

## Full Waveform Inversion with Optimal Basis Functions

Gang Sun,<sup>1,2</sup> Qianshun Chang,<sup>1,3</sup> and Ping Sheng<sup>1,\*</sup>

<sup>1</sup>*Department of Physics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China*

<sup>2</sup>*Institute of Physics, Chinese Academy of Sciences, Beijing, China*

<sup>3</sup>*Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing, China*

(Received 26 March 2002; published 12 March 2003)

Based on the approach suggested by Tarantola, and Gauthier *et al.*, we show that the alternate use of the step (linear) function basis and the block function (quasi- $\delta$  function) basis can give accurate full waveform inversion results for the layered acoustic systems, starting from a uniform background. 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. The physical reason for the success of our approach is illustrated through a simple example.

DOI: 10.1103/PhysRevLett.90.104301

PACS numbers: 43.90.+v, 91.30.-f

Wave inversion means the recovery of the coefficients/parameters of the wave equation from its solution(s). It is one of the most important problems in physical sciences. However, except for some cases of linear inversions (e.g., in 1D inversion) [1–4], most nonlinear inversions still present considerable difficulties. In seismic inversions, involving the imaging of Earth's deep subsurface structures, the traveltime inversion is the most tractable [5]. A more ambitious inversion approach is that of full waveform inversion, in which a common strategy is to retrieve the model parameters by minimizing a misfit function. While simple in concept, the success of full waveform inversion has been rather limited because of the extensive computational requirement and difficulty in realizing target convergence. Some time ago, Tarantola [6] proposed a scheme of full waveform inversion for acoustic systems, which was subsequently implemented by Gauthier *et al.* [7]. If only reflection data were used, the method can map the high spatial frequency components of the model, e.g., the interfaces, but totally fails in recovering the low spatial frequency information, e.g., the layer velocities. This so-called low frequency lacuna problem is rather well known in other seismic imaging approaches as well. Later literature on full wave inversion also met similar convergence problems [8–12]. The use of global minimizing processes [13–16] has shown much better convergence characteristics, but at the expense of computational efficiency. Symes [17] proposed a smooth and convex misfit function by using model parameters which are linear in the inversion calculation. This method has some successful applications [18,19]. Another proposal is to improve the initial model by using the hybrid inversion method, which minimizes a weighted combination of first-arrival traveltime and seismogram misfit functions [20,21].

In this Letter, we consider the full waveform inversion for a 2D layered acoustic system, with point sources and receivers, perhaps the simplest test case for nonlinear inversion and a first-order approximation to the structure

of Earth's subsurface. By using the optimal basis functions alternately in the inversion process, we not only overcame the low frequency lacuna problem encountered before, but also obtained robust and accurate results. The physical reason for this success is elucidated by comparing 1D analytic solutions at a single interface, for both the step-function basis and the block function basis.

The pressure field  $P^s(\vec{x}, t)$  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  $v(\vec{x}) = \sqrt{\kappa(\vec{x})/\rho(\vec{x})}$ . In this work, we limit ourselves to uniform density ( $\rho = 1$ ) and a layered model, 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.,  $v(\vec{x}) = v(y)$ . The nonlinearity of the inversion problem remains under this restriction [22–24], so this simplified problem still retains the most interesting difficulties.

We use the misfit function

$$\begin{aligned} S[v(y)] &= \sum_{r,s,t} \delta P_{r,s,t}^2 \\ &= \sum_{r,s,t} \{P^s[\vec{x}_r, t|v_m(y)] - P^s[\vec{x}_r, t|v(y)]\}^2, \quad (2) \end{aligned}$$

where  $P^s[\vec{x}_r, t|v_m(y)]$  is the measured pressure at receiver position  $\vec{x}_r$  (with the source on the surface, denoted by superscript  $s$ ),  $v_m(y)$  is the target velocity profile,  $P^s[\vec{x}_r, t|v(y)]$  is the calculated pressure at  $\vec{x}_r$  for the current model, characterized by  $v(y)$ , and  $\delta P$  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. 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,  $\delta v(y)$ , is evaluated according to the equation  $(A^T A + \gamma I)\delta \vec{u} = A^T \delta P$ , with  $\delta v(y) = \sum_i \delta u_i b_i(y)$ . Here  $b_i(y)$  denotes the  $i$ th basis function,  $A$  is a  $m \times n$  matrix,  $A^T$  is its transpose,  $\gamma$  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 Letter, 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- $\delta$  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 [1,25–29]. For an arbitrary basis function, we first write  $P^s[\vec{x}, t|v(y) + \delta v(y)] = P^s[\vec{x}, t|v(y)] + \delta P^s(\vec{x}, t)$ . Under the Born approximation,  $\delta P^s(\vec{x}, t)$  satisfies the equation

$$\begin{aligned} \Im \delta P^s(\vec{x}, t) &= \frac{\partial^2 \delta P^s(\vec{x}, t)}{\partial t^2} - v^2(y)\Delta \delta P^s(\vec{x}, t) \\ &= 2v(y)\delta v(y)\Delta P^s[\vec{x}, t|v(y)]. \end{aligned} \quad (3)$$

