Some remarks on nonlinear AVA inversion in anisotropic media

FERNANDO A. NEVES
IAG-Universidade de Sao Paulo, Sao Paolo, Brazil

MAARTEN V. DE HOOP
Center for Wave Phenomena, Colorado School of Mines, Golden CO, USA

April 14, 1999

ABSTRACT

A feasibility study for inverting reflection coefficients for a single-interface model consisting of a shale overlying a fractured sand is presented. To characterize the medium contact from Amplitude Versus scattering Angles (AVA) behavior associated with the reflection from a particular interface, in view of the anisotropy, wide-angle scattering data are required. However, wide-angle scattering turns the inverse problem into an essentially nonlinear one. The inverse problem is hence analyzed with a global optimization procedure, viz. a genetic algorithm (GA) with constraints. The key to make GA work for our AVA inverse problem appears to be a proper parameterization of the direct scattering problem, and introducing a parameter hierarchy. These reduce the complexity of the misfit function occurring in the optimization procedure, and let this function become convex over a relatively large range of medium parameter values. We illustrate how well our approach performs with a single-interface example representing a shale/fractured-sandstone contact. In this example, we assume the overburden, i.e. the shale, to be known; we allow the data to be contaminated with noise.

INTRODUCTION

Reflection and transmission of plane waves incident on a plane boundary between elastic media determine the Amplitude Variations with Offset (AVO). Conventional AVO analysis is based on analytic formulations for P- and S-wave reflection coefficients in isotropic media. Experimental studies (Wright, 1987; Kim *et al.*, 1993) have demonstrated that the presence of anisotropy severely affects the conventional AVO analysis. Several authors (Henneke, 1972; Keith & Crampin, 1977) have derived the expressions for the reflection of plane waves at an interface between two anisotropic media. These expressions have allowed developments of accurate and efficient forward modelling techniques for calculating the AVO response for anisotropic media.

The computation of reflection coefficients, i.e., the forward modelling, can be made sufficiently efficient to permit the application of global optimization methods to address the AVO inverse problem. Optimization methods are associated with the weak formulation of the inverse problem. The weak formulation of the inverse problem is essentially based on comparing theoretical (synthetic) with the observed (recorded) data values. This comparison is accomplished by evaluating an objective (fitness or misfit) function for many different combinations of model parameters (needed to calculate the theoretical values). The objective function integrates the available measurements, hence the adjective 'weak'. Depending upon the nature of the misfit function, a local (linear) or a global (non-linear) method of optimization is used. Linearized methods of optimization depend strongly on the starting model, and hence may be prone to being trapped in local optima. As a result, these methods fail if the initial model is too far from the most likely model. However, this failure might be masked by a regularization procedure commonly applied to local optimization. Linearized methods of optimization require derivatives of the misfit function, the analytical or numerical evaluation of which may be rather involved. A non-linear global optimization, though more costly, avoids various limitations of the linearized methods. For instance, by using an initial population of many randomly chosen

models, one does not require a starting model. Instead of trying to find a 'best' model that explains the observations, the global methods search for a family of velocity models which explain the observations to a desired level of confidence. Both local and global optimization methods provide procedures to quantify the degree of confidence in each of the final estimated models.

In this research, we have used a non-linear global inversion, Genetic Algorithm (GA) scheme, that can provide a statistical error analysis and model resolution. Previous studies have been done on isotropic AVO inversion (Sen & Stoffa, 1992; Mallick, 1995). In standard AVO analysis, an approximate linear relationship between data (reflection coefficient) and certain elastic parameter combinations is established, leading to a representation of the AVO gradient that can be employed in a local optimization procedure. The assumption of linearity is valid only for small angles of incidence (Rüger, 1996, typically less than 30 degrees). For isotropy, narrow-angle data may be sufficient to characterize the medium; for anisotropy, to determine a larger set of parameters, wide-angle data become a necessity. (Of course, a subset of the large set of parameters may still contain valuable information and to determine this, narrow-angle data may be sufficient.) At wide angles linear relations between reflectivity and elastic parameter combinations are no longer accurate and the AVO inverse problem becomes essentially nonlinear. It then depends on the behavior of the misfit function whether a local or a global optimization approach should be followed. It is our experience that the misfit can contain a sparse set of local extrema, i.e. can be non-convex, on which our choice of applying GA is based. Clearly, we could have chosen another global optimization method (Scales *et al.*, 1992) but GA is an attractive choice because it is relatively computationally efficient.

To extract a reflectivity from surface seismic data – compensating for the propagation effects like wave front curvatures in a, possibly complex, background velocity model and for curvature of the reflecting interface – a generalized Radon transform (GRT) (De Hoop & Bleistein, 1997) approach to prestack depth migration has been developed. The reflectivity yields a 1-D,

plane-wave, inverse scattering problem and hence the global optimization approach becomes feasible. In this paper, by example, we discuss what information the optimization can reveal. In the example, we assume that wavelet effects have been removed. We will consider a shale/sand contact, where the shale is transversely isotropic with a vertical axis of symmetry (VTI) and the sand is orthorhombic (ortho) with a horizontal plane of mirror symmetry. Thus the sand can be fractured and, if the fractures are aligned and vertical, the optimization potentially reveals the alignment orientation.

THE INVERSE SCATTERING PROBLEM

The direct scattering problem

We have reduced the direct scattering problem to the evaluation of plane-wave reflection coefficients. The wavefield across an interface separating two anisotropic elastic media should satisfy certain boundary conditions, viz., the particle velocity and the normal traction must be continuous (welded contact). If the incident field is a plane wave, from these boundary conditions follows a linear system of equations (the anisotropic form of the Zoeppritz equations). These equations are solved numerically for each slowness direction $\boldsymbol{\alpha}^{\mathrm{i}}$, say, associated with the incident plane wave, separately. The angles associated with this phase direction are θ (angle of incidence) and ψ (azimuth).

A common goal in AVA analysis is finding and ranking the variations and sensitivities of the reflection coefficient with relative changes in certain stiffness-tensor-component combinations. Since the variation yields differentiation, the validity of such an analysis is limited to narrow angles of incidence. The most general form known by the authors was given by Rüger (1996) for orthorhombic media. The precise component combinations will be given in a later section. That the validity of such an approximation can be limited is illustrated in Figure 1 for a medium contact given in Table 1. Our computations are valid for any direction of incidence, and since the media are anisotropic, are carried out in the complex horizontal slowness plane associated with the incident plane wave. The exact reflection coefficient depends nonlinearly on the me-

dia's stiffness tensors, in the Voigt notation given by $C_{IJ}^{(0),(1)}$, $I, J \in \{1, \cdots, 6\}$: $C^{(0)}$ defines the overburden (upper half space) and $C^{(1)}$ defines the target (lower half space). The discrepancies shown in Figure 1 can thus be attributed to non-linearities associated to the wide-angle scattering regime, justifying the use of optimization in the AVA analysis. Note, from Table 1, that the anisotropy was truly significant in this case.

To illustrate that narrow-angle scattering reveals only limited information about the lower medium, we have linearized the qP-qP reflection coefficient in the contrast $C^{(1)} - C^{(0)}$. In Figure 2 we observe that – in the weak contrast approximation – certain components of the stiffness (like C_{33}) contribute to the reflection coefficient at narrow angles whereas other components (like C_{11}) contribute at wide angles. In the presence of significant anisotropy, to characterize the lower medium, we need wide-angle data indeed.

