Stanford Exploration Project, Report 112, November 11, 2002, pages 181–195

Simultaneous estimation of two slopes from seismic data, applied to signal/noise separation

Morgan Brown[1]

ABSTRACT

I present an efficient new approach to simultaneously estimate two slopes from seismic data. I employ a Newton iteration to overcome the problem's nonlinearity. In spite of my method's theoretical inability to handle aliased data, it robustly estimates two independent slopes in many circumstances. I apply my method to the problem of signal/noise separation on synthetic and real data examples. The estimated slopes provide approximate inverse signal and noise covariance operators good enough to obtain an excellent separation, with only a limited amount of prior information required.

BACKGROUND

For practical purposes, seismic data consists locally of the superposition of n plane waves. Disregarding aliasing effects and the data's wavelet, the slopes of the n plane waves fully and uniquely parameterize the data locally. Claerbout (1992)[2] casts the problem of single-slope estimation as a linear, univariate optimization problem. Fomel (2000; 2001b; 2001a) extends the problem to the estimation of two slopes, and utilizes the estimated slopes for the interpolation of missing data and signal/noise separation. He iteratively solves a linearization of a nonlinear problem, and applies a model regularization term to enforce smoothness of the estimated slopes.

In this paper, I present another method for solving the two-slope estimation problem. It is a nonlinear extension of Claerbout's methodology, and differs from Fomel's in the sense that it is a strictly local method. Because mine is a local method, it runs much faster than Fomel's. Theoretically, my method is sensitive to aliased data, unlike Fomel's. Like Fomel's, the estimated slope depends on the starting guess. The existence of local minima appears to be an inherent weakness of the two-slope estimation problem in general.

Fomel successfully applies "plane-wave destructor" filters, derived from estimated slopes, to the signal/noise separation problem. Analogously, I use the estimated slopes to construct "steering filters" of a form derived by Clapp et al. (1997). Like Fomel, I find that when the signal and noise slopes are too similar, my method converges to (incorrect) local minima, unless the slope estimation is "guided" with a prior model of the signal or the noise. Using

[1]**email:** morgan@sep.stanford.edu

[2]section 4.5, *Dip Picking Without Dip Scanning*

this constrained approach, I obtain excellent separation results on three different real data examples. Most encouragingly, in all cases, very simple, easily-obtained prior models sufficed.

THE METHOD

Claerbout (1992) casts the problem of slope estimation as a univariate optimization problem, based on the observation that the partial differential equation

$$\left(\frac{\partial}{\partial x} + p\frac{\partial}{\partial t}\right) u(t,x) \tag{1}$$

is zero-valued if the wavefield $u(t,x)$ consists only of plane waves with time slope, or "stepout," p. Claerbout also notes that a cascade of two PDEs annihilates data consisting of plane waves with two slopes, p_1 and p_2. The analog to equation (1) is:

$$\left(\frac{\partial}{\partial x} + p_1\frac{\partial}{\partial t}\right)\left(\frac{\partial}{\partial x} + p_2\frac{\partial}{\partial t}\right) u(t,x), \tag{2}$$

or after expansion,

$$\left(\frac{\partial^2}{\partial x^2} + (p_1+p_2)\frac{\partial^2}{\partial x \partial t} + p_1 p_2 \frac{\partial^2}{\partial t^2}\right) u(t,x). \tag{3}$$

Discretizing the problem

Claerbout approximates the derivatives of equation (1) with 2x2 finite difference stencils. Assuming that the grid spacing in both the t and x directions are unity:

$$\frac{\partial}{\partial x} \approx 0.5 * \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, \qquad \frac{\partial}{\partial t} \approx 0.5 * \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}. \tag{4}$$

By convolving together these first-order stencils, we can construct appropriate finite-difference stencils to approximate the second-order differential operators of equation (3):

$$\frac{\partial}{\partial x} * \frac{\partial}{\partial x} = \frac{\partial^2}{\partial x^2} \approx 0.25 * \begin{bmatrix} -1 & 2 & -1 \\ -2 & 4 & -2 \\ -1 & 2 & -1 \end{bmatrix} \tag{5}$$

$$\frac{\partial}{\partial t} * \frac{\partial}{\partial x} = \frac{\partial^2}{\partial x \partial t} \approx 0.25 * \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \tag{6}$$

$$\frac{\partial}{\partial t} * \frac{\partial}{\partial t} = \frac{\partial^2}{\partial t^2} \approx 0.25 * \begin{bmatrix} -1 & -2 & -1 \\ 2 & 4 & 2 \\ -1 & -2 & -1 \end{bmatrix} \tag{7}$$

The stencils of equations (5)-(7) are convolved with the data, $\mathbf{u}$. For simplicity, we can define the following notation:

$$\frac{\partial^2}{\partial x^2} * \mathbf{u} = \mathbf{D}_{xx}; \qquad \frac{\partial^2}{\partial x \partial t} * \mathbf{u} = \mathbf{D}_{xt}; \qquad \frac{\partial^2}{\partial t^2} * \mathbf{u} = \mathbf{D}_{tt}, \tag{8}$$

and rewrite equation (3) in matrix form:

$$\mathbf{r} = \begin{bmatrix} \mathbf{D}_{xx} & \mathbf{D}_{xt} & \mathbf{D}_{tt} \end{bmatrix} \begin{bmatrix} 1 \\ p_1 + p_2 \\ p_1 p_2 \end{bmatrix}. \tag{9}$$

