Geophys. J. Int. (2004) **157,** 975–987 doi: 10.1111/j.1365-246X.2004.02282.x

Frequency-domain finite-difference amplitude-preserving migration

R.-E. Plessix and W. A. Mulder

Shell International E&P, PO Box 60, 2280 AB Rijswijk, the Netherlands. E-mail: ReneEdouard.Plessix@shell.com

Accepted 2004 March 3. Received 2004 February 3; in original form 2003 April 14

SUMMARY

A migration algorithm based on the least-squares formulation will find the correct reflector amplitudes if proper migration weights are applied. The migration weights can be viewed as a pre-conditioner for a gradient-based optimization problem. The pre-conditioner should approximate the pseudo-inverse of the Hessian of the least-squares functional. Usually, an infinite receiver coverage is assumed to derive this approximation, but this may lead to poor amplitude estimates for deep reflectors.

To avoid the assumption of infinite coverage, new amplitude-preserving migration weights are proposed based on a Born approximation of the Hessian. The expressions are tested in the context of frequency-domain finite-difference two-way migration and show improved amplitudes for the deeper reflectors.

Key words: finite differences, frequency domain, migration, pre-conditioner.

1 INTRODUCTION

Migration algorithms produce the depth locations and relative amplitude behaviour of reflectors in the Earth from measured seismic data. The classic imaging principle (Claerbout 1971; Esmersoy 1986) for shot-based migration states that reflectors are located where the forward-propagated wavefield from the source correlates with the backward-propagated wavefield of the receiver data. To find a suitable approximation for the reflector amplitudes, the correlation is divided by the square of the incident field. This approach leads, among others, to the classic reversed-time migration algorithm (Baysal *et al.* 1983).

In the 1980s it was observed that the imaging principle can be obtained by formulating migration as an inverse problem based on a least-squares functional (Lailly 1983; Tarantola 1984; Beylkin 1985). This opened up research on true-amplitude or amplitude-preserving migration. Beylkin (1985) derived an amplitude-preserving migration algorithm by viewing migration as an inverse Radon transform under the assumption of infinite and continuous receiver coverage using high-frequency asymptotes. Lailly (1983) and Tarantola (1984) started with a least-squares error functional that measures the difference between observed and synthetic data. They found that the gradient of this functional with respect to the underlying model corresponds to a migration image. This approach has the advantage that it can be used independently of the numerical scheme that models the synthetic data and independently of the acquisition geometry. In this setting, amplitude-preserving migration is obtained by scaling the gradient by the pseudo-inverse of the Hessian of the least-squares functional. The least-squares approach differs from the imaging principle approach or Beylkin's approach because it can be based on the full data set, resulting in migration weights that are different for multishot multi-offset data. Also, irregular acquisition geometries can be taken into account.

The main difficulty with the least-squares approach is the size of the Hessian, the matrix of the second derivatives of the error functional with respect to the model parameters, which is too large to be directly used in practical applications because an approximation of the pseudo-inverse of this matrix is needed to estimate the migration weights. To avoid this problem, its pseudo-inverse may be approximated by a diagonal matrix (Docherty 1991; Gray 1997; Chavent & Plessix 1999; Shin *et al.* 2001). In the case of high-frequency asymptotics and with an infinite aperture, the Hessian is diagonal in most cases (Beylkin 1985; ten Kroode *et al.* 1994). For a finite range of frequencies, however, the Hessian is no longer diagonal and not even diagonally dominant (Chavent & Plessix 1999; Pratt *et al.* 1998). Therefore, the gradient of the least-squares functional pre-multiplied by a diagonal matrix cannot perfectly recover the true amplitudes. To solve this problem, an *iterative* migration algorithm can be used to find the correct amplitudes (Lambaré *et al.* 1992; Nemeth *et al.* 1999; Østmo *et al.* 2002).

Because iterative migration is relatively costly, it is still useful to derive the best possible approximation to the pseudo-inverse of the Hessian. In this paper, the approximations are derived for a finite-difference or finite-element scheme that solves the *linearized* wave equation. Instead of the common one-way approximations, we use the full or 'two-way' wave equation. The result differs from the asymptotic case (Docherty 1991; Chavent & Plessix 1999) because the Green's functions associated with the receivers are not directly computed when the two-way wave equation is solved by a finite-difference scheme. To avoid the costly computation of the Green's function for each receiver separately, we can make simplifying assumptions. If, for instance, we assume infinite receiver coverage we obtain a pre-conditioner or migration

weighting equivalent to the one in Shin *et al.* (2001). For a shot-based approach, the migration weighting is the same as the one obtained with the imaging principle or for reversed-time migration. Two new formulae are proposed in order to take into account finite receiver coverage and a discrete acquisition geometry.

The paper is organized as follows. First, the least-squares formulation is briefly reviewed. Then, using the Born approximation, analytical formulae for the Hessian are derived and several approximations of the pseudo-inverse are proposed. Finally, the resulting expressions are tested with a two-way finite-difference migration scheme in the frequency domain.

2 GENERAL STATEMENTS

In the frequency domain, the acoustic wave equation for constant density reads

$$(-\omega^2\sigma^2 - \Delta)u(\mathbf{x}_s, \mathbf{x}, \omega, \sigma) = f(\mathbf{x}_s, \mathbf{x}, \omega), \tag{1}$$

with Δ the Laplacian operator, σ the slowness, $u(\mathbf{x}_s, \mathbf{x}, \omega, \sigma)$ the pressure created by a point source at $\mathbf{x}_s$, $f(\mathbf{x}_s, \mathbf{x}, \omega)$ the source term and ω the angular frequency. The synthetics are obtained by projecting the pressure field onto the receiver positions $\mathbf{x}_r$:

$$c(\mathbf{x}_s, \mathbf{x}_r, \sigma, \omega) = P(\mathbf{x}_r, \mathbf{x})u(\mathbf{x}_s, \mathbf{x}, \omega, \sigma), \tag{2}$$

with $P(\mathbf{x}_r, \mathbf{x})$ a projection operator.

Following Tarantola's approach (Tarantola 1987), the imaging algorithm is formulated as a least-squares inverse problem. The error functional is

$$J(\sigma) = \frac{1}{2}\sum_{\omega}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}|c(\mathbf{x}_s, \mathbf{x}_r, \sigma, \omega) - d(\mathbf{x}_s, \mathbf{x}_r, \omega)|^2. \tag{3}$$

The migration operator corresponds to the opposite of the gradient of $J(\sigma)$ with respect to σ left-multiplied by a positive matrix (Lailly 1983; Tarantola 1984). The migration image, m, is

$$m = -K\nabla J(\sigma), \tag{4}$$

with K a positive matrix. The diagonal elements of K are called migration weights in this paper.

Migration aims to find the discontinuities of the model assuming a known smooth background. In a first-order approximation, the error functional, J, is quadratic in the discontinuities of the model. The minimum of a quadratic function is obtained by the gradient pre-multiplied by the (pseudo-) inverse of the Hessian. In this way, the optimal amplitude-preserving migration image is obtained by choosing K to be the pseudo-inverse of the Hessian of the error functional. Unfortunately, it is not feasible to compute the Hessian in practice because of its size. The common approach is to approximate the pseudo-inverse of the Hessian with a diagonal matrix. This is fully valid in the high-frequency limit (Beylkin 1985; ten Kroode *et al.* 1994) but may be too crude an approximation with a finite set of frequencies. Still, it can be useful, in particular for the finite-difference scheme considered in this paper.

Let ν be a parametrization of σ, so $\sigma = \sigma(\nu)$. The derivative of J with respect to ν is

$$\nabla_\nu J(\sigma)(\mathbf{x}) = \frac{\partial J(\sigma)}{\partial \nu(\mathbf{x})} = \sum_{\omega}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}\frac{\partial c(\mathbf{x}_s, \mathbf{x}_r, \sigma, \omega)}{\partial \nu(\mathbf{x})}[c(\mathbf{x}_s, \mathbf{x}_r, \sigma, \omega) - d(\mathbf{x}_s, \mathbf{x}_r, \omega)], \tag{5}$$

with $\nu(\mathbf{x})$ one of the variables describing the slowness model. The Hessian is

$$\frac{\partial^2 J(\sigma)}{\partial \nu(\mathbf{x})\partial \nu(\mathbf{y})} = \sum_{\omega}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}\left(\frac{\partial c(\mathbf{x}_s, \mathbf{x}_r, \sigma, \omega)}{\partial \nu(\mathbf{x})}\frac{\partial c(\mathbf{x}_s, \mathbf{x}_r, \sigma, \omega)}{\partial \nu(\mathbf{y})} + \frac{\partial^2 c(\mathbf{x}_s, \mathbf{x}_r, \sigma, \omega)}{\partial \nu(\mathbf{x})\partial \nu(\mathbf{y})}[c(\mathbf{x}_s, \mathbf{x}_r, \sigma, \omega) - d(\mathbf{x}_s, \mathbf{x}_r, \omega)]\right). \tag{6}$$

The second term in the summation of eq. (6) (after the plus sign) corresponds to the non-linear part of the modelling. This term is usually assumed to be small when the functional is close to its minimum, which is the so-called Gauss–Newton approach (Pratt *et al.* 1998; Shin *et al.* 2001). We also make this assumption, which can be interpreted as being in a multiple-free case. Now, the Hessian can be computed by linearization or Born approximation.

