

APPLICATION OF OPTIMAL BASIS FUNCTIONS IN FULL WAVEFORM INVERSION*

PING SHENG[†], GANG SUN^{†‡}, AND QIANSHUN CHANG^{†§}

Abstract. In full waveform inversion, the lack of low frequency information in the inversion results has been a long standing problem. In this work, we show that by using mixed basis functions this problem can be resolved satisfactorily. Examples of full waveform inversion on layered systems, using surface reflection data from point sources, have shown excellent results nearly indistinguishable from the target model. Our method is robust against additive white noise (up to 20% of the signal) and can resolve layers that are comparable to or smaller than a wavelength in thickness. Physical reason for the success of our approach is illustrated through a simple example.

Key words. Inversion, basis functions

Personal statement by P. S. *I wish to give tribute to Prof. George Papanicolaou on the occasion of his 60th birthday. George is a great mathematician and a good friend. His generosity, strong sense of right and wrong and a deeply-felt sense of humanity have affected positively all those around him. It is especially appropriate for me to choose the topic of layered media inversion for this volume, since I have learned about the mathematics of layered media from George. I wish him a happy birthday and many productive years to come.*

Wave inversion means the recovery of the coefficients/parameters of the wave equation from its solution(s). Except for some cases of linear inversions (e.g., in 1D inversion) [1, 2, 3, 4], however, most nonlinear inversions still present considerable difficulties. A simplest example of nonlinear inversion, with geophysical relevance, is that of layered media, with point sources and receivers on the surface. In such (seismic) inversions the aim is to obtain the interface (reflector) positions as well as the layer velocities. Together they define Earth's subsurface structures.

The layered inversion is intrinsically ill-posed for deep layers. In addition, there is the problem of so-called "low frequency lacuna", i.e., a lack of low (spatial) frequency information in the inversion results. This problem is understandable physically because wave is a differential detector (reflection only from sharp changes in material properties, i.e., interfaces), implying no or very little information on the slow spatial variations (low frequency information) from the reflection data. In this paper, we show that the use of mixed basis functions in full waveform inversion can effectively solve the low frequency lacuna problem, and lead to accurate inversion results starting from a uniform background. In what follows we present our approach based on the framework proposed by Tarantola [5] and Gauthier et al. [6].

Full waveform inversion for a 2D layered acoustic system, with point sources and receivers, perhaps constitutes the simplest test case for nonlinear inversion and a first-order approximation to the structure of the Earth's subsurface. By using the optimal basis functions alternately in the inversion process, we show that not only the low frequency lacuna problem can be resolved, but also robust and accurate results were obtained [7]. Physical reason for this success is elucidated by comparing 1D analytic

*Received July 20, 2004; accepted for publication November 17, 2004.

[†]Department of Physics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China.

[‡]Institute of Physics, Chinese Academy of Sciences, Beijing, China.

[§]Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing, China.

solutions at a single interface, for both the step function basis and the block function basis.

The pressure field $P^s(\vec{x})$ of our model satisfies the equation

$$\frac{\rho(\vec{x})}{\kappa(\vec{x})} \frac{\partial^2 P^s(\vec{x}, t)}{\partial t^2} - \rho(\vec{x}) \nabla \cdot \left[\frac{1}{\rho(\vec{x})} \nabla P^s(\vec{x}, t) \right] = s(\vec{x}, t), \quad (1)$$

where \vec{x} and t denote position and time, respectively, and $s(\vec{x}, t)$ is the source. The model is characterized by bulk modulus $\kappa(\vec{x})$ and density $\rho(\vec{x})$, with velocity $\nu(\vec{x}) = \sqrt{\kappa(\vec{x})/\rho(\vec{x})}$. We limit ourselves to uniform density ($\rho = 1$), thus there is only one model parameter in the system, i.e., the bulk modulus or the velocity, varying from layer to layer as a function of depth y , i.e., $\nu(\vec{x}) = \nu(y)$. The nonlinearity of the inversion problem remains under this restriction [8, 9, 10], so this simplified problem still retains the most interesting difficulties.

We use the misfit function

$$S[\nu(y)] = \sum_{r,s,t} \delta P_{r,s,t}^2 = \sum_{r,s,t} [P^s(\vec{x}_r, t | \nu_m(y)) - P^s(\vec{x}_r, t | \nu(y))]^2, \quad (2)$$

where $P^s(\vec{x}_r, t | \nu_m(y))$ is the measured pressure at receiver position \vec{x}_r (with the source on the surface, denoted by superscript s), $\nu_m(y)$ is the target velocity profile, $P^s(\vec{x}_r, t | \nu(y))$ is the calculated pressure at \vec{x}_r for the current model, characterized by $\nu(y)$, and δP denotes the residual between the measured and calculated pressures. In Eq. (2) the summation indices r, s, t stand for receivers, sources, and time, respectively. Inversion in this case is equivalent to minimizing the misfit function with respect to $\nu(y)$.