The vector $\mathbf{r}$ has the same dimension as the data, $\mathbf{u}$. If the data consists only of plane waves with slopes p_1 and p_2, then equation (9) predicts values of $\mathbf{u}$ from nearby values of $\mathbf{u}$. If the data's slopes change in time and space, however, equation (9) is valid only across local "patches" of the data. We can rewrite equation (9) to reflect this fact:

$$\mathbf{r} = \begin{bmatrix} \mathbf{D}^1_{xx} & \mathbf{D}^1_{xt} & \mathbf{D}^1_{tt} \\ \vdots & \vdots & \vdots \\ \mathbf{D}^n_{xx} & \mathbf{D}^n_{xt} & \mathbf{D}^n_{tt} \end{bmatrix} \begin{bmatrix} 1 \\ p_1 + p_2 \\ p_1 p_2 \end{bmatrix} \tag{10}$$

Equation (10) denotes the convolution of the respective finite-difference stencils over a data patch of size n, where n may be as large as the entire data, or as small as 3×3

While it is tempting to make a change of variables ($a = p_1 + p_2, b = p_1 p_2$) and treat equation (10) as a linear relationship, I have found that this approach produces trivial coupled estimates of the true slopes. This problem is inherently nonlinear.

Dip Estimation

To estimate two local slopes p_1 and p_2, we treat vector $\mathbf{r}$ in equation (10) as a familiar prediction error, and find the p_1 and p_2 which minimize the squared norm of the prediction error. First we define the following shorthand:

$$\sum \mathbf{D}_{xx}\mathbf{D}_{xt} = \sum_{i=1}^{n} \mathbf{D}^i_{xx}\mathbf{D}^i_{xt}.$$

Expanding $\mathbf{r}^T\mathbf{r}$ from equation (10) and collecting terms yields a nonlinear function of p_1 and p_2, which we denote $Q(p_1, p_2)$:

$$\begin{aligned} Q(p_1,p_2) &= \sum \mathbf{D}^2_{xx} + p_1 \cdot 2 \sum \mathbf{D}_{xx}\mathbf{D}_{xt} + p_2 \cdot 2 \sum \mathbf{D}_{xx}\mathbf{D}_{xt} \\ &+ p_1 p_2 \cdot \left(2 \sum \mathbf{D}_{xx}\mathbf{D}_{tt} + 2 \sum \mathbf{D}^2_{xt} \right) + p_1^2 \cdot \sum \mathbf{D}^2_{xt} + p_1^2 p_2 \cdot 2 \sum \mathbf{D}_{xt}\mathbf{D}_{tt} \\ &+ p_2^2 \cdot \sum \mathbf{D}^2_{xt} + p_1 p_2^2 \cdot 2 \sum \mathbf{D}_{xt}\mathbf{D}_{tt} + p_1^2 p_2^2 \cdot \sum \mathbf{D}^2_{tt}. \end{aligned} \tag{11}$$

To find the least-squares-optimal p_1 and p_2, we compute the partial derivatives of $Q(p_1, p2)$, set them equal to zero, and solve a system of two equations.

$$\begin{aligned} \frac{\partial Q(p_1,p_2)}{\partial p_1} = f(p_1,p_2) &= \sum \mathbf{D}_{xx}\mathbf{D}_{xt} + p_2 \sum \mathbf{D}_{xx}\mathbf{D}_{tt} + p_2 \sum \mathbf{D}^2_{xt} \\ &+ 2 p_1 p_2 \sum \mathbf{D}_{xt}\mathbf{D}_{tt} + p_2^2 \mathbf{D}_{xt}\mathbf{D}_{tt} + p_1 p_2^2 \sum \mathbf{D}^2_{tt} = 0 \end{aligned} \tag{12}$$

$$\begin{aligned} \frac{\partial Q(p_1,p_2)}{\partial p_2} = g(p_1,p_2) &= \sum \mathbf{D}_{xx}\mathbf{D}_{xt} + p_1 \sum \mathbf{D}_{xx}\mathbf{D}_{tt} + p_1 \sum \mathbf{D}^2_{xt} \\ &+ 2 p_1 p_2 \sum \mathbf{D}_{xt}\mathbf{D}_{tt} + p_1^2 \mathbf{D}_{xt}\mathbf{D}_{tt} + p_1^2 p_2 \sum \mathbf{D}^2_{tt} = 0 \end{aligned} \tag{13}$$

I use Newton's method for two variables to compute the optimal slopes by updating estimates of p_1 and p_2 with the following iteration:

$$p_{1,k+1} = p_{1,k} + \frac{-f(p_{1,k},p_{2,k})g_{p_2}(p_{1,k},p_{2,k}) + f_{p_2}(p_{1,k},p_{2,k})g(p_{1,k},p_{2,k})}{f_{p_1}(p_{1,k},p_{2,k})g_{p_2}(p_{1,k},p_{2,k}) - f_{p_2}(p_{1,k},p_{2,k})g_{p_1}(p_{1,k},p_{2,k})} \quad (14)$$

$$p_{2,k+1} = p_{2,k} + \frac{-f_{p_1}(p_{1,k},p_{2,k})g(p_{1,k},p_{2,k}) + f(p_{1,k},p_{2,k})g_{p_1}(p_{1,k},p_{2,k})}{f_{p_1}(p_{1,k},p_{2,k})g_{p_2}(p_{1,k},p_{2,k}) - f_{p_2}(p_{1,k},p_{2,k})g_{p_1}(p_{1,k},p_{2,k})} \quad (15)$$

The estimated slopes at iteration k are $p_{1,k}$ and $p_{2,k}$. $f_{p_1}(p_{1,k},p_{2,k})$ is, for example, the partial derivative of $f(p_1,p_2)$ with respect to p_1. While intimidating, equations (14) and (15) result simply from the inversion of a 2-by-2 matrix of second derivatives (of $Q(p_1,p_2)$), the so-called Hessian matrix. Since, the partial derivatives of f and g are non-constant, the problem is non-quadratic, which implies that Newton's method may diverge for certain initial guesses $(p_{1,0},p_{2,0})$, and furthermore, may converge to a local minimum. In practice, however, the method converges to machine precision within 3-5 iterations.

