Application of the medium covariance functions to travel–time tomography

Luděk Klimeš

Department of Geophysics, Charles University, Ke Karlovu 3, 121 16 Praha 2, Czech Republic, Fax: +420-2-2191 1214, E-mail: klimes@seis.karlov.mff.cuni.cz

Summary

The slowness describing the geological structure is assumed to be a representation of a random medium described in terms of the first two statistical moments: the mean value and the medium covariance function. The proposed approach to seismic travel–time tomography enables the equations for the inversion of travel times to be derived without additional subjective a priori information, and the accuracy and resolution of the resulting seismic model of the geological structure to be estimated.

Keywords

Seismic model, travel–time tomographic inversion, covariance functions and matrices, smoothing, resolution.

1 Introduction

The seismic model (macro model) approximates the spatial distribution of the material properties of the geological structure.

It is obvious that geological structures cannot be definitely described in terms of a finite number of model parameters: going from large to small scales, more and more details of the medium may be discovered. From this point of view, the seismic model can never represent all details of the medium, no matter how numerous the seismic data are.

On the other hand, we cannot consider the individual small details of the medium to be mutually independent. Such a medium, with statistical properties of white noise, would be unpredictable from a finite number of seismic data and seismic inversion would be nonsense. We have to assume that the probability of the same or similar material properties at two points increases with decreasing distance of the points. This property of the medium may, in the roughest approximation, be described in terms of medium covariance functions.

This paper, concentrated on the inversion of travel times, follows the approach of Franklin (1970), Tarantola & Valette (1982), Tarantola & Nercessian (1984) and Tarantola (1987), similarly as the related paper of Maurer, Holliger & Boerner (1998). Consideration of the medium covariance functions in the travel–time inversion is inevitable if the ray coverage of the medium is sparse and uneven. Since travel times contain only the information on the average slowness along the corresponding two–point rays, we would have no information on the behaviour of slowness between the rays without the medium covariance functions. The medium covariance functions provide statistically correct interpolation of slowness between the rays, especially over distances larger

In: Seismic Waves in Complex 3-D Structures, Report 11 (Department of Geophysics, Charles University, Prague, 2001), pp. 73–89; to appear in Pure and Applied Geophysics, **159** (2002)

than the dimensions of Fresnel volumes. The properties of this inversion have numerically been demonstrated by Maurer, Holliger & Boerner (1998) by a very nice 2-D synthetic example. Analogous behaviour of the inversion for 1-D electrical resistivities has been shown by Pilkington & Todoeschuck (1990). The medium covariance functions simultaneously enable the information, contained in travel times picked in several solitary stations, to be protected from the superior force of travel times picked along dense receiver profiles, whereas the still often used "naive" least squares, with incorrect travel–time weights, might result in substantial loss of information and considerably bad models. Different inverse methods are designed for different kinds of data and should not be applied in a wrong way. For example, in contrast with this paper, the inverse theory developed by Backus & Gilbert (1970) is designed for the inversion of gross earth data but not for the inversion of data which are considerably sparse and uneven with respect to their resolution. For comparison refer to Box 7.2 of Tarantola (1987).

In this paper, the slowness describing the geological structure is assumed to be a representation of a random medium with known statistical properties. We assume that some travel times corresponding to the geological structure (i.e. to the representation of the random medium) are known, and determine the seismic model as the "best" approximation to the (possibly smoothed) slowness describing the geological structure. The "best" approximation means the smallest Lebesgue L_2 norm of the standard deviations between the model and the (possibly smoothed) slowness describing the geological structure.

This approach enables us to arrive at the final formulae for linearized travel–time tomography and to separate the smoothing effects of (a) the application of the covariance functions in the inversion for the non–smoothed medium, (b) the inversion for the smoothed medium compared to the inversion for the non–smoothed medium, (c) the finite model parametrization, and (d) the additional a priori smoothing information specified in terms of the Sobolev norm, see Section 3.5.

The derived equations apply to the final model resulting from an iterative inversion. The final model should be independent of the iterations, depending only on the mean model, the medium covariance function, and the data. Yet the equations still may be applied to individual linearized iterations in such a way that the ray characteristic functions are approximated by the ones calculated in the initial model of the iteration. The Einstein summation over repetitive indices is used throughout the paper.

2 Formulation

2.1 Random slowness

Assume that slowness $u(\mathbf{x})$ in the geological structure is a realization of the random quantity with the first statistical moment (*mean value*)

$$u^0(\mathbf{x}) = \langle u(\mathbf{x}) \rangle \tag{1}$$

and the second statistical moment

$$\langle u(\mathbf{x}_1)\, u(\mathbf{x}_2) \rangle \quad . \tag{2}$$

We shall call a priori covariance function

$$C(\mathbf{x}_1, \mathbf{x}_2) = \langle u(\mathbf{x}_1)\, u(\mathbf{x}_2) \rangle - u^0(\mathbf{x}_1)\ u^0(\mathbf{x}_2) \tag{3}$$

of the slowness the *medium covariance function*. Here $u^0(\mathbf{x})$ is the a priori known deterministic component of the random medium serving as the background slowness in estimating the medium covariance function $C(\mathbf{x}_1, \mathbf{x}_2)$, which should be determined prior to the proposed tomographic inversion.

We shall see that the determination of $u^0(\mathbf{x})$ by means of very strongly overdetermined least squares is consistent with the proposed tomographic inversion. The least–square fitting of $u^0(\mathbf{x})$ must be very strongly overdetermined in order to leave a sufficient amount of information to estimate medium covariance function $C(\mathbf{x}_1, \mathbf{x}_2)$.

The determination of $u^0(\mathbf{x})$, parametrized by several parameters, and $C(\mathbf{x}_1, \mathbf{x}_2)$, parametrized by one or a few parameters, should be carried out prior to the proposed tomographic inversion. If the medium covariance function depends only on vectorial distance $\mathbf{x}_1 - \mathbf{x}_2$, it may be expressed in terms of *medium correlation function* $c(\mathbf{x})$,

$$C(\mathbf{x}_1, \mathbf{x}_2) = c(\mathbf{x}_1 - \mathbf{x}_2) \tag{4}$$

(Tarantola 1987). One possibility of determining $u^0(\mathbf{x})$ and $c(\mathbf{x})$ has been demonstrated by Klimeš (2002c). Various types of the medium correlation functions $c(\mathbf{x})$ used in geophysics have been systematically listed by Klimeš (2002b).

2.2 Travel times and rays

Field travel times T_I are the integrals of slowness $u(\mathbf{x})$ along rays Ω_I plus the *picking errors*,

$$T_I = \int G_I(\mathbf{x}) u(\mathbf{x}) \, \mathrm{d}\mathbf{x} + \delta T_I \quad , \tag{5}$$

where the integral with ray characteristic function $G_I(\mathbf{x})$ is taken over the model volume. *Uppercase subscripts* are used hereinafter to denote travel times and the corresponding rays. We assume zero mean values of the picking errors,

$$\langle \delta T_I \rangle = 0 \quad , \tag{6}$$

and introduce the covariance matrix of picking errors (*data covariance matrix*)

$$T_{IJ} = \langle \delta T_I \ \delta T_J \rangle \tag{7}$$

describing errors δT_I. We assume no correlation between picking errors δT_I and the exact slowness $u(\mathbf{x})$,

$$\langle \delta T_I \ u(\mathbf{x}) \rangle = 0 \quad . \tag{8}$$

Ray characteristic functions

$$G_I(\mathbf{x}) = \frac{\partial T_I}{\partial u(\mathbf{x})} \quad , \tag{9}$$

also called "travel–time sensitivity functions" or "sampling functions", characterize the sensitivity of travel times to slowness perturbations. Ray characteristic functions $G_I(\mathbf{x})$ should ideally be calculated for the exact medium, but during linearized inversion they are approximated by the ones calculated in the best model of the last iteration.

