Stanford Exploration Project, Report 70, December 18, 1997, pages 1–200

Short Note

Model decomposition with Walsh functions in non-linear traveltime inversion

Carlos A. Cunha Filho[1]

keywords: *modeling, inversion, nonlinear, traveltime, algorithm*

INTRODUCTION

In Cunha (1990) I used a set of *sine*-like and *cosine*-like square functions to describe the velocity model in a traveltime inversion scheme for elliptically anisotropic media. This specific decomposition of the model allowed the implementation of a fast non-linear algorithm that retrieves a different frequency-component of the model at each step. A drawback of this decomposition is that the basic set is neither orthogonal nor complete. Although the non-orthogonality can be somewhat useful (in this nonlinear scheme) for correcting errors in the components estimated by previous steps, the lack of completeness strongly reduces the resolution of the algorithm. It is possible to use instead another set of square functions – the Walsh functions – that form an orthogonal and complete set inside the interval where the model is defined. However, as illustrated in a synthetic example, the results obtained with the Walsh functions are inferior to the results obtained with the use of the *sin*- and *cosine*-like square functions. In addition, both results are inferior to those attained by a linearized inversion that uses the individual layers as the basis parameters for describing the model.

INVERSION SCHEMES

For a medium that can be described (or approximated) by elliptically anisotropic velocities (with vertical and horizontal symmetry axes), the traveltime between two points separated by a vertical distance z and a horizontal distance x is given by

$$t^2 = M^x \, x^2 + M^z \, z^2, \tag{1}$$

where M^x and M^z are the squared horizontal and vertical slownesses. If the medium is heterogeneous, this equation is valid either within each cell or within each layer

[1]**email:** not available

(if a layered description is appropriate) of the model. The inversion problem can be formulated as the search for the model (M_j^x, M_j^z), for all layers j, that minimizes the objective function

$$\mathcal{F} = \sum_{i=1}^{N} (t_i - t_i')^2\,, \tag{2}$$

where N is the number of source-receiver pairs, t_i is the measured traveltime corresponding to source-receiver index i, and t_i' is the traveltime predicted by the perturbed model.

Nonlinear schemes

In a previous report (Cunha, 1990) I used a set of *sine-* and *cosine-* like square functions to describe the model:

$$\mathbf{M} = c + \sum_{j=1}^{N} a_j\,\mathcal{S}_j + \sum_{j=1}^{N} b_j\,\mathcal{C}_j. \tag{3}$$

This choice of model decomposition led to an iterative perturbation scheme that estimates different components of the model at each iteration, starting with the homogeneous solution and successively increasing the frequency of the perturbation. The predicted traveltimes are given by

$$\begin{aligned} t_i' &= \sum_{j=J_s}^{J_r} o_j \sqrt{(M_j^x + s_j\, dM_1^x) x_{ij}^2 + (M_j^z + s_j\, dM_1^z) z_{ij}^2}\ + \\ &\quad e_j \sqrt{(M_j^x + s_j\, dM_2^x) x_{ij}^2 + (M_j^z + s_j\, dM_2^z) z_{ij}^2}, \end{aligned} \tag{4}$$

where o_j, e_j, and s_j are defined as follows:

$$o_j = \frac{1-(-1)^j}{2}, \qquad e_j = \frac{1+(-1)^j}{2}, \qquad s_j = (-1)^{(j+3)/2}. \tag{5}$$

In the above equations j is the layer index, J_s and J_r are the indices of the source and receiver layers, x_{ij} and z_{ij} are the horizontal and vertical distances traveled inside the layer j by the ray associated with source-receiver i. Using equation (4) in the objective function (2), the inversion problem can be easily solved with a nonlinear optimization algorithm (like the Downhill Simplex method) since only four parameters are inverted at each step: $dM_1^x, dM_2^x, dM_1^z, dM_2^z$.

