Estimation of surface-free data by curvelet-domain matched filtering and sparse inversion

Mufeed H. AlMatar, Tim Lin, and Felix J. Herrmann, UBC-Seismic Laboratory for Imaging and Modeling

SUMMARY

Matching seismic wavefields and images lies at the heart of many pre-/post-processing steps part of seismic imaging—whether one is matching predicted wavefield components, such as multiples, to the actual to-be-separated wavefield components present in the data or whether one is aiming to restore migration amplitudes by scaling, using an image-to-remigrated-image matching procedure to calculate the scaling coefficients. The success of these wavefield matching procedures depends on our ability to (i) control possible overfitting, which may lead to accidental removal of energy or to inaccurate image-amplitude corrections, (ii) handle data or images with nonunique dips, and (iii) apply subsequent wavefield separations or migraton amplitude corrections stably. In this paper, we show that the curvelet transform allows us to address all these issues by imposing smoothness in phase space, by using their capability to handle conflicting dips, and by leveraging their ability to represent seismic data and images sparsely. This latter property renders curvelet-domain sparsity promotion an effective prior.

CONVEXIFIED EPSI

Until recently surface-related multiples were mostly seen as source of noise that needs to be removed () from the data prior to imaging. Estimation of primaries by sparse inverion (EPSI-van Groenestijn and Verschuur, 2009; Lin and Herrmann, 2009) changes this paradigm by proposing an inversion procedure during which the source function and surface-free impulse response are directly calculated from the upgoing wavefield using an alternating optimization procedure, which inverts for the source and the Green's function. The formulation of EPSI which involves the inversion derives from the following monochromatic relationship:

$$\underbrace{\widehat{\mathbf{G}}}_{\text{surface-free data}} \quad \overbrace{\left[\widehat{\mathbf{Q}}+\mathbf{R}\widehat{\mathbf{P}}\right]}^{\text{downgoing wavefield}} \approx \overbrace{\widehat{\mathbf{P}}}^{\text{upgoing wavefield}} \tag{1}$$

between the known upgoing wavefield $\widehat{\mathbf{P}}$, and surface reflection operator ($\mathbf{R}$, typically $\mathbf{R}=-\mathbf{I}$ with $\mathbf{I}$ the idendity operator), and the unknown source function $\widehat{\mathbf{Q}}$ and surface-free data $\widehat{\mathbf{G}}$ (for details see e.g. Lin and Herrmann, 2009). We use the symbol$\widehat{\ }$to indicate quantities in the temporal Fourier domain.

The original version of EPSI consists of solving an alternating optimization problem during which (i) sparsity is enforced by the zero (quasi)-norm on the updates for the surface free data; Fourier smoothness (shortness) is enforces on the source function. The surface-free data is assumed to be sparse in the physical domain and the source function is assumed to be omnidirection—i.e., $\widehat{\mathbf{Q}}=\mathbf{I}\widehat{\mathbf{q}}$ with $\widehat{\mathbf{q}}$ the discretized Fourier representation of the source function $\hat{q}(\omega)$.

Even though, the original formulation of EPSI is elegant, its reliance on the zero norm makes it unstable and difficult to use in practice. Recent work by Lin and Herrmann (2009) addresses this important issue by reformulating EPSI in terms of a bi-convex optimization problem where the one-norm on a transform-domain representation for the surface free is minimized alternatingly with the windowed two-norm on the source function. The starting point for this bi-convex inversion procedure is the following bi-linear forward model

$$\mathbf{b}=\mathbf{A}[\widehat{\mathbf{Q}}]\mathbf{x}, \tag{2}$$

where $\mathbf{A}[\widehat{\mathbf{Q}}]:=\text{blockdiag}\left(\left[\widehat{\mathbf{Q}}+\mathbf{R}\widehat{\mathbf{P}}\right]^*_{1\cdots n_f}\otimes\mathbf{I}\right)\mathbf{F}_t\mathbf{S}^*$ with $\mathbf{F}_t=(\mathbf{I}\otimes\mathbf{I}\otimes\mathscr{F}_t)$ the temporal Fourier transform, and $\mathbf{S}:=\mathbf{C}_2\otimes\mathbf{W}$ and $\mathbf{C}$, $\mathbf{W}$ the 2D curvelet transform along the source-receiver coordinates, and the wavelet transform along the time coordinate respectively. The symbol * denotes (conjugate) transpose. The 'data' vector is defined by $\mathbf{b}:=\text{vec}(\widehat{\mathbf{P}}_{1\cdots n_f})$. For now, we will assume $\mathbf{R}=-\mathbf{I}$.

