VELOCITY MODEL BUILDING BY GEOMETRICAL SPREADING FOCUSING

P. Znak, B. Kashtan, and D. Gajewski

email: pavel.znak@uni-hamburg.de **keywords:** tomography, curvatures, geometrical spreading, focusing

ABSTRACT

A reliable and fast workflow for smooth velocity model building is of great interest for subsequent depth migration. Also full-waveform inversion benefits from geologically reasonable initial velocity models. At the same time, seismic data enhancement by wavefront attributes is accepted. These attributes are extracted from seimic data by common-reflection-surface (CRS) stack and locally characterize the wavefronts. We propose a novel tomography based on them. It is valid for reflection, diffraction and passive seismic data inversion. The idea of the new approach is minimizing backpropagated geometrical spreading of diffracted and/or fictitious NIP-waves. During the inversion the wavefront attributes remain fixed at the common midpoints and serve as initial conditions for kinematic and dynamic ray tracing. The inverse problem turns out to be naturally parametrized by the velocity model as the only unknown. It significantly decreases the tomographic matrix dimensions and improves the data/unknowns ratio if compared to the conventional wavefront attributes tomography. The reduction of Fréchet derivatives is also an attractive feature if a generalization to 3D anisotropic media is considered. We provide both Fréchet derivatives and adjoint-state method formulation for the gradient of the new objective function. The algorithm combined with the L-BFGS-B quasi-Newton solver was tested on a salt body synthetic dataset. It converges to a consistent smooth velocity image.

INTRODUCTION

Multiparameter stacking is a powerful tool of seismic processing. It enhances signal-to-noise ratio of zerooffset stack sections (Mayne, 1962; Mann et al., 1999; Jäger et al., 2001) and pre-stack data (Baykulov and Gajewski, 2009) producing physically meaningful wavefront attributes (Hubral, 1983). The stacking quality and parameters search strongly depends on a choice of the traveltime-based operator. Normal moveout (Mayne, 1962) was generalized to common-reflection-surface (CRS) operator (Mann et al., 1999; Jäger et al., 2001) designed to utilize redundancy of reflection data. Later on, to account for the diffraction phenomenon, the double-square-root-based operators were applied. While CRS is robust to overburden heterogeneities, multifocusing (Gelchinsky et al., 1999) is an efficient tool in smooth environments with curved interfaces. Recently, a method of implicit CRS (i-CRS) accounting for the reflector curvature and combining the advantages of CRS and multifocusing was proposed (Schwarz et al., 2014). For strongly non-hyperbolic events a non-hyperbolic CRS (n-CRS) operator was derived by (Fomel and Kazinnik, 2013).

Wavefront attributes have a distinct physical meaning. For the CRS operator they locally describe fictitious waves: a wave emerging from a zero-offset ray normal incidence point and an exploding reflector wave (Hubral, 1983). Namely, they are represented by horizontal component of slowness and curvatures of these fictitious wavefronts. For a point diffractor the wavefront attributes describe the real wavefront and the curvatures coincide. This property, in particular, yields a criterion for diffraction separation (Dell and Gajewski, 2011).

The idea of utilizing wavefront attributes as an input data for a fast and robust smooth velocity model building was first proposed by Duveneck (2004). The slope and curvature inversion was later applied to image rays arising in time-migration by Dell et al. (2014). Recently, the wavefront attributes tomography was extended to passive seismics (Schwarz et al., 2016) and applied to diffractions (Bauer et al., 2017).

Previously, the inverse problem of wavefront attributes tomography was formulated similarly to the pre-stack stereotomography (Lambaré, 2008). The velocity model, diffractors and scattering angles were considered as independent unknowns. The objective function was chosen to minimize misfits of the attributes, one-way traveltimes and arrival coordinates. The number of unknowns in this formulation unnaturally depends on the number of data picks. However, seismic tomography is ill-conditioned and requires a sufficient regularization smoothing the velocity model and reducing resolution (Hansen, 1998; Costa et al., 2008). Usually, a smaller size of the tomographic matrix corresponds to a greater stability of the algorithm. Moreover, a stable and unique inversion usually requires the amount of data significantly exceeding the amount of unknowns, which is not the case in the wavefront attributes tomography. In contrast to the prestack stereotomography, the data of post-stack wavefront attributes tomography contain traveltimes for all involved rays. We fully exploit this fact by developing an inversion workflow based on an original physical principle of geometrical spreading focusing. In addition to the attractive inversion properties, the method takes advantage of a simplified gradient computing through Fréchet derivatives and the adjoint-state method (Plessix, 2006). Complemented by a modern iterative optimization (L-BFGS-B) our approach allowed to build a smooth velocity of a synthetic salt body.

