GEOPHYSICS, VOL. 80, NO. 6 (NOVEMBER-DECEMBER 2015); P. WD11–WD18, 10 FIGS.
10.1190/GEO2015-0016.1

Downloaded 08/20/15 to 138.67.154.174. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Nonminimum phase deconvolution in the log domain: A sparse inversion approach

Antoine Guitton[1] and Jon Claerbout[2]

ABSTRACT

Being disturbed by the discrepancy between the Ricker wavelet and minimum phase wavelets, we wondered if a sparseness criterion could get us deconvolved data with the event polarity being more clearly evident. Five data sets found it does. The sparseness criterion we used is a hyperbolic penalty function. It ranged from ℓ_2 at small residuals to ℓ_1 at large residuals. The main pitfall was that introducing negative filter lags introduced a null space (obviously so for Gaussian data). The null space demanded a regularization. We found a formulation in the domain of the Fourier transform of a log spectrum, in which a Ricker-style regularization appeared. Curiously, this regularization eliminated the leg jumps. A quasi-Newton solver was faster than that of our earlier work, a combination of conjugate directions with a Newton solver.

INTRODUCTION

This paper continues our quest started in Claerbout and Guitton (2014) for the extraction of accurate wavelets from seismic data. In that paper, we claim that the long-standing assumption in reflection seismology that seismic sources are of minimum phase is not valid (Sacchi and Ulrych, 2000): We should be seeing something closer to a symmetrical Ricker wavelet (Ricker, 1953) resulting from water-surface ghosts at the marine gun and hydrophone. To this end, we present an analytical method in the lag-log domain (the Fourier transform [FT] of the log spectrum, also known as *cepstrum*) that estimates a Ricker-like, more symmetrical wavelet by suppressing the phase at small lags. Now, we introduce noncausal filter coefficients (inverse source). These filter coefficients tend to make the problem underdetermined, so we introduce a sparsity goal hoping to determine the correct phase. Indeed, we are delighted to find more plausible polarities in the deconvolved data.

Our earlier work (Zhang et al., 2011) often achieves this goal, but not always. We experience leg jumps (and polarity change with half leg jumps), in which the output spike locations changed with time. It took us considerable time to recognize that regularization was required and even more time to learn how to regularize. We did eventually succeed in the lag-log domain. There, we can understand that gently suppressing the phase at small lags amounts to gently pushing a filter toward a symmetric (Ricker) filter near $t = 0$, which fully resolves the troublesome leg-jump issue.

To achieve sparsity, we move toward the ℓ_1-norm. To preserve rapid solutions, we began using the hyperbolic penalty function (HPF) (Claerbout and Fomel, 2014). It handles small residuals in the usual ℓ_2-like manner but large residuals with an ℓ_1-like penalty. We began fitting to the HPF by a mixture of conjugate directions and the Newton method, but then we found the quasi-Newton method even simpler and faster.

LOG SPACE, SPARSITY, AND GAIN

Traditionally, filters are characterized as coefficients in the time domain, and that is how we started. We experienced a mixture of excellent results and frequent bizarre results (leg jumps) teaching us that sparsity was not always sufficient to control and define phase. We did not get control of the (nonlinear) problem until we cast it in the frequency domain with the logarithm of the filter spectrum expressed back in the time domain. By allowing the filter both sides of zero lag, we inadvertently introduced a null space that requires regularization to manage. Because echoes weaken in time, residuals must be appropriately scaled, so we take some care to do that correctly.

Inversion in the lag-log space

A minimum phase wavelet can be made from any causal wavelet by taking it to Fourier space and exponentiating. The proof is

Manuscript received by the Editor 9 January 2015; revised manuscript received 18 March 2015; published online 8 June 2015.
[1]Geo Imaging Solutions, Inc., San Mateo, California, USA. E-mail: aguitton@gmail.com.
[2]Stanford University, Stanford, California, USA. E-mail: claerbout@stanford.edu.
© 2015 Society of Exploration Geophysicists. All rights reserved.

Downloaded 08/20/15 to 138.67.154.174. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

straightforward: Let $U(Z) = 1 + u_1 Z + u_2 Z^2 + \cdots$ be the Z-transform ($Z = e^{i\omega}$) of any causal function. Then, $e^{U(Z)}$ is the minimum phase. Although we perform such calculations in the Fourier domain, the easy proof is in the time domain. The power series for an exponential is $e^U = 1 + U + U^2/2! + U^3/3! + \cdots$. Each term of the series is causal because the powers of U^n are convolutions of the causal function $U(Z)$ with itself (the convolution of two causal functions is causal). Likewise, e^{-U}, the inverse of e^U, is causal. Thus, the filter and its inverse are causal.

We seek to find two functions, one strictly causal and the other strictly anticausal (nothing at $t = 0$):

$$U^+ = u_1 Z + u_2 Z^2 + \cdots, \tag{1}$$

$$U^- = u_{-1}/Z + u_{-2}/Z^2 + \cdots. \tag{2}$$

Notice that U, U^2, etc., do not contain Z^0. Thus, the coefficient of Z^0 in $e^U = 1 + U + U^2/2! + \cdots$ is unity. Thus, $a_0 = b_0 = 1$:

$$e^{U^+} = A = 1 + a_1 Z + a_2 Z^2 + \cdots, \tag{3}$$

$$e^{U^-} = B = 1 + b_1/Z + b_2/Z^2 + \cdots. \tag{4}$$

We define $U = U^- + U^+$. The decon filter is $AB = e^U$, and the source waveform is its inverse e^{-U}. Defining the FT of the data d_t as $D(\omega) = \mathrm{FT}(d_t)$, the decon output (reflectivity) becomes

$$r_t = \mathrm{FT}^{-1}(D(\omega) e^U), \tag{5}$$

where U is a Z-transform with coefficients parameterizing the filter. We formerly had two causal time functions a_t and b_t combined nonlinearly, whereas with u_t, there is a single noncausal function to be found.

