Geophys. J. Int. (1997) **131**, 618–642

Sensitivities of seismic traveltimes and amplitudes in reflection tomography

Yanghua Wang and R. Gerhard Pratt

Department of Geology, Imperial College of Science, Technology and Medicine, London SW7 2BP, *UK. E-mail: y.h.wang@ic.ac.uk*

Accepted 1997 June 27. Received 1997 June 12; in original form 1996 June 4

SUMMARY

Seismic traveltimes and amplitudes in reflection-seismic data show different dependences on the geometry of reflection interfaces, and on the variation of interval velocities. These dependences are revealed by eigenanalysis of the Hessian matrix, defined in terms of the Fréchet matrix and its adjoint associated with different norms chosen in the model space. The eigenvectors and eigenvalues of the Hessian clearly show that for reflection tomographic inversion, traveltime and amplitude data contain complementary information. Both for reflector-geometry and for interval-velocity variations, the traveltimes are sensitive to the model components with small wavenumbers, whereas the amplitudes are more sensitive to the components with high wavenumbers. The model resolution matrices, after the rejection of eigenvectors corresponding to small eigenvalues, give us some insight into how the addition of amplitude information could potentially contribute to the recovery of physical parameters.

In order to cooperatively invert seismic traveltimes and amplitudes simultaneously, we propose an empirical definition of the data covariance matrix which balances the relative sensitivities of different types of data. We investigate the cooperative use of both data types for, separately, interface-geometry and 2-D interval-velocity variations. In both cases we find that cooperative inversions can provide better solutions than those using traveltimes alone. The potential benefit of including amplitude-data constraints in seismic-reflection traveltime tomography is therefore that it may be possible to resolve the known ambiguity between the reflector-depth uncertainty and the interval-velocity uncertainty better.

Key words: Fréchet derivatives, inversion, perturbation methods, ray tracing, reflection seismology, seismic tomography.

1 INTRODUCTION

Seismic tomography in the reflection configuration with immediate relevance to many exploration problems attempts to recover both the velocity distribution above a reflecting horizon and the reflector geometry from the reflection data. It is generally restricted to the use of traveltimes (e.g. Bishop *et al.* 1985), in which synthetic traveltimes are generated that best match the observed traveltime data. Reflection tomography suffers from several problems, including non-uniqueness, poor resolution and ambiguity between velocity and reflector position (e.g. Farra & Madariaga 1988; Williamson 1990; Stork 1992a,b; Bube, Langan & Resnick 1995). The inclusion of amplitude information might provide better model resolution than is possible with traveltime data alone, without excessive additional computational time.

Ray-amplitude data have previously been used in velocity inversion by Thomson (1983), Nowack & Lutter (1988) and Nowack & Lyslo (1989). Thomson (1983) and Nowack & Lutter (1988) used the amplitudes of the direct arrivals to invert for velocity variation. Using a slightly perturbed model in which the velocity of two smoothly splined velocity heterogeneities is increased by 1 per cent above a constant background, Nowack & Lyslo (1989, Fig. 10b) showed that it is possible to invert for velocity variation using reflection-seismic amplitudes. Wang & Houseman (1994, 1995) have investigated the efficacy of ray-amplitude inversion for, respectively, interface geometry and interval velocities, where traveltime data are excluded in the reflection inversion.

In this paper we investigate and compare the sensitivities of seismic-reflection traveltimes and ray amplitudes with respect to both the interface geometry and the interval velocities (the slowness variations), and we show that traveltime and amplitude data do indeed contain complementary information, being sensitive to different features of the model.

The analysis of sensitivity can be carried out by linearizing the problem in the vicinity of a computed solution. Given a model perturbation around the current estimate, the variations

of the predicted observations are calculated. The Fréchet matrix, F (the partial derivatives of observed data with respect to the model parameters), is sometimes called the sensitivity matrix in forward and inverse problems. Evaluating the eigenvalues of the matrix $\mathbf{F}^T\mathbf{F}$ allows the sensitivity of observations to model parameters to be analysed, since the eigenvalues give a measure of the sensitivity of the corresponding eigenvectors. Most previous work on sensitivity has been done by means of singular value decomposition (SVD) of the Fréchet matrix F, where the positive square root of the eigenvalue of the matrix $\mathbf{F}^T\mathbf{F}$ (or $\mathbf{F}\mathbf{F}^T$) is commonly called the singular value of the matrix F. In the context of crosshole tomography, SVD analysis results were shown by Bregman, Bailey & Chapman (1989). Pratt & Chapman (1992) and Farra & Le Bégat (1995) further applied the SVD analysis method to anisotropic crosshole traveltime inversion. For reflection tomography, Farra & Madariaga (1988) and Stork (1992a,b) showed SVD analysis results in traveltime inversion; Wang & Houseman (1994, 1995) applied the techniques of SVD in amplitude inversion. However, Delprat-Jannaud & Lailly (1992) and Pratt & Chapman (1992) showed, using the SVD of the Fréchet matrix F (that is the evaluation of the eigenvalues of the matrix $\mathbf{F}^T\mathbf{F}$), that data sensitivities are strongly influenced by the way in which we parametrize the model.

In iterative linearized inversion, the Hessian matrix (and not simply the Fréchet matrix) measures the perturbation of the data misfit function resulting from a model perturbation applied to the current solution. Therefore, the sensitivity analysis can be performed by studying the Hessian matrix, which has different forms, dependent on the norm chosen in model space. The matrix $\mathbf{F}^T\mathbf{F}$ above is actually only one form of the Hessian, which follows from the use of the L^2 norm (the Euclidian norm) in the model space. Delprat-Jannaud & Lailly (1992) showed that the eigensolution of the Hessian with L^2 norm in model space is influenced by the chosen model discretization interval rather than by the physical problem under consideration, where they explored a case of traveltime inversion with a model defined by B-spline functions. To ensure the convergence of discrete eigenvalues and eigenvectors towards the solution of a continuum spectral problem, one then needs to introduce a different norm in the model space (supposed to be a Hilbertian norm associated with a scalar product of model components) physically, adding terms to the objective function that penalize large spatial derivatives. For instance, one can define a usual L^2 norm in the Sobolev space H^1, which is the space of L^2 functions such that the spatial derivative of the function is L^2 (Tarantola 1987). Delprat-Jannaud & Lailly (1992) compared the H^1 norm with the L^2 norm in the traveltime inversion, and showed that with the H^1 norm the influence of the model discretization interval is negligible (provided the discretization is sufficiently fine) and the model components are determined intrinsically by the traveltime data. In the following sensitivity investigation we will apply different norms in model space to both cases of traveltime inversion and amplitude inversion. Traveltime and amplitude data will show different sensitivities to different Fourier components of the model.

We are interested in understanding what model parameters most influence reflection traveltimes and amplitudes. Given that this is the principal objective of this study, we shall make a number of simplifications. We assume that the ray theory is valid for a 2-D isotropic earth consisting of smoothly varying velocity regions bounded by discontinuities (interfaces). We consider amplitude variation due to geometrical effects only, with no inelastic attenuation, and we also explicitly exclude the occurrence of caustics. Furthermore, in this paper we treat only a single interface, with a single variable velocity region. This simplified model nevertheless would seem adequate for the purpose of demonstrating the principal sensitivities of reflection-seismic data. As with any geophysical inversion method, we would like to find models that are sufficiently simple to allow computations to be made in a reasonable time, and that are simultaneously sufficiently realistic for those computations to be meaningful.

We begin this paper by presenting the calculation of the sensitivity matrix (the Fréchet matrix F) in Section 2. A Lagrangian formulation of traveltime and its variation due to a model perturbation is reviewed. The Euler–Lagrange equation of ray tracing can be transformed, via the Legendre transform (Kline & Kay 1965), to a Hamiltonian formulation of the ray system. The first-order perturbation to the Hamiltonian formulation (Thomson & Chapman 1985; Farra & Madariaga 1987; Nowack & Lutter 1988) is applied to describe paraxial rays and their perturbations, which in turn are used to estimate the amplitudes of seismic arrivals and their variations due to model perturbations. The Fréchet derivatives are then given by the differences between perturbed and unperturbed observations. The Hessian matrix can be defined in terms of the Fréchet matrix and its adjoint associated with different norms chosen in the model space, which are briefly described in Section 3.

In Sections 4 and 5, we discuss the difference in sensitivities of traveltimes and amplitudes to reflection interfaces and to variations in the slowness field, respectively, by carrying out an eigenanalysis of the Hessian matrix with different norms chosen in the model space. The Hessian represents the local curvature of the data misfit function. Its eigenvalue distribution indicates how the different parameters contribute to the information content of the traveltime and amplitude data, and thus is important for predicting the performance of iterative inversion techniques. Finally, synthetic examples of cooperative inversions, both for interface geometry and for velocity variations separately, are described in Section 6. A cooperative inversion using both types of data, as shown in these sections, can efficiently resolve different model components by balancing the contribution, in terms of relative sensitivities, of both traveltimes and amplitudes.

2 CALCULATION OF THE SENSITIVITY MATRIX

In this section, we describe the forward calculation of traveltimes and ray amplitudes, and their partial derivatives with respect to model perturbations. The sensitivity matrix, F, that is the Fréchet derivatives of reflection traveltimes and/or amplitudes with respect to the model parameters, can be built from these partial derivatives.

2.1 Traveltime and its variation

Introducing the Lagrangian L as

$$L(\mathbf{x}, \dot{\mathbf{x}}, \sigma) = \frac{1}{2}[\dot{\mathbf{x}}^2 + u^2(\mathbf{x})], \tag{1}$$

where σ is an independent variable defined by $d\sigma = u^{-1}ds$ in terms of the slowness $u(\mathbf{x})$ along the ray curve and the

arclength s, and differentiation with respect to σ is denoted with a dot, the traveltime integral along the ray can be expressed as

$$T=\int_{\sigma_0}^{\sigma} L(\mathbf{x},\dot{\mathbf{x}},\sigma)\,d\sigma\,. \tag{2}$$

Its first variation is

$$\delta T=\int_{\sigma_0}^{\sigma}\left[\frac{\partial L}{\partial \dot{\mathbf{x}}}\cdot\delta\dot{\mathbf{x}}+\frac{\partial L}{\partial \mathbf{x}}\cdot\delta\mathbf{x}+\frac{\partial L}{\partial u}\,\delta u\right]d\sigma\,. \tag{3}$$

Integrating the first term by parts this expression can be written as (Snieder & Spencer 1993)

$$\delta T=\left[\frac{\partial L}{\partial \dot{\mathbf{x}}}\cdot\delta\mathbf{x}\right]_{\sigma_0}^{\sigma}+\int_{\sigma_0}^{\sigma}\left[\frac{\partial L}{\partial \mathbf{x}}-\frac{\partial}{\partial\sigma}\left(\frac{\partial L}{\partial\dot{\mathbf{x}}}\right)\right]\cdot\delta\mathbf{x}\,d\sigma+\int_{\sigma_0}^{\sigma}\frac{\partial L}{\partial u}\,\delta u\,d\sigma\,, \tag{4}$$

where $\partial L/\partial u=u(\mathbf{x})$, from eq. (1), in the third integral term.

From the Fermat condition that the traveltime is stationary for perturbations in the ray position, we have $\delta T=0$. For fixed endpoints and assuming no model (slowness and boundary) perturbations, the first and third terms on the right-hand side of eq. (4) are zero. A ray equation is then obtained,

$$\frac{\partial L}{\partial \mathbf{x}}-\frac{\partial}{\partial\sigma}\left(\frac{\partial L}{\partial\dot{\mathbf{x}}}\right)=0\,, \tag{5}$$

known as the Euler–Lagrange equation. Describing the ray trajectory by the canonical vector $\mathbf{y}(\sigma)=[\mathbf{x}(\sigma),\mathbf{p}(\sigma)]^{\mathrm{T}}$ in the position $\mathbf{x}$ and momentum $\mathbf{p}$, the ray equation (5) can also be expressed in a Hamilton form (Chapman & Drummond 1982):

$$\begin{aligned}\dot{\mathbf{x}}&=\nabla_{\mathbf{p}}H\,,\\ \dot{\mathbf{p}}&=-\nabla_{\mathbf{x}}H\,,\end{aligned} \tag{6}$$

where the Hamiltonian function is defined as (Burridge 1976)

$$H(\mathbf{x},\mathbf{p},\sigma)=\frac{1}{2}[\mathbf{p}^2-u^2(\mathbf{x})]\,, \tag{7}$$

corresponding to the definition of the independent variable σ we used in this paper. Transformation between eq. (5) and eq. (6) can be done by means of the Legendre transformation (Kline & Kay 1965, pp. 115–117):

$$L(\mathbf{x},\dot{\mathbf{x}},\sigma)=-H(\mathbf{x},\mathbf{p},\sigma)+\mathbf{p}\cdot\dot{\mathbf{x}}\,. \tag{8}$$

The first-order perturbation to eq. (6) will be used later to describe paraxial rays for the calculation of ray amplitudes.

For a ray with fixed endpoints, the perturbation of traveltime for a variation in material slowness, $u(\mathbf{x})=u_0(\mathbf{x})+\delta u(\mathbf{x})$, is given by the third term on the right-hand side of eq. (4),

$$\delta T=\int_{\sigma_0}^{\sigma} u(\mathbf{x})\delta u\,d\sigma\,, \tag{9}$$

where the integral may be computed along the original unperturbed ray trajectory in the unperturbed reference medium.

Suppose a smooth interface is defined by $f_0(\mathbf{x})=0$. The perturbation of traveltime due to the interface perturbation, $\delta f(\mathbf{x})$, is given by the first term on the right-hand side of eq. (4). From the Legendre transformation eq. (8),

$$\frac{\partial L}{\partial\dot{\mathbf{x}}}=\mathbf{p}\,, \tag{10}$$

the perturbation of traveltime for a ray with fixed endpoints can be expressed as

$$\begin{aligned}\delta T&=[\mathbf{p}\cdot\delta\mathbf{x}]_{\sigma_0}^{\sigma}=[\mathbf{p}\cdot\delta\mathbf{x}]_{\sigma_0}^{\sigma_{\mathrm{a}}^-}+[\mathbf{p}\cdot\delta\mathbf{x}]_{\sigma_{\mathrm{a}}^+}^{\sigma}\\ &=[\mathbf{p}\cdot\delta\mathbf{x}]_{(\sigma_{\mathrm{a}}^-)}-[\mathbf{p}\cdot\delta\mathbf{x}]_{(\sigma_{\mathrm{a}}^+)}\,,\end{aligned} \tag{11}$$

where σ_{a}^- refers to the incident side of the interface and σ_{a}^+ refers to the reflected or transmitted side of the interface. This calculation of the effect on traveltime due to the change of boundary requires that the ray path be retraced through the layers. Approximating to first order and assuming that the effect on traveltime is restricted to the effect of the extra distance travelled, we have

$$\delta T=[\mathbf{p}_0-\hat{\mathbf{p}}_0]\cdot\delta\mathbf{x}\,, \tag{12}$$

where $\mathbf{p}_0$ and $\hat{\mathbf{p}}_0$ are the slowness vectors along the unperturbed ray on the incident side and on the reflected/transmitted side of the interface, respectively. Developing the slowness vectors $\mathbf{p}_0$ and $\hat{\mathbf{p}}_0$ along the normal and the tangent plane to the interface, the difference between them is

$$\begin{aligned}\mathbf{p}_0-\hat{\mathbf{p}}_0&=\frac{\langle\mathbf{p}_0|\nabla f_0\rangle\nabla f_0-(\mathbf{p}_0\times\nabla f_0)\times\nabla f_0}{\langle\nabla f_0|\nabla f_0\rangle}\\ &\quad-\frac{\langle\hat{\mathbf{p}}_0|\nabla f_0\rangle\nabla f_0-(\hat{\mathbf{p}}_0\times\nabla f_0)\times\nabla f_0}{\langle\nabla f_0|\nabla f_0\rangle}\\ &=\frac{\langle\mathbf{p}_0-\hat{\mathbf{p}}_0|\nabla f_0\rangle}{\langle\nabla f_0|\nabla f_0\rangle}\nabla f_0\,,\end{aligned} \tag{13}$$

which follows from the use of Snell's law, $\mathbf{p}_0\times\nabla f_0=\hat{\mathbf{p}}_0\times\nabla f_0$, where $\langle\,|\,\rangle$ and $\times$ denote the inner product and the cross-product. Expanding the perturbed interface $f_0(\mathbf{x})+\delta f(\mathbf{x})=0$ to first order, we obtain

$$\delta f+\langle\nabla f_0|\delta\mathbf{x}\rangle=0\,. \tag{14}$$

Substituting eqs (13) and (14) into eq. (12), we then obtain the variation of traveltime due to the interface perturbation,

$$\delta T=-\frac{\langle\mathbf{p}_0-\hat{\mathbf{p}}_0|\nabla f_0\rangle}{\langle\nabla f_0|\nabla f_0\rangle}\,\delta f\,. \tag{15}$$

This formula is comparable with the one used by Bishop *et al.* (1985) for interface inversion and has been used by Farra, Virieux & Madariaga (1989).

2.2 Ray amplitude and its variation

In the ray approximation, denoting $A(\sigma_0)$ the amplitude at σ_0 along the ray and close to the source point, the amplitude $A(\sigma)$ of a multiply reflected and transmitted ray can be written (Červený 1985)

$$A(\sigma)=A(\sigma_0)\left[\frac{v(\sigma_0)}{v(\sigma)}\right]^{1/2}\frac{C}{D(\sigma,\sigma_0)}\,, \tag{16}$$

where v is the wave velocity, C is the product of reflection and transmission coefficients at the interfaces, calculated by Zoeppritz's equations (Červený & Ravindra 1971), and D is the ray-geometric spreading function, which is derived in this subsection from the ray propagator describing the property of paraxial rays around a reference ray. Eq. (16) here does not consider the frequency-dependent inelastic attenuation factor.

