

# Tomographic Deconvolution of Reflection Tomograms

Tushar Gautam<sup>1</sup>, Yicheng Zhou<sup>1</sup>, Shihang Feng<sup>2</sup> and Gerard T. Schuster<sup>1</sup>

<sup>1</sup>King Abdullah University of Science and Technology (KAUST), Saudi Arabia

<sup>2</sup>Los Alamos National Laboratory, USA

## SUMMARY

We present a tomographic deconvolution procedure for high-resolution imaging of velocity anomalies between reflecting interfaces. The key idea is to first invert reflection or transmission traveltimes for the background velocity model. A convolutional neural network (CNN) model is then trained to estimate the inverse to the blurred tomogram consisting of small scatterers in the background velocity model. We call this CNN a tomographic deconvolution operator because it deconvolves the blurring artifacts in traveltimes tomograms. This procedure is similar to that of migration deconvolution which deconvolves the migration butterfly artifacts in migration images. Results with synthetic examples show the effectiveness of this procedure in significantly sharpening the tomographic images of small scatterers.

## INTRODUCTION

The seismic forward modeling operator  $\mathbf{L}$  computes the reflection traces  $\mathbf{d}$  from the slowness model  $\mathbf{m}$  by

$$\mathbf{L}\mathbf{m} = \mathbf{d}, \quad (1)$$

where  $\mathbf{m} = \mathbf{m}^{scatt} + \mathbf{m}^{smooth}$  is composed of the low-wavenumber smooth velocity model  $\mathbf{m}^{smooth}$  and the high-wavenumber velocity variations  $\mathbf{m}^{scatt}$ . In the context of seismic migration or traveltimes tomography, equation 1 is an overdetermined and inconsistent set of equations.

To reduce the artifacts associated with inconsistent and overdetermined equations, we often compute the least squares solution given by (Yilmaz, 2001; Schuster, 2017)

$$\begin{aligned} \mathbf{m}^* &= [\mathbf{L}^T\mathbf{L}]^{-1}\mathbf{L}^T\mathbf{d}, \\ &= \overbrace{[\mathbf{L}^T\mathbf{L}]^{-1}\mathbf{L}^T\mathcal{L}}^{\text{tomographic PSR}} \sum_i \mathbf{m}_i. \end{aligned} \quad (2)$$

where a regularization term is often used to stabilize the solution and  $\mathcal{L}$  is the Earth's actual modeling operator. Here, we assume that  $\mathbf{d}$  can be decomposed into a summation of modeled slowness anomalies

$$\mathbf{d} = \sum_i \mathbf{L}\mathbf{m}_i, \quad (3)$$

where  $\mathbf{m}_i$  is the  $i^{th}$  localized slowness anomaly. These anomalies can be combined together to form a larger slowness anomaly  $\mathbf{m} = \sum_i \mathbf{m}_i$ .

The Hessian inverse  $[\mathbf{L}^T\mathbf{L}]^{-1}$  is often too expensive to compute so an iterative gradient method is used to estimate the solution:

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} - \alpha \mathbf{C}\mathbf{L}^T \overbrace{\{\mathbf{L}\mathbf{m}^{(k)} - \mathbf{d}^{obs}\}}^{\text{residual}=\mathbf{r}^{(k)}}, \quad (4)$$

where  $k$  is the iteration index, the step length is  $\alpha$ , and  $\mathbf{C}$  is the preconditioning operator that is sometimes chosen to be a diagonal matrix where  $\mathbf{C}_{ii} = [\mathbf{L}^T\mathbf{L}]_{ii}^{-1}$ . In practice, either a preconditioned conjugate gradient or a quasi-Newton method is often used where for large  $k = K$  that  $\mathbf{m}^K = \mathbf{m}^*$  in equation 2.

In the context of traveltimes tomography,  $\mathbf{m}$  is the slowness model and  $\mathbf{d}$  defines the picked traveltimes of transmission and/or reflection traveltimes. The operator  $\mathbf{L}$  represents the forward modeling of traveltimes by either ray tracing or a numerical solution to the wave equation. Despite the efforts to stabilize and precondition the gradient methodology, the final tomogram is always filled with artifacts associated with either incomplete data and/or physics. Such artifacts are often caused by, for example, a limited source-receiver aperture, uneven density of rays in each slowness cell, strong velocity contrasts that prevent many rays from entering areas of interest, and an irregular and coarse distribution of sources and receivers.