The hyperbolic penalty function

For filter estimation, we try to sparsify the reflectivity (Taylor et al., 1979; Chapman and Barrodale, 1983) by minimizing the HPF. The HPF is convex and behaves like the ℓ_1-norm for high residuals and the ℓ_2-norm for small residuals (Li et al., 2012; Kazemi and Sacchi, 2014). Taking advantage of the ℓ_1 part, we can obtain sparse residuals. The filter e^U is parameterized by $U(\omega)$. Expressing U in the time domain is a sequence of Z-transform coefficients u_t. To obtain these coefficients, we use linear algebra, packing the u_t coefficients into a vector $\mathbf{u}$. The penalty function to minimize is

$$f(\mathbf{u}) = \sum_t H(r_t), \tag{6}$$

where $H(r) = \sqrt{r^2 + R^2} - R$, and R is the ℓ_1/ℓ_2 threshold parameter. We show in Appendix A the derivation of the gradient of equation 6 in the lag-log domain τ:

$$\nabla f(\mathbf{u}) = \sum_t r_{t+\tau} H'(r_t), \tag{7}$$

where $H'(r) = \mathrm{softclip}(r) = r/\sqrt{r^2 + R^2}$. The gradient is the correlation of the reflectivity with its softclipped self. Recall that $u(0) = 0$; therefore, $\nabla f(\mathbf{u})_0 = 0$. For the minimization, we use the L-BFGS method, which requires only the value of the penalty function and its gradient to estimate quasi-Newton updates (Liu and Nocedal, 1989).

Introducing gain functions

Seismic data are nonstationary: Their amplitude and spectra vary with time and space. Here, we introduce a gain function g_t that compensates for only the time gain. We apply the HPF to the gained residual $q_t = g_t r_t$ and set out to sparsify q_t. Therefore, we need its derivative by the model parameters u_τ. In Appendix A, we show the following:

$$\nabla f(\mathbf{u}) = \sum_t r_{t+\tau} g_t H'(q_t). \tag{8}$$

Equation 8 says to crosscorrelate the physical residual r_t with the gained softclipped statistical residual. Notice that in reflection seismology, the physical residual r_t generally decreases with time, whereas the gain g_t generally increases to keep the statistical variable q_t roughly constant; therefore, $g_t H'(q_t)$ grows in time. Our default gain function is $g_t = t^{\mathrm{pow}}$, where $\mathrm{pow} = 2$ corresponds to the standard geometric spreading correction augmented by constant Q.

REGULARIZATIONS

As with any model-fitting problem, regularization defines the solution in which the data are unable to or where the model filter is being overfit (very long lags). To this effect, we introduce two regularization terms. First, we do not want the filter as long as a seismic trace, so we penalize coefficients of $\mathbf{u}$ at long lags (beyond the expected bubbles) by a weighting function $\mathbf{P}$ that vanishes until we get beyond those lags where it becomes unity. This weight is a function of frequency (or of lag). The transition from zero to unity is gently done with a sine squared. In operator form, we introduce the vector $\mathbf{r}_p$ as follows:

$$\mathbf{r}_p = \mathbf{P}\mathbf{u}. \tag{9}$$

We occasionally see leg-jump-like behavior in which, after many iterations, the phase of the wavelet changes rapidly. We attribute this change to the null space we introduce by allowing the filter to have extra coefficients before $t = 0$. To control this null space, we need a regularization. We choose to give the filter a slight push toward the Ricker wavelet. Theoretically, the even part of $\mathbf{u}$ controls the amplitude spectrum of the shot waveform, whereas the odd part controls the phase. The near-zero lags in $\mathbf{u}$ control the near-zero lags in the shot waveform and decon filter (Stoffa et al., 1974; Jin and Rogers, 1983). The odd lags represent the phase. We want the odd near-zero lags to be weakened because for the Ricker wavelet, they are zero (Claerbout and Guitton, 2014). Therefore, this regularization is to gently suppress the antisymmetric part of the near-zero lags of $\mathbf{u}$. Thus, for small lags, we want the following:

$$0 \approx u_\tau - u_{-\tau}. \tag{10}$$

Downloaded 08/20/15 to 138.67.154.174. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

In operator form, this is $\mathbf{0} \approx \mathbf{r}_m = \mathbf{J}\mathbf{u}$, where the matrix $\mathbf{J}$ is defined in the following in the ordering required by the periodic boundaries found in FT codes:

$$\mathbf{0} \approx \begin{bmatrix} r_m(1) \\ r_m(2) \\ r_m(3) \\ \vdots \\ r_m(\tau/2) \\ \vdots \\ r_m(\tau-1) \\ r_m(\tau) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ 0 & +1 & 0 & \dots & 0 & \dots & 0 & -1 \\ 0 & 0 & +1 & \dots & 0 & \dots & -1 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & -1 & \dots & 0 & \dots & +1 & 0 \\ 0 & -1 & 0 & \dots & 0 & \dots & 0 & +1 \end{bmatrix} \begin{bmatrix} u(1) \\ u(2) \\ u(3) \\ \vdots \\ u(\tau/2) \\ \vdots \\ u(\tau-1) \\ u(\tau) \end{bmatrix}. \tag{11}$$

Note that the matrix is Hermitian $\mathbf{J}^* = \mathbf{J}$. We add a diagonal matrix of weights $\mathbf{W}$ that allows us to smoothly incorporate this regularization on the small lags with a $\sin^2$ function:

$$\mathbf{r}_m = \mathbf{W}\mathbf{J}\mathbf{u}, \tag{12}$$

where the diagonal elements are symmetric around lag zero and equal to zero at $\tau/2$ (in effect, the zero-frequency component):

$$\mathbf{W} = \mathrm{diag}(1, w_1, w_2, \dots, 0, \dots, w_2, w_1). \tag{13}$$

Putting everything together, we minimize the following functional:

$$f(\mathbf{u}) = \sum_t H(r_t) + \alpha\|\mathbf{P}\mathbf{u}\|^2 + \beta\|\mathbf{W}\mathbf{J}\mathbf{u}\|^2, \tag{14}$$

where $\|.\|^2$ is the ℓ_2-norm. The first regularization term tends to limit the range of filter lags. The second term encourages symmetry near $t = 0$. The gradient search direction becomes

$$\nabla f(\mathbf{u}) = \sum_t r_{t+\tau} g_t H'(q_t) + \alpha\mathbf{P}^*\mathbf{r}_p + \beta\mathbf{J}^*\mathbf{W}^*\mathbf{r}_m. \tag{15}$$

The regularization parameters α and β are selected by trial and error starting from a value of zero (see the section "Parametrization of the inverse problem" for details).