Because Equation 2 is linear both in the transform-domain coefficients $\mathbf{x}$ for the surface-free data and in the Fourier coefficients for the source $\widehat{\mathbf{q}}$, we can invert for these properties by alternatingly solving the following two convex optimization problems, namely the sparsity-promoting one-norm optimization problem

$$\widetilde{\mathbf{x}}=\arg\min_{\mathbf{x}}\|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{A}[\widehat{\mathbf{Q}}]\mathbf{x}-\mathbf{b}\|_2\leq\sigma, \tag{3}$$

for the surface-free data, yielding $\widehat{\widetilde{\mathbf{g}}}:=\text{vec}(\widehat{\widetilde{\mathbf{G}}}_{1\cdots n_F})=\mathbf{F}_t\mathbf{S}^*\widetilde{\mathbf{x}}$ as the current estimate (estimates are denoted by the$\widetilde{\ }$symbol) for the surface-free data, and the regularized least-squares problem

$$\widehat{\widetilde{\mathbf{q}}}=\arg\min_{\widehat{\mathbf{q}}}\frac{1}{2}\|\widetilde{\mathbf{y}}-\mathbf{B}[\widehat{\widetilde{\mathbf{G}}}]\widehat{\mathbf{q}}\|_2^2+\lambda_F\|\mathbf{L}_F\widehat{\mathbf{q}}\|_2^2, \tag{4}$$

with $\mathbf{B}[\widehat{\widetilde{\mathbf{G}}}]:=\text{blockdiag}([\widehat{\widetilde{\mathbf{G}}}\mathbf{I}]_{1\cdots n_f})$, yielding an estimate of the source function, given the current estimate for the surface free data. The 'data' vector $\widetilde{\mathbf{y}}$ in this expression is given by $\widetilde{\mathbf{y}}=\text{vec}([\widehat{\mathbf{P}}-\widehat{\widetilde{\mathbf{G}}}\widehat{\mathbf{P}}]_{1\cdots n_F})$ the current estimate for the surface-free data $\widehat{\widetilde{\mathbf{G}}}$. The ℓ_2 penalty term penalizes non smoothness via the sharpening operator $\mathbf{L}_F$. This smoothness penalty in the Fourier domain corresponds to enforcing decay for the estimated source function in the time domain. The parameter λ_F balances Fourier-domain smoothness versus data misfit and avoids overfitting of the 'data' that may lead on leakage of multiples into the source function.

In words the above formulation, involves two alternating optimization problems that solve for the unknown surface-free Green's function and the unknown source function. Given an estimate for the source ($\widehat{\widetilde{\mathbf{q}}}=\mathbf{0}$ initially), surface-free data is estimated by solving an one-norm optimization problem, which seeks the sparsest set of transform-domain coefficients whose inverse transform after 'convolution' with the downgoing wavefield (cf. Equation 2) explains the total upgoing wavefield. Given this estimate for the surface-free data, the source function is estimated by Fourier matching via regularized least-squares.

For details on the solution of this bi-convex optimization problem, we refer the reader to (Lin and Herrmann, 2009) and another contribution by the authors to the proceedings of this conference.

MOTIVATION

Matched filtering for the purpose of matching the amplitudes of wavefields prior subtraction has been an integral part of the seismic data processor's toolbox (see e.g. Verschuur et al., 1992, where matched-filtering is used within Surface-Related Multiple Elimination, SRME). The recent advent of scaling methods for the restoration of migration amplitudes through image-to-remigrated-image matching (see e.g. Guitton, 2004; Herrmann et al., 2008a; Symes, 2008) represent another instance of matched filtering.