Linearized Inversion

Equation (1) can be linearized for small perturbations around the homogeneous solution as follows:

$$t_i^0 = \sum_{j=J_s}^{J_r} t_{ij}^0, \qquad \text{where} \qquad t_{ij}^0 = \sqrt{M_j^x x_{ij}^2 + M_j^z z_{ij}^2}$$

$$
\begin{aligned}
t'_i &= \sum_{j=J_s}^{J_r} \sqrt{(M^x_j + dM^x_j)x^2_{ij} + (M^z_j + dM^z_j)z^2_{ij}} \\
t'_i &\approx t^0_i + \sum_{j=J_s}^{J_r} \frac{\partial t^0_i}{\partial M^x_j} dM^x_j + \sum_{j=J_s}^{J_r} \frac{\partial t^0_i}{\partial M^z_j} dM^z_j
\end{aligned} \tag{6}
$$

Using this approximation in equation (2) and minimizing it with respect to all dMs leads to the following set of linear equations:

$$
\begin{pmatrix} \mathbf{y^x} \\ \mathbf{y^z} \end{pmatrix} = \begin{pmatrix} \mathbf{A^{xx}} & \mathbf{A^{xz}} \\ \mathbf{A^{zx}} & \mathbf{A^{zz}} \end{pmatrix} \begin{pmatrix} \mathbf{dM^x} \\ \mathbf{dM^z} \end{pmatrix},
$$

where each element $\mathbf{A}^{\eta\zeta}$ is an $L \times L$ matrix, and $\mathbf{y}^\eta$ and $\mathbf{dM}^\zeta$ are column vectors with L elements, L being the number of layers (or cells) of the model. The elements of the matrix and vectors are given by the equations

$$
\begin{aligned}
y^x_j &= \sum_i \frac{(t_i - t^0_i)x^2_{ij}}{t^0_{ij}}, & y^z_j &= \sum_i \frac{(t_i - t^0_i)z^2_{ij}}{t^0_{ij}} \\
A^{xx}_{jk} &= \sum_i \frac{x^2_{ij} x^2_{ik}}{t^0_{ij} t^0_{ik}}, & A^{xz}_{jk} &= \sum_i \frac{x^2_{ij} z^2_{ik}}{t^0_{ij} t^0_{ik}} \\
A^{zz}_{jk} &= \sum_i \frac{z^2_{ij} z^2_{ik}}{t^0_{ij} t^0_{ik}}, & A^{zx}_{jk} &= A^{xz}_{kj}.
\end{aligned} \tag{7}
$$

Michelena (1991) solved the linearized problem using a conjugate gradients routine. He obtained a stable solution by stopping the process after a prespecified number of iterations, thus avoiding the problems arising from the ill-conditioning of matrix **A**. Another way to build a more stable algorithm is to introduce an angle-dependent weighting factor into equation (2) and to apply a weak constraint to the inversion:

$$
\mathcal{F} = \sum_{i=1}^{N} \left(\frac{t_i - t'_i}{t_i} \tan \alpha_i \right)^2 + \lambda^2 \sum_{j=1}^{L} \left(\frac{dM_j}{M_j} \right)^2, \tag{8}
$$

where α_i is the angle of the ray as measured from the horizontal.

WALSH FUNCTIONS

The Walsh functions form an orthogonal and complete set of functions representing a discretized function (Rao, 1983). As in the FFT, the discrete function must be padded before decomposition, so that the total number of samples reaches the nearest power of two. Figure **??**a shows the first sixteen Walsh functions, which form a complete set describing sequences with sixteen samples. Two important characteristics of the Walsh functions are their compactness (representing the lower order functions requires fewer samples), and the simplicity and quickness of their construction.

The following rules are used to generate the subset of Walsh functions corresponding to order **N**:

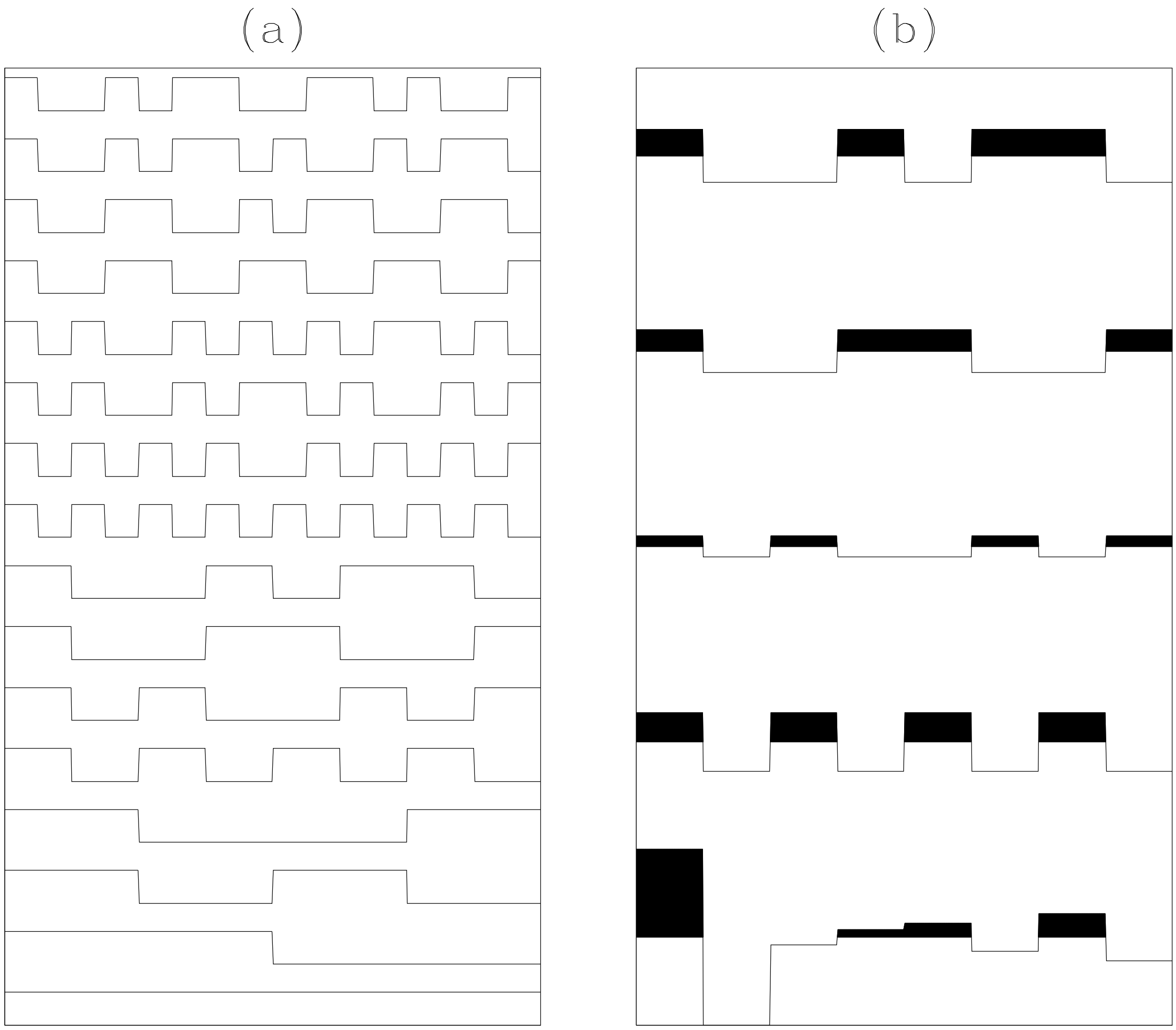

Figure 1: *(a)* The first sixteen Walsh functions. The frequency (order number) increases from bottom to top: W^0, W^1, $W^{2,1}$, $W^{2,2}$, $W^{3,1}$ to $W^{3,4}$, and $W^{4,1}$ to $W^{4,8}$. *(b)* Superposition of the four Walsh functions of order 3 to form the wavelet transform component of the same order. carlos1-walsh [NR]

- The number of samples of the functions in this subset is 2^N.
- Except for order 0 (zero frequency), the number of Walsh functions is given by

$$k_N = 2^{N-1}$$

.

- Each function of order **N-1** generates two functions of order **N**, one by contraction and repetition, the other by contraction and repetition with a change of sign. Thus

$$W_i^{N,2j-1} = W_{i+k_N}^{N,2j-1} = W_i^{N-1,j}, \qquad i = 1, k_N$$

$$W_i^{N,2j} = -W_{i+k_N}^{N,2j} = W_i^{N-1,j}, \qquad i = 1, k_N,$$

where i is the sample index and j is the index of one of the functions of order **N**.

