S©MR

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ
Siberian Electronic Mathematical Reports
http://semr.math.nsc.ru
ISSN 1813-3304

Vol. ??, No. ?, pp ???–??? (2025)
https://doi.org/10.33048/semi.2025.20.???
УДК 550.34.013.4
MSC 86-08

CONTROLLED DIRECTIONAL RECEPTION TOMOGRAPHY BASED ON THE RAY METHOD ASYMPTOTICS OF THE DOUBLE SQUARE ROOT EQUATION

N.N. SHILOV

Communicated by P.P. Petrov

Abstract: Controlled Directional Reception (CDR) is a reflection tomography technique that accepts seismic traveltimes and slopes (traveltime derivatives w.r.t source and receiver coordinates) and returns a velocity model fitting this data. In contrast to other slope-based methods, it uses parsimonious model parametrization and relies on ray tracing thus being computationally efficient and fairly general. However, it is unstable w.r.t. data errors. In this paper we revisit the CDR method and develop a formalism to mitigate its instability. Our approach is based on linearized estimates of ray tracing errors allowing for suboptimal regularization of the inverse problem. We apply our original ray method asymptotics of the pseudodifferential Double Square Root equation to parametrize the wavefield and test our formulation of the CDR method on two benchmark synthetic datasets. We demonstrate that it provides competitive results suitable for depth migration. We restrict ourselves to 2D settings although the approach can be generalized to 3D problems as well.

Shilov, N.N., CDR tomography by the DSR equation asymptotics.

The work is supported by the Ministry of Science and Higher Education of the Russian Federation (grant FSUS-2025-0015).
Received ???, ?, 2025, Published ???, ??, 2025.

Keywords: slope tomography, reflection tomography, Controlled Directional Reception, Double Square Root equation.

1 Introduction

Slope tomography is a distinct group of reflection tomography techniques. They accept a set of so-called events (vectors containing source and receiver coordinates, reflection traveltimes and slopes i.e. traveltime derivatives w.r.t. source and receiver coordinates) and adjust the velocity model to fit this data. The resulting model is suitable for depth migration or full waveform inversion [1]. The earliest slope tomography methods date back to 1950s and were based on straight-rays considerations [2]. In 1987 Charles Sword reformulated the original approach for heterogeneous media with curved rays, dubbing his formulation the Controlled Directional Reception (CDR) method [3]. However, it lacked stability w.r.t. data errors, and in 1998 Frederic Billette et al. introduced more stable and general method of stereotomography [4], which became a classic reference for coming algorithms [5, 6]. Nowadays slope tomography is widely adopted and developed for both 2D and 3D data, isotropic and anisotropic models [7, 8, 9]. Notable industrial examples are Slb's CIP tomography [10] and Viridien's slope tomography [11], and a number of academic studies have been carried out (see [1, 12] for overview).

Each event in seismic data corresponds to a reflecting facet (a fragment of an interface specified by its coordinates, orientation, and the reflection angle) in the subsurface. In classic stereotomography one iteratively traces incident and reflected rays upward from trial facets and then updates both the velocity and the facets to minimize data mismatch [1]. In contrast, the CDR method deals with the velocity model only. Instead of tracing rays upward from unknown reflectors, one traces them downward from fixed sources and receivers, passing the observed slopes as the initial conditions for ray tracing. The tracing is stopped as soon as the total traveltime along the incident and the reflected rays equals the observed one. If the rays do not meet each other as the tracing stops, the velocity model is updated until they converge. The reflecting facets are determined then automatically [3]. On the one hand, this concept benefits from lower dimensionality of the inverse problem. On the other hand, it suffers from instability of one-point ray tracing [1]. However, this drawback formally can be attributed to some more recent slope tomography algorithms [5, 6] as well. Hence, the CDR method still can be a good starting point for new slope tomography algorithms.

Double Square Root (DSR) equation is a one-way wave equation describing singly scattered wavefield as a function of source and receiver coordinates. It was introduced in [13] although its kinematics had been described earlier [14]. It acts on the recorded wavefield and allows one to extrapolate the observed data to another depth level, performing so-called "survey-sinking"

migration [13, 15]. In this migration, one evaluates the extrapolated wavefield at all depths at zero time and extracts zero offset traces (i.e. with coinciding source and receiver) only to obtain the subsurface image. Significant energy at nonzero offsets provides a basis for velocity-building techniques known as migration velocity analysis (MVA) [16, 17]. The "survey-sinking" migration and corresponding MVA much resemble the CDR method where one "sinks" the source and the receiver and updates the velocity model until the rays converge (i.e. the energy focuses at zero offsets). This similarity has been employed in [18] to reformulate the Sword's approach using a ray method asymptotics of the DSR equation [19]. However, the authors limited themselves to noise-free data inversion. In this paper we consider more realistic cases and stress on the method's regularization.

The paper is organized as follows. First, we briefly outline theoretical aspects of the proposed formalism and build up a set of tools for iterative velocity update following the CDR approach. Then we linearize the inverse problem in simplified settings to derive suboptimal estimates of regularization parameters which we use later in the original nonlinear problem. Finally, we provide some technical information on our implementation and proceed with numerical tests. In the Discussion and Conclusion section, we weigh our results versus alternative approaches and summarize the paper's contents. We bring more technical features in three Appendices in the end of the paper.

2 Method's theory

2.1. DSR equation rays. Let us consider a smooth 2D velocity distribution $v(x,z)$ with z axis being the depth, i.e. pointed downward. Let $\tau(x_s,x_r,z)$ be the traveltime of a reflected wave emitted from a source $(x_s,z)^T$ and recorded at a receiver $(x_r,z)^T$. Superscript T denotes transposition. Traveltime τ is described by an equation [13, 14]

$$\frac{\partial\tau}{\partial z}=-\sqrt{\frac{1}{v_s^2}-\left(\frac{\partial\tau}{\partial x_s}\right)^2}-\sqrt{\frac{1}{v_r^2}-\left(\frac{\partial\tau}{\partial x_r}\right)^2}, \tag{1}$$

where $v_s=v\,(x_s,z)$ and $v_r=v\,(x_r,z)$ are the velocities at the source and receiver points respectively. We assume that the expressions under square roots keep positive, which implies nowhere-horizontal wave propagation. We shall briefly touch this restriction in the Discussion.

By solving (1) one can reconstruct the surface $\tau\,(x_s,x_r,z)=0$ which defines all possible source-receiver pairs where the reflected wave is recorded just as the source fires. We shall search for such pairs using ray method [19, 20]. We define a phase space consisting of 6 coordinates:

$$\vec{x}=(x_s,x_r,z)^T,\quad \vec{p}=(p_s,p_r,p_z)^T=\left(\frac{\partial\tau}{\partial x_s},\frac{\partial\tau}{\partial x_r},\frac{\partial\tau}{\partial z}\right)^T. \tag{2}$$

Vector $\vec{x}(\tau)$ defines a source $(x_s, z)^T$ and a receiver $(x_r, z)^T$ in physical space. Thus, each point of the phase space describes a data configuration: a reflected wave will be recorded at the time τ if one places a source-receiver pair at $\vec{x}(\tau)$ while the traveltime derivatives will be described by the vector $\vec{p}(\tau)$.