Paraxial rays $\mathbf{y}(\sigma)=\mathbf{y}_0(\sigma)+\delta\mathbf{y}(\sigma)$, with the perturbation $\delta\mathbf{y}=[\delta\mathbf{x},\delta\mathbf{p}]^T$, can be obtained by solving the ray system of eq. (6). Linearizing the system we have (Thomson & Chapman 1985)

$$\delta\dot{\mathbf{y}}=\mathsf{A}\,\delta\mathbf{y}, \tag{17}$$

where

$$\mathsf{A}=\begin{bmatrix}\nabla_{\mathbf{x}}\nabla_{\mathbf{p}}H & \nabla_{\mathbf{p}}\nabla_{\mathbf{p}}H\\ -\nabla_{\mathbf{x}}\nabla_{\mathbf{x}}H & -\nabla_{\mathbf{p}}\nabla_{\mathbf{x}}H\end{bmatrix}. \tag{18}$$

The solution of eq. (17) can be written in terms of the ray propagator as

$$\delta\mathbf{y}(\sigma)=\Pi(\sigma,\sigma_0)\,\delta\mathbf{y}(\sigma_0), \tag{19}$$

where $\Pi(\sigma,\sigma_0)$ is the 6×6 propagator matrix of the paraxial system (Gilbert & Backus 1966; Aki & Richards 1980). In the case of a ray crossing K interfaces, the propagator along the entire ray is computed using

$$\Pi(\sigma,\sigma_0)=\Pi(\sigma,\sigma_K)\prod_{k=K}^{1}\mathbf{\Sigma}_k\Pi(\sigma_k,\sigma_{k-1}), \tag{20}$$

where $\mathbf{\Sigma}_k$ is the 6×6 transformation matrix representing the geometrical continuity condition for a ray across the kth interface. Different versions of the transformation matrix have been given in Farra *et al.* (1989), Gajewski & Psencik (1990), Wang & Houseman (1995) and Farra & Le Bégat (1995). A generalized expression given by Farra & Le Bégat (1995) keeps the general properties of the propagator matrix (see Thomson & Chapman 1985). Therefore, the ray-geometric spreading can be defined by

$$D=\left\{\det\left[\frac{\delta\mathbf{x}(\sigma)}{\delta\mathbf{x}(\sigma_0)}\right]\right\}^{1/2}. \tag{21}$$

Partitioning the propagator matrix as

$$\Pi(\sigma,\sigma_0)=\begin{bmatrix}\mathsf{Q}_1 & \mathsf{Q}_2\\ \mathsf{P}_1 & \mathsf{P}_2\end{bmatrix}, \tag{22}$$

where Q and P are 3×3 matrices which act on $\mathbf{x}$ and $\mathbf{p}$ separately, the ray-geometric spreading can be written as

$$D=[\det(\mathsf{Q}_1+\mathsf{Q}_2\mathsf{M}_0)]^{1/2}, \tag{23}$$

where M_0 determines the initial shape of the ray beam, $\delta\mathbf{p}(\sigma_0)=\mathsf{M}_0\delta\mathbf{x}(\sigma_0)$. For an initial point source in a constant-slowness medium (Farra *et al.* 1989),

$$\mathsf{M}_0=\frac{u_0}{s_0}\left(\mathsf{I}-\frac{|\mathbf{p}_0\rangle\langle\mathbf{p}_0|}{\langle\mathbf{p}_0|\mathbf{p}_0\rangle}\right), \tag{24}$$

where $u_0(\sigma_0)$ is the slowness at σ_0, and $s_0(\sigma_0)$ is the arclength measured from $\sigma=0$.

The variation of the ray amplitude due to the model perturbation is calculated in terms of the difference between the perturbed and unperturbed observations (logarithm of amplitudes). The perturbed amplitude, in terms of the perturbation to geometric spreading and to the reflection coefficient, is calculated along the perturbed 'two-point' ray. Nowack & Lyslo (1989) gave explicit numerical examples illustrating that, for reflected/transmitted rays, the perturbed reflection/transmission coefficients for the two-point ray must be used in the complete amplitude calculation or for the calculation of partial derivatives. It was also pointed out by Neele, VanDecar & Snieder (1993a) that for amplitude Fréchet derivatives the perturbed ray path must be computed. In the numerical examples shown in this paper, we simply re-do the ray tracing through the perturbed model to determine the perturbed two-point rays. Using the fast, robust ray-tracing algorithm proposed by Wang & Houseman (1995), the process of tracing a perturbed two-point ray needs half the computational time of that of using a ray-perturbation theory, which is somewhat sophisticated. Readers should refer to Farra & Madariaga (1987), Nowack & Lutter (1988), Nowack & Lyslo (1989), Farra *et al.* (1989), Snieder & Sambridge (1992, 1993), Snieder & Spencer (1993) and Farra & Le Bégat (1995), amongst many others, for different versions of the ray-perturbation theory.

The Fréchet derivatives of the observed data with respect to the model parameters are given by $[\mathsf{F}]_{ij}=\delta d_i/\delta m_j$, where δd_i is the variation of the ith observation (traveltime or $\log_{10}$ amplitude) due to the jth model-parameter perturbation, δm_j. The model parameters are described in Sections 4 and 5, where we discuss the relative sensitivity of model components to observations, by means of eigenanalysis of the Hessian matrix discussed in the following section.

3 THE HESSIAN AND THE NORM IN MODEL SPACE

Our sensitivity analysis is carried out by means of eigenanalysis of the Hessian matrix, which relates the influence of each model perturbation to perturbations of the objective function. In this section we discuss the linearized inverse problem in terms of the least-squares formulation, and then form the Hessian matrix in terms of the Fréchet matrix and its adjoint (associated with a chosen norm in the model space).

3.1 The linearized problem and the Hessian

Given a set of observed data, we want to find a model that best matches the data. The least-squares formulation of the problem is to find a model that minimizes the objective function (Tarantola 1987):

$$S(\mathbf{m})=\frac{1}{2}\|\mathbf{f}(\mathbf{m})-\mathbf{d}_{\text{obs}}\|_{\mathsf{D}}^2, \tag{25}$$

where $\mathbf{d}_{\text{obs}}$ is the observed data set and $\mathbf{f}(\mathbf{m})$ is the forward prediction. The objective function $S(\mathbf{m})$ measures the misfit between observed and calculated data. Its definition calls for the choice of a norm $\|\ \|_{\mathsf{D}}^2$ in the data space $\mathscr{D}$. This norm is associated with a scalar product $(\ ,\ )_{\mathsf{D}}$,

$$(\mathbf{a},\mathbf{b})_{\mathsf{D}}=\mathbf{a}^T\mathsf{C}_{\mathsf{D}}^{-1}\mathbf{b},\qquad\forall(\mathbf{a},\mathbf{b})\in\mathscr{D}^2, \tag{26}$$

where C_{D} is a symmetric definite positive matrix normally chosen to describe the covariances of the elements of $\mathscr{D}$. In the following numerical analysis, for inversions using one type of data (traveltimes or amplitudes) performed on modelled data, free of noise, C_{D} is chosen as an identity matrix with dimension $(data)^2$. In cooperative inversions, in which we include both types of data simultaneously, a balancing factor, controlling the contribution to the objective function of observations with different physical dimensions, is used to manipulate the corresponding elements in the identity matrix.

The inverse problem, being non-linear, is solved by successive linearizations. We generally assume that $\mathbf{f}(\mathbf{m})$ is

differentiable around a current estimate $\mathbf{m}_0$. We denote the Fréchet derivative matrix $\mathsf{F}=\nabla_{\mathbf{m}}\mathbf{f}$, which is a linear map such that the approximate equality

$$\mathbf{f}(\mathbf{m}_0+\delta\mathbf{m})\approx\mathbf{f}(\mathbf{m}_0)+\mathsf{F}\,\delta\mathbf{m} \tag{27}$$

is valid for small perturbations, $\delta\mathbf{m}$. The linearized inverse problem consists of minimizing the objective function

$$S(\delta\mathbf{m})=\frac{1}{2}\|\mathsf{F}\,\delta\mathbf{m}-\delta\mathbf{d}\|_{\mathrm{D}}^2, \tag{28}$$

where the data residual, $\delta\mathbf{d}=\mathbf{d}_{\mathrm{obs}}-\mathbf{f}(\mathbf{m}_0)$, is the difference between the data and the prediction of the current model, and the parameter perturbation $\delta\mathbf{m}$ is the 'model' to be solved. We may make a quadratic approximation to this objective function at point $\delta\mathbf{m}_0$ in the form of a Taylor series,

$$S(\delta\mathbf{m})=S(\delta\mathbf{m}_0)+\hat{\mathbf{g}}^{\mathrm{T}}(\delta\mathbf{m}-\delta\mathbf{m}_0)$$
$$+\frac{1}{2}(\delta\mathbf{m}-\delta\mathbf{m}_0)^{\mathrm{T}}\mathsf{H}(\delta\mathbf{m}-\delta\mathbf{m}_0), \tag{29}$$

in terms of the gradient vector $\hat{\mathbf{g}}$ and the Hessian matrix H.

Introducing the adjoint F^+ of F associated with the scalar product $(\ ,\ )_{\mathrm{M}}$ in the model space $\mathscr{M}$ by

$$(\mathsf{C}_{\mathrm{D}}^{-1}\delta\mathbf{d},\ \mathsf{F}\,\delta\mathbf{m})_{\mathrm{D}}=(\mathsf{F}^+\mathsf{C}_{\mathrm{D}}^{-1}\delta\mathbf{d},\ \delta\mathbf{m})_{\mathrm{M}},$$
$$\forall(\delta\mathbf{m},\ \delta\mathbf{d})\in\mathscr{M}\times\mathscr{D}, \tag{30}$$

allows expression of the gradient vector as

$$\hat{\mathbf{g}}=\mathsf{F}^+\mathsf{C}_{\mathrm{D}}^{-1}(\mathsf{F}\delta\mathbf{m}-\delta\mathbf{d}), \tag{31}$$

and the Hessian matrix as

$$\mathsf{H}=\mathsf{F}^+\mathsf{C}_{\mathrm{D}}^{-1}\mathsf{F}. \tag{32}$$

This Hessian matrix measures the influence of a perturbation about the solution model on the objective function (28). The solution of a linearized inverse problem can be obtained from $\hat{\mathbf{g}}(\delta\mathbf{m})=0$, that is

$$\mathsf{F}^+\mathsf{C}_{\mathrm{D}}^{-1}\mathsf{F}\,\delta\mathbf{m}=\mathsf{F}^+\mathsf{C}_{\mathrm{D}}^{-1}\delta\mathbf{d}, \tag{33}$$

and hence the perturbation $\delta\mathbf{m}-\delta\mathbf{m}_0$ makes a change to the objective function (eq. 29) equal to

$$\Delta S=S(\delta\mathbf{m})-S(\delta\mathbf{m}_0)$$
$$=\frac{1}{2}(\delta\mathbf{m}-\delta\mathbf{m}_0)^{\mathrm{T}}\mathsf{H}(\delta\mathbf{m}-\delta\mathbf{m}_0). \tag{34}$$

If a model perturbation gives rise to a large perturbation of the objective function, it is a sensitive component in the model and well determined in the solution. Similarly, a perturbation that has a small or no effect on the objective function is a less sensitive or insensitive component, and correspondingly poorly determined or not determined in the solution.

Ideally such a sensitivity analysis can be evaluated at each iteration of a linearized inversion, where an optimum solution from the linearized problem (33) is $\delta\mathbf{m}=\delta\mathbf{m}_0$, so that ΔS in eq. (34) equals zero. In the following numerical analysis, however, we evaluate the Hessian matrix at the true solution point, **m**, the optimal model found at the global minimum of the non-linear misfit function (25). In the vicinity of the solution point the curvature of the misfit function, defined by the Hessian, shows the sensitivity of the objective function to the traveltime and amplitude data,

$$\Delta S(\mathbf{m})=S(\mathbf{m}+\Delta\mathbf{m})-S(\mathbf{m})$$
$$=\frac{1}{2}\Delta\mathbf{m}^{\mathrm{T}}\mathsf{H}\Delta\mathbf{m}. \tag{35}$$

Since the misfit function (25) is rather non-linear for the model, the relative sensitivities evaluated at the global minimum may not apply for all iterations of the linearized inversion if the initial estimate is remote from the solution **m**.

The Hessian, the local curvature of the data misfit function, describes all possible combinations of parameter perturbations (or model parameters) that give rise to the same change ΔS as seen in eq. (34) (or in eq. 35), and represents an M-dimensional ellipsoid (where M is the total number of model parameters). An eigenanalysis of H will give the lengths and the directions of the principle axes of the ellipsoid,

$$(\mathsf{H}-\lambda_i\,\mathsf{I})\mathbf{v}_i=0, \tag{36}$$

where λ_i is the eigenvalue and $\mathbf{v}_i$ is the associated eigenvector. Eigenvectors corresponding to large and small eigenvalues give, respectively, well- and poorly determined parameter combinations.

If we rewrite eq. (36) by matrices, $\mathsf{H}=\mathsf{V}\Lambda\mathsf{V}^{\mathrm{T}}$, where Λ is a diagonal matrix consisting of the eigenvalues λ_i and V is the matrix of associated eigenvectors, the solution of a linearized inverse problem (eq. 33) can be expressed as

$$\delta\tilde{\mathbf{m}}=(\mathsf{F}^+\mathsf{C}_{\mathrm{D}}^{-1}\mathsf{F})^{-1}\mathsf{F}^+\mathsf{C}_{\mathrm{D}}^{-1}\delta\mathbf{d}$$
$$=\mathsf{H}^{-1}\mathsf{H}\,\delta\mathbf{m}=\mathsf{V}\mathsf{V}^{\mathrm{T}}\delta\mathbf{m}, \tag{37}$$

where the matrix $\mathsf{V}\mathsf{V}^{\mathrm{T}}$ is called the model resolution matrix of the problem. The calculated model resolution matrix, after the rejection of eigenvectors corresponding to small eigenvalues, can give us insight into how the relative sensitivity actually translates into accuracy and efficiency in a real inversion problem.

3.2 The Hessian and the norm in model space

The Hessian matrix used for the sensitivity analysis described above for determining the sensitive and insensitive model components is defined by eq. (32) in terms of the Fréchet matrix and its adjoint, associated by eq. (30) with the scalar product $(\ ,\ )_{\mathrm{M}}$ in the model space $\mathscr{M}$. The scalar product is, however, related to the choice of the norm $\|\ \|_{\mathrm{M}'}$ in the dual space $\mathscr{M}'$ (see Appendix A). Following Delprat-Jannaud & Lailly (1992), we define the norm $\|\ \|_{\mathrm{M}'}$ in $\mathscr{M}'$ by a weighted sum of two integral terms, in which one term is the sum of the L^2 scalar product of the model components and the other is the sum of the L^2 scalar product of the first spatial derivatives of the model components (see eq. A1). The scalar product depends on the weighting parameter α, which can take values within [0, 1). When $\alpha=0$ it defines the L^2 norm. When $\alpha=0.5$ it corresponds to the usual norm in the Sobolev space H^1, which is equal to the sum of the usual L^2 norm of the function and of the L^2 norm of its spatial derivative (Tarantola 1987). This process can be physically understood to penalize both the data misfit and the first derivatives of the solution in the inversion, that is we search for model perturbations with small spatial derivatives. This is consistent with the ray theory, which requires a locally smooth model.

Denoting $\{\beta_i(\mathbf{x}), \forall i(1 \le i \le M)\}$ to be basis functions (harmonic functions in our case, as shown in the following sections) defining the model, we build an $M \times M$ symmetric positive definite matrix D_α with elements defined by

$$[D_{ij}]_\alpha = (\beta_i(\mathbf{x}), \beta_j(\mathbf{x}))_\alpha ,$$
$$\forall i(1 \le i \le M), \forall j(1 \le j \le M), \tag{38}$$

where the scalar product $(\ ,\ )_\alpha$ is defined by eq. (A5), and the subscript α records the dependency on the value of α chosen to define the scalar product. As shown in the appendix, with the operator D_α the Hessian matrix in eq. (32) can then be expressed as

$$\mathsf{H} = \mathsf{D}_\alpha^{-1}\mathsf{F}^{\mathrm{T}}\mathsf{C}_{\mathrm{D}}^{-1}\mathsf{F}, \tag{39}$$

in terms of the Fréchet matrix and its transposition. The matrix C_{D} is assumed equal to the identity matrix in this paper, for one type of noise-free data.

Delprat-Jannaud & Lailly (1992) have concluded that one cannot define a continuum spectral problem associated with the operator $\mathsf{F}^{\mathrm{T}}\mathsf{F}$, and that the numerical solutions are then strongly dependent on the discretization and do not converge as the discretization is refined. However, one can define a continuum spectral problem for the Hessian $\mathsf{F}^{+}\mathsf{F}$ associated with the scalar product $(\ ,\ )_\alpha$, $(0 < \alpha < 1)$, and obtain accurate numerical solutions of this continuum spectral problem. They have shown the results of eigenanalysis of $\mathsf{F}^{+}\mathsf{F}$ with the H^1 norm, compared to that with the L^2 norm, in their traveltime inversion with a model parametrized as a B-spline interpolation of discretized points. In the following sections these two norms are applied also in the case of traveltime inversion, but with a different model parametrization, and are extended to the case of amplitude inversion for the analysis of traveltime and amplitude sensitivities to interface and slowness distributions.

4 SENSITIVITIES TO INTERFACE GEOMETRY

4.1 Interface parametrization and the Hessian operator

In this section we attempt to investigate the sensitivities of reflection traveltime and ray-amplitude data to the geometry of a reflecting interface. The relative sensitivity is measured by the magnitude of eigenvalues of the Hessian matrix for the linearized equation. The analyses for traveltimes and amplitudes are initially considered separately. The relative sensitivity of the amplitudes compared to the traveltimes is then evaluated by means of a cooperative inversion including both types of data. The model resolution matrix for the pseudo-inverse using a truncation of the SVD is also given in each case.

We consider a 2-D stratified velocity structure consisting of variable-thickness layers, and assume that the depth of the interface dividing the layers varies continuously. A continuous interface, band-limited in wavenumber, may be approximated by the Fourier series for its periodic continuation,

$$Z(x) = \sum_{n=0}^{N} [a_n \cos(n\pi k_0 x) + b_n \sin(n\pi k_0 x)], \tag{40}$$

where a_n and b_n are amplitude coefficients of the nth harmonic term (a basis function used in eq. 38) with a wavenumber equal to an integer multiple of fundamental wavenumber k_0 (the reciprocal of the fundamental wavelength), and N is the number of harmonic terms. We chose this parametrization in order to study carefully the scale dependence of the traveltime and amplitude variation.

In order to illustrate the calculation of the Fréchet matrix for the study of sensitivities, we consider a model with a constant velocity (2500 m s^{-1}) and a single flat reflector at a depth of 2000 m as a reference model for the following perturbation analysis. The velocity within the layer below the reflector is set equal to 2800 m s^{-1}. (We will return to this geometry and this reference model for a number of synthetic examples, which are shown in Section 6.) We generate a synthetic experiment with a realistic reflection-acquisition geometry in which 10 shots are located on the surface at intervals of 1000 m, with data recorded at 25 receivers for each shot. The minimum and maximum shot–receiver offsets are equal to 100 and 2500 m. Thus, we have 250 traveltime and 250 amplitude observations. The Hessian matrix, determined with the model estimate equal to the solution, is calculated using eq. (39), in terms of the Fréchet matrix and the operator D_α, relating to different norms in the model space.