The misfit

In an AVA analysis, to obtain the local reflection coefficient at the target as a function of directivity $\boldsymbol{\alpha}^{\mathrm{i}}$ from the seismic measurements, we have to compensate for propagation effects in the overburden and for radiation patterns of the sources and receivers at the surface. We have developed a generalized Radon transform (GRT) inversion procedure (De Hoop & Bleistein, 1997) to carry out the appropriate, migration-like, corrections for these phenomena. Here, we will assume that this procedure has been applied and that we have obtained 'observed' reflectivities $R^{\mathrm{obs}}(\boldsymbol{\alpha}^{\mathrm{i}})$. From our understanding of the direct problem we can simulate such observations as $R^{\mathrm{syn}}(\boldsymbol{\alpha}^{\mathrm{i}}; C^{(0)}, C^{(1)})$ as function of the media properties.

We cast the inverse problem – aiming to determine the properties $C^{(1)}$ of the (perturbed) medium in the lower half space given the ones $C^{(0)}$ of the (background) medium occupying the upper half space – in a common weak form: find $C^{(1)}$ that minimizes the misfit function

$$\mathcal{E}(C^{(1)}) = \int |R^{\mathrm{obs}}(\boldsymbol{\alpha}^{\mathrm{i}}) - R^{\mathrm{syn}}(\boldsymbol{\alpha}^{\mathrm{i}}; C^{(0)}, C^{(1)})| \, \mathrm{d}\boldsymbol{\alpha}^{\mathrm{i}} \ , \tag{1}$$

for a fixed point at the interface and dip. The reason for choosing the L^1 criterion is suppressing

the influence of isolated amplitude outliers. The proposed misfit can be associated with a Laplacian distribution of experimental uncertainties.

In practice, only a small number N^{i} of judiciously chosen directions is evaluated so that Eq.(1) must be replaced by

$$\mathcal{E}(C^{(1)}) \simeq \frac{1}{N^{\mathrm{i}}} \sum_{k=1}^{N^{\mathrm{i}}} |R^{\mathrm{obs}}(\boldsymbol{\alpha}_k^{\mathrm{i}}) - R^{\mathrm{syn}}(\boldsymbol{\alpha}_k^{\mathrm{i}}; C^{(0)}, C^{(1)})| \; . \tag{2}$$

Due to the nonlinearity of the direct scattering formula, the function $\mathcal{E}(C^{(1)})$ may be nonconvex and contain local extrema.

GENETIC ALGORITHMS

Genetic Algorithm is a global optimization technique which incorporates a stochastic search procedure. GA is a simulation of natural genetics. Initially applied to machine learning problems (Holland, 1975), it has been applied to many different problems in engineering, computer science and other fields (Goldberg, 1989) and to a range of geophysical problems. In seismology it has been applied to background velocity estimation (Jin & Madariaga, 1993), Rayleigh group dispersion estimation (Lomax & Snieder, 1994), vertical seismic profiling (Nolte & Frazer, 1994), seismic waveform inversion (Stoffa & Sen, 1991; Sambridge & Drijkoningen, 1992) and isotropic AVO analysis (Mallick, 1995). The successful application of GA to a wide range of geophysical problems (Whitley *et al.*, 1995) lead us to employ GA in anisotropic AVA inversion of wide-angle seismic data.

Basically, all the applications referenced above employed the same fundamentals of the genetic algorithm. Initially, each model is discretized by coding its parameters as a binary string, i.e., a series of 0's and 1's defining the value of each parameter in a model space. These strings are the building blocks of GA. The different ordering of bits within the string give different values allowed for each anisotropic parameter. The medium model $\mathbf{m}$, defined by $C^{(1)}$, may be described as

$$m_j = m_j^{\min} + k_j \, \Delta m_j \; , \qquad\qquad j = 1, \cdots, n \; , \tag{3}$$

where $m_j^{\min}$ is the minimum possible value of the j^{th} anisotropic parameter, $\Delta\mathbf{m}$ is the parameter discretization and k_j is the integer value stored in a binary string, there being a total of n such values in the string. With M bits we allow 2^M values for k_j.

A genetic algorithm works simultaneously with a group of different strings called a population. The size Q of the population should be approximately equal to the length of the binary string (Goldberg, 1989). Each iteration of GA aims to find an optimal model by manipulating this population of Q models using a three-stage procedure of selection, crossover and then mutation. All three stages are controlled by predefined probability values, named probability of selection, P_s, probability of crossover, P_c, and probability of mutation P_m. These probability values, which range between zero and one, may vary as the algorithm iterates. The key parameters controlling a GA are Q, P_s and P_c.

GA starts by randomly generating a population of Q models from a predefined model space. This procedure produces models according to a uniform distribution but is also able to incorporate any *a priori* information or constraints applied to the model space. Without constraints, any model defined within the model space is as likely to be generated as any other. Several authors (Cary & Chapman, 1988; Sambridge & Drijkoningen, 1992) assign a high penalizing objective value to models that do not satisfy constraints. These models are not eliminated from the pool of possible models. In this way, the creation of forbidden regions in the model space is avoided. This approach, which keeps these regions as possible to be searched (although less likely), permits a pathway towards the global optimum. However, the choice of a value for this high penalty is rather subjective. Davis (1987) discusses the disadvantages of using high penalties. He states that the incorporation of a high penalty into a GA scheme leads to a GA in which generation of models violating constraints is likely. Consequently, it may spend most of its time evaluating unacceptable models (which do not satisfy the constraints). Besides, it is also possible that when an acceptable model is found, it drives the search towards itself, since the likely paths to other acceptable models require the generation of unacceptable models

as intermediate structures, and the penalties for not satisfying the constraint make it unlikely that such structures will survive throughout GA evolution. We have followed the approach of Michalewicz and Janikow (1991), where constraints are incorporated into the GA scheme. This approach avoids the creation and evaluation of unacceptable models. Hence, regions in the model space associated with high penalty are not created.

For each of the $\mathbf{m}^J$, $J = 1, \cdots, Q$ models in the current population a numerical objective (raw fitness or misfit) value, $\mathcal{E}(\mathbf{m}^J)$, is computed (cf. Eq.(1)). In summary, the three-stage GA evolution is carried out as follows:

Selection

Selection creates a new off-spring population from the current (parent) population of models. The likelihood of selection of each member of the population is determined by a probability depending on its misfit function value,

$$P(\mathbf{m}^I) = \frac{\mathsf{S}[\mathcal{E}(\mathbf{m}^I)]}{\sum_{J=1}^{Q} \mathsf{S}[\mathcal{E}(\mathbf{m}^J)]} , \qquad I = 1, \cdots, Q . \tag{4}$$

Here, S is a scaling function that we have chosen to be exponential (Sen & Stoffa, 1992).

We have opted for the stochastic remainder selection without replacement method. In this method, the selection operation is repeated until the number of off-springs equals the size of the parents' population. In the stochastic remainder selection without replacement, first the expected number c_I of copies is calculated for each model I (Nolte & Frazer, 1994),

$$c_I = Q \, \frac{\mathsf{S}[\mathcal{E}(\mathbf{m}^I)]}{\sum_{J=1}^{Q} \mathsf{S}[\mathcal{E}(\mathbf{m}^J)]} , \qquad I = 1, \cdots, Q , \tag{5}$$

see Eq.(4). Initially, each model I of the parent population is copied $N_I = \mathrm{int}[c_I]$ times. Let us denote the fractional remainder by ϕ_I; we have $c_I = N_I + \phi_I$. Then a model is selected with a probability proportional to ϕ. However, once it has been selected it cannot be selected again.

