Nonlinear Processes in Geophysics (1995) 2: 228 - 240

Nonlinear Processes in Geophysics

3D nonlinear inversion by entropy of image contrast optimization

G. Ryzhikov, M. Biryulina and A. Hanyga

Institute of Solid Earth Physics, University of Bergen, Allégaten 41, 5007 Bergen, Norway

Received 10 September 1994 - Accepted 11 January 1995 - Communicated by P. C. Sabatier

Abstract. An approach to solve 3D inverse problem associated with inverting seismic reflection data is presented. It exploits the a priori assumption that the reflection data, reduced properly, can be interpreted as a perturbation of a dynamical response of a certain 'reference background'. It is supposed that the corresponding perturbation of medium parameters can be treated in terms of the 'Ray + Born'-, or 'Rytov + Born'- set of medium functions. This for the reflection data means that 'kinematical' part does not generate reflections, while proper reflections are caused by single-scattering perturbations. Moreover, it is guessed that the latters are cooperated in a vicinity of a certain unknown 2D smooth surface ('interfaces'). When this a priori information is adequate, the approach allows to recover both the low-frequency features of the medium (the background) and its discontinuities. The approach involves a new optimization criterion, called the Entropy of Image Contrast (EnIC), and a new global optimization algorithm, called Regularized Global Approximation algorithm (RGA-algorithm). It allows to choose such a background that the linerized inversion provides the most focused image of interfaces. In other words, it yields the maximum-contrast, or minimum-entropy, interface image. The method takes into account the large amounts of data that have to be processed in 3D inversion and the sparseness of input data. It is also robust with respect to the noise in the data.

1 Introduction.

Probably it is a propagation of sound waves in a medium that is the best-known nonlinear process in geophysics. Parameters of the process are distributed in a three-dimensional physical space and have an excellent physical interpretation, e.g. density, slowness, impedance, Lamé parameters. To extract these parameters from observed data means to solve 3D nonlinear inverse problem that creates additional difficulties as compared with any nonlinear incorrect problem. Indeed, it is impossible to apply a conventional iterative algorithm to solve the nonlinear problem because of extremely high computational costs of 3D forward modeling that is necessary on each step of the iteration procedure. In view of general computational difficulties arising even when solving a 3D linear inverse problem we try to save the general structure of the suggested algorithm as much as possible omitting details that depend on an operator.

The problem of numerical inversion of 3D seismic reflection data remains to be a challenge because of the huge amount of unknowns and data that are to be handled. At the same time it is impossible to ignore the three-dimensional geometry of real geological structures, i.e. one can't reduce sufficiently the set of unknown medium parameter functions.

It is also impossible to decouple migration or linear inversion with respect to an interface from the determination of the elastic parameters of the overlying layers. It has been shown by numerical experiments that linear inversion is extremely sensitive to noise in the travel times (e.g. Versteeg, 1993) but sufficiently robust with respect to amplitude noise. It is therefore important to couple linear inversion with velocity analysis (e.g. Luo and Schuster, 1992).

We start from realistic assumptions about the data taking into account their incompleteness, sparseness and spatial randomness. The data represent an unstacked digitized finite-band noisy seismogram. It is supposed a source-receiver array has an arbitrary configuration.

In order to handle incomplete data it is necessary to introduce an adequate a priori information (Franklin, 1970), (Tikhonov and Arsenin, 1977). We have based inversion on the assumption that the deviation of model parameters $m(\mathbf{x})$ $\mathbf{x} \in \Omega \subset R^3$ from a reference model s_0 can be decomposed into two components $m - s_0 = \delta s + \mu$. The low frequency component $\delta s(\mathbf{x})$ corresponds to the parameter variation inside the layers. The high

Correspondence to: G. Ryzhikov

frequency component $\mu(\mathbf{x})$ describes the jumps of the parameters at interfaces. The high frequency component is assumed to be concentrated in a neighborhood of some two-dimensional manifolds, while the low frequency component δs is assumed to be such that relative perturbation of a reference medium $s_0(\mathbf{x})$ is rather smooth, i.e. to belong to a small ball in the Sobolev space of second order.

Taking into account sparseness of the data we reduce the size of the problem without any loss of information. Spatial randomness of the acquisition system does not cause additional difficulties since we do not interpolate in the data space.

Updating the low frequency component δs is based on solving an optimization problem. The objective function is a measure of the sharpness of an image μ of the high-frequency component obtained by linear inversion of the data residuals with respect to the background $s = s_0 + \delta s$. It is defined in terms of a new function of the gradient of μ, which we call entropy of image contrast (EnIC), and an appropriate regularizing term depending on s. It is conjectured that a "good" background allows successful migration of high frequency model features. In other words, we suppose that a background $s = s_0 + \delta s$ is responsible for nonlinear part of inversion problem, so we suggest the measure to estimate focusing features of unknown $s(\mathbf{x})$ by analysis of a contrast of corresponding image $\mu(\mathbf{x})$.

In order to reduce the number of optimization parameters we represent the low frequency component in terms of a truncated expansion in terms of the eigenfunctions of the "anisotropic Laplacian" $\nabla \cdot \mathbf{S} \nabla$ (Sect. 4). For example, the anisotropy can assign different a priori weights to horizontal and vertical variations of model parameters.

The optimization problem is solved by a global optimization algorithm. In order to reduce the number of evaluations of the objective function a new algorithm (RGA-algorithm) has also been introduced by Ryzhikov and Biryulina (1994) (Appendix C). It has been applied to optimization in a 12-parameter space.

Each evaluation of EnIC requires interface imaging, i.e. determination of the high-frequency component μ while the background model M is held fixed. Interface imaging is implemented by a wave field inversion based on the Born approximation. In order to reduce the costs of each evaluation of EnIC the image is constructed in a small volume Ω, typically a thin vertical layer in the physical space. The inversion is based on a relatively crude diagonalization of the Hessian to facilitate computation of its inverse. More accurate inversion may be needed after the final update of the background model.

The general structure of nonlinear inversion is shown in Fig. 1.

The method described below is illustrated by applying it to the scalar wave equation. The model to be reconstructed is represented by the slowness field $s(\mathbf{x})$.

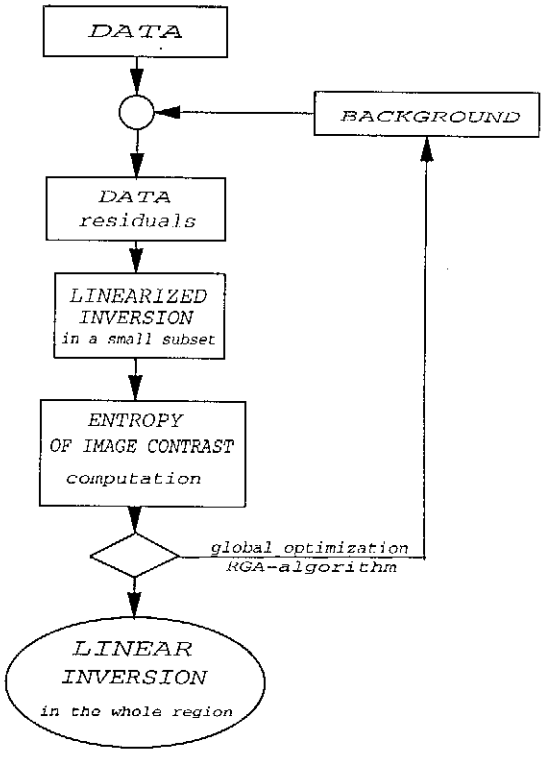

Fig. 1. Block diagram of nonlinear inversion.

2 Linearized inversion

We assume that the digitized reflection seismic data d can be expressed in terms of a sounding field φ and an additive noise e:

$$d = D\varphi + e \tag{1}$$

where D is the compact operator representing convolving features of a registration channel. For example, data $d(x_r; t)$, registered with the single receiver at a point $\mathbf{x}_r$, can be modeled by $d(\mathbf{x}_r; t) = \int dx \delta(\mathbf{x} - \mathbf{x}_r) \mathcal{D}(t) * \varphi(\mathbf{x}; t)$, where the operator D is represented with the integral kernel $\mathcal{D}\delta(\mathbf{x} - \mathbf{x}_r)$, and $\mathcal{D}(t)$ is an impulse response of the '$\mathbf{x}_r$'-channel. It means an inverse operator D^{-1} does not exist and we are to take into account that we are dealing with narrow time-frequency band data given in a set of sparse spatial points. Here we try to avoid any intermediate transformation of data, e.g. deconvolution, interpolation in space-time, extrapolation. Solving of these incorrect problems supposes to use a priori information in data space which induces indirectly a priori information in medium parameter space. It is rather difficult to control this kind of a priori information because of nonlinear mapping of medium parameters into data.

The sounding field φ, generated by source f, satisfies a partial differential equation

$$L_m \varphi = f \tag{2}$$

where $m = m(\mathbf{x})$, $\mathbf{x} \in \Omega \subset R^3$, represents a set of parameters of the medium (e.g. $L_m = m\,\partial_t^2 - \nabla^2$ for the scalar wave equation, and $m(\mathbf{x})$ is an inverse squared velocity).