A PSEUDOCODE

A pseudocode illustrates how to find the best single filter for a group of seismograms, in which the gradient is obtained by averaging the contribution of all the traces to be deconvolved. For the sake of clarity, we omit the regularization terms, but keep in mind that they are used in practice. Lowercase letters are used for variables in time and space like `d` $= d(t,x)$, `r` $= r(t,x)$, `g` $= g(t,x)$, and `q` $= q(t,x)$, whereas uppercase letters are used for functions of frequency `D` $= D(\omega,x)$, `R` $= R(\omega,x)$, `dR` $= \Delta R(\omega,x)$, `U` $= U(\omega)$, and `du` $= \Delta U(\omega)$. An asterisk (*) means to multiply within an implied loop on t or ω.

```
D = FT(d)
U = 0.
Iteration -
  dU =0
  f(u)=0 # penalty function
  For all x positions
    r = IFT( D * exp(U))
    q = g * r
    f(u) = f(u) + H(q)
    dU = dU + conjg(FT(r)) * FT(g*softclip(q))
  Remove the mean from dU(omega)
  du= IFT(dU)
  du_update <- QuasiNewton(du,f(u))
  u = u + du_update
  U = FT(u).
```

EXAMPLES

Now, we illustrate our deconvolution method on one synthetic and four field data sets. The results demonstrate the ability of our technique to effectively estimate seismic wavelets, while unraveling reflectivities. Note that the results we present herein are the actual q_t outputs (gained residuals). Some ringing in the deconvolved sections is present because of the flatter spectra close to the Nyquist frequency. This ringing is less troublesome on full-screen displays.

A synthetic data example

In 2012, Chevron released a synthetic data set for testing model-building tools. Shot gathers and a synthetic source waveform used for the modeling were made available. Our goal is to deconvolve the zero-offset section of this data set with our deconvolution in the lag-log domain. Figure 1 shows a subset of the zero-offset section of interest multiplied by the gain function $g_t = t$: Some familiar events, such as a water-bottom, sedimentary layers, a top salt between 3.2 and 3.4 s, and even a bottom salt at 3.5 s and $x = 65{,}000$ m are visible. After deconvolution, the polarity of all these events is clearly visible in the gained residual panel (Figure 1b). Notice the strong top and bottom salt reflections with opposite polarities.

The synthetic waveform provided by Chevron serving as a source function for modeling is ghost free. To compare with our estimated wavelet using inversion, we add the source and receiver ghosts to this synthetic waveform (taking into account depth and offset information) to form a "total" source function. Figure 2 displays the total source and the estimated wavelet with deconvolution. In this plot, the total source is displayed 5% above the estimated wavelet. They

Downloaded 08/20/15 to 138.67.154.174. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

are remarkably close. Our deconvolution was able to find a very accurate wavelet.

Four marine data examples

Our first field data example demonstrates that our deconvolution method can estimate wavelets far removed from the minimum phase. Figure 3a shows a subset of the near-offset section of a data set from offshore California acquired in the 1980s (Fowler, 1988). Apparently, this data set was gained before we got it, so we use $g_t = 1$. There is a strong precursor at 0.05 s before the main pop. The water bottom shows two reflections (a weak one above a stronger one). After deconvolution with our method (Figure 3b), the water-bottom reflection is reduced to one event, and the frequency content has increased noticeably. The complex wavelet is shown in Figure 4a. We notice the precursor at negative lags before the main event arrives. It turns out that this data set was acquired with a watergun source. Our estimated wavelet resembles quite well the far-field pressure signatures of such sources seen, for instance, in Safar (1985a) (e.g., his Figure 7) and Safar (1985b) (e.g., his Figure 6). Also, the water gun invented for no bubble shows no bubble.

Our second example is from the Gulf of Mexico. Figure 5a shows a subset of the near-offset section that we wish to deconvolve with a gain of $g_t = t^{1.5}$ applied to it (and used in the inversion). On the right side of the section, the top salt reflection between 2.2 and 2.4 s as well as the bottom salt reflection between 2.6 and 2.8 s are visible. Inside the salt, the marine bubbles (120 and 240 ms after the top salt reflection) are also quite strong. Deconvolution

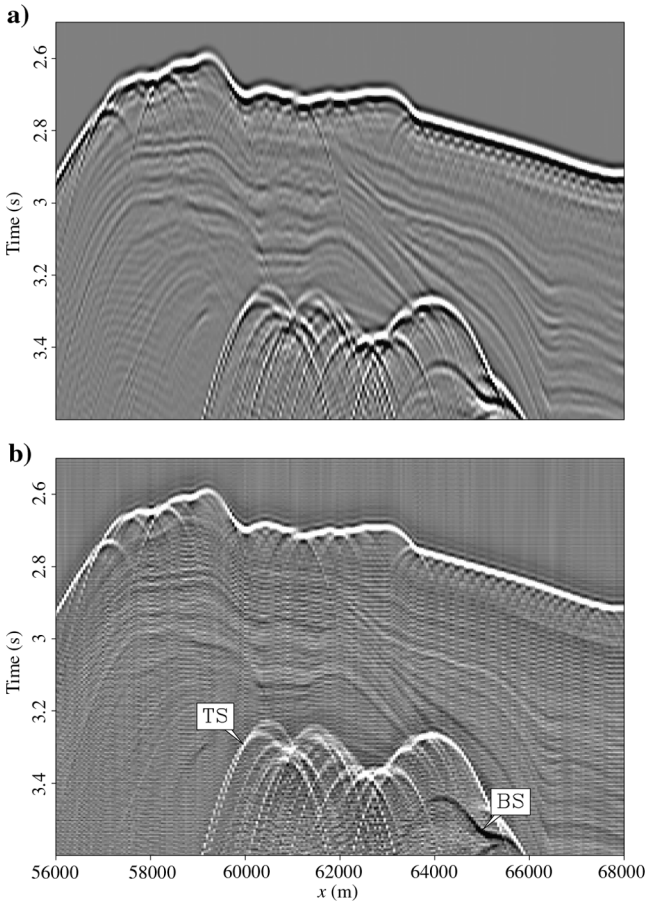

Figure 1. A zero-offset section of the Chevron synthetic data set (a) before $g_t d_t$ and (b) after $q_t = g_t r_t$ deconvolution. After deconvolution, the resolution improves significantly and the polarities are cleaner. The arrows show the top (TS) and bottom salt (BS) reflections.