Crossover

From the parent population of Q models a new generation of Q models is generated, each of which is obtained by mixing bit-strings from two parents. All the Q parents are randomly paired to produce $Q/2$ couples. A probability P_c (usually greater than 0.5) for performing this step is assigned. The GA employed in this research made use of a single-point crossover (Sen & Stoffa, 1992; Eshelman *et al.*, 1989).

Mutation

This final process allows any bit in an individual string to flip between 0 and 1. The main goal of this process is to add and maintain some degree of local diversity (since only individual bits are affected) into the whole inversion process, so that no genetic feature is permanently lost. In the GA used in this research the mutation probability is set to a value equal to the reciprocal of length of the binary string (Allan, 1995).

Our code makes use of the GENESIS package (Grefenstette, 1986). GENESIS is a system for function optimization based on Genetic Algorithms.

OPTIMIZATION STRATEGY

We impose orthorhombic symmetries for the media on both sides of the interface; the two media have in common the interface as a plane of mirror symmetry. The complexity of the misfit function – dependent on the angular range – reveals that a GA can be an appropriate choice for the optimization procedure. To make the GA viable, however, we have made the following adaptions:

Pre- and post-critical modelling

Though we consider pre-critical data, the search through model space may generate parameter combinations that yield post-critical events. Hence, during the search, the reflection

coefficient may become complex as a function of horizontal slownesses (associated with scattering angle and azimuth) and in our code is treated as such.

Constraints for sedimentary rocks

To guarantee obtaining an acceptable model, we impose constraints following from the conditions of *mild* anisotropy, applicable to most sedimentary rocks. They are (Helbig & Schoenberg, 1986): the lowest P-wave velocity along any coordinate axis is greater than the highest S-wave velocity along any coordinate axis, i.e.,

$$\min(C_{11}, C_{22}, C_{33}) > \max(C_{44}, C_{55}, C_{66}) \ , \tag{6}$$

and the longitudinally polarized wave is always faster than the transversely polarized wave in any of the coordinate planes. In fact no anomalous polarizations can occur if

$$C_{23} + C_{44} > 0 \ , \ \ C_{13} + C_{55} > 0 \ , \ \ C_{12} + C_{66} > 0 \ . \tag{7}$$

It is important to emphasize that these conditions for mild anisotropy do not imply weak anisotropy as defined by Thomsen (1986). It does imply that the 'longitudinal' sheet of the slowness surface is disconnected from the two 'transverse' sheets.

Parametrization matching the scattering process

From studying the AVA gradient in the symmetry planes (Rüger, 1996), we have determined a hierarchy of stiffness-tensor-component combinations that can, in principle, be resolved. Such parameter combinations, given below, were introduced by Tsvankin (1996) for describing seismic signatures in orthorhombic media. We split them into a hierarchy of four sets. In parentheses we indicate what subset of the observations is sensitive to the parameter set under consideration.

Set A (vertical ($3-$ direction) incidence):

$$\alpha = \sqrt{C_{33}/\rho} \ ; \tag{8}$$

set B (narrow angle):

$$\beta = \sqrt{C_{55}/\rho}\,, \quad \gamma^{\rm S} = \frac{\gamma^{(1)} - \gamma^{(2)}}{1 + 2\gamma^{(2)}}\,, \tag{9}$$

in which

$$\gamma^{(1)} = \frac{C_{66} - C_{55}}{2\,C_{55}}\,, \quad \gamma^{(2)} = \frac{C_{66} - C_{44}}{2\,C_{44}}\,,$$

control the shear wave propagation along the vertical direction, and

$$\delta^{(1)} = \frac{(C_{23} + C_{44})^2 - (C_{33} - C_{44})^2}{2\,C_{33}(C_{33} - C_{44})}\,, \quad \delta^{(2)} = \frac{(C_{13} + C_{55})^2 - (C_{33} - C_{55})^2}{2\,C_{33}(C_{33} - C_{55})}\,. \tag{10}$$

Set C (wide angle):

$$\begin{aligned} \epsilon^{(1)} &= \frac{C_{22} - C_{33}}{2\,C_{33}}\,, & \epsilon^{(2)} &= \frac{C_{11} - C_{33}}{2\,C_{33}}\,, \\ \delta^{(3)} &= \frac{(C_{12} + C_{66})^2 - (C_{11} - C_{66})^2}{2\,C_{11}(C_{11} - C_{66})}\,, \end{aligned} \tag{11}$$

and *set* D:

$$\sqrt{C_{66}/\rho}\,. \tag{12}$$

The parameters $\epsilon^{(2)}, \delta^{(2)}$ are associated with 'P-SV' waves in the $[x_1, x_3]$-plane; the parameters $\epsilon^{(1)}, \delta^{(1)}$ are associated with 'P-SV' waves in the $[x_2, x_3]$-plane. The shear-wave splitting at vertical incidence is represented by $\gamma^{\rm S}$. The coefficient $\delta^{(3)}$ plays the role of Thomsen's 'δ' for the $[x_1, x_2]$-symmetry plane.

The parameters $\delta^{(3)}$ and C_{66} have to be determined from reflections at azimuths away from the symmetry planes. Note that α, β parametrize isotropic AVO behavior.

Narrow angle 'preconditioning'

Narrow-angle scattering is sensitive to parameter sets A and B, while wide-angle scattering is particularly sensitive to parameter set C and also to parameter set D. We make use of this

hierarchy by first performing vertical-incidence-angle inversion for set A (setting the perturbations in the parameters of sets B, C and D equal to zero). The outcome of this inversion provides us with an optimal *range* of values for the parameters in set A which information is passed on as a prior distribution in the narrow-angle inversion for set B (setting the perturbations in the parameters of sets C and D equal to zero). The procedure is repeated to find set C, and eventually set D. Essentially, this 'preconditioning' reduces the complexity of the misfit function: it will become convex over a larger range of parameter values. With regard to this, observe that the scattering problem becomes *linear* in our parameters for *narrow* scattering angles.

Sharing

To circumvent the problem of genetic drift (Goldberg & Segrest, 1987) caused by finite population sizes, the method of sharing (Holland, 1975) is employed. In sharing, the objective fitness of the neighborhood of a particular point in model space is degraded in proportion to the number of members sharing that neighborhood.

ERROR ANALYSIS

The goal of non-linear seismic inversion is to find the optimum model for given seismic data by exploring favorable regions of the model space. However, several models might satisfy the observed data to a preset accuracy level. In our case, assuming that the preprocessing – including amplitude balancing, extracting reflectivity from the data with the GRT, and wavelet deconvolution – does not introduce a coherent noise contribution, this is generally due to the presence of incoherent noise in the observed reflectivities or numerical errors in the forward modelling scheme. Therefore, a Bayesian interpretation of the inversion procedure is desirable (Tarantola, 1987), an interpretation which can be described by *a posteriori* probability density functions (PPD's) for the different parameters in model space. The method of sharing guarantees the survival of a reasonably homogeneous distribution of models from which the

PPD's are determined. The computation of error associated with the optimal model is necessary to quantify the degree of confidence in this model, and thus providing information on alternative models that would fit the data equally well.

