Chapter 3
Separating signal and noise for reliable inversions of acoustic parameters from VSP's

3.1. SUMMARY

An inversion of a vertical seismic profile (VSP) should encourage a homogeneous acoustic-impedance function. The one-dimensional acoustic wave equation successfully models the vertical displacement of transmitted and reflected waves in a zero-offset VSP; however, impedance functions with very different statistics can model the data equally well. I prefer impedance functions having non-Gaussian first derivatives, and as few non-zero derivatives as possible. Least-squares methods of inversion encourage Gaussian statistics. If a first derivative is only slightly more probable with a non-zero value, I prefer a zero that will not distract attention from more reliable details. I treat the impedance derivatives as statistically independent parameters.

Let us model the data as a sum of two random processes: noise and (non-linearly) transformed signal—each with statistically independent parameters. A "reliable" signal parameter should not create an event in the data that can be easily described by a chance combination of noise events, and vice versa. An inversion should describe the data efficiently with a small number of reliable parameters.

From the data are estimated four signal functions: the acoustic impedance, the seismic source, the geophones' receptivity, and the geophones' depths. Linearized modeling equations and a quadratic objective function define linearized least-squares (LS) perturbations of the model parameters. Signal and noise remain additive in the perturbations, so their probability distribution functions (pdf's) convolve. Because of the statistical independence, noise becomes more Gaussian than signal in these perturbations. I take an amplitude histogram of the LS perturbations and of perturbations that treat all data as noise. A maximum-likelihood deconvolution of histograms then estimates pdf's for the signal and noise.

Using these pdf's let us calculate the Bayesian expectation of signal in the LS perturbations. I calculate the reliability of the estimates and reject impedance-derivative perturbations that contain too much noise with too great a probability. Less reliable perturbations are allowed only after the global objective function has been relinearized. The resulting simple impedance function models the data as well as do complicated functions produced by conventional gradient algorithms.

3.2. NON-UNIQUENESS OF VSP INVERSION

Let us begin with a demonstration that seismic data may poorly determine physical parameters, even when there is much redundancy. Vertical seismic profiles (VSP's), recorded as two-dimensional matrices, can correspond to one-dimensional acoustic-impedance functions that differ considerably. Because the VSP has much more redundancy than do surface reflection profiles, such news is disturbing.

Figures 3.1a and 3.1b show alternative acoustic-impedance functions; some details correlate, but on the whole they are strikingly different. Figures 3.2a and 3.2b show the corresponding synthetic VSP's. For each VSP a one-dimensional acoustic wave equation models the vertical wave displacement with depth in a stratified earth. Also used were three other one-dimensional model functions—a source waveform at the earth's surface, the depths of geophones measuring the seismic field, and the amplification of the field by individual geophones (receptivity). The two synthetic VSP's have negligible differences, as shown when the second is subtracted from the first (Figure 3.2c). The difference between the two impedance functions belongs to the non-linear analog of a "null space": it does not greatly affect the modeled section. Such VSP's need not be mathematically identical to be numerically and statistically indistinguishable.

The first impedance function (Figure 3.1a) appears to be more easily interpreted than the second (Figure 3.1b). The first is "blockier," with wide intervals of constant impedance: a simple, layered earth model can be imagined. The second appears more periodic and mono-chromatic. Few details stand out for interpretation. The low-frequency trend in Figure 3.1a is poorly determined and can be replaced by changes in other physical parameters, such as the geophones' receptivity, dispersion of waves, etc.

The derivative of impedance with depth is roughly a measure of reflectivity; this derivative better displays the qualitative difference between the two functions (Figure 3.3). The derivative of the first (Figure 3.3a) shows fewer non-zero values than does that of the second (Figure 3.3b). An interpreter, given no evidence to the contrary,

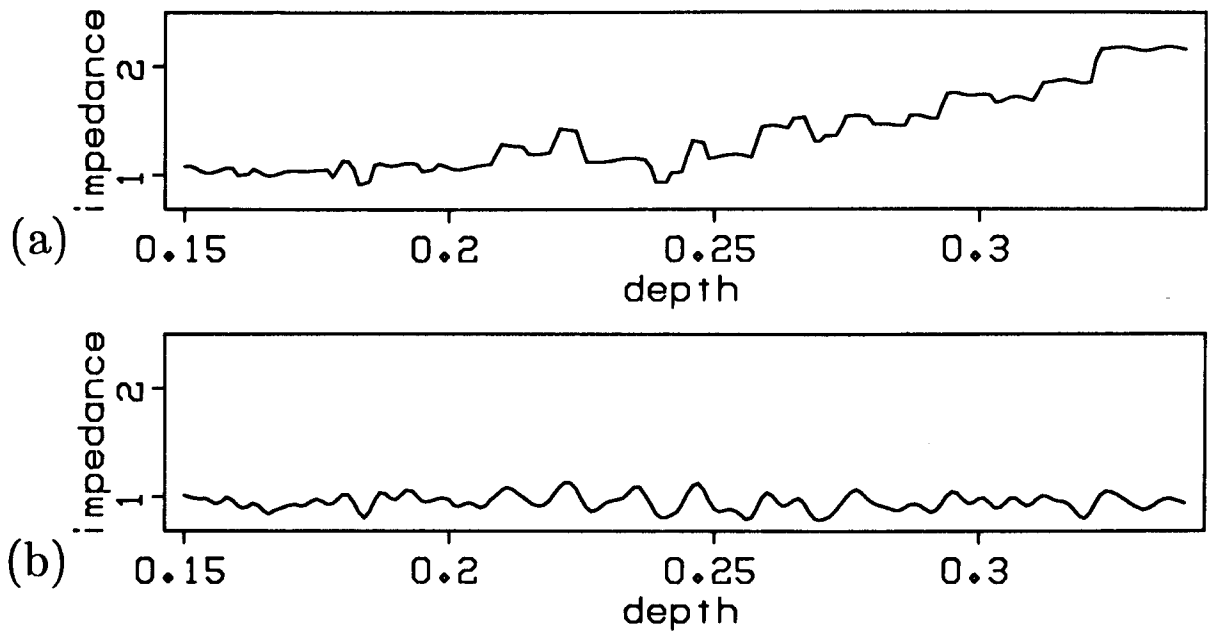

FIG. 3.1. These two acoustic-impedance functions correlate somewhat, but (a) is much easier to interpret than (b). (a) is "blockier," with wide intervals of constant impedance: a simple layered earth can be imagined. (b) appears more periodic and mono-chromatic; few details stand out for interpretation.

prefers to assume an earth as homogeneous as possible, or equivalently, with a small number of non-zero derivatives. Therefore, the first impedance function appears simpler; fewer non-zero samples in the derivative are required for creation of the VSP. The non-zero values contain all unpredictable details. Zero values could express either a lack of information or a genuinely homogeneous interval. In either case, an interpreter will not be distracted by zeros. Because the first impedance function introduced fewer non-zero derivatives to describe the same data, we are more inclined to trust and interpret the details.

The histograms of Figure 3.4 show the frequency with which various amplitudes appear in the differentiated functions of Figure 3.3. Figure 3.4a corresponds to Figure 3.3a, and Figure 3.4b to Figure 3.3b. Zero derivatives appear more often in the first histogram (Figure 3.4a) than in the second (Figure 3.4b). In the first histogram, the largest and smallest derivatives appear more often, and the intermediate derivatives less often. The first histogram has a narrower inner peak with longer "tails."

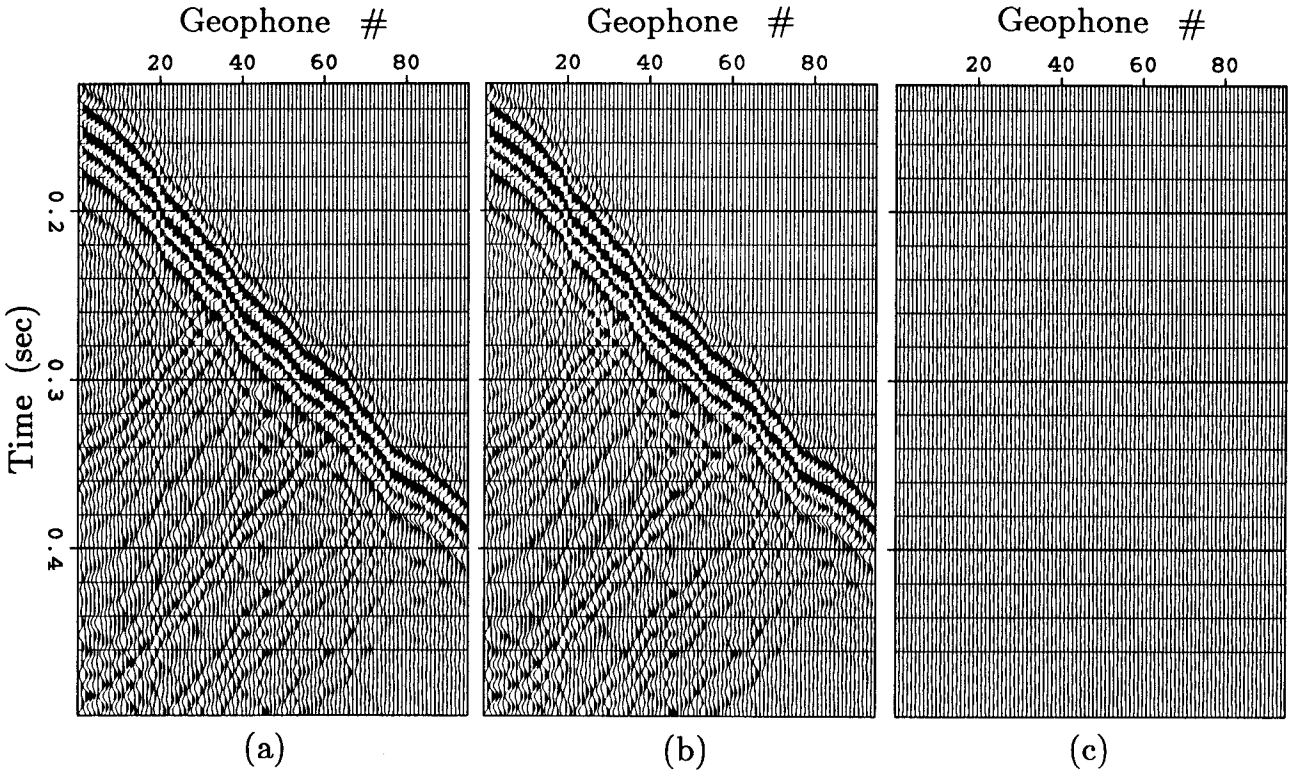

FIG. 3.2. Two very different impedance functions can create almost identical synthetic vertical seismic profiles (VSP's). The 1D acoustic wave-equation and a source function determines these two VSP's: (a) from the impedance function of Figure 3.1a, (b) from Figure 3.1b. Subtracting these two VSP's shows their negligible differences in (c).