SLOPE ESTIMATION TESTS

Figures 1 and 2 illustrate tests of the my nonlinear two-slope estimation algorithm. The "textures" (Brown, 1999) were computed by constructing nonstationary steering filters (the 9-point Lagrange filter derived by Clapp et al. (1997)) with the estimated slopes, and then using those filters to deconvolve random noise. The textures, used also by Fomel (2000; 2001b), provide a quick check of the accuracy of the estimated slopes.

In Figure 1, a simple crossing-plane-wave dataset is tested. The slope panels shown have been smoothed with a sliding weighted mean filter (4-by-4 analysis window). The program used to compute the slopes also computes the weights, which are either 1 or 0. If the estimated slopes at a single point in (t,x) are equal, then the result is assumed to be trivial and the weight at that point is set to 0. Otherwise, the weight is set to 1.

The textures in Figure 1 illustrate that the estimated slopes are not totally accurate. The steep positive slope in particular seems smaller than the true positive slopes in the data, while the shallower negative slope seems better represented.

Figure 2 illustrates a more difficult test dataset, a "CMP gather" overlain by upward-sloping linear "noise." Notice that some regions of the data contain either one signal or the other. My slope estimation program, through the use of mask operators, allows the user to specify regions where only one slope is present in the data. In those regions, I use Claerbout's (1992) univariate "puck" method to estimate that single slope.

We notice some discontinuity in the textures in Figure 2 at one-slope/two-slope boundaries in the data. My two-slope algorithm slightly underestimates the positive "noise" slope, while in some sections of the data, it overestimates the magnitude of the "CMP gather" slope. Still, the general trend of both slopes honors those present in the data.

At the right and bottom edges of the estimated slopes in Figure 2, notice the constant-valued regions. Because the finite-difference templates of equation (7) run off the right and

bottom edges of the data, the slope cannot (easily) be computed in these regions. In this case, the slope remains unchanged from the starting guesses, which in this case were -0.5 and 0.5. We expect the estimated slopes to exhibit some sensitivity to starting guess. I have experimented qualitatitvely, and indeed found some sensitivity, though it is not generally severe.

SIGNAL/NOISE SEPARATION

Consider the recorded data to be the simple superposition of signal and noise events: $\mathbf{d} = \mathbf{s} + \mathbf{n}$. The so-called *Wiener estimator* is a filter, which when applied to the data, produces an optimal (least-squares sense) estimate of the embedded signal (Castleman, 1996). For the special case of uncorrelated signal and noise, the frequency response of this filter is

$$\mathbf{H} = \frac{\mathbf{P_s}}{\mathbf{P_n} + \mathbf{P_s}}, \tag{16}$$

where $\mathbf{P_s}$ and $\mathbf{P_n}$ are the signal and noise power spectra, respectively. Define operators $\mathbf{N}$ and $\mathbf{S}$, as convolution with filters which decorrelate the unknown noise $\mathbf{n}$ and signal $\mathbf{s}$, respectively. Brown and Clapp (2000), for example, show that the following least-squares optimization problem is approximately equivalent to Wiener estimation:

$$\begin{aligned} \mathbf{Nn} &\approx 0 \\ \epsilon \mathbf{Ss} &\approx 0 \\ \text{subject to} &\leftrightarrow \mathbf{d} = \mathbf{s} + \mathbf{n} \end{aligned} \tag{17}$$

Equation (17) is a regularized linear least-squares problem. The scalar parameter ϵ is related to the data's signal-to-noise ratio.

The conceptual model of seismic data as *n* locally-crossing plane waves lends itself well to parameterization by a few parameters. The multidimensional prediction-error filter (PEF) is a particularly popular option (see, for example, (Claerbout, 1998)). Estimated by autoregression against the data, the PEF encodes hidden multiplicity in the data with a few filter coefficients. It has the approximate inverse spectrum of the data from which it was estimated.

By using a model of the noise to obtain a nonstationary noise PEF and deconvolving a PEF estimated from the data by the noise PEF to obtain a signal PEF (Spitz, 1999), many authors have solved equation (17) to successfully separate coherent noise from signal (Spitz, 1999; Brown et al., 1999; Clapp and Brown, 2000; Brown and Clapp, 2000; Guitton et al., 2001).

As noted by Fomel (2000), however, the considerable amount of parameter tuning required to create stable nonstationary PEFs (a requirement for the deconvolution step) remains a significant obstacle to their use in industrial-scale processing environments.