To trace the rays we introduce a Hamiltonian [19]:

$$H(\vec{x},\vec{p}) = -\frac{p_z + \sqrt{\frac{1}{v_s^2} - p_s^2} + \sqrt{\frac{1}{v_r^2} - p_r^2}}{\frac{1}{v_s^2\sqrt{\frac{1}{v_s^2} - p_s^2}} + \frac{1}{v_r^2\sqrt{\frac{1}{v_r^2} - p_r^2}}} = \\ = -\left(p_z + \sqrt{\frac{1}{v_s^2} - p_s^2} + \sqrt{\frac{1}{v_r^2} - p_r^2}\right) C_\tau(\vec{x},\vec{p}). \tag{3}$$

The DSR equation ray tracing system reads

$$\begin{cases} \dfrac{d\vec{x}}{d\tau} = \nabla_{\vec{p}} H, \\ \dfrac{d\vec{p}}{d\tau} = -\nabla_{\vec{x}} H, \end{cases} \tag{4}$$

where ∇ is the gradient operator and its subscript indicates the variables of differentiation. We impose the initial conditions for this system at the Earth' upper surface $\hat{z} = 0$:

$$\begin{cases} \vec{x}|_{\tau=\hat{\tau}} = (\hat{x}_s, \hat{x}_r, \hat{z} = 0)^T = \widehat{\vec{x}}, \\ \vec{p}|_{\tau=\hat{\tau}} = \left(\hat{p}_s, \hat{p}_r, \hat{p}_z = -\sqrt{\frac{1}{\hat{v}_s^2} - \hat{p}_s^2} - \sqrt{\frac{1}{\hat{v}_r^2} - \hat{p}_r^2}\right)^T = \widehat{\vec{p}}, \end{cases} \tag{5}$$

with $\hat{v}_s = v(\hat{x}_s, \hat{z} = 0)$ and $\hat{v}_r = v(\hat{x}_r, \hat{z} = 0)$. The "capped" variables represent conventional slope tomography data including actual source and receiver coordinates $\hat{x}_s$ and $\hat{x}_r$, observed traveltime $\hat{\tau}$ and its derivatives $\hat{p}_s = \dfrac{\partial\hat{\tau}}{\partial\hat{x}_s}$ and $\hat{p}_r = \dfrac{\partial\hat{\tau}}{\partial\hat{x}_r}$.

System (4) with conditions (5) define a DSR ray, i.e. a trajectory $(\vec{x}(\tau), \vec{p}(\tau))$. By tracing it backwards in time one can reconstruct a source-receiver pair $\vec{x}(\tau = 0)$ where the reflected wave is recorded just as the source fires. We introduce the subsurface midpoint and half-offset to describe this pair:

$$m = \left.\frac{x_r + x_s}{2}\right|_{\tau=0}, \quad h = \left.\frac{x_r - x_s}{2}\right|_{\tau=0}. \tag{6}$$

A physical wave cannot instantly get across a finite distance between the source and the receiver. Thus, assuming error-free data and accurate velocity model one can expect h to be zero and m to be the reflection point's horizontal coordinate.

2.2. DSR equation ray perturbations. Let us assume that the velocity model is inaccurate so that $h \neq 0$. We are going to build a framework for iterative velocity update aiming to reduce h to zero. To do so, we parametrize the model by a parameters' matrix $\mathbf{V}$:

$$v = v\left(x, z; \mathbf{V}\right), \quad \mathbf{V} = \begin{pmatrix} v_{11} & \cdots & v_{1N} \\ \vdots & \ddots & \vdots \\ v_{M1} & \cdots & v_{MN} \end{pmatrix}. \tag{7}$$

We assume that $v\left(x, z; \mathbf{V}\right)$ is twice smooth w.r.t. spatial coordinates and linear w.r.t. $\mathbf{V}$. By differentiating (4) and (5) w.r.t. a parameter v_{mn} we obtain:

$$\begin{cases} \dfrac{d}{d\tau}\dfrac{\partial \vec{x}}{\partial v_{mn}} = \nabla_{\vec{p}}\nabla_{\vec{x}}H\dfrac{\partial \vec{x}}{\partial v_{mn}} + \nabla_{\vec{p}}\nabla_{\vec{p}}H\dfrac{\partial \vec{p}}{\partial v_{mn}} + \nabla_{\vec{p}}\dfrac{\partial H}{\partial v_{mn}}, \\ \dfrac{d}{d\tau}\dfrac{\partial \vec{p}}{\partial v_{mn}} = -\nabla_{\vec{x}}\nabla_{\vec{x}}H\dfrac{\partial \vec{x}}{\partial v_{mn}} - \nabla_{\vec{x}}\nabla_{\vec{p}}H\dfrac{\partial \vec{p}}{\partial v_{mn}} - \nabla_{\vec{x}}\dfrac{\partial H}{\partial v_{mn}}, \end{cases} \tag{8}$$

and

$$\begin{cases} \left.\dfrac{\partial \vec{x}}{\partial v_{mn}}\right|_{\tau=\hat{\tau}} = \vec{0}, \\ \left.\dfrac{\partial \vec{p}}{\partial v_{mn}}\right|_{\tau=\hat{\tau}} = \begin{pmatrix} 0 \\ 0 \\ \dfrac{\dfrac{\partial \hat{v}_s}{\partial v_{mn}}}{\hat{v}_s^3\sqrt{\dfrac{1}{\hat{v}_s^2} - \hat{p}_s^2}} + \dfrac{\dfrac{\partial \hat{v}_r}{\partial v_{mn}}}{\hat{v}_r^3\sqrt{\dfrac{1}{\hat{v}_r^2} - \hat{p}_r^2}} \end{pmatrix}, \end{cases} \tag{9}$$

where the symbol $\vec{0}$ stands for a vector of zeros. We refer the reader to [18, 19] for the Hamiltonian derivatives from (4) and (8) and discuss the model parametrization in the Practical implementation section.

The solution of (8) determines the ray's partial derivatives w.r.t. velocity perturbations caused by the parameter v_{mn}. By solving it for all the elements of $\mathbf{V}$, one can obtain the gradient of the subsurface half-offset (6) w.r.t. velocity parameters:

$$\nabla_{\mathbf{V}} h = \left.\frac{\nabla_{\mathbf{V}} x_r - \nabla_{\mathbf{V}} x_s}{2}\right|_{\tau=0} = \left(\frac{\partial h}{\partial v_{11}}, \frac{\partial h}{\partial v_{12}}, \ldots, \frac{\partial h}{\partial v_{MN}}\right)^T \tag{10}$$

and perform an iterative velocity model update aiming to minimize the absolute value of h.

3 Inverse problem statement

3.1. Linearized CDR method for a single ray in homogeneous medium. Let us restrict ourselves to the class of homogeneous velocity models parametrized by a single parameter $\hat{v}$:

$$v\left(x, z; \mathbf{V}\right) \equiv \hat{v} = const., \quad \mathbf{V} = \left(\hat{v}\right). \tag{11}$$

In such media the subsurface half-offset (6) reads:

$$\hat{h} = \frac{\hat{x}_r - \hat{x}_s}{2} - \left(\frac{\hat{p}_r}{\sqrt{\frac{1}{\hat{v}^2} - \hat{p}_r^2}} - \frac{\hat{p}_s}{\sqrt{\frac{1}{\hat{v}^2} - \hat{p}_s^2}} \right) \frac{C_\tau \left(\widehat{\vec{x}}, \widehat{\vec{p}} \right) \hat{\tau}}{2}, \tag{12}$$

with $C_\tau \left(\widehat{\vec{x}}, \widehat{\vec{p}} \right)$ defined in (3) and "capped" variables in (5). We introduced $\hat{h}$ to distinguish it from the general case and emphasize its dependence on $\hat{v}$.