Following the definition of the scalar product in the model space (given explicitly in the appendix), the scalar product of the model perturbations can be written in terms of the correlation matrix of the perturbations of interface depth and the correlation matrix of their slopes, correlating along the spatial direction. Considering a reflector parametrized by eq. (40), these two correlation matrices, denoted as $\hat{\mathsf{B}}_0$ and $\hat{\mathsf{B}}_x$ respectively, can be explicitly written in terms of basis functions:

$$(\hat{\mathsf{B}}_0)_{ij} = \int_{I_x} \beta_i(x)\beta_j(x)\,dx ,$$
$$(\hat{\mathsf{B}}_x)_{ij} = \int_{I_x} \frac{d\beta_i(x)}{dx}\frac{d\beta_j(x)}{dx}\,dx , \tag{41}$$

where I_x is that part of the interface from which rays are reflected. In the analysis to follow, we will show only the results associated with the coefficients of cosine terms, and thus set $\beta(x) = \cos(n\pi k_0 x)$ in eq. (41). Because a sine function is a phase-shifted cosine function, the sensitivity of observations to the coefficients of sine terms is similar to that of the coefficients of cosine terms. The question of the independence of the sine and cosine terms, related to the phase (that is the lateral position of the anomalies), will be discussed later.

In the following numerical experiments the two correlation matrices in eq. (41) are scaled as follows:

$$\mathsf{B}_0 = \frac{\varepsilon}{\mathrm{trace}(\hat{\mathsf{B}}_0)}\hat{\mathsf{B}}_0 ,$$
$$\mathsf{B}_x = \frac{\varepsilon}{\mathrm{trace}(\hat{\mathsf{B}}_x)}\hat{\mathsf{B}}_x , \tag{42}$$

where ε can be arbitrarily set in the range 1.0 to M for the $M \times M$ matrix to avoid floating-point overflow in the matrix inverse calculation. The effect of the scaling operation in eq. (42) is to normalize the two matrices so that they are relatively balanced and dimensionless. Thus, the operator D_α is built by

$$\mathsf{D}_\alpha = (1-\alpha)\mathsf{B}_0 + \alpha\mathsf{B}_x . \tag{43}$$

The first of the two terms in eq. (43) penalizes large perturbations and the second penalizes solutions with large gradients

(slopes). The scalar product used in eqs (41)–(43) can alternatively be understood as a working definition of the model covariance matrix, $\mathbf{C}_M^{-1}$, if we set $\mathbf{C}_M^{-1}$ as D_x times an identity matrix with dimension $(\textit{model parameter})^{-2}$ (see Section 6.1 for a further discussion of this point). Penalizing solutions with a non-zero value of α implies that a near-horizontal solution is to be preferred and non-horizontal slopes are undesirable. The choice of parameter α for the scalar product in model space allows emphasis on the model perturbations or on their spatial derivatives.

Note that the model parametrization in eq. (40) should include high-order components so that the discretized solution approximates the continuum solution. This differs from a possible inversion practice in which one may try to obtain a long-wavelength solution by truncating the high-wavenumber components in the parametrization. In the following sensitivity analysis, we use wavenumbers up to $k_{\max} = 5\ \text{km}^{-1}$. The lateral resolving power of any reflection method based on ray theory will be limited to the first Fresnel zone width, given by $[(1/2)\lambda h]^{1/2}$, where λ is the seismic wavelength and h is the depth of the reflector (Sheriff & Geldart 1995). Assuming a dominant wavelength of 100 m (for realistic exploration frequencies), the Fresnel-zone width will be of the order of 300 m for this example. Thus, a wavenumber of 5 km^{-1}, which corresponds to an interface wavelength of 200 m, is an adequate high-wavenumber limit.

4.2 Interface inversion with the L^2 norm chosen in model space

Consider an interface with wavenumbers ranging from zero to 5 km^{-1} and discretized by eq. (40) with the number of harmonic terms $N = 80$, corresponding to a discretization interval k_0 equal to $1/16\ \text{km}^{-1}$. The eigenvalues and associated eigenvectors of the Hessian matrix, with the L^2 norm chosen in the model space, and the model resolution matrix of reflection-traveltime and amplitude inversions for the interface geometry are shown in Fig. 1. Resultant eigenvalues shown in the upper part of the figure are normalized relative to the maximum value and ordered in decreasing size. Each associated eigenvector shown in the middle part of the figure is represented as a column of pixels, ranging from zero wavenumber at the top to maximum wavenumber at the bottom. The eigenvector is normalized for each eigenvalue, with darker shades representing positive values (up to +1) of the element in an eigenvector and lighter shades representing negative values (down to −0.6). The model resolution matrices of traveltime and amplitude inversion are shown in the bottom of the figure.

Let us first examine the traveltime sensitivity to the reflection interface (left-hand side of Fig. 1). Since we only have a finite number of data, the solution will be underdetermined. The sharp decrease in the magnitude of eigenvalues of the Hessian (after the 53rd) clearly indicates that a null space exists for such an over-fine discretization. To determine the geometry of a reflection interface with such densely sampled wavenumber components, more observations would naturally be helpful for a traveltime inversion. However, for a fixed experimental geometry, one has little choice but to deal with this null space in some manner. Re-parametrization using a coarser set of basis functions would be one choice; alternatively, one could just ignore the null space, as in a pseudo-matrix inversion by SVD with truncation.

If we ignore the null space by truncated SVD, biases in the solution are inevitable (Ory & Pratt 1995). The model resolution matrix indicates the reliability of the solution. To obtain the model resolution matrix, we arbitrarily ignore the eigenvectors with relative eigenvalues smaller than 10^{-3}, set approximately corresponding to the obvious discontinuity appearing in the eigenvalue curve. The diagonal elements of the model resolution matrix shown in Fig. 1, for traveltime, are

$$\left\{ \underset{(1)}{1.00};\ \underset{(2)}{0.98};\ \underset{(3)}{0.81};\ \underset{(4)}{0.63};\ \underset{(5-65)}{\ldots};\ \underset{(66)}{0.67}\right.$$
$$\left. \underset{(67-73)}{0.75-0.88};\ \underset{(74-80)}{0.90-0.97};\ \underset{(81)}{0.98};\right\},$$

and the magnitude for elements 5–65 is around 0.63–0.67. Each component $\delta\tilde{m}_j$ is the combination of its true solution δm_j and neighbouring components. The magnitudes show that traveltime inversion can relatively better determine the interface components with wavenumbers equal to or close to zero and the components with very short wavelengths (step changes causing reflections), compared to the rest of the components.

From the eigenvalues and the associated eigenvectors we see that the objective function in traveltime inversion is most sensitive to the components with wavenumbers of zero and k_0. The first eigenvector for traveltime is given as a normalized vector of $\{0.52, 0.38, \cdots,$ [with the remaining elements$< O(0.01)$]$\}$. The most significant elements are the first and second elements. The 53rd–55th eigenvectors (representing higher-wavenumber components) have small eigenvalues (close to 10^{-3}) and indicate that high-wavenumber components have relatively weak sensitivities in traveltime inversion. For the remaining eigenvalues, there is no clear dependence of reflection traveltime on the wavenumber.

In the case of amplitude inversion, the ray amplitudes are significantly more sensitive to interface components with higher wavenumbers (shorter wavelengths), as demonstrated by the distribution of the eigenvalues and associated eigenvectors of the Hessian matrix shown in the right-hand side of Fig. 1. The eigenvector pattern shows that the amplitude sensitivity decreases quasi-exponentially with wavenumber. This may be expected as the curvature of interface has a large effect on the recorded amplitude because of consequent focusing or defocusing of the beam. However, in the case of traveltime inversion with the L^2 norm chosen, there is no such dependency.

In amplitude inversion, the eigenvalues decrease smoothly without the obvious discontinuity that appears in traveltime inversion. We then follow the traveltime-inversion case and truncate the eigenvectors $\mathbf{v}_i$ for $i \geq 56$ associated with the smallest eigenvalues, representing those components with the weakest sensitivities, in the calculation of the model resolution matrix. The diagonal elements of the model resolution matrix are

$$\left\{ \underset{(1-13)}{0.00};\ \underset{(14)}{0.01};\ \underset{(15)}{0.02};\ \underset{(16)}{0.16};\ \underset{(17)}{0.58};\ \underset{(\ldots)}{\ldots};\right.$$
$$\left. \underset{(41)}{0.70};\ \underset{(42)}{0.86};\ \underset{(43)}{0.98};\ \underset{(44-81)}{1.00}\right\},$$

where the values for elements 18–40 gradually increase from 0.60 to 0.70. The model resolution matrix clearly shows that amplitude inversion with the L^2 norm can constrain all wavenumber components well except those components with small wavenumbers.

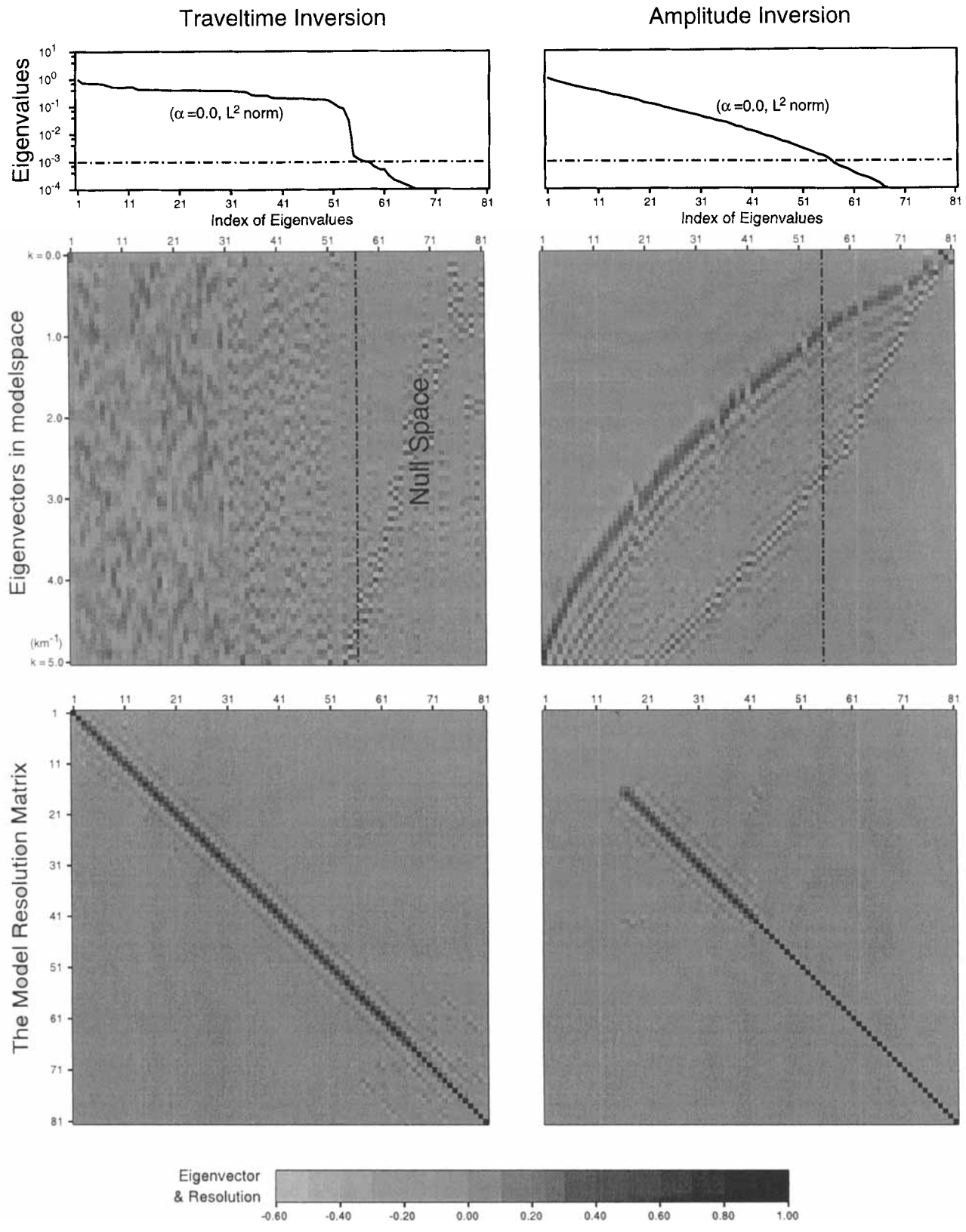

Figure 1. Independent sensitivities of reflection traveltimes and amplitudes to interface wavenumber components, in terms of the eigenvalues and eigenvectors of the Hessian matrices and the model resolution matrices in linearized inversion, where the L^2 norm ($\alpha=0$) in model space is chosen. The eigenvalues, in order of decreasing magnitude, are normalized relative to the maximum value. Each eigenvector is represented as a vertical column below the associated eigenvalue. In each eigenvector column, wavenumbers of the model components increase from top to bottom. The model resolution matrices for the problems are shown in the lower half of the figure. Integers along the sides refer to the wavenumber indices.

4.3 Interface inversion with the constraint of spatial derivatives

The same sensitivity analysis but with the H^1 norm (as $\alpha=0.5$ in eq. 43) chosen in the model space, which takes into account not only the depth variation but also the slope of the reflector, is shown in Fig. 2.

In the previous subsection we have seen that, if we choose the L^2 norm in model space for the calculation of the Hessian matrix for traveltime inversion, the objective function is most sensitive to the mean depth of a reflection horizon and the component with fundamental wavenumber k_0, but there is no simple dependence on the other non-zero wavenumbers. The clear dependence of the amplitude L^2 objective-function sensitivity on wavenumber is provided in traveltime inversion with the choice of H^1 norm in the model space, as shown in Fig. 2.

Although the choice of a norm in model space is essential to the process of traveltime inversion, as can be seen by comparing the trends of eigenvectors of the Hessian with the L^2 norm (Fig. 1) and the H^1 norm (Fig. 2), for the case of amplitude inversion we see that the value of α has only a slight influence on the trend of eigenvalues. Even with the L^2 norm in model space, a strong dependence of reflection ray amplitude on the wavenumber of interface components exists. Therefore, the constraint term of the reflector slope is not necessarily required in the objective function for amplitude inversion, although the introduction of an H^1 norm in the model space slightly improves the condition number, as seen from the comparison of eigenvalue curves. In the case of amplitude inversion the condition number of the Hessian matrix is 10^{+7} for the L^2 norm and 10^{+6} for the H^1 norm with $\alpha=0.5$.

Figs 1 and 2 also depict the model resolution matrix for each problem. These matrices show the extent to which each model parameter (that is each Fourier coefficient) is uniquely resolved; a perfectly resolved parameter has a single non-zero value of 1.0 on the main diagonal of the corresponding row of the resolution matrix, and off-diagonal values show the extent to which the parameter is non-unique with respect to other model parameters. The model resolution matrices show some apparent improvement in resolution as the H^1 norm is introduced, when compared with the results from the L^2 norm. Traveltime inversion using the H^1 norm appears to resolve the first two components, as the value of the first two diagonal elements of the model resolution matrix are equal to 1.0. In the case of amplitude inversion the model resolution matrix is also improved for components with intermediate and small wavenumber. These improvements are due to the use of the gradient penalization term in the inversion; they indicate that amplitude inversion, coupled with a constraint on the gradient of the solution, will apparently resolve the low wavenumbers. It should be noted, however, that the additional information does not come from the data, but from the constraints.

In practice, the choice of α is perhaps arbitrary, depending only on the fact that one desires to minimize the magnitude of the perturbations, or their slopes. We now consider a case of interface inversion in which interface gradients are more heavily penalized in the inversion. Note that $\alpha=1.0$ does not constrain the magnitude of the model perturbations at all, and therefore results in a singular matrix. For the results shown in Fig. 3, we set $\alpha=0.99$ in eq. (43). (Experiments have shown that if α takes a value of 0.9 the result is similar to Fig. 3; a choice of $\alpha=0.999$ results in a singular matrix.) With this level of penalization on the spatial derivatives, the small-wavenumber components are emphasized and the corresponding eigenvalues (at small indexes for traveltime and at large indexes for amplitude) are increased relative to the rest. In this case, the eigenvector pattern of the reflection traveltime clearly shows linear dependence on the wavenumber. Traveltime inversion can now apparently determine the 17 lowest-wavenumber components, as the first 17 diagonal elements of the model resolution matrix are equal to 1.0. Also evident is the suppression of the high-wavenumber components, as seen from a comparison between model resolution matrices in the case of $\alpha=0.99$ and in the cases of $\alpha=0.5$ (Fig. 2) and $\alpha=0$ (Fig. 1). In the case of amplitude inversion the model resolution matrix is almostly an identity matrix, except for 13 components with the lowest wavenumbers. Once again, the large level of penalization on the interface gradient will cause models with no gradient to be generated perhaps, in spite of the data.

In summary, with an L^2 norm in the model space, reflected traveltimes are most sensitive to the mean depth of a reflector, whereas reflected amplitudes are most sensitive to the shorter wavelengths. By introducing the constraint of spatial derivatives in the definition of the Hessian ($\alpha\neq0.0$), the traveltime sensitivities will have a dependence on the wavenumber of the interface components. Such a dependence is in the opposite direction to the sensitivity of ray amplitudes to the interface wavenumber components. Therefore, the information contents in reflection traveltimes and reflection amplitudes are indeed complementary in linearized inversion, being sensitive to different components of the interface.

4.4 Cooperative inversion of amplitudes and traveltimes (for interface geometry)

The numerical examples shown above have treated two types of data separately. Another point that we should address in this paper is the relative value of including amplitude and traveltime data in a cooperative inversion. Let us see what will happen if we take the examples of Fig. 1 and Fig. 2 and use both types of data simultaneously, $\mathbf{d}=[\mathbf{d}_{\text{ampl}}, \mathbf{d}_{\text{time}}]^{\text{T}}$, in the definition of the objective function.

If both physical data types are included, the relative effects will be determined, at least partly, by the statistical reliability of the two data types. Ideally, one would seek to incorporate such statistical information. In practice, such information is usually unavailable or unreliable. As a working solution, we characterize these statistics using a covariance matrix defined as a diagonal matrix, $\mathbf{C}_{\text{D}}=\text{diag}\{\sigma_j^2\}$, in terms of the estimated uncertainty of the jth measurement σ_j. Without data noise, we first set the standard deviation σ_j as the rms values of observation data, denoted as σ_{ampl} and σ_{time}, two constants corresponding to amplitudes and traveltimes, respectively. This procedure can be physically understood to remove the physical dimensions so that both types of data make a similar contribution in a cooperative inversion. With this model we carried out a number of experiments which revealed that reflection amplitudes absolutely dominate the inversion for interface geometry, and traveltimes only weakly influence the mean depth of interfaces. Based on this observation we then propose an alternative to the definition for the data covariance matrix, which balances the data contributions not directly by the data magnitudes but by the data sensitivities.