We consider a shot record from a synthetic line, generated by an acoustic finite-difference code for a velocity model that consists of a high-velocity layer, which represents salt, surrounded by sedimentary layers and a water bottom that is not completely flat . In Figure 1, the results for correctly and overmatching. Figures 1(a)-1(c) include the total input data with multiples, the SRME-predicted multiples and the "multiple-free" data, respectively. The predicted multiples are the result of conventional matching in a single window. The "multiple-free" data were modeled with an absorbing boundary condition, removing the surface-related multiples. Results for the matched multiples Figures 1(d)-1(e). Comparison of these results shows a significant improvement for the multiple predicion obtained with smoothness regularized curvelet-domain amplitude scaling (calculated by for $\lambda = 0.5$). For this choice of λ, the multiples are not over fitted and the amplitude correction leads to a removal of remnant multiple energy, in particular for the events annotated by the arrows. The value for λ was found experimentally.

CURVELET-MATCHED EPSI

Equation 4 corresponds to Fourier matching, which is also part of Surface-Related Multiple Elimination. However, the difference in this case is that the surface-free data ('primaries') are matched to the total data instead of the predicted multiples as in SRME. This means that the output of Equation 4 produces estimates for the source wavelet and not its 'inverse'. With the appropriate choices for regularization parameters, the above alternating optimization algorithm converges and produces reliable estimates for the surface-free data and the source function.

The success of EPSI hinges on the delicate nonlinear relationship between the source wavelet and the upgoing total wavefield on the one hand and the surface-free impulse response on the other. EPSI leverages this relationship and this explains its success in solving the non-convex (read no global minimum) blind deconvolution problem where the source and surface-free data are both unknown. However, this success comes at a price, namely a strict validity of this nonlinear relationship. This means that we are implicitly assumming that the sources are omnidirectional and colocated, which can be accomplished relatively easily by proper regulariation and extension of Fourier matching to include directivity; absence of 3D effects; infinite aperture; and ideal total reflection at the surface $\mathbf{R} = -\mathbf{I}$. Even though EPSI can in principle be extended to 3D, the latter

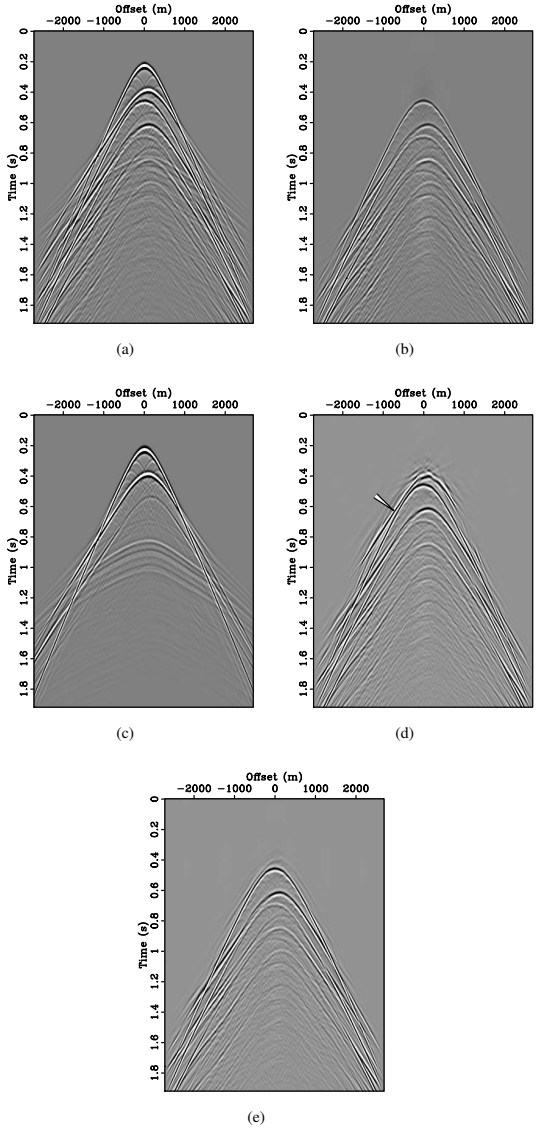


Figure 1: Primary-multiple separation on a synthetic shot record. **(a)** The total data, **p**, including primaries and multiples. **(b)** Single-term SRME-predicted multiples wavelet-matched within a global window ($\check{s}_2$). **(c)** Reference surface-related multiple-free data modeled with an absorbing boundary condition. **(d)** Estimate for the multiples using without and **(e)** with curvelet-domain smoothing. The latter clearly prevents overfitting.

assumptions are more problematic to handle in practice as experience with SRME-based methods have shown (). This may have a detrimental effect on the quality of the estimated surface-free data and because we exploit a nonlinear relationship, there is a good chance that these problem will be even exacerbated compared to single-term SRME.