This paper proposes to apply a non-linear deconvolution filter  $F$  that can suppress many of these artifacts in the slowness tomogram  $\mathbf{m}^*$  in equation 2, i.e.

$$\mathbf{m} \approx F(\mathbf{m}^*), \quad (5)$$

where  $\mathbf{m}$  is the actual slowness model and the deconvolved tomogram  $F(\mathbf{m}^*)$  is a more accurate rendering of  $\mathbf{m}$  than  $\mathbf{m}^*$ . According to equation 2, the deconvolution filter in equation 5 is given by

$$F = ([\mathbf{L}^T\mathbf{L}]^{-1}\mathbf{L}^T\mathcal{L})^{-1}, \quad (6)$$

where  $\mathcal{L}$  can be roughly approximated by the computer-based operator  $\mathbf{L}$ . The deconvolution filter  $F(\cdot)$  is also denoted as the inverse to the tomographic point scatterer response (PSR). This filter is computed by training a convolution neural network (CNN), where the input is the tomogram  $\mathbf{m}^{tomo}$  and the output is the sum of localized slownesses  $\mathbf{m} \approx \sum_i \mathbf{m}_i$ . Here, the index  $i$  indicates different locations of the localized anomalies  $\mathbf{m}_i$ .

The first part of this paper provides the theory of tomographic deconvolution. The fundamental concept of the tomographic point scatterer response (PSR) is defined, and then its inverse is obtained using a neural network model. A convolutional neural network is used and the workflow for its implementation is described. The next section presents the numerical results for synthetic data. The last section presents the conclusions.

## WAVEPATHS AND WORKFLOW FOR TOMOGRAPHIC DECONVOLUTION

The first Fresnel zones for transmission and reflection arrivals are illustrated in Figure 1b for a single source-receiver pair. These zones are defined as the points visited by rays that can be traced from the source to the receiver with a traveltime greater than or equal to the central ray traveltime at time  $t$  and less than  $t + T_0/2$ . Here,  $T_0$  is the dominant period of the source (Schuster, 2017). These first Fresnel zones are also denoted as wavepaths (Woodward, 1989, 1992; Luo, 1991).

For the tomographic deconvolution problem, we assume that the location of the reflector and the background velocity is known, and the unknown is the location and extent of the green velocity anomaly in Figure 1b. This unknown anomaly is the only part of the model that generates a traveltime residual in the data. For a single source-receiver pair, this means that the velocity is updated all along the rabbit-ear wavepath in Figure 1b because this is the zone which largely influences the phase of the reflection arrival. The green velocity anomaly is at the intersection of the rabbit ears on the reflector so that the resulting slowness tomogram  $\mathbf{m}^{(k+1)} = \mathbf{m}^{(k+1)} - \alpha \mathbf{L}^T \Delta \mathbf{r}^{(k)}$  will take the appearance of a pair of rabbit ears. This is what we call the *point scatterer response* of the tomographic operator for a single traveltime residual. The shape of this response will change depending on the location of the scatterer relative to the reflector boundary and the number of traces as illustrated in Figures 2a-2b. Figure 2c is the ideal deconvolved tomogram  $F(\mathbf{m}^{tomo})$  in equation 5, where the rabbit-ear artifacts are removed by the deconvolution operator  $F$  to obtain the actual shape of the scatterer.

### Tomographic Deconvolution Workflow

The tomographic deconvolution workflow consists of 5 steps.

- Collect traveltime reflection or refraction data and invert it by traveltime tomography to get the smooth background slowness model  $\mathbf{m}^{tomo}$ .
- Train the CNN model with synthetic reflection or refraction data to find  $F(\mathbf{m}^{tomo})$  in Figure 2b. The background slowness model is obtained by smoothing  $\mathbf{m}^{tomo}$ . The input to the training consist of small 'point' slowness anomalies located at different positions in the model. The size of these point anomalies is no smaller than a wavelength if ray tracing is used for forward modeling.
- Once the CNN model is properly trained and validated, it is applied to the tomogram  $\mathbf{m}^{tomo}$  to get  $\mathbf{m}^{decon} = F(\mathbf{m}^{tomo})$ , where the tomographic artifacts are suppressed and  $\mathbf{m} \approx \mathbf{m}^{decon}$ .
- Predicted data are computed  $\mathbf{d}^{pred} = \mathbf{L}\mathbf{m}^{decon}$  as a sanity test is used to ensure that  $\mathbf{m}^{decon}$  accurately predicts the recorded traveltime data  $\mathbf{d}$ .