These differences in the two histograms may be quantified by a comparison of their Gaussianity. With each histogram of Figure 3.4 are plotted two curves: the best-fitting Gaussian and the best-fitting generalized Gaussian. The generalized Gaussian has the following probability distribution function (pdf).

$$p(x) \equiv C \cdot \exp(-|x/\beta|^{\alpha}) \;\; ; \;\; C \equiv \alpha/[2\beta\Gamma(\alpha^{-1})] \tag{3.1}$$

When α is 2, the generalized Gaussian distribution is equal to the standard Gaussian distribution. Best-fitting values of α and β were chosen by a minimization of the cross-entropy of the histogram with respect to the distribution. (A similar procedure is discussed in Appendix C.) Minimizing the cross-entropy equivalently maximizes the probability of the measured histogram and gives a maximum-likelihood estimate of the pdf. The best-fitting values of α measure the Gaussianity of histograms. The first histogram gave an α of 0.26, a highly non-Gaussian fit to the numerous zeros and long

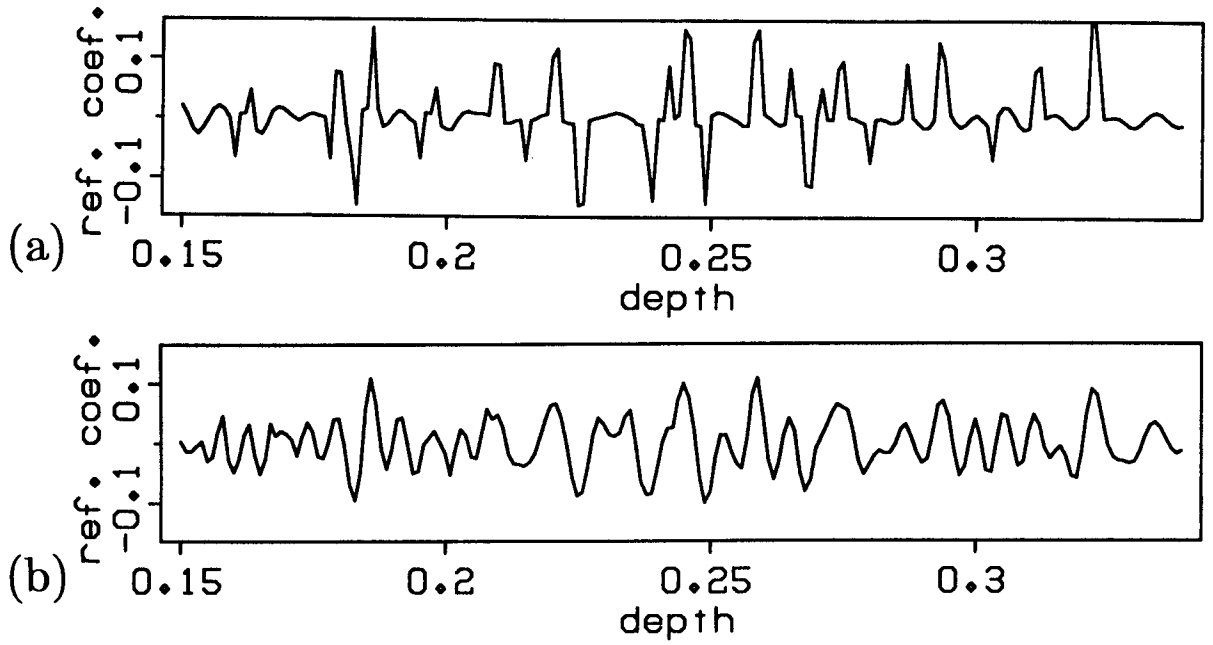

FIG. 3.3. The first derivative of impedance with depth, roughly a measure of reflectivity, shows statistical differences between the two functions better than before: (a) corresponds to Figure 3.1a, and (b) to Figure 3.1b. (a) is sparser, with a higher number of near-zero values. If free to choose, an interpreter prefers to assume an earth as homogeneous as possible. So (a) is simpler; fewer non-zero derivatives are necessary for creation of the VSP.

tails. The second histogram gave an α of 2.02, indistinguishable from the best-fitting Gaussian.

What makes one impedance function appear simpler than the other? The simpler one has fewer unexpected details—in this case, fewer large non-zero samples in the first derivative. The second impedance function (Figure 3.1b) is a least-squares inversion of a VSP and so shows L2 or Gaussian statistics. The first (Figure 3.1a) is from a more reliable inversion that I shall derive in this paper; it shows very non-Gaussian statistics. How much non-Gaussianity is desirable in inverted model parameters depends on the reliability of perturbations from a simpler structure. We must know how well these perturbations can be determined from the noisy data. Let us now use our observations to define an optimum geophysical inversion.

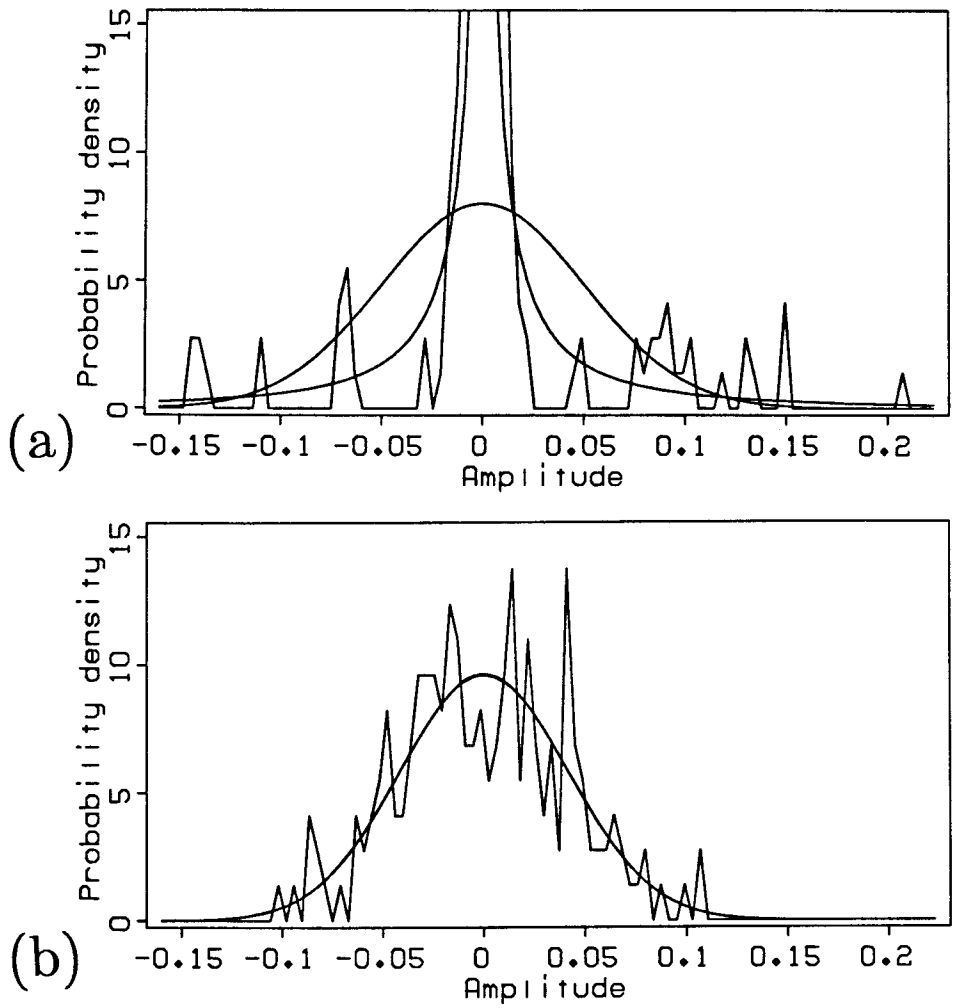

FIG. 3.4. Histograms show the frequency with which various amplitudes appear in the impedance derivatives: (a) is taken from Figure 3.3a, (b) from Figure 3.3b. I also plot the best fitting generalized Gaussian distribution, $p(x) = C \cdot \exp(-|x/\beta|^{\alpha})$, and the best fitting Gaussian, for which $\alpha = 2$. (a) is very non-Gaussian, with $\alpha = 0.26$; (b) is very Gaussian with $\alpha = 2.02$.

3.3. DEFINING AN INVERSION

Non-uniqueness makes the inversion of physical parameters an act of creation or of selection. If the result of the inversion is to be examined and interpreted by an person, not by another program, then the inversion should determine not only what details are most probable but which are most reliable. Unreliable details should not draw attention from the reliable. Interpretation of a non-unique inversion should begin with the most reliable details.

Model a data set as a sum of two random processes: noise and (non-linearly) transformed signal. Define the signal and noise parameters as statistically independent random variables. Let the "signal" be the unknown parameters we wish to estimate in our physical modeling equations. Other components are assumed to act as unpredictable, additive noise. The possible resemblance of noise to signal will be considered.

Define inversion as a choice of reliable parameters that efficiently describe the data. Let us assume that the modeling equations have been defined so that default values of parameters describe the simplest structure, least distracting to an interpreter's eye. I define an "event" in the data as the differences that appear when a parameter is changed from its default value. A reliable signal parameter should not create an event that can be easily described by a chance combination of noise events, and vice versa. Again, an inversion should describe the data efficiently with reliable parameters. Not all data need be inverted as signal or noise. An iterative inversion should converge on the most reliable parameters first because these determine the reliability of later perturbations.

3.4. CHOOSING A GLOBAL OBJECTIVE FUNCTION

I shall modify in this first section the non-linear VSP inversions of Bamberger et al. (1982) and Macé and Lailly (1984) for a least-squares objective function. The least-squares form, though still a non-quadratic function of impedance, will make the formulation of an iterative least-squares linearization much easier later. These two papers use a constraint on the L^1 norm of the differentiated impedance. An L^1 norm corresponds to an exponential pdf. The L^1 constraint (or alternatively, a penalty function) has consistently required that a great many iterations be performed before the impedance function is sufficiently simplified. Constraints on parameter perturbations will affect the appearance of the impedance functions far more than would such changes in the global objective function. The algorithm given in this paper will converge in a few iterations.

3.4.1. The differential system

As do Macé and Lailly, I choose the one-dimensional acoustic wave equation to model zero-offset VSP's. All interbed multiple reflections are modeled, but there is no thin-bed dispersion of wavelets. The effect of spherical spreading will be removed from VSP data by a scaling of amplitudes with the arrival time.

$$\sigma \frac{\partial^2 y}{\partial t^2} - \frac{\partial}{\partial x}\left[\sigma \frac{\partial y}{\partial x}\right] = 0 \tag{3.2}$$

$$\sigma \frac{\partial y}{\partial x} = -g \quad \text{for } x = 0 \tag{3.3}$$

$$y = \frac{\partial y}{\partial t} = 0 \quad \text{for } t = 0 \tag{3.4}$$

$y(x,t)$ is the time derivative of the vertical displacement measured by geophones at time t and a pseudo-depth x (see Macé and Lailly, 1984).

$$x = \int_0^z \frac{dz}{v_p(z)} \tag{3.5}$$

x is the depth measured by the travel-time from the wave source; z is the true depth; and $v_p(z)$ is the p-wave velocity as a function of depth. $\sigma(x) = v_p(x)\cdot\text{density}(x)$ is the acoustic impedance. $g(t)$ is the seismic source, the time derivative of the vertical traction (force per unit area) at the surface.

Functions are assumed to be invariant over the horizontal spatial dimensions. Assume that data are recorded at a set of depth points $\{x_i\}$ for $0 \leq t \leq T$. These depths can be irregular and sparse. I assume that the first geophone is slightly below zero depth and invert for an appropriately shifted source function.

Note that σ and g may be scaled by any two constants as long as the product of the constants is 1. I shall give σ an expected value of 1.

3.4.2. Geophone coupling and non-uniqueness

Inversions of recorded VSP's showed that the coupling and amplification of geophones can vary considerably and unpredictably. The addition of this coupling to the inversion introduced some non-uniqueness into the estimates of acoustic impedance.

The inversion of the data set used in this paper revealed that amplitudes could vary as much as 30% for equivalent events in neighboring traces. Inverted impedances were obliged first to model these strong variations non-physically and only then were allowed to model the reflected waves. An examination of reflected waves, explained only by impedance anomalies, can distinguish rapid changes in impedance and coupling.

I define, in addition to impedance and source functions, a third function for inversion. After modeling the data $y(x,t)$ from $\sigma(x)$ and $g(t)$ with equations (3.2),(3.3), and (3.4), I scale each trace by the receptivity $r(x)$.

$$y'(x,t) \equiv [1 + r(x)] \cdot y(x,t) \tag{3.6}$$

Smooth changes in impedance affect the amplitudes of transmitted waves without affecting their energy; they do not create reflected waves. For these modeling equations, smooth changes in impedance are not distinguishable from slow changes in the geophones' receptivity. Low frequencies should be subtracted from inverted impedance logs if deceptive structure is to be avoided. The depths of strong reflection coefficients are more reliable and more important than their strengths. I shall keep the low frequencies for examination.

3.4.3. A least-squares objective function

The maximum *a posteriori* (MAP) estimate finds the signal (and hence the noise) that maximizes the probability of the recorded data. The MAP estimate requires knowledge of prior probability distributions for the signal and noise.

Let us assume that we know very little about the prior distributions. If we assume that the signal and noise parameters are Gaussian and stationary, then the MAP estimate equivalently minimizes a simple least-squares objective function. To invert the data $d(x,t)$ for the model parameters $\sigma(x)$, $g(t)$, and $r(x)$, we may then minimize

$$J_1 = C_n^{-2} \sum_{x \in \{x_i\}} \int_0^T [d(x,t) - y'(x,t)]^2 \, dt$$

$$+ C_g^{-2} \int_0^T g^2 dt + C_\sigma^{-2} \int_0^X \left(\frac{d\sigma}{dx}\right)^2 dx \quad . \tag{3.7}$$

C's are the assumed standard deviations of the random variables. $y'(x,t)$ is an implicit function of the three arrays of model parameters. $(0 \leq x \leq X)$ is the interval over which $\sigma(x)$ is to be inverted.