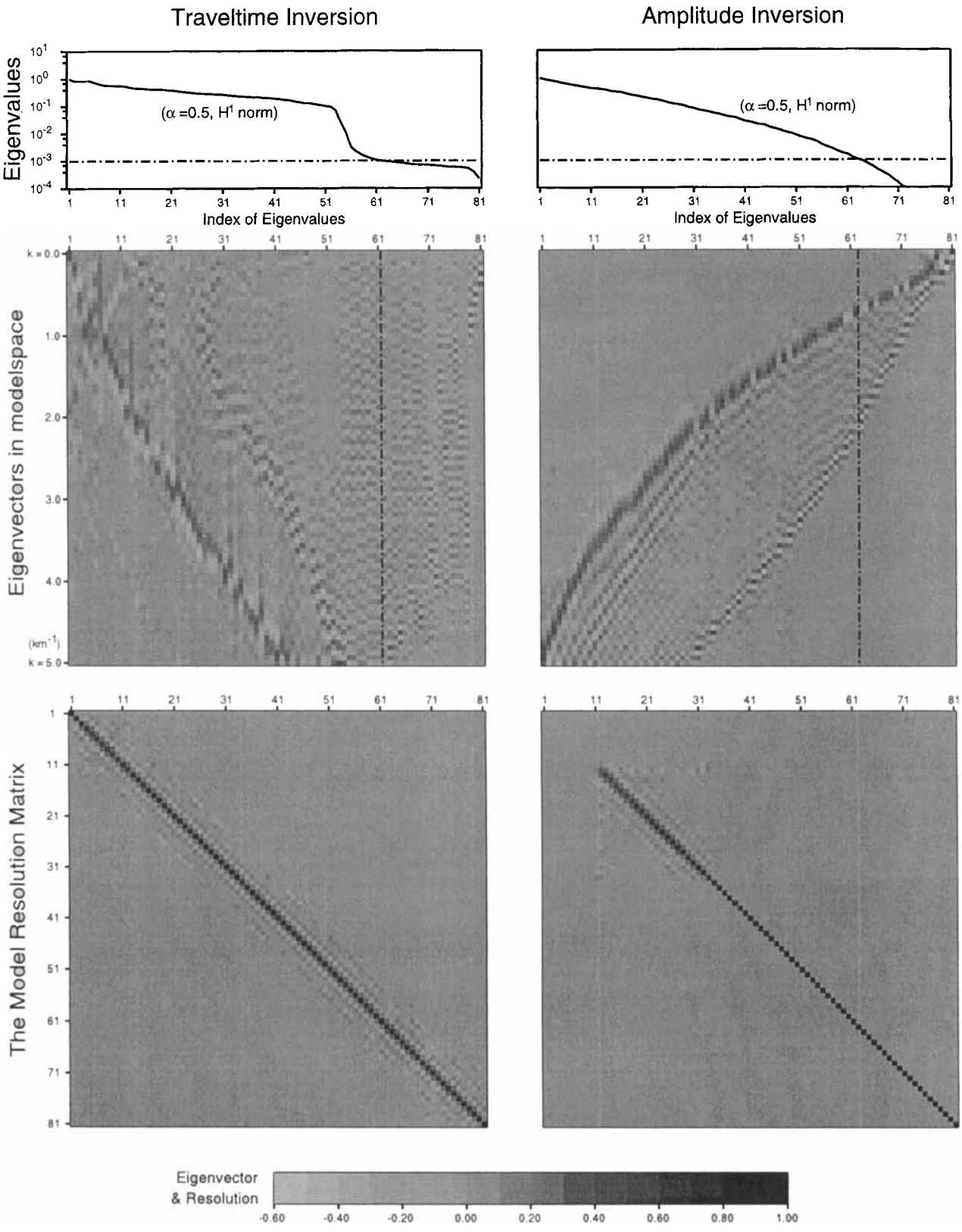

Figure 2. Independent sensitivities of reflection traveltimes and amplitudes to the interface wavenumber components, where the H^1 norm ($\alpha = 0.50$) in model space is chosen. See Fig. 1 caption for comments.

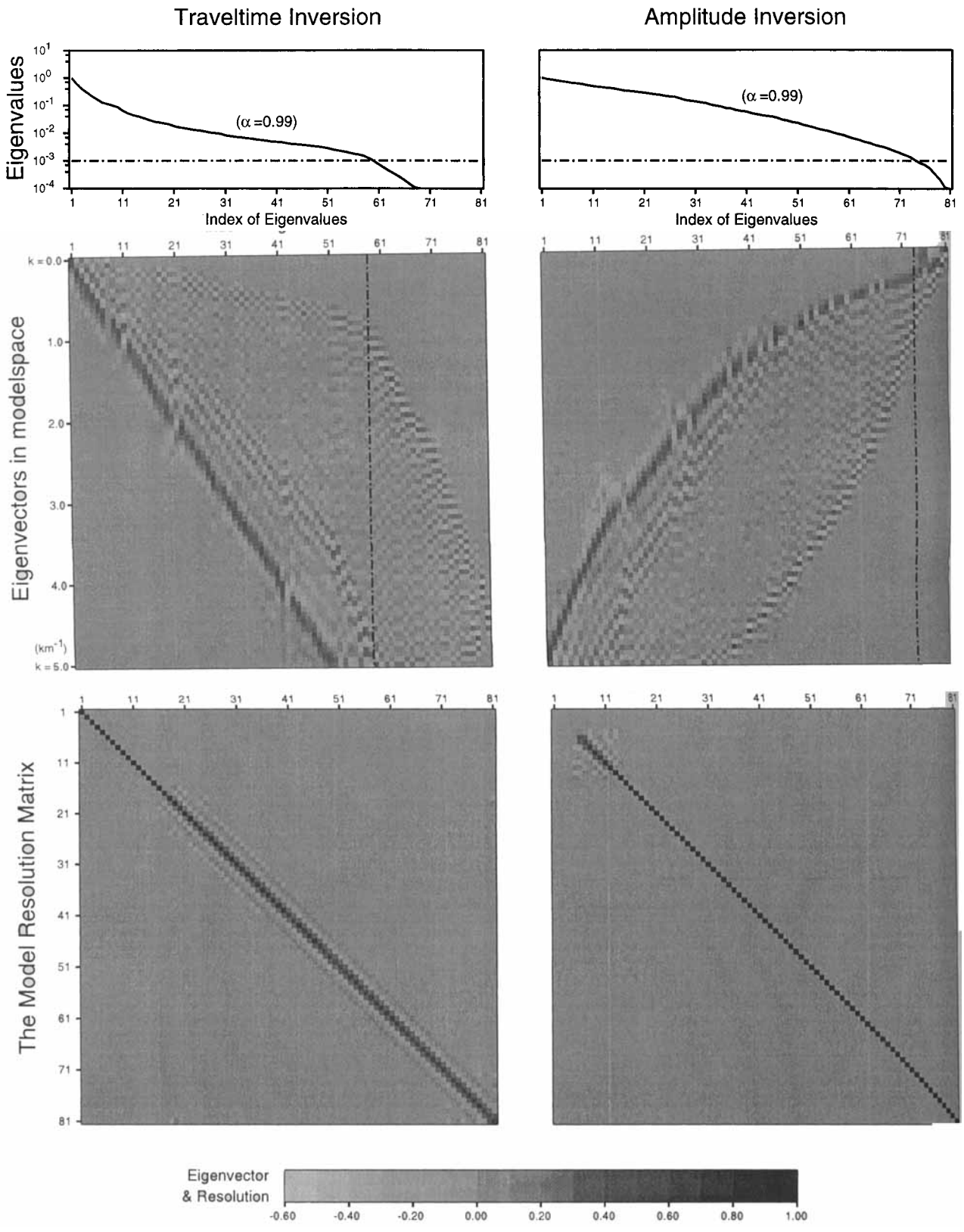

Figure 3. Independent sensitivities of reflection traveltimes and amplitudes to the interface wavenumber components, where $\alpha = 0.99$ is set in the scalar product in the model space. See Fig. 1 caption for comments.

We denote the data covariance matrices for amplitudes and traveltimes respectively as $\mathbf{C}_{\text{ampl}}^{-1}$ and $\mathbf{C}_{\text{time}}^{-1}$. The combined covariance matrix for including both types of data is then defined as

$$\mathbf{C}_{\text{D}}^{-1} = \begin{bmatrix} \kappa^{2p}\mathbf{C}_{\text{ampl}}^{-1} & 0 \\ 0 & \mathbf{C}_{\text{time}}^{-1} \end{bmatrix}, \tag{44}$$

where κ is a dimensionless balancing factor, and p can be understood as a second, somewhat arbitrary, weighting factor. If we normalize $\mathbf{C}_{\text{ampl}}^{-1}$ and $\mathbf{C}_{\text{time}}^{-1}$ to the maximum value, or set them to identity matrices after we suppress data errors [see the real data examples shown in Wang, White & Pratt (1998)], the dimensionless balancing factor κ is given by

$$\kappa = \frac{\text{trace}(\mathbf{F}_{\text{time}})}{\text{trace}(\mathbf{F}_{\text{ampl}})}, \tag{45}$$

where the ratio of matrix traces is an empirical quantity indicating the relative sensitivities of the traveltimes compared to the amplitudes. In our numerical example, the ratio of the matrix traces is $\kappa = 1.685 \times 10^{-2}$, which indicates that the amplitudes are much more sensitive to variations in the interface model than the traveltimes are. The larger p is, the more the influence of the traveltimes is included. Experiments have shown that the optimal value of p is in the range 0.25 to 0.75 for the interface inversion.

Fig. 4 shows the eigenanalysis results of the Hessian for the cooperative inverse problem using both data types, in which we chose $p = 0.75$. The sensitivity patterns for both the L^2 norm and the H^1 norm are modified from those observed in Figs 1 and 2. Comparing the patterns of eigenvectors obtained for the cooperative problem with those observed for the individual traveltime and amplitude problems, we see that the amplitudes appear to dominate the eigenvector patterns for the components with high wavenumbers and the traveltimes are dominant for those with low wavenumbers. With the L^2 norm the components with small wavenumbers are less sensitive than those using the H^1 norm in the model space, shown by the comparison of the eigenvalues corresponding to the combination of the small-wavenumber components. The model resolution matrices for the cooperative inversion appear to be a combination of those in the previous (individual) inversions. With the H^1 norm, the resolution of inversion to model components of small wavenumber is enhanced, since the penalization of spatial derivatives equally emphasizes the model components with small wavenumbers in the inversion. Therefore, if we choose $\alpha = 0.99$, as in Fig. 3, the model resolution matrix of the cooperative inversion will be close to an identity matrix.

In the case of cooperative inversion including amplitude data and traveltime data, or in the case of inversion using only reflection amplitudes, the largest eigenvalue of the Hessian corresponds to the interface component with the highest wavenumber. A situation in which the maximum sensitivity is for the shortest wavelength would not necessarily lead to a stable inversion using iterative matrix inversion methods. A requirement of stability is that it is possible to truncate the parametrization at an arbitrary high wavenumber without significantly changing the answer. If maximum sensitivity is at the short-wavelength parameter, the answer will depend to a high degree on the choice of truncation wavenumber. The amplitude-inversion examples given by Wang & Houseman (1994) have shown, however, that declaring different groups of interface components with different ranges of sensitivity magnitude into separate subspaces can effectively stabilize the inversion procedure, where it is assumed that components with sufficiently high wavenumbers to fit the continuous interface (up to the limit of the Fresnel-zone width) are included in the model parametrization. An alternative approach is also possible: a traveltime inversion is first performed to reconstruct the longer-wavelength components, and then amplitude data are used in the inversion to constrain high-wavenumber components of the model.

4.5 Determination of the phase

In the parametrization of eq. (40), we did not explicitly express the phases of the harmonic terms, which control the lateral position of any anomalies on a reflector. Naturally, the smaller the coefficient, the less important the phase, defined by $\varphi_i = \tan^{-1}(b_i/a_i)$, in describing the overall reflector geometry. If both a_i and b_i are independently sensitive components of the model, then the corresponding phase φ_i would be well determined by inversion. Fig. 5 shows the eigenanalysis of the Hessian matrix with $\alpha = 0.5$ (the H^1 norm), in which the coefficients of both sine and cosine terms (a total of 161 parameters) have been included. Only 61 of the largest eigenvalues and the corresponding eigenvectors of the Hessian are shown in the figure. This plot is similar to previous figures, but in this case each eigenvector here is plotted with two columns: the left-hand column shows the coefficients of the cosine terms and the right-hand column shows the coefficients of the sine terms (note that the zero wavenumber has no sine coefficient). This result is again for the uniform single-layer model with a single reflector.

From Fig. 5 we can see that the eigenvalues and associated eigenvectors of the sine functions have the same trends as those of the cosine terms in each of the three cases of traveltime inversion, amplitude inversion and cooperative inversion using both types of data. The eigenvectors for traveltime data contain a natural progression from small wavenumbers (associated with large eigenvalues) to large wavenumbers (associated with small eigenvalues). The trend is reversed for amplitude data and for cooperative amplitude and traveltime inversion. In the case of traveltime inversion, the cosine and sine terms do not appear to be independently resolved. For the amplitude data, there is some mixing of sine and cosine terms at the largest eigenvalues, associated with the short wavelengths. As the eigenvalues decrease, the cosine and sine terms become better resolved from each other. In the case of cooperative inversion, at low wavenumbers the model components are not distinct either. However, the inversion example in Section 6 will show that the cooperative inversion can better resolve the interface structure.

5 SENSITIVITIES TO 2-D SLOWNESS VARIATION

5.1 Slowness representation and the Hessian operator

In this section we investigate the sensitivities of seismic-reflection traveltimes and amplitudes to 2-D, smoothly varying slowness variations within the layers, again by analysing the eigenvalues in the model space of the Hessian described by

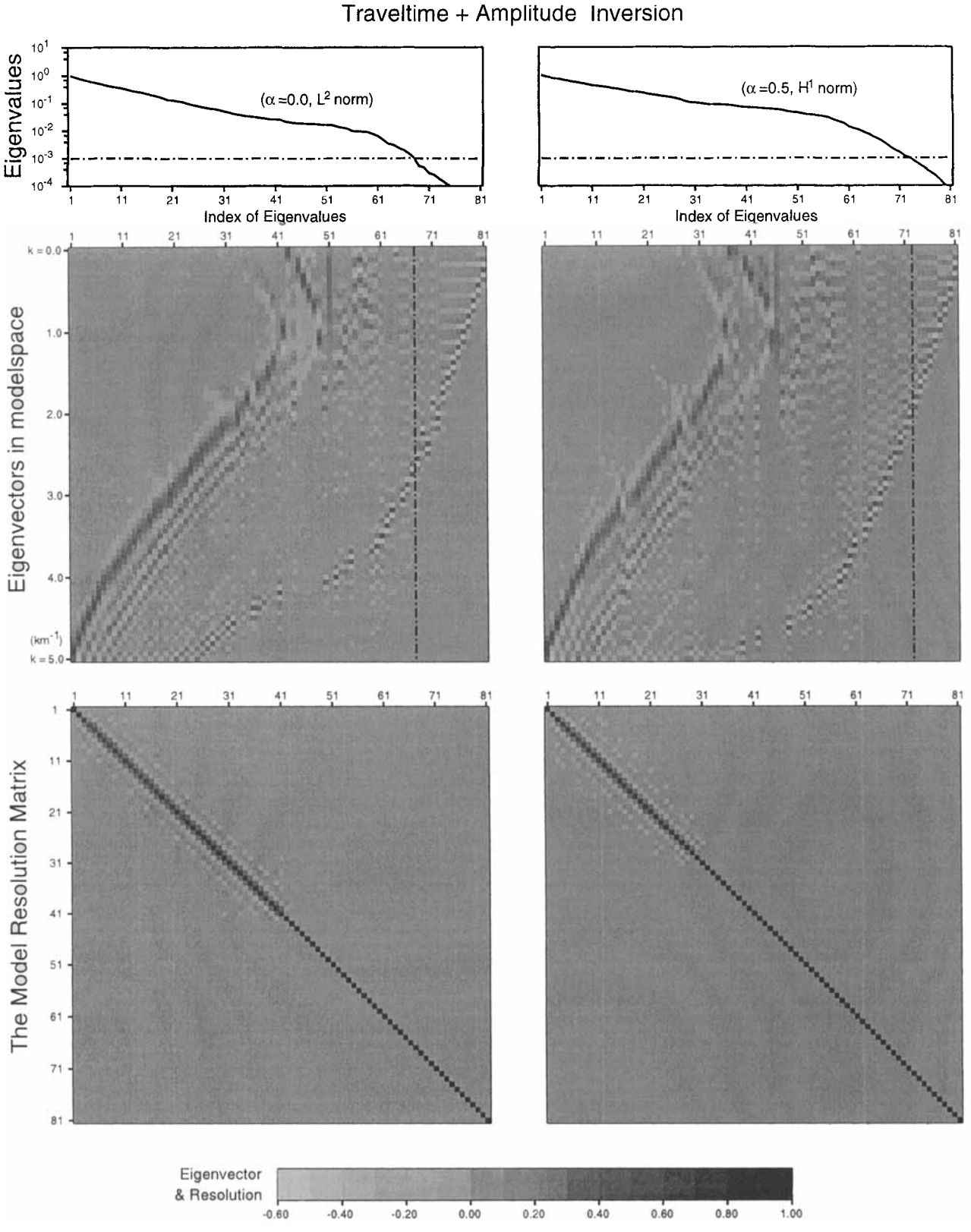

Figure 4. Joint sensitivity analysis of reflection traveltimes and amplitudes to the interface wavenumber components. See Fig. 1 caption for comments.