### CNN Architecture

The CNN model is that of a U-Net (Schuster, 2021) and consists of a contracting path and an expansive path (which gives it a U-shaped architecture). The training of the CNN model

uses synthetic reflection traveltime data computed by a reflection ray-tracing code, and then inverted by an iterative gradient method that minimizes the  $L^1$  data misfit function. About ten thousand training pairs  $(\mathbf{m}^{tomo}, \mathbf{m})$  are used to train the CNN. Each training model consists of several localized slowness anomalies with a different location for each anomaly.

## NUMERICAL RESULTS

Numerical results are now presented for deconvolving tomograms by the tomographic deconvolution operator  $F$  computed by a CNN. The reflection traveltime data are computed by reflection ray tracing in a two-layer model with embedded slowness anomalies. A source and receiver spacing of 10 m is used, and there are 199 sources located on the surface. Each source shoots into 200 receivers so 39,800 traveltimes are generated for each simulation. It is assumed that the background medium is a two-layer model with a homogeneous velocity in each layer, and the depth of the horizontal reflector is known. The localized slowness anomalies have a small contrast in velocity so ray bending is negligible, and the training is computed for localized slowness anomalies located randomly between the reflecting interface and the recording surface.

The CNN trained on 10,000 pairs of slowness tomograms and deconvolved tomograms. The training dataset contains multiple anomalies at random locations. The close comparison (not shown) between the loss functions for the training and validation sets suggest that the model is well trained without overfitting. After training, the CNN is tested on the out-of-the-training set models shown in the first column of images in Figure 3. The deconvolved tomograms in the third column suggest that tomographic deconvolution can significantly improve the resolution of the slowness tomograms in the middle column of images. Many other examples (not shown) are tested and the all of the results show that the deconvolved tomograms provide significant improvement in resolution compared to the original tomograms.

### Models that Break the Deconvolution Filter

It is well known that the best resolution of a raypath is perpendicular to the ray (Schuster, 2017). Therefore slowness variations in depth are not well resolved for reflection data with narrow incident angles. This is illustrated in Figure 4 where two oscillatory interfaces at different depths in Figures 4a and 4d give about the same traveltime data. This means that there are many models that satisfy nearly the same data so there is not a unique solution to the traveltime equations, as illustrated by the tomograms in Figures 4b and 4e. Thus, it is not surprising that the deconvolution operator gives poor results in Figures 4c and 4f.

## CONCLUSIONS

We presented the theory of tomographic deconvolution where the tomogram is deconvolved of artifacts to give a high resolution estimate of the localized velocity anomalies in a reflecting layer. The assumptions are that ray bending can be neglected in

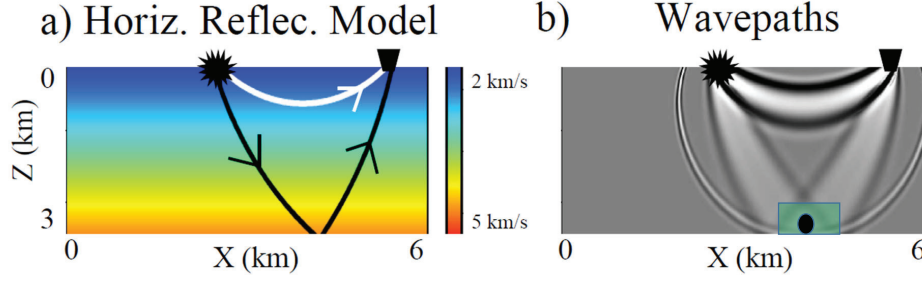


Figure 1: a) Horizontal reflector model along with the white direct and black reflection raypaths. b) Specular reflection and transmission wavepaths obtained by reverse time migration of a trace recorded at the geophone (quadrilateral) and excited by a bandlimited source (star). The gray-shaded migration ellipse has foci at the source and receiver. Figures courtesy of Ge Zhan.