The least-squares form does not make the minimization of objective function (3.7) by conventional gradient methods any easier: the objective function remains a non-quadratic function of the model parameters. If more were assumed or known about the prior pdf's, other MAP objective functions could serve better than this one (see equation [A.3]). Gray (1979) found that the L^1 norm is good for the optimization of reflection coefficients. An L^1 norm corresponds to a generalized Gaussian pdf with $\alpha = 1$ (see equation [A.4]). We shall see, however, that the least-squares form is best for iterative linearization.

Numerically, the first term of (3.7) minimizes the error between the measured data and the modeled data; the second and third encourage simplicity (less energy) in the inverted parameters.

In fact, $\{x_i\}$ must be treated as a fourth one-dimensional array requiring inversion. I estimate this function once in an preliminary process. The observed arrival time of the earliest downgoing wave is more reliable an estimate of the geophone depth than are field measurements converted to time units. To calculate these arrival times I window and cross-correlate the wave that arrives first, as do automatic statics programs.

3.4.4. Constraints on impedance

Differentiating impedance with depth creates a function whose samples have negligible statistical dependence. One change in impedance tells nothing to an interpreter about the location or magnitude of other changes (statistical independence). In addition few interpreters would care to prejudice their inversions with prior predictions about the overall distribution of transitions with depth (stationarity). Because the earth is made of largely homogeneous packages, large transitions are few: a majority will be small (non-Gaussianity). These observations can be summarized in statistical terms: the samples of a differentiated impedance function should be an independent, identically distributed (IID), non-Gaussian random process (an ordered sequence of random variables).

The least-squares constraint given in objective function (3.7) treats $d\,\sigma(x\,)/dx$ as IID but wrongly assumes that the samples are Gaussian. Real impedance well logs are strongly non-Gaussian. Accordingly, an inversion with this objective function increases the probability of the impedances by an exchange of numerous small derivatives for sparse strong derivatives. The result is more Gaussian and difficult to interpret.

I shall not attempt then to reach the global minimum of the least-squares objective function (3.7). We found in the first section that the bottom can be too flat for a reliable minimum. Rather, to severely constrain the impedance function, I shall refuse all non-zero derivatives whose effect on the data is poorly distinguishable from noise.

3.5. RELIABLE INVERSION

I now turn to a more compact and general notation to explain the algorithm. Let the data, the recorded VSP, be a sum of noise and transformed signal.

$$d_i = f_i(\mathbf{s}) + n_i$$

$$\text{or } \mathbf{d} = \mathbf{f}(\mathbf{s}) + \mathbf{n} \tag{3.8}$$

The signal array $\mathbf{s}$ contains the three one-dimensional arrays of $d\,\sigma(x)/dx$, $g(t)$, and $r(x)$. The non-linear transform f is implicitly defined by a finite-difference solution of equations (3.2), (3.3), (3.4), and (3.6). The noise includes all data components not modeled by this system of equations; such components are Gaussian background noise, bad traces, ground-roll, dispersion of wavelet frequencies, etc.

To anticipate the strengths and shortcomings of the iterative signal and noise extractions described in later sections, I first present results showing which details can appear when parameters are perturbed reliably. Next I explain a conventional linearized least-squares method of optimization. I discuss how the statistics of signal and noise can be distinguished in linear perturbations of the model parameters. These statistics are used to recognize the reliable perturbations.

3.5.1. An example of signal extraction

Inversion of a known synthetic VSP model shows best the sort of impedance derivatives that may be reliably recognized. VSP data are non-linear functions only of the acoustic impedance, the only signal treated here. Ordinary linear least-squares methods invert Gaussian source and receptivity functions very well.

Figure 3.5a contains a synthetic VSP prepared from an impulsive source with a constant receptivity for geophones. The corresponding impedance function (Figure 3.6a) was created by the integration of strongly non-Gaussian, independent random numbers. A least-squares inversion of the data attempts only to minimize objective function (3.7) with a conventional Fletcher-Reeve's algorithm (see Luenberger [1984]). See Appendix H for the appropriate gradients. This inversion fits the data very well and has a small percentage error. To show incorrect events, I plot the fit and the error (Figures 3.5b and 3.5c) at a low clip.

The least-squares inverted impedance function (Figure 3.6b) differs drastically from the original model. Low frequencies have been lost. Strong high frequencies do not change the modeled data much, but do increase the Gaussianity of the differentiated impedance. The major impedance changes are correctly estimated, but

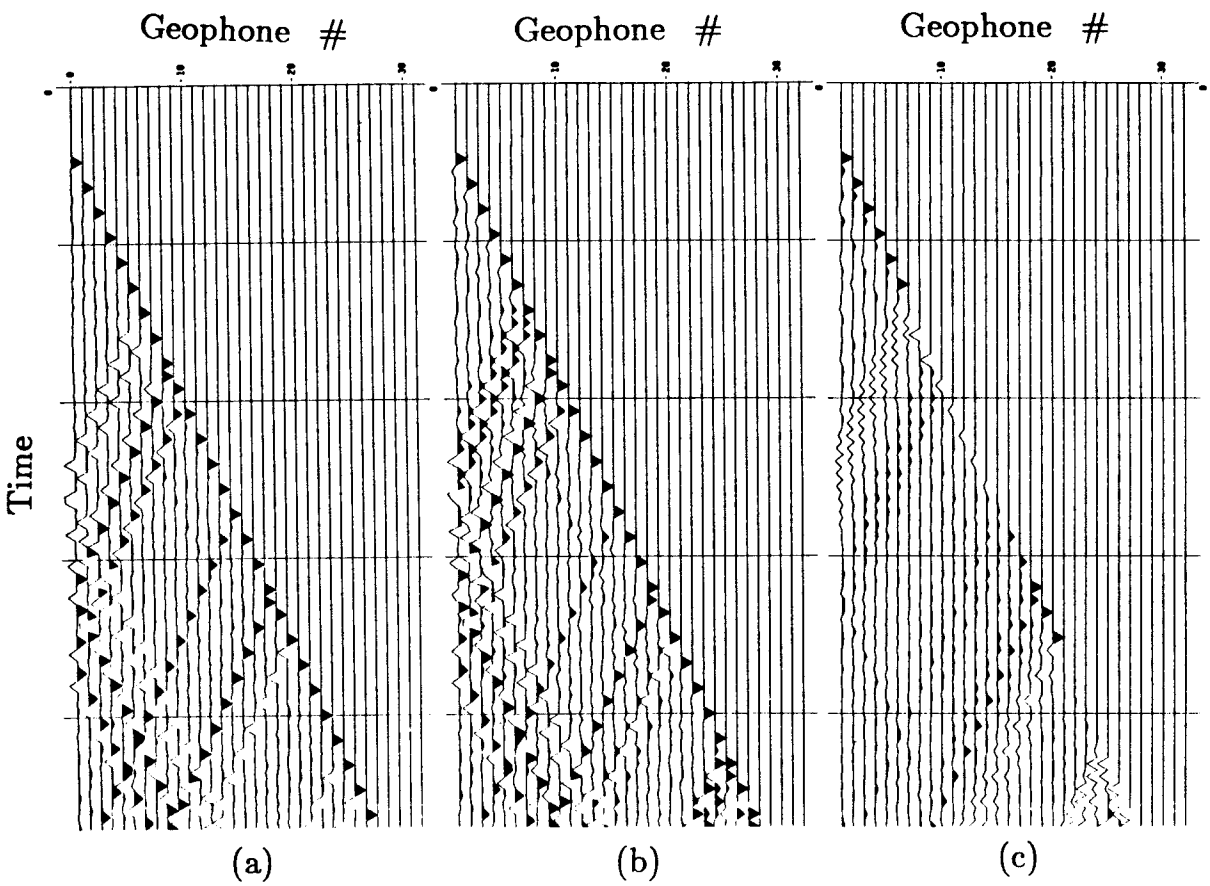

FIG. 3.5. A synthetic VSP from an impulsive source (a) is inverted by conventional least-squares methods. The VSP modeled from the inversion (b) is shown at a low amplitude clip so that the numerical artifacts are emphasized. When plotted at a more reasonable clip, the artifacts appear negligible. Subtracting the modeled VSP from the original VSP in (c) shows the residual artifacts clearly.

they are largely obscured by the unreliable details.

Figure 3.6c displays an impedance function obtained when only the most reliable perturbations of the first derivative were accepted. The modeled data are negligibly different from those in Figure 3.5b, so this is also an excellent fit to the original data. The major transitions in the impedance function are recovered very well. The smaller transitions are eliminated as too easily confused with noise. We should anticipate similar differences in statistics for the recorded VSP.

In both Figure 3.6b and 3.6c, the magnitude of impedance derivatives is approximately half the original correct magnitude. This error results from tradeoffs between the various damping terms in the objective function (3.7), particularly between the

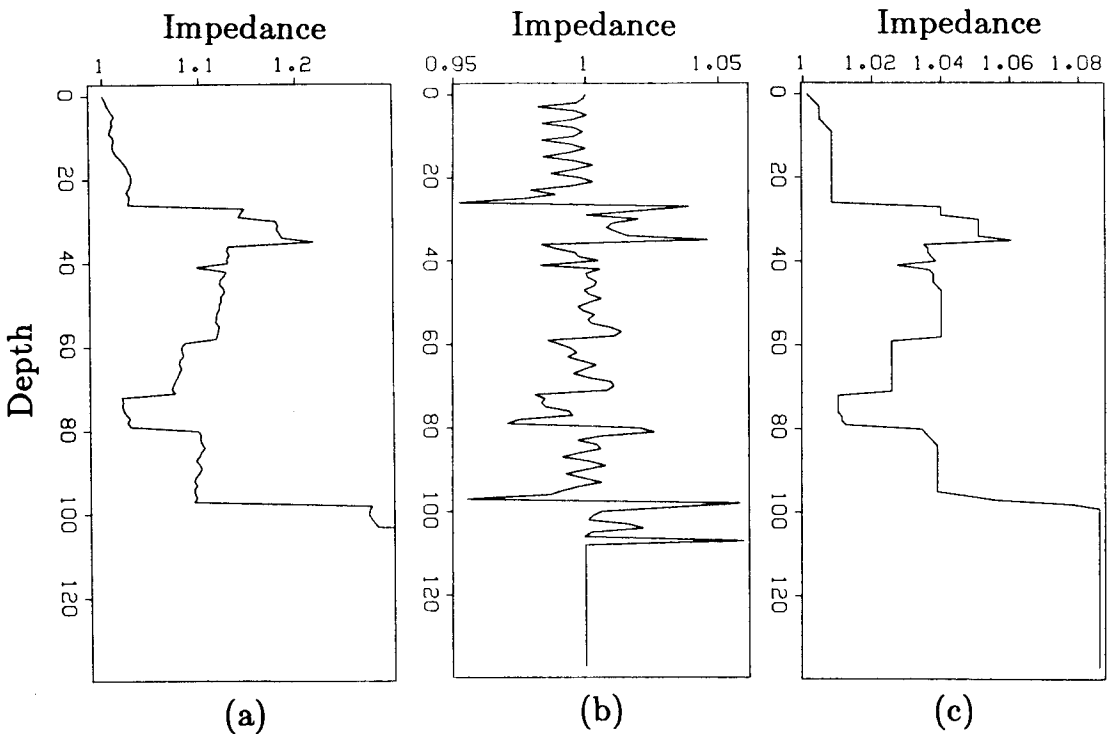

FIG. 3.6. Non-Gaussian, independent random numbers were integrated for impedance function (a), used to create the synthetic data of Figure 3.5a. Part (b) shows a conventional least-squares inversion of the impedance function; the objective function of equation (3.7) was minimized by steepest descent. The least-squares assumption encourages Gaussianity. Sparse strong derivatives are replaced by numerous weak derivatives. Important details are obscured. Part (c) shows the impedance function from an inversion that accepts only the most reliable perturbations of the first derivative, those details that cannot be explained by random alignments of noise.

source and impedance derivatives. A stronger source function can increase the strength of all modeled waves, but derivatives of impedance must change at many depths to increase equivalently the strength of all reflected waves. No effort has been made to optimize the scaling of damping terms. As a result, the impedance derivatives clearly have been overpenalized, and the source function underpenalized. The impulsive source occupied only one sample and adjusted its amplitude more freely than could the impedance derivatives. One can see from the data residuals of Figure 3.6 that the relative strengths of transmitted and reflected waves have suffered. Fortunately, we can expect damping to be more effective for a realistic source that must occupy a large number of samples.