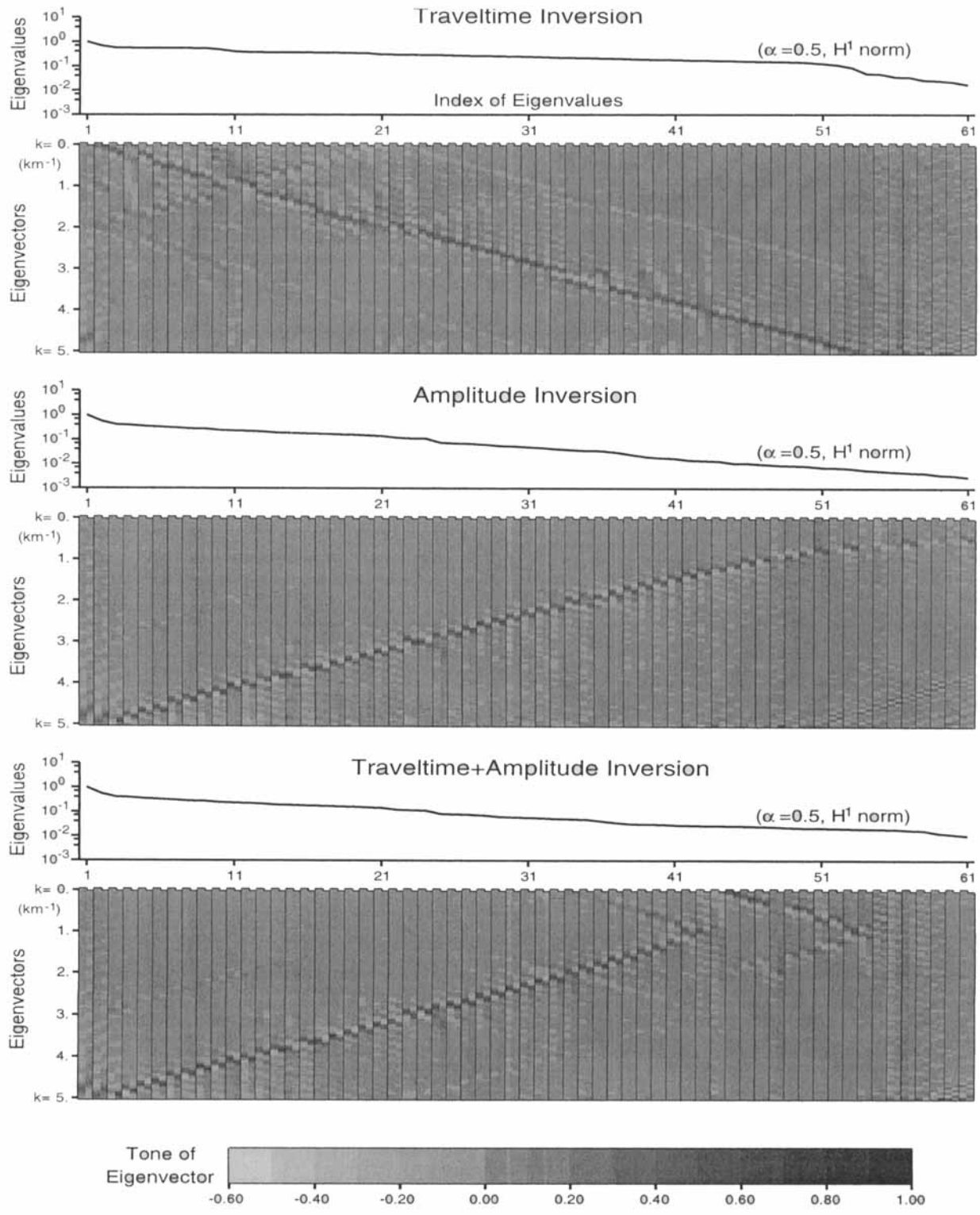

Figure 5. Eigenresults of the Hessian matrices ($\alpha=0.5$, H^1 norm) in traveltime inversion, amplitude inversion and the cooperative inversion for interface geometry, when both sine and cosine terms are included. Each eigenvector, including the 81 cosine coefficients $\{a_n, (n=0, 80)\}$ and the 80 sine coefficients $\{b_n, (n=1, 80)\}$ in the interface parametrization, is diagrammatically shown as two columns, the left column showing the cosine coefficients and the right showing the corresponding sine terms. Only 61 of the largest eigenvalues and associated eigenvectors are shown.

different norms in the model space. We represent the slowness distribution within a layer by a truncated 2-D Fourier series,

$$u(\mathbf{r}) = a_{00} + \sum_{m=1}^{N} [a_{m0} \cos(\mathbf{k}\cdot\mathbf{r}) + b_{m0} \sin(\mathbf{k}\cdot\mathbf{r})] + \sum_{m=-N}^{N} \sum_{n=1}^{N} [a_{mn} \cos(\mathbf{k}\cdot\mathbf{r}) + b_{mn} \sin(\mathbf{k}\cdot\mathbf{r})], \quad (46)$$

where the location-vector $\mathbf{r} = x\mathbf{i} + z\mathbf{j}$, the wavenumber-vector $\mathbf{k} = m\pi k_0\mathbf{i} + n\pi k_0\mathbf{j}$ and a_{mn} and b_{mn} are amplitude coefficients of the (m,n)th harmonic term. We assume that slowness varies continuously, and that any discontinuity is explicitly represented by a smooth interface. For the calculation of the Fréchet derivatives that follows, we use the same uniform single-layer model defined in the previous section. Writing the perturbation of ray amplitudes with respect to the velocity perturbation within the layer above the reflector, we keep the velocity below the interface constant.

Considering only 2-D slowness variations, the correlation matrices of the model variation and its derivatives in horizontal and vertical directions are

$$(\hat{\mathsf{B}}_0)_{ij} = \int_{I_x \times I_z} \beta_i(x,z)\beta_j(x,z)\, dx\, dz,$$

$$(\hat{\mathsf{B}}_x)_{ij} = \int_{I_x \times I_z} \frac{\partial \beta_i(x,z)}{\partial x} \frac{\partial \beta_j(x,z)}{\partial x}\, dx\, dz, \quad (47)$$

$$(\hat{\mathsf{B}}_z)_{ij} = \int_{I_x \times I_z} \frac{\partial \beta_i(x,z)}{\partial z} \frac{\partial \beta_j(x,z)}{\partial z}\, dx\, dz,$$

where $I_x \times I_z$ is the rectangular area through which rays propagate. The operator D_α can then be expressed as

$$\mathsf{D}_\alpha = (1-\alpha)\mathsf{B}_0 + \alpha(\mathsf{B}_x + \mathsf{B}_z), \quad (48)$$

where $\mathsf{B}_0 = (1/w_0)\hat{\mathsf{B}}_0$, $\mathsf{B}_x = (1/w_x)\hat{\mathsf{B}}_x$ and $\mathsf{B}_z = (1/w_z)\hat{\mathsf{B}}_z$, with scale factors, w, defined by the trace of the matrices as in eq. (42). Similar to the analysis of the interface sensitivities in the previous section, in the following numerical experiments we only show the results associated with the coefficients of cosine terms. The symmetry inherent in the model parametrization of 2-D Fourier series implies that only a quarter of the k-plane ($k_z \geq 0$, $k_x \geq 0$) is required. However, we will show the result on half the k-plane ($k_z \geq 0$) in the following analysis as we wish to gain an insight into the dependence on the phase (spatial location) of the slowness variation.

5.2 Sensitivity analysis with the L^2 norm

Fig. 6 shows the eigenvalues and associated eigenvectors of the Hessian matrix and the model resolution matrices for traveltime and amplitude inversions. The L^2 norm ($\alpha = 0.0$) is chosen in the model space. For the slowness model, the maximum wavenumber under consideration is 5.0 km^{-1}. The discretization interval of wavenumber $k_0 = 1.0$ km^{-1} ($N = 5$) is set in eq. (46). Therefore, 61 ($= 5 \times 5 + 6 \times 6$) cosine coefficients a_{mn} are determined in the eigenanalysis. In each eigenvector square, horizontal wavenumbers of model components increase from left to right (with the wavenumber range from -5 to 5 km^{-1}) and vertical wavenumbers increase from top to bottom (from zero to 5 km^{-1}). Only those 20 eigenvectors associated with the largest eigenvalues of the Hessian are shown in the figure. These eigenvectors are the linear combinations of the model parameters which have the most influence on the traveltime and amplitude data. The model resolution matrices for each inversion are displayed as a 61×61 matrix split into several panels, each corresponding to a constant horizontal wavenumber k_x. The diagonal elements of the resolution matrices are displayed in an inset in these figures.

From the eigenvalues and associated eigenvectors of the Hessian shown in Fig. 6 we see that in a traveltime inversion the objective function is most sensitive to the model components with smaller wavenumber, typically to the zero-valued component. However, in amplitude inversion the most sensitive components seem to prefer large k_x values but mid-range k_z values, as seen from the first and second eigenvectors associated with the largest eigenvalues. For the surface geometry that was employed, both mid-range k_z and large k_x cause large transverse slowness derivatives, which focus or defocus energy. Neele *et al.* (1993a) explicitly showed the dependence of amplitudes on the higher transverse derivatives of the slowness field and, therefore, on the higher-wavenumber components. This dependence is also illustrated in their inversions using real data (Neele, VanDecar & Snieder 1993b).

To calculate the model resolution matrix of a pseudo-matrix inversion, we truncate eigenvalues at 10^{-3}, as in the previous section with interface inversion, indicated by a dash-dotted line in Fig. 6. The model resolution matrices shown in the bottom of the figure clearly tell us which slowness component can or cannot be constrained by the pseudo-inverse. Slowness components of vertical variation at higher (absolute) horizontal wavenumber k_x can be resolved by both traveltime and amplitude inversions with the L^2 norm chosen in model space. When the (absolute) horizontal wavenumber decreases, the constraint on vertical variation becomes weaker and weaker. When $k_x = 0$, neither inversion can constrain the vertical variation. These observations are summarized by the 6×11 matrix-like rectangle shown in the inset on each model resolution matrix, consisting of the diagonal elements of the model resolution matrix. When the diagonal element is equal to 1, the corresponding component of the solution is the true model parameter. If the element is zero, the inversion cannot constrain the model component at all. The diagonal elements clearly show that neither inversion can resolve wavenumbers in the purely vertical direction and that traveltimes cannot constrain any vertical variation with high wavenumber k_z, but that amplitude inversion can apparently constrain the vertical variation of slowness at a high horizontal wavenumber k_x.

5.3 Sensitivity analysis with the penalization of spatial derivatives

Let us now examine the sensitivities shown in Fig. 7, in which the H^1 norm was chosen in the model space. In the traveltime inversion, the sensitivities to the wavenumber appear almost identical to the case with the L^2 norm shown in Fig. 6. For the amplitudes, the eigenvectors corresponding to the largest eigenvalues show that they are still sensitive to the combination of the components with high-wavenumber $\mathbf{k}$, as with the L^2 norm, but that additional sensitivity to the components of intermediate wavenumber has been created. The sensitivity to the intermediate components is due to the use of the smoothing regularization. If the degree of smoothing increases, the inversion is still more sensitive to the small-wavenumber

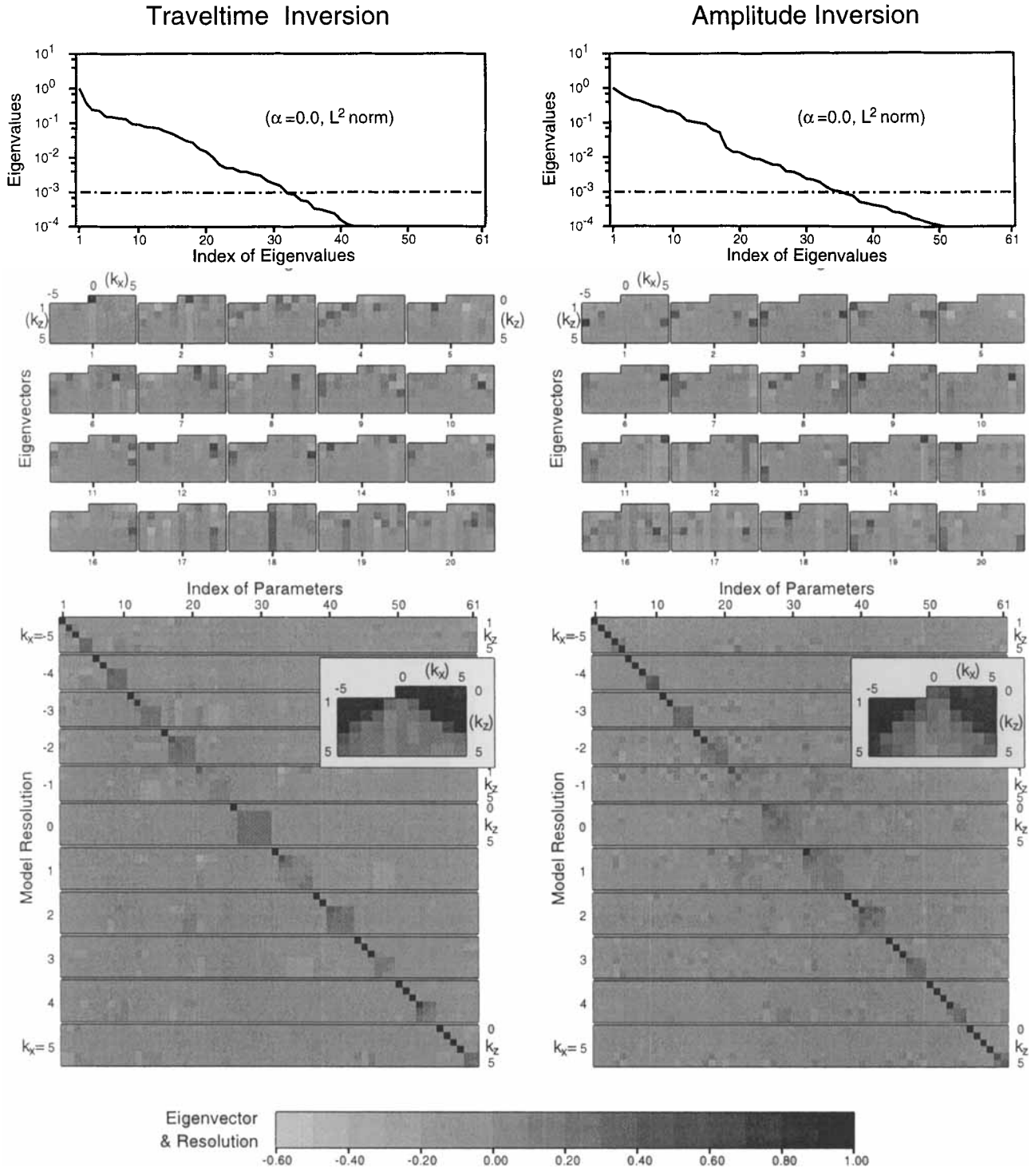

Figure 6. Eigenvalues and eigenvectors of the Hessian matrix and the model resolution matrices for traveltime and amplitude inversions for slowness variation. The L^2 norm is chosen in the model space. In each eigenvector image, horizontal wavenumbers of the model components increase from left to right (with the wavenumber ranging from -5 to 5 km^{-1}) and the vertical wavenumbers increase from top to bottom (from zero to 5 km^{-1}). Only those 20 eigenvectors associated with the largest eigenvalues of the Hessian are shown. The diagonal elements of each resolution matrix, computed using a pseudo-inverse by SVD truncation, are shown in the matrix-like rectangular inset.

components, as can be seen from the result with $\alpha = 0.99$ shown in Fig. 8.

In a reflection amplitude inversion, when we set $\alpha = 0.99$ in the scalar product of model perturbation, the objective function is sensitive to perturbations in the background slowness a_{00}, as seen from the first and second eigenvectors. This apparent sensitivity is due to the constraints. The amplitudes are still sensitive, however, to the slowness components with larger (absolute) wavenumbers (shorter wavelengths), as shown in eigenvectors 1 and 2. For the traveltime inversion,

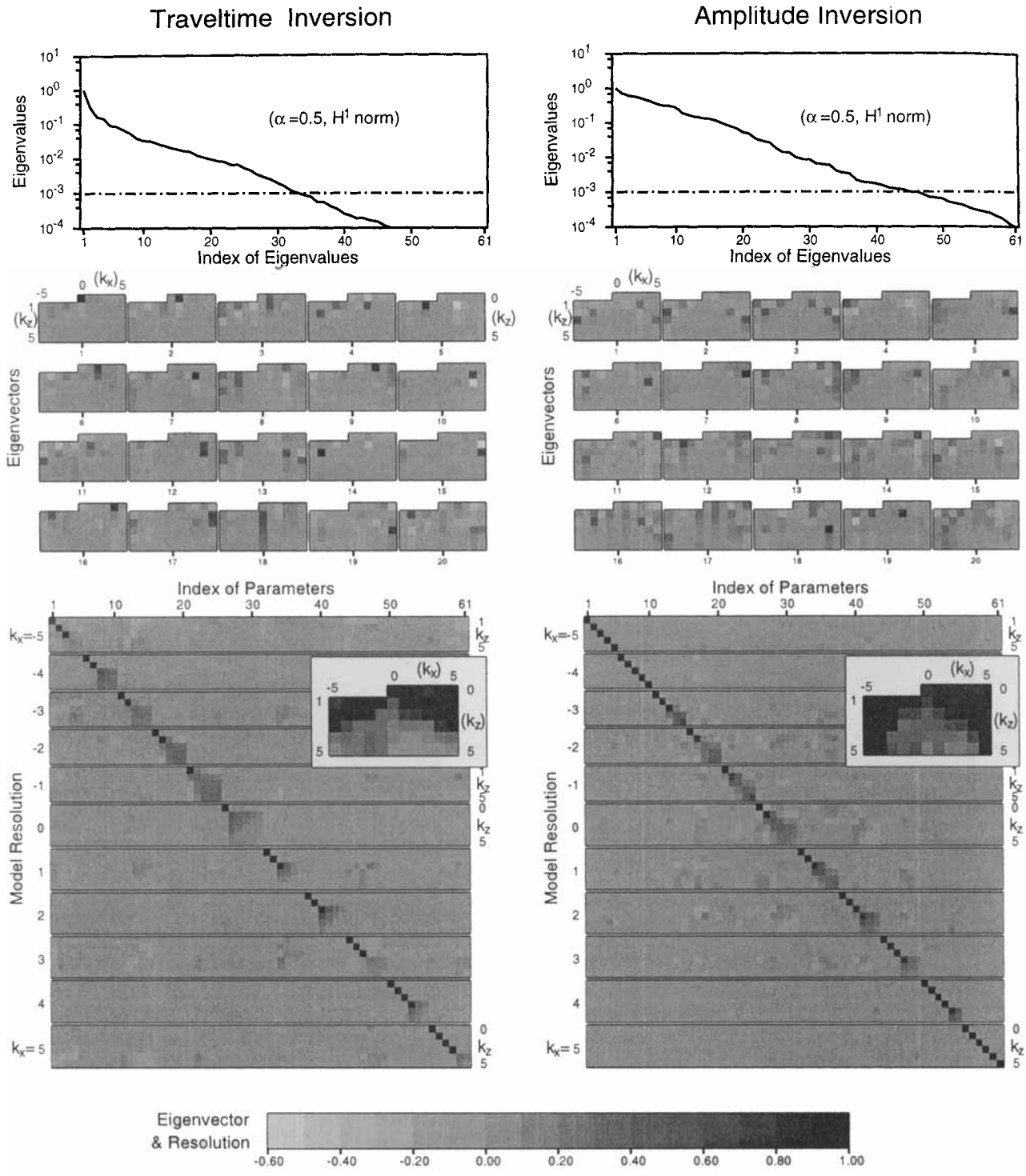

Figure 7. Eigenvalues, associated eigenvectors of the Hessian matrix, and the model resolution matrices for traveltime and amplitude inversions for slowness variation. The H^1 norm ($\alpha = 0.50$) is chosen in the model space. See Fig. 6 caption for comments.

the absolute value of the largest eigenvalue is increased, due to the stronger penalization of the spatial derivatives. The model resolution matrices, after the rejection of eigenvectors associated with small eigenvalues (Fig. 8), show that the traveltime inversion can only constrain the slowness components defining horizontal variation with $n=0$ and $n=1$. However, the amplitudes provide information on the high-wavenumber components. The model resolution matrices for the amplitude inversion, for both $\alpha=0.50$ and $\alpha=0.99$, are almost an identity matrix. The different information content of amplitudes and traveltimes is again evident.