THEORY AND METHOD

Wavefront attributes tomography

Processing of streamer data by the hyperbolic zero-offset CRS stacking operator (Mann et al., 1999; Jäger et al., 2001)

$$t^{2}(\triangle x,h) = (t_{0} + 2p_{0}\triangle x)^{2} + 2t_{0}\cos^{2}\alpha_{0}(M^{N}\triangle x^{2} + M^{NTP}h^{2})$$
(1)

at a traveltime sample t_0 results in a set of wavefront attributes: horizontal slowness $p_0 = \sin \alpha_0/v_0$, wavefront curvature of normal incidence point (NIP) wave $M^{NIP} = \kappa^{NIP}/v_0$ and wavefront curvature of normal (N) exploding reflector wave $M^N = \kappa^N/v_0$. For a specific system of rays M is a ratio of slowness spreading P and geometrical spreading Q (Cervený, 2001). CRS attributes locally describe wavefronts of fictitious NIP waves radiated from reflector elements. Utilizing them for the velocity model building is wavefront attributes tomography (Duveneck, 2004). If a point diffractor wavefield is processed the wavefront attributes describe the actual wavefront and $M^N = M^{NIP}$ (Dell and Gajewski, 2011). This gives a way to joint reflection and diffraction tomography (Bauer et al., 2017). Moreover, analogous processing workflow exists for passive seismic data and results in excitation time, horizontal slowness and curvature of actual wavefronts (Schwarz et al., 2016). The wavefront attributes tomography is especially suitable for these new applications since both point diffraction and passive seismic event rays coverage doesn't benefit from large offsets of stereotomography.

Inverse problem formulation

An objective function introduced by Duveneck (2004) to utilize the wavefront attributes in the poststack inversion

$$J(\mathbf{m}) = \sum_{i=1}^{N_{data}} (x_i - x_i(\mathbf{m}_i))^2 + \sum_{i=1}^{N_{data}} (t_i - t_i(\mathbf{m}_i))^2 + \sum_{i=1}^{N_{data}} (p_i - p_i(\mathbf{m}_i))^2 + \sum_{i=1}^{N_{data}} (M_i - M_i(\mathbf{m}_i))^2$$
(2)

is a sum of squared attributes misfits over all the automatically picked N_{data} data points. A vector of unknowns **m** comprises diffractor (NIP) coordinates x_i^d , z_i^d and zero-offset ray emergence angles α_i^d for each pick $i = 1, 2, ..., N_{data}$ together with coefficients v_{ij} defining a smooth velocity model in terms of B-splines,

$$v = \sum_{j=1}^{N_z} \sum_{k=1}^{N_x} v_{jk} \beta_j(z) \beta_k(x),$$
(3)

with $N_z \times N_x$ nodes, $\mathbf{m}_i = (x_i^d, z_i^d, \alpha_i^d, \mathbf{v}).$

The inverse problem has $4N_{data}$ data points and $3N_{data} + N_z N_x$ unknowns. The number of unknowns depends on the number of picks. At best the N_{data} picks number significantly exceeds the number of basis functions nodes. This would lead to a ratio of the number of data points and the number of unknowns equal to $\frac{4}{3}$. Other cases with higher velocity sampling leads to even smaller ratios close to one. Our research was initially motivated by having a velocity model as the only unknown with the free choice of grid nodes spacing for a given data amount. The first idea that came to mind was to replace the objective function (2) with one minimizing only wavefronts curvatures:

$$J(\mathbf{v}) = \sum_{i=1}^{N_{data}} (M_i - M_i(\mathbf{m}_i(\mathbf{v})))^2,$$
(4)