The inverse problem aims at a reconstruction of the function $m(\mathbf{x})$ from the data registered at a a finite set of receivers and from a finite set of sources. The problem is underdetermined and requires a proper regularization.

The inverse problem can be expressed in terms of the following optimization problem (e.g. Tarantola and Valette, 1982):

$$m^* = \arg \inf_{m \in M} \; J^d(m) \tag{3}$$

where set M has to be chosen properly, and

$$J^d(m) \;=\; \frac{1}{2}\langle\, d - d_{cal}(m) \mid d - d_{cal}(m)\rangle \tag{4}$$

$d_{cal}(m)$ are the data calculated for a model m and d are the observed data. Assuming that C_e is the covariance matrix of the noise e a natural inner product in the *data space* $\mathcal{D}$ is given by

$$\langle\, d' \mid d''\,\rangle \;\stackrel{\text{def}}{=}\; \sum_{n'}^{N}\sum_{n''}^{N} (C_e^{-1})_{n'n''} d'_{n'} d''_{n''}, \tag{5}$$

while d_{cal} is nonlinear operator on M. The data space is assumed to be a finite-dimensional space. The vector components in the data space are indexed by a multi-index $N = (r, s, t)$, where the indices r, s, t specify the receiver, the source and the discretized time delay. It should be mentioned that real data are discrete and an interpolation can lead to a loss of information. The norm in the data space

$$\| d \|_{\mathcal{D}} \;\stackrel{\text{def}}{=}\; \langle\, d \mid d\,\rangle^{1/2} \tag{6}$$

induced a discriminating norm in the *wave-field space* Φ

$$\| \varphi - \tilde{\varphi} \|_{\Phi} \;\stackrel{\text{def}}{=}\; \| (D^{\dagger} C_e^{-1} D)^{1/2} (\varphi - \tilde{\varphi}) \| \tag{7}$$

where the compact operator : $(D^{\dagger} C_e^{-1} D)^{1/2}$ is the weight in the $\mathcal{L}^2$-norm, and $D^{\dagger}$ is adjoint of D: $\langle\, d \mid D\varphi\,\rangle = (\, D^{\dagger} d, \varphi)$, where the last expression denotes the integration over space-time region. Such a discriminating norm (7) allows us not to take care of accurate approximation of φ very much, at least it concerns high-frequency components of the wave-field approximation (Ryzhikov and Troyan , 1992b)

Linear inversion is based on an approximate solution of the optimization problem (3), that can be applied if the unknown $m(x)$ is close to a known model $\tilde{m}(\mathbf{x})$

$$\| d_{res} - \int d^{(B)}(\mathbf{x}) \mu(\mathbf{x})\, \mathrm{d}x \|_{\mathcal{D}}^2 + \varepsilon \|\mu\|^2 \to \min \tag{8}$$

where residuals $d_{res} = d - d_{cal}(\tilde{m})$, $\mu(\mathbf{x}) = m(\mathbf{x}) - \tilde{m}(\mathbf{x})$, $d^{(B)}(\mathbf{x})$ is the integral kernel corresponding to the linear response of the data to a model perturbation in the Born approximation, i.e. to the Fréchet-derivative of $d_{cal}(m)$ with respect to m (Appendix A), $\|\mu\|^2 = \int_{\Omega} \mu^2(\mathbf{x})\, \mathrm{d}\mathbf{x}$, and $\varepsilon > 0$ is a regularization parameter.

The solution of eq.(8) can be written in terms of the integral kernel of the Hessian of (8)

$$H(\mathbf{x}, \mathbf{x}') = \sum_{n'}^{N}\sum_{n''}^{N} (C_e^{-1})_{n'n''} d_{n'}^{(B)}(\mathbf{x}) d_{n''}^{(B)}(\mathbf{x}') + \varepsilon \delta(\mathbf{x} - \mathbf{x}')$$

A further simplification (Ryzhikov and Troyan , 1992b)

$$H(\mathbf{x}, \mathbf{x}') \cong [h(\mathbf{x}) + \varepsilon]\, \delta(\mathbf{x} - \mathbf{x}') \tag{9}$$

where

$$h(\mathbf{x}) = \sum_{n'}^{N}\sum_{n''}^{N} (C_e^{-1})_{n'n''} d_{n'}^{(B)}(\mathbf{x})\, d_{n''}^{(B)}(\mathbf{x}')\, \delta(\mathbf{x} - \mathbf{x}')$$

is introduced to make the problem tractable in 3D. A similar approximation has beeb made by Beydoun and Mendes (1989) in elastic inversion problem.

The solution of eq. (8) is given by *RT*-algorithm (Appendix A)

$$\mu(\mathbf{x}) = \langle\, \tilde{d}^{(B)}(\mathbf{x}) \mid d\,\rangle \tag{10}$$

where *weighted Born response* is defined by

$$\tilde{d}^{(B)}(\mathbf{x}) \;\stackrel{\text{def}}{=}\; \frac{d^{(B)}(\mathbf{x})}{[h(\mathbf{x}) + \varepsilon]} \tag{11}$$

To specify the expression for localized inversion (10) one needs to construct Fréchet-derivative of the operator L_m (e.g. Tarantola (1987), Beylkin and Burridge (1990), Chapman and Coates (1994)) and to solve forward problems: to reconstruct wave fields in a chosen inhomogeneous background. In other words, to recover a component of $\mu(\mathbf{x})$ in a given point $\mathbf{x}$ it is necessary to calculate all *single-scattered* in the point $\mathbf{x}$ wave fields generated by all sources and convolved with impulse responses of corresponding receivers.

3 Entropy of image contrast.

The definition of the set of admissible solutions m in (3), must account for the specific features of the 3D inverse problem. The inversion problem of reflection seismics involves sources and receivers located on the surface of registration and with a relatively small aperture. Dealing with reflection data means we can make an a priori assumption that the medium m has two well-separated components. The first component is the background $s(\mathbf{x})$, which does not cause a significant back-scattering effect and therefore does not contribute to reflection. The second one corresponds to interfaces $\mu(\mathbf{x})$ and is responsible for reflections. The component $\mu(\mathbf{x})$ can be assumed to be located in a vicinity of 2D-surfaces.

The entropy of image contrast (EnIC) is a functional which quantifies contrast of a function $\mu(\mathbf{x})$ defined over a region Ω. The EnIC is defined in terms of the entropy

$$E = -\int_{\Omega} p(\mathbf{x})\, \ln p(\mathbf{x})\, \mathrm{d}x \tag{12}$$

of the pseudo-probability density function (*pseudo-pdf*)

$$p(\mathbf{x}) = (\nabla\mu(\mathbf{x}))^2 / \int_{\Omega} (\nabla\mu(\mathbf{x}))^2 \mathrm{d}x \tag{13}$$

Large contrast of $\mu(\mathbf{x})$ — or high concentration of $|\nabla\mu(\mathbf{x})|^2$ — corresponds to low values of EnIC. EnIC is non-negative and assumes minima on suitably constrained sets of functions. A few entropy features are described in Appendix B. In 3D inversion the function $\mu(\mathbf{x})$ appears as the image of the the high frequency component of the model. A background model is assumed to be acceptable if the solution μ of the linearized inversion problem for the residuals is a high contrast function.

The entropy of image contrast $\mathcal{E}(\mu)$ is obtained by substituting eq. (13) in eq. (12). It is sufficient to test the contrast of μ in a small subset Ω of the model space $\mathcal{V}$ provided the rays joining Ω to the sources and receivers cover a sufficient part of $\mathcal{V}$. This makes the 3D problem tractable if the linear inversion/migration algorithm is sufficiently simple (e.g. 10) and based on ray tracing.

The high-frequency component reconstruction is based upon the set of functions $\{d^B(\mathbf{x})\}$ (8). Each of them has a support in a vicinity of an isochron if the sounding signal is sufficiently short (e.g. Miller, Oristaglio and Beylkin (1987), Ryzhikov and Troyan (1992a)) Because of the reflection seismic geometry the normals to the isochrons are contained in a narrow vertical cone, which leads the set $\mu(\mathbf{x};\, s\,)$ (10) belongs to a near-one-parametrical family of images, and consequently the set of pseudo-pdf's forms the near-one-parametrical family also.

On the other hand, in contrary to Symes and Carazzone (1991), Jin and Madariaga (1992) we apply linear inversion to the full set of data d (1). This yields a better resolution due to a larger set of isochrons at each point $\mathbf{x}$. Besides it allows us not to claim the images associated with different sources to be similar. It should be expected that these images do differ because of different illumination of the medium.

4 Regularized inverse problem

The inverse problem for the data

$$d \;=\; d_{cal}(m) \;+\; e \tag{14}$$

with $m = m(\mathbf{x})$, $\mathbf{x} \subset \Omega \in R^3$ while no spatial symmetry is assumed a priori, can be formulated in terms of the following optimization problem:

$$m^* = \arg\inf \{\, \mathcal{E}(m) \quad | \quad m \in M = \mathcal{S} \oplus \mathcal{M}\} \tag{15}$$

where $m(\mathbf{x}) = s(\mathbf{x}) + \mu(\mathbf{x})$, $s \in \mathcal{S}$, $\mu \in \mathcal{M}$.