Beginning with the functions of order 0 (W^0 =1) and 1 (W^1 =1,-1), it is possible to generate a complete set by using this iterative process.

There is a direct relationship between Walsh functions and the particular wavelet transform described by Ottolini (1990), which had a two-sample signal function as the basis wavelet. In contrast to the Walsh set, the wavelet set has only one component of order **N**, but the number of parameters associated with this component is equal to the number of Walsh functions of the same order. Figure **??***b* shows that adding all Walsh components of a given order results in the wavelet component of the same order. In general, any wavelet component can be obtained as a superposition of all the Walsh components of same order.

The Walsh functions have some properties that make them suitable for certain geophysical applications. Lanning and Johnson (1983) used the Walsh decomposition to develop an automated algorithm to identify rock boundaries in well-log data. They claim that the discrete transitions in signal level of these functions make them ideal for describing layered media. Figure **??***a* shows a sonic-log from the Gulf of Mexico and its Walsh spectrum, while Figure **??***b* compares the sonic smoothed with a triangle filter, with a band-limited (order 0 to 6 for a maximum order of 12) Walsh composition of the same sonic-log. Both curves show the same basic features, but the Walsh composition manifest a "blocky" structure that can be better correlated with the geologic layers.

In traveltime inversion the advantage of using Walsh instead of Fourier decomposition lies in the simplicity of the forward modeling computation, since the samples of the basis functions can only assume values 1 or -1, and only a few samples are required to describe the lower frequency components. Replacing square functions by Walsh functions in the nonlinear scheme described in the preceding section leads to a solution similar to the set of equations (7).