3 LINEARIZED CASE

We define

$$u = u_0 + \varepsilon u_1 \quad \text{and} \quad \sigma^2 = \sigma_0^2 + \varepsilon r. \tag{7}$$

The parameter ν, defined in the previous section, now corresponds to r. The zero- and first-order equations, obtained from the wave eq. (1), are

$$\left(-\Delta - \omega^2\sigma_0^2\right)u_0 = f \quad \text{and} \quad \left(-\Delta - \omega^2\sigma_0^2\right)u_1 = \omega^2 r u_0. \tag{8}$$

The Green's function, $G(\mathbf{x}_s, \mathbf{x}, \omega)$, is defined by

$$\left[-\Delta - \omega^2\sigma_0^2(\mathbf{x})\right]G(\mathbf{x}_s, \mathbf{x}, \omega) = \delta(\mathbf{x}_s - \mathbf{x}). \tag{9}$$

The incident field, u_0, at the subsurface point $\mathbf{x}$ becomes

$$u_0(\mathbf{x}_s, \mathbf{x}, \omega) = f_s(\omega) G(\mathbf{x}_s, \mathbf{x}, \omega), \tag{10}$$

and the secondary field, u_1, at the subsurface point $\mathbf{y}$ is

$$u_1(\mathbf{x}_s, \mathbf{y}, \omega) = \omega^2 \sum_{\mathbf{x}} r(\mathbf{x}) u_0(\mathbf{x}_s, \mathbf{x}, \omega) G(\mathbf{x}, \mathbf{y}, \omega), \tag{11}$$

with $r = (\sigma^2 - \sigma_0^2)$ for $\varepsilon = 1$, which is related to the reflectivity $(\sigma - \sigma_0)/\sigma_0$. For $\mathbf{y} = \mathbf{x}_r$, we obtain the Born approximation of the synthetic field at the receiver position, $\mathbf{x}_r$:

$$c(\mathbf{x}_s, \mathbf{x}_r, \omega) = \omega^2 \sum_{\mathbf{x}} r(\mathbf{x}) f_s(\omega) G(\mathbf{x}_s, \mathbf{x}, \omega) G(\mathbf{x}, \mathbf{x}_r, \omega). \tag{12}$$

The error functional is

$$J(\sigma) = \frac{1}{2} \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} |c(\mathbf{x}_s, \mathbf{x}_r, \omega) - d(\mathbf{x}_s, \mathbf{x}_r, \omega)|^2. \tag{13}$$

The gradient with respect to $v = r$ is

$$g(\mathbf{x}) = \nabla_r J(\sigma)(\mathbf{x}) = \mathrm{Re}\left(\sum_{\omega} \omega^2 \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} f_s(\omega) G(\mathbf{x}_s, \mathbf{x}, \omega) G(\mathbf{x}, \mathbf{x}_r, \omega) (\bar{c}(\mathbf{x}_s, \mathbf{x}_r, \omega) - \bar{d}(\mathbf{x}_s, \mathbf{x}_r, \omega)) \right), \tag{14}$$

and the second derivative with respect to $r(\mathbf{x})$ and $r(\mathbf{y})$ is

$$H(\mathbf{x}, \mathbf{y}) = \mathrm{Re}\left(\sum_{\omega} \omega^4 \sum_{\mathbf{x}_s} |f_s(\omega)|^2 G(\mathbf{x}_s, \mathbf{x}, \omega) \bar{G}(\mathbf{x}_s, \mathbf{y}, \omega) \sum_{\mathbf{x}_r} G(\mathbf{x}, \mathbf{x}_r, \omega) \bar{G}(\mathbf{y}, \mathbf{x}_r, \omega) \right). \tag{15}$$

For small models, it is feasible to compute a line of the Hessian by using the reciprocity $G(\mathbf{x}, \mathbf{x}_r, \omega) = G(\mathbf{x}_r, \mathbf{x}, \omega)$. We have computed three lines of the Hessian for a small 2-D salt dome model assuming a land-type acquisition geometry. Note that for a fixed point $\mathbf{x}$, each line $H(\mathbf{x}, \mathbf{y})$ of the Hessian can be mapped to a 2-D grid. Fig. 1 shows a velocity model with the three points for which the lines were computed. Point 1 has coordinates $x = 3$ km, $z = 1.25$ km, point 2 has $x = 3$ km, $z = 1.75$ km, and point 3 has $x = 3.75$ km, $z = 1.5$ km. We observe that the main contribution to the Hessian occurs on the diagonal, but is spread out into an elliptical shape due to finite-frequency effects. Note that the computational domain was larger than the domain used for the display of the line of the Hessian. This explains why the pictures are not symmetric around the diagonal point. The amount of spreading, or the size of the ellipse, depends on the depth of the point and on the background velocity. In this way, the ellipse around point 2 is larger than the one around point 1 because point 2 lies deeper. The ellipse around point 3 is smaller because point 3 lies in a low-velocity zone where the energy refocuses. These observations show that the size of the ellipse appears to be proportional to z^α and to $1/\sigma^\beta$, where z is the depth, σ the slowness and α and β are positive numbers. To compensate for this spreading, a mass-lumping approach has been proposed (Chavent & Plessix 1999). The idea of the mass-lumping is to form a diagonal

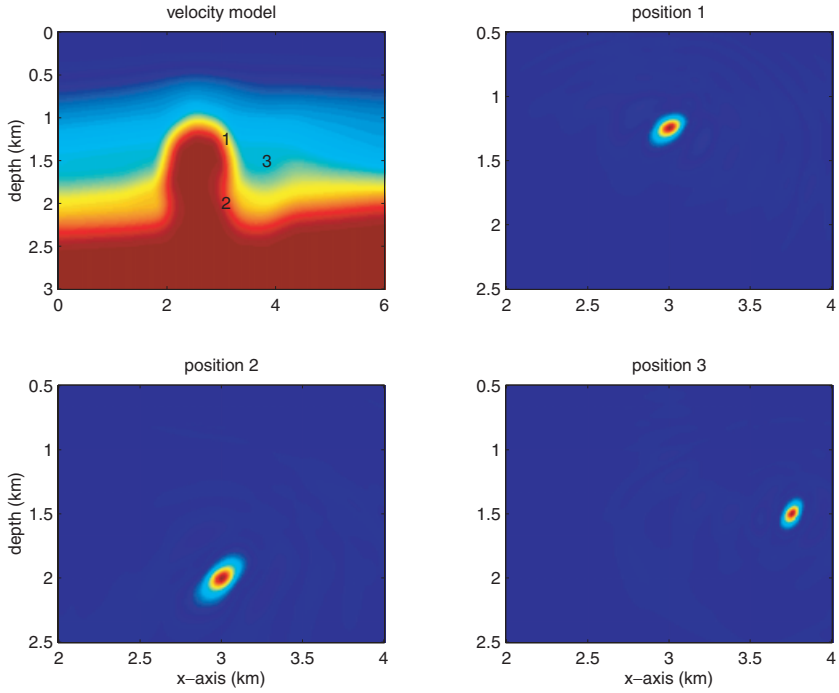

Figure 1. Several lines of the Hessian. The top left image shows the background velocity with three points for which a line of the Hessian is computed (see text). The result for the first point is plotted on the top right, for the second on the bottom left, and for the third on the bottom right.

approximation of the matrix by stacking the absolute values of (some of) the off-diagonal terms of a given line of the matrix with the diagonal term and to use this value in the diagonal approximation. With ray-tracing modelling it is feasible to compute a number of off-diagonal elements and add them up to carry out the mass-lumping. With a finite-difference scheme this is too expensive for practical applications, so we just divide the Hessian by $z^\alpha \sigma^{-\beta}$, assuming that the area of the ellipse is proportional to $z^\alpha \sigma^{-\beta}$. In two space dimensions, we observed that $\alpha = \beta = 1/2$ provided the best result. This value may be related to the Fresnel zone. According to Berkhout (1985), the Fresnel zone is proportional to $\sqrt{z/2f\sigma}$ with f the frequency. The factor $1/\sqrt{f}$ is not taken into account in the examples because the migration was carried out over a limited range of frequencies for which this factor is almost constant. This factor would probably have be needed for a larger frequency range.