The set of admissible background models $\mathcal{S}$ is defined by

$$\mathcal{S} \;=\; \{\, s \;\; | \;\; (\sigma, [\, I - \nabla \cdot \mathbf{S}\nabla]\sigma) \le \eta\} \tag{16}$$

with $\sigma = (s - s_0)/s_0$, $s_0(\mathbf{x}) > 0$.

The set (16) constrains the relative perturbation of the background $\sigma = (s - s_0)/s_0$ to belong to a small ball in an anisotropic Sobolev space (Ryzhikov and Troyan, 1991). This allows to be sure that the wave field φ in such a background can be treated in terms of the ray/Rytov/WKBJ - approach. while the set of backgrounds is sufficiently representative.

The set of admissible interface images $\mathcal{M}$ is defined by

$$\mathcal{M} \;=\; \{\, \mu \;=\; R_s\, [\, d \;-\; d_{cal}(s)]\, \} \tag{17}$$

where $s \in \mathcal{S}$, $R_s\, w := \langle\, \tilde{d}^{(B)}(\mathbf{x}; s) \;|\; w\, \rangle$. Since $\mathcal{S}$ (16) is compact in $\mathcal{L}^2$ and R_s (10) is continuous, the mapping $r : \mathcal{S} \mapsto \mathcal{M}$ is continuous, therefore $\mathcal{M}$ as an image of r on $\mathcal{S}$ is compact also.

By using of Lagrange multipliers $\{\eta_k\}$ the solution (15-17) can be expressed as an approximate solution of the global optimization problem (3):

$$m^* \;=\; \arg \inf_{m \in M}\; \mathcal{G}(m)$$

$$\mathcal{G} = \|\, d \;-\; d_{cal}(m)\|^2_{\mathcal{D}} \;+\; \eta_2 S(s) \;+\; \eta_3 \mathcal{B}(\mu) \tag{18}$$

where $m \;=\; s + \mu$, $S(s)$ is a squared norm in the anisotropic Sobolev space W with a weight $s_0^{-2}(\mathbf{x})$:

$$S \;\; (s) \;=\; \|\, s \;-\; s_0\|^2_W \;=$$

$$=\; \int s_0^{-2}(\mathbf{x})\mathrm{d}x\; \{(s - s_0)[I - \nabla \cdot \mathbf{S}\nabla](s - s_0)\} \tag{19}$$

and $\mathcal{B}(\mu)$ is an appropriate regularizing (quadratic) term.

Supposing the a priori representation of m to be valid, we have reduced the problem (18) to the following:

$$s^* = \arg\inf\{\, \mathcal{E}(s) + \eta(\sigma, [I - \nabla \cdot \mathbf{S}\nabla]\sigma) \;\; | \;\; s \in \mathcal{S}\,\} \tag{20}$$

where $\mathcal{E}(s) = \mathcal{E}(\mu(s))$. The functional in equation (20) will be referred to as REnIC (*regularized EnIC*). The solution of (20) can be interpreted as a generalized ray tomography, if the linearized inversion/migration $\mu(s)$ is treated in terms of the ray theory (see Sect. 5.3).

5 Numerical experiments.

5.1 The forward problem.

In order to test the applicability of EnIC to 3D inversion of reflection seismic data some preliminary numerical tests have been carried out. Because of computational costs of the numerical experiment we have reduced each step of it as much as possible.

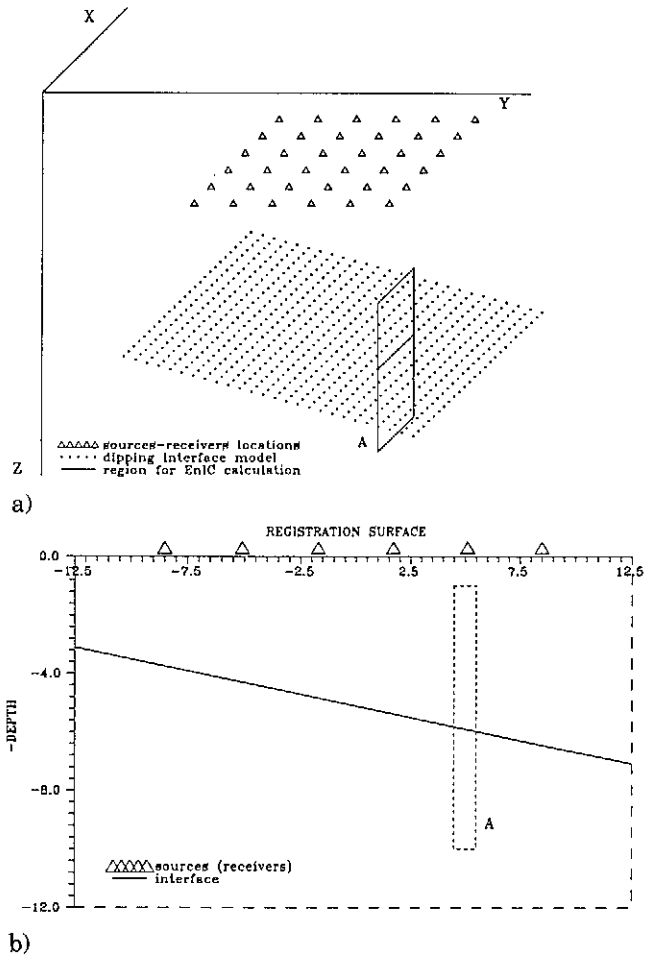

Fig. 2. Geometry of 3D inversion experiment. a) general view; b) slice x=0 . Characteristic wavelength of the sounding signal is the unit for lateral variables and depth. Cross-section A indicates a region for EnIC calculations

A block scheme of the experiment was represented by Fig. 1.

The synthetic data were calculated or by applying the ray-theory approximation for the scalar-wave equation to imitate reflections caused by plane interfaces, or by applying the Born-approximation in the case of arbitrary smooth interfaces. Example of a geometry for simulating of 3D reflection seismic data is given in Fig. 2 The Born-approximation of the incoming wave field $\varphi_{in}(\mathbf{x},t)$ (Appendix A for a homogeneous medium $m(\mathbf{x}) = s = const$ as the background was constructed by convolution of a low-frequency signal $f(t)$ with the Green function for the unbounded medium:

$$\varphi_{in}(\mathbf{x},t) = A(\mathbf{x}|\mathbf{x}_s)\,\ddot{f}(t-\tau(\mathbf{x}|\mathbf{x}_s)) = \\ = \frac{1}{4\pi}\frac{1}{|\mathbf{x}-\mathbf{x}_s|}\ddot{f}(t-s_0|\mathbf{x}-\mathbf{x}_s|) \qquad (21)$$

Note here that the second derivative in the eq. 21 is an example of Fréchet derivative of operator L_m mentioned in eq. 2 with respect to m: $L_m = m\partial_t^2 - \nabla^2$, so $\partial_m L_m = \partial_t^2$. Examples of Fréchet derivatives of a response of data (Appendix A) for elastic inversion are given in the paper by Ryzhikov and Troyan (1992a).

For computational efficiency, travel times in a non-homogeneous background $s(\mathbf{x})$ were calculated using linear (with respect to slowness) approximation. The wave field was expressed in terms of an approximate asymptotic Green function of an amplitude $1/r$ and a time-shift calculated by integrating slowness $s(\mathbf{x})$ over straight-line ray. More specifically:

$$\varphi_{in}(\mathbf{x},t) \doteq \frac{1}{4\pi}\frac{1}{|\mathbf{x}-\mathbf{x}_s|}\ddot{f}(t-\int s(\mathbf{x}')\,\mathrm{d}\sigma) \qquad (22)$$

where $\mathbf{x}' = \mathbf{x}'(\alpha) = \mathbf{x}_s + \alpha(\mathbf{x}-\mathbf{x}_s)$; $\alpha \in [0,1]$, and $\mathrm{d}\sigma = |\mathbf{x}-\mathbf{x}_s|\,\mathrm{d}\alpha$ is an arclength.

Thus, for example the response from a plane $P_q = \{\mathbf{x}^* \mid (\mathbf{n}\ ,\ \mathbf{x}^*) = q\}$ in the experiment with ξ-th pair of source $\mathbf{x}_s$ and receiver $\mathbf{x}_r$ and chosen background $s(\mathbf{x})$ can be written in the form

$$d^\xi(t) = \int_\Omega \mathrm{d}x^*\,\delta((\mathbf{n}\ ,\ \mathbf{x}^*)-q)\,A^\xi(\mathbf{x}^*)\,\ddot{f}(t-\tau^\xi(\mathbf{x}^*)) \\ \doteq \int_{P_q}\mathrm{d}x^*\tilde{A}^\xi(\mathbf{x}^*)\ddot{f}(t-\int s(\mathbf{x}')\mathrm{d}\sigma_1 \\ -\int s(\mathbf{x}'')\mathrm{d}\sigma_2\ ) \qquad (23)$$

where $A^\xi(\mathbf{x}^*) \doteq \tilde{A}^\xi(\mathbf{x}^*) = (4\,\pi)^{-2}\,|\mathbf{x}_r-\mathbf{x}^*|^{-1}|\mathbf{x}^*-\mathbf{x}_s|^{-1}$, $\mathbf{x}' = \mathbf{x}'(\alpha) = \mathbf{x}_s + \alpha(\mathbf{x}_s-\mathbf{x}^*)$, $\mathrm{d}\sigma_1 = |\mathbf{x}_s-\mathbf{x}^*|\,d\alpha$; $\mathbf{x}'' = \mathbf{x}^* + \alpha(\mathbf{x}_r-\mathbf{x}^*)$, $\mathrm{d}\sigma_2 = |\mathbf{x}_r-\mathbf{x}^*|\,\mathrm{d}\alpha$, $\alpha \in [0,1]$, and $(\mathbf{n}\ ,\ \mathbf{x}^*)$ denotes the scalar product in $\mathcal{R}^3$.