This synthetic example, with a known solution, should guide our interpretation of inverted data. Though some degrees of non-uniqueness remain, the robust techniques recover structural details that are most valuable to the physical interpretation.

3.5.2. An example of noise extraction

Estimates of reliable signal will improve if reliable noise is iteratively estimated and removed from the data. Noise estimates, likewise, improve after estimations of signal have been made. I shall now illustrate this outermost loop of the inversion and again postpone to following sections the explanation of the non-linear estimations.

Figure 3.7 contains a portion of a VSP provided by L'Institut Français du Petrole. This section contains considerable Gaussian noise. A strong tube wave violates the physical assumptions of the modeling equations (3.2), (3.3), (3.4), and (3.6) and acts as strong additive non-Gaussian noise.

To first estimate the signal present in this data (Figure 3.8), I use a conventional Fletcher-Reeve's descent on objective function (3.7) (see Luenberger [1984]). Because noise is the object, the appearance of the impedance function is unimportant. Subtracting the modeled signal from the original data reveals the Gaussian noise, tube wave, and weak uninverted signal (Figure 3.9).

Now that reliable signal has been removed, the most reliable noise may be estimated. Using the statistical techniques explained in the following sections, I extract those samples of the residual data that contain, with a sufficient probability, a high percentage of noise. Figure 3.10 displays those samples that contain, with greater than 50% probability, more than 96% noise. Only the most non-Gaussian noise, the tube wave, has been successfully extracted. Too much of the Gaussian noise is easily explained as the sum of coherent Gaussian signal. I leave this noise behind rather than inadvertently extract too much signal. The remaining Gaussian noise is more easily ignored because Gaussian noise satisfies the assumptions of the least-squares objective functions. Subtracting the most reliable noise from the original data (Figure 3.11) removes a source of error in previous estimates of signal. I shall use this cleaner VSP in place of the recorded data in the following iterations.

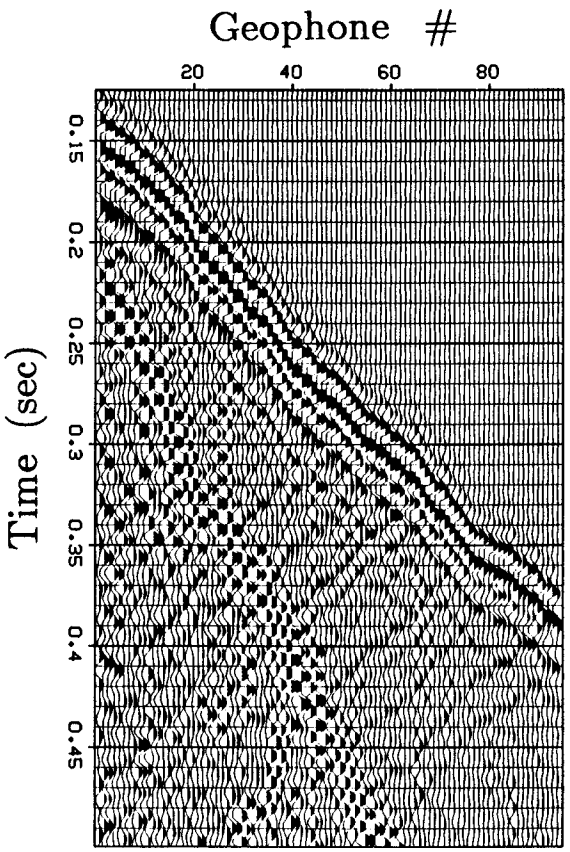

FIG. 3.7. A VSP provided by L'Instut Français du Petrole contains 96 traces. The geophones were spaced irregularly. Geophone depths were parametrized by the relative arrival times of the first recorded wave. The source function and the recorded traces were shifted so that the first geophone could be treated as if it were slightly below zero depth. A tube wave appears at about twice the slope of the first arriving wave.

3.5.3. Linearized least-squares estimation

The MAP estimate of the signal (see Appendix A) is a linear function of the data when

1. the prior distributions are assumed to be Gaussian,

2. the forward transformation is linear.

If we write a linearized transformation as

$$f_i(\mathbf{s}^0 + \Delta\mathbf{s}) \approx f_i(\mathbf{s}^0) + \sum_j F_{ij}^0 \Delta s_j \quad , \tag{3.9}$$

then the MAP estimate of $\mathbf{s}$ equivalently minimizes a quadratic least-squares objective function:

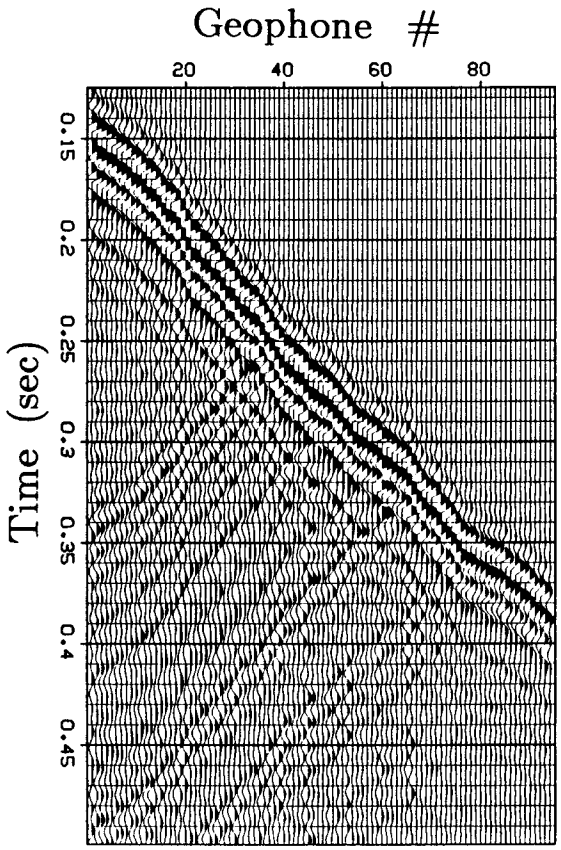

FIG. 3.8. To extract from Figure 3.7 those events easily described by the one-dimensional wave equation, I minimize the objective function of equation (3.7).

$$\min_{\Delta \mathbf{s}} \{ \sum_i C_{s_i}^{-2} (s_i{}^0 + \Delta s_i)^2 + \sum_i C_{n_i}^{-2} [d_i - f_i(\mathbf{s}^0) - \sum_j F_{ij}^0 \Delta s_j]^2 \} \quad . \tag{3.10}$$

The C's are appropriate Gaussian standard deviations for the statistically independent parameters.

The quadratic objective function (3.10) may be minimized by gradient techniques such as the conjugate gradient algorithm. Application to the acoustic equations requires the adjoint and linearized equations of (3.2), (3.3), (3.4), and (3.6) (found in Appendix H). If the global minimum of objective function (3.7) is to be found, then this linearization must be repeated.

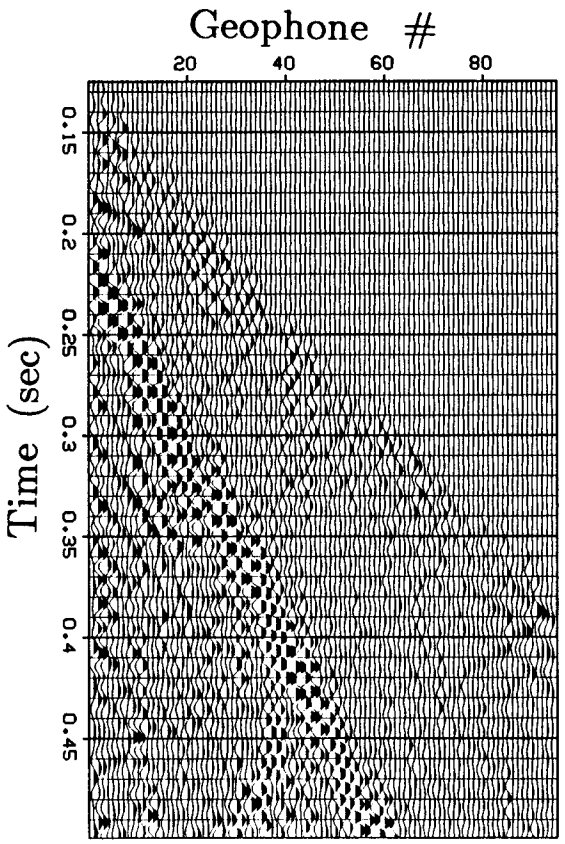

FIG. 3.9. Subtracting the signal of Figure 3.8 from the original data of Figure 3.7 reveals Gaussian noise, a tube wave, and weak uninverted signal.

3.5.4. Insufficiency of the linear estimate

When beginning a linear inversion of the VSP of Figure 3.11, I assume no knowledge of the signal parameters. Impedance is given a constant value of 1, and its derivatives a variance of 0.01—a strong constraint. The geophones' receptivity, $r(x)$, is set to zero, with a variance of 10. The variance of the noise is pessimistically set equal to that of the recorded data. The source is set to zero, and its variance to that of the recorded data.

The first minimization of the quadratic objective function (3.10) perturbs only the source function (Figure 3.12) and accounts only for downgoing waves without scattering. This function changes little in later iterations, though it is free to. The corresponding full-equation model and residuals appear in Figure 3.13. The first and higher-order scattered waves appear clearly in the residuals.

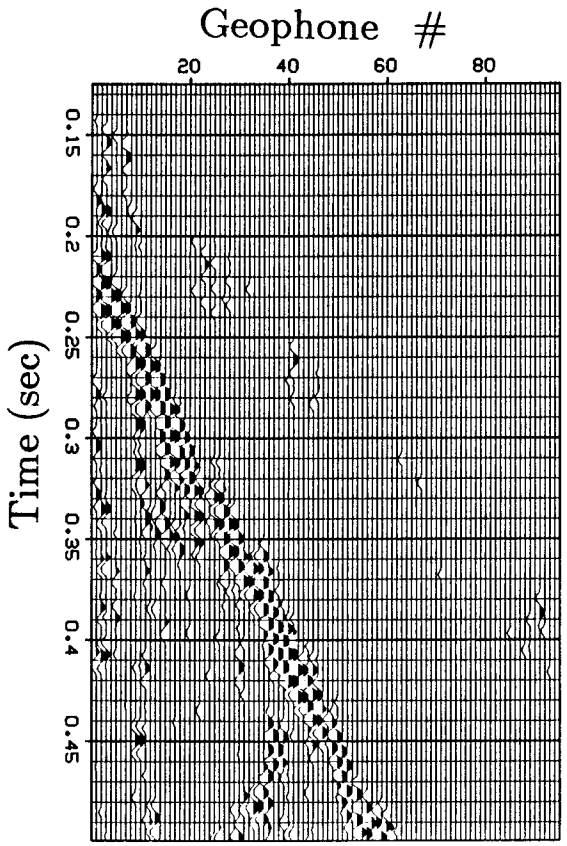

FIG. 3.10. To extract the most reliable noise from Figure 3.9, I accept those samples containing, with greater than 50% probability, at least 96% noise. Only the most non-Gaussian noise, the tube wave, has been successfully extracted. The remaining Gaussian noise is too easily described as the superposition of residual signal.

Because the background wavefield is known, the second least-squares linearization can perturb impedance and account for first-order scattering. Each linearization allows the inversion to consider an additional order of scattering. Figure 3.14a displays the new events modeled by the linearized acoustic wave equation (Equations [I.2] through [I.5]). All three signal functions receive perturbations. The perturbations of the impedance derivatives appears in Figure 3.14b, and the geophones' receptivity in Figure 3.15. The receptivity, like the source function, changes little in later iterations.