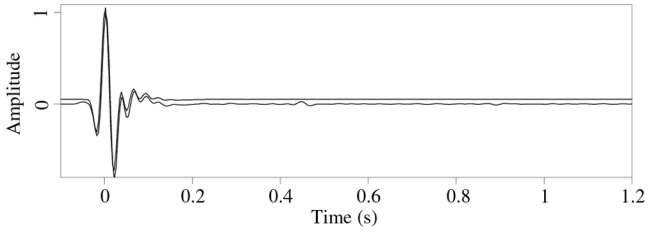

Figure 2. The total source and estimated wavelet for the data in Figure 1. The total source is amplitude shifted by 5% above the estimated wavelet for clarity. Both are remarkably close.

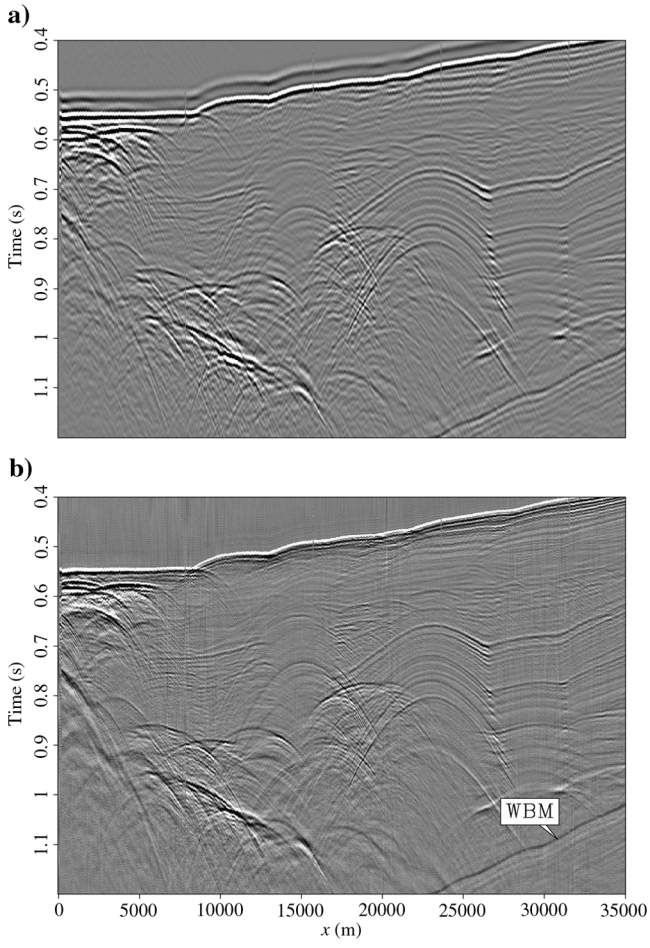

Figure 3. Near-offset section from offshore California (a) before $g_t d_t$ and (b) after $q_t = g_t r_t$ deconvolution. The source waveform is characteristic of a far-field pressure signature of a watergun source (Safar, 1985a). The first-order water-bottom multiple (black) is labeled WBM.

Downloaded 08/20/15 to 138.67.154.174. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

(Figure 5b) clarifies the polarity. The water bottom and top salt are white, whereas the bottom salt is black, showing the change of impedance. The bubbles have also been attenuated inside the salt (and throughout the section). The estimated wavelet is shown in Figure 4b: The first and second bubbles appear clearly at the expected lags. Note also that the wavelet is not exactly symmetric: The third lobe is larger than the first lobe.

Our third example takes us offshore Australia. Figure 6a shows a subset of the near-offset section to be deconvolved with a gain of $g_t = t^2$ applied to it. Figure 6b displays the deconvolution result (i.e., gained residual q_t). The water bottom shows up as a strong white reflection, and an unconformity at 1.9 s shows up as a strong black event. Thanks to deconvolution, the amount of fine detail in the shallowest part (less than 1.6 s) has increased. The estimated wavelet is shown in Figure 4c. Compared with the previous example, the marine bubble is very weak. Similar to the previous example, the wavelet is not symmetric either: In this case, the first lobe is slightly bigger than the third one.

Our final example comes from the Campos Basin, offshore Brazil. Figure 7a shows a subset of the near-offset section to be deconvolved, and Figure 7b shows the deconvolution result ($g_t = t$ in this case). Similar to our previous results, the deconvolution allows us to see polarities of seismic events more clearly. Notice the detail at $t = 3.2$ s and $x = 10,000$ m, where two small events of opposite polarities are well separated after deconvolution. The bubble is also lifted off from the section as seen 100 ms below the water bottom. The estimated wavelet in Figure 4d shows the strong bubble clearly. In this case, the wavelet is quite asymmetric with a small first lobe and two strong second and third lobes.

The synthetic and field data examples show how well our proposed inversion strategy works. Now, we present some computational aspects of this method detailing our choice of parameters.

PARAMETRIZATION OF THE INVERSE PROBLEM

Each of the two regularization terms introduces two parameters for the data analyst to determine. One parameter is physical, and the other is the strength (α or β) of the regularization. Naturally, we would like to do a better job of automating such decisions.