In order to generate synthetic reflected data we modeled the interface as a set of point diffractors on a smooth surface. Numerically the interface was modeled by "inserting" point scatterers $\{\mathbf{x}^*\}$ in a regular grid located on the plane surface, which is equivalent to a Riemann sum for (23).

The Born-approximation was also applied to the calculation of the virtual response $\tilde{d}^{(B)}(\mathbf{x}^*)$ (10).

5.2 EnIC and REnIC behavior

The low-frequency component of the slowness is expressed in terms of the reference slowness $s_0(\mathbf{x})$ and a truncated expansion in terms the eigenfunctions $\xi_\gamma(\mathbf{x})$ of the regularizing operator $\eta[\,I - \nabla\cdot\mathbf{S}\nabla]$ (Sect.4):

$$s(\mathbf{x}) = s_0(\mathbf{x})[1 + \sum_{\gamma=1}^{\gamma_{\max}}\alpha_\gamma\lambda_\gamma^{-1/2}\xi_\gamma(\mathbf{x})] \qquad (24)$$

where $\{\lambda_\gamma\}$ are eigenvalues of the operator $\eta[I-\nabla\cdot\mathbf{S}\nabla]$ and $\mathbf{S}$ is a positive definite symmetric matrix that allows for an anisotropic a priori scaling. The optimization parameters are given by the coefficients α_γ, $\gamma = 1,\ldots,\gamma_{\max}$. Truncation a priori excludes high and middle frequencies in s. Eq. (24) can also be interpreted as an expansion in terms of the principal components of

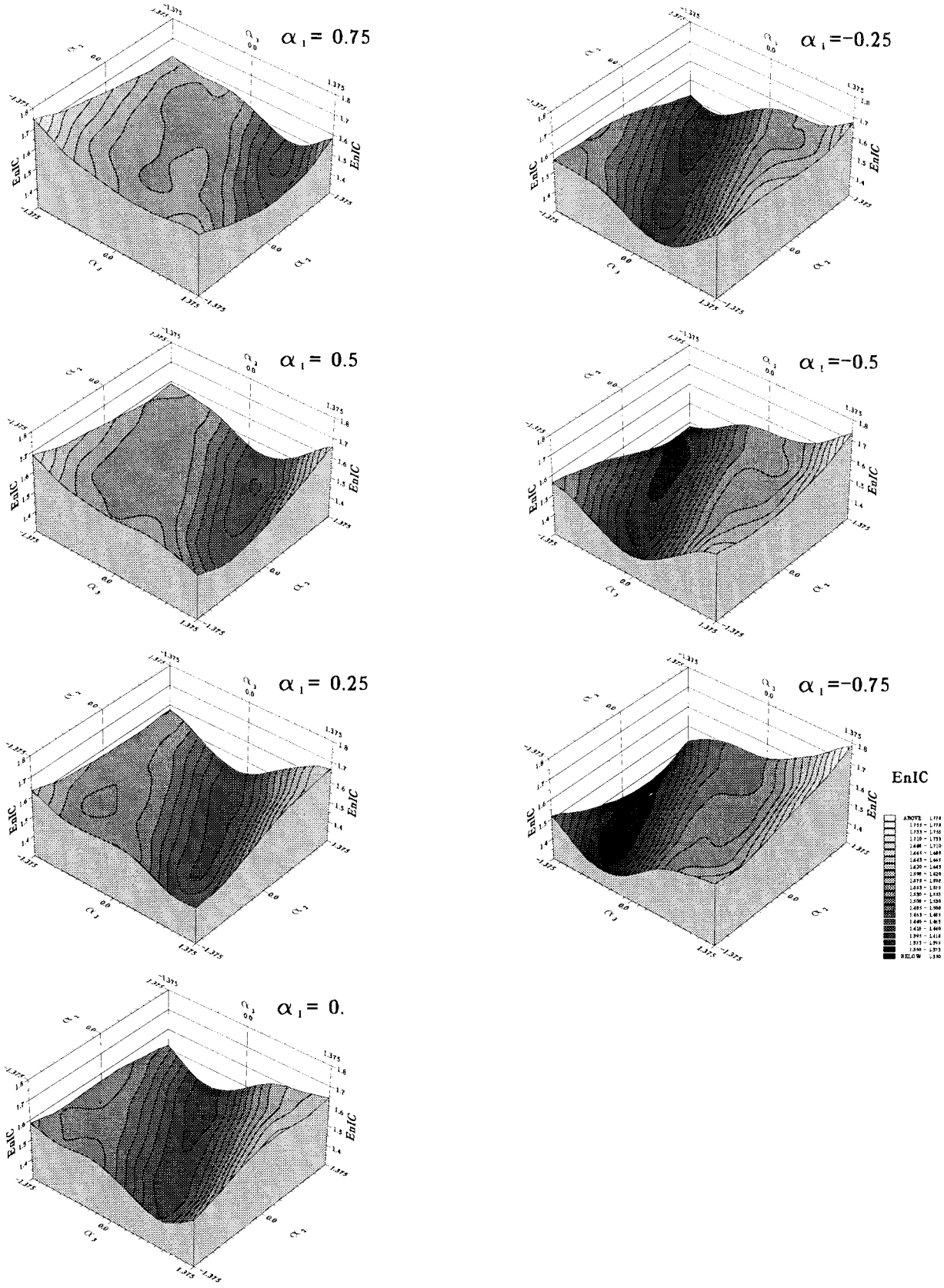

Fig. 3. EnIC for a 3-parameter slowness model

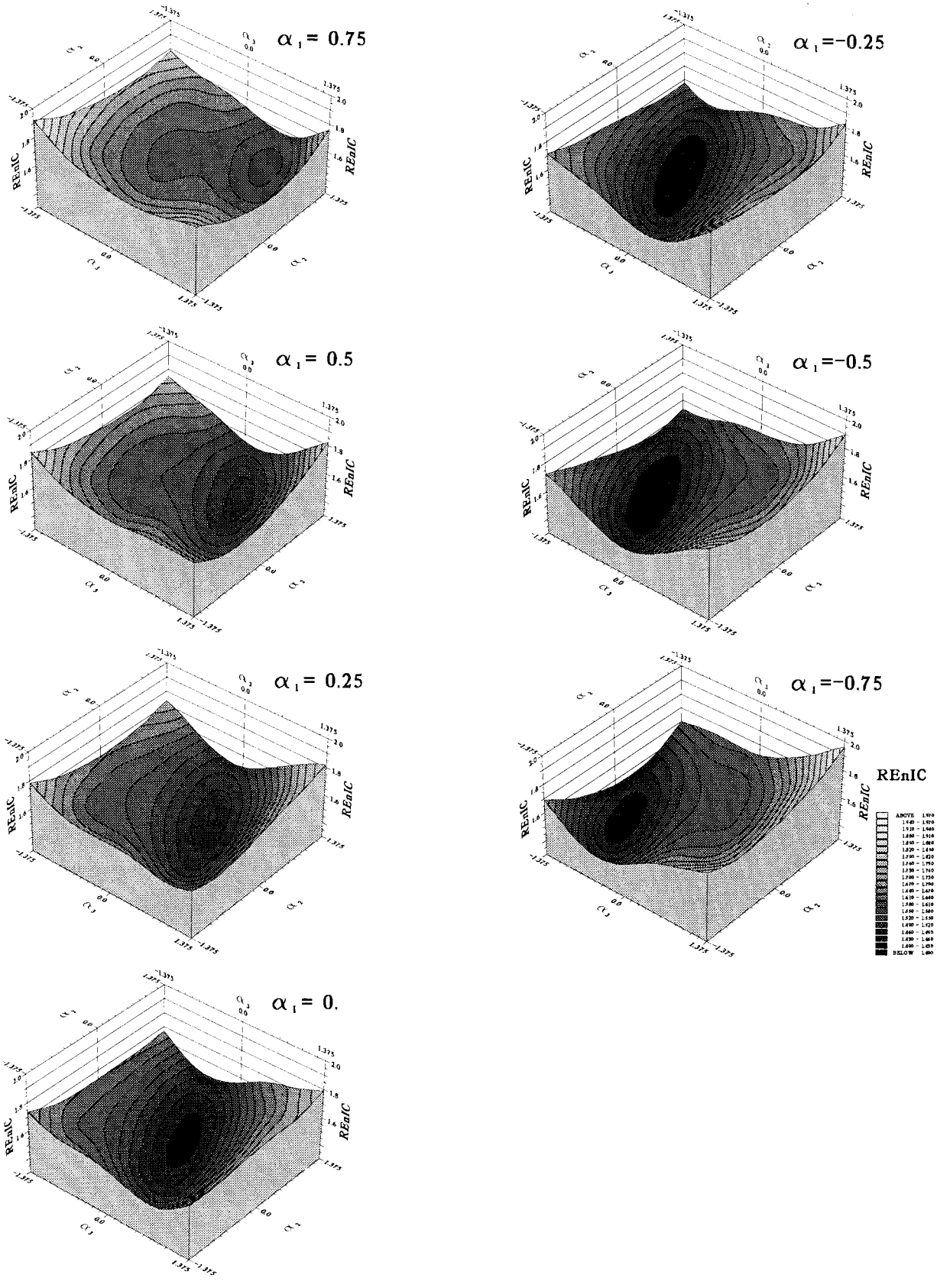

Fig. 4. Regularized EnIC (REnIC) for a 3-parameter slowness model.