This equation can be solved by using the Green's 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

$$\begin{aligned} \delta P^s(\vec{x}, t) &= 2 \int_0^t \int_V G(\vec{x}, t - t'; \vec{x}') v(y') \delta v(y') \\ &\quad \times \Delta P^s(\vec{x}', t') d\vec{x}' dt', \end{aligned} \quad (4)$$

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

$$\begin{aligned} \delta P^s(\vec{x}_r, t) &= 2 \int_0^t \int_V G(\vec{x}', t - t'; \vec{x}_r) v(y') \delta v(y') \\ &\quad \times \Delta P^s(\vec{x}', t') d\vec{x}' dt'. \end{aligned} \quad (5)$$

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

$$\begin{aligned} \delta P^s(\vec{x}_r, t) &= 2 \int_0^t \int_V G(\vec{x}', t - t'; \vec{x}_r) \frac{\delta v(y')}{v(y')} \\ &\quad \times \frac{\partial^2 P^s(\vec{x}', t')}{\partial t'^2} d\vec{x}' dt' \\ &= -2 \int_0^t \int_V \frac{\dot{G}(\vec{x}', t - t'; \vec{x}_r) \delta v(y') \dot{P}^s(\vec{x}', t')}{v(y')} \\ &\quad \times d\vec{x}' dt'. \end{aligned} \quad (6)$$

Here the overdot denotes time derivative. Consider changing  $v(y)$  to  $v(y) + \delta u_i b_i(y)$ . The partial derivative with

respect to the basis function  $b_i(y)$  is then given by

$$\frac{\delta P^s(\vec{x}_r, t)}{\delta u_i} = -2 \int_0^t \int_V \frac{\dot{G}(\vec{x}', t - t'; \vec{x}_r) b_i(y') \dot{P}^s(\vec{x}', t')}{v(y')} d\vec{x}' dt'. \quad (7)$$

It is seen 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. 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  $\delta u(y_i) \propto \sum_{r,s,t} [\partial P^s(\vec{x}_r, t)/\partial u(y_i)] \delta P_{r,s,t}$ , where  $\delta u(y_i) = \delta u_i$ . We now use an exactly solvable 1D model [5,30,31] to elucidate the rationale for optimal basis selection in the context of the steepest descent method. Consider an interface at  $y = a > 0$ , where  $v(y) = v_0$  for  $y < a$ , and  $v(y) = v_1$  for  $y \geq a$ . A single pulse, in the form of  $\{\cos[2\pi(y - v_0 t)/d] + 1\}^4$  in the region of  $-1/2 < (y - v_0 t)/d < 1/2$  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  $v_0$ . 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,  $\delta P(t) = \gamma p_s(t - 2a/v_0)$ , where  $\gamma = (v_1 - v_0)/(v_1 + v_0)$ . The partial derivative in the block function (quasi- $\delta$  function) basis is

$$\frac{\partial P}{\partial u(y')} = \frac{1}{2v_0} \left[ p_s \left( t - \frac{2y'}{v_0} \right) - p_s \left( t - \frac{2y'}{v_0} - \frac{2\delta y}{v_0} \right) \right]. \quad (8)$$

The resulting  $\delta u(y')$ , for the case  $v_1 > v_0$ , 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  $\delta v(y) = \int \delta u(y') b(y', y) dy'$ , 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.

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. (7). Then the partial derivative is given by

$$\frac{\partial P(t)}{\partial u(y')} = \frac{1}{2v_0} p_s \left( t - \frac{2y'}{v_0} \right). \quad (9)$$

The resulting  $\delta u(y')$  is also shown in Fig. 1(a) (solid line). It leads to the corresponding velocity correction shown by the 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.