We iterate 400 times for all our results, and the computation of each wavelet took less than 30 s on an Intel Xeon X5650 processor. This computing time is fast enough to test different parameteriza-

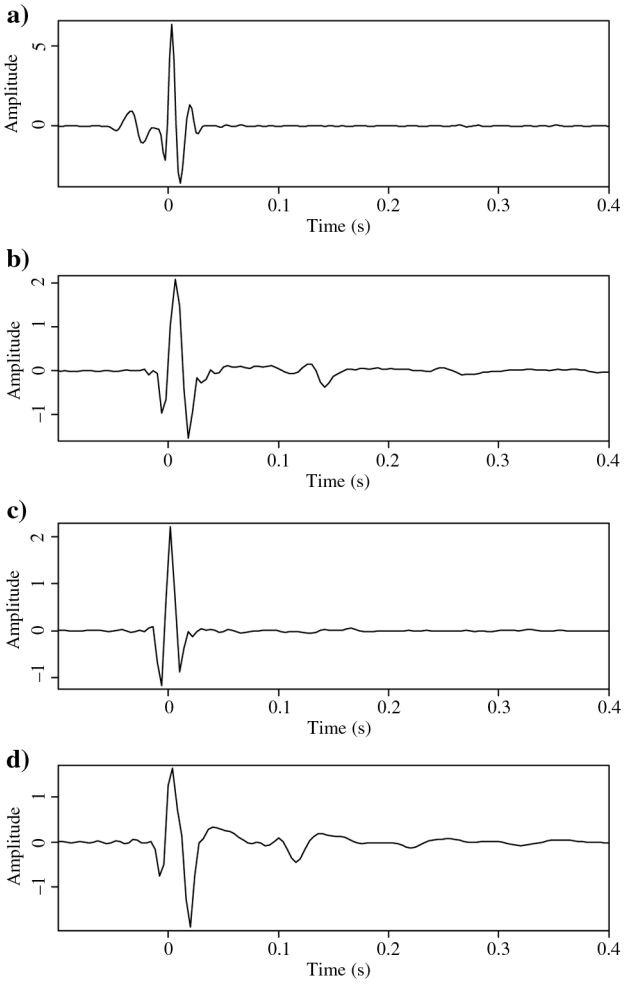

Figure 4. Estimated shot waveforms for the (a) offshore California (Figure 3), (b) Gulf of Mexico (Figure 5), (c) offshore Australia (Figure 6), and (d) Campos Basin (Figure 7) near-offset sections.

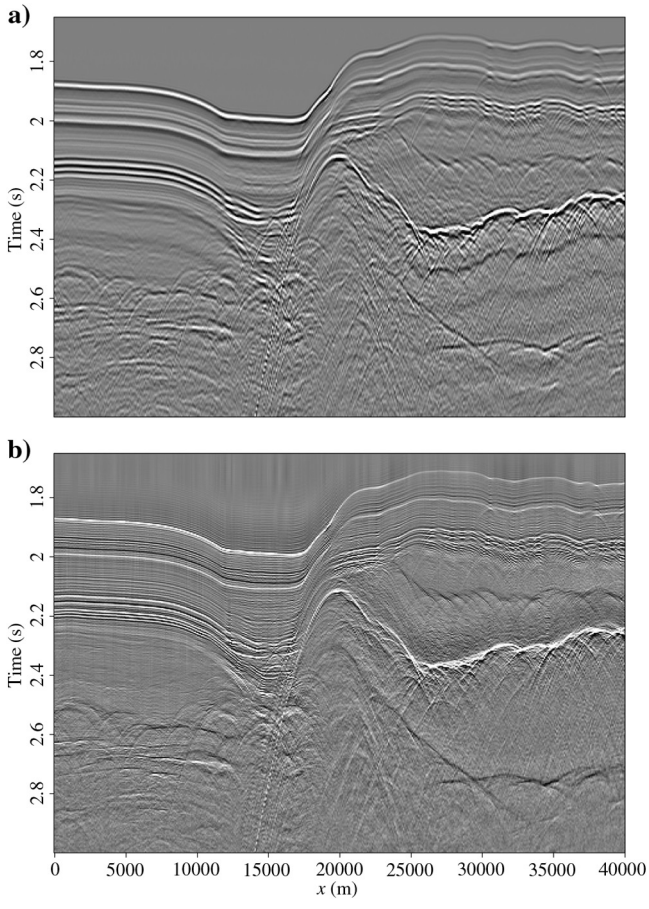

Figure 5. A near-offset section from the Gulf of Mexico (a) before $g_t d_t$ and (b) after $q_t = g_t r_t$ deconvolution. The polarity of seismic events, such as the top and bottom salt, is highlighted after deconvolution. The marine bubble is also lifted off everywhere, as seen within the salt boundaries.

Downloaded 08/20/15 to 138.67.154.174. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

tions. In our experience, the "Ricker" regularization term encouraging near zero-lag symmetry is the least needed. In all our examples, we have to use it only for the synthetic data set. Our assessment comes from looking at strong events, e.g., water-bottom reflections and seeing if the deconvolution reduces it to one phase. Ricker regularization helps with data sets in which the wavelet seems symmetrical but the deconvolution does not yield one pulse and leg jumps. As with many regularization terms in any inverse problem, these decisions are rather subjective and open to interpretation. The regularization term attenuating long and negative lags is needed often. It suppresses the tendency to "overfit" (fit a filter as long as a trace). We always start with $\alpha = 0$ in equation 14 and look at the results. This regularization is needed if some filter coefficients grow abnormally at these unwanted lags. Then, we define the regularization zone (where lags are attenuated) as being beyond the bubble (if present) and before zero lag (with a smooth transition). Having defined the regularization zone, we start increasing α keeping everything else equal until all growth stops entirely. The next parameter is the power involved in the gain function $g_t = t^{\text{pow}}$. We select the power that makes the input data the most identically distributed when multiplied with g_t, in other words, when seismic amplitudes at all times are roughly equivalent. We keep the gain function constant for all iterations. The last important parameter is the one involved in the HPF, defining the transition zone between the ℓ_1 and ℓ_2 regions. Our code defines this parameter as a percentile of the input data. We always start with the 50th percentile, which corresponds to the median of the data. Then, we change this percentile value by observing the results: More sparseness requires a smaller value, less sparseness requires a bigger value. We keep this value constant during the inversion.