provided that x, t and p are fixed. The coordinates of diffractors (NIPs) and the zero-offset ray emergence angles are determined by the velocity model coefficients v with time reversal ray tracing as soon as the surface values x, t and p are given, $\mathbf{m}_i(\mathbf{v}) = (x_i^d(\mathbf{v}), z_i^d(\mathbf{v}), \mathbf{v})$. M turns out to be a composition of reverse time kinematic ray tracing to the diffractor and forward dynamic ray tracing from the diffractor to the surface. Hence, to get a gradient of this objective function we could apply a chain rule for Fréchet derivatives. However, we find this approach too cumbersome. Especially frightening seems a perspective of generalizing it to the 3D anisotropic case. Therefore, we present another natural solution which preserves all the advantages and doesn't require a hard derivation. The idea is to apply a concept of focusing: to minimize a third-party quantity at the diffractor position while keeping the data being fixed boundary conditions.

Geometrical spreading focusing

We simultaneously minimize geometrical spreading of all the diffracted and/or NIP waves presented in the dataset. Wherein the wavefront attributes t, p, M are considered fixed for corresponding CMP coordinates. Since the diffracted wavefield and the fictitious NIP waves are kinematically equivalent to point sources, in the true velocity model geometrical spreading vanishes when back propagated up to one-way traveltime. This imaging principle is valid for the passive seismic events as well. After a velocity model is retrieved reflectors and/or diffractors are localized by ray tracing.

All the attributes picks are considered independent. In order to simplify the notations let us think all diffracted (NIP) waves emerges at the same time t = 0 and arrives to the surface at CMP coordinate x_i and time t_i . To be solved in depth the dynamic ray tracing system (Cervený, 2001)

$$\frac{d}{dt}\begin{pmatrix}Q\\P\end{pmatrix} = S\begin{pmatrix}Q\\P\end{pmatrix}, \quad S = \begin{pmatrix}0 & v^2\\-\frac{1}{v}\frac{\partial^2 v}{\partial q^2} & 0\end{pmatrix}$$
(5)

requires initial conditions for Q and P. Unfortunately, they are not presented in the set of CRS attributes. The only available after processing dynamic quantity is the wavefront curvature related quantity M. Although dynamic tracing of M is theoretically possible by solving the Riccati equation (Cervený, 2001), it faces difficulties due to the unlimited growth in the vicinity of point sources and caustics. The geometrical spreading Q is defined on a ray only up to a certain constant quotient determined by ray fan parametrization. With the help of M we are able to retrieve the geometrical spreading along the whole ray up to a constant multiplier. This is done by choosing the initial conditions of dynamic ray tracing satisfying $\hat{P}(t_{init})/\hat{Q}(t_{init}) = M(t_{init})$. There exist an infinite number of such choices. We fix one with a condition for geometrical spreading to be normalized at the surface. A downward propagator of dynamic ray tracing system (5) for the *i*-th pick:

$$\Pi_i(t,t_i) = \begin{pmatrix} Q_i^{(1)}(t,t_i) \ Q_i^{(2)}(t,t_i) \\ P_i^{(1)}(t,t_i) \ P_i^{(2)}(t,t_i) \end{pmatrix}, \quad \Pi_i(t_i,t_i) = E,$$
(6)

has to be traced back in time $(t < t_i)$ to the subsurface. We define normalized dynamic ray tracing

quantities to be a special linear combination of the propagator columns

$$\begin{pmatrix} Q_i \\ \hat{P}_i \end{pmatrix}(t) = \Pi_i(t, t_i) \begin{pmatrix} 1 \\ M_i \end{pmatrix},$$
(7)

where M_i is the experimental value of the wavefront curvature attribute. $Q_i^{(1)}$ and $P_i^{(1)}$ are multiplied by one with the units of M_i . Therefore, units of \hat{Q}_i coincide with the units of P. As a linear combination of solutions the normalized vector satisfies the dynamic ray tracing system (5). Indeed, for normalized dynamic quantities we obtain $\hat{Q}_i(t_i) = 1$ and $\hat{P}_i(t_i) = M_i$. Therefore,

$$\frac{P_i(t_i)}{\hat{Q}_i(t_i)} = \frac{P_i(t_i)}{Q_i(t_i)} = M_i,$$
(8)

where Q_i and P_i are the actual dynamic quantities along the ray. Due to the linearity of the system the actual Q_i, P_i pair and the pair of normalized \hat{Q}_i, \hat{P}_i are proportional at any traveltime with the quotient equal to the actual geometrical spreading value at the surface:

$$Q_i(t) = Q_i(t_i)\hat{Q}_i(t), \quad P_i(t) = Q_i(t_i)\hat{P}_i(t).$$
 (9)