The exact ray characteristic functions, describing "thick rays", depend on the definition of the arrival times picked for the inversion. For examples refer, e.g., to Woodward & Rocca (1988), Coates & Chapman (1990), Woodward (1992), Pratt, Song, Williamson & Warner (1996) or Marquering, Nolet & Dahlen (1998).

From the point of view of the high–frequency asymptotic ray theory, the ray characteristic function $G_I(\mathbf{x})$ of "thin" ray Ω_I has the form of the 2-D Dirac distribution

$\delta(q^1, q^2)$ in ray–centred coordinates (q^1, q^2, q^3), with q^1 and q^2 being the Cartesian coordinates in the plane perpendicular to the ray parametrized by arclength q_3,

$$G_I(\mathbf{x}(q^1, q^2, q^3)) = \delta(q^1, q^2) \tag{10}$$

for the values of q^3 between the source and the receiver, zero for other q_3. For the two–point ray tracing refer, e.g., to Bulant (1996, 1999).

If the spacing between two–point rays exceeds the dimensions of the relevant Fresnel volumes considerably, the sensitivity of the resulting seismic model to the shape of the ray characteristic functions vanishes and the differences between "thick" and "thin" rays practically disappear.

2.3 Seismic model

The slowness distribution $\widetilde{u}(\mathbf{x})$ in the seismic model is considered to be an approximation to the smoothed version $\overline{u}(\mathbf{x})$ of the slowness $u(\mathbf{x})$ of the medium,

$$\overline{u}(\mathbf{x}) = \int D(\mathbf{x}, \mathbf{x}')\, u(\mathbf{x}')\, \mathrm{d}\mathbf{x}' \quad , \tag{11}$$

where $D(\mathbf{x}_1, \mathbf{x}_2)$ is the given *smoothing function.* The smoothing function is included here for the sake of generality. It may express a priori ideas on the resolution of the model and may strongly influence the local standard deviation between the model and the geological structure, because the small details in the medium are much worse resolved than larger inhomogeneities. The *smoothing function may be taken in the form of the Dirac distribution*, $D(\mathbf{x}_1, \mathbf{x}_2) = \delta(\mathbf{x}_1 - \mathbf{x}_2)$. In such a case, no a priori smoothing is applied during the tomographic travel–time inversion and the slowness distribution $\widetilde{u}(\mathbf{x})$ in the seismic model is an approximation to the slowness $u(\mathbf{x})$ of the medium. Smoothing may then be applied to the model a posteriori. This approach may be, to some extent, limited only by an insufficiently rough parametrization of the model or other functions.

Model slowness $\widetilde{u}(\mathbf{x})$ depends on the exact slowness $u(\mathbf{x})$ through travel times T_I. We assume the model to be linearly dependent on travel times,

$$\boxed{\widetilde{u}(\mathbf{x}) = R^0(\mathbf{x}) + R_K(\mathbf{x}) \left[T_K - T_K^0\right] \quad ,} \tag{12}$$

where

$$\boxed{T_I^0 = \int G_I(\mathbf{x}) u^0(\mathbf{x})\, \mathrm{d}\mathbf{x}} \tag{13}$$

is used to denote the integrals with the ray characteristic functions $G_I(\mathbf{x})$ corresponding to exact rays Ω_I (*reference travel times*), although we know that this linearity may be violated by inaccurate approximations to $G_I(\mathbf{x})$ during the iterative linearized inversion. Note that, because of (5) and (13) with (1) and (6),

$$\langle T_K - T_K^0 \rangle = 0 \quad . \tag{14}$$

We assume here that the differences between $u(\mathbf{x})$, $\overline{u}(\mathbf{x})$ and $\widetilde{u}(\mathbf{x})$ are sufficiently small to be able to calculate the corresponding travel times along the same rays within an accuracy which would not distort the complete covariance matrix S_{IJ} of travel times, introduced later on by equation (34). In other words, we assume that the *linearization* of travel times with respect to medium perturbations may be applied in the *vicinity of the exact model.*

After inserting (5) and (13), equation (12) reads

$$\widetilde{u}(\mathbf{x}) = R^0(\mathbf{x}) + \int R(\mathbf{x},\mathbf{x}') \left[u(\mathbf{x}') - u^0(\mathbf{x}')\right] \, \mathrm{d}\mathbf{x}' + R_K(\mathbf{x})\delta T_K \quad , \tag{15}$$

where

$$R(\mathbf{x}_1,\mathbf{x}_2) = R_K(\mathbf{x}_1)G_K(\mathbf{x}_2) \quad , \tag{16}$$

is the *resolving kernel* of the *resolving operator*, linearly projecting the exact slowness $u(\mathbf{x})$ onto model $\widetilde{u}(\mathbf{x})$.

For the sake of generality, we allow the resolving kernel to be constrained by zero to several a priori conditions indexed here with α,

$$\Delta^\alpha R^0(\mathbf{x}) + \int R(\mathbf{x},\mathbf{x}')\, u^\alpha(\mathbf{x}') \, \mathrm{d}\mathbf{x}' = \int D(\mathbf{x},\mathbf{x}') \left[\Delta^\alpha u^0(\mathbf{x}') + u^\alpha(\mathbf{x}')\right] \, \mathrm{d}\mathbf{x}' \quad , \tag{17}$$

where $0 \le \Delta^\alpha \le 1$ are the weighting coefficients. For weighting coefficients $\Delta^\alpha = 1$, these projection constraints express the condition that the travel–time inversion yields exact results in the structures whose differences from $u^0(\mathbf{x})$ lie in a linear subspace generated by a few functions $u^\alpha(\mathbf{x})$. For weighting coefficients $\Delta^\alpha = 0$, we do not require the travel–time inversion to yield exact results in the structure described by $u^0(\mathbf{x})$, but enforce the exact results for perturbations $u^\alpha(\mathbf{x})$ of a structure from $u^0(\mathbf{x})$, allowing $u^0(\mathbf{x})$ to be smoothed by means of the Sobolev scalar products. We shall see later in Section 3.4 that these conditions are closely related to adjusting mean slowness $u^0(\mathbf{x})$ by a linear combination of functions $u^\alpha(\mathbf{x})$, fitted by least squares.

2.4 Parametrization of the model and resolving kernel

Assume that the slowness distribution $\widetilde{u}(\mathbf{x})$ in the seismic model, which is the best possible approximation to $\overline{u}(\mathbf{x})$, is composed of a finite set of *basis functions* $B_i(\mathbf{x})$ indexed hereinafter with *lowercase subscripts*,

$$\boxed{\widetilde{u}(\mathbf{x}) = B_i(\mathbf{x})\; u_i \quad .} \tag{18}$$

For example, tricubic B-splines may be employed for the basis functions $B_i(\mathbf{x})$. Comparing now (18) with (12), we see that functions $R^0(\mathbf{x})$ and $R_K(\mathbf{x})$ should also be composed of the basis functions,

$$R^0(\mathbf{x}) = B_i(\mathbf{x})R^0_i \quad , \tag{19}$$

$$R_K(\mathbf{x}) = B_i(\mathbf{x})R_{iK} \quad , \tag{20}$$

that the resolving kernel (16) takes the form

$$R(\mathbf{x}_1,\mathbf{x}_2) = B_i(\mathbf{x}_1)R_{iI}G_I(\mathbf{x}_2) \quad , \tag{21}$$