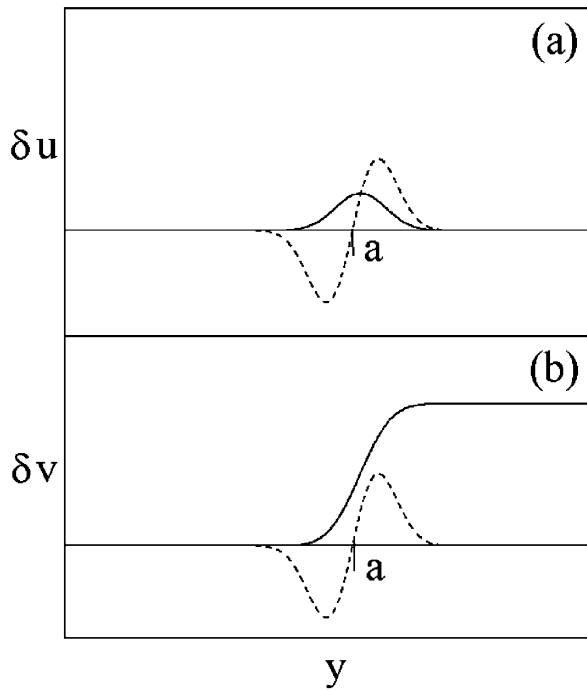


FIG. 1. The velocity correction calculated with respect to the block function (quasi- $\delta$  function) basis (dashed line) and the step basis function (solid line) basis for the 1D system. The amplitudes  $\delta u(y_i)$  are shown in (a); the corresponding velocity corrections  $\delta v(y)$  are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

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.

While the block function basis has its deficiency, 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 a sixth-order finite difference on a  $80 \times 80$  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.

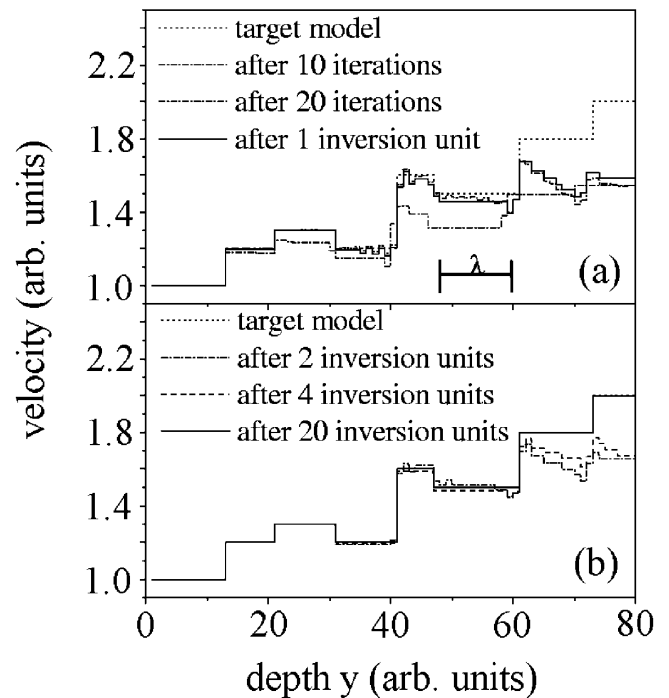


FIG. 2. Results at different stages of the inversion process for a 2D layered model. The target model is shown by the dotted line in both (a) and (b). In (a), the result after the first ten iterations of the damped least square with the step basis function is shown by the dash-dotted line. The result after ten 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. The 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.

The initial model is a constant profile with 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, the current model was scanned from the surface to the bottom, and 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 final result is almost indistinguishable from the target model,

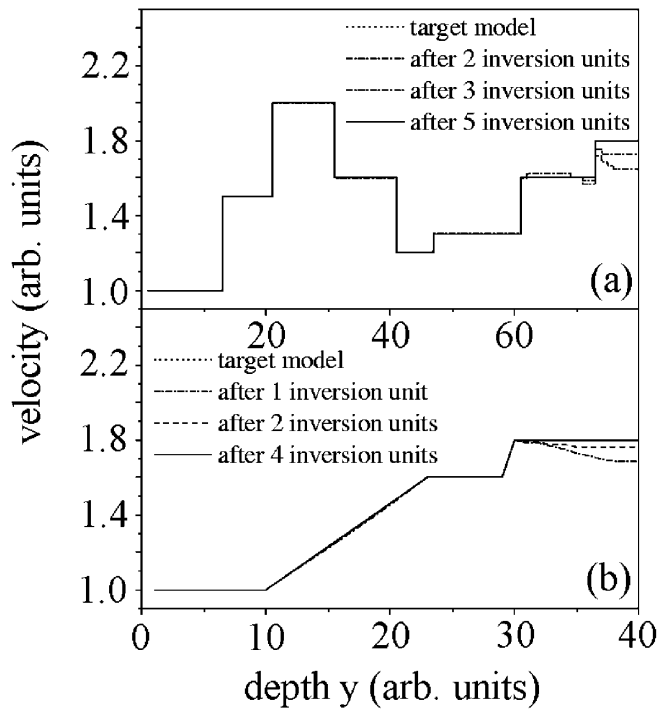


FIG. 3. Results at different stages of the inversion process for another 2D layered target model (a), where there is a bump in the middle, and a mixed 2D piecewise linear and piecewise constant target model (b). The initial model for both is the constant profile.