a *generalized Karhunen-Loève basis* generated by an a priori Gaussian distribution of the relative perturbation σ (16):

$$p(\sigma) \propto \exp\{-\frac{1}{2}\eta(\sigma, [I - \nabla \cdot \mathbf{S}\nabla]\sigma)\} \tag{25}$$

where $(\cdot,\cdot)$ denotes the scalar product in $\mathcal{L}^2$. Note, that in terms of α_γ eq. 25 can be rewritten as

$$p(\sigma(\alpha)) \propto \exp\{-\frac{1}{2}\sum_{\gamma=1}^{\infty}\alpha_\gamma^2\} \tag{26}$$

therefore a sphere $\{\sum_\gamma \alpha_\gamma^2 = const\}$ represents a isosurface of a priori probability in the slowness space.

In Figs. 3 - 4 we present plots of EnIC and REnIC for some inversion problems mimicking inversion of 3D reflection data. The synthetic model involves principal components $\xi_\gamma(\mathbf{x})$ corresponding to lower values of $\{\lambda_\gamma\}$, where $\gamma = \{\gamma_x, \gamma_y, \gamma_z\}$. In order to avoid extra calculations of eigenvalues and eigenfunctions of anisotropic Laplacian in a given domain Ω we took eigenfunctions of Laplacian in a unit box scaled along lateral coordinates and depth by factors a, b, c under homogeneous Neumann boundary conditions (the unit box is chosen rather arbitrary and it has not to create artificial interfaces because of nonzero gradient of slowness on boundaries). So, eigenvalues λ_γ and eigenfunctions $\xi_\gamma(\mathbf{x})$ (24) playing the role of generalized Karhunen-Loève basis in our simplified version were the following:

$$\lambda_\gamma = \eta\pi^2\,(\frac{\gamma_1^2}{a^2} + \frac{\gamma_2^2}{b^2} + \frac{\gamma_3^2}{c^2}),\ \gamma_1,\gamma_2,\gamma_3 = 0,1,2,\ldots$$
$$\xi_\gamma(\mathbf{x}) = \cos\frac{\gamma_1\pi x}{a}\cos\frac{\gamma_2\pi y}{b}\cos\frac{\gamma_3\pi z}{c} \tag{27}$$

The corresponding indexes of coefficients α (24) in Figs. 3-4 are following: $\alpha_1 \to \{0,0,0\}$, $\alpha_2 \to \{0,0,1\}$, $\alpha_2 \to \{0,1,0\}$. Real slowness model is generated with $\alpha_1 = -0.5, \alpha_2 = 0.0, \alpha_3 = -0.5, \alpha_4 = 0.0$. The corresponding value of REnIC is 1.41. The value of α_4 in the plots is fixed at 0. The geometry of the model and acquisition is shown in Fig. 2. A 3–5 percent noise was included in the synthetic data. The test volume consisted of a single vertical plane, as shown with the cross-section A in Fig. 2.

The minimum of EnIC in Fig. 3 has the form of a valley, which indicates that the solution of eq. (20) without regularization is non-unique. This kind of non-uniqueness results from the fact that the mapping $\mu = F(s)$ is not injective.

5.3 ReNIC optimization as a generalized ray tomography

Numerical experiments indicate that REnIC is a well-behaved smooth function with a unique global minimum.

The RGA-algorithm (Appendix C) requires a relatively small number of evaluations of REnIC. It has been applied in 4- and 12-parametric (with $\{0,0,0\}$ – $\{1,1,2\}$ principal components) inversion by REnIC optimization. Reconstruction of of REnIC in the 4-parametric case is shown in Fig. 5. The values of α_2 and α_4 in the plots are fixed at 0.

It took just 107 evaluations of REnIC in the 4-parametric case and 256 evaluations in the 12-parametric case to get a REnIC value error of ≤ 2 %. In Fig. 6 the reconstructed slowness is represented by isolines of relative differences. One of the "bad" points in the slowness space which was used by the algorithm is shown in Fig. 7. It is easy to see that the accuracy of slowness reconstruction is influenced by the choice of region for EnIC testing (cross-section A in Fig. 2). Indeed, contrast images are obtained only if the kinematics, being responsible for the reconstruction in a chosen region for EnIC testing, was true. It gave us the base to call the method *generalized ray tomography* as far as *virtual* rays as well as real rays are involved in reconstruction of velocity model.

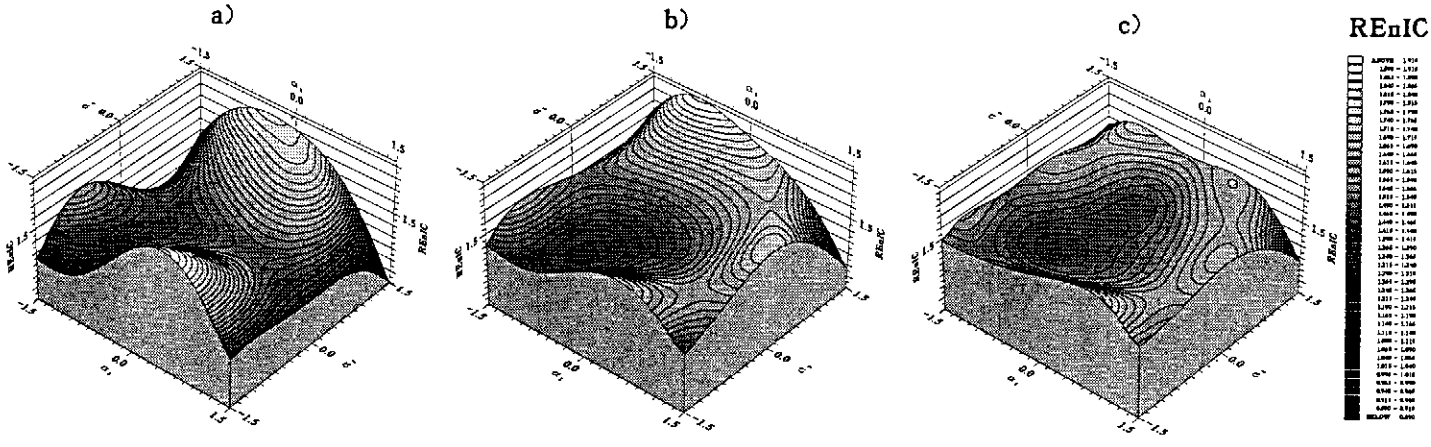

Fig. 5. Reconstruction of 3D four-parameter slowness by global optimization of REnIC. a) 0-th iteration. 51 points. b) 1-st iteration. 70 points. c) 2-nd iteration. 107 points.

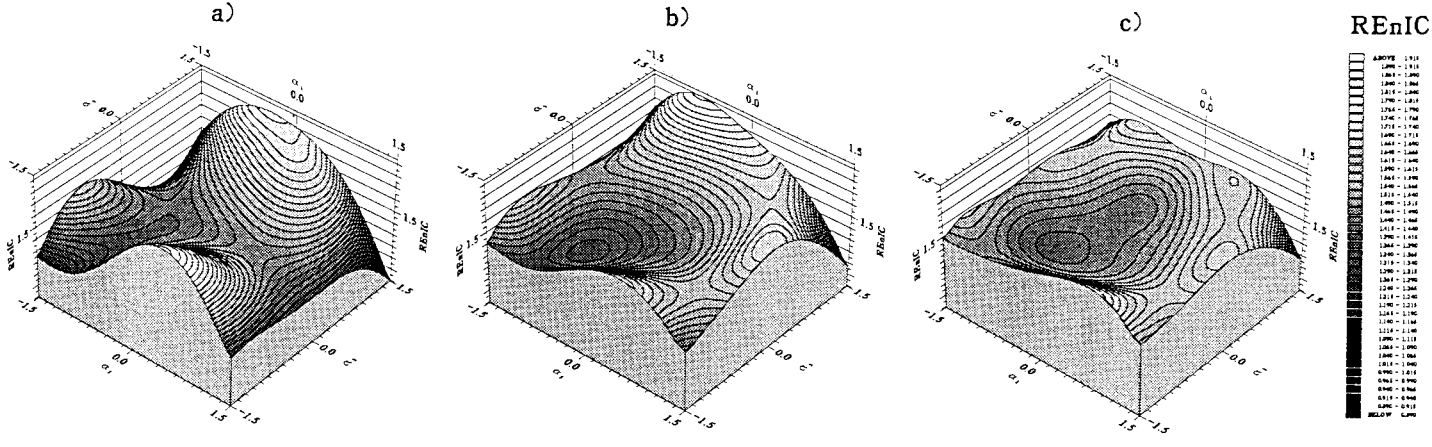

Fig. 5. Reconstruction of 3D four-parameter slowness by global optimization of REnIC. a) 0-th iteration. 51 points. b) 1-st iteration. 70 points. c) 2-nd iteration. 107 points.