A simple test shows the effect of noise on the impedance perturbations. I randomly reorder the traces of Figure 3.13b in Figure 3.16a. This reordering does not affect the statistics of the incoherent noise. The coherent signal, however, becomes incoherent and behaves as noise in an inversion. I repeat the linearized least-squares inversion and find the perturbed impedance derivatives of Figure 3.16b. Note that the

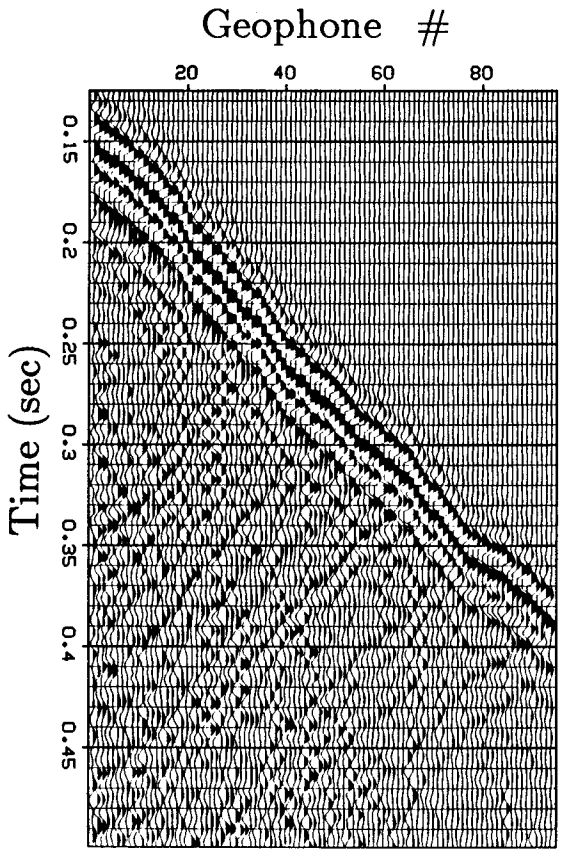

FIG. 3.11. I subtract the extracted noise of Figure 3.10 from the original data of Figure 3.7. I use this cleaner VSP in place of the original data in later estimations of signal.

perturbations are far from zero. These false details appear from random alignments of noise. We can expect this much amplitude to be unreliable in the signal perturbation of Figure 3.14b. More complicated schemes of generating noisy data sets (with the same data amplitude distributions) gave the same results. This simple exercise tells us much about the reliability of the linear perturbation.

3.5.5. Distinguishing transformed signal from noise

The advantage of the least-squares linearization is that the optimum perturbation $\Delta\mathbf{s}$ remains a linear function of the uninverted events in the data. We may write the linear perturbation $\Delta\mathbf{s}$ as the following transformation.

$$d_i' = \Delta s_i = \sum_j F_{i,j}^{-1}[d_j - f_j(\mathbf{s}^0)] \tag{3.11}$$

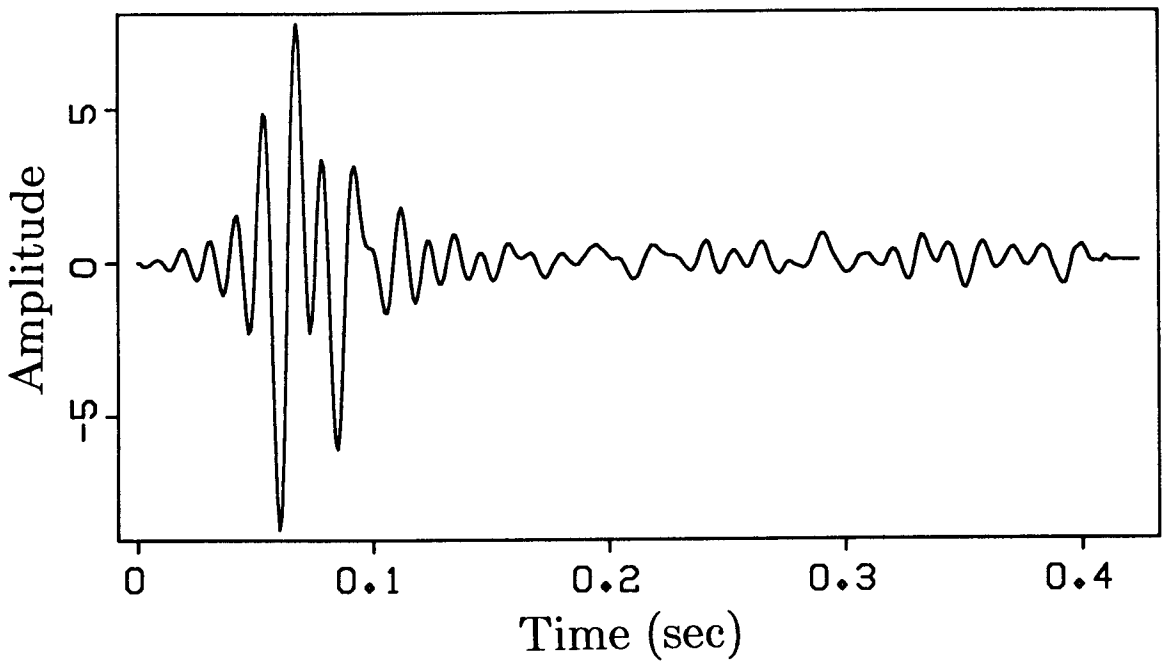

FIG. 3.12. The first minimization of the quadratic objective function (10) perturbed only the source function. The wavelet, very far from minimum phase, incorporates all reverberations and reflections between the physical source and the first geophone. This wavelet changed little in later iterations, though it was free to do so.

I suppress some additive constants because I have defined the VSP signal as having a zero mean.

Let us now make an assumption that our parameters are stationary over some dimension: many parameters should result from the same pdf. This redundancy is essential to any estimation of non-Gaussian pdf's. I now drop subscripts from individual samples. Dimensions without stationarity may be treated as silently indexed. See Appendix C for further discussion of stationarity.

The transformed signal and noise remain additive in a transformed data sample: $d' = s' + n'$. The array $\mathbf{s}'$ is strictly a function of $\mathbf{s}$, and $\mathbf{n}'$ of $\mathbf{n}$. When independent random variables add, their probability distribution functions convolve.

$$p_{d'}(x) = p_{s'}(x) * p_{n'}(x) \tag{3.12}$$

To estimate the pdf for the transformed data (Figure 3.14b) I take a histogram (Figure 3.17a). A transformation of the artificially incoherent data (Figure 3.16b) effectively overestimates the statistics of the incoherent noise; a histogram gives a pessimistic

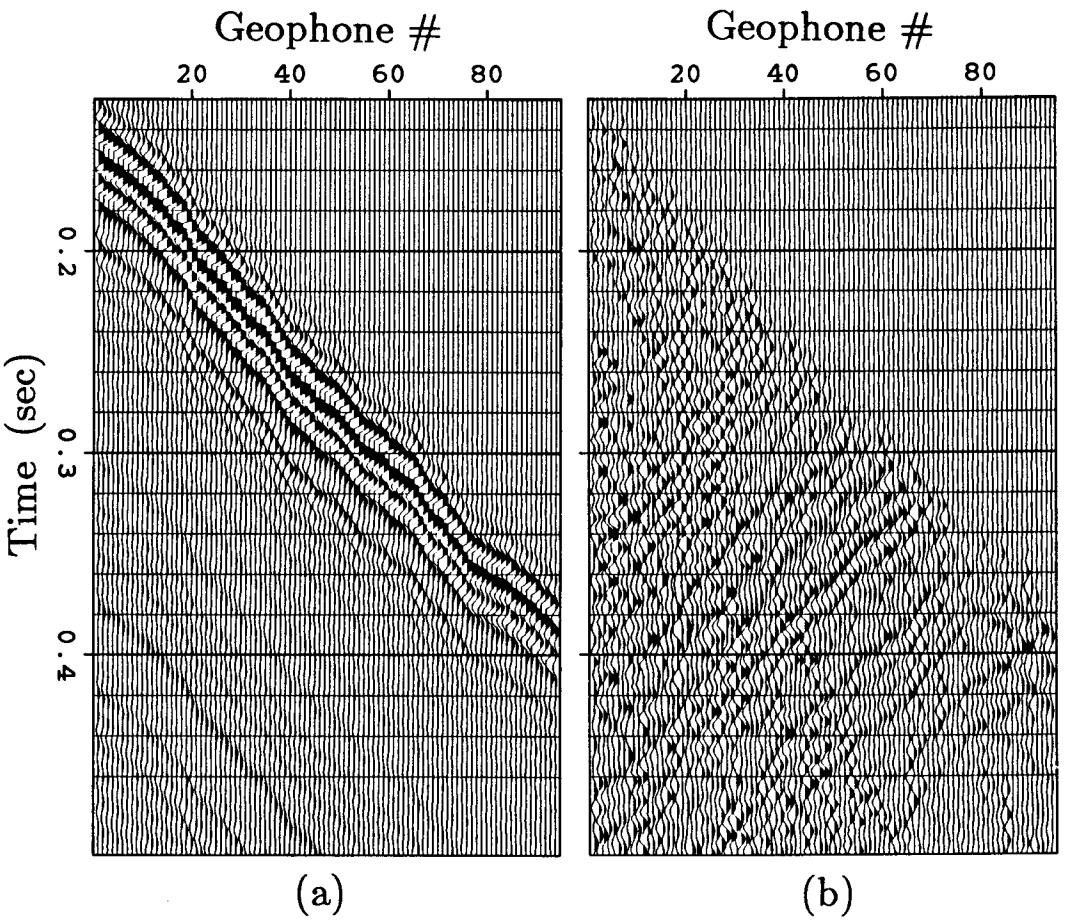

FIG. 3.13. The source function of Figure 3.12 created the modeled wave field of (a). Subtracting (a) from the data of Figure 3.11 left all scattered waves in the residuals of (b). When there is a non-zero background wavefield later linearizations are able to perturb the impedance and the geophones' receptivity functions.

estimate of the noise pdf (Figure 3.17b).

A deconvolution of equation (3.12) will estimate a pdf for the transformed signal. Assuming that the pessimistic estimate of $p_n(\cdot)$ is correct, I define the deconvolved pdf as the $p_s(\cdot)$ that maximizes the probability of the data histogram $p_d(\cdot)$. This MAP estimate equivalently minimizes the following cross entropy function, with appropriate constraints of positivity and unit area.

$$\min_{p_s'(\cdot)} \int p_{d'}(x) \ln[p_{d'}(x)/p_{s'}(x) * p_{n'}(x)]dx \tag{3.13}$$

Figure 3.17c displays the signal pdf obtained from such a deconvolution. Figure 3.17d displays the convolution $p_{s'}(\cdot) * p_{n'}(\cdot)$ of the result. This convolution cannot equal the data pdf $p_{d'}(\cdot)$ because the inner peak of the noise distribution is wider than

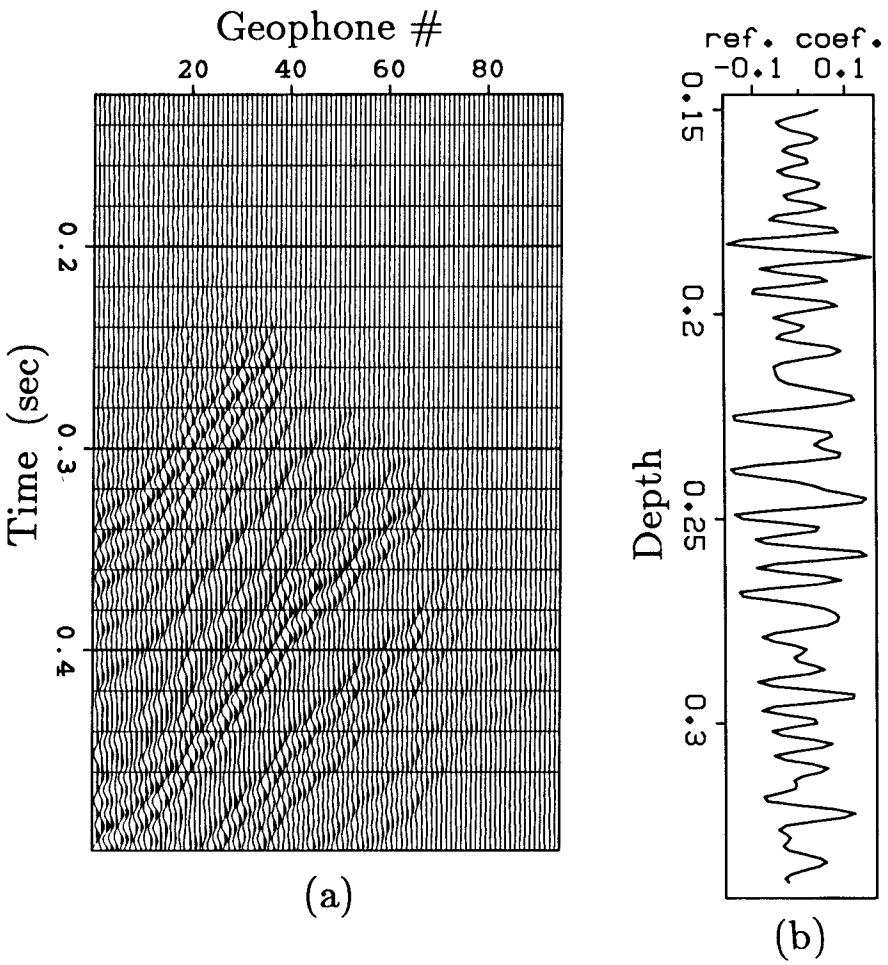

FIG. 3.14. A relinearized inversion of the residuals of Figure 3.13b models the first-order scattered waves in (a). I use the linearized modeling equations (I.2), (I.3), (I.4), and (I.5). The impedance-derivative perturbation (b) accounts for these new events.

that of the data distribution. The tails of the convolved pdf match those of the data pdf very well. See Appendix C for more information on the techniques used.