and that equation (12) can be expressed in component form

$$\boxed{u_i = R^0_i + R_{iI}\left[T_I - T^0_I\right] \quad .} \tag{22}$$

Assume that optional functions $u^\alpha(\mathbf{x})$ may also be composed of the same basis functions,

$$u^\alpha(\mathbf{x}) = u^\alpha_i \; B_i(\mathbf{x}) \quad . \tag{23}$$

Optional constraints (17) with (19), (21) and (23), L_2–projected on the basis functions, then take the form

$$\Delta^\alpha R_i^0 + R_{iI}\, T_I^\alpha = \Delta^\alpha \overline{u}_i^0 + \overline{u}_i^\alpha \quad , \tag{24}$$

where

$$T_I^\alpha = \int G_I(\mathbf{x}) u^\alpha(\mathbf{x})\, \mathrm{d}\mathbf{x} \quad , \tag{25}$$

are the travel–time variations with respect to functions $u^\alpha(\mathbf{x})$,

$$\overline{u}_i^\alpha = B_{ik}^{-1} \iint B_k(\mathbf{x}_1) D(\mathbf{x}_1, \mathbf{x}_2) u^\alpha(\mathbf{x}_2)\, \mathrm{d}\mathbf{x}_1\, \mathrm{d}\mathbf{x}_2 \tag{26}$$

are the components of the Lebesgue L_2 projections of smoothed perturbations $u^\alpha(\mathbf{x})$ onto the basis functions, and

$$\boxed{\overline{u}_i^0 = B_{ik}^{-1} \iint B_k(\mathbf{x}_1) D(\mathbf{x}_1, \mathbf{x}_2) u^0(\mathbf{x}_2)\, \mathrm{d}\mathbf{x}_1\, \mathrm{d}\mathbf{x}_2 \qquad (27)}$$

are the components of the Lebesgue L_2 projections of the smoothed mean slowness onto the basis functions.

2.5 Model covariance function

The posterior covariance function describing the deviation of model slowness $\widetilde{u}(\mathbf{x})$ from the smoothed exact slowness $\overline{u}(\mathbf{x})$ is

$$\widetilde{C}(\mathbf{x}_1, \mathbf{x}_2) = \langle [\widetilde{u}(\mathbf{x}_1) - \overline{u}(\mathbf{x}_1)][\widetilde{u}(\mathbf{x}_2) - \overline{u}(\mathbf{x}_2)] \rangle \quad . \tag{28}$$

We shall call it the *model covariance function*. Although the model covariance function is frequently neglected in the geophysical literature, it represents together with the model the most important result of the inversion. The model without the model covariance function is worthless from the scientific point of view. We thus derive here the equations for the full posterior covariance function $\widetilde{C}(\mathbf{x}_1, \mathbf{x}_2)$, although only the posterior variance function $\widetilde{C}(\mathbf{x}, \mathbf{x})$ appears in the objective function for the inversion.

Inserting (11) and (15) into (28), and considering (1), (6) and (8), we arrive at

$$\begin{aligned} \widetilde{C}(\mathbf{x}_1, \mathbf{x}_2) = & \iint [R(\mathbf{x}_1, \mathbf{x}_1') - D(\mathbf{x}_1, \mathbf{x}_1')]\, [R(\mathbf{x}_2, \mathbf{x}_2') - D(\mathbf{x}_2, \mathbf{x}_2')] \\ & \qquad \times \langle [u(\mathbf{x}_1') - u^0(\mathbf{x}_1')]\; [u(\mathbf{x}_2') - u^0(\mathbf{x}_2')] \rangle\, \mathrm{d}\mathbf{x}_1'\, \mathrm{d}\mathbf{x}_2' \\ & + R_I(\mathbf{x}_1) R_J(\mathbf{x}_2) \langle \delta T_I\; \delta T_J \rangle \\ + \left[R^0(\mathbf{x}_1) - \int D(\mathbf{x}_1, \mathbf{x}_1') u^0(\mathbf{x}_1')\, \mathrm{d}\mathbf{x}_1' \right] & \left[R^0(\mathbf{x}_2) - \int D(\mathbf{x}_2, \mathbf{x}_2') u^0(\mathbf{x}_2')\, \mathrm{d}\mathbf{x}_2' \right] \quad . \end{aligned} \tag{29}$$

Using then (3), (7), and (20), we arrive at

$$\begin{aligned} \widetilde{C}(\mathbf{x}_1, \mathbf{x}_2) = & \iint [R(\mathbf{x}_1, \mathbf{x}_1') - D(\mathbf{x}_1, \mathbf{x}_1')]\, [R(\mathbf{x}_2, \mathbf{x}_2') - D(\mathbf{x}_2, \mathbf{x}_2')]\, C(\mathbf{x}_1', \mathbf{x}_2')\, \mathrm{d}\mathbf{x}_1'\, \mathrm{d}\mathbf{x}_2' \\ & + B_i(\mathbf{x}_1) B_j(\mathbf{x}_2) R_{iI} R_{jJ} T_{IJ} \\ + \left[R^0(\mathbf{x}_1) - \int D(\mathbf{x}_1, \mathbf{x}_1') u^0(\mathbf{x}_1')\, \mathrm{d}\mathbf{x}_1' \right] & \left[R^0(\mathbf{x}_2) - \int D(\mathbf{x}_2, \mathbf{x}_2') u^0(\mathbf{x}_2')\, \mathrm{d}\mathbf{x}_2' \right] \quad . \end{aligned} \tag{30}$$

Insertion of (19) and (21) into (30) yields

$$
\begin{aligned}
\widetilde{C}(\mathbf{x}_1,\mathbf{x}_2) = & \iint D(\mathbf{x}_1,\mathbf{x}_1')D(\mathbf{x}_2,\mathbf{x}_2')C(\mathbf{x}_1',\mathbf{x}_2')\ \mathrm{d}\mathbf{x}_1'\ \mathrm{d}\mathbf{x}_2' \\
& -B_i(\mathbf{x}_1)R_{iI}\iint D(\mathbf{x}_2,\mathbf{x}_2')C(\mathbf{x}_2',\mathbf{x}_1')G_I(\mathbf{x}_1')\ \mathrm{d}\mathbf{x}_1'\ \mathrm{d}\mathbf{x}_2' \\
& -B_j(\mathbf{x}_2)R_{jJ}\iint D(\mathbf{x}_1,\mathbf{x}_1')C(\mathbf{x}_1',\mathbf{x}_2')G_J(\mathbf{x}_2')\ \mathrm{d}\mathbf{x}_1'\ \mathrm{d}\mathbf{x}_2' \\
& +B_i(\mathbf{x}_1)B_j(\mathbf{x}_2)R_{iI}R_{jJ}\left[\iint G_I(\mathbf{x}_1')C(\mathbf{x}_1',\mathbf{x}_2')G_J(\mathbf{x}_2')\ \mathrm{d}\mathbf{x}_1'\ \mathrm{d}\mathbf{x}_2' + T_{IJ}\right] \\
& +\left[B_i(\mathbf{x}_1)R_i^0 - \int D(\mathbf{x}_1,\mathbf{x}_1')u^0(\mathbf{x}_1')\ \mathrm{d}\mathbf{x}_1'\right]\left[B_j(\mathbf{x}_1)R_j^0 - \int D(\mathbf{x}_2,\mathbf{x}_2')u^0(\mathbf{x}_2')\ \mathrm{d}\mathbf{x}_2'\right] \quad . \qquad (31)
\end{aligned}
$$