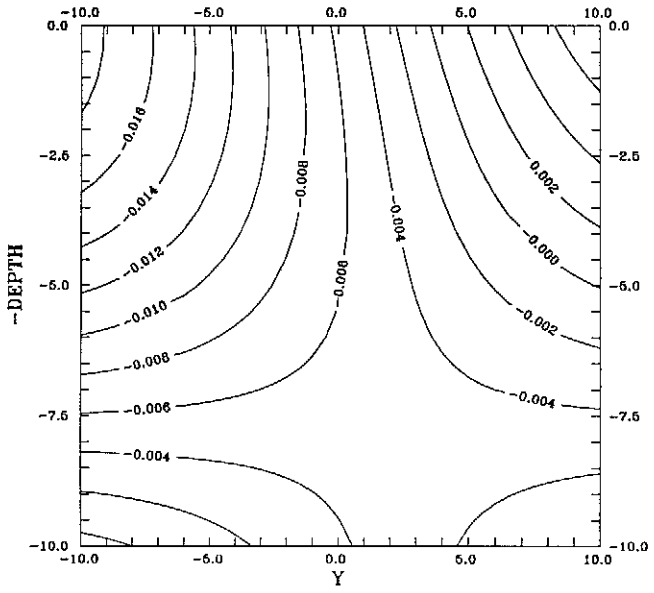

Fig. 6. Relative difference between reconstructed and real slownesses. Slice X=0. Reconstruction has been carried out for 12-parameter slowness model. Ratio of REnIC values for reconstructed and real slownesses is equal to 1.015 .

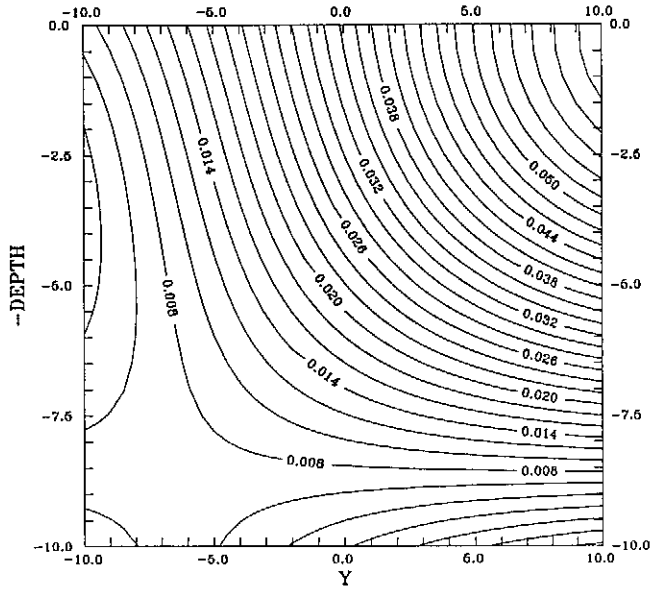

Fig. 7. The same as Fig.6, except one of the current random slowness realizations obtained by optimization procedure is presented instead of reconstructed slowness. Ratio of REnIC values in this case is equal to 1.440.

5.4 Interface image reconstruction

Typical artifacts appear while image reconstructing because of limited aperture and signal shape even in the case when the reference model is reconstructed properly. Interface images have been processed using 2D Radon-type filtering of 2D slices of $\mu(\mathbf{x})$ followed the regularized approximation. The filtering has been carried out applying the RGA-algorithm (Appendix C) to get an approximation of Radon-projection function in the vicinity of its global maximum.

In Fig.. 8 the result of global inversion for an inhomogeneous model involving a dipping interface is shown. The picture shows a vertical cross-section of the reconstructed part of the interface. The result of global inversion with 8-parameter slowness space (with $\{0,0,0\}$ – $\{1,1,1\}$ principal components) and two curved interfaces is presented in Fig.. 9. To estimate slowness function more accurately than in the previous experiment we took two small boxes to calculate EnIC (cross-sections A in Fig.. 9). The error of reconstructed slowness is less than 2 % in the whole region (less than 0.5% inside the boxes), as compared with "bad" slowness model that was detected while optimization searching when the ratio of EnIC values was 1.42, the error reached 7%-level inside the boxes and 11 % in the whole region.

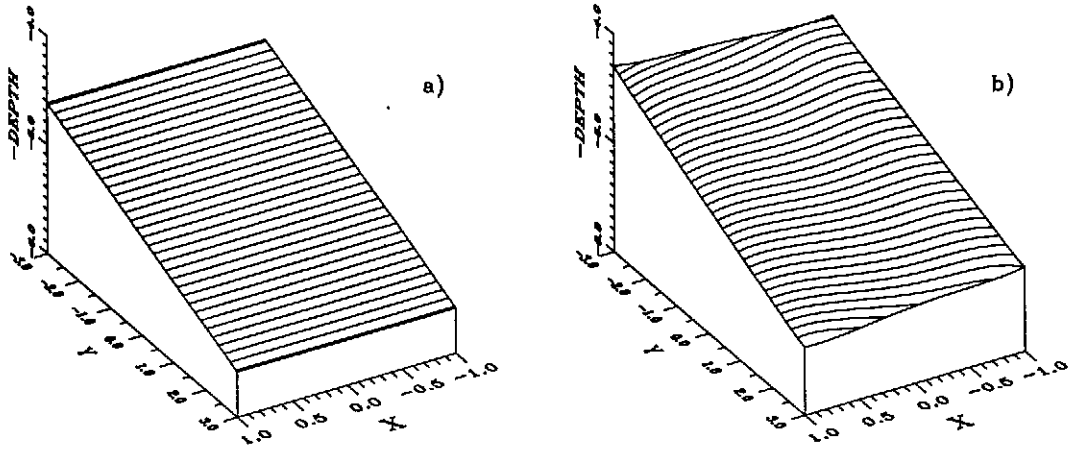

Fig. 8. Reconstruction of the interface. a) real interface; b) filtered image of the interface. Section $[-1 \leq x \leq 1] \cap [-3 \leq y \leq 3]$.

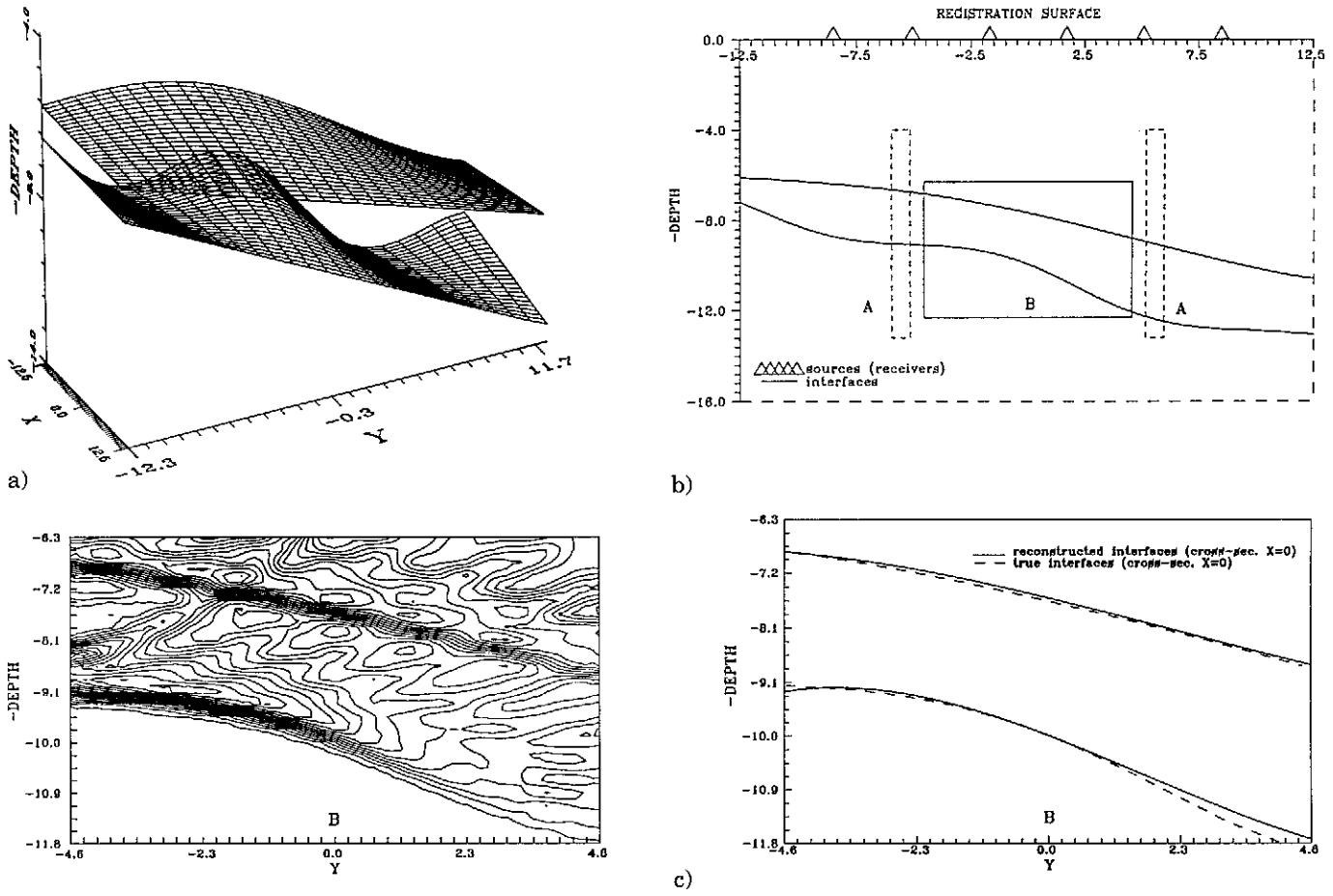

Fig. 9. 3D inversion experiment for two curved dipping interfaces. a)general view of the interfaces; b) geometry of the experiment (slice X=0). Cross-sections A indicate a region for EnIC calculations. Cross-section B indicates a region for image reconstruction; c) reconstruction of interface image for cross-section B. Upper - unfiltered image. Lower - filtered image of interfaces.