If the signal and noise consist of distinct slopes everywhere, then it is in theory possible to implicitly separate signal from noise in the slope domain with a two-slope estimation algorithm. Fomel uses estimated slope to construct plane-wave destructor filters which are used

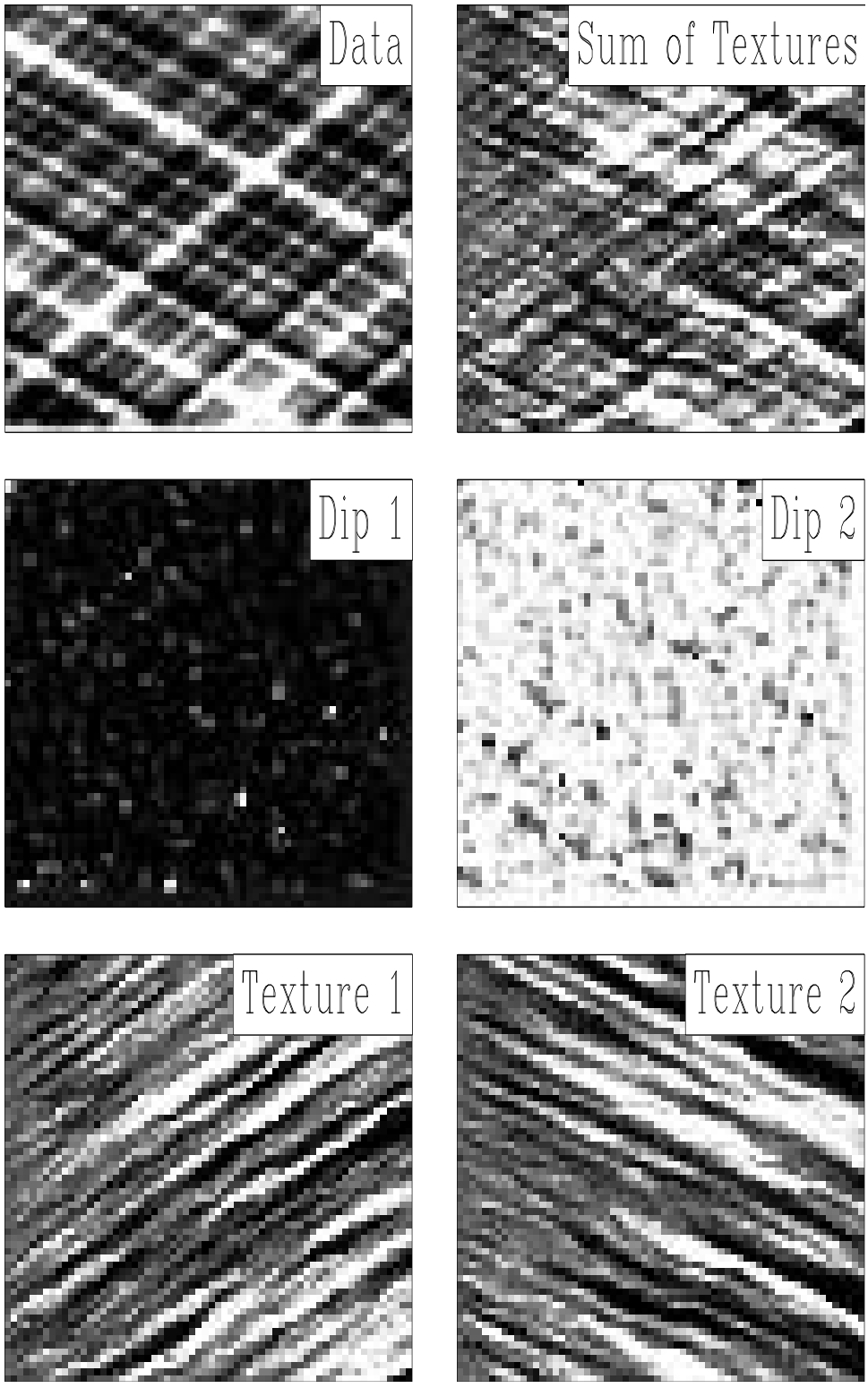

Figure 1: Test of my nonlinear two-slope estimation algorithm on a simple synthetic test case. morgan1-pucknl.dumb2 [ER,M]