Introducing the *ray back–projective functions* (or ray resolution functions)

$$
\boxed{\Psi_I(\mathbf{x}) = \langle T_I\ [u(\mathbf{x}) - u^0(\mathbf{x})]\rangle = \int C(\mathbf{x},\mathbf{x}')G_I(\mathbf{x}')\ \mathrm{d}\mathbf{x}' \quad , \qquad (32)}
$$

representing the covariances between slowness and field travel times, the *a priori geometrical travel–time covariance matrix*

$$
\boxed{\Theta_{IJ} = \iint G_I(\mathbf{x}_1)C(\mathbf{x}_1,\mathbf{x}_2)G_J(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2 = \int G_I(\mathbf{x}_1)\Psi_J(\mathbf{x}_1)\ \mathrm{d}\mathbf{x}_1 \qquad (33)}
$$

(Chernov 1960; Tarantola 1987), and the *complete travel–time covariance matrix*

$$
\boxed{S_{MN} = \langle\left[T_M - T_M^0\right]\ \left[T_N - T_N^0\right]\rangle = T_{MN} + \Theta_{MN} \quad , \qquad (34)}
$$

equation (31) for the model covariance function becomes

$$
\boxed{
\begin{aligned}
\widetilde{C}(\mathbf{x}_1,\mathbf{x}_2) = & \iint D(\mathbf{x}_1,\mathbf{x}_1')D(\mathbf{x}_2,\mathbf{x}_2')C(\mathbf{x}_1',\mathbf{x}_2')\ \mathrm{d}\mathbf{x}_1'\ \mathrm{d}\mathbf{x}_2' \\
& -B_i(\mathbf{x}_1)R_{iI}\int D(\mathbf{x}_2,\mathbf{x}_2')\Psi_I(\mathbf{x}_2')\ \mathrm{d}\mathbf{x}_2' \\
& -B_j(\mathbf{x}_2)R_{jJ}\int D(\mathbf{x}_1,\mathbf{x}_1')\Psi_J(\mathbf{x}_1')\ \mathrm{d}\mathbf{x}_1' \\
& +B_i(\mathbf{x}_1)B_j(\mathbf{x}_2)R_{iI}R_{jJ}S_{IJ} \\
& +\left[B_i(\mathbf{x}_1)R_i^0 - \int D(\mathbf{x}_1,\mathbf{x}_1')u^0(\mathbf{x}_1')\ \mathrm{d}\mathbf{x}_1'\right]\left[B_j(\mathbf{x}_1)R_j^0 - \int D(\mathbf{x}_2,\mathbf{x}_2')u^0(\mathbf{x}_2')\ \mathrm{d}\mathbf{x}_2'\right] \quad . \qquad (35)
\end{aligned}
}
$$

3 Inversion

3.1 Sobolev scalar product

Model covariance function (35) describes the deviation of the resulting model slowness $\widetilde{u}(\mathbf{x})$ from the smoothed exact slowness $\overline{u}(\mathbf{x})$. The deviation causes the differences between the synthetic wavefields calculated in the model and exact wavefields measured in geological structures. To minimize the deviation, the objective function minimizing the squared standard deviation $\widetilde{C}(\mathbf{x},\mathbf{x})$ may be considered.

However, there is also another source of the differences between the synthetic and exact wavefields: inaccuracy of the wavefield modelling methods in too complex models. The accuracy of all numerical methods increases with the smoothness of the model. The smoothness of the model may be conveniently expressed in terms of the square

$$\left(\|\widetilde{u}-u^{\Phi}\|^{\Phi}\right)^2 = ((\widetilde{u}-u^{\Phi},\widetilde{u}-u^{\Phi}))^{\Phi} \tag{36}$$

of the *Sobolev norm* $\|\bullet\|^{\Phi}$ of the difference between $\widetilde{u}(\mathbf{x})$ and given function

$$u^{\Phi}(\mathbf{x}) = B_i(\mathbf{x})\; u_i^{\Phi} \quad . \tag{37}$$

We may combine several smoothing expressions (36), indexed by superscript $^{\Phi}$. Here

$$((f,g))^{\Phi} = \sum_{\alpha_1}\sum_{\alpha_2}\sum_{\alpha_3}\sum_{\beta_1}\sum_{\beta_2}\sum_{\beta_3}\int W^{\Phi}_{\alpha_1\alpha_2\alpha_3\beta_1\beta_2\beta_3}(\mathbf{x})$$

$$\times\left[\left(\frac{\partial}{\partial x^1}\right)^{\alpha_1}\left(\frac{\partial}{\partial x^2}\right)^{\alpha_2}\left(\frac{\partial}{\partial x^3}\right)^{\alpha_3}f(\mathbf{x})\right]\left[\left(\frac{\partial}{\partial x^1}\right)^{\beta_1}\left(\frac{\partial}{\partial x^2}\right)^{\beta_2}\left(\frac{\partial}{\partial x^3}\right)^{\beta_3}g(\mathbf{x})\right]\mathrm{d}\mathbf{x} \tag{38}$$

is the *Sobolev scalar product* of functions $f(\mathbf{x})$ and $g(\mathbf{x})$, defined in terms of weighting functions $W^{\Phi}_{\alpha_1\alpha_2\alpha_3\beta_1\beta_2\beta_3}(\mathbf{x})$, see Tarantola (1987). The order of the Sobolev scalar product is determined by non–zero coefficients $W^{\Phi}_{\alpha_1\alpha_2\alpha_3\beta_1\beta_2\beta_3}$. The coefficients are chosen according to the application of the velocity model being constructed. The minimization of the derivatives has already been suggested by Menke (1984), but his equations for the first derivatives are applicable only to the properly normalized 1-D triangular basis functions on a regular grid, and his equations for the second derivatives only to the properly normalized 1-D quadratic B-splines on a regular grid.

As the relative standard deviation $\sqrt{\widetilde{C}(\mathbf{x},\mathbf{x})}$ quantifies the local travel–time error per the small distance of wave propagation due to the inaccuracy of the model, it is desirable to adjust weighting functions $W^{\Phi}_{\alpha_1\alpha_2\alpha_3\beta_1\beta_2\beta_3}(\mathbf{x})$ in such a way that the Sobolev norm $\|\widetilde{u}-u^{\Phi}\|^{\Phi}$ approximates the analogous local travel–time errors due to the inaccuracies of the numerical methods expected to be applied in the resulting model. For examples of constructing particular Sobolev norms for various numerical forward–modelling methods refer to Klimeš (2000), Bulant (2002) and Žáček (2002). Let us only mention here two Sobolev scalar products relevant to the ray tracing. The complicated behaviour of rays depends on the second velocity derivatives (Klimeš 2002a) and may thus be restricted by minimizing the second–order Sobolev norm $\|\bullet\|^{\Phi}$ given by a completely symmetric tensor of coefficients (Klimeš 2000), with $u^{\Phi}(\mathbf{x}) = 0$. The second–order travel–time perturbation depends on the perturbation of the first velocity derivatives (Farra 1999; Červený 2001). It may thus be useful to limit the velocity model updates within individual iterations of the linearized inversion by means of the isotropic first–order Sobolev norm $\|\bullet\|^{\Phi}$, with $u^{\Phi}(\mathbf{x})$ equal to the initial model of the respective iteration of the linearized inversion.

In order to improve the efficiency and accuracy of numerical forward–modelling algorithms to be applied in the resulting model, it may thus be desirable to constrain the spatial slowness variations by means of adding squared Sobolev norm $\left(\|\widetilde{u}-u^{\Phi}\|^{\Phi}\right)^2$, quantifying the roughness of the resulting slowness distribution $\widetilde{u}(\mathbf{x})$, to the objective function y minimizing the squared standard deviation $\widetilde{C}(\mathbf{x},\mathbf{x})$ integrated over the model volume.