From the sensitivities shown in Figs 6, 7 and 8 we see that the eigenvectors clearly show a symmetrcal pattern for $k_x > 0$ and $k_x < 0$ in each case. Therefore, ambiguities between different harmonic components inevitably exist in the inversion.

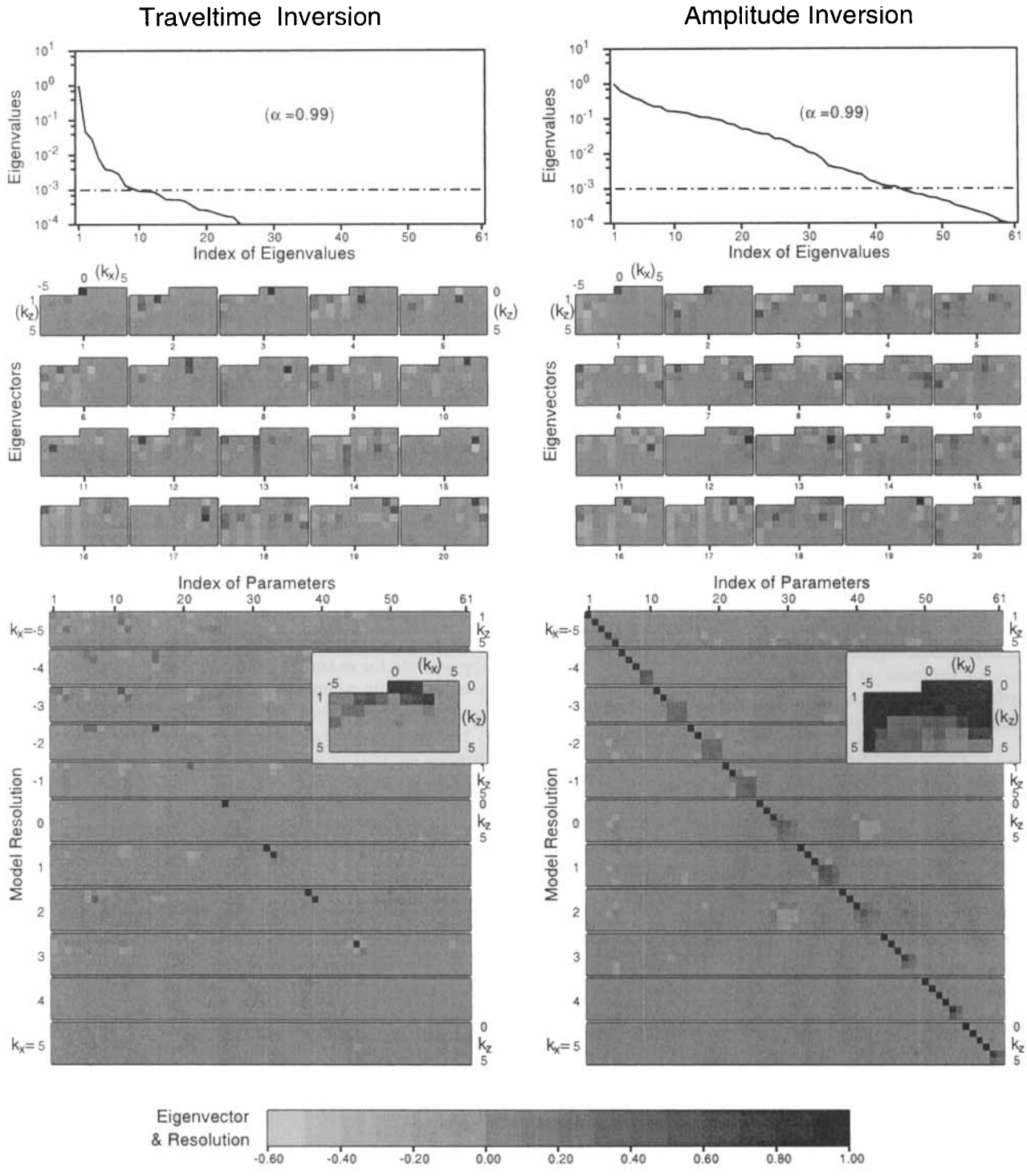

Figure 8. Eigenvalues, associated eigenvectors of the Hessian matrix, and the model resolution matrices for traveltime and amplitude inversions for slowness variation, where $\alpha = 0.99$ is used. See Fig. 6 caption for comments.

5.4 Cooperative inversion of amplitudes and traveltimes (for slowness variations)

For a cooperative inversion using both amplitude data and traveltime data simultaneously, the same definition of the data covariance matrix as eq. (44) is used here for slowness variation. In eq. (44) the parameter κ is used to balance the relative contributions of the traveltime and the amplitude data and is determined by using the ratio of the traces of the two sensitivity matrices, F_{time} and F_{ampl}.

For the cooperative inversion of both traveltimes and amplitudes for interface geometry (shown in previous section), we determined a value of $\kappa = 1.685 \times 10^{-2}$; the traveltimes were not as sensitive to the geometry as the amplitudes were. In contrast, for this example in which we are inverting the same data for the interval slowness variation, we find the ratio of

traces of the sensitivity matrices is $\kappa = 16.81$; the traveltimes are significantly more sensitive to the slowness model than the amplitudes are. The lower sensitivity of amplitudes to the slowness variations is probably due to the fact that amplitudes sense slowness gradients and curvatures, rather than the absolute values of the slownesses. Once again, this is an indication that amplitudes and traveltimes are sensitive to different physical parameters. This is an important result, with the potential benefit that it may be possible to resolve the known ambiguity between reflector depth uncertainty and interval velocity uncertainty better by including amplitude data in seismic-reflection traveltime tomography.

We performed experiments that show that the range 0.75–1.25 for the second balancing factor p is appropriate in this case. For $p \leq 1.0$, the contributions of traveltimes and amplitudes when the L^2 norm is used are balanced, whereas traveltimes dominate the inversion with the H^1 norm. For $p > 1.0$, amplitudes dominate the inversion with the L^2 norm, whereas traveltimes and amplitudes are balanced in the case of inversion with the H^1 norm. The result with the balancing factor $p = 1.0$ is shown in Fig. 9.

From Fig. 9 we see that, with the L^2 norm used in the model space, the objective function is most sensitive to the background slowness a_{00}, dominated by traveltime data, and then to components with high-wavenumber $\mathbf{k}$, dominated by amplitude data. With the H^1 norm in the inversion, the most significant parameters in the slowness model are the background slowness a_{00} and those components defining the horizontal variation of the slowness. The sensitivities to the variation with high-wavenumber $\mathbf{k}$ are reduced relative to the background slowness.

6 INVERSION EXAMPLES

In previous sections we have considered a cooperative inversion using both traveltimes and amplitudes simultaneously. In this section, we will show some synthetic examples of inversion for interface geometry and velocity variation separately, and explore how the inclusion of amplitudes in the tomography improves the solution of traveltime inversion.

6.1 Inversion formula

The solution for the linearized inverse problem (eq. 33) may be given by

$$\delta\mathbf{m} = (\mathsf{F}^+ \mathsf{C}_\mathrm{D}^{-1} \mathsf{F} + \mu \hat{\mathsf{I}})^{-1} \mathsf{F}^+ \mathsf{C}_\mathrm{D}^{-1} \delta\mathbf{d}, \tag{49}$$

where μ is a dimensionless damping factor introduced to stablize the inverse procedure, $\hat{\mathsf{I}}$ is the identity matrix with units of *(model parameter)*$^{-2}$, and $\mathsf{C}_\mathrm{D}^{-1}$ is the data covariance matrix containing the two weighting parameters, κ and p, discussed earlier. Eq. (49) contains the full Hessian matrix $\mathsf{H} = \mathsf{F}^+ \mathsf{C}_\mathrm{D}^{-1} \mathsf{F} = \mathsf{D}_\alpha^{-1} \mathsf{F}^\mathrm{T} \mathsf{C}_\mathrm{D}^{-1} \mathsf{F}$. In spite of the use of the additional constraint D_α ($\alpha \neq 0.0$), we still require a strategy for selecting the damping factor, μ; as we have seen, even with the H^1 norm, the Hessian is a singular matrix. The choice of μ is addressed below.

Because the adjoint F^+ of the matrix F is given by $\mathsf{F}^+ = \mathsf{D}_\alpha^{-1} \mathsf{F}^\mathrm{T}$ (eq. A11), eq. (49) can be expressed as

$$\delta\mathbf{m} = (\mathsf{F}^\mathrm{T} \mathsf{C}_\mathrm{D}^{-1} \mathsf{F} + \mu \mathsf{D}_\alpha \hat{\mathsf{I}})^{-1} \mathsf{F}^\mathrm{T} \mathsf{C}_\mathrm{D}^{-1} \delta\mathbf{d}, \tag{50}$$

where the operator D_α is introduced to penalize the first spatial derivatives of the model, and can then be alternatively understood as a working definition of the model covariance matrix, $\mathsf{C}_\mathrm{M}^{-1}$, expressing expected correlations between different model parameters. In fact, if we set $\mathsf{C}_\mathrm{M}^{-1} = \mathsf{D}_\alpha \hat{\mathsf{I}}$, eq. (50) will be the solution of well-known stochastic inversion (Franklin 1970; Jackson 1979; Tarantola & Valette 1982; Tarantola 1987), which stablizes the inverse problem by adding a term to the data misfit function that depends on the *a priori* model, $\mathbf{m}_\mathrm{ref}$, and its covariance matrix,

$$S(\mathbf{m}) = \frac{1}{2} \{ (\mathbf{f}(\mathbf{m}) - \mathbf{d}_\mathrm{obs})^\mathrm{T} \mathsf{C}_\mathrm{D}^{-1} (\mathbf{f}(\mathbf{m}) - \mathbf{d}_\mathrm{obs}) + \mu (\mathbf{m} - \mathbf{m}_\mathrm{ref})^\mathrm{T} \mathsf{C}_\mathrm{M}^{-1} (\mathbf{m} - \mathbf{m}_\mathrm{ref}) \}, \tag{51}$$

where the scalar μ acts as a trade-off parameter that controls the relative weights of *a priori* information and the observed data. The model covariance matrix has been extensively used to remove numerical instabilities by damping poorly constrained parameters towards the reference model, and allowing only well-constrained parameters to be controlled by the data. However, the dependence of the second term in eq. (51) on the reference model $\mathbf{m}_\mathrm{ref}$ may result in unwarranted structure. In contrast, the model regularization given by the operator D_α is defined by the scalar product of the basis functions $\{\beta_i(\mathbf{x}), \forall i (1 \leq i \leq M)\}$ (and not by the model $\mathbf{m}$ or the perturbation $\delta\mathbf{m}$, see eq. 38). Thus, the form of the constraint we used in eq. (50) does not depend on the starting model or on any arbitrary reference model. This would appear to be more appropriate for examining lateral variations in seismic structure, especially when no strong *a priori* model exists for the region (Constable, Parker & Constable 1987; Sambridge 1990).

In the following synthetic examples we use the inversion formula eq. (50), in which $\alpha = 0.50$ (the H^1 norm) is arbitrarily chosen in the scalar product D_α, which contains an intermediate degree of penalization of the spatial derivatives. The parameters κ and p used for $\mathsf{C}_\mathrm{D}^{-1}$ are the same as in the sensitivity analyses. The trade-off parameter μ is determined using the relation $\mu = \rho / r$, where

$$r = \frac{\mathrm{trace}(\mathsf{D}_\alpha \hat{\mathsf{I}})}{\mathrm{trace}(\mathsf{F}^\mathrm{T} \mathsf{C}_\mathrm{D}^{-1} \mathsf{F})} \tag{52}$$

is an empirical quantity used to normalize the matrices. Eq. (50) is thus rewritten as

$$\delta\mathbf{m} = r (r\, \mathsf{F}^\mathrm{T} \mathsf{C}_\mathrm{D}^{-1} \mathsf{F} + \rho\, \mathsf{D}_\alpha \hat{\mathsf{I}})^{-1} \mathsf{F}^\mathrm{T} \mathsf{C}_\mathrm{D}^{-1} \delta\mathbf{d}, \tag{53}$$

where the dimensionless factor ρ easily controls the total amount of regularization in the inversion.

6.2 Interface inversion

Consider an interface (shown in Fig. 10a) which is simply defined by eq. (40), where the amplitude coefficients a_0, a_1, a_2, a_{20} and a_{40} are 1980, 50, 15, 15 and 5 m, respectively, the remaining coefficients are zero and the fundamental wavenumber k_0 is 1/16 km^{-1}. The amplitude coefficients a_n of the cosine terms and b_n of the sine terms are also shown, plotted against wavenumbers in the figure. For the calculation of synthetic data, we used the same observation configuration as in the previous sections. In the inversion, the starting model is given by a flat straight reflector at a depth of 2000 m (the same uniform structure used in sensitivity analyses). The number of harmonics in the parametrization is set at $N = 80$ in eq. (40).

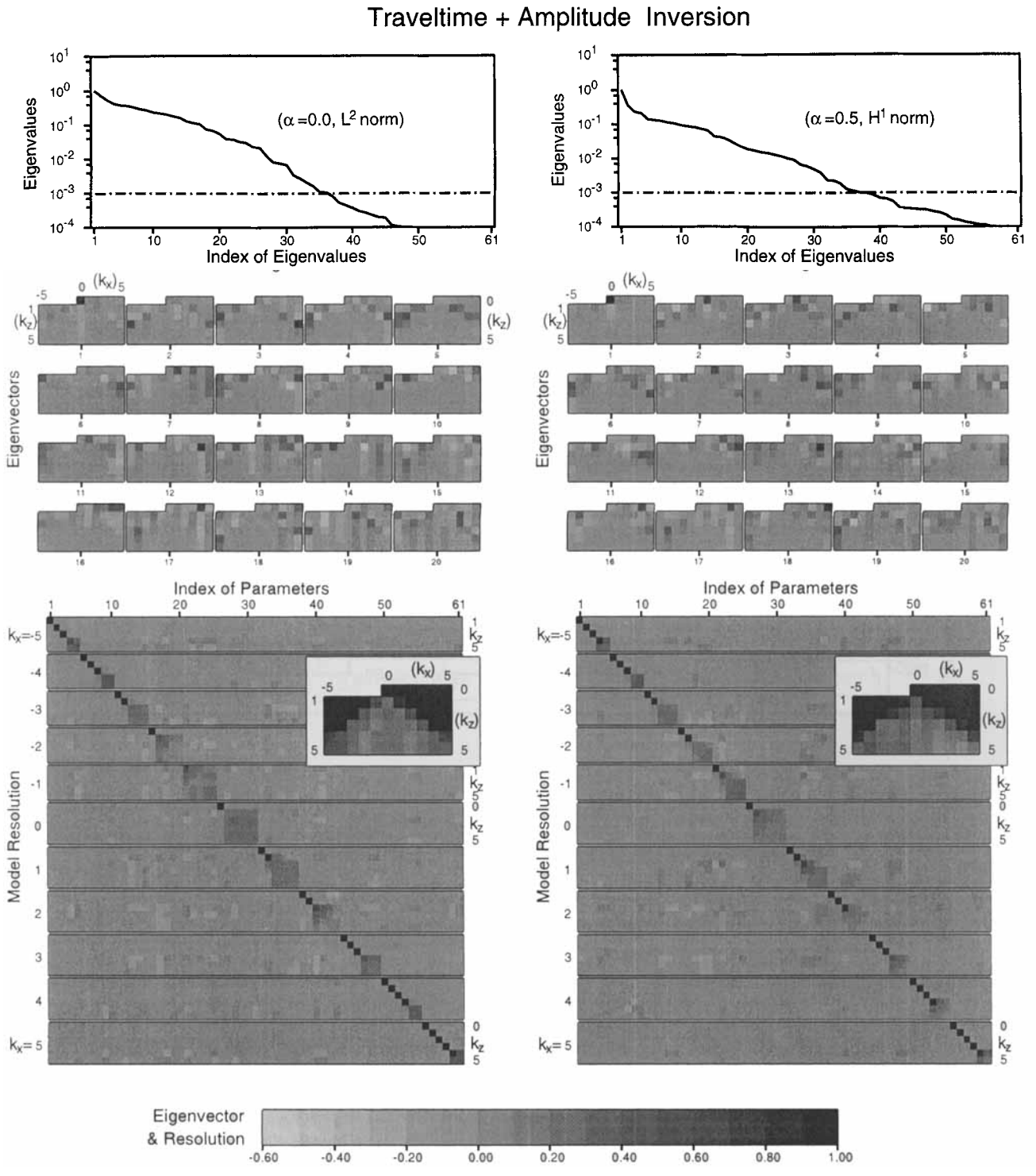

Figure 9. Eigenvalues and eigenvectors of the Hessian matrix, and the model resolution matrices in the case of a cooperative inversion of traveltime and amplitude data for slowness variation. See Fig. 6 caption for comments.

We therefore have 161 parameters (81 cosine and 80 sine terms) in the inversion.