Following the succesful application of curvelet-domain matched filtering during the SRME processing flow, we propose to incorporate this step in our methodology. This will allow us to mitigate possible adverse effects that step from amplitude mismatched that vary smoothly as a function of position and dip along the predicted wavefronts of the current estimate for the surface-free data. We correct for these errors by including a non-trivial surface-reflection operator in the formulation—i.e., $\mathbf{R} \neq -\mathbf{I}$. This operator will be able to absorb amplitude errors of non-ideal reflections at the surface, finite-aperture, and other unknown effects.

To keep the formulation relatively simple, we assume that this operator is frequency independent, even though our formulation can easily be extended to include angular frequency dependence. (For this reason, we dropped the hat in the above equations.) We also assume $\mathbf{R}$ to be symmetric positive definite and pseudo local (no kinematic shifts). As in our earlier work on the curvelet-domain matched filter, we model the amplitude errors—i.e., the operator $\mathbf{R}$—by a zero-order pseudodifferential operator, which can be interpreted as a space-dip-dependent filter. This operator can be approximated by a simple curvelet-domain scaling—i.e, we have

$$\mathbf{R} \approx \mathbf{C}^* \operatorname{diag}(z)\mathbf{C}, \tag{5}$$

During curvelet-domain matched filtering the diagonal in this expression is calculated by solving a regularized least-squares problem, possibly supplemented with a positivity constraint, i.e, $\mathbf{z} > 0$, during which the two wavefield are matched after conservative global Fourier-domain matching. This Fourier matching removes the time characteristics of the source function. During SRME, the two wavefields that are matched are the total data and the source-corrected predictions for the surface-related multiples.

Similarly, we can extend the surface-free data estimation by substitution Equation 5 for $\mathbf{R}$ in our formulation and rearrange terms, yielding

$$\operatorname{vec}([\widehat{\mathbf{P}} - \widetilde{\widehat{\mathbf{G}}}\widehat{\widetilde{\mathbf{Q}}}]_i) = \widehat{\mathbf{G}}_i \mathbf{C}^* \operatorname{diag}(\mathbf{C}\operatorname{vec}(\widehat{\mathbf{P}}_i))\mathbf{z}. \tag{6}$$

Incorporation of this expression into our formulation, yields a tri-convex optimization problem that now includes curvelet-domain matching.

$$\min_{\mathbf{z}} \frac{1}{2}\|\widetilde{\mathbf{u}} - \mathbf{M}\mathbf{z}\|_2^2 + \lambda_C \|\mathbf{L}_c \mathbf{z}\|_2^2, \tag{7}$$

with $\mathbf{M} := \widehat{\mathbf{G}}_i \mathbf{C}^* \operatorname{diag}(\mathbf{C}\operatorname{vec}(\widehat{\mathbf{P}}_i))$, and $\widetilde{\mathbf{u}} := \operatorname{vec}([\widehat{\mathbf{P}} - \widetilde{\widehat{\mathbf{G}}}\widehat{\widetilde{\mathbf{Q}}}]_i)$. The index i corresponds to the frequency for which $\widetilde{\widehat{\mathbf{q}}}$ is maximum.

As in the case of Fourier-domain matching, this inversion procedure can be regularized by imposing smoothness. Again, the smoothness is promoted by the sharpening operator $\mathbf{L}_c$. This formulation seeks a solution fitting the total data with a smoothness constraint imposed by the sharpening operator $\mathbf{L}_c$, which for each scale penalize fluctuations amongst neighboring curvelet coefficients in the space and angle directions (see Herrmann et al., 2008b, for a detailed description). The amount of smoothing is controlled by the parameter λ. For increasing λ, there is more emphasis on smoothness at the expense of overfitting the data (i.e., erroneously fitting the primaries), see Figure 3(d). For a specific λ, the penalty functional in Equation 7 is minimized with respect to the vector $\mathbf{z}$ with least squares (Nocedal and Wright, 1999).