6 Conclusions.

The algorithm based on REnIC optimization significantly reduces the computational effort involved in 3D inversion. The linear inversion (RT-algorithm) takes into account all the available data rather than their subsets, in contradistinction to Symes and Carazzone (1991), Jin and Madariaga (1992). Linear inversion appears as an element of the evaluation of the EnIC and, at the last step, it allows imaging of the interfaces. In the tests described in Sect. 5 we have applied linear inversion based on the Born approximation and its generalizations. It involves two Green functions which have to be calculated for the background medium. A cheap way of computing Green functions in an inhomogeneous medium involves a high-frequency asymptotic wave field approximation based on ray tracing methods. This approach is also well adapted to the reconstruction of the background because it involves ray tracing from the sources and receivers to a relatively small subset $\mathcal{T}$ of the physical space. It is clear that the time delays along the rays depend on the slowness field in a much larger subset of the physical space, coupling the background model to interface imaging in $\mathcal{T}$.

To keep the calculations as cheap as possible we applied the approach just to the scalar wave equation, using rather rough approximations. But the main goal of our computer experiments was to check the ability of the EnIC to measure the focusing feature of velocity model.

In fact the approach does not need in the additive representation of the medium function $m = s_0 + \delta s + \mu$. The only condition is a priori assumption that data can be treated as linearized response if the background is taken properly (e.g. Clayton and Stolt, 1981).

The approach can be applied, for example, to the elastic inversion problem. Elastic inversion deals with two background models, say so m_P and m_S, and gives two images μ_P and μ_S. Assuming a priori that jumps of these two images occur in the same spatial regions, it is possible to suggest a regularization likewise the entropy

of image contrast. Namely, introducing two pseudo-probability density functions

$$p_P(\mathbf{x}) = (\nabla\mu_P(\mathbf{x}))^2 / \int_\Omega (\nabla\mu_P(\mathbf{x}))^2 \mathrm{d}x$$
$$p_S(\mathbf{x}) = (\nabla\mu_S(\mathbf{x}))^2 / \int_\Omega (\nabla\mu_S(\mathbf{x}))^2 \mathrm{d}x \qquad (28)$$

one can use Kullback-Leibler measure as the regularizing functional

$$I\ (p_P \mid p_S) = \\ = \int p_P(\mathbf{x}) \ln \frac{p_P(\mathbf{x})}{p_S(\mathbf{x})} \mathrm{d}x + \int p_S(\mathbf{x}) \ln \frac{p_S(\mathbf{x})}{p_P(\mathbf{x})} \mathrm{d}x \qquad (29)$$

It allows , for example, the ratio of P- and S-velocities to be arbitrary while treating of backgrounds but to get near-cooperative jumps (in the same spatial regions).

We suppose the EnIC can be useful for detecting of any near-singular events, for example for detecting the signal from time series, for location of single source.

Appendix A Linearized inversion by the modified gradient method.

The nonregularized solution of the problem (3) can be calculated by an iterative procedure (the modified gradient method):

$$m_{k+1}(x) = m_k(x) - \mathcal{R} \frac{\delta}{\delta m(x)} |_{m=m_k} J^d(m) \qquad (A1)$$

where $\mathcal{R}$ is a positive operator, and the gradient of $J^d(m)$ is given by

$$\frac{\delta}{\delta m} J^d(m) = -\langle \frac{\delta}{\delta m} d_{cal} \mid d - d_{cal}(m) \rangle \qquad (A2)$$

The derivative of d_{cal} with respect to m is given by

$$\frac{\delta}{\delta m} d_{cal} = \frac{\delta}{\delta m} D\, \varphi_{in} = D [\frac{\delta}{\delta m} L_m^{-1}] f = \\ = -D\, L_m^{-1} \mathcal{V} L_m^{-1} f = -D\, L_m^{-1} \mathcal{V} \varphi_{in} \qquad (A3)$$

The *incoming* sounding field φ_{in} is the solution of the forward problem

$$L_m\, \varphi_{in} = f \qquad (A4)$$

where L_m denotes the propagation operator for the given medium. The medium is specified by a set of functions m, e.g. $\lambda(x), \mu(x), \rho(x)$ for an elastic isotropic medium.

More specifically, the formal equation (A4) can be formulated in terms of a pseudo-differential operator $\mathcal{L}_m$ representing the propagation in the medium m and the boundary/initial conditions:

$$\mathcal{L}_m \varphi_{in} = \chi$$
$$\Gamma_m \varphi_{in} = \gamma \qquad (A5)$$

The operator L_m has an inverse $G_m = L_m^{-1}$ mapping source function f to a sounding signal field φ under given initial/boundary conditions.

The operator $\mathcal{V}$

$$\mathcal{V}_{\mathbf{x}} = \frac{\delta}{\delta m(\mathbf{x})} L_m \qquad (A6)$$

is the Fréchet-derivative of operator L_m with respect to m at $\mathbf{x}$.

Let us now analyze a step of the iteration procedure (A1), omitting the indices k.

In terms of the data residual $d_{res} = d - d_{cal}$ the model update $m_{res}(\mathbf{x}) = m_{k+1}(\mathbf{x}) - m_k(\mathbf{x})$ (A1) can be rewritten as follows:

$$m_{res}(\mathbf{x}^*) = -\mathcal{R} \frac{\delta}{\delta m(\mathbf{x}^*)} J^d(m) = \\ = \langle\, d_{res} \mid \mathcal{R} D\, L_m^{-1} \mathcal{V}_{\mathbf{x}^*}\, \varphi_{in} \,\rangle \qquad (A7)$$

The Born response of the data is given by

$$d^{(B)} \stackrel{\text{def}}{=} \frac{\delta}{\delta m} d_{cal} = D L^{-1} \mathcal{V} \varphi_{in} \qquad (A8)$$

For $\mathcal{R} = \beta I$ in (A1) we can write a representation of the restored perturbation m_{res} in the *data space*:

$$m_{res}(\mathbf{x}^*) = \beta \langle\, d_{res} \mid d^{(B)}_{x^*} \rangle_D \qquad (A9)$$

Supposing for simplicity that data d_{res} are deconvolved we can rewrite eq. (A7) in terms of the wave fields φ_{in} and φ_{out}.

$$m_{res}(x^*) = \beta \langle\, \varphi_{out} \mid \mathcal{V}_{x^*} \mid \varphi_{in} \,\rangle_\Phi \qquad (A10)$$

where the scalar product in the wave-field space Φ is defined by integration over space-time, and φ_{out} denotes an *outgoing* wave field (e.g.Tarantola and Valette (1982), Ryzhikov and Troyan (1992a)). The outgoing wave field obeys the equation:

$$L^\dagger \varphi_{out} = \tilde{d}_{res} \qquad (A11)$$

where $\tilde{d}_{res} = D^\dagger d_{res}$ and $D^\dagger$ denotes the operator, conjugate to D in the data space;. $L^\dagger$ denotes the operator, conjugate to L in the wave-field space. For a differential operator L we have the local operator $\mathcal{V}$: $\mathcal{V}_{\mathbf{x}^*} = V\, \delta(\mathbf{x} - \mathbf{x}^*)$.

The eq.(A9), or eq. (A10), can be interpreted as the approximate linear inversion of the data residuals. The inversion algorithm (10) one can obtain by substituting $(h(\mathbf{x}) + \varepsilon)^{-1} \delta(\mathbf{x} - \mathbf{x}')$ (9) as an integral kernel of the positive operator $\mathcal{R}$ in (A1), which represent a step of the quasi-Newton iteration procedure.