Therefore, for our problem, the main parameters to set are those affecting the regularization of long and negative lags, the gain function, and the ℓ_1-ℓ_2 behavior of the HPF. The number of iterations can be set almost automatically (in all our results, 400 worked well), and the Ricker regularization is seldom needed, but you know you need it when you see a full-cycle jump or a half-cycle jump with a polarity flip. For a long while, we believed that the nonlinearity was causing the aberrant phase behavior, thus implying that a measured near-field shot waveform might be helpful to improve the estimated wavelet (Ziolkowski et al., 1982). But, we could start from a pretty good filter and watch the iteration smoothly bring it away to a leg

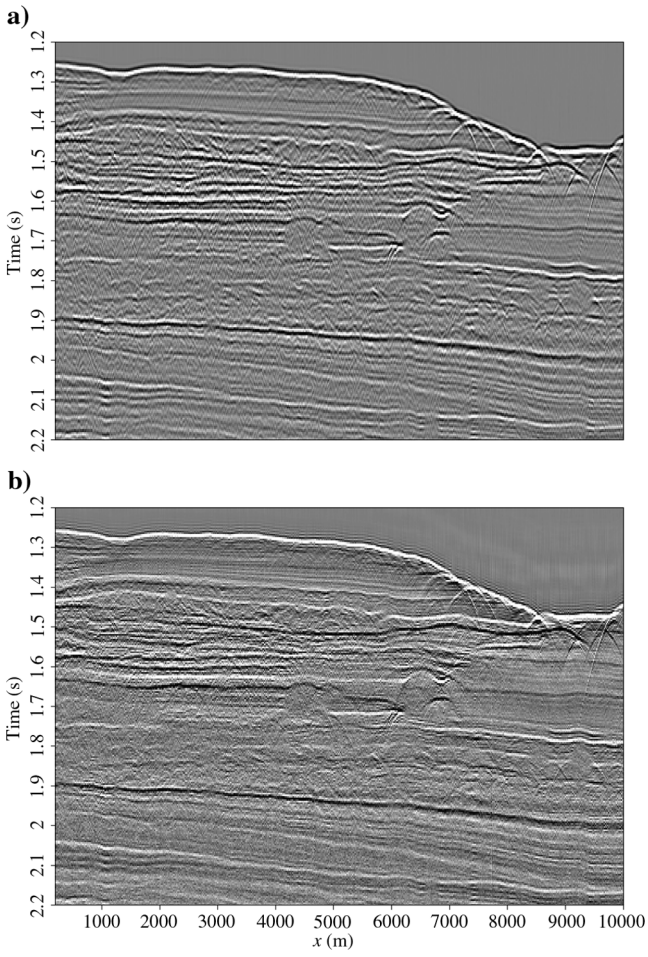

Figure 6. A near-offset section from offshore Australia. After deconvolution in panel (b), we can observe that many events are now easily labeled "white" or "black."

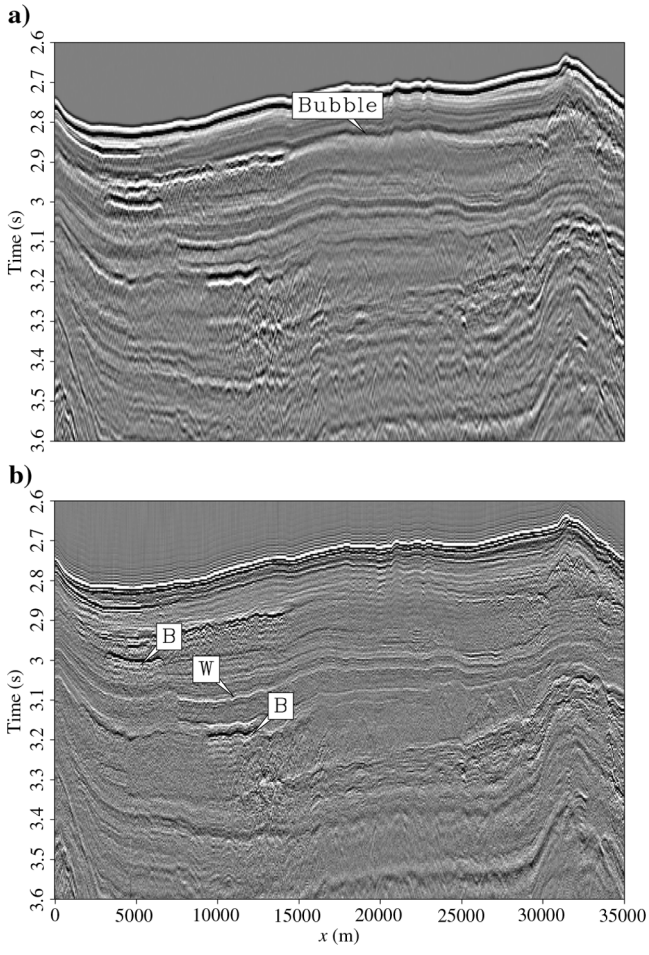

Figure 7. A near-offset section from the Campos Basin, Brazil, (a) before $g_t d_t$ and (b) after $q_t = g_t r_t$ deconvolution. Events labeled B and W (for black and white) have clear polarity after deconvolution. The source bubble is also gone.

Downloaded 08/20/15 to 138.67.154.174. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

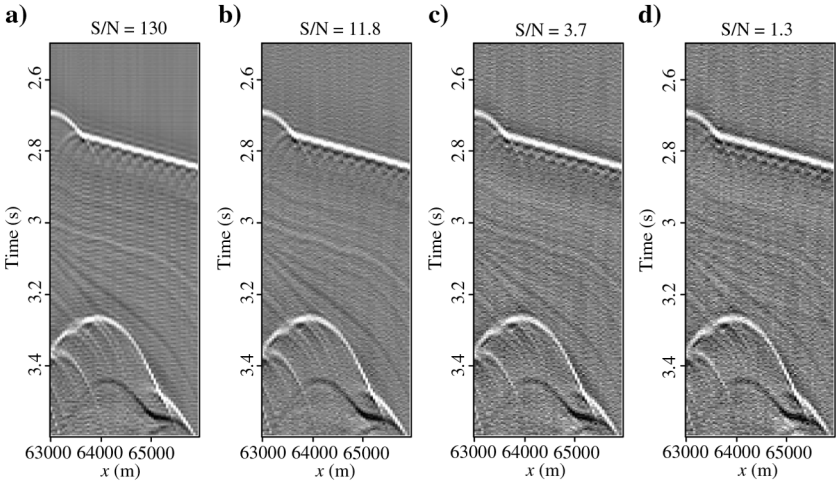

Figure 8. Subsets of the synthetic data set in Figure 1 showing deconvolution results for decreasing S/N.