This expression is similar to Fourier-domain matching

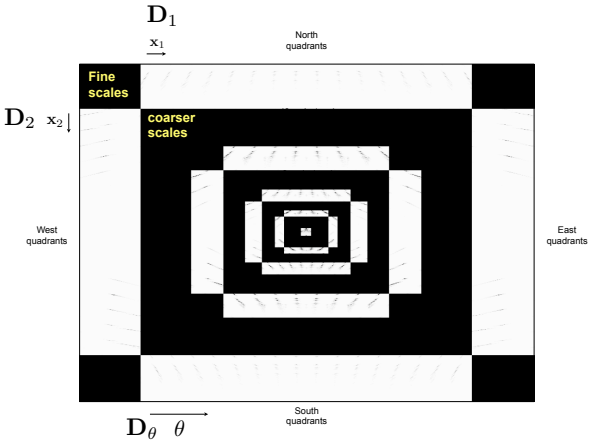


Figure 2: Curvelet decomposition at different frequencies (scale) and angles (dip). Five scales are used for this decomposition. The centre (coarsest scale) shows the DC and low frequency components . The 2nd coarsest scale has 16 angles. The number of angles is doubled each other scale.

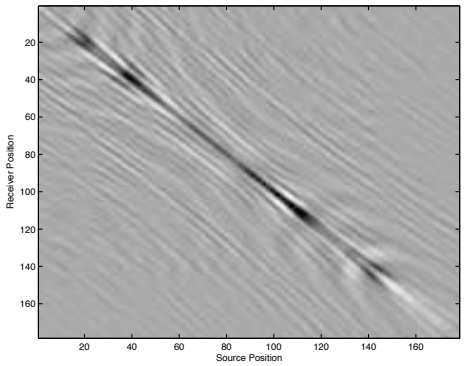
Result: Estimate for $\mathbf{Sx}$, $\mathbf{Q}$ and $\mathbf{R}$
choose noise level σ
$\mathbf{x}_0 \longleftarrow 0$, $\mathbf{Q}_0 \longleftarrow 0$, $k \longleftarrow 1$
while $\|\mathbf{A}[\widetilde{\mathbf{Q}}]\mathbf{x} - \mathbf{b}\|_2 \leq \sigma$ **do**
  $(\mathbf{x}_k, \tau_k) \leftarrow$ solve #3 using initial guess
  $\mathbf{x}_{k-1}, \tau_{k-1}, \mathbf{Q}_{k-1}, \mathbf{R}_{k-1}, \mathbf{z}_{k-1}$ with SPGℓ_1.

  $\mathbf{Q}_k \leftarrow$ solve #4 to match wavelet $(\mathbf{Q})'$ with LSQR
  $\mathbf{z}_k \leftarrow$ solve #7 with LSQR
  $\mathbf{R}_k \leftarrow \mathbf{C}^* \operatorname{diag}(\mathbf{z_k})\mathbf{C}$
end
$\widetilde{\mathbf{Q}} \leftarrow \mathbf{Q}_k$
$\widetilde{\mathbf{R}} \leftarrow \mathbf{C}^* \operatorname{diag}(\mathbf{z_k})\mathbf{C}$

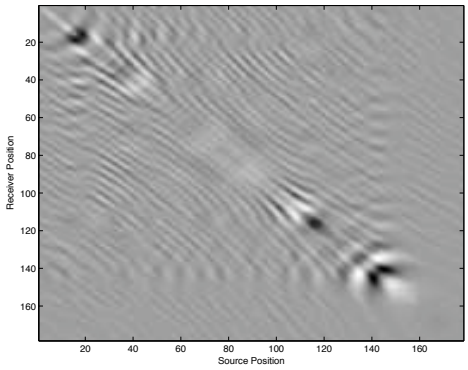
Algorithm 1: Stabilized Estimation of Primaries by Sparse Inversion