So far, we have seen a Hessian that is not diagonally dominant, but still resembles a smeared-out version of a diagonal matrix. In some cases, however, there may be non-negligible contributions further away from the diagonal. Multiples are one example (Pratt *et al.* 1998), but these are not considered here.

We continue with the expression for the diagonal term of the Hessian. In the shot-based case, the least-squares functional is

$$J = \sum_{\mathbf{x}_s} J_s, \quad \text{with} \quad J_s(\sigma) = \frac{1}{2} \sum_{\omega} \sum_{\mathbf{x}_r} |c(\mathbf{x}_s, \mathbf{x}_r, \omega) - d(\mathbf{x}_s, \mathbf{x}_r, \omega)|^2, \tag{16}$$

and its gradient is

$$\nabla_r J(\sigma) = \sum_{\mathbf{x}_s} \nabla_r J_s(\sigma), \tag{17}$$

with the gradient per shot

$$\nabla_r J_s(\sigma)(\mathbf{x}) = \mathrm{Re}\left(\sum_{\omega} \omega^2 \sum_{\mathbf{x}_r} f_s(\mathbf{x}_s, \mathbf{x}, \omega) G(\mathbf{x}_s, \mathbf{x}, \omega) G(\mathbf{x}, \mathbf{x}_r, \omega) [\bar{c}(\mathbf{x}_s, \mathbf{x}_r, \omega) - \bar{d}(\mathbf{x}_s, \mathbf{x}_r, \omega)] \right). \tag{18}$$

Here the overbar denotes the complex conjugate.

The diagonal of the Hessian becomes

$$H(\mathbf{x}, \mathbf{x}) = \sum_{\omega} \omega^4 \sum_{\mathbf{x}_s} |f_s(\mathbf{x}_s, \mathbf{x}, \omega)|^2 |G(\mathbf{x}_s, \mathbf{x}, \omega)|^2 \sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega)|^2. \tag{19}$$

With an infinite and continuous receiver coverage per shot, the factor $\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega)|^2$ is almost constant in the lateral direction. Moreover, the dependency on the lateral direction is small far from the source with an infinite coverage and can be neglected inside the migration domain (see Appendix B). Therefore it can be assumed that $\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega)|^2$ is independent of the spatial coordinates, at least for a homogeneous Earth.

With this assumption, and in the case of a single shot at $\mathbf{x}_s$, the migrated image m_s obtained by scaling the gradient with the diagonal of the Hessian is

$$m_s(\mathbf{x}) = -\frac{\nabla_r J_s(\sigma)}{\sum_{\omega} \omega^4 |u_0(\mathbf{x}_s, \mathbf{x}, \omega)|^2}. \tag{20}$$

This result is the same as the one obtained with the classic imaging condition. If we use this result to define a stacked image, we obtain

$$m = \sum_{\mathbf{x}_s} m_s = -\sum_{\mathbf{x}_s} \frac{\nabla_r J_s(\sigma)}{\sum_{\omega} \omega^4 |u_0(\mathbf{x}_s, \mathbf{x}, \omega)|^2}. \tag{21}$$

This is rather different from the least-squares result for all shots, using the full data set:

$$m = -\frac{\sum_{\mathbf{x}_s} \nabla_r J_s(\sigma)}{\sum_{\omega} \omega^4 \sum_{\mathbf{x}_s} |u_0(\mathbf{x}_s, \mathbf{x}, \omega)|^2}. \tag{22}$$

We conclude that the least-squares approach leads to a different way to compute the migration weights.

If the source coverage is also infinite, as we previously assumed for the receivers, the factor $\sum_{\mathbf{x}_s} |u_0(\mathbf{x}_s, \mathbf{x}, \omega)|^2$ is also independent of $\mathbf{x}$. In that case, the unweighted gradient $\sum_{\mathbf{x}_s} \nabla_r J_s(\sigma)$ already provides a good image because the diagonal of the Hessian is almost proportional to the identity matrix. This is the main difference between the imaging principle, which works on one shot gathered at a time, and the more general least-squares approach, which handles directly multishot and multi-offset data.

In practice, we do not have an infinite and uniform source and receiver coverage. Then, the acquisition geometry can be taken into account by the least-squares approach using the migration weights from eq. (19). Further simplifying approximations are needed to make its use computationally tractable. Some options are discussed in the next section.

4 AMPLITUDE-PRESERVING MIGRATION WEIGHTS

The computation of the migration weights requires an approximation to the inverse of the diagonal of the Hessian given in eq. (19) if excessive computational cost is to be avoided. With a finite-difference scheme, the term $f(\mathbf{x}_s, \mathbf{x}, \omega) G(\mathbf{x}_s, \mathbf{x}, \omega)$ has to be computed anyway and is

therefore known, but the terms $G(\mathbf{x}, \mathbf{x}_r, \omega)$ are not (see Appendix C). Computing the latter for each receiver separately is too expensive even in a frequency-domain finite-difference code. We therefore need some approximations for $\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega|^2$.

The first option is to *assume* that $\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega|^2$ is independent of $\mathbf{x}$, so that it can be dropped, being a constant scaling factor. This is the infinite continuous aperture approximation already mentioned in the previous section. The resulting migration weights, $K^{(1)}$, called type 1, and the corresponding migrated image, $m^{(1)}$, are

$$K^{(1)}(\mathbf{x}) = \left(\sum_{\omega} \omega^4 \sum_{\mathbf{x}_s} |u(\mathbf{x}_s, \mathbf{x}, \omega)|^2 \right)^{-1}, \quad m^{(1)}(\mathbf{x}) = -K^{(1)}(\mathbf{x}) \nabla J(\mathbf{x}). \tag{23}$$

This approach is similar to the amplitude-preserving migration proposed in (Shin *et al.* 2001), see Appendix A.

The second approximation is to replace $\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega|^2$ by $\sum_{\mathbf{x}_s} |G(\mathbf{x}_s, \mathbf{x}, \omega|^2$. This would be valid only if the receiver locations coincide with the source locations. The resulting migration weights, $K^{(2)}$, called type 2, and the migrated image, $m^{(2)}$, are

$$K^{(2)}(\mathbf{x}) = \left(\sum_{\omega} \omega^4 \left(\sum_{\mathbf{x}_s} |u(\mathbf{x}_s, \mathbf{x}, \omega)|^2 \right)^2 \right)^{-1}, \quad m^{(2)}(\mathbf{x}) = -K^{(2)}(\mathbf{x}) \nabla J(\mathbf{x}). \tag{24}$$

The third approximation is to compute $\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega|^2$ for a constant model and continuous aperture. In two space dimensions, let $x_r^{\min}(\mathbf{x}_s)$ and $x_r^{\max}(\mathbf{x}_s)$ be the minimum and maximum receiver positions for a shot located at $\mathbf{x}_s$. Then (see Appendix B)

$$\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega|^2 \propto \operatorname{asinh}\left(\frac{x_r^{\max}(\mathbf{x}_s) - x}{z} \right) - \operatorname{asinh}\left(\frac{x_r^{\min}(\mathbf{x}_s) - x}{z} \right). \tag{25}$$

This defines the migration weights, $K^{(3)}$, called type 3, and the migrated image, $m^{(3)}$:

$$K^{(3)}(\mathbf{x}) = \left\{ \sum_{\omega} \omega^4 \sum_{\mathbf{x}_s} |u(\mathbf{x}_s, \mathbf{x}, \omega)|^2 \left[\operatorname{asinh}\left(\frac{x_r^{\max}(\mathbf{x}_s) - x}{z} \right) - \operatorname{asinh}\left(\frac{x_r^{\min}(\mathbf{x}_s) - x}{z} \right) \right] \right\}^{-1}, \quad m^{(3)}(\mathbf{x}) = -K^{(3)}(\mathbf{x}) \nabla J(\mathbf{x}). \tag{26}$$

The fourth approximation is based on the assumption that $\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega|^2$ is independent of frequency, so it can be computed for only a single frequency ω_0 in order to reduce the computational cost. The migrated image becomes

$$m^{(4)}(\mathbf{x}) = -\left(\sum_{\omega} \omega^4 \sum_{\mathbf{x}_s} |u(\mathbf{x}_s, \mathbf{x}, \omega)|^2 \sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega_0)|^2 \right)^{-1} \nabla J(\mathbf{x}). \tag{27}$$

Because this approach is still slower than the three earlier ones, it is not considered in the following examples.