A complete error analysis would involve calculation of the PPD by evaluating the misfit for every possible model allowed by the parameterization of the model space and satisfying the constraints. Obviously, this task is computationally untractable for a model space consisting of millions of acceptable models. Furthermore, there is the problem of how to interpret a multi-parameter PPD. Consequently, a feasible and still accurate method must be adopted to estimate the PPD for the AVA inverse problem. The method used in this paper is a graph binning technique, proposed by Basu and Frazer (1990), Frazer and Basu (1990) and Sen and Stoffa (1992). The latter authors fully describe the theory for this technique, here we just outline the method. Graph binning is performed by assigning a bin to each discrete value that a parameter might take. Once a probability has been computed for a particular model, that value is summed into the bins corresponding to that model's parameters.

The choice of the probability function depends upon the distribution of the noise or error in the observed data (Cary & Chapman, 1988). This is an important issue since in many situations it is quite difficult to obtain an estimate of noise statistics. The error can be due to instrumental errors or due to the use of an approximate theory in the prediction of the data (Tarantola, 1987). Taking the exponential scaling function in Eq.(4), a model $\mathbf{m}$ with a misfit $\mathcal{E}(\mathbf{m})$ is assigned a probability (PPD)

$$\Pi(\mathbf{m}^I) = \frac{\exp[-\mathcal{E}(\mathbf{m}^I)/2\sigma_{\mathcal{E}}^2]}{\sum_{\mathbf{m}^J \in \{\text{all populations generated}\}} \exp[-\mathcal{E}(\mathbf{m}^J)/2\sigma_{\mathcal{E}}^2]} \,, \tag{13}$$

where $\sigma_{\mathcal{E}}$ is the standard deviation of the distribution of misfit values of the initial population. In principle, this expression is an $n \times 2^M$ matrix, where n is the number of anisotropy parameters and 2^M is the number of values they can take. We have assumed that the data are statistically independent.

For display purposes, we will always choose all parameters except one equal to their optimum values. The PPD curve thus obtained, spans the range of models for a single parameter variation. From such a PPD curve, the uncertainty (auto-covariance) in the final model parameter can be estimated. To obtain all the covariances, the full matrix in Eq.(13) has to be computed.

SYNTHETIC AMPLITUDE DATA EXAMPLES

To mimic a reservoir setting, viz. a shale/fractured-sandstone contact, in our example we have chosen a VTI medium overlying an orthorhombic medium, assuming that the two media have the interface as a plane of mirror (up-down) symmetry in common. The media are described by the stiffness tensors in Table 1. We thus restrict our analysis to a single-interface case. In the inverse problem, we assume the overburden – the upper VTI medium – to be known. Not only do we have to determine the parameters associated with the second row of Table 1, but also the azimuthal orientation ψ_C of the lower ortho medium.

We parameterize the unknown ortho medium according to sets A, B, C and D, and discretize our model space into 10^{15} different media. The search domain is bounded by the parameter values given in Table 2. We illustrate the performance of our inversion procedure by running four cases: narrow- and wide-angle, noise-free and noisy data. The noise-free data are the exact qP-qP reflection coefficients, computed from the values of Table 1, as function of angle of incidence (0 to 40 degrees) and azimuth (15, 75 and 105 degrees). The noise added to the data is chosen from a random sequence, has a magnitude as high as 30% of the maximum amplitude, and has zero-mean. The noise-free and noisy data are shown in Figure 3. Note that the acquisition geometry does not coincide with the true symmetry planes of the lower medium (in fact rotated by 15 degrees). The narrow-angle data are sampled at incidence angles of $\theta \in \{0,5,10,15,20\}$ degrees, whereas the wide-angle data are sampled at incidence angles of $\theta \in \{20,25,30,35,40\}$ degrees.

In all cases we allow the GA to run for 10 000 iterations (generations). We repeat this search 50 times. In Tables 3 and 4 below the 10 best solutions for the lower medium are given in the two noise-free cases (the true solution is slanted). In the narrow-angle data inversion we have set the contrasts in the parameters of sets C and D to zero, since they hardly influence the simulations. The results given in Table 3 are passed on to the wide-angle inversion through a prior distribution. With our choices of parameters, number of iterations, and population size, modelling the *exact* reflection coefficients, we find that the GA converges consistently. This confirms that our single-interface inverse scattering problem with sufficient noise-free data has a unique solution.

How well-determined the different parameters of our inversion example are, is illustrated by adding random noise to our synthetic data. The results are given in Table 5 from narrow-angle and in Table 6 from wide-angle data. It is no surprise that the parameter combination associated with the smallest misfit (indicated by an arrow) no longer represents the true solution, but a subset of parameters is close. From Table 5, we observe that certain parameters, like α, seem to be well-determined whereas others, like $\delta^{(1)}$, are not. The wide-angle data do help in this respect, but do not resolve instabilities.

To illustrate the stability aspects of our inversion example, we have computed approximate marginal PPD's both from the noise-free and the noisy data of Figure 3 in accordance with Eq.(13). The PPD's were estimated from the same 50 GA searches as were carried out to estimate the parameters of the lower medium. As conjectured by Sen and Stoffa (1996) repeating the GA search a number of times leads to a reasonable estimate of marginal PPD's. A more rigorous approach to find the PPD's is discussed, for example, by Davis and Principe (1991). The underlying idea is to form a generalized Markov chain each element of which is an entire population. After evolving the GA for a while, a Boltzmann factor for the entire population is computed; based on a Metropolis criterion, it is then decided whether to accept or reject the entire population. Also, see Tierney (1994).

With the vertical P-wave velocity (α) contrast, determined from narrow-angle data, corresponds the PPD shown in Figure 4 (noise-free) and the PPD shown in Figure 5 (noisy data). As expected, the PPD in Figure 5 is wider than the one in Figure 4, but the mean is well-determined in both cases.

Analyzing, on the other hand, the PPD's associated with $\epsilon^{(2)}$ reveals a different behavior alltogether. The PPD determined from noise-free narrow-angle data is shown in Figure 6 and is slightly wider than the PPD determined from noise-free wide-angle data, shown in Figure 7. The PPD's determined from the noisy data, see Figures 8 and 7, are far from well behaved. Though one can identify a peak at the true value, the confidence in the mean is truly low. Note, however, that the PPD determined from noisy narrow-angle data is considerably wider than the one determined from noisy wide-angle data. The contrast in a parameter like $\epsilon^{(2)}$ contributes to the reflection coefficient significantly at large scattering angles only (a leading behavior proportional to the sines of the scattering angle in the fourth power) – hence, the relatively large sensitivity to noise revealed by the PPD.

Proceeding along those lines, we are able to rank the levels of confidence in the different parameter reconstructions prior to the AVA analysis itself.

One may wonder whether it would be useful to add to the wide-angle body-wave data the head waves. The problem with this is that the order of singularity of the head wave front is one lower (smoother) than the body wave front. The head wave will thus be much harder to detect comparatively accurately. Also, the head wave is significantly influenced by any lateral medium variations along the interface: its interaction with the interface is non-local.