One of the most generally used methods in minimizing the misfit function is the damped least square approach, in which the correction to the velocity, $\delta\nu(y)$, is evaluated according to the equation $(A^T A + \gamma I) \delta\vec{u} = A^T \delta P$, with $\delta\nu(y) = \sum_i \delta u_i b_i(y)$.

Here $b_i(y)$ denotes the i th basis function, A is a $m \times n$ matrix, A^T is its transpose, γ is a damping coefficient, n is the number of model parameters and $m = n_r \times n_s \times n_t$ denotes the product of the numbers of receivers, source and time steps. The matrix elements of A in the (r, s, t) th row and i th column are given by $A(\{r, s, t\}, i) = \delta P^s(\vec{x}_r, t) / \partial u_i$, defined below. In this work we are concerned mainly with three sets of basis functions: the step function basis $b_i(y) = b(y_i, y) = H(y - y_i)$, where $H = 1$ if the argument is positive, and $H = 0$ otherwise; the quasi- δ function (block function) basis, given by $b_i(y) = b(y_i, y) = H(y - y_i)H(y_i + \delta y - y)$; and the linear basis $b_i(y)$, defined to be $b_i(y) = b(y_i, y) = y - y_i$ if $y > y_i$, and 0 otherwise.

In all the prior full waveform inversions, the partial derivatives were calculated with respect to the usual block function basis [1, 11, 12, 13, 14, 15], i.e., the delta function in the continuum case. Here we want to generalize the calculation of partial derivatives with other basis functions as well. That is, for an arbitrary basis function, if $\nu(y)$ is changed to $\nu(y) + \delta u_i b_i(y)$, then the partial derivative with respect to the basis function $b_i(y)$ is given by [4]

$$\frac{\delta P^s(\vec{x}_r, t)}{\delta u_i} = -2 \int_0^t \int_V \frac{\dot{G}(\vec{x}', t - t'; \vec{x}_r) b_i(y') \dot{P}^s(\vec{x}', t')}{\nu(y')} d\vec{x}' dt'. \quad (3)$$

For completeness, the derivation of Eq. (3) is given in the Appendix. Here we observe that only $P^s(\vec{x}, t)$ for all sources, and $G(\vec{x}, t - t'; \vec{x}_r)$ for all receivers are needed to

obtain all the matrix elements of A . Thus the calculation of the matrix elements can be very efficient.

Steepest descent is a special case of the damped least square method when $\gamma \gg \|A^T A\|$. In that limit we have $\delta u(y_i) \propto \sum_{r,s,t} [\partial P^s(\bar{x}_r, t) / \partial u(y_i)] \delta P_{r,s,t}$, where $\delta u(y_i) = \delta u_i$. We now use an exactly solvable 1D model [16, 17, 18] (that is, use normally incident plane wave as the source instead of point sources) to elucidate the rationale for optimal basis selection. Consider an interface at $y = a > 0$, where $\nu(y) = \nu_0$ for $y < a$, and $\nu(y) = \nu_1$ for $y \geq a$. A single pulse, in the form of $[\cos(2\pi(y - \nu_0 t)/d) + 1]^4$ in the region of $-1/2 < (y - \nu_0 t)/d < 1/2$ and 0 otherwise, is incident on the interface at $y = a$. From the reflected signal retrieved at the receiver located at $y = 0$, the aim is to recover the model profile, starting from a uniform initial model with ν_0 . If the pressure signal of the incident wave at the receiver is denoted by $p_s(t)$, then the difference between the measured and calculated pressures is simply the reflected signal, $\delta P(t) = \gamma p_s(t - 2a/\nu_0)$, where $\gamma = (\nu_1 - \nu_0)/(\nu_1 + \nu_0)$. The partial derivative in the block function (quasi- δ function) basis is

$$\frac{\partial P}{\partial u(y')} = \frac{1}{2\nu_0} \left[p_s\left(t - \frac{2y'}{\nu_0}\right) - p_s\left(t - \frac{2y'}{\nu_0} - \frac{2\delta y}{\nu_0}\right) \right]. \quad (4)$$

The resulting $\delta u(y')$, for the case $\nu_1 > \nu_0$, is plotted in Fig. 1(a) by the dashed line. It is seen that the correction is oscillatory in the neighborhood of $y = a$, with a net integrated area of zero. The corresponding velocity correction, given by $\delta \nu(y) = \int \delta u(y') b(y', y) dy'$, shown by the dashed line in Fig. 1(b), bears no resemblance to a step function. Thus it would be difficult, if not impossible, to recover the true velocity profile by using the block function basis. This is the underlying cause of the low frequency lacuna problem encountered before.