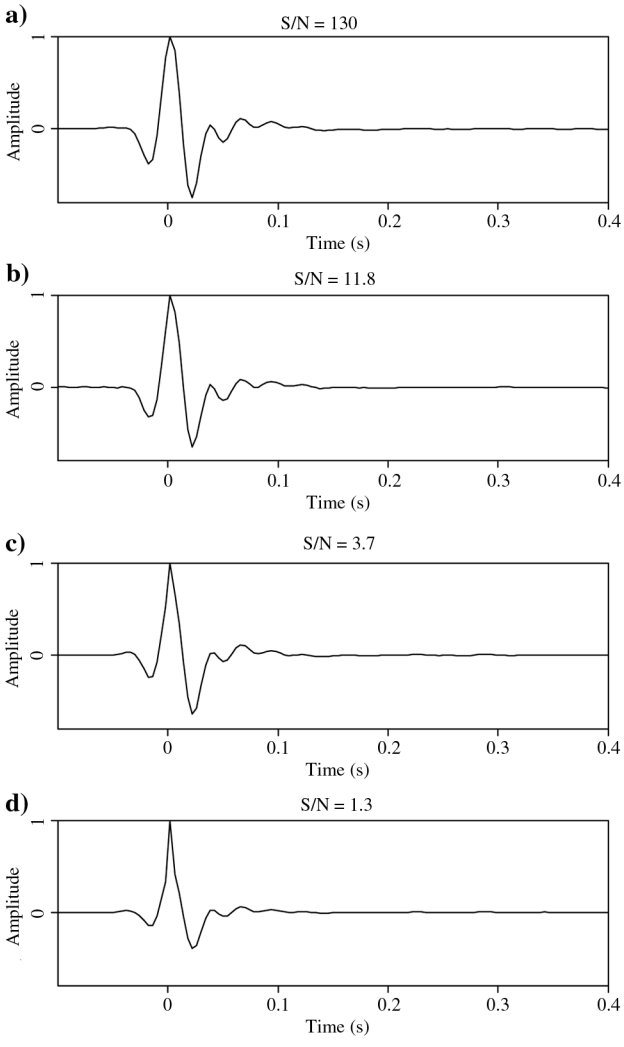

Figure 9. Estimated wavelets for the deconvolutions results of Figure 8.

jump. Therefore, we came to realize that the problem is fundamental, not computational. It results wholly from the null space introduced by the double-sided filter.

ROBUSTNESS TO NOISE AND PHASE

First, we test the robustness of our method with respect to the noise level present in the data. To do so, we add Gaussian random noise to the synthetic data set of Figure 1. Figure 8 shows the deconvolution results for four different signal-to-noise ratios (S/Ns) with the corresponding wavelets in Figure 9. Figures 8a and 9a are equivalent to a noise-free result (S/N = 130). As the S/N decreases, the noise levels in the deconvolved panels of Figure 8b–8d increase as well. The associated wavelets shown in Figure 9b–9d deteriorate accordingly: Notice how the negative lobe at $t \approx 0.2$ s shrinks with the increased noise content. However, the deconvolution results for S/N = 2.7 and S/N = 1.3 in Figure 8c and 8d, respectively, prove that our method can still deliver satisfying results by removing most of the wavelet's footprint. A more thorough investigation, similar to the work of Kazemi and Sacchi (2014), should be done to better understand the behavior of our method with respect to the S/N.

Last, we briefly test the robustness of our method with respect to phase. To do so, we rotate by 90° the Santos data in Figure 7a and estimate a wavelet from it. Then, we compare the estimated wavelet from the rotated data set with the original estimated wavelet of Figure 4d rotated by the same amount (i.e., 90°). We do not show the deconvolution results because they are nearly identical to Figure 7b. Figure 10a shows the rotated wavelet, whereas 10b shows the wavelet from the rotated data set. Both are remarkably close: This test indicates that our method delivers consistent phase estimates.

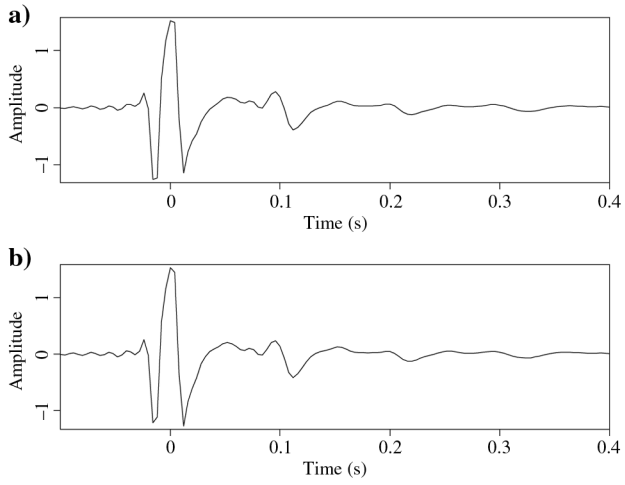

Figure 10. A phase-consistency experiment using the Santos data set. (a) The same as in Figure 4d but rotated by 90° and (b) estimated wavelet from the rotated Santos data set.

Downloaded 08/20/15 to 138.67.154.174. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

CONCLUSION

In practice, the theory works. On each of the five data sets examined, seismogram polarity becomes more apparent. The source wavelet on the watergun data set is far from the minimum phase with a large precursor, but our method nicely compresses the water bottom and its multiple to pulses of opposite polarity. Results with the Gulf of Mexico and the Australia data sets are particularly gratifying with many events acquiring clear polarities. However, the estimation of the fitting parameters remains challenging.

ACKNOWLEDGMENTS

We thank associate editor M. Sacchi and three anonymous reviewers for their comments and suggestions. J. Claerbout would like to thank Y. Zhang, Y. Shen, and Q. Fu, who worked with him during the frustrating time that we wrongly understood the instabilities to arise from nonlinearity. A. Guitton thanks Repsol Sinopec Brasil SA and Geo Imaging Soluções Tecnológicas em Geociências Ltda for permission to publish these results. We thank P. Fowler for helping us to understand the causes of the precursor seen in the California data set. We thank the sponsors of the Stanford Exploration Project, Chevron for the synthetic data set, F. Herkenhoff and Chevron Australia for the Australian-shelf data line, BP for the California data set, O. Kirstetter at Repsol Sinopec and the Agência Nacional do Petróleo, Brazil, for the Campos basin data set, and WesternGeco for the Gulf of Mexico data set.