The CDR method consists of zeroing out $\hat{h}$ by appropriate choice of $\hat{v}$. However, errors in the observed data will cause a perturbation of the subsurface half-offset $\delta\hat{h}$, making it nonzero in the precise velocity model. Thus, we propose regularizing the inverse problem to avoid overfitting. Let us search for a small perturbation $\delta\hat{v}$ minimizing a linearized and regularized loss functional:

$$\hat{L}\left(\delta\hat{v}\right) = \left(\hat{h} + \delta\hat{h} + \frac{\partial \hat{h}}{\partial \hat{v}} \delta\hat{v} \right)^2 + \hat{\alpha} \frac{\delta\hat{v}^2}{\hat{v}^2} \to \min_{\delta\hat{v}}, \tag{13}$$

where $\hat{\alpha}$ is the Tikhonov's regularization parameter [21]. Given $\delta\hat{h}$, its optimal value can be found analytically:

$$\hat{\alpha} = \sup \left| \delta\hat{h} \right| \hat{v} \frac{\partial \hat{h}}{\partial \hat{v}}. \tag{14}$$

To estimate $\sup \left| \delta\hat{h} \right|$, we suggest considering the underlying data errors as independent normally distributed random variables with zero expectations and known variances:

$$\delta\hat{\tau} \sim \mathcal{N}\left(0, \hat{\sigma}_\tau^2\right), \quad \delta\hat{p}_s \sim \mathcal{N}\left(0, \hat{\sigma}_s^2\right), \quad \delta\hat{p}_r \sim \mathcal{N}\left(0, \hat{\sigma}_r^2\right). \tag{15}$$

Assuming small data errors, $\delta\hat{h}$ also can be estimated as a normal variable:

$$\delta\hat{h} \sim \mathcal{N}\left(0, \hat{\sigma}_h^2\right), \quad \hat{\sigma}_h^2 = \left(\frac{\partial \hat{h}}{\partial \hat{\tau}} \hat{\sigma}_\tau \right)^2 + \left(\frac{\partial \hat{h}}{\partial \hat{p}_s} \hat{\sigma}_s \right)^2 + \left(\frac{\partial \hat{h}}{\partial \hat{p}_r} \hat{\sigma}_r \right)^2. \tag{16}$$

Finally, we apply the "three sigma rule" to estimate the optimal $\hat{\alpha}$ value:

$$\hat{\alpha} \approx 3\hat{\sigma}_h \hat{v} \frac{\partial \hat{h}}{\partial \hat{v}}. \tag{17}$$

We conclude that given a single event in homogeneous medium one can estimate the subsurface offset's variance (16) and the optimal regularization parameter (17). We suggest considering the homogeneity restriction (11) formally and interpret $\hat{v}$ just as a realistic velocity estimate. We expect (16) and (17) to be accurate at least in order of magnitude in heterogeneous models. We provide more details on the errors' distribution in the Numerical test section and list the explicit formulae for the $\hat{h}$'s derivatives in the Appendix A.

3.2. Nonlinear CDR method for multiple rays in heterogeneous media. Let us return to the general case of heterogeneous velocity models parametrized by a matrix $\mathbf{V}$ (7). We shall build an iterative algorithm of velocity updates and denote the parameters' matrix at i-th iteration by $\mathbf{V}^{(i)}$. In particular, $\mathbf{V}^{(0)}$ will denote the initial guess matrix. Unlike the previous section, we shall deal with K events, each defining a DSR ray. We shall mark all the quantities related to k-th event with a corresponding sub- or superscript. For instance, h_k will denote the k-th subsurface half-offset.

We suggest minimizing weighted and regularized loss functional:

$$L(\mathbf{V}) = \sum_{k=1}^{K} w_k^2 h_k^2 + \alpha \frac{\| v\left(x,z;\mathbf{V}-\mathbf{V}^{(0)}\right)\|^2}{\| v\left(x,z;\mathbf{V}^{(0)}\right)\|^2} \to \min_{\mathbf{V}} \tag{18}$$

with the weights w_k and the regularization parameter α chosen so that the k-sum is dominated by high-accuracy events while the regularizing term prohibits excessively complex models.

We are going to use estimates (16) and (17) in their "formal" interpretation. For each ray we prescribe data errors (15) and define $\hat{v}_k$ as the average of the source and receiver initial velocities:

$$\hat{v}_k = \frac{v\left(\hat{x}_s^k, \hat{z}=0; \mathbf{V}^{(0)}\right) + v\left(\hat{x}_r^k, \hat{z}=0; \mathbf{V}^{(0)}\right)}{2}, \tag{19}$$

Then we proceed with the following multipliers for (18):

$$w_k = \frac{\dfrac{1}{\hat{\sigma}_h^k}}{\sqrt{\sum_{k'=1}^{K}\left(\dfrac{1}{\hat{\sigma}_h^{k'}}\right)^2}}, \quad \alpha = \sum_{k'=1}^{K} w_{k'}^2 \hat{\alpha}_{k'}. \tag{20}$$

Essentially, we divided each h_k by its expected variance $\hat{\sigma}_h^k$ (16) and formed up α by individual rays' contributions $\hat{\alpha}_k$ (17). Sub- and superscripts k and k' refer to the ray's number, not power. Such weights w_k favor high-accuracy rays while α decreases with data errors. We expect the loss functional (18) to be well-behaved, allowing one to minimize it by gradient-based methods. We emphasize that w_k and α are estimated prior to inversion.

4 Practical implementation

4.1. Model parametrization. We required the model $v(x,z;\mathbf{V})$ (7) to be twice smooth w.r.t. x and z and linear w.r.t. $\mathbf{V}$. We shall use gridded velocity models supplied with bicubic interpolation [22]. We introduce a fixed inversion grid and interpret $\mathbf{V}$ as a matrix of grid velocities

$$\vec{X} = (x_1, \ldots, x_M)^T, \quad \vec{Z} = (z_1, \ldots, z_N)^T, \quad v_{mn} = v(x_m, z_n; \mathbf{V}). \tag{21}$$

In this paper we assume that the grid step is constant:

$$\Delta x = x_{m+1} - x_m = const., \quad \Delta z = z_{n+1} - z_n = const. \tag{22}$$

A bicubic spline at $x \in [x_m, x_{m+1}]$, $z \in [z_n, z_{n+1}]$ is given by

$$v\left(x, z; \mathbf{V}\right)|_{\substack{x\in[x_m,x_{m+1}]\\ z\in[z_n,z_{n+1}]}} = \Delta\vec{X}_m^T \mathbf{C}_{mn} \Delta\vec{Z}_n \tag{23}$$

with

$$\Delta\vec{X}_m = \begin{pmatrix} x - x_m \\ x_{m+1} - x \\ (x - x_m)^3 \\ (x_{m+1} - x)^3 \end{pmatrix}, \quad \Delta\vec{Z}_n = \begin{pmatrix} z - z_n \\ z_{n+1} - z \\ (z - z_n)^3 \\ (z_{n+1} - z)^3 \end{pmatrix}, \tag{24}$$
$$\mathbf{C}_{mn} = \begin{pmatrix} c_{mn}^{11} & \cdots & c_{mn}^{14} \\ \vdots & \ddots & \vdots \\ c_{mn}^{41} & \cdots & c_{mn}^{44} \end{pmatrix}.$$

Coefficient matrices $\mathbf{C}_{mn}$ linearly depend on the grid velocities $\mathbf{V}$, see the Appendix B for details.

Bicubic splines belong to the Sobolev space of twice differentiable functions with a norm defined as an integral of their squared values and derivatives. We propose a finite-sum variant of this norm:

$$\begin{aligned} \|v(x, z; \mathbf{V})\|^2 = \sum_{m=1}^{M}\sum_{n=1}^{N} v_{mn}^2 + \left(\left.\frac{\partial v}{\partial x}\right|_{\substack{x=x_m\\z=z_n}} \Delta x\right)^2 + \\ + \left(\left.\frac{\partial v}{\partial z}\right|_{\substack{x=x_m\\z=z_n}} \Delta z\right)^2 + \left(\left.\frac{\partial^2 v}{\partial x^2}\right|_{\substack{x=x_m\\z=z_n}} \Delta x^2\right)^2 + \\ + \left(\left.\frac{\partial^2 v}{\partial z^2}\right|_{\substack{x=x_m\\z=z_n}} \Delta z^2\right)^2 + 2\left(\left.\frac{\partial^2 v}{\partial x \partial z}\right|_{\substack{x=x_m\\z=z_n}} \Delta x \Delta z\right)^2. \end{aligned} \tag{25}$$

When inserted in (18), this norm will penalize large deviations from the initial model, fast growth and fast oscillations in both directions, like the "damping" and "smoothing" terms in other slope tomography algorithms [3, 4, 23]. In homogeneous models this norm simplifies to the squared velocity used in (13) which partially justifies the use of (16) and (17) in heterogeneous media.