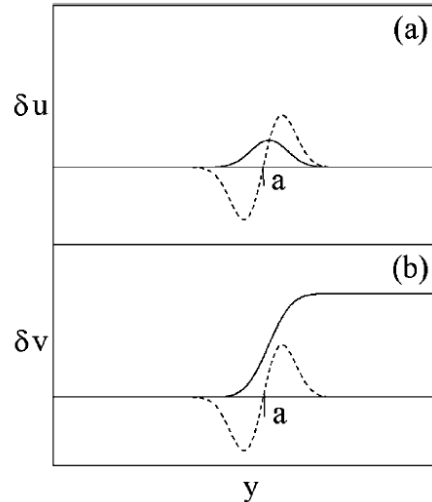


FIG. 1. The velocity correction calculated with respect to the block function (quasi- δ function) basis (dash line) and the step basis function (solid line) basis for the 1D system. The amplitudes $\delta u(y_i)$ are shown in (a); the corresponding velocity corrections $\delta v(y)$ are shown in (b) [7].

The reason for the oscillatory correction is easily traced to the interference of waves scattered by y' and $y' + \delta y$. To suppress the interference, we switch to the

step-function basis in the calculation of the derivatives as stipulated by Eq. (3). Then the partial derivative is given by

$$\frac{\partial P(t)}{\partial u(y')} = \frac{1}{2\nu_0} p_s(t - \frac{2y'}{\nu_0}). \quad (5)$$

The resulting $\delta u(y')$ is also shown in Fig. 1(a) (solid line). It leads to the corresponding velocity correction shown by solid line in Fig. 1(b), seen to be close to a step function at $y = a$. The step function basis is thus more efficient in recovering the target.

In real systems, a point source generally generates an oscillatory waveform, and a more complex form of $\delta u(y)$ than that shown in Fig. 1 would result. In these cases we identify the relevant $\delta u(y)$ (calculated from the step function basis) to be only those part(s) which are greater in magnitude than 80% the maximum of $|\delta u(y)|$. The rest of $\delta u(y)$ was set to zero. In other words, we used only the largest part(s) of the correction.

To avoid biasing the inversion process, we alternately use block function basis and the step function basis. While the block function basis has its deficiency as demonstrated above, yet it can recover the positions of the interfaces efficiently. Thus an inversion process involving the alternate use of both basis functions (plus the linear basis as well if the model has piecewise linear regions) would be optimal in recovering the interface positions as well as the layer velocities.

Figure 2 shows a layered 2D acoustic model with a velocity versus depth (defined as $y > 0$) profile given by the dotted line. Point sources and receivers were used. Forward calculations used sixth-order finite difference on a 80×80 grid to generate the data. The pressure release boundary condition was used on the upper surface $y = 0$. For the other boundaries wave reflection was minimized by adding an additional damping layer. The Ricker wavelet was used as the source pulse: $s(t) \propto [1 - 2\pi^2(t - t_s)^2/t_p^2] \exp(-2\pi^2(t - t_s)^2/t_p^2)$, with $t_p = 12.0$ and $t_s = 20.0$ in our time unit. Ten source positions and ten receiver positions were interlaced and evenly distributed on the surface.

The initial model is a constant profile with the velocity equal to that at the surface. It was first updated by the damped least square method using the step function basis, iterated ten times. The result is shown in Fig. 2(a) (dash-dotted line). The convergence was found to be much faster than that by using the block function basis. To obtain the positions of the deeper interfaces, we further updated the current model by using ten iterations using the block function basis. The new result is shown in Fig. 2(a) by the dashed line. It is seen that there are small ‘‘incorrect’’ variations. We introduced a threshold, small compared to interfacial velocity jumps, to further update the model. That is, only those velocity jumps larger than the threshold were retained. This re-configuration process led to the result shown by the solid line in Fig. 2(a). The previous three processes constitute an ‘‘inversion unit.’’ The results after two, four, and ten iterations of the inversion unit are shown by dash-dotted, dashed, and solid lines in Fig. 2(b), respectively. The convergence to the target configuration is observed to be from the surface downward, as expected from intuition. The final result is almost indistinguishable from the target model, and our inversion has resolved layers with thicknesses comparable to or smaller than the wavelength. We have confirmed the robustness of our method for several target models. In Fig. 3(a), we show the results for a model with a high bump in the middle. The final inversion result is also excellent.