3.5.6. Extracting reliable signal

Using the three distributions of equation (3.12), we may calculate the most probable amount of signal in every sample of the transformed data. A Bayesian estimate gives the expected value of s' given d'.

$$\hat{s}' = E(s' \mid d') = \int x \; p_{s'|d'}(x \mid d')dx$$

$$= \frac{\int x \; p_{s'}(x)\, p_{n'}(d'-x)\, dx}{p_{d'}(d')} \tag{3.14}$$

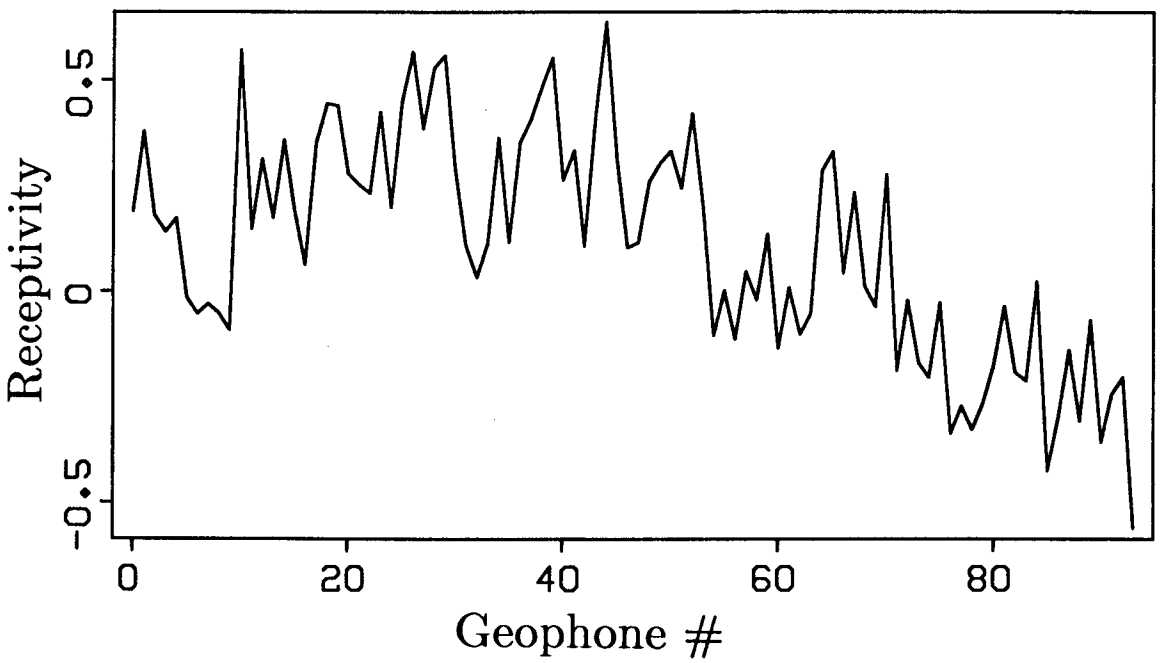

FIG. 3.15. The second linearization also perturbs the geophones' receptivity function, the fractional change from an amplification of 1. Amplification for adjacent traces varies as much as 30%; the acoustic impedance cannot model these rapid changes without creating false reflections. The low-frequency trend, however, can also be modeled by slow changes in the impedance function.

Figure 3.18 shows the estimate $\hat{s}'$ as a function of d'. This Bayesian estimate very nearly equals d': for most amplitudes, noise is expected to be zero. I have found in other applications that the Bayesian estimate can alter perturbations considerably. Godfrey (1979) uses the Bayesian estimate to estimate non-Gaussian reflection coefficients from his stochastic deconvolution. In most of the applications of Harlan et al. (1983) the Bayesian estimate did much to separate signal and noise. The Bayesian estimate of signal, however, may not be much more probable than other values. If an estimated value is unreliable, I prefer not to use it.

Now let us calculate the reliability of the Bayesian estimates. Accept samples of the Bayesian estimate as reliable perturbations of the signal if these samples contain less than, say, 5% noise with greater than 95% probability. Define the reliability by

$$\begin{aligned} \text{reliability} &\equiv P\left[-c\hat{s}' < s' - \hat{s}' < c\hat{s}' \mid d'\right] \\ &= \frac{\int_{(1-c)\hat{s}'}^{(1+c)\hat{s}'} p_{s'}(x)p_{n'}(d'-x)dx}{\int_{-\infty}^{\infty} p_{s'}(x)p_{n'}(d'-x)dx} . \end{aligned} \tag{3.15}$$

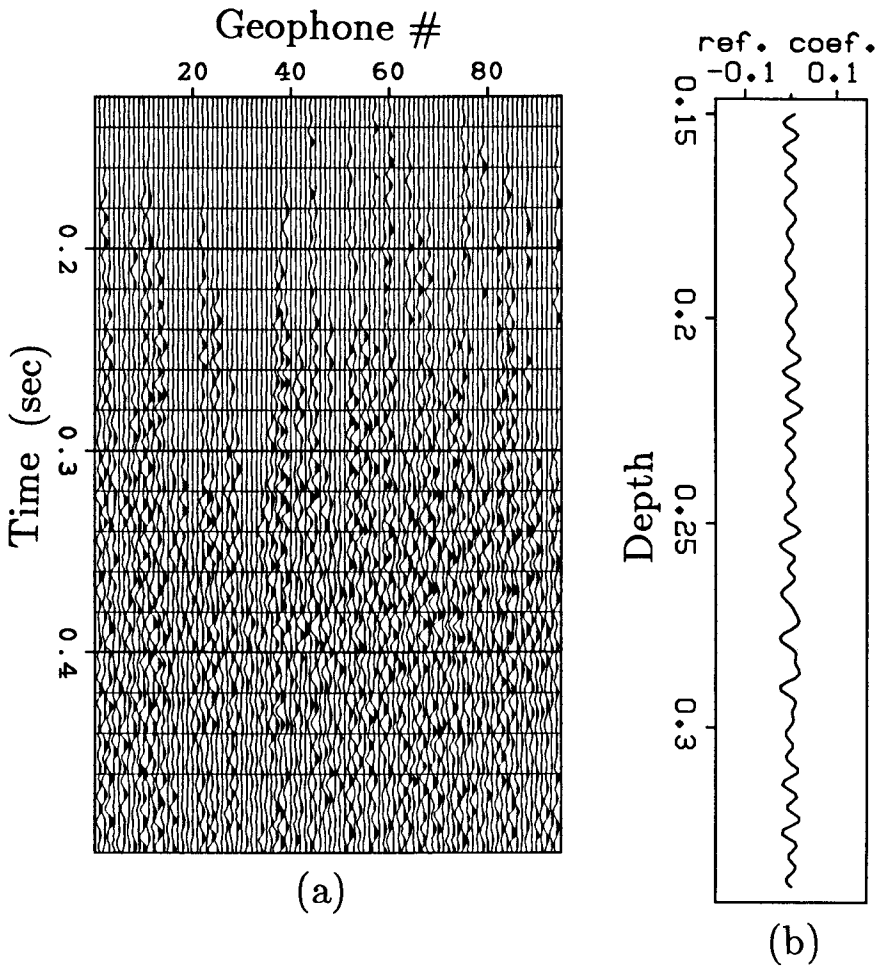

FIG. 3.16. (a) To test the influence of noise on the impedance perturbation of Figure 3.14b, I destroy the coherence of the residual data of Figure 3.13b. I calculate the impedance-derivative perturbations that linearly model the incoherent residuals best. By making all residuals behave as noise, this perturbation effectively over-estimates the noise in the perturbation of Figure 3.14b. If the residuals had contained nothing but noise, we would expect (b) to have the same amplitude distribution as Figure 3.14b.

$c = 0.05$ or some other small fraction of allowable noise. See Appendix B for further discussion.

Figure 3.19 shows the reliability of the Bayesian estimate as a function of the amplitude d'. The high amplitudes are the most reliable, and the low amplitudes the least. Figure 3.20a displays the Bayesian estimate of the signal in the linear signal perturbations. Figure 3.20b shows the reliability of each sample. Quite a few samples show greater than 95% reliability. Deciding to examine a slow convergence, I accept only the eight most reliable perturbations and set the others to zero (Figure 3.20c). Figure 3.21a shows the corresponding impedance function. Figures 3.21b, 3.21c, and

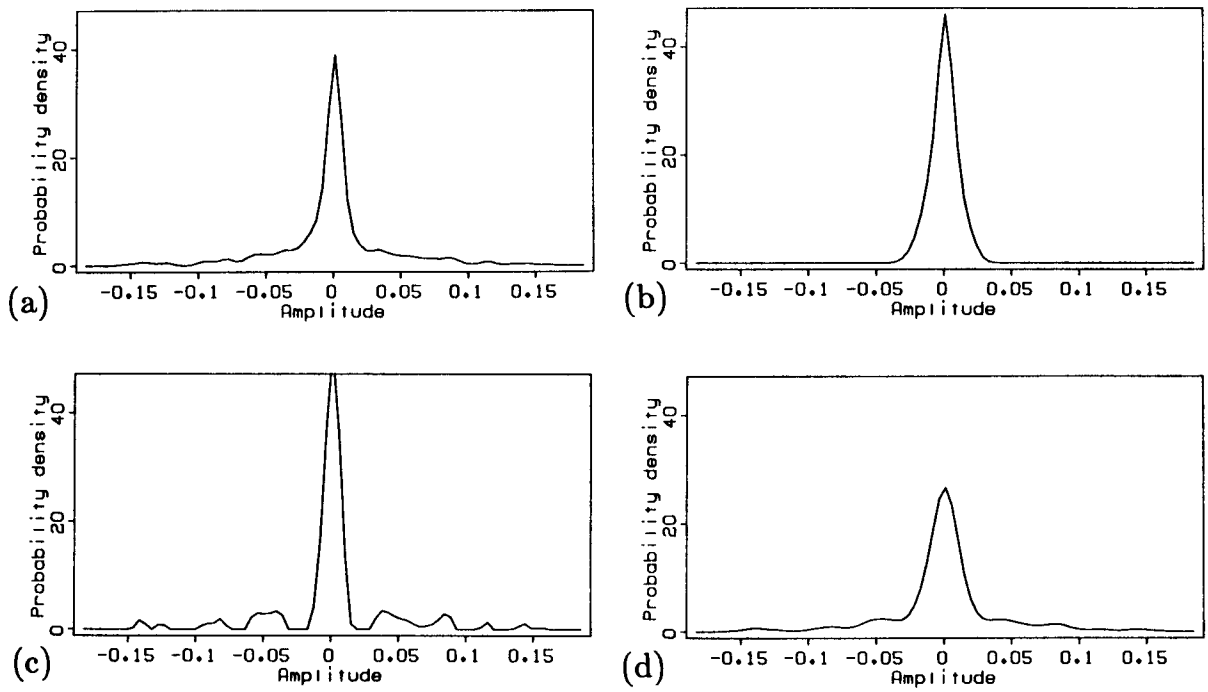

FIG. 3.17. (a) To estimate the probability distribution function (pdf) of the transformed data residuals, I take a histogram of the perturbation of Figure 3.13b. Signal and noise remain additive in the linear perturbation. (b) I over-estimate the transformed noise pdf with a histogram of Figure 3.16b. Note that the highest amplitudes of the data pdf do not appear in the noise pdf. If samples of transformed signal and noise add, then their pdf's should convolve. To estimate the pdf for signal (c), I deconvolve (a) with (b); we use a cross-entropy method that equivalently maximizes the probability of the data. Part (d) shows the convolution of (b) with (c); (a) is imperfectly reproduced because (b) has a wide central peak.