Appendix B Gateaux derivative of entropy

To illustrate the entropy E (12) behavior in the probability density function (pdf) space let introduce p^*:

$$p^* = \arg\sup E(p) \qquad (B1)$$

and for a pdf $p(\mathbf{x}) \geq p_0 > 0$ define a κ-parameter family of pdf's:

$$p_\kappa(\mathbf{x}) = p(\mathbf{x}) + \kappa\,(p(\mathbf{x}) - p^*(\mathbf{x})) \tag{B2}$$

Using the derivative of the entropy $E(p_\kappa)$ with respect to κ in the point $\kappa = 0$

$$E'_\kappa = -\int_\Omega (p(\mathbf{x}) - p^*(\mathbf{x}))\,\ln p(\mathbf{x})\,\mathrm{d}x \tag{B3}$$

we can write the evolution equation

$$\partial_\kappa E(p) = -\int_\Omega p^*(\mathbf{x})\,\ln\frac{p^*(\mathbf{x})}{p(\mathbf{x})}\,\mathrm{d}x + (E(p) - E^*) \tag{B4}$$

where $E^* = E(p^*) = \sup E$ and therefore $(E(p) - E^*) \leq 0$. The integral $I(p^*;\,p) = \int_\Omega p^*(\mathbf{x})\,\ln\frac{p^*(\mathbf{x})}{p(\mathbf{x})}\,\mathrm{d}x$ is Kullback-Leibler measure, that is the first part in (B4) is negative and $\partial_{p_\kappa} E < 0$. It means that the value of entropy $E(p_\kappa)$ is decreasing while changing of pdf along the line $p(\mathbf{x}) + \kappa\,(p(\mathbf{x}) - p^*(\mathbf{x}))$, i.e. at least linear increasing of any small perturbation of $p^* = \arg\sup E(p)$ leads to monotonous decreasing of the entropy.

As another example to illustrate the entropy structure one can take a 3-parametrical family of Gaussian pdf's:

$$\tilde{p}(\mathbf{x}) = \frac{1}{\sqrt{2\pi\det H}}\exp\{-1/2(\mathbf{x},\,H^{-1}\mathbf{x})\} \tag{B5}$$

where parameters of pdf family $\tilde{p}(\mathbf{x})$ are eigenvalues of the covariance matrix H. As far as the entropy of Gaussian pdf's can be written as

$$E(\tilde{p}) = const + \frac{1}{2}\ln\det H \tag{B6}$$

it easy to see that decreasing of any eigenvalue of the covariance matrix H decreases the value of entropy.

Appendix C Global optimization algorithm

Global optimization algorithm is based on Regularized Global Approximation of a objective function (RGA-algorithm). Using values of the objective function (OF) calculated on a set of points a differentiable global approximation of the OF is constructed, which can be interpreted as prediction of OF-value at any point of the domain of searching. Next step consists in expansion of the initial set of points by extremal points of the OF-approximation. The OF is to be calculated at these points. Renewing set of OF values gives a renewing OF-approximation. The termination of the algorithm is defined by condition that minimum of calculated values of OF is reached namely at the point of global minimum of OF-approximation provided the OF-value is well-predicted.

In our case the OF $\Re$ is regularized entropy of image contrast (eq. 20) that can be rewritten in terms of slowness parameters α (24) as follow

$$\Re = \tilde{\mathcal{E}}(\alpha) = \mathcal{E}(\alpha) + \langle\,\alpha \mid \alpha\,\rangle$$

where $\alpha = \{\alpha_\gamma,\ \gamma = \mathbf{1},\mathbf{2},\ldots\}$ and $\langle\,\alpha \mid \alpha\,\rangle = \sum_\gamma \alpha_\gamma^2$.

So the optimization problem is to find global minimum of $\Re$ in the region $\mathcal{A}$, foe example $\mathcal{A} = \{\,\|\,\alpha\,\| \leq \mathbf{C}\}$

$$\alpha^* = \arg\inf_{\alpha\in\mathcal{A}}\ \Re(\alpha) \tag{C1}$$

RGA-algorithm can be expressed in terms of the following iteration procedure:

$$\alpha^{\mathbf{k+1}} = \arg\inf_{\alpha\in\tilde{\mathcal{A}}^{\mathbf{k+1}}\subset\mathcal{A}}\ \Re(\alpha) \tag{C2}$$

The subset $\tilde{\mathcal{A}}^{k+1}$ is obtained using the regularized approximation $\tilde{\Re}^k$ of objective function $\Re$ calculated at the set of [K] points $\tilde{\mathcal{A}}^k = \{\alpha^k = (\alpha^1, \alpha^2, \ldots, \alpha^{[K]})\}$:

$$\tilde{\Re}^k = \arg\inf_{\tilde{\Re}}\ \{\|\,\tilde{\Re} - \mathbf{r}(\alpha^k)\,\|^2 + (\tilde{\Re}, [\partial_t - D^2\nabla^2]\,\tilde{\Re}\} \tag{C3}$$

where $\mathbf{r}(\alpha^k) = \Re(\alpha^k)$, $\|\,\tilde{\Re} - \mathbf{r}\,\|^2 := \sum_{k=1}^{[K]} (\tilde{\Re}(\alpha^k) - \mathbf{r}(\alpha^k))^2$, and $(\ ,\ )$ denotes squared $\mathcal{L}^2$-norm. The operator $[\partial_t - D^2\nabla^2]$ in the regularizing functional in (C3) stands for the positive exponential operator $\exp\{-tD^2\nabla^2\}$ with parameter of regularization tD^2. We called this regularization *diffusion regularization*. The second functional in (C3) represents a squared norm in the space that can be called as *Sobolev space of infinite order*.

The subset $\tilde{\mathcal{A}}^{k+1}$ is constructed as

$$\tilde{\mathcal{A}}^{k+1} = \{\alpha \in \tilde{\mathcal{A}}^{\mathbf{k}} \cup \delta\tilde{\mathcal{A}}^{\mathbf{k}}\} \tag{C4}$$

where $\delta\tilde{\mathcal{A}}^k = \{\ \alpha \quad \mid \quad \nabla\tilde{\mathcal{A}}^{\mathbf{k}}(\alpha) = \mathbf{0}\ \}$. The subset $\tilde{\mathcal{A}}_0$ is generated in random manner.

The point of global extremum α^* of $\Re$ and the global approximation $\tilde{\Re}^*$ are defined by

$$\tilde{\Re}^*(\alpha^*) \doteq \Re(\alpha^*) \tag{C5}$$

Acknowledgements. The research was conducted under the EU projects "Seismic Tomography Based on Advanced Ray Tracing" (associated with the Geoscience II Project " Stratigraphic Modelling & Seismic Inversion") and "3D Asymptotic Seismic Imaging" (JOU2-93-0321). Both projects were financed by the Norwegian Research Foundation, Oslo.

The authors (G.R. and M.B.) would like to thank the anonymous reviewer for many helpful comments, which improved the exposition.

References

Beydoun, W. B. and Mendes, M., Elastic ray-Born l_2-migration/inversion, *Geophys. J. Int.*, *97*, 151-160, 1989.

Beylkin, G., and Burridge, R., Linearized inverse scattering problems in acoustics and elastisity, *Wave Motion*, *12*, 15-52, 1990.

Chapman, C. H., Coates, R. T., Generalized Born scattering in anysotropic media, *Wave Motion*, *19*, 309-341, 1994.

Clayton, R. W., Stolt R. H., A Born-WKBJ inversions method for acoustic reflection data, *Geophysics*, *46*, 1559-1567, 1981.

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

Jin, S., Madariaga, R., Virieux, J., and Lambare, G., Two-dimensional asymptotic iterative elastic inversion, *Geophys. J. Int.*, *108*, 575-588, 1992

Luo, Y. and Schuster, G. T., Wave equation traveltime inversion, *Geophysics*, *56*, 645-653, 1992.

Miller, D., Oristaglio, M. and Beylkin, G., A new slant on seismic imaging: Migration and integral geometry *Geophysics*, *52*, 943-964, 1987

Ryzhikov, G. and Biryulina, M., RGA-algorithm for global optimization with applications in seismic inversion, in preparation, 1994

Ryzhikov, G., and Troyan, V., On regularization methods in 3-D ray tomography, in *Geophysical Data Interpretation by Inverse Modelling. Proc. of the 9-th International Seminar on Model Optimization in Exploration Geophysics. Berlin 1991. Edited by A.Vogel et al, Vieweg. Wiesbaden*, 53-61, 1991.

Ryzhikov, G. and Troyan, V., 3D diffraction tomography. Part 1: Construction and interpretation of tomography functionals, in *Expanded Abstracts, Russian - Norwegian Oil Exploration Workshop II, Voss, May 5-7, 1992.*

Ryzhikov, G. and Troyan, V., 3D diffraction tomography. Part 2: Reconstruction algorithm with statistical regularization, in *Expanded Abstracts, Russian - Norwegian Oil Exploration Workshop II, Voss, May 5-7, 1992*, pp. 1-14, 1992.

Symes, W.and Carazzone, J., Velocity inversion by differential semblance optimization, *Geophysics*, *56*, 654-663, 1991

Tarantola, A., *Inverse problem theory.* Elsevier,1987.

Tarantola, A. and Valette, B., Generalized nonlinear inverse problems solved using the least squares criterion *Rev. Geophys. Space Phys.*, *50*, 219-232, 1982

Tikhonov, A. N., Arsenin, V. Y., *Solution of ill-posed problems.* John Wiley and Sons Inc., 1977.

Versteeg, R. J., Sensitivity of prestack depth migration to the velocity model, *Geophysics*, *58*, 873-882, 1993.