The above inversion approach is slightly modified for a model shown in Fig. 3(b), consisting of a piecewise linear region mixed with piecewise constant regions. Now in

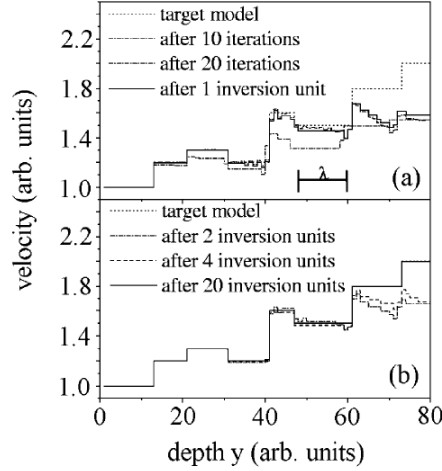


FIG. 2. Results at different stages of the inversion process for a 2D layered model. The target model is shown by the dotted line in both (a) and (b). In (a), the result after first 10 iterations of the damped least square with the step basis function is shown by the dash-dotted line. The result after 10 further iterations by using the block function basis is shown by the dashed line. The solid line shows the result after a whole inversion unit. In (b), the results after two, four, and ten inversion units are shown by the dash-dotted, dashed, and the solid lines, respectively. Definition of an inversion unit is described in the text. Convergence to the target model is seen to initiate at the surface and proceed downward [7].

our inversion unit ten iterations of the linear basis function precedes the rest of the steps in the inversion unit defined above. Otherwise the process remains the same. It is seen that even in this case, the model can still be accurately recovered starting from a uniform background. Our inversion is also robust with respect to additive white noise. Appreciable deviations from the target model appear only with white noise amplitude greater than $\sim 20\%$ of the signal.

In order to relate our models to the geophysical context, we note that the p-wave velocity in sandstone and shale ranges from 0.8-3.4 km/sec. If we take the value of 3 km/sec as the mean value and frequency ~ 50 Hz, then in Fig. 2 the depth is only about 0.45 km, or 0.3 sec in terms of 2-way travel time. Typical seismic sections can be 3 seconds or more in terms of the 2-way travel time. Hence for realistic applications the efficiency of the present approach must be further optimized. One potential avenue that might enhance the inversion efficiency is by exploiting the causal nature of the reflected signals. In fact, from the convergence pattern of our model calculations, i.e., from top layer proceeding downward, it is already clear that inverting the layer parameters successively could be fruitful.

Appendix. If we write $P^s(\vec{x}, t|\nu(y) + \delta\nu(y)) = P^s(\vec{x}, t|\nu(y)) + \delta P^s(\vec{x}, t)$, then under the Born approximation, $\delta P^s(\vec{x}, t)$ satisfies the equation

$$\Im \delta P^s(\vec{x}, t) = \frac{\partial^2 \delta P^s(\vec{x}, t)}{\partial t^2} - \nu^2(y) \Delta \delta P^s(\vec{x}, t) = 2\nu(y) \delta\nu(y) \Delta P^s(\vec{x}, t|\nu(y)). \quad (\text{A1})$$

This equation can be solved by using the Green function, i.e., by defining $G(\vec{x}, t - t'; \vec{x}')$ as the solution of $\Im G(\vec{x}, t - t'; \vec{x}') = \delta(\vec{x} - \vec{x}') \delta(t - t')$ so that

$$\delta P^s(\vec{x}, t) = 2 \int_0^t \int_V G(\vec{x}, t - t'; \vec{x}') \nu(y') \delta\nu(y') \Delta P^s(\vec{x}', t') d\vec{x}' dt', \quad (\text{A2})$$