3.21d show the impedance functions of later iterations. Figure 3.22 shows the VSP's that are modeled for each of these four iterations. Convergence is largely reached by the last iteration. Further perturbations are weak and have less than 95% reliability (for less than 5% noise). Each iteration adds details of decreasing reliability to the impedance; an interpreter can emphasize the details accordingly.

To obtain the impedance function of Figure 3.1a, I perform one last iteration that accepts all perturbations, whether reliable or unreliable. The resulting model (Figure 3.2a) is little better, and the structure of the impedance function is only slightly blurred. Objective function (3.7) has been effectively minimized. The impedance function of Figure 3.1b minimizes the objective function equally well in the same number of iterations but does not eliminate unreliable signal perturbations. These very different impedance functions show that the objective function has a relatively flat bottom. Our

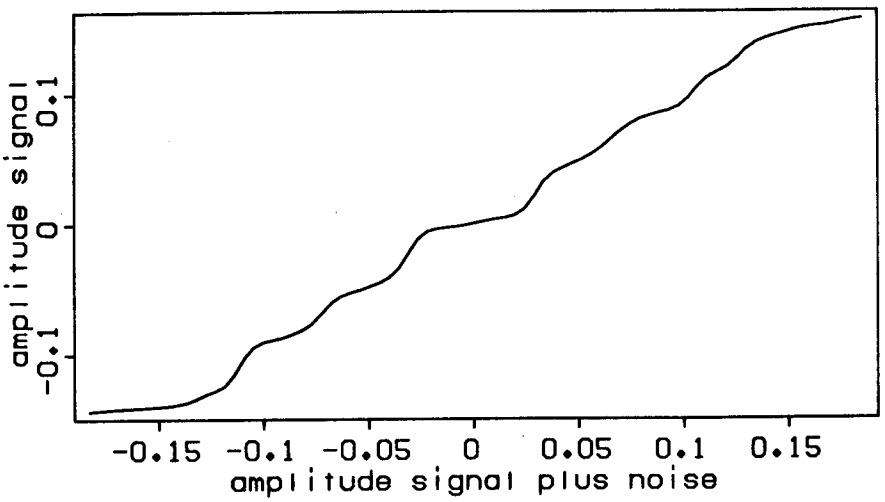

FIG. 3.18. From the signal and noise pdf's of Figure 3.17, I calculate the Bayesian expectation of transformed signal present in a sample of transformed data (the linear perturbation of impedance). This expectation changes perturbations very little, though we cannot assume it will always do so.

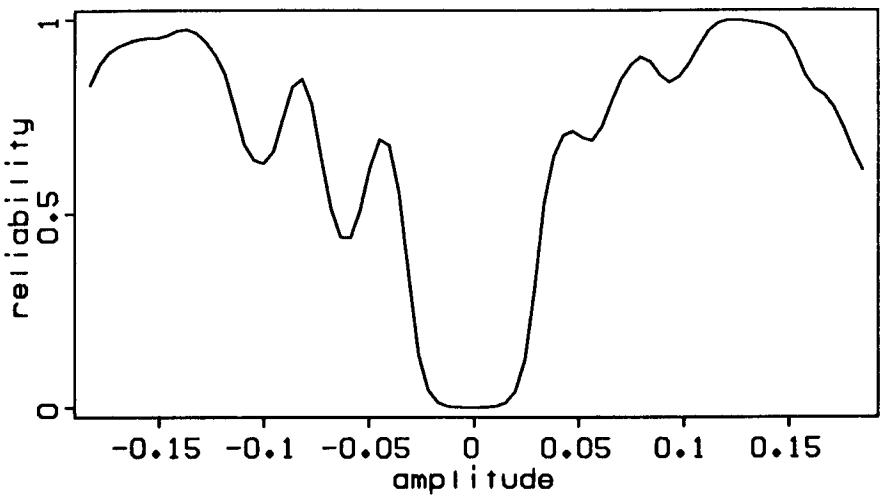

FIG. 3.19. From the pdf's of Figure 3.17, I calculate the reliability of the Bayesian estimate. I calculate the probability that the error in this estimate is less than 5% . Note that the largest amplitudes are the most reliable. The largest amplitudes were present in the transformed data and signal pdf's of Figures 3.17a and 3.17c, but not in the noise pdf of Figure 3.17b.

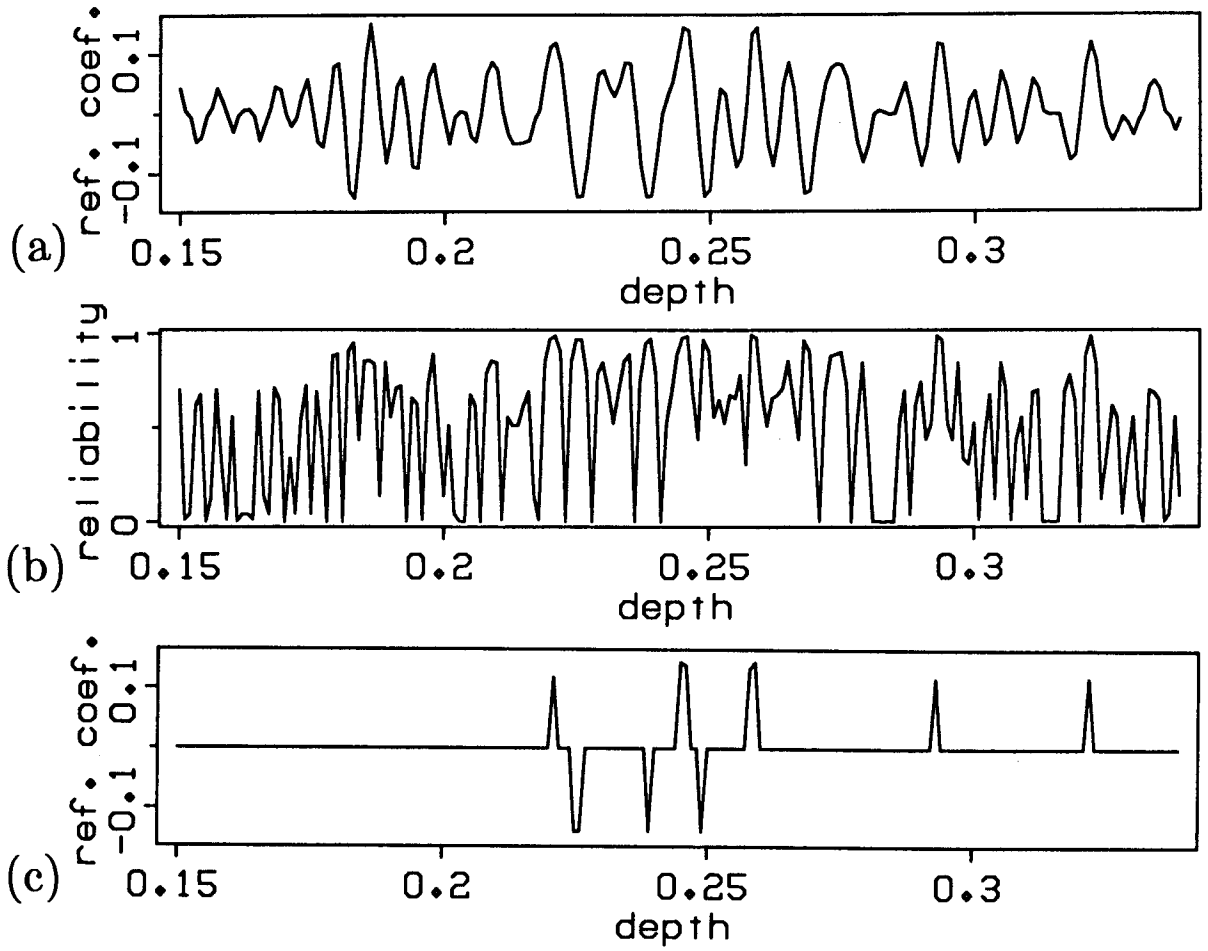

FIG. 3.20. (a) contains the Bayesian estimate of the signal in the impedance-derivative perturbations of Figure 13.4b. (b) shows the reliability of the samples of this Bayesian estimate. For (c) I accept the eight most reliable perturbations, those with greater than 95% reliability.

inversion of reliable parameters effectively limits the class of impedance functions to be considered.

3.5.7. Extracting reliable noise

The extraction of VSP noise uses the same statistical tools used in the extraction of signal. The estimation of signal and noise pdf's, however, can avoid the cross-entropy deconvolution. Subtract the most reliable signal from the data and take a histogram to estimate the pdf $p_{d^{res}}(\cdot)$ of the residual data, as shown in Figure 3.9. Let us pessimistically define the residual signal to be all events that can be modeled by the linearized modeling equation (3.9).

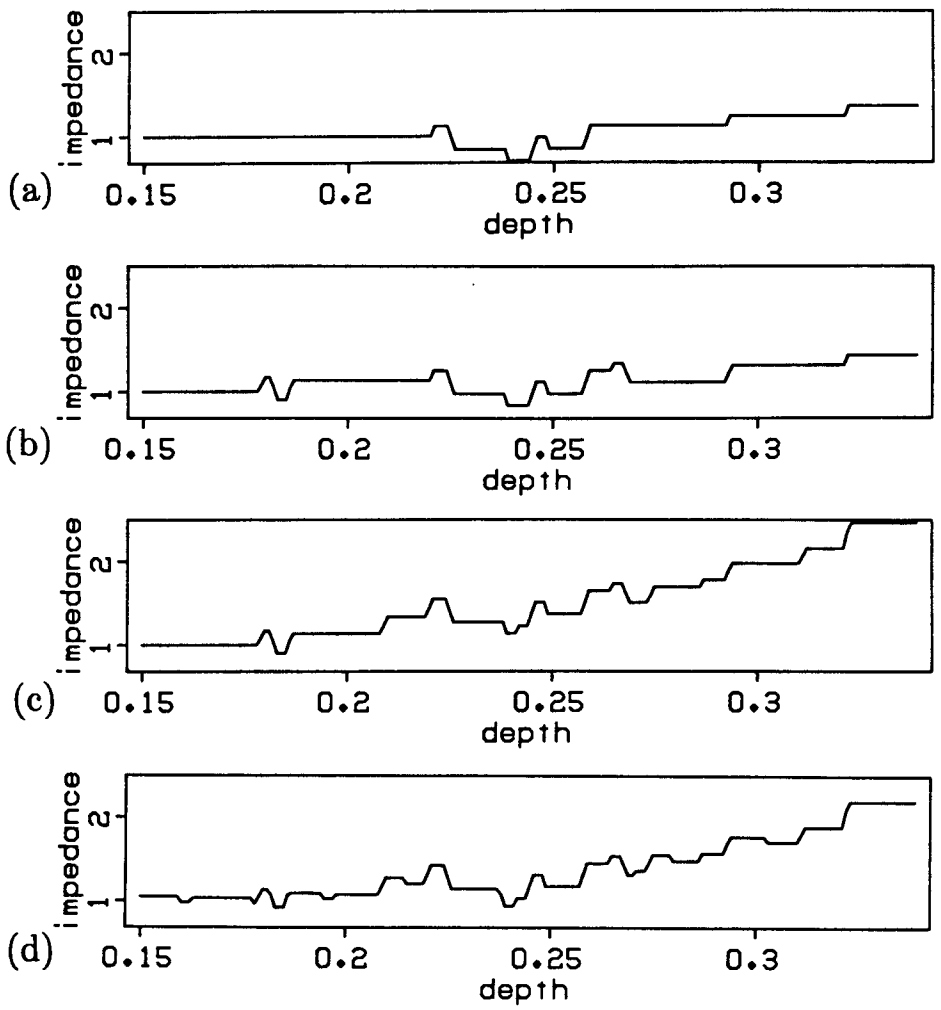

FIG. 3.21. (a) shows the impedance function perturbed with the reliable derivatives of Figure 3.20c. (b), (c), and (d) show the results of further iterations. Further perturbations did not have 95% reliability (for less than 5% noise).