Geometrical spreading for a given traveltime and M value can be retrieved only up to a constant. However, if at some traveltime the actual geometrical spreading Q vanishes the normalized geometrical spreading \hat{Q} necessarily does the same. In particular at a diffractor (source) position $\hat{Q}_i(0) = Q_i(0) = 0$. In accordance to this we introduce an objective function as a sum of the squared normalized geometrical spreading spreading propagated back up to the one-way traveltime over all the data picks:

$$J = \frac{1}{2} \sum_{i=1}^{N_{data}} \hat{Q}_i^2(0).$$
 (10)

Finally, the wavefront attributes tomography is formulated as an inversion in velocities as the only unknown with N_{data} data dimension and the $N_z \times N_x$ - dimensional space of unknowns.

Fréchet derivatives and adjoint-state method

The gradient of the objective function can be expressed through the Fréchet derivatives of the state variables:

$$\Delta J = \sum_{i=1}^{N_{data}} \hat{Q}_i(0) \Delta \hat{Q}_i(0).$$
(11)

To calculate the perturbations of dynamic quantities we apply, following Duveneck (2004), the equations of Farra and Madariaga (1987) for a paraxial to a perturbed ray in ray-centered coordinates:

$$\frac{d}{dt} \begin{pmatrix} \triangle \hat{Q} \\ \triangle \hat{P} \end{pmatrix} = S \begin{pmatrix} \triangle \hat{Q} \\ \triangle \hat{P} \end{pmatrix} + \triangle S \begin{pmatrix} \hat{Q} \\ \hat{P} \end{pmatrix}, \quad \triangle S = S_{\triangle q, \triangle p} + S_{\triangle v}, \tag{12}$$

$$S_{\Delta q,\Delta p} = \begin{pmatrix} 2v \frac{\partial v}{\partial q} \Delta p & 2v \frac{\partial v}{\partial q} \Delta q \\ \left(\frac{3}{v^2} \frac{\partial v}{\partial q} \frac{\partial^2 v}{\partial q^2} - \frac{1}{v} \frac{\partial^3 v}{\partial q^3}\right) \Delta q - 2v \frac{\partial v}{\partial q} \Delta p \end{pmatrix}, \quad S_{\Delta v} = \begin{pmatrix} 0 & v \Delta v \\ -\frac{\partial^2}{\partial q^2} \left(\frac{\Delta v}{v}\right) + \frac{\Delta v}{v^2} \frac{\partial^2 v}{\partial q^2} & 0 \end{pmatrix}, \quad (13)$$

where $\triangle q(t)$ and $\triangle p(t)$ are ray-centered coordinate and slowness of the perturbed ray, which in turn have to be evaluated by

$$\frac{d}{dt} \begin{pmatrix} \Delta q \\ \Delta p \end{pmatrix} = S \begin{pmatrix} \Delta q \\ \Delta p \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{\Delta v}{v^2} \frac{\partial v}{\partial q} - \frac{1}{v} \frac{\partial \Delta v}{\partial q} \end{pmatrix}.$$
(14)

The initial conditions $\triangle \hat{Q}_i(t_i)$ and $\triangle \hat{P}_i(t_i)$ are zeros because the downward propagator is the identity matrix at $t = t_i$ regardless of the velocity model. We fix the arrival point of the ray, which means $\triangle q_i(t_i)$ is also zero. However, $\triangle p_i(t_i)$ differs:

$$\Delta p_i(t_i) = \frac{\Delta v}{v_0} \frac{p_i}{\sqrt{1 - v_0^2 p_i^2}}.$$
(15)



Figure 1: (a) Stack section. (b) Coherence section with overlaid picks.