DISCUSSION

The analysis given in this paper is in a sense an extension of the work of Mallick (1995) to anisotropic media. Our primary concern was to make the GA a tractable option to carry out the AVA analysis for anisotropic parameters. The incorporation of anisotropy enforces one to

consider wide-angle data and hence to recognize the nonlinearity of the inverse problem. By example, introducing an appropriate parameterization and parameter hierarchy, we have shown that GA is a useful tool to deal with the nonlinearity. Also, we have briefly illustrated how to infer confidence levels for the different parameters from uncertainties in the data.

We are aware of many remaining issues. We mention the few that can be analyzed with GA:

- *Residual moveout*. The background medium, i.e. overburden, may not be known accurately enough. Nonhyperbolic moveout due to heterogeneity or anisotropy is likely to play a role in wide-angle data. In this respect, it is noted that from the reflection moveout $\delta^{(1)}$ and $\delta^{(2)}$ of the overburden (or background) can be determined, provided that α is known. Compensating for the residual moveout can be made part of the GA approach if the parameterization of the background is sparse enough.

- *Tuning*. A reservoir layer can be thin or consist of many thin layers, in which case primary reflections can no longer be separated from the (internal) multiples. Tuning is an increasing complication at wider angles in view of the apparent vertical dominant wave length. By adding multiple interfaces in the direct scattering problem, this issue can be addressed in our procedure (Scales *et al.*, 1992). The reflection coefficient is then replaced by the reflectivity following from an invariant imbedding procedure.

- *Mode conversions*. With the development of multi-component ocean bottom acquisition, the singly mode-converted waves are increasingly succesfully observed. These mode conversions carry additional information about the target layer and the associated measurements can be incorporated in our misfit function. The challenge, however, will be identifying *where* in the subsurface a particular mode conversion occurred, bearing in mind that a pair of conversions may be hidden in the single mode primary reflection.

- *Transmission*. The amplitude and arrival times for wide-angle events may be severely

affected by scattering due to fine-scale heterogeneity in the overburden and in the near surface. One way of addressing this issue is by incorporating a stochastic component in the constitutive parameters appearing in the modelling equations, and sampling the stochastic component appropriately.

Concerning the first issue, we would like to point out that the parametrizations of AVA analysis for wide-angle scattering and non-hyperbolic moveout analysis show many similarities (Grechka & Tsvankin, 1998).

ACKNOWLEDGMENT

This work was supported in part by the Consortium Project at the Center for Wave Phenomena, Colorado School of Mines, and in part by the Brazilian Ministry of Science and Technology, through the PRONEX–Reservoir Engineering Project. The authors would like to thank Dr. Tsvankin and Dr. Scales for many helpful discussions. They would also like to thank the referees Dr. Mallick and Dr. Sen for their valuable comments and suggestions.

References

Allan, W. 1995. *Stochastic waveform inversion of wide-angle seismic data*. Ph.D. dissertation, University of Cambridge, England.

Basu, A., & Frazer, L.N. 1990. Rapid determination of the critical temperature in simulated annealing inversion. *Science*, **249**, 1409–1412.

Cary, P.W., & Chapman, C.H. 1988. Automatic 1-D waveform inversion of marine seismic refraction data. *Geophys. J.*, **93**, 527–546.

Davis, L. (Editor). 1987. *Genetic Algorithms and Simulated Annealing*. Pitman.

Davis, T. E., & Principe, J. C. 1991. A simulated annealing like convergence theory for the simple genetic algorithm. *In:* Belew, R. K., & Booker, L.B. (eds), *Proceedings of the fourth international conference on genetic algorithms*. Morgan Kaufmann.

De Hoop, M.V., & Bleistein, N. 1997. Generalised Radon Transform inversions for reflectivity in anisotropic media. *Inv. Problems*, **13**, 669–690.

Eshelman, L.J., Caruana, R.A., & Schaffer, J.D. 1989. Bias in the crossover landscape. *Pages 10–19 of: Genetic algorithms and their applications: Proc. of the 3rd Int. Conf. on Genetic Algorithms*. Kauffmann Publishers Inc.

Frazer, L.N., & Basu, A. 1990. Freeze-bath inversion. *Pages 1123–1125 of: 60th Ann. Internat. Mtg., Soc. Expl. Geophys.*

Goldberg, D.E. 1989. *Genetic algorithms in search for optimization and machine learning*. Addison-Wesley Pub. Co.

Goldberg, D.E., & Segrest, P. 1987. Finite Markov chain analysis of genetic algorithms. *Pages 1–8 of: Genetic algorithms and their applications: Proc. of the 2nd Int. Conf. on Genetic Algorithms*.

Grechka, V., & Tsvankin, I. 1998. Feasibility of nonhyperbolic moveout inversion in transversely isotropic media. *Geoph.*, **63**, 957–969.

Grefenstette, J.J. 1986. Optimization of control parameters for genetic algorithms. *IEEE Trans. Systems, Man and Cybernetics*, **SMC-16**, 122–128.

Helbig, K., & Schoenberg, M. 1986. Anomalous polarization of elastic waves in transversely isotropic media. *J. Acoust. Soc. of Am.*, **81**, 1235–1245.

Henneke, E.G. 1972. Reflection-refraction of a stress wave at a plane boundary between anisotropic media. *J. Acoust. Soc. of Am.*, **51**, 210–217.

Holland, J.H. 1975. *Adaptation in Natural and Artificial Systems*. University of Michigan Press.

Jin, S., & Madariaga, R. 1993. Background velocity inversion with a genetic algorithm. *Geophys. Res. Lett.*, **20**, 93–96.

Keith, C.M., & Crampin, S. 1977. Seismic body waves in anisotropic media: Reflection and refraction at a plane interface. *Geophys. J. R. A. S.*, **49**, 181–208.

Kim, K.Y., Wrolstad, K.H., & Aminzadeh, F. 1993. Effects of transverse isotropy on P-wave AVO for gas sands. *Geoph.*, **58**, 883–888.

Lomax, A., & Snieder, R. 1994. Finding sets of acceptable solutions with a genetic algorithm with application to surface wave group dispersion in Europe. *Geophys. Res. Lett.*, **21**, 2617–2620.

Mallick, S. 1995. Model-based inversion of amplitude-with-offset data using a genetic algorithm. *Geoph.*, **60**, 939–954.

Michalewicz, Z., & Janikow, C. 1991. Handling constraints in Genetic Algorithms. *Pages 151–157 of: Genetic algorithms and their applications: Proc. of the 4th Int. Conf. on Genetic Algorithms*.

Nolte, B., & Frazer, L.N. 1994. Vertical seismic profile inversion with genetic algorithms. *Geophys. J. Int.*, **117**, 162–178.

Rüger, A. 1996. Variation of P-wave reflectivity with offset and azimuth in anisotropic media. *Pages 1810–1813 of: 66th Ann. Internat. Mtg., Soc. Expl. Geophys.*

Sambridge, M., & Drijkoningen, G. 1992. Genetic algorithms in seismic waveform inversion. *Geophys. J. Int.*, **109**, 323–342.

Scales, J.A., Smith, J.A., & Fischer, T.L. 1992. Global optimization methods for multimodal inverse problems. *J. Comp. Phys.*, **103**, 258–268.

Sen, M.K., & Stoffa, P.L. 1992. Rapid sampling of model space using genetic algorithms: Examples from seismic waveform inversion. *Geophys. J. Int.*, **108**, 281–292.