4.2. Gauss-Newton minimization. Given a set of events and an initial model, we shall use the Gauss-Newton method to minimize the loss functional (18), following [3, 4]. In this method, the problem is linearized at each iteration, and a perturbation $\delta\vec{V} = (\delta v_{11}, \delta v_{12}, \ldots, \delta v_{MN})^T$ is searched as the best-fit solution for a system of linear equations:

$$\mathbf{G}\delta\vec{V} = -\vec{l}, \tag{26}$$

where $\vec{l}$ consists of the loss terms and $\mathbf{G}$ consists of their gradients w.r.t. velocity parameters, see the Appendix C for explicit expressions. System (26) is solved in the least-squares sense:

$$\delta\vec{V} = -\left(\mathbf{G}^T\mathbf{G}\right)^{-1}\mathbf{G}^T\vec{l} \tag{27}$$

and the model is updated in the estimated direction. However, due to nonlinearity of the original problem, a line search step is required. Assuming i-th iteration, matrix $\mathbf{V}$ is updated as follows:

$$\mathbf{V}^{(i+1)} = \mathbf{V}^{(i)} + \varepsilon \cdot \delta\mathbf{V},$$
$$\varepsilon = \underset{0\le\varepsilon\le 1}{\operatorname{argmin}}\, L\left(\mathbf{V}^{(i)} + \varepsilon \cdot \delta\mathbf{V}\right), \quad \delta\mathbf{V} = \begin{pmatrix} \delta v_{11} & \cdots & \delta v_{1N} \\ \vdots & \ddots & \vdots \\ \delta v_{M1} & \cdots & \delta v_{MN} \end{pmatrix}. \tag{28}$$

The iterations are halted as soon as the loss stops decreasing.

4.3. Data picking and selection. We shall use the slant stacks to pick the events in seismic data, following [7] with some minor modifications. This method scans the data and returns events $\vec{d}_k$ specified by source and receiver coordinates, traveltimes, their derivatives, so-called semblances (events' coherency [24]) and the numbers of traces falling in the analysis window (the folds):

$$\vec{d}_k = \left(\hat{x}_s^k, \hat{x}_r^k, \hat{\tau}_k, \hat{p}_s^k, \hat{p}_r^k, \hat{s}_k, \hat{Q}_k\right), \quad \hat{s}_k \in [0,1], \quad \hat{Q}_k > 1. \tag{29}$$

While the first five components form up the initial conditions for the ray tracing system (5), the last two bring information on the picking error. We expect the events with higher semblances and folds to be more accurate, see the next section for details.

As highlighted in [1], it is important to sort out some events during the inversion due to non-uniform ray coverage of the model. In our parametrization the model's variability is controlled by the grid spacing (22). We suggest an ad-hoc selection algorithm:

(1) Trace all the DSR equation rays downward and estimate the subsurface midpoints' positions (6);
(2) Select at most 10 events whose subsurface midpoints fall into each model cell, prioritizing higher weight (20) events;
(3) Proceed by the Gauss-Newton scheme using the selected events only;
(4) Repeat the selection at the beginning of each inversion iteration.

The factor 10 at the second step is heuristic and implies that the number of processed events should be greater than that of velocity parameters. We emphasize that the proposed selection is performed at each iteration, leading to possible instability of the loss-based stopping criterion.

4.4. Programming issues. We implemented our algorithm using NumPy [25], SciPy [26] and Numba [27] libraries of the Python language. We employed FteikPy [28] library to compute traveltimes for the Kirchhoff migration and Matplotlib [29] library to plot the results, see the next section for details.

5 Numerical Tests

5.1. Tests' layout. To test our formulation of the CDR method we designed a hierarchy of velocity models revolving around the Marmousi dataset [30]

which is a classic benchmark for reflection seismology and slope tomography in particular [5, 6, 7]. The models' hierarchy is displayed on the Fig. 1. It contains four models: the original Marmousi model, its smoothed version (Marmousoft), its best linear fit (Marmoulin) and a step-like approximation of the latter one (Marmouflat). The first model is publicly available [31], the second one was proposed in [7] and kindly provided to us by Gilles Lambaré, and the last two models were created by ourselves. We shall test our method on the Marmousi and Marmousoft models, passing the Marmoulin model as the initial guess, and using the Marmouflat model to estimate the picking errors.

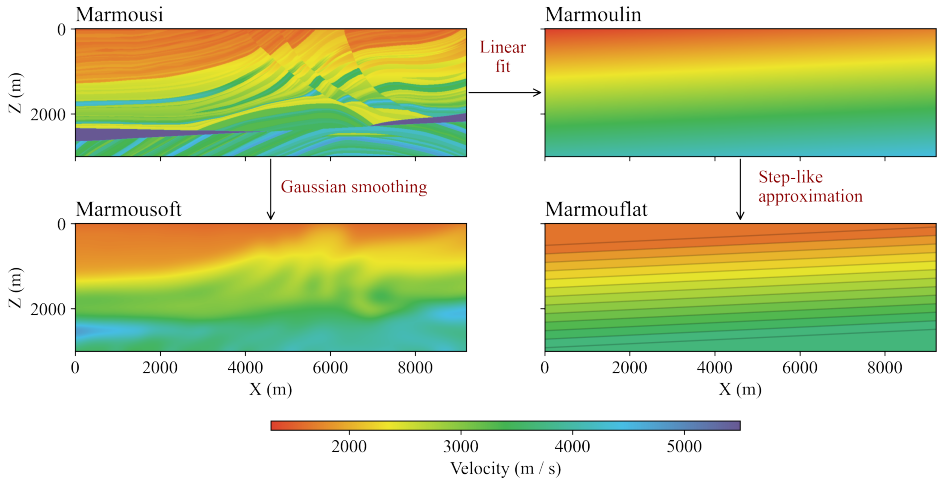

FIG. 1. Models' hierarchy

We borrowed the survey geometry and time sampling from [7]. The survey line consisted of 261 sources with 96 receivers per source. The sources and receivers spacing was 25 m, with sources ranging from 2600 to 9100 m and source-receiver offsets ranging from −2475 to −100 m. Time sampling interval was 4 ms. We simulated the Marmousi data in the Madagascar package [32], received the Marmousoft data from Gilles Lambaré [7] and simulated the Marmouflat data using kinematic ray tracing followed by a convolution with the Ricker wavelet [33]. Three common-shot gathers are displayed on the Fig. 2. Despite being oversimplified, the Marmouflat model exhibits certain similarities with its more complex counterparts and thus can be used to estimate the picking accuracy.

5.2. Picking accuracy. We picked events in the data and filtered out low-semblance and low-fold events (29). We display the picking results on the Fig. 3. We can describe the picking quality as poor in the Marmousi model, satisfying in the Marmousoft model and good in the Marmouflat model.