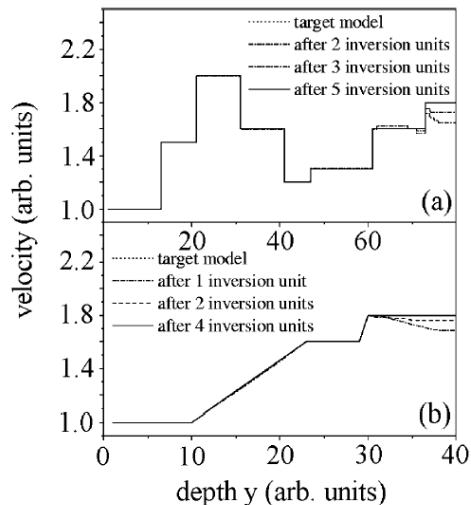


FIG. 3. Results at different stages of the inversion process for another 2D layered target model (a), where there is a bump in the middle, and a mixed 2D piecewise linear and piecewise constant target model (b). The initial model for both is the constant profile [7].

where $\vec{x}' = (x', y', z')$. For real $G(\vec{x}, t - t'; \vec{x}')$, it is a property of the Green function that $G(\vec{x}, t - t'; \vec{x}') = G(\vec{x}', t - t'; \vec{x})$. Thus we have

$$\delta P^s(\vec{x}_r, t) = 2 \int_0^t \int_V G(\vec{x}', t - t'; \vec{x}_r) \nu(y') \delta \nu(y') \Delta P^s(\vec{x}', t') d\vec{x}' dt'. \quad (\text{A3})$$

If we restrict $\delta \nu(\vec{x}_s) = 0$, i.e., no velocity correction at the source position, then

$$\delta P^s(\vec{x}_r, t) = 2 \int_0^t \int_V G(\vec{x}', t - t'; \vec{x}_r) \frac{\delta \nu(y')}{\nu(y')} \frac{\partial^2 P^s(\vec{x}', t')}{\partial t'^2} d\vec{x}' dt' \quad (\text{A4})$$

$$= -2 \int_0^t \int_V \frac{\dot{G}(\vec{x}', t - t'; \vec{x}_r) \delta \nu(y') \dot{P}^s(\vec{x}', t')}{\nu(y')} d\vec{x}' dt'. \quad (\text{A5})$$

Here the over dot denotes time derivative. Noting $\delta \nu(y) = \delta u_i b_i(y)$ then yields the desired Eq. (3).

REFERENCES

- [1] R. BURRIDGE AND G. PAPANICOLAOU AND P. SHENG AND B. WHITE, *SIAM J. Appl. Math.*, 49 (1989), p. 582.
- [2] K. P. BUBE AND R. BURRIDGE, *SIAM Rev.*, 25 (1983), p. 497.
- [3] R. BURRIDGE, *Wave Motion*, 12 (1990), p. 15.
- [4] G. BEYLKIN AND R. BURRIDGE, *Wave Motion*, 12 (1990), p. 15.
- [5] A. TARANTOLA, *Geophys.*, 48 (1984), p. 1259.
- [6] O. GAUTHIER AND J. VIRIEUX AND A. TARANTOLA, *Geophys.*, 51 (1986), p. 1387.
- [7] G. SUN AND Q. CHANG AND P. SHENG, *Phys. Rev. Lett.*, 90 (2003), p. 104301.
- [8] F. JURADO AND M. CUER AND V. RICHARD, *Geophys.*, 60 (1995), p. 11857.
- [9] A. J. BERKHOUT, *Proc. IEEE*, 4 (1986), p. 415.
- [10] M. JANNANE AND W. BEYDOUN AND E. CRASE AND D. CAO AND Z. KOREN AND E. LANDA AND M. MENDES AND A. PICA AND M. NOBLE AND G. ROETH AND S. SINGH AND R. SNIEDER AND A. TARANTOLA AND D. TREZEGUET AND M. XIE, *Geophys.*, 54 (1989), p. 906.

- [11] A. TARANTOLA, *Geophys.*, 51 (1986), p. 1893.
- [12] A. TARANTOLA, *Pure Appl. Geophys.*, 128 (1988), p. 365.
- [13] R. J. GELLER AND T. OHMINATO, *Geophys. J. Int.*, 116 (1994), p. 421.
- [14] R. J. GELLER AND T. HARA, *Geophys. J. Int.*, 115 (1993), p. 699.
- [15] T. HARA AND S. TSUBOI AND R. J. GELLER, *Geophys. J. Int.*, 115 (1993), p. 667.
- [16] J. G. BERRYMAN, *Lecture notes on nonlinear inversion and tomography*, Lawrence Livermore National Laboratory, California, 1991.
- [17] R. BURRIDGE AND G. S. PAPANICOLAOU AND B. S. WHITE, *Wave Motion*, 10 (1988), p. 19.
- [18] R. BURRIDGE AND H. W. CHANG, *Wave Motion*, 11 (1989), p. 231.