Sen, M.K., & Stoffa, P.L. 1996. Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion. *Geophys. Prosp.*, **44**, 313–350.

Stoffa, P.L., & Sen, M.K. 1991. Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane-wave seismograms. *Geoph.*, **56**, 1794–1810.

Tarantola, A. 1987. *Inverse Problem Theory*. Elsevier.

Thomsen, L. 1986. Weak elastic anisotropy. *Geoph.*, **51**, 1954–1966.

Tierney, L. 1994. Markov chains for exploring posterior distributions. *The Annals of Statistics*, 1701–1728.

Tsvankin, I. 1996. Effective parameters and P-wave velocity for azimuthally anisotropic media. *Pages 1850–1853 of: 66th Ann. Internat. Mtg., Soc. Expl. Geophys.*

Whitley, D., Mathias, K., & Stork, C. 1995. Global optimization for geophysical applications using genetic algorithms. *Col. State University*, **CASI-TR-95-09**.

Wright, J. 1987. The effect of transverse isotropy on reflection amplitude versus offset. *Geoph.*, **52**, 564–567.

FIGURE CAPTIONS

FIG. 1. Linearized (dashed) versus exact (solid) reflection coefficients for qP waves in both symmetry planes.

FIG. 2. Linearized contributions to the qP-qP reflection coefficients in a symmetry plane.

FIG. 3. Amplitude data contaminated with (zero mean) random noise at two different azimuths, dots at 15 (or 0) deg and triangles at 105 (or 90) deg (solid and dashed lines represent the associated noise-free data).

FIG. 4. Approximate marginal PPD for the vertical P-wave velocity from noise-free narrow-angle data.

FIG. 5. Approximate marginal PPD for the vertical P-wave velocity from narrow-angle data contaminated with noise.

FIG. 6. Approximate marginal PPD for $\epsilon^{(2)}$ from noise-free narrow-angle data.

FIG. 7. Approximate marginal PPD for $\epsilon^{(2)}$ from noise-free wide-angle data.

FIG. 8. Approximate marginal PPD for $\epsilon^{(2)}$ from narrow-angle data contaminated with noise.

FIG. 9. Approximate marginal PPD for $\epsilon^{(2)}$ from wide-angle data contaminated with noise.

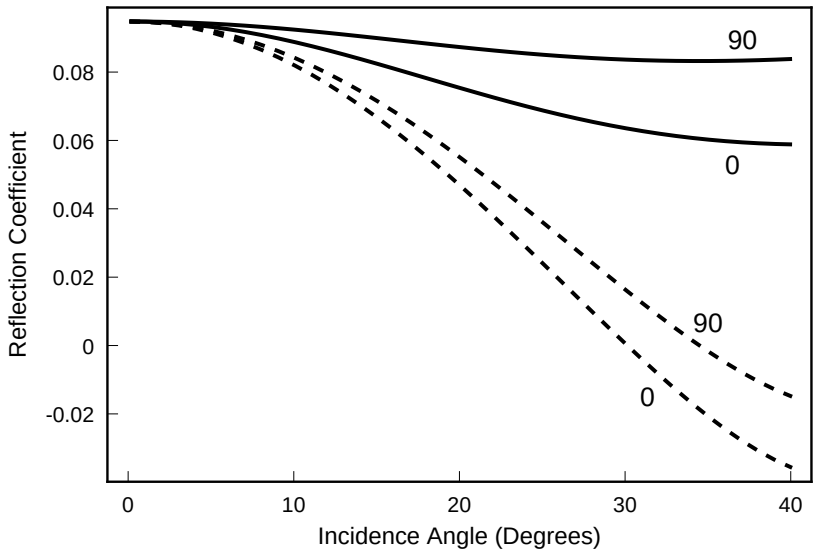

FIG. 1. Linearized (dashed) versus exact (solid) reflection coefficients for qP waves in both symmetry planes.

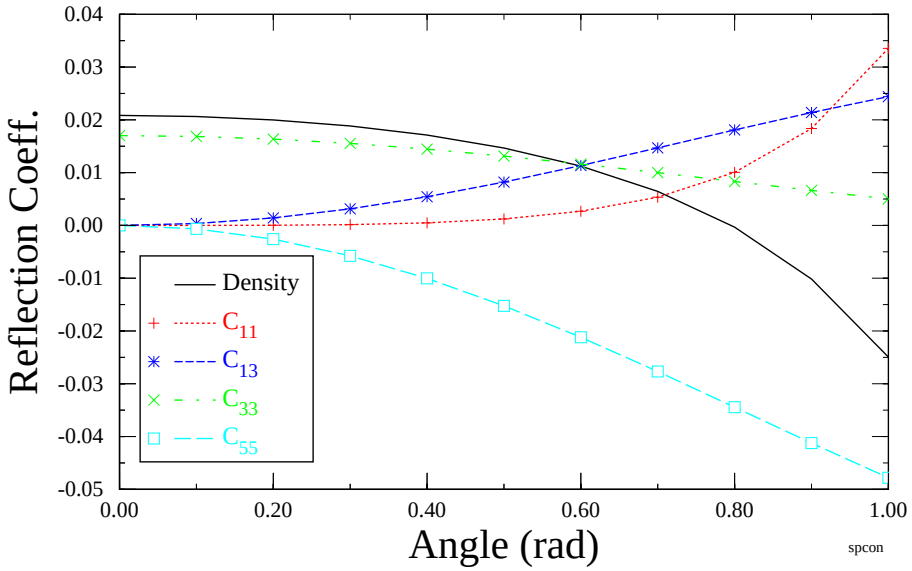

FIG. 2. Linearized contributions to the qP-qP reflection coefficients in a symmetry plane.

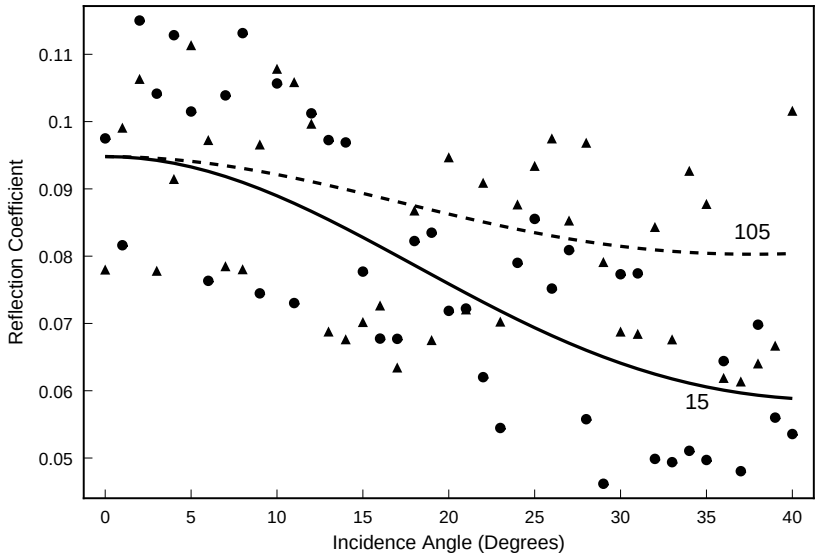

FIG. 3. Amplitude data contaminated with (zero mean) random noise at two different azimuths, dots at 15 (or 0) deg and triangles at 105 (or 90) deg (solid and dashed lines represent the associated noise-free data).