and our inversion has resolved layers with thicknesses comparable to or smaller than the wavelength, shown in Fig. 2(a). 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 our inversion unit ten iterations of the linear basis function precede 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.

We thank Professor Z. B. Su for his support in the initial stage of this research.

\*Author to whom correspondence should be addressed.

Email address: sheng@ust.hk

- [1] R. Burridge, G. Papanicolaou, P. Sheng, and B. White, *SIAM J. Appl. Math.* **49**, 582 (1989).
- [2] K. P. Bube and R. Burridge, *SIAM Rev.* **25**, 497 (1983).
- [3] R. Burridge, *Wave Motion* **2**, 305 (1980).
- [4] G. Beylkin and R. Burridge, *Wave Motion* **12**, 15 (1990).
- [5] J. G. Berryman, *Lecture notes on nonlinear inversion and tomography*, Lawrence Livermore National Laboratory, California, 1991.
- [6] A. Tarantola, *Geophys.* **48**, 1259 (1984).
- [7] O. Gauthier, J. Virieux, and A. Tarantola, *Geophys.* **51**, 1387 (1986).
- [8] A. Pica, J. P. Diet, and A. Tarantola, *Geophys.* **55**, 284 (1989).
- [9] P. Mora, *Geophys.* **52**, 1211 (1987).
- [10] P. Mora, *Geophys.* **53**, 750 (1988).
- [11] P. Mora, *Geophys.* **54**, 1575 (1989).
- [12] E. Landa, D. Kosloff, S. Keydar, Z. Koren, and M. Reshef, *Geophys. Prosp.* **36**, 223 (1988).
- [13] R. Snieder, M. Xie, A. Pica, and A. Tarantola, *Geophys.* **54**, 991 (1989).
- [14] D. Cao, W. Beydoun, S. Singh, and A. Tarantola, *Geophys.* **55**, 458 (1990).
- [15] M. K. Sen and P. L. Stoffa, *Geophys.* **56**, 1624 (1991).
- [16] P. L. Stoffa and M. K. Sen, *Geophys.* **56**, 1794 (1991).
- [17] W. W. Symes and J. J. Carazzone, *Geophys.* **56**, 654 (1991).
- [18] M. S. Gockenbach, W. W. Symes, and R. A. Tapia, *Inverse Probl.* **11**, 501 (1995).
- [19] S. E. Minkoff and W. W. Symes, *Geophys.* **62**, 540 (1997).
- [20] C. Zhou, W. Cai, Y. Luo, G. T. Schuster, and S. Hassanzadeh, *Geophys.* **60**, 765 (1995).
- [21] C. Zhou, G. T. Schuster, S. Hassanzadeh, and J. M. Harris, *Geophys.* **62**, 853 (1997).
- [22] F. Jurado, M. Cuer, and V. Richard, *Geophys.* **60**, 11857 (1995).
- [23] A. J. Berkhout, *Proc. IEEE* **4**, 415 (1986).
- [24] M. Jannane, W. Beydoun, E. Crase, D. Cao, Z. Koren, E. Landa, M. Mendes, A. Pica, M. Noble, G. Roeth, S. Singh, R. Snieder, A. Tarantola, D. Trezeguet, and M. Xie, *Geophys.* **54**, 906 (1989).
- [25] A. Tarantola, *Geophys.* **51**, 1893 (1986).
- [26] A. Tarantola, *Pure Appl. Geophys.* **128**, 365 (1988).
- [27] R. J. Geller and T. Ohminato, *Geophys. J. Intl.* **116**, 421 (1994).
- [28] R. J. Geller and T. Hara, *Geophys. J. Intl.* **115**, 699 (1993).
- [29] T. Hara, S. Tsuboi, and R. J. Geller, *Geophys. J. Intl.* **115**, 667 (1993).
- [30] R. Burridge, G. S. Papanicolaou, and B. S. White, *Wave Motion* **10**, 19 (1988).
- [31] R. Burridge and H. W. Chang, *Wave Motion* **11**, 231 (1989).