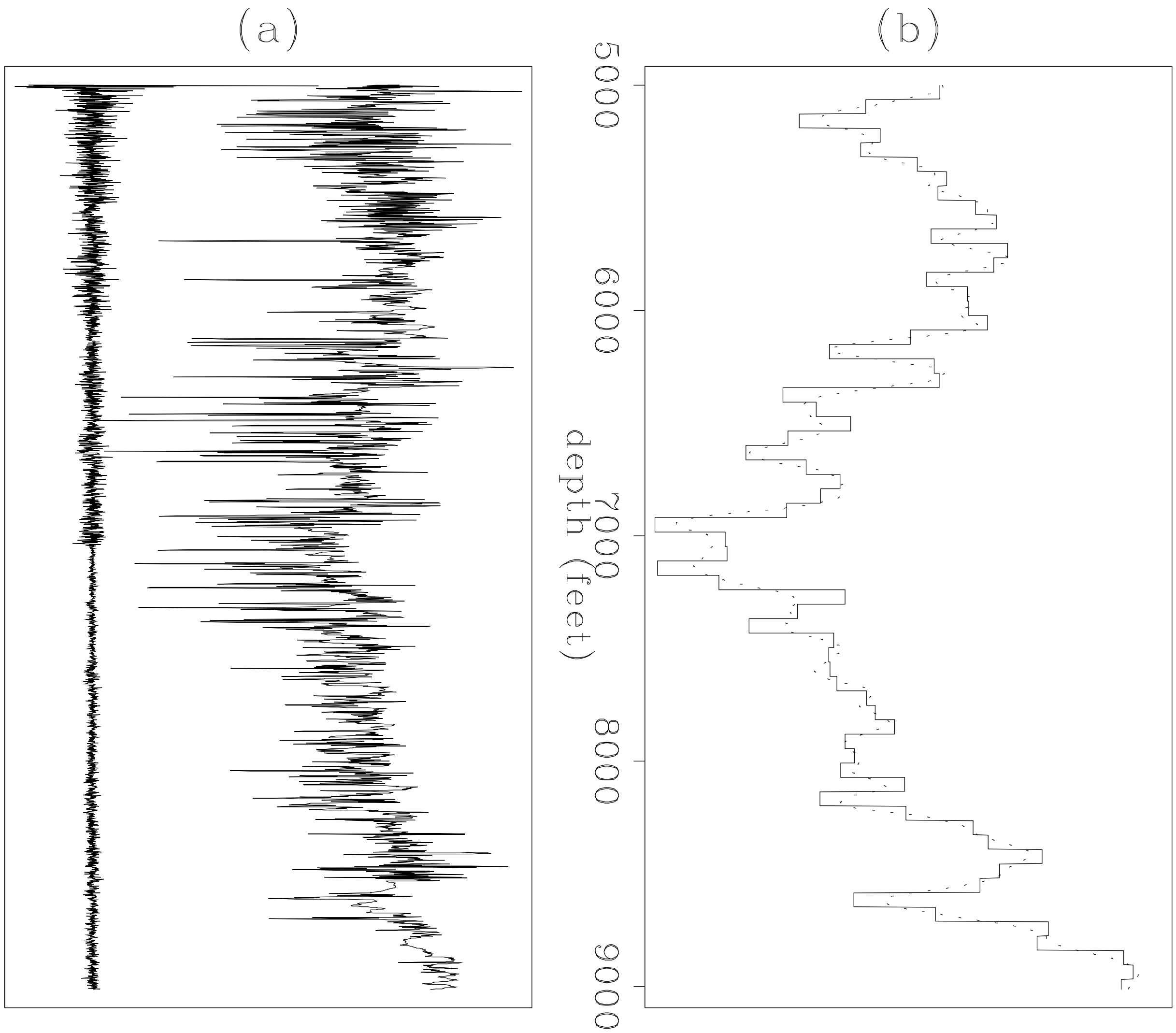

Figure 2: *(a)* A sonic log (left), and its Walsh spectrum (right). The zero frequency component was removed. *(b)* A Walsh band-limited version (continuous line), and smoothed version (dotted line) of the sonic log. carlos1-well [NR]

COMPARING THE DIFFERENT SCHEMES

I use the same simple synthetic model as in Cunha (1990) to compare the results obtained with different inversion schemes. The model consists of an isotropic layer in an isotropic background ($V^z = V^x$ in both the layer and the background). The layer is not centered, and its size is not a multiple of the layer size used in the inversion schemes.

Figures **??***a* and **??***b* show that, contrary to what we would expect, the results obtained with the Walsh functions are inferior to the ones obtained with the *sine*- and *cosine*-like square functions. Although Walsh functions form a complete set, their orthogonality impedes further steps from correcting errors in previously estimated components of the model in this kind of iterative scheme. Moreover, both nonlinear schemes are inferior to the linearized algorithms shown in Figures **??***c* and **??***d*. The result in **??**c was obtained with an iterative scheme that uses the least-squares solution of equation (8) by successively increasing (by a power of 2) the number of layers at each step. The difference between **??**c and **??**d is the introduction into the objective function of an extra term that corresponds to the difference between the traveltime-derivatives relative to the receiver positions. As a result the solution for V^z in **??**d becomes more stable since with this term the inversion will try to fit not only the traveltimes, but also the form of the traveltime curves.

CONCLUSIONS

Solving the nonlinear traveltime problem in successive steps of increasing resolution, through a frequency decomposition of the model space, has diverse limitations that depend on the specific choice of velocity basis functions. Walsh functions form an orthonormal complete set for representing discrete functions on a finite interval but, because of their orthogonality, the estimation of components associated with later steps of the algorithm cannot correct the errors in the previously estimated components. Using *sine*- and *cosine*-like square functions allow for these corrections, but the resulting set is not complete and the resolution is therefore (severely) limited. For the specific synthetic model tested in this study, the direct linearized solution proves to be much more stable, and to have a better resolution than the iterative nonlinear solutions based on the two types of square functions decomposition.

ACKNOWLEDGMENTS

Francis Muir was helpful in many constructive discussions that were essential for the development of this work.