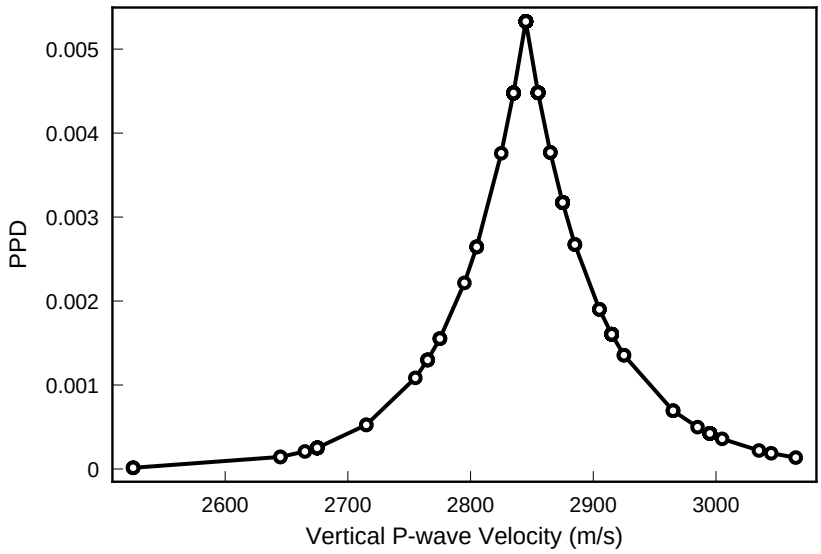

FIG. 4. Approximate marginal PPD for the vertical P-wave velocity from noise-free narrow-angle data.

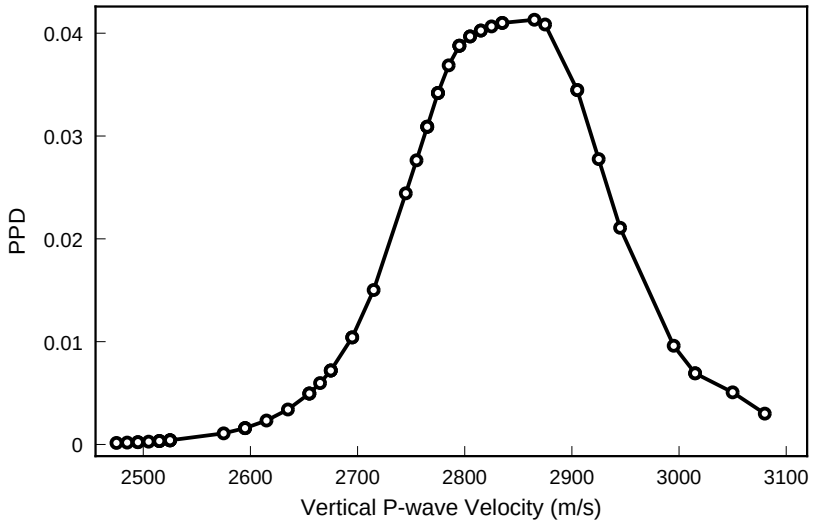

FIG. 5. Approximate marginal PPD for the vertical P-wave velocity from narrow-angle data contaminated with noise.

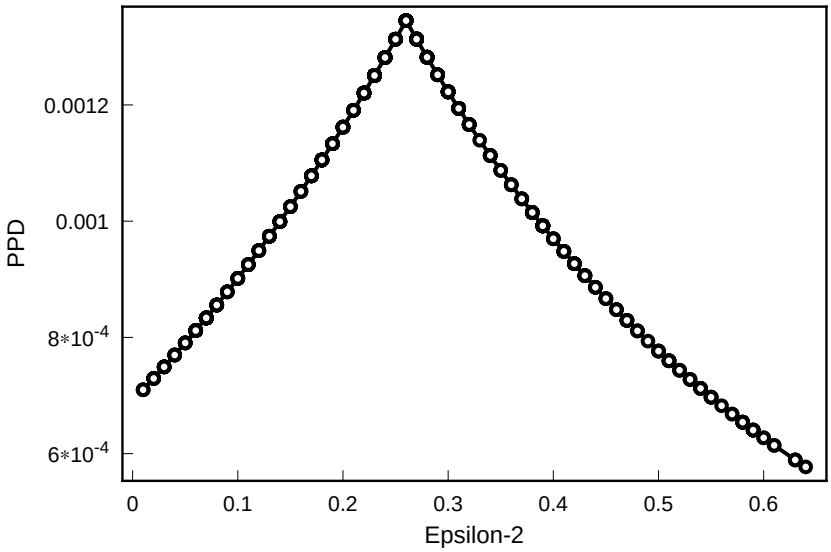

FIG. 6. Approximate marginal PPD for $\epsilon^{(2)}$ from noise-free narrow-angle data.

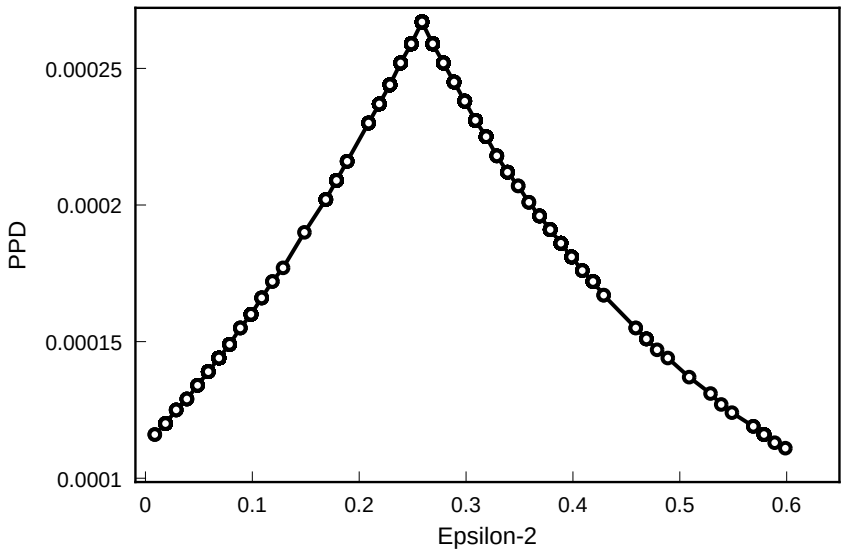

FIG. 7. Approximate marginal PPD for $\epsilon^{(2)}$ from noise-free wide-angle data.

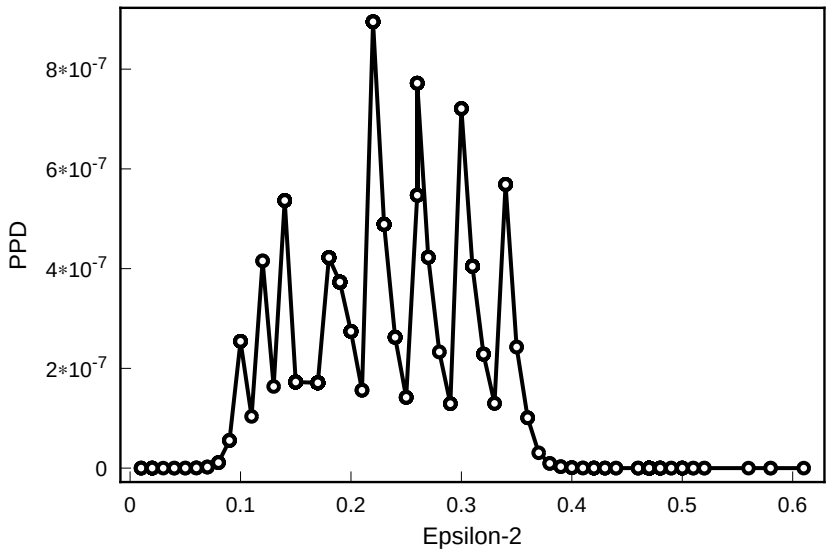

FIG. 8. Approximate marginal PPD for $\epsilon^{(2)}$ from narrow-angle data contaminated with noise.