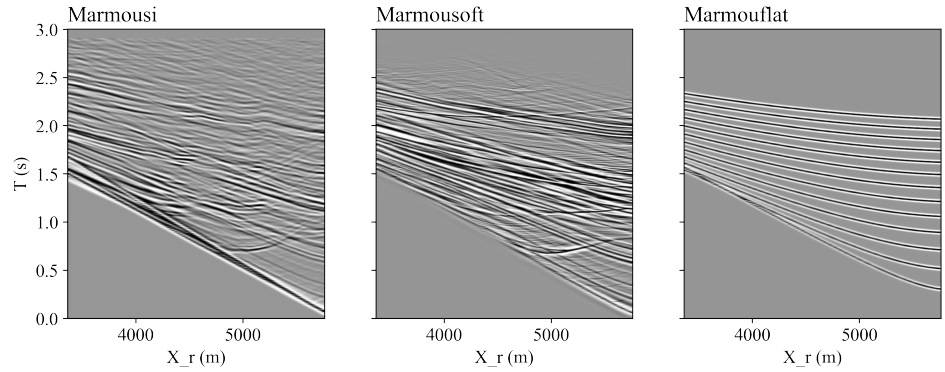

FIG. 2. Common-shot gathers in different models

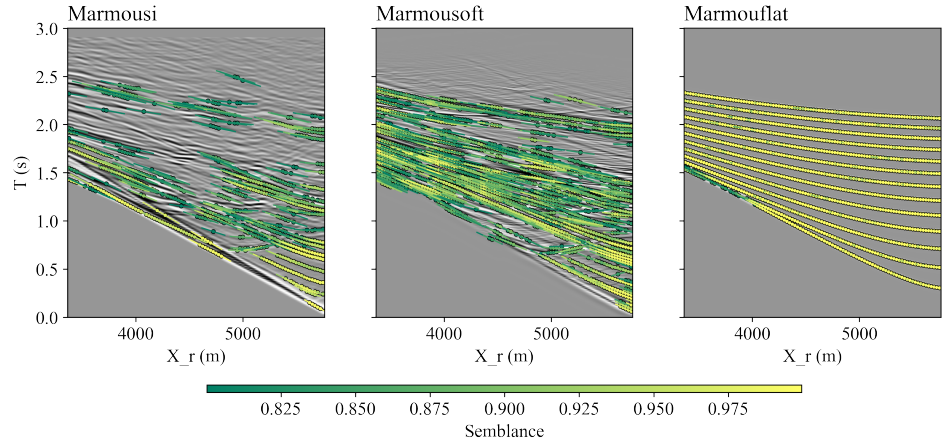

FIG. 3. Common-shot gathers in different models with picking results

We compared the picked Marmouflat events to the ray-modelled ones. We estimate the traveltime error's variance (15) as a single time sampling interval for all the events at once. In contrast, we link the slopes' errors to events' semblances and folds (29), individual for each event. We propose the following estimates:

$$\hat{\sigma}_\tau^k \approx 4 \cdot 10^{-3} \ [s], \quad \hat{\sigma}_s^k = \hat{\sigma}_r^k \approx \frac{1}{\hat{v}_k} \frac{34.2 \hat{s}_k}{\left(-4.9 + 1.6 \hat{Q}_k\right)} \ \left[\frac{s}{m}\right], \tag{30}$$

where the sub- and superscripts k refer to the event's number, not power. We obtained numeric multipliers in $\hat{\sigma}_s^k = \hat{\sigma}_r^k$ using maximum likelihood method. To check them, we divided the actual slope errors by their estimated variances and built a histogram fitting theoretical probability density function of the standard normal distribution, see the Fig. 4. We multiplied the obtained

estimates $\hat{\sigma}_s^k = \hat{\sigma}_r^k$ by a factor of 3 while working with the Marmousi and Marmousoft models to consider their complexity. Average estimated slope error variance in these models amounted for $3.6 \cdot 10^{-5}$ s/m and $1.5 \cdot 10^{-5}$ s/m respectively which accords with the a-priori estimate $1 \cdot 10^{-5}$ s/m used in [7].

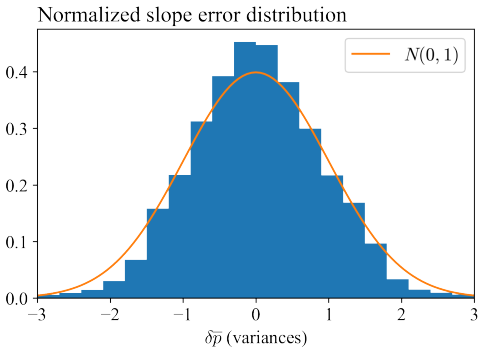

FIG. 4. Normalized slope error distribution

5.3. Inversion results. We passed the Marmousi and Marmousoft events to nonlinear inversion, using inversion grids (21) with 100 m spacing in both directions. In both cases we started from the Marmoulin model and performed 6 Gauss-Newton iterations to reach the stopping criterion. We present the results on the Fig. 5 and Fig. 6.

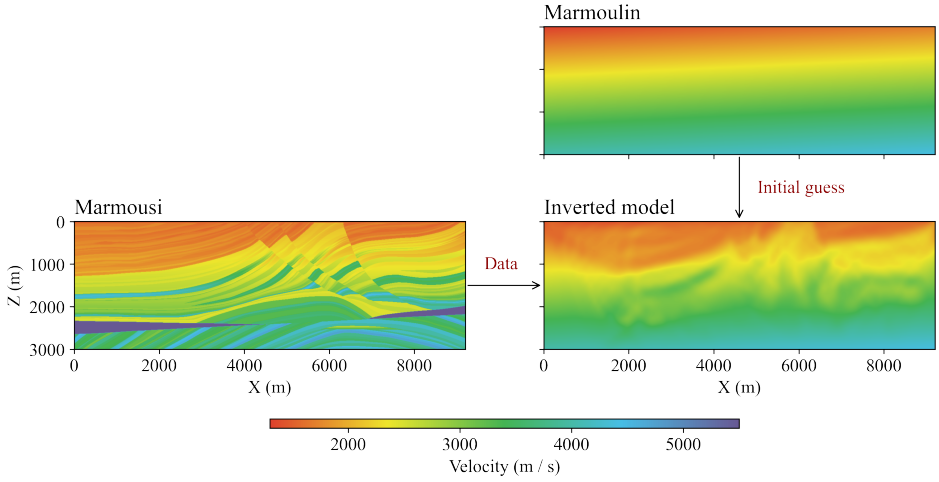

FIG. 5. Marmousi model inversion

Evidently, our tomography failed to reconstruct the Marmousi model, although it did image some of its principal structures. We attribute this

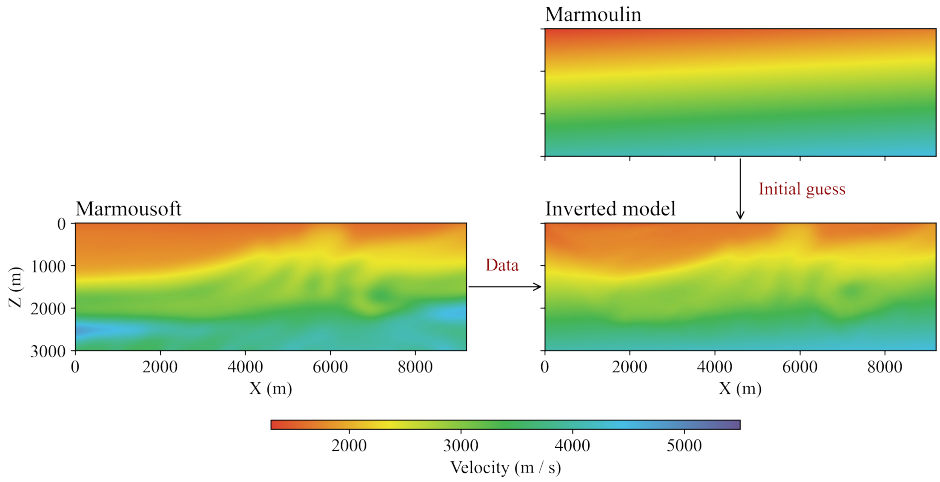

FIG. 6. Marmousoft model inversion

failure to the model's blocky structure, which is not representable in the class of bicubic splines (23), and poor picking quality. The bottom part of the model was almost unilluminated by rays, obstructing its successful resolution. In contrast, smooth Marmousoft model is fairly well-resolved down to its bottom part, thanks to its dense ray coverage. In this model our results are competitive with [5, 6, 7].

To further assess the tomography results, we performed first-arrival Kirchhoff depth migration [15] of the Marmousi and Marmousoft datasets. We present the resulting images on the Fig. 7, along with the true model migration results as a reference. Although being inexact, the inverted models allow for interpretative imaging even in the Marmousi case. We admit that the bottom part of the Marmousi image remains unfocused after inversion, but we remind that we used first-arrival migration, which does not consider all the raypaths and cannot image reflectors in such complex media [34], as the true model migration suggests.