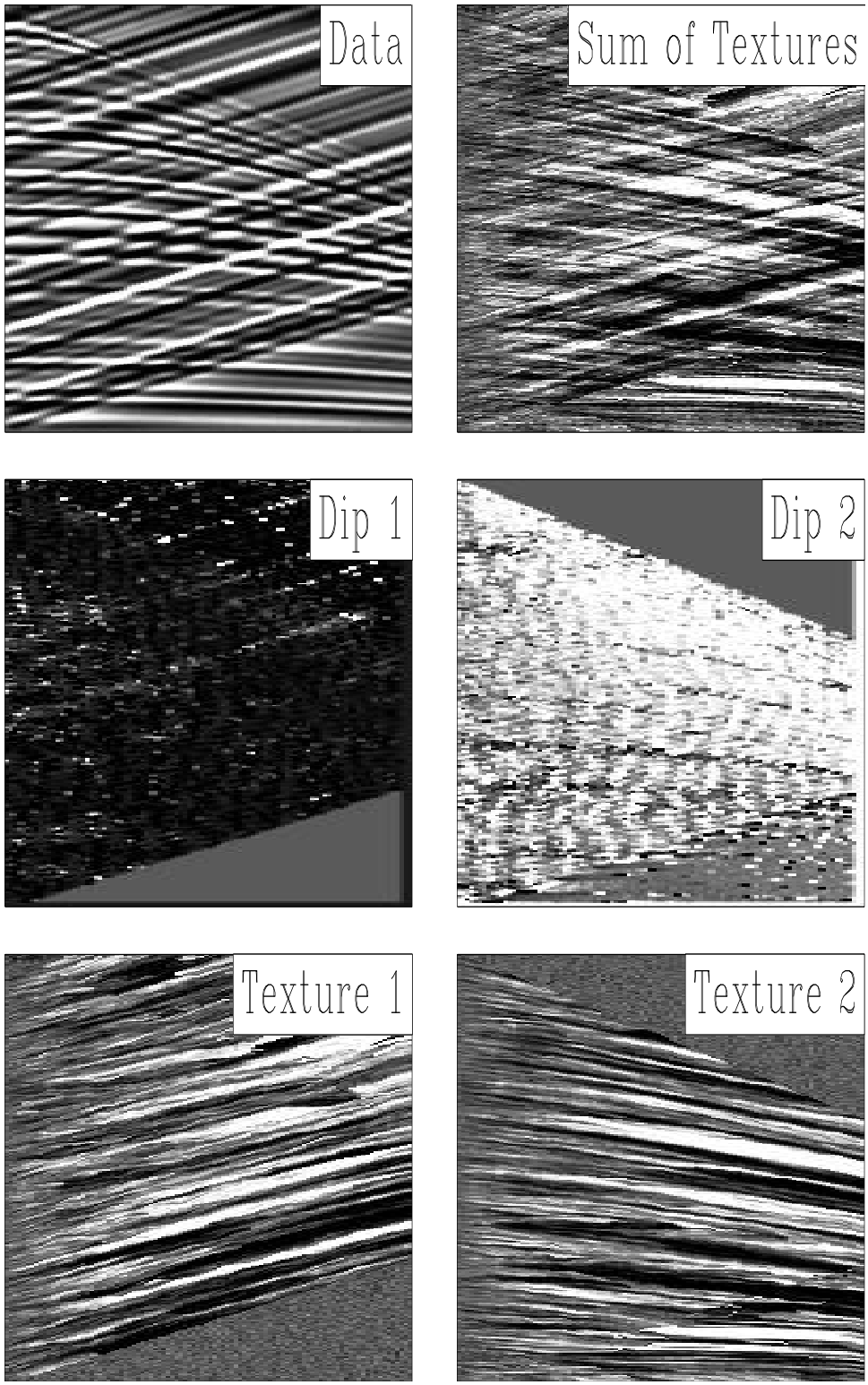

Figure 2: Test of nonlinear 2-slope "puck" algorithm on a more realistic synthetic test. morgan1-pucknl.dumb [ER,M]

directly as **N** and **S** in equation (17), without any deconvolution. The filters are guaranteed stable and insensitive to spatially aliased data. Fomel obtains an independent estimate the noise slope from a prior noise model, and then fixes the noise slope as the signal slope is estimated.

I take a slightly different tack at the problem. Like Fomel, I use my two-slope estimation technique to directly obtain signal and noise slope estimates. I also exploit a prior noise model and also a prior signal model, in cases where the signal is simpler to model than the noise. Most importantly, I find that very simple, easily-obtainable signal or noise models suffice. To overcome aliasing, I apply normal moveout (NMO) to the data. Rather than plane-wave destructor filters, I (again) use 9-point Lagrange steering filters derived by Clapp et al. (1997).

Constrained signal/noise separation results

I tested my signal/noise separation approach on three real data examples: two common-midpoint (CMP) gathers infested with multiples and a 2-D slice from a 3-D terrestrial shot gather with strong ground roll. In each case, I generate prior models of the signal and the noise (details described below), which are used in the following two-stage procedure:

1. Use single-slope estimation to obtain noise/signal slope from the prior noise/signal model.
2. Use two-slope estimation, with the slope from step 1 fixed, to estimate the other slope.
3. Initialize signal and noise steering filters and solve equation (17).

The first test, shown in Figure 3, was performed on a CMP gather from a 2-D dataset acquired by WesternGeco in the Gulf of Mexico. Characterized by strong water-bottom and top-of-salt multiple reflections, this dataset was the focus of the 1997 SEG multiple attenuation workshop. To model the signal and noise, I adopted a simple approach: given a random zero-offset section and a stacking velocity, apply inverse NMO and inverse NMO for first-order water-bottom multiples (Brown, 2002) to create a signal and noise models, respectively.

The panels on the top row of Figure 3 have been NMO-corrected with the stacking velocity, to facilitate comparison. Estimated primary reflections should be flat. We can see from the estimated signal and noise panels that my approach has produced an excellent separation result. Many totally obscured primaries now appear from beneath the multiple train. To dealias the data, I applied NMO with water velocity (4900 ft/sec). The effects of this step can be seen in the panels on the prior models and estimated dips panels of Figure 3. In this case, I fixed the *signal* slope, not the noise slope.