5 EXAMPLES

To test the various migration weights, the two-way wave equation, eq. (1), was solved by a finite-difference scheme (Pratt 1999). We only considered 2-D problems.

The first synthetic example is a simple layered model. The true velocity model used to generate the data and the smooth velocity model used to migrate the data are both plotted in Fig. 2(a). The data set consists of 62 shots and 63 receivers. The receivers are located between 25 m and 1575 m at a 25 m interval. The shots are located between 32.5 m and 1562.5 m at a 25 m interval. This mimics a land-type acquisition survey. For the migration, we used frequencies between 15 and 30 Hz.

The migration formulae described above provide only relative amplitudes. The normalized migration weights by the maximum value are shown in Fig. 2(b). We notice that the type 2 (dashed black line) and the type 3 (dotted red line) migration weights are very similar with this land-type acquisition example.

Next, we compare the migration images, divided by the square of the smooth slowness, with the true reflectivity. The latter corresponds to the difference between the smooth slowness and the true one divided by the smooth one. The division of the migrated images, $m^{(1)}$, $m^{(2)}$ and $m^{(3)}$, by the square of the slowness is needed because the variable $r = \sigma^2 - \sigma_0^2$ defined in Section 3 does not represent the reflectivity $(\sigma - \sigma_0)/\sigma_0$. Therefore, the gradient with respect to r should be divided by σ_0^2 (at $r = 0$) to obtain the gradient with respect to the reflectivity. Because we only have relative amplitudes, the migration images have been calibrated to the first layer corresponding to the bottom of the water. The result is displayed in Fig. 3. It turns out that the reflectors below the bottom of the water have amplitudes which are too small. In fact, the velocity contrast between the water and the first Earth layer is too high to be perfectly handled by the Born approximation. We have therefore decided to calibrate the migration images on the second layer. The result is shown in Fig. 4. Note that we have plotted only the results between a depth of 0.6 and 1.6 km which excludes the water bottom reflection. We observe that migration weights of type 1 still underestimate the amplitudes of the deep reflectors, whereas the migration weights of type 2 and 3 correctly predict the relative amplitudes. This is clearly seen for the reflectors at depths of 1.15 and 1.4 km. The migration type 2 is not plotted because it gives results similar to migration type 3 for the land-type acquisition used.

The second synthetic example was based on the same layered model, but now a marine-type acquisition geometry was chosen. Each shot had 61 receivers with a minimum offset (distance between source and receiver) of 75 m and a maximum offset of 1575 m. The results are similar to the first example. The migration weights of type 1 underestimate the amplitudes of the deep events. A comparison between the

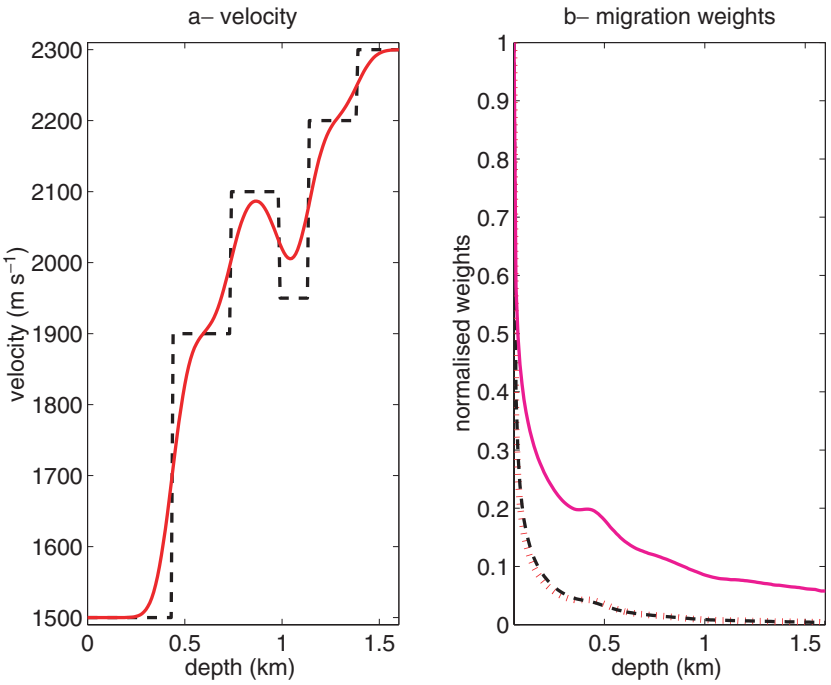

Figure 2. (a) The true velocity (dashed black line) and the smooth velocity (solid red). (b) Migration weights of type 1 (solid magenta line), type 2 (dashed black line) and type 3 (dotted red line).

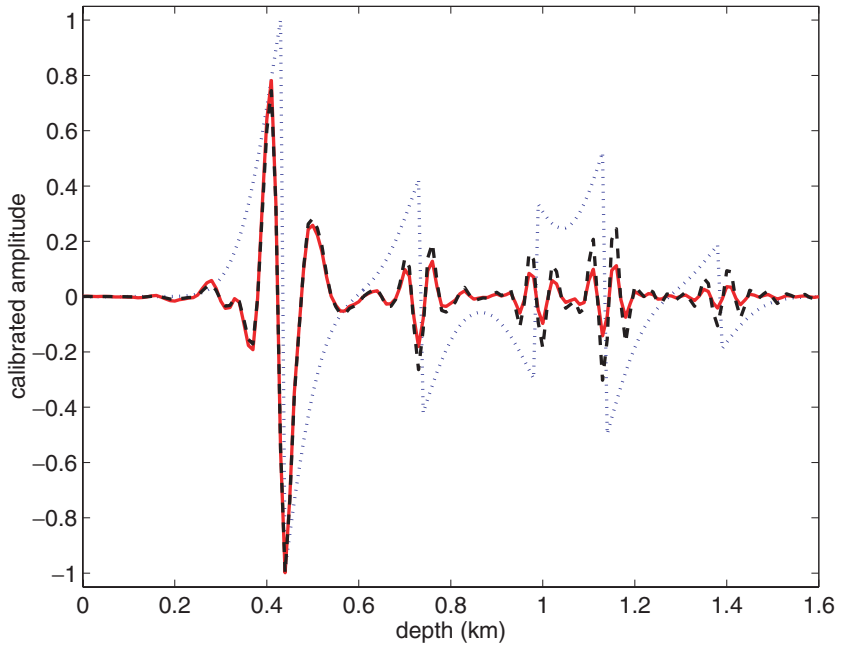

Figure 3. A trace of the migrated image after calibration on the first layer (water bottom). The true reflectivity is the dotted blue line, the migration result of type 1 is the solid red and the migration of type 3 is dashed-dotted black line.

results obtained for the migration weights type 2 and 3 is shown Fig. 5. The two migrated images are fairly similar, although it seems that the migration weights of type 3 give a slightly better result (see for instance the event at 1.4 km depth). In fact, the acquisition geometry is better taken into account by the migration weights of type 3 for this marine-type survey.

The third example is based on the SEG/EAGE salt-dome model (Aminzadeh *et al.* 1997). The velocity model used to generate the data is plotted in Fig. 6. The model for the migration is a smooth version of the true model. A marine-type acquisition geometry has been chosen. A time-domain finite-difference code was used to generate 237 shots with 65 receivers each. The distance between two shots was 80 m and the distance between two receivers was 40 m with a maximum offset of 2.7 km.