Unfortunately, the addition of $((\widetilde{u}-u^{\Phi},\widetilde{u}-u^{\Phi}))^{\Phi}$ would violate the assumption (15) of the linear dependence of the model on travel times. Since the squared standard deviation $\widetilde{C}(\mathbf{x},\mathbf{x})$ between the medium and its model is minimized, instead of minimizing the difference corresponding to a particular representation, we should treat the Sobolev norm consistently and minimize the mean value $\langle((\widetilde{u}-u^{\Phi},\widetilde{u}-u^{\Phi}))^{\Phi}\rangle$ instead of $((\widetilde{u}-u^{\Phi},\widetilde{u}-u^{\Phi}))^{\Phi}$.

3.2 The objective function

The objective function is thus assumed in the form

$$y=\int\widetilde{C}(\mathbf{x},\mathbf{x})\ \mathrm{d}\mathbf{x}+\sum_{\Phi}\langle((\widetilde{u}-u^{\Phi},\widetilde{u}-u^{\Phi}))^{\Phi}\rangle \quad , \tag{39}$$

where $((\bullet,\bullet))^{\Phi}$ are zero to several Sobolev scalar products, quantifying the roughness of the resulting slowness distribution. Inserting (35) and (18) with (22) into (39), we arrive at

$$\begin{aligned}
y=&\iiint D(\mathbf{x},\mathbf{x}_1)D(\mathbf{x},\mathbf{x}_2)C(\mathbf{x}_1,\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2\ \mathrm{d}\mathbf{x}\\
&-2\,R_{iI}\iint B_i(\mathbf{x}_1)D(\mathbf{x}_1,\mathbf{x}_2)\Psi_I(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2+B_{ij}R_{iI}R_{jJ}S_{IJ}\\
&+B_{ij}R_i^0R_j^0-2\,R_i^0\iint B_i(\mathbf{x}_1)D(\mathbf{x}_1,\mathbf{x}_2)u^0(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2\\
&+\iiint D(\mathbf{x},\mathbf{x}_1)D(\mathbf{x},\mathbf{x}_2)u^0(\mathbf{x}_1)u^0(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2\ \mathrm{d}\mathbf{x}\\
&+\sum_{\Phi}((B_i,B_j))^{\Phi}\langle[R_i^0-u_i^{\Phi}+R_{iI}\ \ (T_I-T_I^0)][R_j^0-u_j^{\Phi}+R_{jJ}\ \ (T_J-T_J^0)]\rangle\ ,
\end{aligned}\tag{40}$$

where

$$\boxed{B_{ij}=\int B_i(\mathbf{x})B_j(\mathbf{x})\ \mathrm{d}\mathbf{x}} \tag{41}$$

is the L_2 scalar product of the basis functions, and $((B_i,B_j))^{\Phi}$ are the Sobolev scalar products of basis functions $B_i(\mathbf{x})$ and $B_j(\mathbf{x})$.

Considering (14) and (34), equation (40) reads

$$\begin{aligned}
y=&\iiint D(\mathbf{x},\mathbf{x}_1)D(\mathbf{x},\mathbf{x}_2)C(\mathbf{x}_1,\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2\ \mathrm{d}\mathbf{x}-2\,\overline{\Psi}_{iI}B_{ij}R_{jI}\\
&+B_{ij}\left[R_{iI}R_{jJ}S_{IJ}+(R_i^0-\overline{u}_i^0)(R_j^0-\overline{u}_j^0)\right]-\overline{u}_i^0B_{ij}\overline{u}_j^0
\end{aligned}$$

$$+\sum_{\Phi}((B_i,B_j))^{\Phi}\left[R_{iI}R_{jJ}S_{IJ}+(R_i^0-u_i^{\Phi})(R_j^0-u_j^{\Phi})\right]$$
$$+\iiint D(\mathbf{x},\mathbf{x}_1)D(\mathbf{x},\mathbf{x}_2)u^0(\mathbf{x}_1)u^0(\mathbf{x}_2)\,\mathrm{d}\mathbf{x}_1\,\mathrm{d}\mathbf{x}_2\,\mathrm{d}\mathbf{x}\ , \tag{42}$$

where components $\overline{u}_i^0$ of the Lebesgue L_2 projections of the smoothed mean slowness onto the basis functions are defined by (27), and

$$\boxed{\begin{aligned}\overline{\Psi}_{iJ} &= B_{ik}^{-1}\iint B_k(\mathbf{x}_1)D(\mathbf{x}_1,\mathbf{x}_2)\Psi_J(\mathbf{x}_2)\,\mathrm{d}\mathbf{x}_1\,\mathrm{d}\mathbf{x}_2\\ &= B_{ik}^{-1}\iiint B_k(\mathbf{x}_1)\,D(\mathbf{x}_1,\mathbf{x}_1')\,C(\mathbf{x}_1',\mathbf{x}_2)\,G_J(\mathbf{x}_2)\,\mathrm{d}\mathbf{x}_1\,\mathrm{d}\mathbf{x}_1'\,\mathrm{d}\mathbf{x}_2 \qquad (43)\end{aligned}}$$

are the components of the Lebesgue L_2 projections of the smoothed ray back–projective functions (covariances between the smoothed slowness and travel times) onto the basis functions.

3.3 The minimum of the objective function with respect to the resolving kernel

Differentiating (42), under constraints (24), with respect to R_i^0, we obtain the minimum objective function for

$$B_{ij}[R_j^0-\overline{u}_j^0]+\sum_{\Phi}((B_i,B_j))^{\Phi}[R_j^0-u_j^{\Phi}]=\Lambda_i^{\alpha}\Delta^{\alpha}\quad , \tag{44}$$

where Λ_i^{α} are Lagrange multipliers corresponding to optional constraints (24). Note that the Einstein summation over repetitive indices applies also to superscripts $^{\alpha}$. The coefficients of the inverted mean value $u^0(\mathbf{x})$ thus have the form

$$R_i^0=\left[B_{ij}+\sum_{\Phi}((B_i,B_j))^{\Phi}\right]^{-1}\left[B_{jk}\overline{u}_k^0+\sum_{\Phi}((B_i,B_j))^{\Phi}u_j^{\Phi}+\Lambda_i^{\alpha}\Delta^{\alpha}\right]\quad , \tag{45}$$

where $\left[B_{ji}+\sum_{\Phi}((B_jB_i))^{\Phi}\right]^{-1}$ are the components of the matrix inverse to matrix $\left[B_{ij}+\sum_{\Phi}((B_iB_j))^{\Phi}\right]$.

Differentiating (42), under constraints (24), with respect to R_{iI}, we see that the minimum objective function is obtained for

$$\left[B_{ij}+\sum_{\Phi}((B_i,B_j))^{\Phi}\right]R_{jJ}S_{JI}-B_{ik}\overline{\Psi}_{kI}=\Lambda_i^{\alpha}T_I^{\alpha} \tag{46}$$

with the same Lagrange multipliers Λ_i^{α} as in (44). The coefficients of the resolving kernel thus have the form

$$R_{jJ}=\left[B_{ji}+\sum_{\Phi}((B_jB_i))^{\Phi}\right]^{-1}\left[B_{ik}\overline{\Psi}_{kI}+\Lambda_i^{\alpha}T_I^{\alpha}\right]S_{IJ}^{-1}\quad , \tag{47}$$

where S_{NM}^{-1} are the components of the matrix inverse to the complete travel–time covariance matrix S_{MN} defined by (34).