The inhomogeneous system (12) is solved with the propagator of homogeneous one (see, e.g, Gilbert and Backus, 1966; Bellman, 1997):

$$\Delta \hat{Q}_i(0) = \begin{pmatrix} Q_i^{(1)}(0,t_i) \\ Q_i^{(2)}(0,t_i) \end{pmatrix}^T \int_{t_i}^0 \Pi_i^{-1}(t',t_i) \Delta S_i(t') \begin{pmatrix} \hat{Q}_i(t') \\ \hat{P}_i(t') \end{pmatrix} dt'.$$
(16)

We also derived a formula for the gradient in form of the adjoint-state method (Plessix, 2006; Chavent, 2010):

$$\Delta J = \sum_{i=1}^{N_{data}} \int_{0}^{t_i} \left(-\lambda_i^P(t') \,\lambda_i^Q(t') \right) \Delta S_i(t') \begin{pmatrix} \hat{Q}_i(t') \\ \hat{P}_i(t') \end{pmatrix} dt'. \tag{17}$$

Adjoint-state variables λ_i^Q and λ_i^P are uniquely defined as a solution of dynamic ray tracing system (5) with initial conditions at emergence time $\lambda_i^Q(0) = -\hat{Q}_i(0)$, $\lambda_i^P(0) = 0$. Combining the unperturbed downward propagator values we are able to compute an upward one

$$\Pi_i(t,0) = \Pi_i(t,t_i)\Pi_i^{-1}(0,t_i)$$
(18)

and, therefore, the adjoint-state variables.

The adjoint-state method formula (17) assumes an integration of the scalar function, whereas the gradient expressed through the Fréchet derivatives (11, 16) needs an integration of the two element vector. However, the computing time of this vector is not equivalent to a doubled computing time by the formula



Figure 2: (a) Slowness section. (b) Wavefront curvature section.

(17). For the synthetic setup given in the following section we found that the adjoint-state method is approximately 1.5 times faster than the Fréchet derivatives approach. Increasing the number of nodes in the ray propagation direction and the basis function support (e.g., by taking the B-splines of higher order) makes this advantage even more significant. We expect it to grow also with a transition to 3D (Plessix, 2006).

SYNTHETIC DATA EXAMPLE

An acoustic land streamer dataset of primary reflections for a salt body velocity model shown in Figure 3(a) was modeled by the Seismic Unix Gaussian beam method routine. The source function was a 40 Hz Ricker wavelet. Gaussian noise with SNR = 40 was added to the traces. The receivers were located at 50 m spacing leading to a total CMP number of 641. We performed the CRS stack (Figure 1(a)) with a maximum offset aperture of 2 km, which implies a maximum CMP fold of 21. The coherence section is given in Figure 1(b). The corresponding slowness and wavefront curvature attributes (Figure 2) were picked with a coherence threshold criterion following a condition of one pick per boundary. As an input for the inversion we used picks from odd CMP numbers. They are depicted in Figure 1(b) by black dots.

These 2952 picks were utilized for the inversion by means of geometrical spreading focusing in the frame of 11×17 4th order B-spline nodes grid with 1000 m lateral and 350 m vertical nodes spacing. We apply the limited-memory constrained Broyden-Fletcher-Goldfarb-Shanno (L-BFGS-B) optimization algorithm (Byrd et al., 1995). The initial model was a constant gradient growing from 2.2 km/s surface velocity to 4.4 km/s velocity of the salt body bottom (Figure 3(a)). We forced the nodes values to stay within this velocity interval and added an additional term to the objective function and the gradient keeping the surface velocity value fixed. The only regularization was due to the B-splines smoothness. No additional regularizing terms were added to the objective function. The inversion process presented in Figure 4 stopped after 81 iterations satisfying the convergence criterion. Figure 3(b) illustrates the retrieved tomographic image. The normal incidence points were localized and superimposed on the retrieved velocity image. The image being smoothed, nonetheless, highly correlates with the synthetic model, as well as the retrieved reflector elements.

CONCLUSIONS

We have developed a new approach for the macro-velocity model building. Our method utilizes the local wavefront attributes for focusing geometrical spreading. The traveltime, slowness and wavefront curvature are picked during CRS analysis. We provide the Fréchet derivatives and the adjoint-state method formulation for the gradient of the objective function. The geometrical spreading focusing was tested on the reflection only synthetic dataset driven by the L-BFGS-B optimization algorithm. The retrieved to-mographic image highly correlates with the synthetic salt body model and, therefore, could serve as a relevant initial model for the detailed imaging by full-waveform inversion or as a subsurface model for the



Figure 3: a) Synthetic salt body velocity model. b) Retrieved tomographic image with overlaid reflector elements.