Finally, we checked our regularization parameter's estimate (20) using the L-curve method [35]. To do so, we performed a sequence of Marmousi data inversions, setting α to different fractions of its estimated value. Then we evaluated the models' norms (25) and residuals $\sum_{k=1}^{K} w_k^2 h_k^2$ at final iteration of each inversion run and plotted the squared norms as a function of the residuals, see Fig. 8. We attribute the non-monotonous behaivour of the curve to the events selection algorithm. The residuals were computed for different sets of rays, producing such a non-regular output. However, our α estimate appears to be close to the curve's bending point, thus being at least suboptimal.

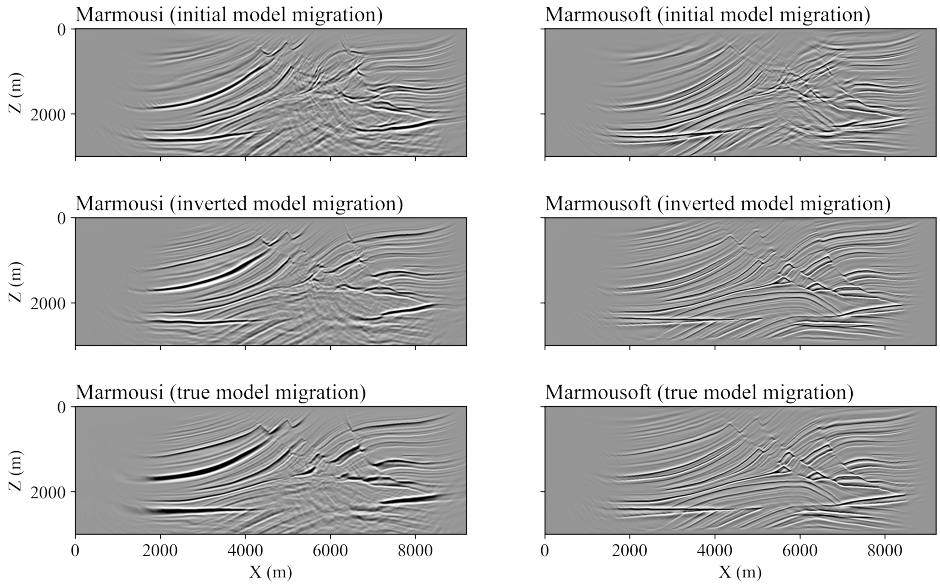

FIG. 7. Migration results in different models

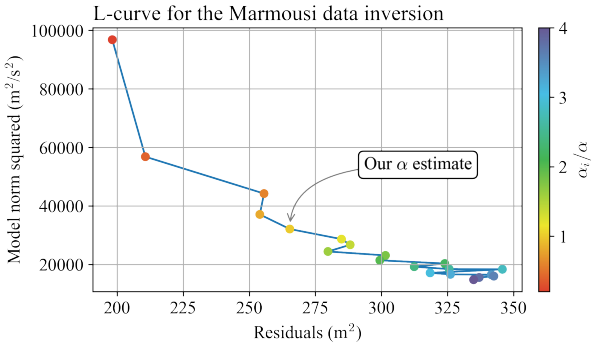

FIG. 8. L-curve for the Marmousi model

6 Discussion and conclusion

In our paper we have presented a reformulation of the CDR method based on our asymptotics of the DSR equation. We provided theoretical aspects of our approach and developed several heuristic or semi-heuristic tricks to obtain a stable slope tomography algorithm competitive with other techniques. Our most important contribution is the suboptimal regularization of the CDR inverse problem which mitigates the notorious instability of the Sword's method. The key steps towards this result were the DSR equation-based wavefield parametrization, homogeneous medium assumption and linearization w.r.t. both velocity model and data errors. Our approach is parsimonious (compared to stereotomography [4]), does not rely on the single-arrival assumption

(compared to adjoint-state method [5, 6]) and does not require multiple data migrations (compared to residual moveout-based methods [36, 37]). On the other hand, it does not process horizontal rays which limits its applicability in complex media with steep reflectors and contrastive layers [1]. However, it manages to construct acceptable velocity models for depth migration and allows for improvements like more elaborated picking engine [38, 39, 30], directional smoothing [23] and multiscale inversion [1, 5, 6]. In addition, the DSR parametrization theoretically allows for lifting the nowhere-horizontal rays restriction by means of a special coordinate system [41]. As a conclusion, we add that it is scalable to 3D settings and anisotropic problems [42] as well.

7 Acknowledgements

This research is supported by grant FSUS-2025-0015 of the Ministry of Science and Higher Education of the Russian Federation. I thank my scientific supervisor, Anton Duchkov, for the problem statement.

References

[1] G. Lambaré, *Stereotomography*, GEOPHYSICS, **73** (2008), VE25–VE34.
[2] L.A. Ryabinkin, *Voprosy Reguliruemogo Napravlennogo Priyoma (RNP) Sejsmicheskih Voln*, Gostoptekhizdat: Moskva, 1957
[3] C.H. Sword, *Tomographic Determination of Interval Velocities from Reflection Seismic Data: The Method of Controlled Directional Reception*, Stanford University, 1987.
[4] F. Billette, G. Lambaré, *Velocity Macro-Model Estimation from Seismic Reflection Data by Stereotomography*, Geophys. J. Int., **135** (1998), 671–690.
[5] F.B. Tavakoli, S. Operto, A. Ribodetti, J. Virieux, *Slope Tomography Based on Eikonal Solvers and the Adjoint-State Method*, Geophys. J. Int., **209** (2017), 1629–1647.
[6] S. Sambolian, S. Operto, A. Ribodetti, F.B. Tavakoli, J. Virieux, *Parsimonious Slope Tomography Based on Eikonal Solvers and the Adjoint-State Method*, Geophys. J. Int., **218** (2019), 456–478.
[7] F. Billette, S.le Bégat, P. Podvin, G. Lambaré, *Practical Aspects and Applications of 2D Stereotomography*, Geophysics, **68** (2003), 1008–1021.
[8] E. Chalard, P. Podvin, S.le Bégat, P. Berthet, B. David, *3D Stereotomographic Inversion On Real Data Set*, SEG Technical Program Expanded Abstracts, (2002), 946-948.
[9] B.S.S. Barbosa, J.C. Costa, E.N.S. Gomes, J. Schleicher, *Sensitivity Analysis for Stereotomography in Elliptic and Anelliptic Media*, SEG Technical Program Expanded Abstracts, (2006), 3413-3417.
[10] SLB, *Common Image Point (CIP) Tomography* (accessed on 18 July 2025).
[11] Viridien, *Geovation* (accessed on 18 July 2025).
[12] S. Sambolian, *Kinematically Consistent Slope Tomography Using Eikonal Solvers and the Adjoint-State Method: Theory and Applications to Velocity Model Building and Event Location*, Université Côte d'Azur, 2021.
[13] J.F. Claerbout, *Imaging the Earth's Interior*, Blackwell Scientific Publications, 1985.
[14] A.V. Belonosova, A.S. Alekseev, *Ob Odnoj Postanovke Obratnoj Kinematicheskoj Zadachi Sejsmiki Dlya Dvumernoj Neodnorodnoj Sredy*, Nekotorye metody i algoritmy interpretacii geofizicheskih dannyh, (1967), 137–154.
[15] B.L. Biondi, *3D Seismic Imaging*, Society of Exploration Geophysicists, 2006.