The migration result without any pre-conditioner is shown in Fig. 7. This is the gradient of the cost functional, eq. (3). The migration weights of type 2 are shown in Fig. 8 and of type 3 in Fig. 9. The amplitude-preserving migration images are shown in Fig. 10 for type 2 and

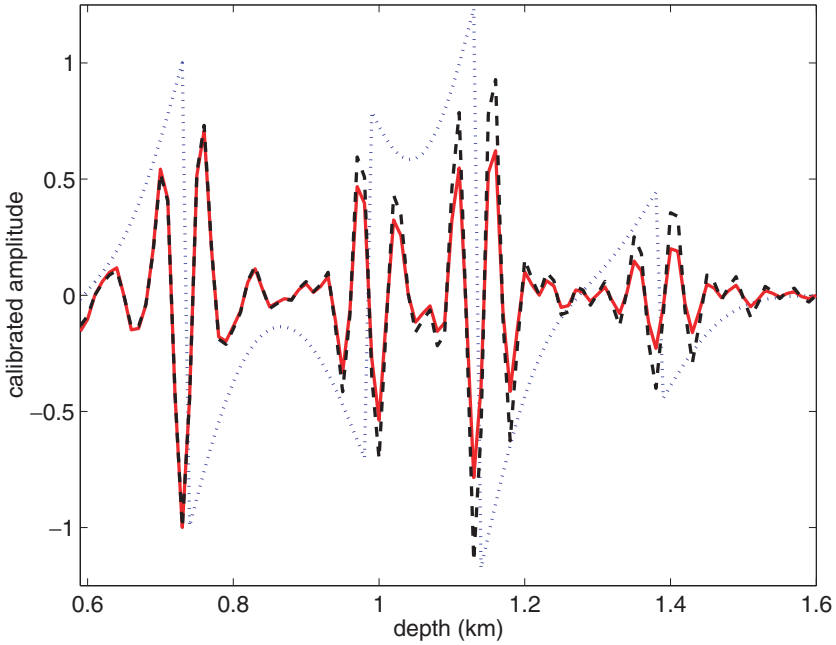

Figure 4. A zoom of the trace of the migrated image after calibration on the second layer (the water bottom is not included in the zoom). The true reflectivity is the dotted blue line, the migration result of type 1 is the solid red and the migration of type 3 is dashed-dotted black line.

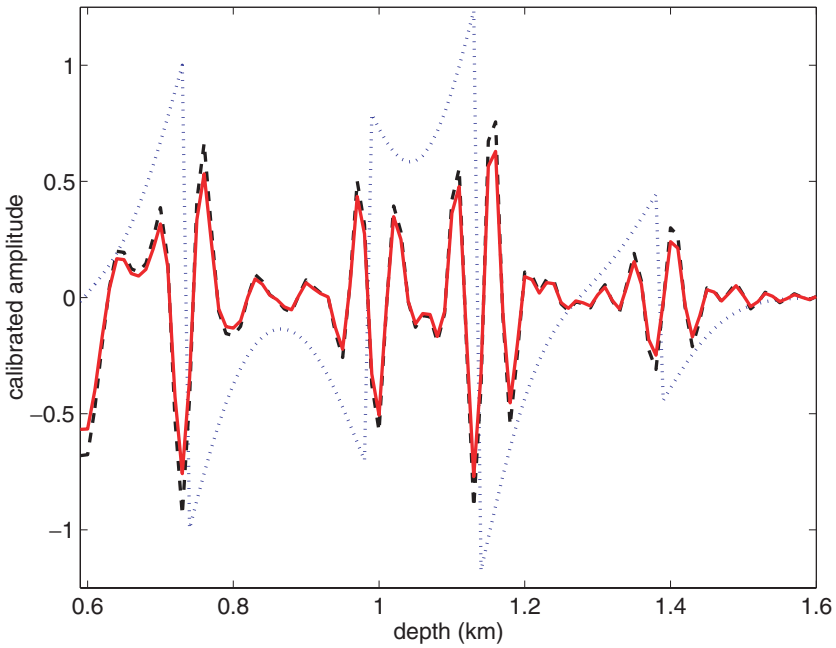

Figure 5. The migrated images for a marine acquisition after calibration on the second layer. The true reflectivity is the dotted blue line, the migration result for type 2 is the solid red line and type 3 is the dashed-dotted black line.

in Fig. 11 for type 3. A comparison is made in Figs 12 and 13 by plotting vertical cross-sections of the migration results with type 2 and 3 weighting next to the true reflectivity for $x = 2.5$, 8.0 and 11.5 km. The first section at $x = 2.5$ km crosses the sediments away from the salt body. We observe that the type 2 pre-conditioner overcorrects the amplitudes of the deep reflectors, whereas the type 3 pre-conditioner allows us to correctly retrieve them. The next section at $x = 8.0$ km crosses the salt body. Due to the high velocity contrasts, some of the amplitudes are underestimated underneath the salt body. This effect is more pronounced with the type 2 pre-conditioner than with the type 3 one. The amplitude of the reflector around 3.5 km depth is better retrieved with the type 3 pre-conditioner (Fig. 13) than with the type 2 one (Fig. 12).

The example demonstrates the importance of taking the receiver aperture into account. For a marine-type acquisition this is the main advantage of the type 3 pre-conditioner over the type 2 one.

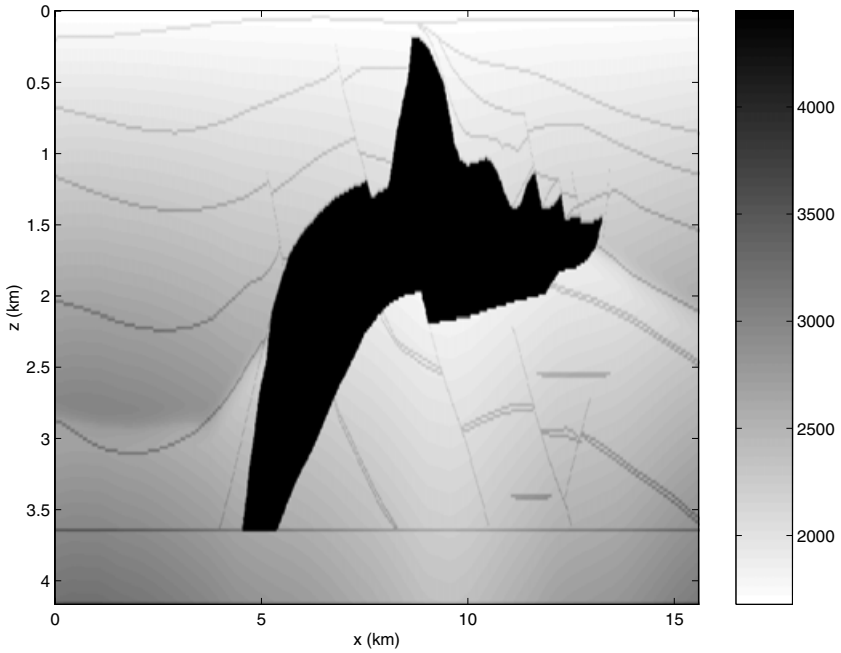

Figure 6. True velocity for the salt dome example.

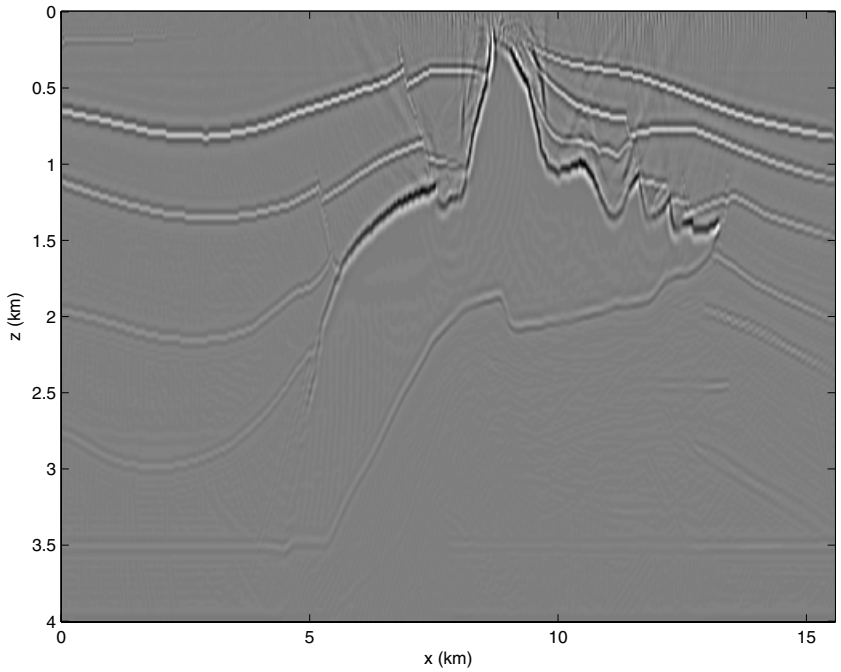

Figure 7. Migration result for the salt dome example without pre-conditioner.