APPENDIX A

DERIVATION OF GRADIENTS

We present the gradients of equation 6, with and without gain g_t

Gradient without gain

To better understand the gradient derivation, we first assume that there is only one unknown variable u_3 giving a single regression equation (the extension to all variables is straightforward):

$$\nabla f(\mathbf{u}) = \sum_t \frac{\partial H}{\partial r}\frac{\partial r}{\partial u_3}, \tag{A-1}$$

$$\nabla f(\mathbf{u}) = \sum_t \left(\mathrm{FT}^{-1} D(\omega)\frac{\partial}{\partial u_3} e^{U(Z)}\right)_t H'(r_t), \tag{A-2}$$

$$\nabla f(\mathbf{u}) = \sum_t (\mathrm{FT}^{-1} D(\omega) Z^3 e^{U(Z)})_t H'(r_t), \tag{A-3}$$

so the deconvolution output selected at time $t+3$ multiplies $H'(r_t)$. Now, if we had another unknown variable u_4, we would find an analogous expression this time shifted by Z^4 instead of Z^3, and so on. This shifting is nothing other than a correlation of r_t with $H'(r_t)$. Thus, the gradient for all nonzero lags is

$$\nabla f(\mathbf{u}) = \sum_t r_{t+\tau} H'(r_t), \tag{A-4}$$

where τ measures the filter lag. Actually, equation A-3 is wrong as it stands: We also need to have $\nabla f(\mathbf{u})$ at zero lag set to zero because we have no interest in uniformly scaling the filter. More simply, the mean can be removed in the Fourier domain (as shown in the pseudocode).

Gradient with gain

Now, it is the gained residual $q_t = g_t r_t$ that we are trying to sparsify. Therefore, we need its derivative by the model parameters u_τ:

$$q_t = g_t r_t = r_t g_t, \tag{A-5}$$

$$\frac{dq_t}{du_\tau} = \frac{dr_t}{du_\tau} g_t = r_{t+\tau} g_t. \tag{A-6}$$

To find the update direction at nonzero lags, take the derivative of the HPF $\sum_t H(q_t)$ by u_τ:

$$\nabla f(\mathbf{u}) = \sum_t \frac{dH(q_t)}{du_\tau} \qquad \tau \neq 0, \tag{A-7}$$

$$\nabla f(\mathbf{u}) = \sum_t \frac{dq_t}{du_\tau}\frac{dH(q_t)}{dq_t}, \tag{A-8}$$

$$\nabla f(\mathbf{u}) = \sum_t r_{t+\tau} g_t H'(q_t) \qquad \tau \neq 0. \tag{A-9}$$

This last equation says to crosscorrelate the physical residual r_t with the statistical residual $g_t H'(q_t)$.

REFERENCES

Chapman, N. R., and I. Barrodale, 1983, Deconvolution of marine seismic data using the ℓ_1 norm: Geophysical Journal of the Royal Astronomical Society, **72**, 93–100, doi: 10.1111/j.1365-246X.1983.tb02806.x.

Claerbout, J., and S. Fomel, 2014, Geophysical image estimating by example, www.lulu.com, accessed 18 April 2015.

Claerbout, J., and A. Guitton, 2014, Ricker-compliant deconvolution: Geophysical Prospecting, **63**, 615–625, doi: 10.1111/1365-2478.12213.

Fowler, P. J., 1988, Seismic velocity estimation using prestack time migration: Ph.D. thesis, Stanford University.

Jin, D., and J. Rogers, 1983, Homomorphic deconvolution: Geophysics, **48**, 1014–1016, doi: 10.1190/1.1441512.

Kazemi, N., and M. D. Sacchi, 2014, Sparse multichannel blind deconvolution: Geophysics, **79**, no. 5, V143–V152, doi: 10.1190/geo2013-0465.1.

Li, Y., Y. Zhang, and J. Claerbout, 2012, Hyperbolic estimation of sparse models from erratic data: Geophysics, **77**, no. 1, V1–V9, doi: 10.1190/geo2011-0099.1.

Liu, D., and J. Nocedal, 1989, On the limited memory BFGS method for large scale optimization: Mathematical Programming, **45**, 503–528, doi: 10.1007/BF01589116.

Ricker, N., 1953, The form and laws of propagation of seismic wavelets: Geophysics, **18**, 10–40, doi: 10.1190/1.1437843.

Sacchi, M., and T. Ulrych, 2000, Nonminimum-phase wavelet estimation using higher order statistics: The Leading Edge, **19**, 80–83, doi: 10.1190/1.1438466.

Safar, M., 1985a, On the calibration of water gun pressure signature: Geophysical Prospecting, **33**, 97–109, doi: 10.1111/j.1365-2478.1985.tb00423.x.

Safar, M., 1985b, Single water gun far-field pressure signatures estimated from near-field measurements: Geophysics, **50**, 257–261, doi: 10.1190/1.1441916.

Stoffa, P. L., P. Buhl, and G. M. Bryan, 1974, The application of homomorphic deconvolution to shallow-water marine seismology. Part I: Models: Geophysics, **39**, 401–416, doi: 10.1190/1.1440438.

Taylor, H., S. Banks, and J. McCoy, 1979, Deconvolution with the *l*1 norm: Geophysics, **44**, 39–52, doi: 10.1190/1.1440921.

Zhang, Y., A. Guitton, and J. Claerbout, 2011, A new bi-directional sparse/spike deconvolution method that overcomes the minimum phase assumption: 73rd Annual International Conference and Exhibition, EAGE, Extended Abstracts, F001.

Ziolkowski, A., G. Parkes, L. Hatton, and T. Haugland, 1982, The signature of an air gun array: Computation from near-field measurements including interactions: Geophysics, **47**, 1413–1421, doi: 10.1190/1.1441289.