Figure 4: (a) Objective function values. (b) Maximum norm values of the objective function gradient.

subsequent depth migration.

ACKNOWLEDGMENTS

The authors appreciate the support of the Wave Inversion Technology (WIT) Consortium sponsors and the Federal Ministry for Economic Affairs and Energy of Germany (project number 03SX427B). We also sincerely thank the members of the applied seismic group of Hamburg University for the assistance and fruitful discussions. The synthetic dataset was modeled by means of Seismic Unix.

REFERENCES

- Bauer, A., Schwarz, B., and Gajewski, D. (2017). Utilizing diffractions in wavefront tomography. *Geophysics*, 82(2):R65–R73.
- Baykulov, M. and Gajewski, D. (2009). Prestack seismic data enhancement with partial common-reflection-surface (CRS) stack. *Geophysics*, 74(3):V49–V58.
- Bellman, R. (1997). *Introduction to matrix analysis*. Society for Industrial and Applied Mathematics, 2nd edition.
- Byrd, R., Lu, P., and Nocedal, J. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208.
- Cervený, V. (2001). Seismic ray theory. Cambridge university press.
- Chavent, G. (2010). Nonlinear least squares for inverse problems: theoretical foundations and step-by-step guide for applications. Springer Science & Business Media.
- Costa, J., Silva, F., Gomes, E., Schleicher, J., Melo, L., and Amazonas, D. (2008). Regularization in slope tomography. *Geophysics*, 73(5):VE39–VE47.
- Dell, S. and Gajewski, D. (2011). Common-reflection-surface-based workflow for diffraction imaging. *Geophysics*, 76(5):S187–S195.
- Dell, S., Gajewski, D., and Tygel, M. (2014). Image-ray Tomography. *Geophysical Prospecting*, 62(3):413–426.
- Duveneck, E. (2004). Velocity model estimation with data-derived wavefront attributes. *Geophysics*, 69(1):265–274.
- Farra, V. and Madariaga, R. (1987). Seismic waveform modeling in heterogeneous media by ray perturbation theory. *Journal of Geophysical Research*, 92(B3):2697–2712.
- Fomel, S. and Kazinnik, R. (2013). Non-hyperbolic common reflection surface. *Geophysical Prospecting*, 61(1):21–27.
- Gelchinsky, B., Berkovitch, A., and Keydar, S. (1999). Multifocusing homeomorphic imaging, Part 1: Basic concepts and formulae. *Journal of Applied Geophysics*, 42(3):229–242.
- Gilbert, F. and Backus, G. (1966). Propagator matrices in elastic wave and vibration problems. *Geophysics*, 31(2):326–332.
- Hansen, P. (1998). *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. Society for Industrial and Applied Mathematics.
- Hubral, P. (1983). Computing true amplitude reflections in a laterally inhomogeneous earth. *Geophysics*, 48(8):1051–1062.
- Jäger, R., Mann, J., Höcht, G., and Hubral, P. (2001). Common-reflection-surface stack: Image and attributes. *Geophysics*, 66(1):97–109.

Lambaré, G. (2008). Stereotomography. Geophysics, 73(5):VE25-VE34.

- Mann, J., Jäger, R., Müller, T., Höcht, G., and Hubral, P. (1999). Common-reflection-surface stack a real data example. *Journal of Applied Geophysics*, 42(3-4):283–300.
- Mayne, W. (1962). Common reflection point horizontal data stacking techniques. *Geophysics*, 27(6):927–938.
- Plessix, R. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503.
- Schwarz, B., Bauer, A., and Gajewski, D. (2016). Passive seismic source localization via common-reflection-surface attributes. *Studia Geophysica et Geodaetica*, 60(3):531–546.
- Schwarz, B., Vanelle, C., Gajewski, D., and Kashtan, B. (2014). Curvatures and inhomogeneities: An improved common-reflection-surface approach. *Geophysics*, 79(5):S231–S240.