6 CONCLUSIONS

We have presented and studied a number of pre-conditioners or migration weights that yield acceptable reflector amplitudes when a frequency-domain finite-difference migration algorithm is used. The migration weights are found by constructing suitable approximations to the diagonal of the Hessian of the least-squares functional. Contrary to the classic imaging principle or reverse-time migration formula, the receiver aperture is not assumed to be infinite.

Two new approximations to the diagonal of the Hessian have been proposed and tested. The numerical examples show that the new amplitude-preserving migration weights derived from these approximations better predict the relative amplitudes, especially for deeper reflectors. For retrieving the amplitudes of the deeper reflectors, it is necessary to take into account the limited aperture.

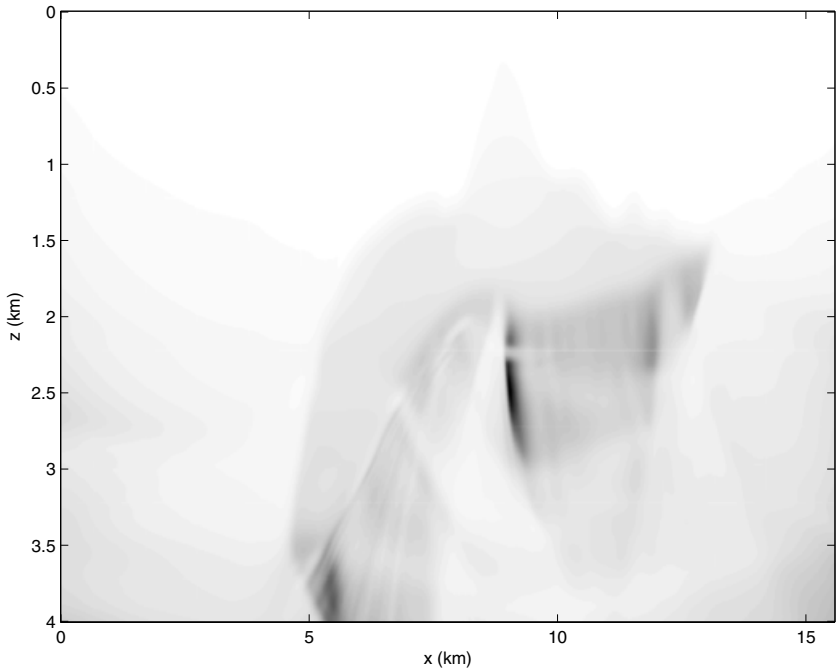

Figure 8. Migration weights of type 2.

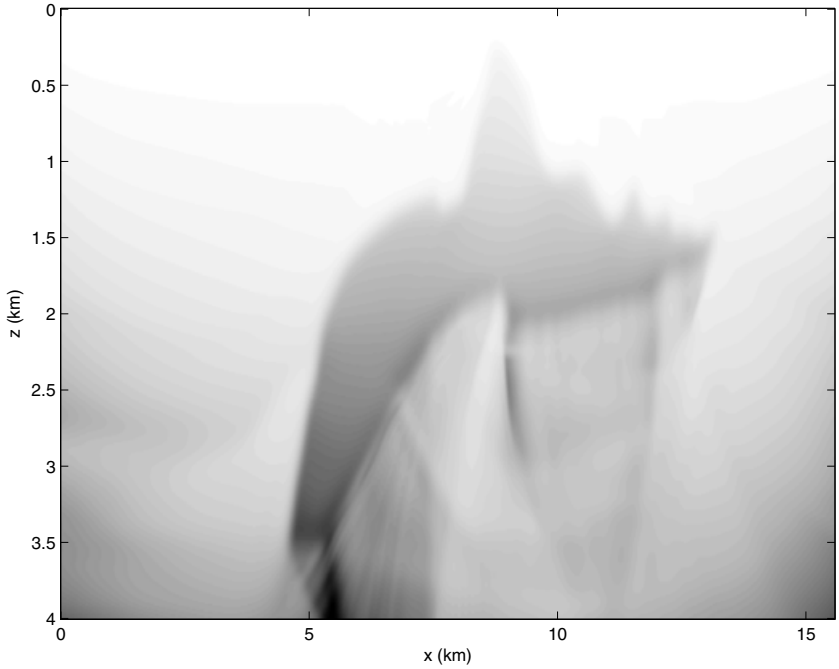

Figure 9. Migration weights of type 3.

REFERENCES

Aminzadeh, F., Brac, J. & Kunz, T., 1997. *3-D Salt and Overthrust Model,* Society of Exploration Geophysicists, Tulsa, Oklahoma.

Baysal, E., Kosloff, D.D. & Sherwood, J.W.C., 1983. Reverse time migration, *Geophysics,* **48,** 1514–1524.

Berkhout, A.J., 1985. Seismic resolution: a key to detailed geological information, *World Oil,* **201**(7), 47–51.

Beylkin, G., 1985. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform, *J. Math. Phys.,* **26,** 99–108.

Chavent, G. & Plessix, R.-E., 1999. An optimal true-amplitude least-squares prestack depth-migration operator, *Geophysics,* **64,** 508–515.

Claerbout, J.F., 1971. Towards a unified theory of reflector mapping, *Geophysics,* **36,** 467–481.

Docherty, P., 1991. A brief comparison of some Kirchhoff integral formula for migration and inversion, *Geophysics,* **56,** 1164–1169.

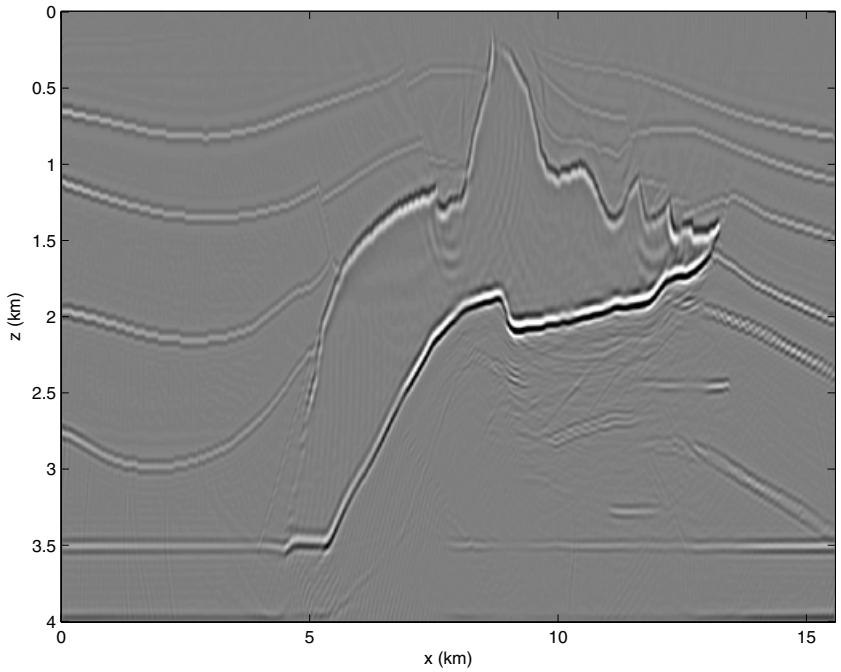

Figure 10. Amplitude-preserving migration result for type 2.