The method performed less impressively at near offsets, where the signal and noise dips both tend to 0. In these regions, the inversion simply "splits the difference," according to the ϵ parameter in equation (17). I conducted a second multiple separation test on CMP gather taken from the "Mobil AVO" dataset (Lumley et al., 1994). The gather is characterized by a strong train of first order water-bottom multiples, as well as strong primary events partially

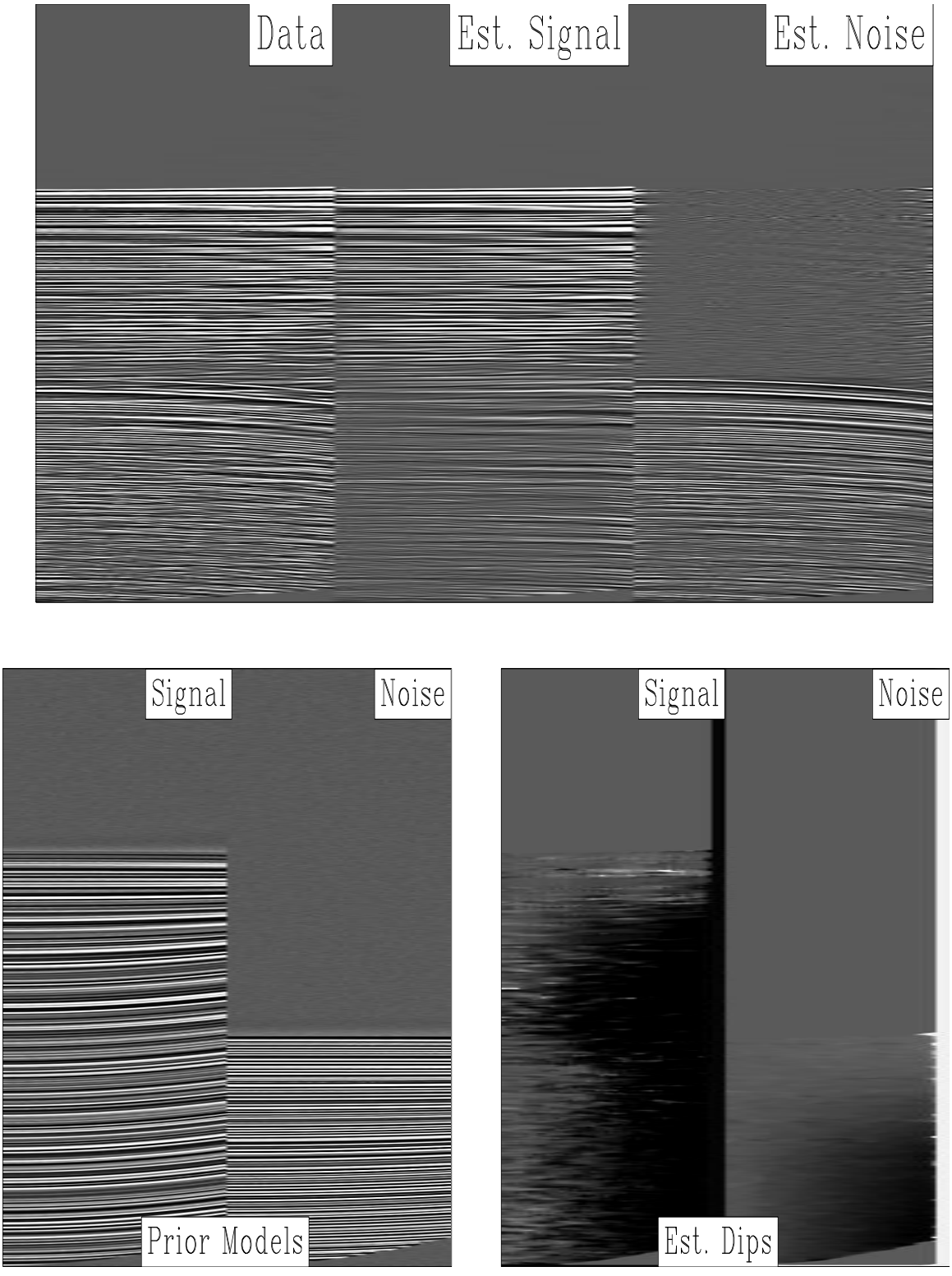

Figure 3: Signal/Noise separation tested on Gulf of Mexico CMP gather. Top row, left to right: original data, estimated signal, estimated noise. Bottom row, left to right: Prior signal and noise models, estimated signal and noise dips. morgan1-sn.gulfcmp [ER,M]

concealed under the noise. These deep primaries provide an excellent benchmark to test the signal preservation characteristics of my method.

The separation results are shown in Figure 4. As with the previous example (Figure 3), inverse NMO was used to create the prior models, and the prior signal slope was fixed in the slope estimation step. Again, the separation results are quite impressive. The embedded primary events appear to be perfectly preserved, and most of the reverberations are segregated to the noise panel. We expect the same near-offset behavior as the previous example, though the results in this region look plausible. Like before, the data were dealiased with an NMO-correction of 2.0 km/sec. My final test was conducted on a 2-D receiver line extracted from a 3-D shot gather acquired by Saudi Aramco. While the ground roll looks impossibly strong to conceal any extractable information, there is indeed no shortage of primaries under the noise cone.

The signal and noise are approximately, but not perfectly, separable in temporal frequency (Brown et al., 1999). To obtain an approximate noise model, I applied a lowpass filter with a cutoff of 10 Hz. Conversely, I obtained a signal model by applying a highpass filter with a 35 Hz cutoff. Unlike the previous two examples, I treated the noise slope as fixed in the slope estimation step. In this case, the noise is simpler than the signal. We have confidence in the noise slope; everything else is treated as signal.

While the results are not as impressive, they are good nontheless. It is not difficult to find many coherent primary events that have been unmasked from under the strong ground roll. Notice that near zero offset, some noise has leaked into the signal model. Although the data were again dealiased with an NMO correction (decreasing velocity), the noise is still spatially aliased at far offsets, a fact confirmed by a look at the estimated noise slope. Th separation results are visibly compromised in those regions. Furthermore, a persistent "ringing" is present around zero offset. Fomel (2000, 2001b,a) solved the problem by supplementing the noise decorrelation filter [**N** in equation (17)] with a 3-point notch filter.