Without constraints (24), there are no Lagrange multipliers Λ_i^α in equations (45) and (47). The model coefficients then depend on travel times through equation (22) with

$$\boxed{R_i^0 = \left[B_{ij} + \sum_\Phi ((B_i, B_j))^\Phi \right]^{-1} \left[B_{jk}\overline{u}_k^0 + \sum_\Phi ((B_i, B_j))^\Phi u_j^\Phi \right]} \tag{48}$$

and

$$\boxed{R_{jJ} = \left[B_{ji} + \sum_\Phi ((B_j B_i))^\Phi \right]^{-1} B_{ik}\overline{\Psi}_{kI} S_{IJ}^{-1} \quad .} \tag{49}$$

In this case of no constraints (24), Section 3.4 may be skipped, continuing with Section 3.5.

3.4 Constraints adjusting the mean slowness

Let us now discuss the meaning of constraints (24). Inserting (45) and (47) into constraints (24), we obtain equations

$$\begin{aligned} &\left[B_{ij} + \sum_\Phi ((B_i, B_j))^\Phi \right]^{-1} \\ \times &\left[\left(B_{jk}\overline{u}_k^0 + \sum_\Phi ((B_i, B_j))^\Phi u_j^\Phi + \Lambda_i^\alpha \Delta^\alpha \right) \Delta^\beta + \left(B_{ik}\overline{\Psi}_{kI} + \Lambda_i^\alpha T_I^\alpha \right) S_{IJ}^{-1} T_J^\beta \right] \\ = &\Delta^\alpha \overline{u}_i^0 + \overline{u}_i^\alpha \end{aligned} \tag{50}$$

for Lagrange multipliers Λ_i^α. The Lagrange multipliers are then

$$\begin{aligned} \Lambda_i^\alpha = &\left[T_M^\alpha S_{MN}^{-1} T_N^\beta + \Delta^\alpha \Delta^\beta \right]^{-1} \\ &\times \left[B_{ij}\overline{u}_j^\beta + \sum_\Phi ((B_i, B_j))^\Phi \left[\overline{u}_j^\beta + (\overline{u}_j^0 - u_j^\Phi)\Delta^\beta \right] - B_{ij}\overline{\Psi}_{jJ} S_{JK}^{-1} T_K^\beta \right] \quad . \end{aligned} \tag{51}$$

Inserting (45) and (47) with Lagrange multipliers (51) into (22) and subtracting (22) with (48) and (49), we see that the increment in u_i due to constraints (24) is

$$\Delta u_i = \left[B_{ij}\overline{u}_j^\beta + \sum_\Phi ((B_i, B_j))^\Phi \left[\overline{u}_j^\beta + (\overline{u}_j^0 - u_j^\Phi)\Delta^\beta \right] - B_{ij}\overline{\Psi}_{jJ} S_{JK}^{-1} T_K^\beta \right] U^\beta \tag{52}$$

with

$$U^\beta = \left[T_M^\beta S_{MN}^{-1} T_N^\alpha + \Delta^\beta \Delta^\alpha \right]^{-1} \left[T_P^\alpha S_{PR}^{-1} (T_R - T_R^0) + \Delta^\alpha \right] \quad . \tag{53}$$

The same increment in u_i can be achieved by updating the mean medium to

$$u^{0\mathrm{new}}(\mathbf{x}) = u^0(\mathbf{x}) + u^\alpha(\mathbf{x}) U^\alpha \quad , \tag{54}$$

replacing functions (37) by

$$u_i^{\Phi\mathrm{new}} = u_i^\Phi + \left[\overline{u}_i^\alpha + (\overline{u}_i^0 - u_i^\Phi)\Delta^\alpha \right] U^\alpha \quad , \tag{55}$$

and then performing inversion (22) without constraints (24). Note that coefficients (53) can be obtained by minimizing objective function

$$y^{(U)} = [T_I - T_I^0 - T_I^\alpha U^\alpha] S_{IJ}^{-1} [T_J - T_J^0 - T_I^\beta U^\beta] + [U^\alpha \Delta^\alpha - 1]^2 \quad , \tag{56}$$

where the perturbed reference travel times $T_J^{0\text{new}} = T_J^0 + T_J^\alpha U^\alpha$ coincide with the reference travel times (13) calculated in the perturbed mean medium (54).

The first term of addition in objective function (56) corresponds to constraints (24) with $\Delta^\alpha = 0$, and its interpretation is straightforward. It represents the least–square fitting of given travel times T_I by perturbed mean medium (54). On the other hand, the second term in objective function (56), present if at least one Δ^α is non–zero, is much more difficult to interpret. The second term tends to keep smoothed perturbation (54) of the mean medium close to the linear subspace generated by the right–hand sides of constraints (24). It then seems natural to choose $\Delta^\alpha = 0$ when constraints (24) are applied.

The impact of the application of constraints (24) on model $\widetilde{u}(x)$ resulting from the inversion is equivalent to adjusting the a priori mean slowness $u^0(\mathbf{x})$ through a linear combination (54) of functions $u^\alpha(\mathbf{x})$, fitted by least squares according to equation (53) prior to the inversion, and then performing inversion (22) with perturbed mean slowness (54) but without constraints (24). The least–square fitting (56) should be very strongly overdetermined. We may consider this procedure as a part of the determination of the statistical moments $u^0(\mathbf{x})$ and $C(\mathbf{x},\mathbf{x}')$ of the medium and assume no constraints (24) during the travel–time inversion.

For example, imagine that we have no idea about the mean slowness and thus choose $u^0(\mathbf{x}) = 0$, but we wish the inversion to be exact in a homogeneous space. We may thus introduce one constraint (24) with $u^1(\mathbf{x}) = 1$ and $\Delta^1 = 0$. Inversion with such a constraint yields the same model as the determination of $u^0(\mathbf{x}) = \mathit{constant}$ by means of least squares (54) and its use in the inversion without constraints.

We shall not discuss the impact of constraints (24) on the model covariance matrix (35) in this paper.

3.5 Discussion

Let us now decompose equation (22) with coefficients (48) and (49) into several steps to get a better insight into the proposed travel–time tomographic inversion.

The travel–time differences with respect to the reference travel times in mean slowness are back–projected onto the model by the operator

$$\left[T_I - T_I^0\right] \to f(\bullet) = u^0(\bullet) + r_J(\bullet) \left[T_J - T_J^0\right] \quad , \tag{57}$$

where

$$r_J(\mathbf{x}) = \Psi_K(\mathbf{x}) \; S_{KJ}^{-1} \quad , \tag{58}$$

with ray back–projective functions $\Psi_K(\mathbf{x})$ given by (32). Considering either exact ray characteristic functions or their approximations from the last iteration of the linearized inversion, equation (57) is identical to equation (7.127) of Section 7.5.2 [or equation (18) of Problem 7.1] by Tarantola (1987). The back–projection by means of functions $\Psi_K(\mathbf{x})$ is not only statistically correct, but is also much smoother than the back–projection by means of the ray characteristic functions $G_K(\mathbf{x})$, especially if the ray characteristic functions are narrow in comparison with the distances between two–point rays. The application of the medium covariance functions has thus a significant smoothing effect.

The resulting back–projection of the weighted travel–time differences is then optionally smoothed with the smoothing function,

$$f(\bullet) \rightarrow \overline{f}(\bullet) = \int D(\bullet, \mathbf{x})\, f(\mathbf{x})\; \mathrm{d}\mathbf{x} \quad . \tag{59}$$

There is no reason to apply this smoothing if the ray characteristic functions $G_K(\mathbf{x})$ correctly account for the sensitivity of travel times to the medium at finite frequencies.

The optionally smoothed back–projection is then projected onto the basis functions

$$\overline{f}(\bullet) \rightarrow B_i(\bullet)\, \overline{f}_i = B_i(\bullet)\; B_{ij}^{-1} \int B_j(\mathbf{x})\, \overline{f}(\mathbf{x})\; \mathrm{d}\mathbf{x} \quad . \tag{60}$$

This projection should modify the velocity model as little as possible. The projection onto the basis functions should never substitute smoothing by means of the Sobolev scalar products.