ACKNOWLEDGMENTS

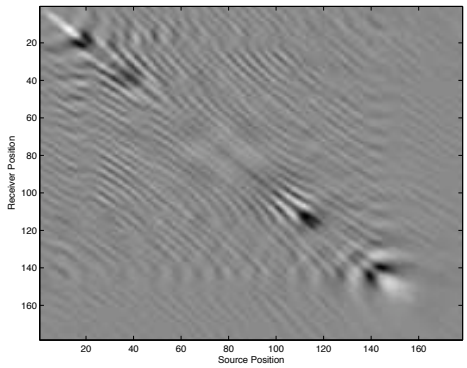
This work was financed through the SINBAD II project with support from BG, BP, Chevron, Petrobras, and WesternGeco and was partly supported by a NSERC Discovery Grant (22R81254) and CRD Grant DNOISE (334810-05).



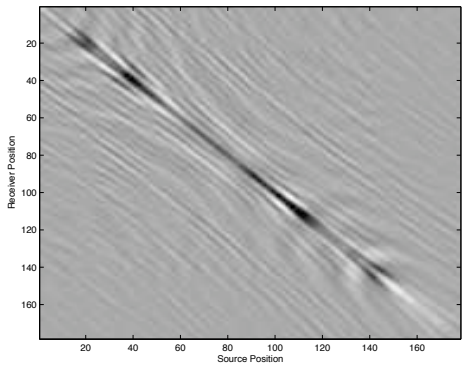
(a)



(b)

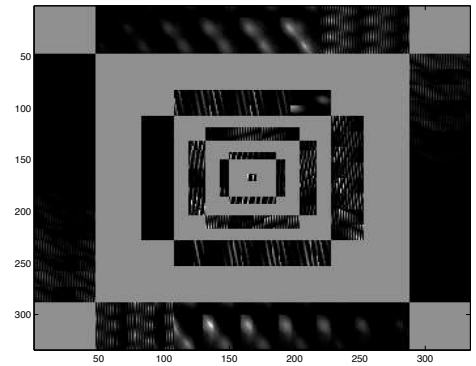


(c)

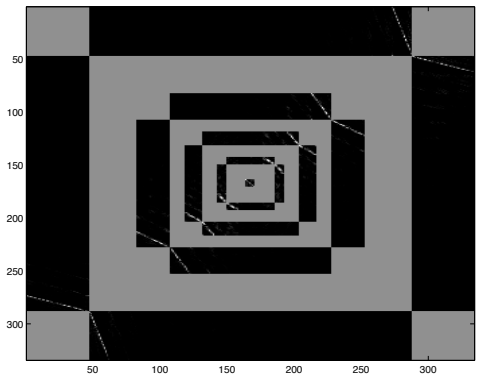


(d)

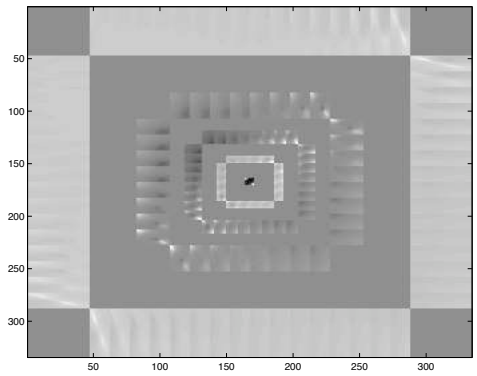
Figure 3: A source receiver slides that correspond to the frequency for which $\widehat{\widetilde{\mathbf{q}}}$ is maximum. **(a)** $\widehat{\mathbf{P}} - \widehat{\widetilde{\mathbf{G}}}\widehat{\widetilde{\mathbf{Q}}}$ in equation (6). **(b)** $\widehat{\mathbf{G}}\mathbf{R}\widehat{\mathbf{P}}$ with $\mathbf{R} = \mathbf{I}$ and **(c)** with $\mathbf{R}$ approximated by a simple curvelet-domain scaling. **(d)** Increasing λ yields more smoothing at the expense of overfitting the data.



(a)



(b)



(c)

Figure 4: Mosaic plots of the estimated diagonal weighting $\mathbf{z}$ at different scales. **(a)** For large λ, the coefficients are smooth at the expense of over fitting the data. **(b)** For small λ, the coefficients are rough and the matching under-fits the data. **(c)** The λ used to generate Figure 3(c).