[16] P.C. Sava, S. Fomel, *Angle-domain Common-image Gathers by Wavefield Continuation Methods*, GEOPHYSICS, **68** (2003), 1065–1074.
[17] A. Khoury, W. Symes, P. Williamson, P. Shen, *DSR Migration Velocity Analysis by Differential Semblance Optimization*, SEG Technical Program Expanded Abstracts, (2006), 2450–2454.
[18] N.N. Shilov, A.A. Duchkov, *Migration Velocity Analysis Using a Ray Method Asymptotics of the Double Square Root Equation*, J. Appl. Ind. Math., **18** (2024), 150–166.
[19] N.N. Shilov, A.A. Duchkov, *Asymptotic Ray Method for the Double Square Root Equation*, J. Mar. Sci. Eng., **12** (2024), 636.
[20] A.A. Duchkov, M.V. de Hoop, *Extended Isochron Rays in Prestack Depth (Map) Migration*, Geophysics, **75** (2010), S139–S150.
[21] A.N. Tikhonov, B.Ya. Arsenin, *Metody Resheniya Nekorrektnyh Zadach, 3-e ispr. izd.*, Nauka. Glavnaya redakciya fiziko-matematicheskoj literatury: Moskva, 1986.
[22] G.I. Marchuk, *Metody Vychislitel'noj Matematiki*, Nauka. Glavnaya redakciya fiziko-matematicheskoj literatury: Moskva, 1980.
[23] J.C. Costa, F.J. da Silva, E.N. Gomes, J. Schleicher, L.A. Melo, D. Amazonas, *Regularization in Slope Tomography*, GEOPHYSICS, **73** (2008), VE39–VE47.
[24] M.T. Taner, F. Koehler, R.E. Sheriff, *Complex Seismic Trace Analysis*, GEOPHYSICS, **44** (1979), 1041–1063.
[25] C.R. Harris, K.J. Millman, S.J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith et al., *Array Programming with NumPy*, Nature, **585** (2020), 357–362.
[26] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright et al., *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, Nat. Methods, **17** (2020), 261–272.
[27] S.K. Lam, A. Pitrou, S. Seibert, *Numba: a LLVM-based Python JIT compiler*, LLVM '15: Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC, (2015), 1–6.
[28] GitHub, *Fteikpy: Accurate Eikonal Solver for Python* (accessed on 18 July 2025).
[29] J.D. Hunter, *Matplotlib: A 2D Graphics Environment*, Comput. Sci. & Eng., **9** (2007), 90–95.
[30] A. Brougois, M. Bourget, P. Lailly, M. Poulet, P. Ricarte, R. Versteeg, *Marmousi, Model and Data*, Proceedings of the EAEG Workshop - Practical Aspects of Seismic Data Inversion; European Association of Geoscientists & Engineers, 1990.
[31] SW3D, *Marmousi model and data set* (accessed on 18 July 2025).
[32] S. Fomel, P. Sava, I. Vlad, Y. Liu, V. Bashkardin, *Madagascar: Open-Source Software Project for Multidimensional Data Analysis and Reproducible Computational Experiments*, J. Open Res. Softw., **1** (2013), e8.
[33] V. Červený, *Seismic Ray Theory*, Cambridge University Press: Cambridge, 2001.
[34] M.S. Operto, S. Xu, G. Lambaré, *Can We Quantitatively Image Complex Structures with Rays?*, GEOPHYSICS, **65** (2000), 1223–1238.
[35] P.C. Hansen, D.P. O'Leary, *The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems*, SIAM J. Sci. Comput., **14** (1993), 1487–1503.
[36] H. Chauris, M.S. Noble, G. Lambaré, P. Podvin, *Migration Velocity Analysis from Locally Coherent Events in 2-D Laterally Heterogeneous Media, Part I: Theoretical Aspects*, GEOPHYSICS, **67** (2002), 1202–1212.
[37] H. Chauris, M.S. Noble, G. Lambaré, P. Podvin, *Migration Velocity Analysis from Locally Coherent Events in 2-D Laterally Heterogeneous Media, Part II: Applications on Synthetic and Real Data*, GEOPHYSICS, **67** (2002), 1213–1224.
[38] S. Nguyen, R. Baina, M. Alerini, G. Lambaré, V. Devaux, M. Noble, *Stereotomography Assisted by Migration of Attributes*, Geophys. Prospect., **56** (2008), 613–625.

[39] T. Barros, R. Lopes, H. Chauris, *Towards a More Robust Input for Stereotomography*, Geophys. Prospect., **70** (2022), 502–524.
[40] F. Brandolin, M. Ravasi, T. Alkhalifah, *PINNslope: Seismic Data Interpolation and Local Slope Estimation with Physics Informed Neural Networks*, GEOPHYSICS, **89** (2024), V331–V345.
[41] C.C. Stolk, M.V. de Hoop, W.W. Symes, *Kinematics of Shot-Geophone Migration*, GEOPHYSICS, **74** (2009), WCA19–WCA34.
[42] C. Wu, B. Feng, H. Wang, T. Wang, *Three-Dimensional Angle-Domain Double-Square-Root Migration in VTI Media for the Large-Scale Wide-Azimuth Seismic Data*, Acta Geophys., **68** (2020), 1021–1037.

Appendix A

In this Appendix we provide explicit formulae for the subsurface offset's $\hat{h}$ (12) derivatives appearing in (13)–(17). We define

$$\hat{C}_\tau = C_\tau\left(\widehat{\vec{x}}, \widehat{\vec{p}}\right). \tag{1A}$$

With this shorthand:

$$\frac{\partial \hat{h}}{\partial \hat{\tau}} = -\left(\frac{\hat{p}_r}{\sqrt{\frac{1}{\hat{v}^2} - \hat{p}_r^2}} - \frac{\hat{p}_s}{\sqrt{\frac{1}{\hat{v}^2} - \hat{p}_s^2}}\right)\frac{\hat{C}_\tau}{2}, \tag{2A}$$

$$\frac{\partial \hat{h}}{\partial \hat{p}_s} = \left(\frac{1}{\sqrt{\frac{1}{\hat{v}^2} - \hat{p}_s^2}} + \frac{\hat{p}_s^2}{\left(\frac{1}{\hat{v}^2} - \hat{p}_s^2\right)^{\frac{3}{2}}}\right)\frac{\hat{C}_\tau \hat{\tau}}{2} + \frac{\hat{p}_s}{\sqrt{\frac{1}{\hat{v}^2} - \hat{p}_s^2}}\frac{\partial \hat{C}_\tau}{\partial \hat{p}_s}\frac{\hat{\tau}}{2}, \tag{3A}$$

$$\frac{\partial \hat{h}}{\partial \hat{p}_r} = -\left(\frac{1}{\sqrt{\frac{1}{\hat{v}^2} - \hat{p}_r^2}} + \frac{\hat{p}_r^2}{\left(\frac{1}{\hat{v}^2} - \hat{p}_r^2\right)^{\frac{3}{2}}}\right)\frac{\hat{C}_\tau \hat{\tau}}{2} - \frac{\hat{p}_r}{\sqrt{\frac{1}{\hat{v}^2} - \hat{p}_r^2}}\frac{\partial \hat{C}_\tau}{\partial \hat{p}_r}\frac{\hat{\tau}}{2}, \tag{4A}$$

$$\begin{aligned}\frac{\partial \hat{h}}{\partial \hat{v}} = &-\frac{1}{\hat{v}^3}\left(\frac{\hat{p}_r}{\left(\frac{1}{\hat{v}^2} - \hat{p}_r^2\right)^{\frac{3}{2}}} - \frac{\hat{p}_s}{\left(\frac{1}{\hat{v}^2} - \hat{p}_s^2\right)^{\frac{3}{2}}}\right)\frac{\hat{C}_\tau \hat{\tau}}{2} - \\ &-\left(\frac{\hat{p}_r}{\sqrt{\frac{1}{\hat{v}^2} - \hat{p}_r^2}} - \frac{\hat{p}_s}{\sqrt{\frac{1}{\hat{v}^2} - \hat{p}_s^2}}\right)\frac{\partial \hat{C}_\tau}{\partial \hat{v}}\frac{\hat{\tau}}{2},\end{aligned} \tag{5A}$$