The solutions of traveltime inversion, amplitude inversion and the cooperative inversion using both types of data are shown in Figs 10(b), 10(c) and 10(d), respectively. Although the problem would normally be solved iteratively, the results after only one iteration are sufficient to show the dramatic differences between the inversions. These inversions were run with the factor ρ in eq. (53) set equal to 0.01. Recall that in the case of traveltime inversion with the H^1 norm, the zero- and low-wavenumber components are the most sensitive components (see Fig. 5a). However, from Fig. 10(b) we see that the solutions, after the first iteration, for the components with zero wavenumber and low wavenumber are compromised. In practice, more iterations may be needed, and the Fréchet matrix for each iteration will need to be recalculated. In the

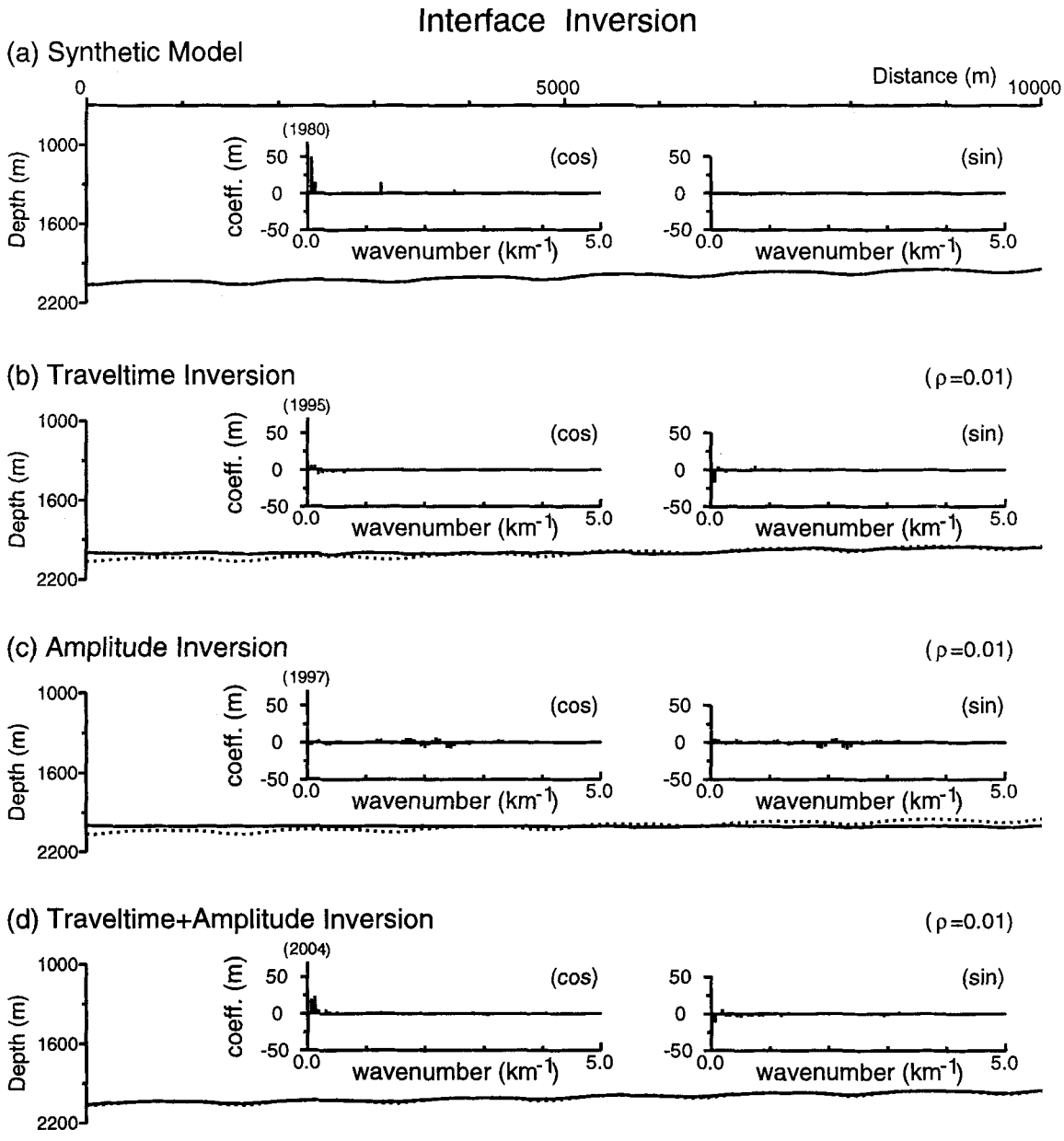

Figure 10. An example of interface inversion: (a) synthetic model; (b) traveltime inversion; (c) amplitude inversion; and (d) the cooperative inversion using both traveltime and amplitude data simultaneously. Dotted lines in (b), (c) and (d) are the true interface, which is compared with the inversion solution (solid lines) after the first iteration. Model parameters, the amplitude coefficients of the Fourier components, are also depicted in the figure.

case of the amplitude inversion shown in Fig. 10(c), we see that slight changes to interface components with intermediate wavenumber have been obtained and that the long-wavelength components were not obtained. In the cooperative inversion using both types of data, however, the solution (Fig. 10d) is dramatically improved in a single iteration. The significance of these figures is that neither inversion alone (Figs 10b and c) is successful in resolving the two smallest wavenumbers from each other, but that the cooperative inversion (Fig. 10d) does begin to resolve these uniquely.

Fig. 11 shows the comparison of cooperative inversions with different values of ρ applied. Penalizing first derivatives, introduced by C_M^{-1} or the operator D_α ($\alpha=0.5$, the H^1 norm) in eq. (53), implies that a near-horizontal solution is to be preferred. As the value of ρ decreases from 1.0 (Fig. 11a) to 0.001 (d), the inversion gradually allows the inclusion of larger slopes. In this case, the slope of the interface is dominated by the amplitude coefficients of harmonic terms with small wavenumbers. It is clearly seen from the figure that those amplitude coefficients gradually increase as ρ decreases.

6.3 Slowness inversion

To demonstrate the inversion for velocity variation, we show a synthetic model in Fig. 12(a), simply defined by eq. (46), in which we set $a_{00}=0.4$, $a_{10}=-0.001$, $a_{20}=0.005$, $a_{13}=-0.004$, $a_{23}=-0.005$, $b_{10}=0.004$, $b_{20}=-0.006$, $b_{12}=-0.008$ and $b_{22}=0.002$ (s km^{-1}) and the remaining coefficients are zero-valued and $k_0=1.0$ km^{-1}. A graphical representation of the amplitude coefficients a_{mn} and b_{mn} is also shown in the figure.

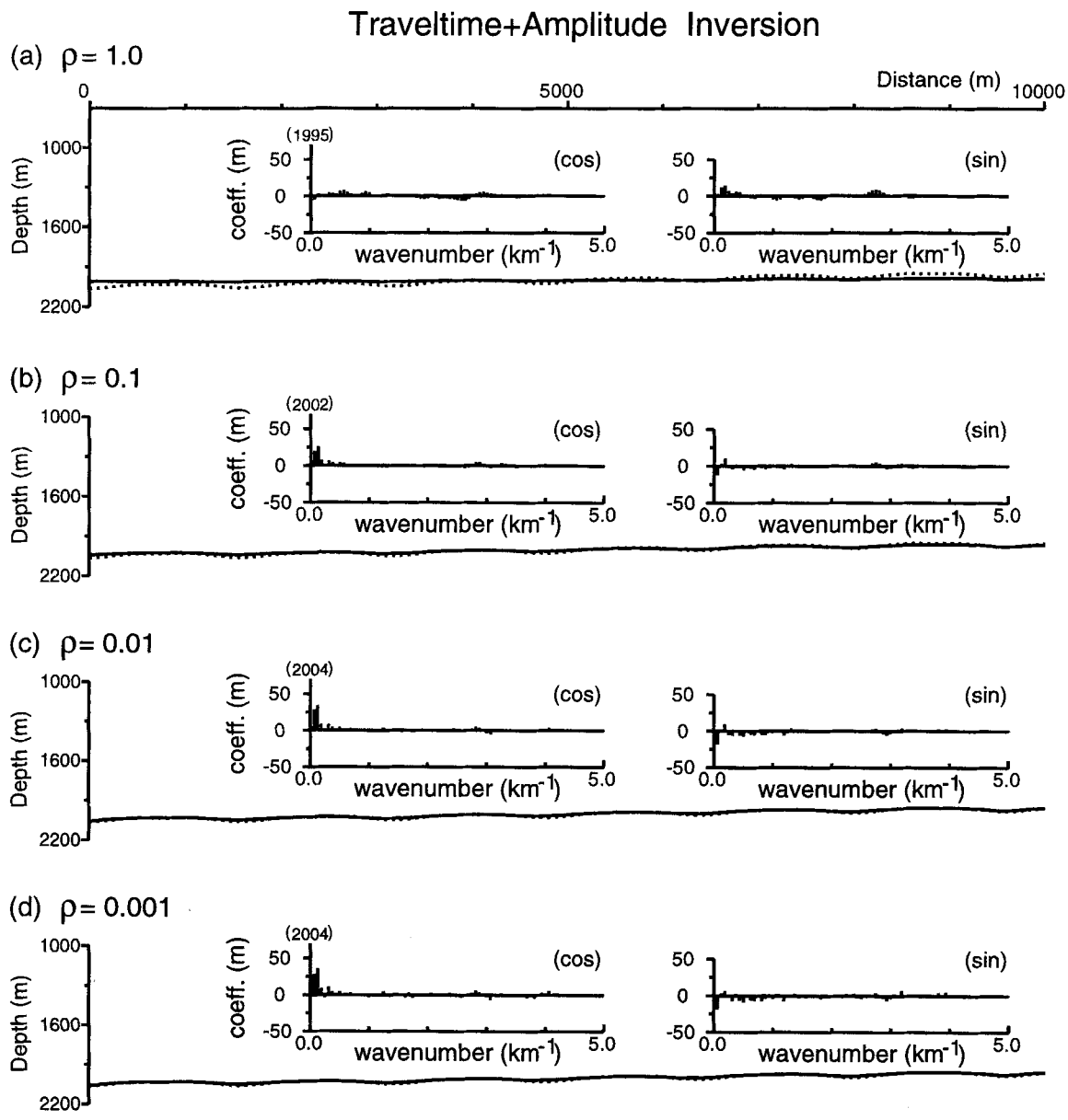

Figure 11. Four results of cooperative inversion using both traveltime and amplitude data simultaneously for interface geometry which differ in the value of ρ used in inversion formula eq. (53).

In the inversion we set $N=5$ in the model parametrization, therefore a total of 121 parameters (61 cosine coefficients a_{mn} and 60 sine coefficients b_{mn}) will be determined. The starting model is given by the constant background slowness (a_{00}). The Fréchet matrix based on this starting model was used in the previous section for the sensitivity analysis.

The results of traveltime and amplitude inversions and the cooperative inversion, after the first iteration, are shown in Figs 12(b), 12(c) and 12(d), respectively. In those inversions $\rho=0.01$ is used. From Fig. 12 we see that both traveltime and amplitude inversions show some image of velocity variation, but far from the true model. It is interesting to note that the amplitude inversion seems to recover anomalies close to the surface, and close to the interface, whereas the traveltime inversion seems to recover the anomalies in between. The cooperative inversion shows an encouraging result, in which the basic features of the velocity variation are reconstructed.

Fig. 13 presents four solutions of the cooperative inversion using both traveltime and amplitude data simultaneously, which differ in the value of ρ used. Compared with the synthetic model (Fig. 12a), we see that the anomalous structure is clearly reconstructed by using the constraint of both types of data when we set $\rho=0.1$ (b) or $\rho=0.01$ (c). When $\rho=1.0$ (a) the inversion gives the smoothest model (in the sense of the first derivatives of the slowness), whereas when $\rho=0.001$ (d) it gives the roughest structure. Clearly, the value of ρ cannot be indefinitely reduced; it is required in order to stabilize the inversion in the presence of noise.

In this section we have considered a simultaneous inversion for the slowness variation, using both types of data (traveltimes and amplitudes). However, the numerical analyses carried out by Neele *et al.* (1993a) suggest that, when amplitudes and traveltimes are combined in an inversion for the velocity structure, a non-linear traveltime inversion must be performed

before amplitude data are included, ensuring that the non-linear behaviour of amplitudes due to ray shift induced by slowness perturbations is minimized. This is because traveltimes show a more linear dependence on slowness perturbations than amplitudes do, whereas amplitudes depend on the spatial derivatives of slowness distribution. The success in the simultaneous inversion shown in this paper is partially due to the application of operator D_α, which penalizes the first derivatives of slowness variation, and partially due to the data covariance matrix that we proposed, which balances the contribution, in terms of relative sensitivities, of different types of data.

7 CONCLUSIONS

The sensitivities of reflection-seismic amplitudes and traveltimes to the variation of interface geometry and slowness distributions have been compared. In general, the information content of the traveltimes and the amplitudes are complementary, being sensitive to different features of the model. Traveltimes are more sensitive to model components with smaller wavenumbers, but amplitudes are more sensitive to the components defined by the larger wavenumbers.

The ray theory used in this paper does not account for the finite extent of the Fresnel zone. Under the ray approximation, the sensitivity of amplitudes to lateral slowness or interface variations is then always dominated by the high-end k_x values. The amplitude depends on the second derivatives of the traveltime function; short-wavelength structure has the largest effect on these derivatives, causing large amplitude fluctuations. The way around this is to include the Fresnel zone in the computations of the Fréchet derivatives, removing the local character of ray theory. Snieder & Lomax (1996) present a way to do this in an intuitive fashion, but more exact methods are possible, although expensive. In this way, the sensitivity of amplitudes would become a maximum for scale-lengths comparable to the size of the Fresnel zone, with lower sensitivity for both larger and smaller scale-lengths, without the need for elaborate weighting schemes or roughness penalization terms in the misfit function. Our solution of cutting off the parametrization of the model at scale-lengths comparable to the size of the Fresnel zone is a first step towards a more complete and accurate description of the dependence of amplitudes on slowness and discontinuity structure and, therefore, the main conclusions of the present paper would be unaltered if a more exact forward method were used. However, the size of the amplitude Fréchet derivatives would naturally be different.

The investigation here of amplitude sensitivity confirms the observations from the inversion examples reported by Wang & Houseman (1994, 1995). Although the previous work showed that seismic amplitude data contain information which independently constrains aspects of both reflector geometry and slowness variation, the amplitude inversion should not be viewed as an alternative to reflection traveltime tomography. It should be viewed as a complementary set of constraints on the inversion problem for reflector geometry and interval slowness. For the cooperative inversion using both traveltime and amplitude data simultaneously, an empirical definition of the data covariance matrix which balances the relative sensitivities of different types of data has been proposed. The inversion results shown in this paper suggest that in this manner a cooperative inversion can provide a much better solution than that using one type of data alone.

Amplitudes and traveltimes are also sensitive to different physical parameters. The ratio of traces of the sensitivity matrices has shown that traveltimes are more sensitive to variations in the slowness model than amplitudes are, whereas amplitudes are more sensitive to variations in interface geometry than traveltimes are. Therefore, we may design a cost-effective joint inversion for slowness variation and interface geometry, using both types of data, in which traveltimes dominate the determination of the slowness variation and amplitudes dominate the determination of the interface geometry. Because amplitudes can provide better constraints on model components with shorter wavelengths, and traveltimes are more sensitive to longer-wavelength components, the cooperative inversion may resolve the known ambiguity between reflector depth uncertainty and interval slowness uncertainty better in seismic-reflection traveltime tomography. This prospect needs to be investigated by further inversion tests.

ACKNOWLEDGMENTS

The work presented in this paper was a project undertaken by the London University Seismic Research Consortium. We sincerely thank Prof. Michael Worthington for his support of this research. Thanks also to the project sponsors, Amoco (UK) Exploration Company, Elf UK plc, Enterprise Oil plc, Fina Exploration Ltd, Mobil North Sea Ltd and Texaco Britain Ltd. Constructive comments from Dr Greg Houseman of Monash University and two anonymous reviewers are gratefully acknowledged.

REFERENCES

Aki, K. & Richards, P., 1980. *Quantitative Seismology: Theory and Method,* W. H. Freeman, San Francisco, CA.

Bishop, T.N. *et al.*, 1985. Tomographic determination of velocity and depth in laterally varying media, *Geophysics,* **50,** 903–923.

Bregman, N.D., Bailey, R.C. & Chapman, C.H., 1989. Ghosts in tomography: the effects of poor angular coverage in 2-D seismic traveltime inversion, *Can. J. expl. Geophys.*, **25,** 7–27.

Bube, K.P., Langan, R.T. & Resnick, J.R., 1995. Theoretical and numerical issues in the determination of reflector depths in seismic reflection tomography, *J. geophys. Res.*, **100,** 12 449–12 458.

Burridge, R., 1976. *Some Mathematical Topics in Seismology,* Courant Institute of Mathematical Sciences, New York University, New York, NY.

Červený, V., 1985. The application of ray tracing to numerical modelling of seismic wave fields in complex structures, in *Seismic Shear Waves, Part A: Theory,* pp. 1–124, ed. Dohr., G.P., Geophysical Press, London.

Červený, V. & Ravindra, R., 1971. *Theory of Seismic Head waves,* University of Toronto Press, Toronto, Canada.

Chapman, C.H. & Drummond, R., 1982. Body-wave seismograms in inhomogeneous media using Maslov asymptotic theory, *Bull. seism. Soc. Am.*, **72,** S277–S317.

Constable, S.C., Parker, R.L. & Constable, C.G., 1987. Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics,* **52,** 289–300.

Delprat-Jannaud, F. & Lailly, P., 1992. What information on the Earth model do reflection traveltimes provide?, *J. geophys. Res.*, **97,** 19 827–19 844.

Delprat-Jannaud, F. & Lailly, P., 1993. Ill-posed and well-posed formulations of the reflection traveltime tomography problem, *J. geophys. Res.*, **98,** 6589–6605.

Farra, V. & Le Bégat, S., 1995. Sensitivity of *qP*-wave traveltimes and polarization vectors to heterogeneity, anisotropy and interfaces, *Geophys. J. Int.*, **121,** 377–390.

