and Pao, Y.H. [1982] Seismic migration—imaging of acoustic energy by wave field extrapolation. *Journal of Applied Mechanics*, **49**(3), 682–683, doi:10.1115/1.3162563.

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.

Donoho, D. [2006] For most large underdetermined systems of linear equations the minimal l(1)-norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics*, **59**(6), 797–829.

Herrmann, F.J. and Bernabé, Y. [2004] Seismic singularities at upper-mantle phase transitions: a site percolation model. *Geop. J. Int.*, **159**, 949–960.

Lin, T. and Herrmann, F.J. [2010] Stabilized estimation of primaries via sparse inversion. *72nd EAGE Conference & Exhibition*.

van den Berg, E. and Friedlander, M.P. [2008] Probing the pareto frontier for basis pursuit solutions. *SIAM Journal on Scientific Computing*, **31**(2), 890–912, doi:10.1137/080714488.

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

van Groenestijn, G.J.A. and Verschuur, D.J. [2009b] Estimation of primaries and near-offset reconstruction by sparse inversion: Marine data applications. *Geophysics*, **74**(6), R119–R128, doi:10.1190/1.3213532.