First find the perturbation $\Delta \mathbf{s}$ that minimizes the quadratic objective function (3.10) for the residual data. To estimate the pdf of the residual signal $p_{s\,res}(\cdot)$ take a histogram of all events that the minimization may attempt to fit to the residuals—a histogram that overestimates the remaining signal. As before, we could deconvolve histograms for the residual noise pdf $p_{n\,res}(\cdot)$. However, if the estimate of residual signal is statistically independent of the noise, then we may estimate the noise pdf directly from a histogram of the remaining residuals (after subtraction of the overestimated signal). The pessimistic estimates of residual signal and noise will be sufficiently independent only if the linear optimization does not attempt to fit too many noise events as signal.

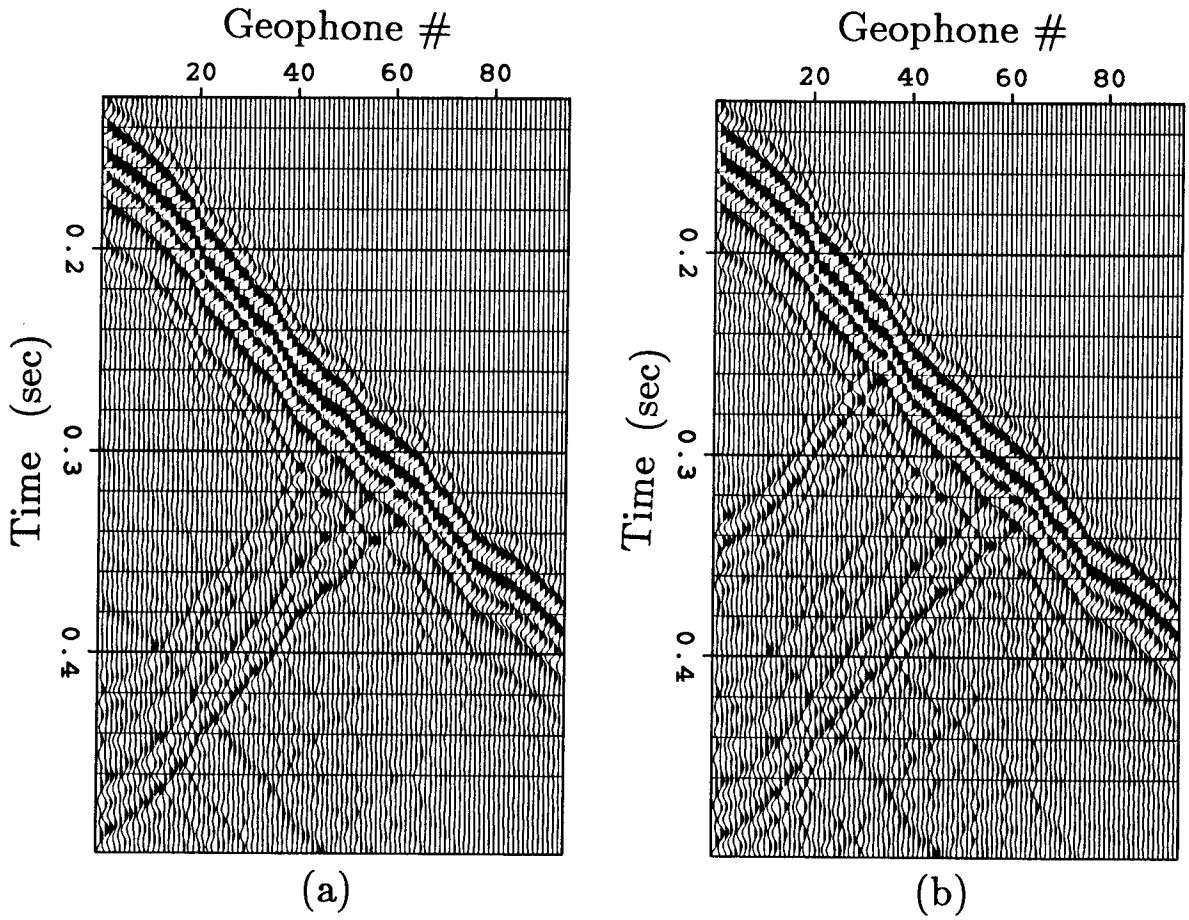

FIG. 3.22a,b. These modeled VSP's correspond to the impedance logs of Figures 3.21a and 3.21b. The full acoustic wave equation is used. The strongest reflections appear to have been inverted first.

With the necessary pdf's, we may proceed as before. Calculate a Bayesian estimate of the residual noise. Extract those samples of the residuals that with greater than, say, 50% probability contain less than 4% signal (Figure 3.10). Trial and error has suggested that a small percentage of contamination is preferable to a large probability of certainty.

The coherence of the noise requires that the extraction be somewhat smooth over time. Otherwise the intervals between the peaks of noise wavelets are broadened and the frequency content distorted. One simple method to achieve smoothness is simply to smooth a multiplier that selects the reliable samples. Instead, I take advantage of the fact that elastic wave motion tends to be locally monochromatic (the noise is from wave motion).

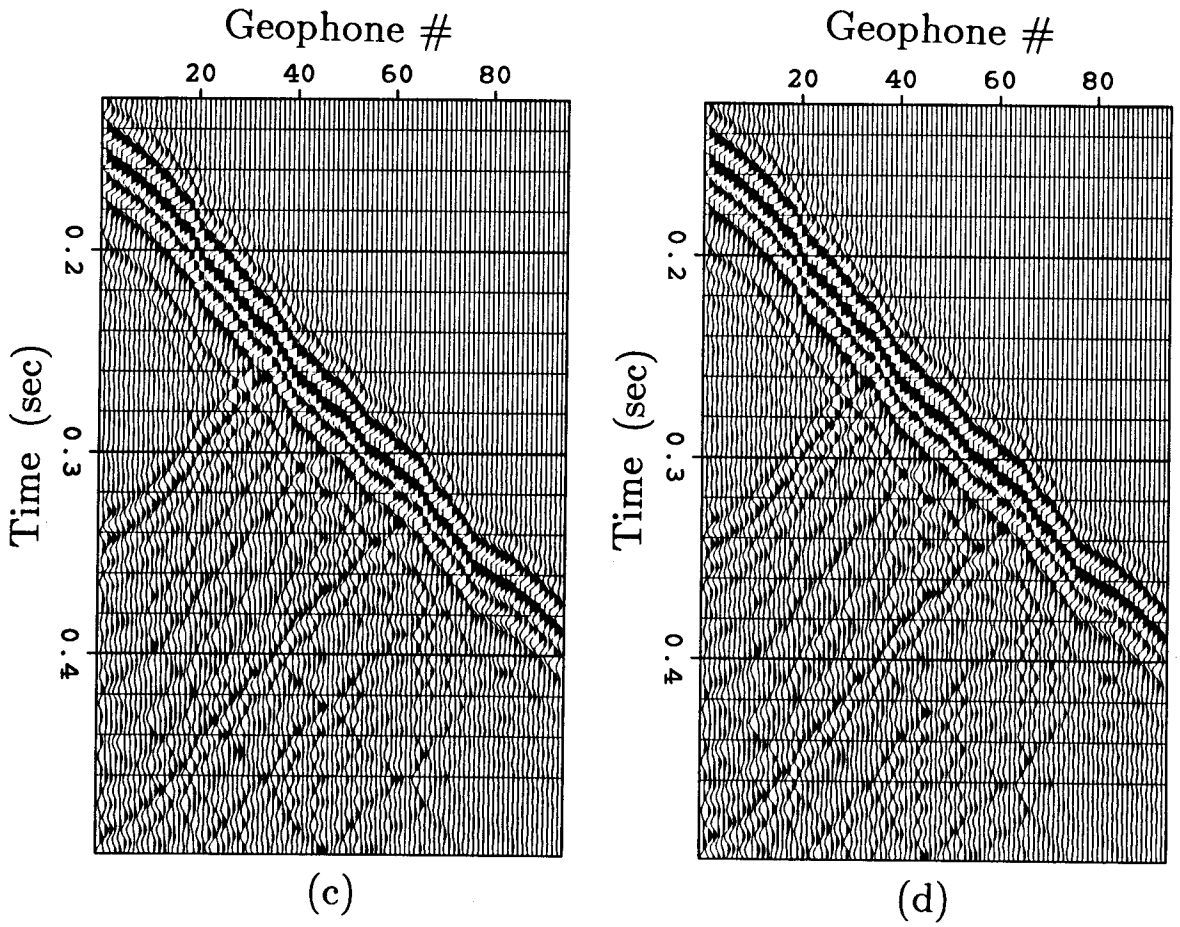

FIG. 3.22c,d. These modeled VSP's correspond to the impedance logs of Figures 3.21c and 3.21d. Convergence is essentially reached by the last iteration. A final iteration accepted all perturbations, reliable and unreliable, to create the impedance function and VSP of Figures 3.1a and 3.2a .

The true amplitude of a nearly monochromatic displacement is given by the analytic envelope, the magnitude of the complex analytic trace. Calculate analytic traces with the Hilbert transform. With every real trace of the residuals, couple an imaginary trace that is 90° out of phase. To apply this transformation, Fourier transform, zero negative frequencies, double positive frequencies, and inverse Fourier transform. Signal and noise remain additive; moreover, histograms do not change noticeably. The zeros of the real wavelets correspond to the peaks of the imaginary. The analytic envelope is the magnitude of the analytic trace, the square root of the summed squares of the real and imaginary parts.

As we saw before, discrimination between signal and noise is highest in samples with the highest amplitudes. For all phase shifts between plus and minus 90°, the maximum amplitude is equal to the analytic envelope, with the sign of the original

trace. Although the analytic envelope is not a linear function of the data, any single phase-shifted trace is. Let us imagine an enlarged data space containing traces of all phase shifts. We could estimate the reliability of noise for all trace coordinates and phase shifts. Let us consider any particular coordinate to be reliable if the phase shift with the highest amplitude is reliable. Again, this highest amplitude is just the signed analytic envelope.

In summary, take histograms from the unphase-shifted traces (or from analytic traces) for signal and noise statistics. Estimate reliability from the signed analytic envelope, and extract the noise from corresponding samples of the data residuals.

3.6. CONCLUSIONS

The inversion of VSP's features many problems common to seismic inversions. The simple acoustic wave equation ignores many physical parameters that must affect the data, yet the equation models the transmitted and reflected P waves very well. Moreover, very different one-dimensional acoustic-impedance functions can model the two-dimensional data equally well. Given no information to the contrary, I prefer the earth model that contains the smallest number of unpredictable details—in this case, non-zero derivatives in the impedance. A tube wave acts as strong additive non-Gaussian noise in VSP's, as does groundroll in surface recordings. Many varieties of coherent noise, such as sea-floor diffractions in ocean data, can be extracted if they are defined as signal with their own distinctive modeling equations.

The statistical estimates and extractions are effective only if the events of physical interest can be modeled by statistically independent parameters, the signal. Independence requires that each parameter be physically meaningful when taken alone. Furthermore, enough signal parameters must have similar statistics for histograms to be taken. Only non-Gaussian signal parameters will be distinguishable from noise. Gaussian signal remains Gaussian after a linear transformation has been performed, as does any unextracted noise.

Of course, not all signal and noise should be encouraged toward sparseness and non-Gaussianity: least-squares methods give the maximum-likelihood estimates of genuinely Gaussian signal and noise. Once the non-linear estimations have found all reliable perturbations of non-Gaussian parameters, least-squares methods will sufficiently invert the Gaussian residuals.

We do not need to estimate earth parameters directly; we do not even need to assume that values measured by different tools contain the same information. The VSP

measures information that is distinctly different from an acoustic well log. In processing the data, we should aim to put the information in the form most usable for interpretation. An inversion should choose a set of parameters that describe the important events in the data most simply. Many sets of parameter values may describe the data equally well, so we should not complicate the values without good reason.

A data set always contains some events, noise, that should not be described by the signal parameters. Before perturbing any signal parameter, we should ask if the corresponding data event can be explained by noise. Using the assumptions of statistical independence, we can recognize constructive and destructive interference in the signal and noise. We can estimate the distributions of signal and noise in the proposed perturbations directly from the data. It is not necessary that the form of probability distribution functions be simplified—histograms describe them very well. The reliability of perturbations can be calculated explicitly from the histograms.

The results of this approach to inversion seem to justify the generalization of its description. Indeed, other applications have already been explored by Harlan et al. (1983) with useful results. Because the inversion has been defined largely by what the interpreter considers a simple model, the results should be unlikely to mislead anyone with hidden assumptions.