where

$$\frac{\partial \hat{C}_\tau}{\partial \hat{p}_s} = -\frac{\hat{C}_\tau^2}{\hat{v}^2}\frac{\hat{p}_s}{\left(\frac{1}{\hat{v}^2}-\hat{p}_s^2\right)^{\frac{3}{2}}}, \quad \frac{\partial \hat{C}_\tau}{\partial \hat{p}_r} = -\frac{\hat{C}_\tau^2}{\hat{v}^2}\frac{\hat{p}_r}{\left(\frac{1}{\hat{v}^2}-\hat{p}_r^2\right)^{\frac{3}{2}}},$$
$$\frac{\partial \hat{C}_\tau}{\partial \hat{v}} = -\frac{2\hat{C}_\tau^2}{\hat{v}^3}\left(\frac{1}{\sqrt{\frac{1}{\hat{v}^2}-\hat{p}_s^2}}+\frac{1}{\sqrt{\frac{1}{\hat{v}^2}-\hat{p}_r^2}}\right) - \tag{6A}$$
$$-\frac{\hat{C}_\tau^2}{\hat{v}^5}\left(\frac{1}{\left(\frac{1}{\hat{v}^2}-\hat{p}_s^2\right)^{\frac{3}{2}}}+\frac{1}{\left(\frac{1}{\hat{v}^2}-\hat{p}_r^2\right)^{\frac{3}{2}}}\right).$$

We emphasize that (3A)–(5A) linearly grow with the reflection traveltime $\hat{\tau}$. Thus, when inserted in (16), (17) and (20), these multipliers will reduce the weights of deeper events and magnify corresponding regularization contributions.

Appendix B

In this Appendix we present the remaining formulae for bicubic splines' construction. For the theory see [22]. We shall follow the main text's notation (21)–(24).

The coefficient matrices $\mathbf{C}_{mn}$ from (24) linearly depend on the grid velocities:

$$\mathbf{C}_{mn} = \mathbf{D}_{x,m}^T\ \mathbf{V}\mathbf{D}_{z,n}, \tag{1B}$$

with

$$\mathbf{D}_{x,m} = \left(\frac{1}{\Delta x}\begin{pmatrix}\vec{e}_{m+1}^T\\ \vec{e}_m^T\\ 0\\ 0\end{pmatrix} + \frac{\Delta x}{6}\begin{pmatrix}-\vec{e}_{m+1}^T\\ -\vec{e}_m^T\\ \vec{e}_{m+1}^T\\ \vec{e}_m^T\end{pmatrix}\right)^T, \tag{2B}$$
$$\vec{e}_m = \begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix} \leftarrow m^{th}\ row.$$

Matrices $\mathbf{A}_x$ and $\mathbf{B}_x$ are sparse matrices with nonzero elements defined as

$$a_{ij}^x = \begin{cases} 1, & i=j=1, \\ 1, & i=j=M, \\ \dfrac{2\Delta x}{3}, & 1<i=j<M, \\ \dfrac{\Delta x}{3}, & 2<i=j+1<M, \\ & 1<i=j-1<M-1, \end{cases}$$

$$b_{ij}^x = \begin{cases} -\dfrac{2}{\Delta x}, & 1<i=j<M, \\ \dfrac{1}{\Delta x}, & 2<i=j+1<M, \\ & 1<i=j-1<M-1. \end{cases} \tag{3B}$$

Analogous expressions for $\mathbf{D}_{z,n}$, $\mathbf{A}_z$ and $\mathbf{B}_z$ are obtained by a change of indices and subscripts.

The coefficient matrices $\mathbf{C}_{mn}$ are required to evaluate the spline or its derivatives w.r.t spatial coordinates. Matrices $\mathbf{D}_{x,m}$ and $\mathbf{D}_{z,n}$ are required to evaluate its derivatives w.r.t. grid velocities. The coefficients $\mathbf{C}_{mn}$ are updated during the inversion along with the grid velocities while $\mathbf{D}_{x,m}$ and $\mathbf{D}_{z,n}$ remain fixed.

Appendix C

In this Appendix we provide explicit expressions for the matrix $\mathbf{G}$ and vector $\vec{l}$ from (26). They read

$$\mathbf{G} = \nabla_{\vec{\mathbf{V}}} \vec{l} = \left(\frac{\partial \vec{l}}{\partial v_{11}}, \frac{\partial \vec{l}}{\partial v_{12}}, \ldots, \frac{\partial \vec{l}}{\partial v_{MN}} \right), \quad \vec{l} = \begin{pmatrix} \vec{l}_h \\ \vec{l}_v \\ \vec{l}_{\partial_x v} \\ \vec{l}_{\partial_z v} \\ \vec{l}_{\partial_x^2 v} \\ \vec{l}_{\partial_z^2 v} \\ \vec{l}_{\partial_{xz}^2 v} \end{pmatrix}. \tag{1C}$$

We note that matrix $\mathbf{G}$ is formed by vector-columns, and no transposition is required. We shall list all the $\vec{l}$'s subvectors below. To avoid large repetitive multipliers, we denote

$$w_\alpha = \frac{\sqrt{\alpha}}{\| v\left(x, z; \mathbf{V}^{(0)}\right) \|}. \tag{2C}$$

With this shorthand we have:

$$\vec{l}_h = \begin{pmatrix} w_1 h_1 \\ \vdots \\ w_K h_K \end{pmatrix}, \quad \vec{l}_v = w_\alpha \begin{pmatrix} v_{11} \\ \vdots \\ v_{MN} \end{pmatrix},$$

$$\vec{l}_{\partial_x v} = w_\alpha \Delta x \begin{pmatrix} \left.\dfrac{\partial v}{\partial x}\right|_{\substack{x=x_1 \\ z=z_1}} \\ \vdots \\ \left.\dfrac{\partial v}{\partial x}\right|_{\substack{x=x_M \\ z=z_N}} \end{pmatrix}, \quad \vec{l}_{\partial_z v} = w_\alpha \Delta z \begin{pmatrix} \left.\dfrac{\partial v}{\partial z}\right|_{\substack{x=x_1 \\ z=z_1}} \\ \vdots \\ \left.\dfrac{\partial v}{\partial z}\right|_{\substack{x=x_M \\ z=z_N}} \end{pmatrix},$$

$$\vec{l}_{\partial_x^2 v} = w_\alpha \Delta x^2 \begin{pmatrix} \left.\dfrac{\partial^2 v}{\partial x^2}\right|_{\substack{x=x_1 \\ z=z_1}} \\ \vdots \\ \left.\dfrac{\partial^2 v}{\partial x^2}\right|_{\substack{x=x_M \\ z=z_N}} \end{pmatrix}, \quad \vec{l}_{\partial_z^2 v} = w_\alpha \Delta z^2 \begin{pmatrix} \left.\dfrac{\partial^2 v}{\partial z^2}\right|_{\substack{x=x_1 \\ z=z_1}} \\ \vdots \\ \left.\dfrac{\partial^2 v}{\partial z^2}\right|_{\substack{x=x_M \\ z=z_N}} \end{pmatrix}, \tag{3C}$$

$$\vec{l}_{\partial_{xz}^2 v} = \sqrt{2} w_\alpha \Delta x \Delta z \begin{pmatrix} \left.\dfrac{\partial^2 v}{\partial x \partial z}\right|_{\substack{x=x_1 \\ z=z_1}} \\ \vdots \\ \left.\dfrac{\partial^2 v}{\partial x \partial z}\right|_{\substack{x=x_M \\ z=z_N}} \end{pmatrix}.$$

Evidently, the squared Euclidean norm of $\vec{l}$ gives the loss functional (18).

Nikolay Nikolaevich Shilov
Novosibirsk State University,
ul. Pirogova, 2,
630090, Novosibirsk, Russia
Email address: n.shilov@g.nsu.ru