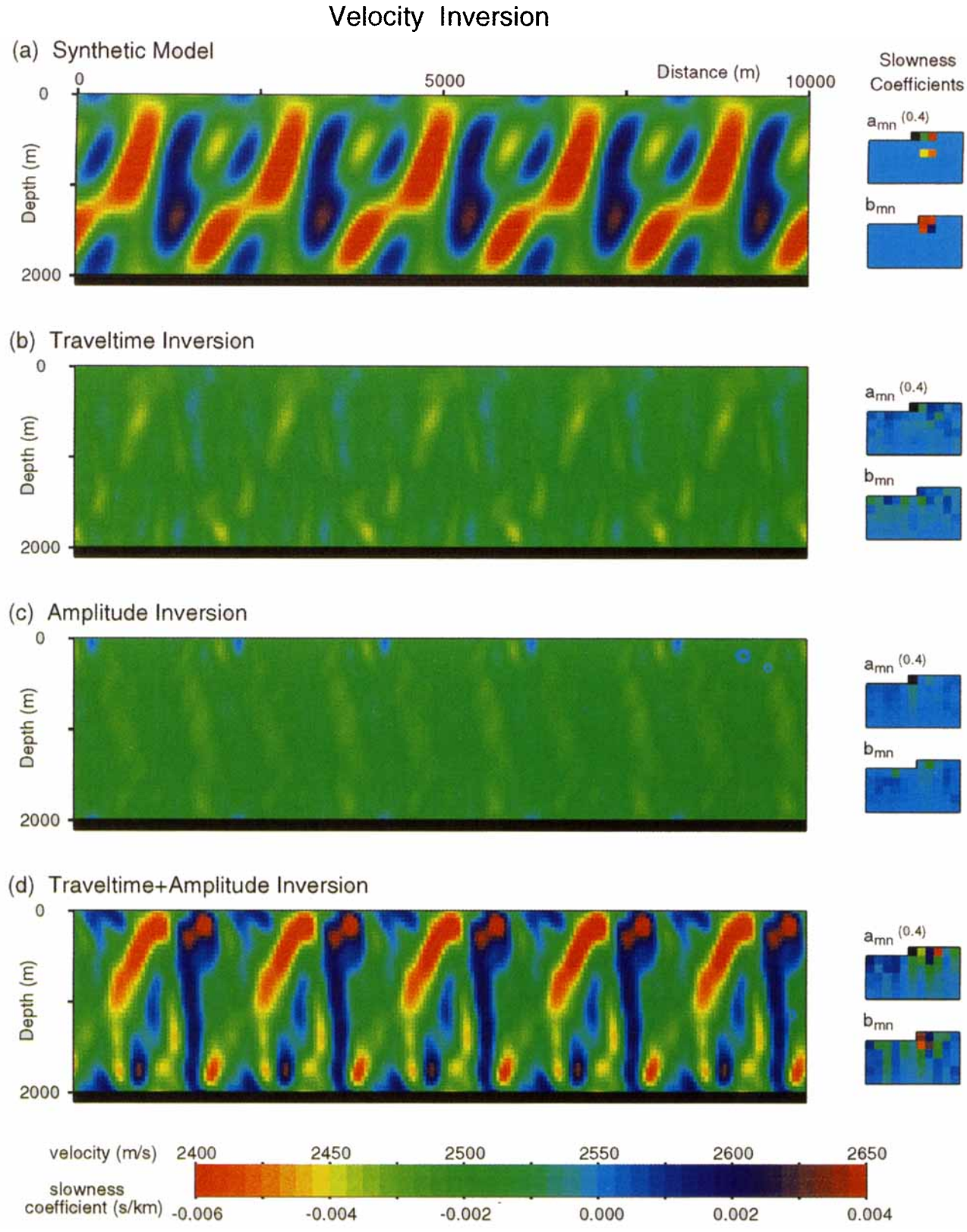

Figure 12. An example of velocity inversion: (a) synthetic model; (b) traveltime inversion; (c) amplitude inversion; and (d) the cooperative inversion using both traveltime and amplitude data simultaneously, where $\rho = 0.01$ is used in the inversion (eq. 53) to control the amount of regularization in the inversion. Model parameters, the coefficients of the Fourier components, are graphically represented as images on the right side of the figure.

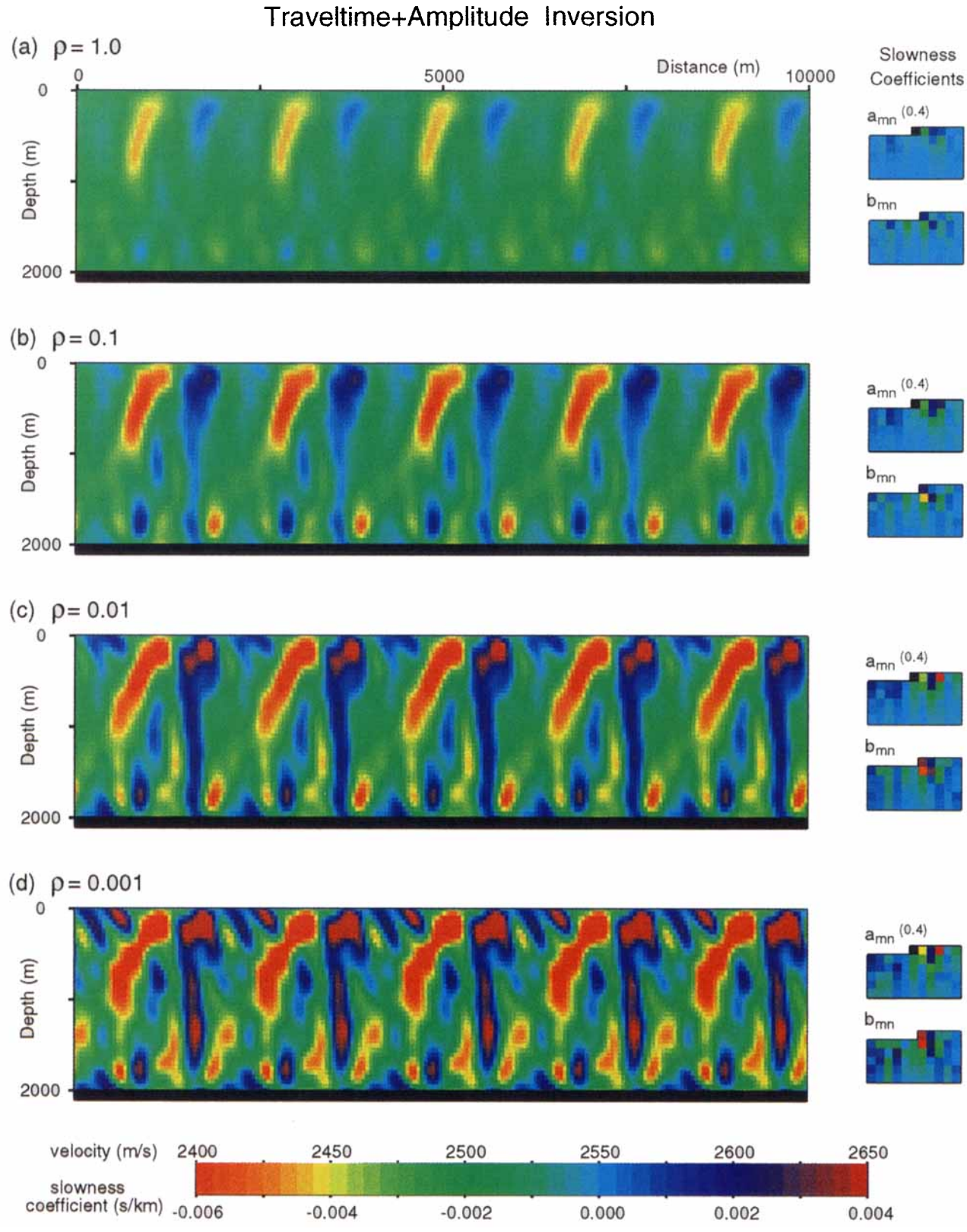

Figure 13. Four results of the cooperative inversion using both traveltime and amplitude data simultaneously for velocity variation which differ in the value of ρ used in inversion formula (eq. 53).

Farra, V. & Madariaga, R., 1987. Seismic waveform modelling in heterogeneous media by ray perturbation theory, *J. geophys. Res.*, **92,** 2697–2712.

Farra, V. & Madariaga, R., 1988. Non-linear reflection tomography, *Geophys. J.*, **95,** 135–147.

Farra, V., Virieux, J. & Madariaga, R., 1989. Ray perturbation theory for interfaces, *Geophys. J. Int.*, **99,** 377–390.

Franklin, J. N., 1970. Well-posed stochastic extensions of ill-posed linear problems, *J. Math. Anal. Appl.*, **31,** 682–716.

Gajewski, D. & Psencik, I., 1990. Vertical seismic profile synthetics by dynamic ray tracing in laterally varying layered anisotropic structures, *J. geophys. Res.*, **95,** 11 301–11 315.

Gilbert, F. & Backus, G.E., 1966. Propagator matrices in elastic wave and vibration problems, *Geophysics,* **31,** 326–333.

Jackson, D.D., 1979. The use of *a priori* data to resolve nonuniqueness in linear inversion, *Geophys. J. R. astr. Soc.*, **57,** 137–157.

Kline, M. & Kay, I.W., 1965. *Electromagnetic Theory and Geometrical Optics,* John Wiley & Sons, New York, NY.

Neele, F., VanDecar, J.C. & Snieder, R., 1993a. A formalism for including amplitude data in tomographic inversions, *Geophys. J. Int.*, **115,** 482–496.

Neele, F., VanDecar, J.C. & Snieder, R., 1993b. The use of *P*-wave amplitude data in a joint tomographic inversion with travel times for upper-mantle velocity structure, *J. geophys. Res.*, **98,** 12 033–12 054.

Nowack, R.L. & Lutter, W.J., 1988. Linearized rays, amplitude and inversion, *Pure appl. Geophys.*, **128,** 401–421.

Nowack, R.L. & Lyslo, J.A., 1989. Fréchet derivatives for curved interfaces in the ray approximation, *Geophys. J.*, **97,** 497–509.

Ory, J. & Pratt, R.G., 1995. Are our parameter estimators biased? The significance of finite-difference regularization operators, *Inverse Problem* **11,** 397–424.

Pratt, R.G. & Chapman, C.H., 1992. Traveltime tomography in anisotropic media—II. Application, *Geophys. J. Int.*, **109,** 20–37.

Sambridge, M.S., 1990. Non-linear arrival time inversion: constraining velocity anomalies by seeking smooth models in 3-D, *Geophys. J. Int.*, **102,** 653–677.

Sheriff, R.E. & Geldart, L.P., 1995. *Exploration Seismology,* 2nd edn, Cambridge University Press, Cambridge.

Snieder, R. & Sambridge, M., 1992. Ray perturbation theory for traveltimes and ray paths in 3-D heterogeneous media, *Geophys. J. Int.*, **109,** 294–322.

Snieder, R. & Sambridge, M., 1993. The ambiguity in ray perturbation theory, *J. geophys. Res.*, **98,** 22 012–22 034.

Snieder, R. & Lomax, A., 1996. Wavefield smoothing and the effect of rough velocity perturbations on arrival times and amplitudes, *Geophys. J. Int.*, **125,** 796–812.

Snieder, R. & Spencer, C., 1993. A unified approach to ray bending, ray perturbation and paraxial ray theories, *Geophys. J. Int.*, **115,** 456–470.

Stork, C., 1992a. Singular value decomposition of the velocity-reflector depth tradeoff, Part 1: Introduction using a two-parameter model, *Geophysics,* **57,** 927–932.

Stork, C., 1992b. Singular value decomposition of the velocity-reflector depth tradeoff, Part 2: High-resolution analysis of a generic model, *Geophysics,* **57,** 933–943.

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

Tarantola, A. & Vallette B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys. Space Phys.*, **20,** 219–232.

Thomson, C.J., 1983. Ray-theoretical amplitude inversion for laterally varying velocity structure below NORSAR, *Geophys. J. R. astr. Soc.*, **74,** 525–558.

Thomson, C.J. & Chapman, C.H., 1985. An introduction to Maslov's asymptotic method, *Geophys. J. R. astr. Soc.*, **83,** 143–168.

Wang, Y. & Houseman, G.A., 1994. Inversion of reflection seismic amplitude data for interface geometry, *Geophys. J. Int.*, **117,** 92–110.

Wang, Y. & Houseman, G.A., 1995. Tomographic inversion of reflection-seismic amplitude data for velocity variation, *Geophys. J. Int.*, **123,** 355–372.

Wang, Y., White, R.E. & Pratt, R.G., 1998. Geometrical inversion of seismic amplitudes from real, migrated reflection gathers, *Geophysics,* submitted.

Williamson, P.R., 1990. Tomographic inversion in reflection seismology, *Geophys. J. Int.*, **100,** 255–274.

APPENDIX A: THE OPERATOR D_α AND THE HESSIAN MATRIX H

A1 The operator D_α

We have seen that the form of the Hessian H depends on the definition of the scalar product $(\ ,\ )_M$ in the model space $\mathcal{M}$ (see eq. 30). We will relate this scalar product in $\mathcal{M}$ to a norm $\|\ \|_{M'}$ in $\mathcal{M}'$, the dual space of $\mathcal{M}$. Following Delprat-Jannaud & Lailly (1992, 1993), we define the Hilbertian norm for $\|\ \|_{M'}$; $\mathcal{M}'$ is thus a Hilbert space. Considering two model perturbations $\delta\mathbf{U}^{(1)}(\mathbf{x})$ and $\delta\mathbf{U}^{(2)}(\mathbf{x})$, which are the perturbations of reflector geometry or slowness variation or their combination (in general, not simply the perturbations of the model parameters $\delta\mathbf{m}$), the norm $\|\delta\mathbf{U}(\mathbf{x})\|_{M'}$ in the dual space $\mathcal{M}'$ is defined as

$$(\delta\mathbf{U}^{(1)}(\mathbf{x}),\ \delta\mathbf{U}^{(2)}(\mathbf{x}))_\alpha = (1-\alpha)\int_\Omega \delta\mathbf{U}^{(1)}(\mathbf{x})\delta\mathbf{U}^{(2)}(\mathbf{x})d\Omega + \alpha\int_\Omega \nabla_\mathbf{x}\delta\mathbf{U}^{(1)}(\mathbf{x})\nabla_\mathbf{x}\delta\mathbf{U}^{(2)}(\mathbf{x})d\Omega,$$
$$\forall(\delta\mathbf{U}^{(1)},\ \delta\mathbf{U}^{(2)})\in\mathcal{M}'\times\mathcal{M}', \qquad \text{(A1)}$$

where Ω is a generic term consisting of the reflector-geometry component and the slowness component of the model, and α is a weighting parameter which can take values within [0, 1).

Suppose we adopt the model parametrization defined by

$$\mathbf{U}(\mathbf{x}) = \sum_{i=1}^{M} c_i\beta_i(\mathbf{x}), \qquad \text{(A2)}$$

where c_i is the amplitude coefficient of the ith basis function $\beta_i(\mathbf{x})$. Denoting the projections on basis functions $\{\beta_i(\mathbf{x}),\ \forall i(1\le i\le M)\}$ as vectors (of the model parameters),

$$\delta\mathbf{m}^{(1)} = (\delta c_i^{(1)})_{(i=1,M)},$$
$$\delta\mathbf{m}^{(2)} = (\delta c_i^{(2)})_{(i=1,M)}, \qquad \text{(A3)}$$

for $\delta\mathbf{U}^{(1)}(\mathbf{x})$ and $\delta\mathbf{U}^{(2)}(\mathbf{x})$ respectively, the scalar product eq. (A1) is then given by

$$(\delta\mathbf{U}^{(1)}(\mathbf{x}),\ \delta\mathbf{U}^{(2)}(\mathbf{x}))_\alpha = \langle \mathsf{D}_\alpha\delta\mathbf{m}^{(1)},\ \delta\mathbf{m}^{(2)}\rangle,$$
$$\forall(\delta\mathbf{U},\ \delta\mathbf{m})\in\mathcal{M}'\times\mathcal{M}, \qquad \text{(A4)}$$

where D_α is an $M\times M$ symmetric definite positive matrix with tensors defined by

$$[D_{ij}]_\alpha = (1-\alpha)\int_\Omega \beta_i(\mathbf{x})\beta_j(\mathbf{x})\,d\Omega + \alpha\int_\Omega \nabla_\mathbf{x}\beta_i(\mathbf{x})\nabla_\mathbf{x}\beta_j(\mathbf{x})\,d\Omega,$$
$$\forall i(1\le i\le M),\ \forall j(1\le j\le M). \qquad \text{(A5)}$$

Note that the correlation matrices in the definition for the operator D_α are scaled in this paper so that the quantities are dimensionless and matrices are balanced with respect to each other. In a particular case of model parametrization using

regular discrete grids, the vector $\delta\mathbf{U}$ and the vector $\delta\mathbf{m}$ in eq. (A4) would be identical.

A2 The Hessian matrix H

Denoting the transpose F^{T} of F associated with the inner product $\langle\ ,\ \rangle$ defined by

$$(\mathsf{C}_{\mathrm{D}}^{-1}\delta\mathbf{d},\ \mathsf{F}\,\delta\mathbf{m})_{\mathrm{D}}=\langle\mathsf{F}^{\mathrm{T}}\mathsf{C}_{\mathrm{D}}^{-1}\delta\mathbf{d},\ \delta\mathbf{m}\rangle,$$
$$\forall(\delta\mathbf{m},\ \delta\mathbf{d})\in\mathscr{M}\times\mathscr{D}, \tag{A6}$$

and comparing with eq. (30),

$$(\mathsf{C}_{\mathrm{D}}^{-1}\delta\mathbf{d},\ \mathsf{F}\,\delta\mathbf{m})_{\mathrm{D}}=(\mathsf{F}^{+}\mathsf{C}_{\mathrm{D}}^{-1}\delta\mathbf{d},\ \delta\mathbf{m})_{\mathrm{M}},$$
$$\forall(\delta\mathbf{m},\ \delta\mathbf{d})\in\mathscr{M}\times\mathscr{D}, \tag{A7}$$

we have

$$\langle\mathsf{F}^{\mathrm{T}}\mathsf{C}_{\mathrm{D}}^{-1}\delta\mathbf{d},\ \delta\mathbf{m}\rangle=(\mathsf{F}^{+}\mathsf{C}_{\mathrm{D}}^{-1}\delta\mathbf{d},\ \delta\mathbf{m})_{\mathrm{M}},$$
$$\forall(\delta\mathbf{m},\ \delta\mathbf{d})\in\mathscr{M}\times\mathscr{D}. \tag{A8}$$

We now define the scalar product $(\ ,\ )_{\mathrm{M}}$ in the model space $\mathscr{M}$ as

$$(\delta\mathbf{m}^{(1)},\ \delta\mathbf{m}^{(2)})_{\alpha}=\langle\mathsf{D}_{\alpha}\delta\mathbf{m}^{(1)},\ \delta\mathbf{m}^{(2)}\rangle,$$
$$\forall(\delta\mathbf{m}^{(1)},\ \delta\mathbf{m}^{(2)})\in\mathscr{M}^{2}, \tag{A9}$$

where $\langle\ ,\ \rangle$ is the duality product between $\mathscr{M}$ and $\mathscr{M}'$ (Delprat-Jannaud & Lailly 1992). We then have

$$\langle\mathsf{F}^{\mathrm{T}}\mathsf{C}_{\mathrm{D}}^{-1}\delta\mathbf{d},\ \delta\mathbf{m}\rangle=\langle\mathsf{D}_{\alpha}\mathsf{F}^{+}\mathsf{C}_{\mathrm{D}}^{-1}\delta\mathbf{d},\ \delta\mathbf{m}\rangle,$$
$$\forall(\delta\mathbf{m},\ \delta\mathbf{d})\in\mathscr{M}\times\mathscr{D}, \tag{A10}$$

and the following equality:

$$\mathsf{F}^{+}=\mathsf{D}_{\alpha}^{-1}\mathsf{F}^{\mathrm{T}}. \tag{A11}$$

The Hessian matrix (eq. 32) can thus be expressed as

$$\mathsf{H}=\mathsf{D}_{\alpha}^{-1}\mathsf{F}^{\mathrm{T}}\mathsf{C}_{\mathrm{D}}^{-1}\mathsf{F}. \tag{A12}$$