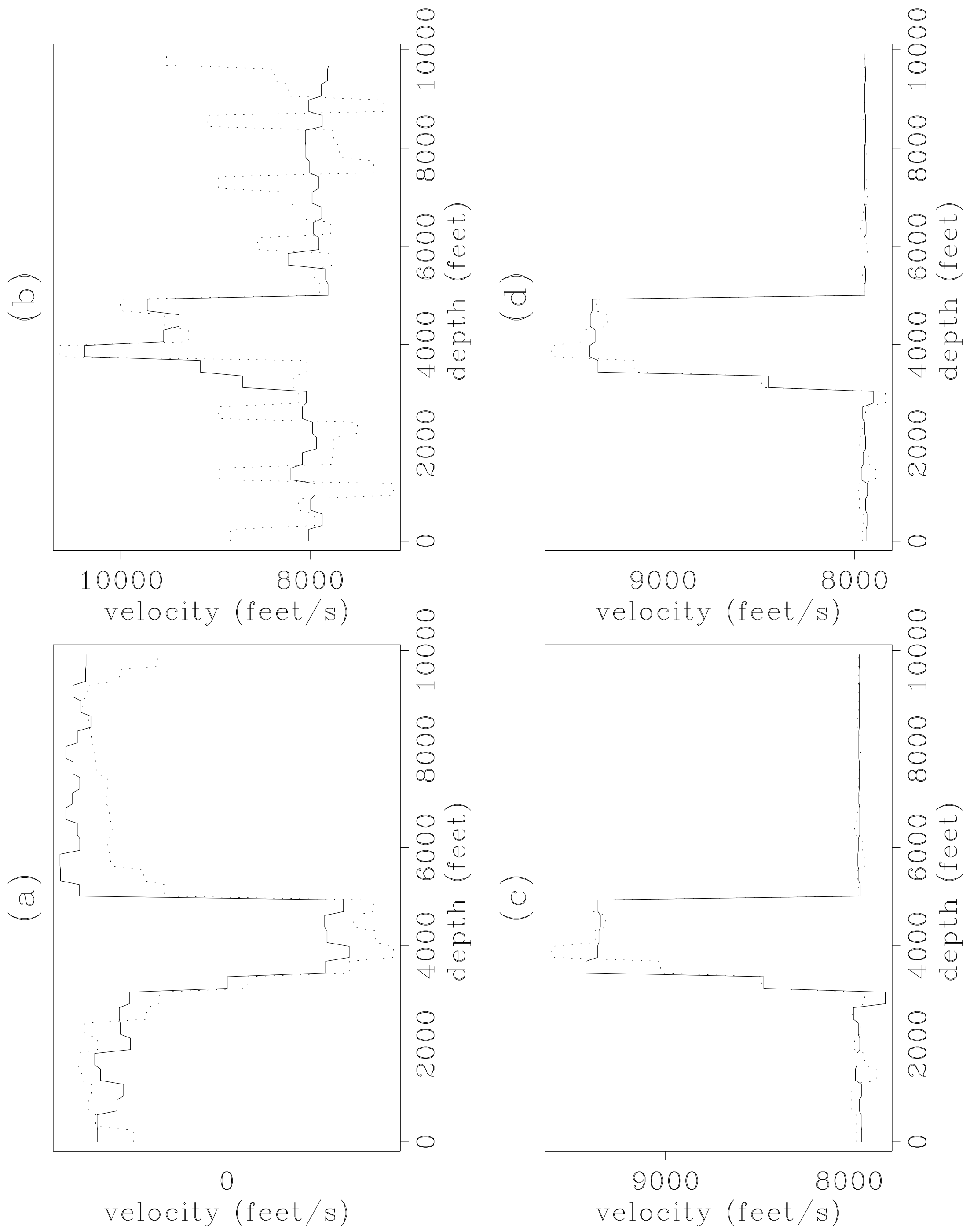

Figure 3: Inversion results for the *one-layer isotropic model*, using different algorithms: *(a)* square function decomposition (nonlinear), *(b)* Walsh function decomposition (nonlinear), *(c)* linearized inversion with layer-parametrization, and *(d)* the same as in *c*, but with the derivative term introduced into the objective function.**continuous** line represents V_x and **dotted** line represents V_z. carlos1-isoinv [NR]

REFERENCES

Cunha, C.A., 1990, Traveltime inversion of a cross-well dataset for elliptically anisotropic media: SEP–**67**, 31-40.

Lanning, E.N., and Johnson, D.M., 1983, Automated identification of rock boundaries: An application of the Walsh transform to geophysical well-log analysis: Geophysics, **48**, 197-205.

Michelena, R., and Muir, F., 1991, Anisotropic tomography: SEP–**70**, 201-226.

Ottolini, R., 1990, Wavelet decomposition of seismic data: SEP–**65**, 295-304.

Rao, G.P., 1983, Piecewise constant orthogonal functions and their application to systems and control: Springer-Verlag, Berlin.