CONCLUSIONS

I presented a nonlinear two-slope extension to Claerbout's (1992) single-slope estimation method. Compared to Fomel's approach, (2000; 2001b; 2001a), my method is less accurate and theoretically sensitive to aliasing, but also faster. I demonstrated very encouraging results in the application of signal/noise separation, where the two slopes in the data were associated to signal and noise steering filters.

REFERENCES

Brown, M., and Clapp, R., 2000, T-x domain, pattern-based ground-roll removal: 70th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 2103–2106.

Brown, M., Clapp, R. G., and Marfurt, K., 1999, Predictive signal/noise separation of groundroll-contaminated data: SEP–**102**, 111–128.

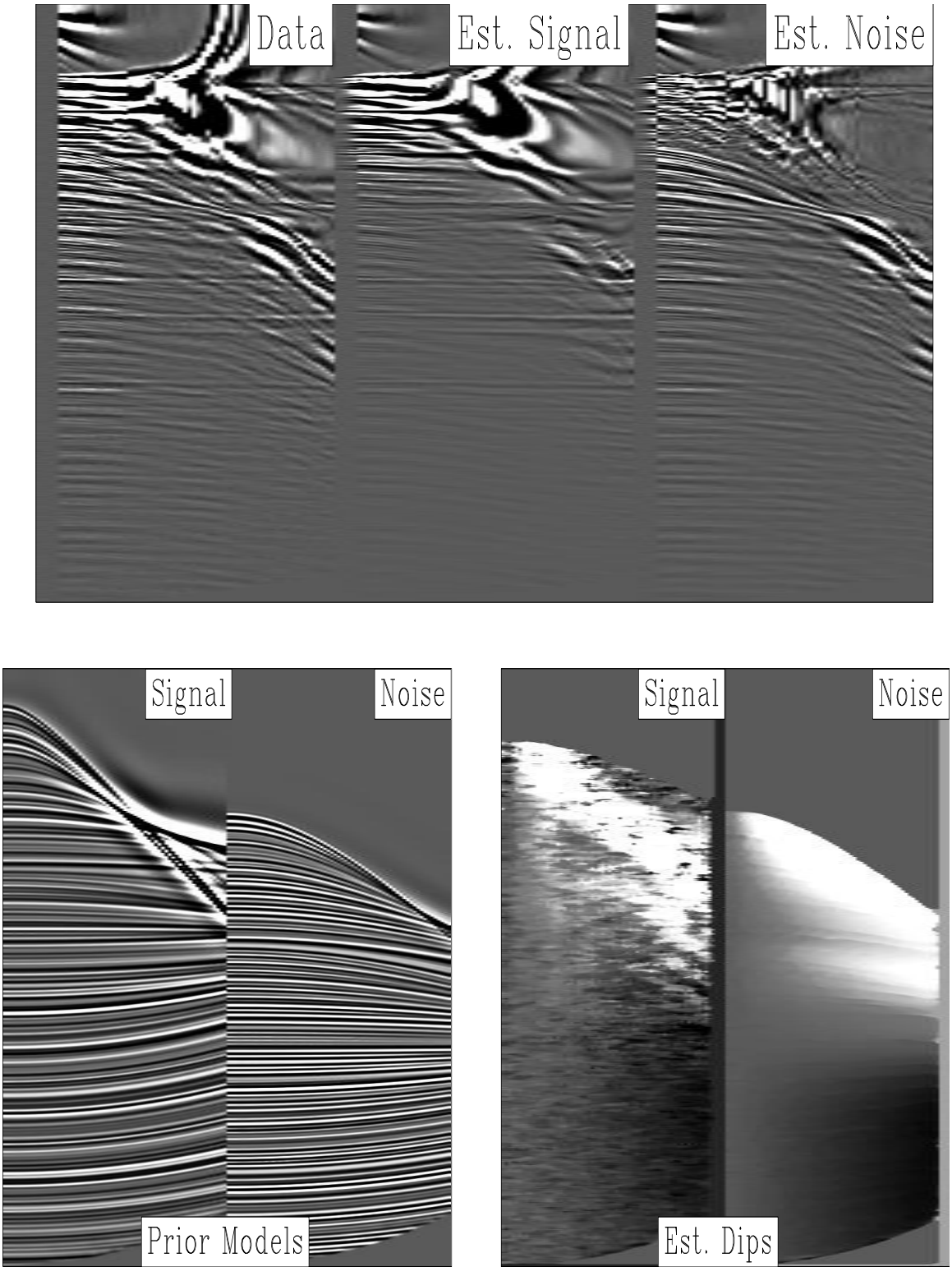

Figure 4: Signal/Noise separation tested on Mobil AVO CMP gather. Top row, left to right: original data, estimated signal, estimated noise. Bottom row, left to right: Prior signal and noise models, estimated signal and noise dips. morgan1-sn.haskreal [ER,M]