The projection onto the basis functions is then smoothed with the optional Sobolev scalar products,

$$B_i(\bullet)\, \overline{f}_i \rightarrow \widetilde{u}(\bullet) = B_i(\bullet) \left[B_{ij} + \sum_{\Phi} ((B_i, B_j))^{\Phi} \right]^{-1} \left[B_{jk}\, \overline{f}_k + \sum_{\Phi} ((B_j, B_k))^{\Phi} u_k^{\Phi} \right] \quad . \tag{61}$$

Step (57) is independent of the a priori information contained within the smoothing function $D(\mathbf{x}_1, \mathbf{x}_2)$, model parametrization, and smoothing expressed in terms of the Sobolev scalar products $((\bullet, \bullet))^{\Phi}$. These subjective a priori parts of information are applied step–by–step a posteriori as expressed in equations (59), (60) and (61).

Equation (61) shows that the a priori additional smoothing term $\sum_{\Phi} \langle ((\widetilde{u} - u^{\Phi}, \widetilde{u} - u^{\Phi}))^{\Phi} \rangle$ to the objective function (39) acts just as a convolutional low–pass filter applied a posteriori to the result of the tomographic travel–time inversion performed without the Sobolev scalar products. It is thus possible to perform the inversion without the Sobolev scalar products. A posteriori smoothing is then at the discretion of the seismologist.

4 Special cases

4.1 Special choice of the smoothing function

Let us discuss here the very important special case of the smoothing function $D(\mathbf{x}_1, \mathbf{x}_2)$ coinciding with the integral kernel of the projection operator onto the model subspace generated by the basis functions,

$$D(\mathbf{x}_1, \mathbf{x}_2) = B_i(\mathbf{x}_1) B_{ij}^{-1} B_j(\mathbf{x}_1) \quad . \tag{62}$$

Equation (43) then simplifies to

$$\overline{\Psi}_{iJ} = B_{ij}^{-1} \iint B_j(\mathbf{x}_1) C(\mathbf{x}_1, \mathbf{x}_2) G_J(\mathbf{x}_2)\; \mathrm{d}\mathbf{x}_1\; \mathrm{d}\mathbf{x}_2 \quad . \tag{63}$$

Note that the same applies if $D(\mathbf{x}_1, \mathbf{x}_2)$ takes the form of the Dirac distribution $\delta(\mathbf{x}_1 - \mathbf{x}_2)$ or, more generally, if $D(\mathbf{x}_1, \mathbf{x}_2)$ coincides with the Dirac distribution on the model subspace.

In a self–affine random medium, the squared standard deviation $\widetilde{C}(\mathbf{x}, \mathbf{x})$ of the resulting slowness $\widetilde{u}(\mathbf{x})$ may be very large for very narrow smoothing function $D(\mathbf{x}_1, \mathbf{x}_2)$, and may go to infinity for $D(\mathbf{x}_1, \mathbf{x}_2)$ approaching the Dirac distribution $\delta(\mathbf{x}_1 - \mathbf{x}_2)$. On the other hand, for the smoothing function in the form of (62), the model covariance function (35) takes the form

$$\widetilde{C}(\mathbf{x}_1,\mathbf{x}_2) = B_i(\mathbf{x}_1)\widetilde{C}_{ij}B_j(\mathbf{x}_1) \tag{64}$$

with

$$\widetilde{C}_{ij} = C_{ij} - R_{iK}\overline{\Psi}_{jK} - \overline{\Psi}_{iK}R_{jK} + R_{iI}S_{IJ}R_{jJ} + [R_i^0 - \overline{u}_i^0][R_j^0 - \overline{u}_j^0] \quad , \tag{65}$$

where

$$C_{ij} = B_{ik}^{-1}B_{jl}^{-1}\iint B_k(\mathbf{x}_1)C(\mathbf{x}_1,\mathbf{x}_2)B_l(\mathbf{x}_2)\,\mathrm{d}\mathbf{x}_1\,\mathrm{d}\mathbf{x}_2 \tag{66}$$

are the components of the projection

$$C^{(BB)}(\mathbf{x}_1,\mathbf{x}_2) = B_i(\mathbf{x}_1)C_{ij}B_j(\mathbf{x}_2) \tag{67}$$

of the medium covariance function $C(\mathbf{x}_1,\mathbf{x}_2)$ onto the model space generated by the basis functions.

4.2 Special choice of the ray characteristic functions

Assume here that the ray characteristic functions $G_I(\mathbf{x})$ are replaced by their projections $G_I^{(B)}(\mathbf{x})$ onto the the model subspace generated by basis functions $B_i(\mathbf{x})$,

$$G_J^{(B)}(\mathbf{x}) = B_i(\mathbf{x})B_{ik}^{-1}\int B_k(\mathbf{x})G_J(\mathbf{x})\,\mathrm{d}\mathbf{x} = B_i(\mathbf{x})B_{ik}^{-1}G_{kJ} \quad , \tag{68}$$

where

$$G_{iJ} = \int B_i(\mathbf{x})G_J(\mathbf{x})\,\mathrm{d}\mathbf{x} \tag{69}$$

is the variation of the J^{th} synthetic travel time with respect to the i^{th} model slowness parameter. This simplification may be justified, e.g., if the model parametrization is so fine that the "thick rays" (68) are mostly situated within the Fresnel volumes corresponding to frequencies significant for travel–time picking.

Integral (33) may then be expressed in terms of integrals (66) and (69),

$$\Theta_{IJ} = G_{iI}C_{ij}G_{jJ} \quad , \tag{70}$$

and integration along rays in integral (43) may be replaced with the projection onto the basis functions,

$$\overline{\Psi}_{iJ} = B_{ij}^{-1}\iiint B_j(\mathbf{x}_1)D(\mathbf{x}_1,\mathbf{x}_1')C(\mathbf{x}_1',\mathbf{x}_2)B_l(\mathbf{x}_2)\,\mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_1'\mathrm{d}\mathbf{x}_2 \;\; B_{lj}^{-1}\;G_{jJ} \quad . \tag{71}$$

Note that (70) and (71) are also valid for general $G_{iI}(\mathbf{x})$ if $C(\mathbf{x}_1,\mathbf{x}_2)$ is replaced by $C^{(BB)}(\mathbf{x}_1,\mathbf{x}_2)$, see (67).