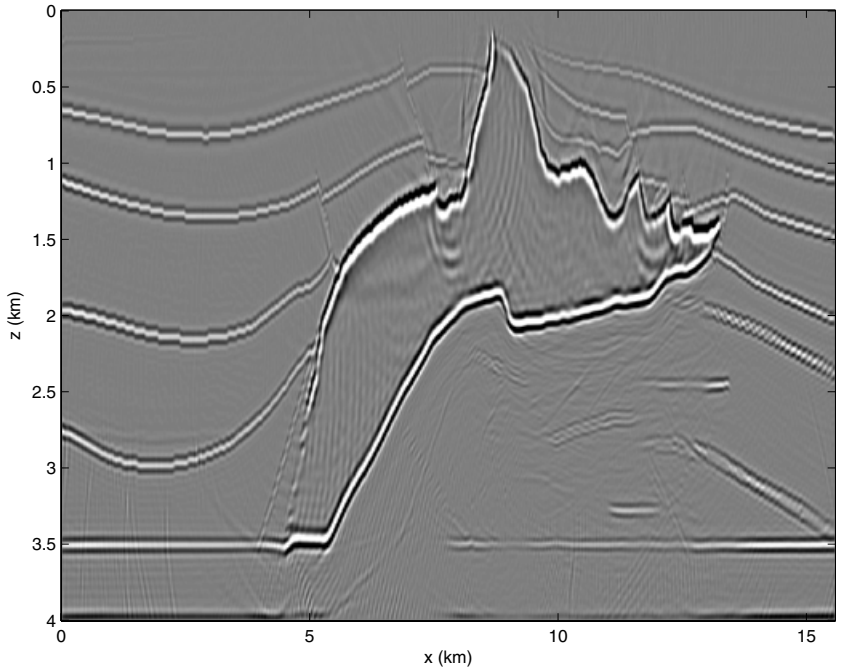

Figure 11. Amplitude-preserving migration image for type 3.

Esmersoy, C., 1986. Inversion by reverse-time extrapolation and a new imaging principle, *Society of Exploration Geophysicists 56th Annual International Meeting,* Houston, Texas, Expanded Abstracts, S16.4.

Gray, S., 1997. True amplitude migration: a comparison of three approaches, *Geophysics,* **62,** 929–936.

Lailly, P., 1983. The seismic inverse problem as a sequence of before stack migration, in *Proc. Conf. on Inverse Scattering, Theory and Applications,* SIAM, Philadelphia.

Lambaré, G., Virieux, J., Madariaga, R. & Jin, S., 1992. Iterative asymptotic inversion in the acoustic approximation, *Geophysics,* **57,** 1138–1154.

Nemeth, T., Wu, C. & Schuster, G.T., 1999. Least-squares migration of incomplete reflection data, *Geophysics,* **64,** 208–221.

Østmo, S., Mulder, W.A. & Plessix, R.-E., 2002. Finite-difference iterative migration by linearized waveform inversion in the frequency domain, *Society of Exploration Geophysicists 72nd Annual International Meeting, Salt Lake City, Utah,* Expanded Abstracts, MIG P1.6.

Pratt, R.G., 1999. Seismic waveform inversion in the frequency domain. Part I: Theory and verification in a physical scale model, *Geophysics,* **64,** 888–901.

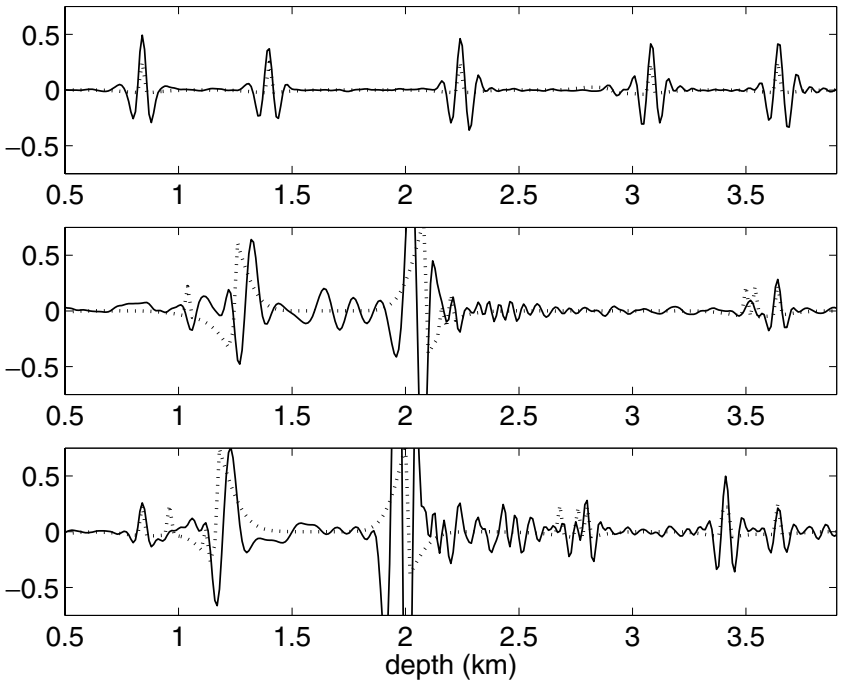

Figure 12. Comparison between the true reflectivity and the amplitude-preserving migration result for type 2 for three traces at $x = 2.5$, 8.0 and 11.5 km. The dotted line corresponds to the true reflectivity and the solid line to the migrated trace.

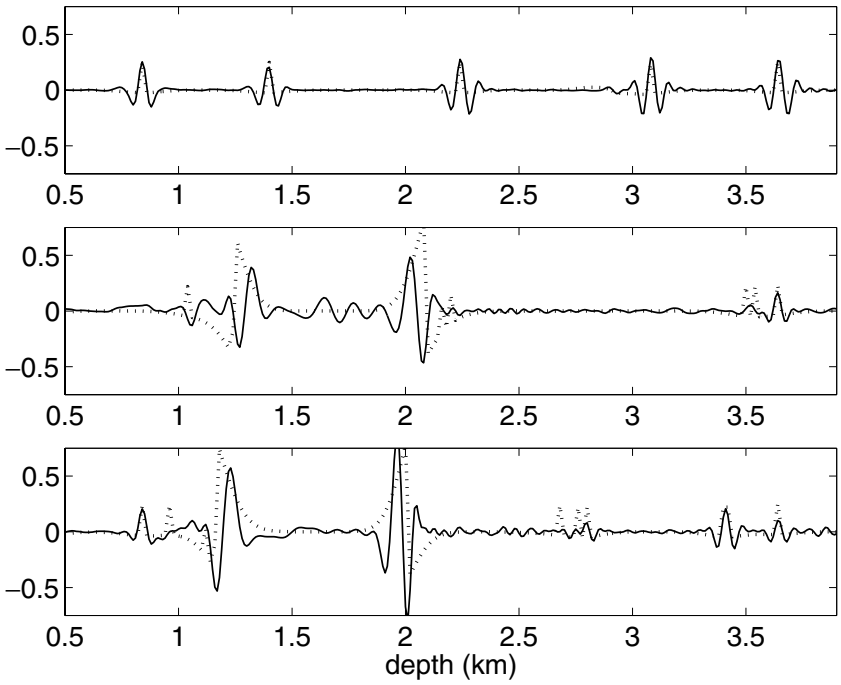

Figure 13. Comparison between the true reflectivity and the amplitude-preserving migration result for type 3 for three traces at $x = 2.5$, 8.0 and 11.5 km. The dotted line corresponds to the true reflectivity and the solid line to the migrated trace.

Pratt, R.G., Shin, C. & Hicks, G.J., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion, *Geophys. J. Int.,* **133,** 341–362.

Shin, C., Jang, S. & Min, D.J., 2001. Improved amplitude preservation for prestack depth migration by inverse scattering theory, *Geophys. Prosp.,* **49,** 592–606.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics,* **49,** 1259–1266.

Tarantola, A., 1987. *Inverse Problem Theory: Method for Data Fitting and Model Parameter Estimation,* Elsevier, Amsterdam.

ten Kroode, A.P.E., Smit, D.J. & Verdel, A.R., 1994. A microlocal analysis of migration, *Wave Motion,* **28,** 149–172.