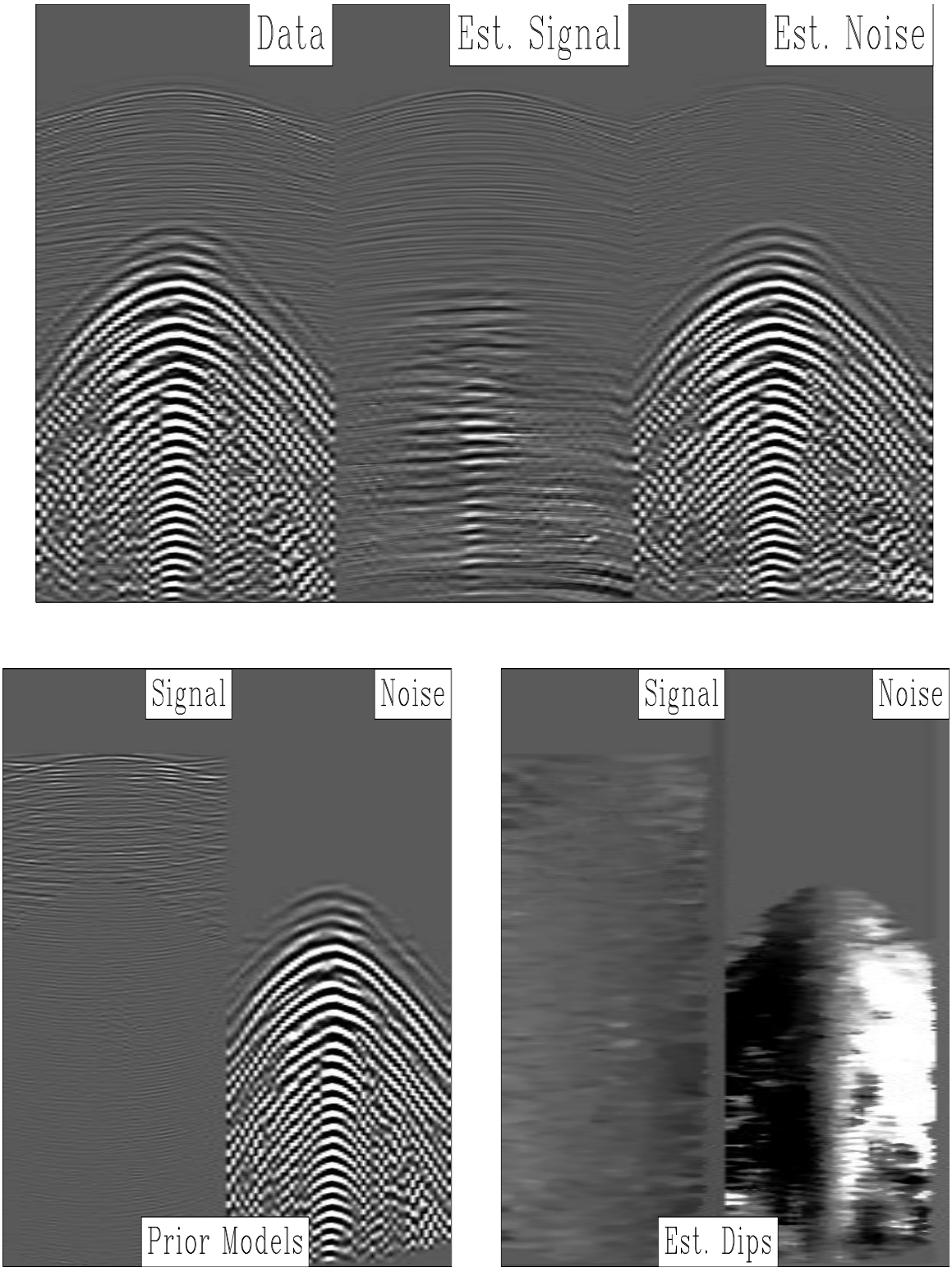

Figure 5: Signal/Noise separation tested on Saudi ground roll data. Top row, left to right: original data, estimated signal, estimated noise. Bottom row, left to right: Prior signal and noise models, estimated signal and noise dips. morgan1-sn.dune [ER,M]

Brown, M., 1999, Texture synthesis and prediction error filtering: SEP–**100**, 211–222.

Brown, M., 2002, Least-squares joint imaging of primaries and multiples: SEP–**111**, 33–47.

Castleman, K. R., 1996, Digital image processing: Prentice-Hall.

Claerbout, J. F., 1992, Earth Soundings Analysis: Processing Versus Inversion: Blackwell Scientific Publications.

Claerbout, J. Geophysical estimation by example: Environmental soundings image enhancement:. http://sepwww.stanford.edu/sep/prof/, 1998.

Clapp, R. G., and Brown, M., 2000, $(t - x)$ domain, pattern-based multiple separation: SEP–**103**, 201–210.

Clapp, R. G., Fomel, S., and Claerbout, J., 1997, Solution steering with space-variant filters: SEP–**95**, 27–42.

Fomel, S., 2000, Applications of plane-wave destructor filters: SEP–**105**, 1–26.

Fomel, S., 2001a, Seismic data interpolation and noise attenuation with plane-wave destructor filters: 71st Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1985–1988.

Fomel, S., 2001b, Three-dimensional seismic data regularization: Ph.D. thesis, Stanford University.

Guitton, A., Brown, M., Rickett, J., and Clapp, R., 2001, Multiple attenuation using a t-x pattern-based subtraction method: 71st Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1305–1308.

Lumley, D., Nichols, D., and Rekdal, T., 1994, Amplitude-preserved multiple suppression: SEP–**82**, 25–45.

Spitz, S., 1999, Pattern recognition, spatial predictability, and subtraction of multiple events: The Leading Edge, **18**, no. 1, 55–58.