4.3 Special choice of the smoothing function and ray characteristic functions

In the special case of both options (62) and (68), equation (70) may be used to calculate integral (33). In addition, integral (43) takes the simple form

$$\overline{\Psi}_{iJ} = C_{ij}G_{jJ} \quad , \tag{72}$$

compared with (63) and (71). In consequence, equation (65) simplifies to

$$\widetilde{C}_{ij} = [\delta_{ik} - R_{iI}G_{kI}]\,C_{kl}\,[\delta_{jl} - R_{jJ}G_{lJ}] + R_{iI}T_{IJ}R_{jJ} + [R_i^0 - \overline{u}_i^0][R_j^0 - \overline{u}_j^0] \quad . \tag{73}$$

5 Some consequences

5.1 Residual travel–time covariance matrix

The covariance matrix of the travel–time residuals is

$$\widetilde{S}_{IJ} = \langle (\widetilde{T}_I - T_I)(\widetilde{T}_J - T_J) \rangle \quad , \tag{74}$$

where

$$\widetilde{T}_I = \int G_I(\mathbf{x}) \widetilde{u}(\mathbf{x}) \, \mathrm{d}\mathbf{x} \tag{75}$$

and T_I are given by (5). Inserting (5) and (75), equation (74) reads

$$\widetilde{S}_{IJ} = \iint G_I(\mathbf{x}_1) G_J(\mathbf{x}_1) \langle [\widetilde{u}(\mathbf{x}_1) - u(\mathbf{x}_1)][\widetilde{u}(\mathbf{x}_2) - u(\mathbf{x}_2)] \rangle \, \mathrm{d}\mathbf{x}_1 \, \mathrm{d}\mathbf{x}_2$$

$$- \int G_I(\mathbf{x}) \langle \widetilde{u}(\mathbf{x}) \delta T_J \rangle \, \mathrm{d}\mathbf{x} - \int \langle \delta T_I \widetilde{u}(\mathbf{x}) \rangle G_J(\mathbf{x}) \, \mathrm{d}\mathbf{x} \; + \langle \delta T_I \; \delta T_J \rangle \quad . \tag{76}$$

Here $\langle [\widetilde{u}(\mathbf{x}_1) - u(\mathbf{x}_1)][\widetilde{u}(\mathbf{x}_2) - u(\mathbf{x}_2)] \rangle$ coincides with (28) for $\overline{u}(\mathbf{x}) = u(\mathbf{x})$, i.e. for smoothing function $D(\mathbf{x}_1, \mathbf{x}_2)$ in the form of the Dirac distribution $\delta(\mathbf{x}_1 - \mathbf{x}_2)$ (no smoothing). Inserting (35) with $D(\mathbf{x}_1, \mathbf{x}_2) = \delta(\mathbf{x}_1 - \mathbf{x}_2)$ for $\langle [\widetilde{u}(\mathbf{x}_1) - u(\mathbf{x}_1)][\widetilde{u}(\mathbf{x}_2) - u(\mathbf{x}_2)] \rangle$, and

$$\langle \widetilde{u}(\mathbf{x}) \delta T_J \rangle = B_i(\mathbf{x}) R_{iK} \langle \delta T_K \; \delta T_J \rangle = B_i(\mathbf{x}) R_{iK} T_{KJ} \tag{77}$$

into (76), we arrive at

$$\widetilde{S}_{IJ} = \Theta_{IJ} - G_{iI} R_{iK} \Theta_{KJ} - \Theta_{IL} R_{jL} G_{jJ} + G_{iI} R_{iK} S_{KL} R_{jL} G_{jJ}$$

$$+ [R_i^0 G_{iI} - T_I^0][R_j^0 G_{jJ} - T_J^0] - G_{iI} R_{iK} T_{KJ} - T_{IK} R_{jK} G_{jJ} + T_{IJ} \; . \tag{78}$$

Thus the statistical expectation of the travel–time residuals is described by the covariance matrix

$$\widetilde{S}_{IJ} = [\delta_{IK} - G_{iI} R_{iK}] \, S_{KL} \, [\delta_{LJ} - R_{jL} G_{jJ}] + [R_i^0 G_{iI} - T_I^0][R_j^0 G_{jJ} - T_J^0] \; , \tag{79}$$

where S_{KL} is given by (34).

6 Conclusions

In this paper, we have discussed the approach to seismic travel–time tomography enabling the accuracy and resolution of the resulting model of a geological structure (macro model) to be estimated. The medium covariance functions yield correct travel–time weights for uneven ray coverage of the medium and provide statistically correct interpolation of slowness between two–point rays.

The presented approach already enables us to consider a priori information such as the obviously much greater probability of slow materials close to the earth surface than deeper in the earth, or greater probability of slow materials in the low and concave parts of the earth periphery than in the high–elevated and convex parts.

On the other hand, we have still made various simplifications. We have considered only smooth models without structural interfaces, and have restricted the study to a single material parameter — slowness. However, we believe that this approach to seismic travel–time tomography may further be generalized.

Acknowledgements

The author is indebted to Sergei Shapiro for fruitful discussions essential for this study.

This research has been supported by the Grant Agency of the Czech Republic under Contracts 205/95/1465 and 205/01/0927, by the Grant Agency of the Charles University under Contract 237/2001/B-GEO/MFF, by the Ministry of Education of the Czech Republic within Research Project J13/98 113200004, and by the members of the consortium "Seismic Waves in Complex 3-D Structures" (see "http://seis.karlov.mff.cuni.cz /consort/main.htm").

References

Backus, G.E. & Gilbert, F. (1970): Uniqueness in the inversion of inaccurate gross Earth data. Phil. Trans. R. Soc. London, **266A**, 123–192.

Bulant, P. (1996): Two–point ray tracing in 3-D. Pure appl. Geophys., **148**, 421–447.

Bulant, P. (1999): Two–point ray tracing and controlled initial–value ray tracing in 3-D heterogeneous block structures. J. seism. Explor., **8**, 57–75.

Bulant, P. (2002): Sobolev scalar products in the construction of velocity models — application to model Hess and to SEG/EAGE Salt Model. Pure appl. Geophys., **159**, in press.

Červený, V. (2001): Seismic Ray Theory. Cambridge Univ. Press, Cambridge.

Chernov, L.A. (1960): Wave Propagation in a Random Medium. McGraw-Hill, New York.

Coates, R.T. & Chapman, C.H. (1990): Ray perturbation theory and the Born approximation. Geophys. J. int., **100**, 379–392.

Farra, V. (1999): Computation of 2nd–order travel–time perturbation by Hamiltonian ray theory. Geophys. J. int., **136**, 205–217.

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

Klimeš, L. (2000): Sobolev scalar products in the construction of velocity models. In: Seismic Waves in Complex 3-D Structures, Report 10, pp. 15–40, Dep. Geophys. Charles Univ., Prague, online at "http://seis.karlov.mff.cuni.cz/consort/main.htm".

Klimeš, L. (2002a): Lyapunov exponents for 2-D ray tracing without interfaces. Pure appl. Geophys., **159**, in press.

Klimeš, L. (2002b): Correlation functions of random media. Pure appl. Geophys., **159**, in press.

Klimeš, L. (2002c): Estimating the correlation function of a self–affine random medium. Pure appl. Geophys., **159**, in press.

Marquering, H., Nolet, G & Dahlen, F.A. (1998): Three–dimensional waveform sensitivity kernels. Geophys. J. int., **132**, 521–534.

Maurer, H, Holliger, K. & Boerner, D.E. (1998): Stochastic regularization: Smoothness or similarity? Geophys. Res. Lett., **25**, 2889–2892.

Menke, W. (1984): Geophysical Data Analysis: Discrete Inverse Theory. Academic Press, Orlando.

Pilkington, M. & Todoeschuck, J.P. (1990): Stochastic inversion for scaling geology. Geophys. J. int., **102**, 205–217.

Pratt, R.G., Song, Z.-M., Williamson, P. & Warner, M. (1996): Two–dimensional velocity models from wide–angle seismic data by wavefield inversion. Geophys. J. int., **124**, 323–340.

Tarantola, A. (1987): Inverse Problem Theory. Elsevier, Amsterdam.

Tarantola, A. & Nercessian, A. (1984): Three–dimensional inversion without blocks. Geophys. J. R. astr. Soc., **76**, 299–306.

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

Woodward, M.J. (1992): Wave–equation tomography. Geophysics, **57**, 15–26.

Woodward, M.J. & Rocca, F. (1988): Wave–equation tomography. Expanded Abstracts of 58th Annual Meeting (Anaheim), pp. 1232–1235, Soc. Explor. Geophysicists, Tulsa.

Žáček, K. (2002): Smoothing the Marmousi model. Pure appl. Geophys., **159**, in press.