APPENDIX A: TYPE 1 MIGRATION WEIGHTS

Here we show that the type 1 migration weights are the same as those derived by Shin *et al.* (2001). In that paper, it is proposed to estimate the diagonal of the Hessian via the so-called virtual sources. The Gauss–Newton approximation of the Hessian of J is (Pratt *et al.* 1998)

$$H = \sum_{\omega} \sum_{\mathbf{x}_s} A_J(\mathbf{x}_s, \omega)^T A_J(\mathbf{x}_s, \omega), \tag{A1}$$

with $A_J(\mathbf{x}_s, \omega)$ the Jacobian matrix associated with one shot and one frequency. The element $\mathbf{x}$, $\mathbf{y}$ of A_J is

$$A_J(\mathbf{x}_s, \omega)(\mathbf{x}, \mathbf{y}) = \frac{\partial u(\mathbf{x}_s, \mathbf{x}, \omega)}{\partial \nu(\mathbf{y})}. \tag{A2}$$

This element can be computed from

$$L(\sigma, \omega) \frac{\partial u(\mathbf{x}_s, \mathbf{x}, \omega)}{\partial \nu(\mathbf{y})} = f_{\mathbf{y}}^*(\mathbf{x}_s, \mathbf{x}, \omega), \tag{A3}$$

where $L = -\omega^2 \sigma^2 - \Delta$ and $f_{\mathbf{y}}^*(\mathbf{x}_s, \mathbf{x}, \omega)$ is the virtual source defined by

$$f_{\mathbf{y}}^*(\mathbf{x}_s, \mathbf{x}, \omega) = -\frac{\partial L(\sigma, \omega)}{\partial \nu(\mathbf{y})} u(\mathbf{x}_s, \mathbf{x}, \omega). \tag{A4}$$

However, the solution of the linear systems (A3) is too expensive when the number of parameters $\nu(\mathbf{y})$ is large. In Shin *et al.* (2001), it is proposed to replace $A_J(\mathbf{x}_s, \omega)^T A_J(\mathbf{x}_s, \omega)$ by $F^*(\mathbf{x}_s, \omega)^T F^*(\mathbf{x}_s, \omega)$, where

$$F^*(\mathbf{x}_s, \omega)(\mathbf{x}, \mathbf{y}) = f_{\mathbf{y}}^*(\mathbf{x}_s, \mathbf{x}, \omega). \tag{A5}$$

The migration image becomes

$$m = \left[\operatorname{diag} \left(\sum_{\omega} \sum_{\mathbf{x}_s} F^*(\mathbf{x}_s, \omega)^T F^*(\mathbf{x}_s, \omega) \right) \right]^{-1} \nabla J, \tag{A6}$$

where it should be noted that there is a summation over the sources and over the frequencies in ∇J. Since

$$\frac{\partial L(\sigma, \omega)}{\partial \nu(\mathbf{y})} = -\omega^2 \frac{\partial \sigma^2}{\partial \nu(\mathbf{y})}, \tag{A7}$$

and because of the parametrization

$$\sigma^2(\mathbf{y}) = \sigma_0^2(\mathbf{y}) + \nu(\mathbf{y}), \tag{A8}$$

we obtain

$$f_{\mathbf{y}}^*(\mathbf{x}_s, \mathbf{x}, \omega) = \omega^2 u(\mathbf{x}_s, \mathbf{x}, \omega) \delta(\mathbf{x}, \mathbf{y}), \tag{A9}$$

where $\delta(\mathbf{x}, \mathbf{y}) = 1$ if $\mathbf{x} = \mathbf{y}$ and 0 otherwise. Therefore,

$$\left(\sum_{\omega} \sum_{\mathbf{x}_s} F^*(\mathbf{x}_s, \omega)^T F^*(\mathbf{x}_s, \omega) \right) (\mathbf{x}, \mathbf{x}) = \sum_{\omega} \omega^4 \sum_{\mathbf{x}_s} |u(\mathbf{x}_s, \mathbf{x}, \omega)|^2. \tag{A10}$$

Eq. (A6) is then identical to eq. (22) or eq. (23).

APPENDIX B: FINITE APERTURE

For a 2-D constant velocity model, we have at large distances from the source

$$G(\mathbf{x}, \mathbf{x}_r, \omega) \propto \frac{1}{\sqrt{||\mathbf{x}_r - \mathbf{x}||}}. \tag{B1}$$

For a continuous receiver coverage from $x_r^{\min}(\mathbf{x}_s)$ to $x_r^{\max}(\mathbf{x}_s)$ at zero depth $z_r = 0$, we find

$$\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega|^2 \propto \int_{x_r^{\min}(\mathbf{x}_s)}^{x_r^{\max}(\mathbf{x}_s)} \frac{1}{\sqrt{(x_r - x)^2 + z^2}} \, dx_r \propto \operatorname{asinh} \left(\frac{x_r^{\max}(\mathbf{x}_s) - x}{z} \right) - \operatorname{asinh} \left(\frac{x_r^{\min}(\mathbf{x}_s) - x}{z} \right). \tag{B2}$$

If $x_r^{\max} = h_r$ and $x_r^{\min} = -h_r$, for large (almost infinite) h_r we obtain

$$\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega|^2 \propto 2 \operatorname{asinh}(h_r/z) \tag{B3}$$

Then, for an infinite coverage, $\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega|^2$ is constant in the lateral direction. This quantity depends on the vertical coordinate; however if h_r is very large, the dependence can be neglected. Indeed if h_r is equal to 10^{20} m and z varies from 1 to 10 km and $\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega|^2$ varies from 88.9 to 93.5. Therefore, inside the migration domain and far from the source, the term $\sum_{\mathbf{x}_r} |G(\mathbf{x}, \mathbf{x}_r, \omega|^2$ can be considered as constant for an infinite coverage.

APPENDIX C: FINITE-DIFFERENCE SCHEME AND MIGRATION

The discretization of the two-way wave equation with a finite-difference scheme leads to a linear system:

$$L(\mathbf{x}, \omega, \sigma)u(\mathbf{x}, \mathbf{x}_s, \omega, \sigma) = f(\mathbf{x}_s, \mathbf{x}, \omega) \tag{C1}$$

with $L(\mathbf{x}, \omega, \sigma)$ the discretization of $-(\omega^2/\sigma^2) - \nabla$, $u(\mathbf{x}, \mathbf{x}_s, \omega, \sigma)$ the pressure field on a grid and $f(\mathbf{x}_s, \mathbf{x}, \omega)$ the source term. By definition, $u(\mathbf{x}, \mathbf{x}_s, \omega, \sigma)$ is equal to the Green's function multiplied by the source function.

The error least-squares functional, J, is

$$J(\sigma) = \frac{1}{2} \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} ||P(\mathbf{x}_r, \mathbf{x})u(\mathbf{x}, \mathbf{x}_s, \omega, \sigma) - d(\mathbf{x}_s, \mathbf{x}_r, \omega)||^2. \tag{C2}$$

with $P(\mathbf{x}_r, \mathbf{x})$ the projection onto the receiver positions of the pressure field and $d(\mathbf{x}_s, \mathbf{x}_r, \omega)$ the data.

The gradient of $J(\sigma)$ with respect to $v = 1/\sigma^2$ is

$$\nabla_v J(\sigma) = \mathrm{Re} \sum_{\omega} \sum_{\mathbf{x}_s} \left\langle \sum_{\mathbf{x}_r} P^* e, \frac{\partial u}{\partial v} \right\rangle_{\mathbf{x}} \tag{C3}$$

with $e = Pu - d$. The asterisk denotes the conjugate transpose and the term in the subscripted angle brackets is the scalar product in the space domain. Re is the real part of a complex number. To simplify the notation, the dependences have been removed.

From eq. (C1) we find

$$\frac{\partial u}{\partial v} = -L^{-1} \frac{\partial L}{\partial v} u. \tag{C4}$$

Then

$$\nabla_v J(\sigma) = \mathrm{Re} \sum_{w} \sum_{\mathbf{x}_s} \left\langle (L^{-1})^* \sum_{\mathbf{x}_r} P^* e, \frac{\partial L}{\partial v} u \right\rangle_{\mathbf{x}}. \tag{C5}$$

If we now define

$$L(\mathbf{x}, \omega, \sigma)^* q(\mathbf{x}, \mathbf{x}_s, \omega, \sigma) = \sum_{\mathbf{x}_r} P(\mathbf{x}_r, \mathbf{x})^* e(\mathbf{x}_s, \mathbf{x}_r, \omega), \tag{C6}$$

we obtain

$$\nabla_v J(\sigma) = \mathrm{Re} \sum_{w} \sum_{\mathbf{x}_s} \omega^2 \langle q, u \rangle_{\mathbf{x}} \tag{C7}$$

since $\partial L/\partial v = -\omega^2$. An identical result can be obtained by the adjoint-state method.

If we assume that no reflections are generated in a sufficiently smooth background velocity model, then $Pu = 0$ and $q(\mathbf{x}, \mathbf{x}_s, \omega, \sigma)$ is equal to the back-propagated wavefield of the shot gather. The gradient is minus the cross-correlation of the incident wavefield from the source position with the back-propagated wavefield of the shot gather. This proves that the gradient gives a result similar to the imaging principle. Note that the summation over the receivers has disappeared from eq. (C7), because it has already been done in the right-hand side of eq. (C6). This is the reason why, in a finite-difference approach, the Green's functions at the receiver locations are not available, as a separate computation for each receiver would increase the computational cost enormously.