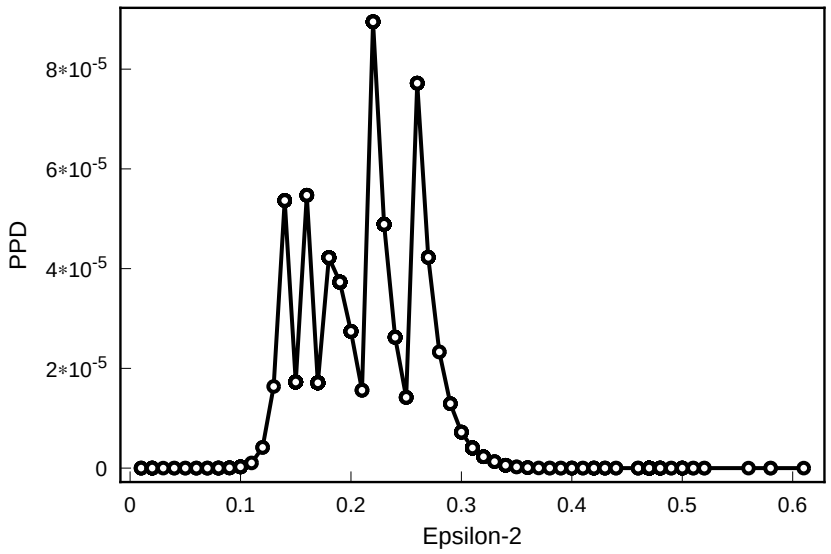

FIG. 9. Approximate marginal PPD for $\epsilon^{(2)}$ from wide-angle data contaminated with noise.

	C_{11}	C_{12}	C_{13}	C_{22}	C_{23}	C_{33}	C_{44}	C_{55}	C_{66}
VTI (upper)	17.35	0	6.75	C_{11}	C_{13}	10.71	3.08	C_{44}	4.12
ortho (lower)	27.00	10.72	6.70	29.56	7.29	17.81	5.94	4.80	6.53

Table 1. Shale and sand parameters in GPa.

	α	β	$\delta^{(1)}$	$\delta^{(2)}$	γ^{S}	ψ_C	$\epsilon^{(1)}$	$\epsilon^{(2)}$	$\delta^{(3)}$	C_{66}
actual	2845	1475	0.08	-0.08	0.12	15	0.33	0.26	-0.11	6.53
minimum	2000	1000	-0.20	-0.20	0.00	0	0.00	0.00	-0.20	5.00
maximum	4000	2000	0.20	0.20	0.40	180	0.40	0.40	0.20	7.00

Table 2. Sand parameter bounds.

α	misfit	β	$\delta^{(1)}$	$\delta^{(2)}$	γ^{S}	ψ_C	misfit
2850	3.9375e-05	1490	0.07	-0.09	0.14	15	2.6734e-03
2850	3.9375e-05	1450	0.09	-0.07	0.15	15	4.6293e-05
2845	0.0000e-00	1485	0.12	-0.08	0.13	15	5.1322e-05
2845	2.7554e-19	1465	0.11	-0.09	0.15	15	2.9698e-04
2840	5.7767e-04	1470	0.09	-0.08	0.12	15	7.6464e-05
2850	3.9375e-05	*1475*	*0.08*	*-0.08*	*0.12*	*15*	5.5325e-13
2850	3.9375e-05	1495	0.10	-0.07	0.14	15	6.6464e-04
2845	2.7554e-19	1475	0.12	-0.09	0.10	15	3.1178e-02
2840	5.7767e-04	1485	0.10	-0.07	0.13	15	2.8585e-05
2845	2.7554e-19	1500	0.06	-0.08	0.11	15	2.5354e-02

Table 3. Inversions from narrow-angle noise-free data.

$\epsilon^{(1)}$	$\epsilon^{(2)}$	$\delta^{(3)}$	misfit	C_{66}	misfit
0.33	0.27	-0.11	2.6204e-04	6.51	3.6816e-05
0.32	0.25	-0.11	5.5170e-04	6.52	1.1458e-07
0.32	0.26	-0.11	4.9498e-04	6.57	3.8099e-05
0.33	*0.26*	*-0.11*	9.6855e-12	6.49	1.5380e-05
0.34	0.25	-0.11	5.9089e-04	6.54	7.5396e-07
0.32	0.27	-0.11	8.8066e-04	6.56	7.6417e-05
0.31	0.26	-0.11	9.7855e-04	*6.53*	1.0586e-19
0.35	0.26	-0.11	1.0265e-05	6.51	2.6490e-06
0.31	0.25	-0.11	1.5284e-05	6.61	1.8112e-04
0.34	0.26	-0.11	3.0697e-04	6.52	1.1458e-07

Table 4. Inversions from wide-angle noise-free data using the values of Table 3.

	α	β	$\delta^{(1)}$	$\delta^{(2)}$	γ^{S}	ψ_C	misfit
	2840	1460	0.14	-0.07	0.14	15	3.7647e-02
	2830	1455	0.15	-0.08	0.13	15	2.3138e-01
	2850	1430	0.13	-0.08	0.15	15	4.6485e-01
	2860	1435	0.09	-0.09	0.14	15	7.6109e-02
$\rightarrow$	2845	1475	0.09	-0.08	0.12	15	2.1185e-10
	2870	1450	0.11	-0.08	0.13	15	5.9932e-02
	2860	1485	0.11	-0.09	0.14	15	4.9736e-03
	2865	1495	0.12	-0.09	0.13	15	1.1850e-01
	2855	1515	0.13	-0.07	0.17	15	3.0077e-02
	2845	1510	0.16	-0.08	0.15	15	2.1698e-01

Table 5. Inversions from narrow-angle data contaminated with noise.

	$\epsilon^{(1)}$	$\epsilon^{(2)}$	$\delta^{(3)}$	C_{66}	misfit
	0.31	0.27	-0.11	6.45	3.5354e-01
	0.36	0.25	-0.10	6.34	4.1185e-01
$\rightarrow$	0.33	0.26	-0.11	6.53	4.1323e-10
	0.33	0.28	-0.11	6.49	3.3790e-01
	0.30	0.27	-0.11	6.60	8.7729e-02
	0.35	0.25	-0.10	6.57	3.8357e-01
	0.33	0.25	-0.11	6.31	4.9017e-01
	0.37	0.26	-0.11	6.53	2.9540e-01
	0.31	0.26	-0.12	6.50	5.9812e-04
	0.33	0.27	-0.11	6.59	6.8464e-08

Table 6. Inversions from wide-angle data contaminated with noise using the values of Table 3.