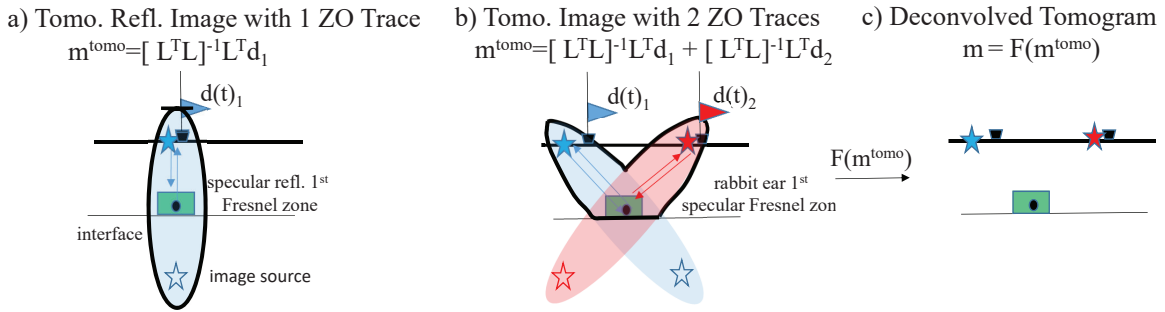


Figure 2: Tomographic models  $\mathbf{m}^{tomo} = [\mathbf{L}^T \mathbf{L}]^{-1} \mathbf{L}^T \mathbf{d}$  inverted from specular reflection traveltimes picked from a) one and b) two ZO traces. The slowness model is updated by smearing the traveltimes residuals to be within the Fresnel zone bounded by the thick black line in b) that forms the rabbit ears. c) is the deconvolved tomogram  $F(\mathbf{m}^{tomo})$ .

the slowness anomalies and that the anomalies are no smaller than a wavelength in order to satisfy the high-frequency approximation. To go beyond this approximation we should use a wave equation tomography method. If the model with the anomalies is in the null space of  $[\mathbf{L}^T \mathbf{L}]$  then this method will not be able to accurately recover these components. A sanity test is recommended after obtaining the deconvolved tomogram. As an example, the final deconvolved tomogram should be used to generate synthetic traveltimes, which can be compared to the recorded ones to ensure that the  $L^2$  residual of the deconvolved tomogram is equal to or less than that from the input tomogram.

The computational requirements for training the CNN are high in this feasibility study, but it is straightforward to reduce costs by reusing the forward modeling Green's functions as described by the generalized migration procedure in Schuster (2017). Other short cuts can be used as well such as transfer training. Tomographic deconvolution can also be used for transmission tomography and refraction tomography.

## ACKNOWLEDGEMENTS

The research reported in this publication was supported by the King Abdullah University of Science and Technology (KAUST)

in Thuwal, Saudi Arabia. We are grateful to the sponsors of the Center for Subsurface Imaging and Modeling Consortium for their financial support. For computer time, this research used the resources of the Supercomputing Laboratory at KAUST and the IT Research Computing Group. We thank them for providing the computational resources required for carrying out this work.

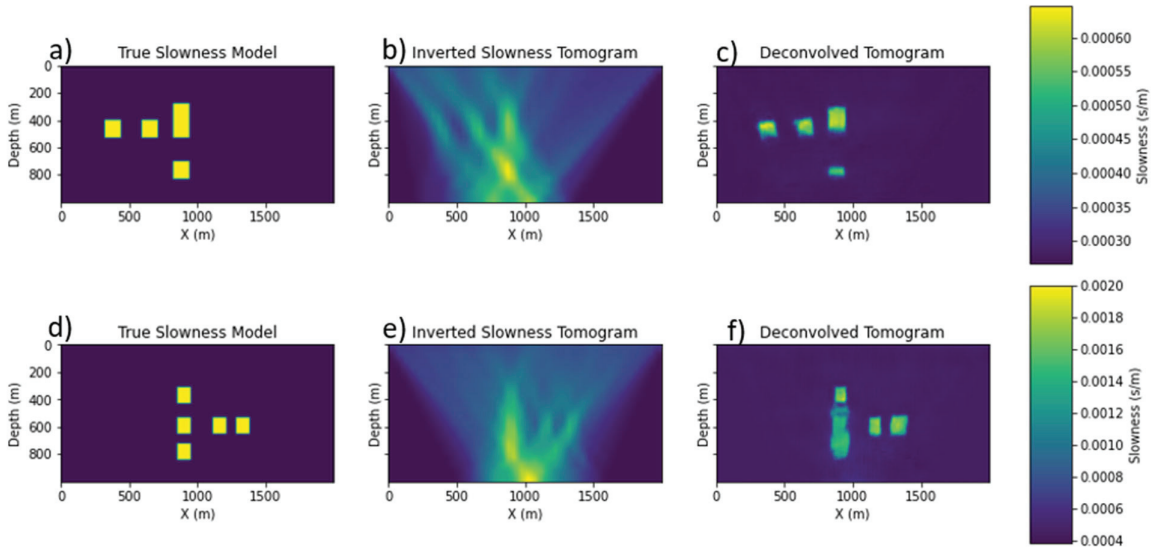


Figure 3: First column depicts the actual slowness models, the second column show the traditional slowness tomograms and the third column shows the deconvolved tomograms.

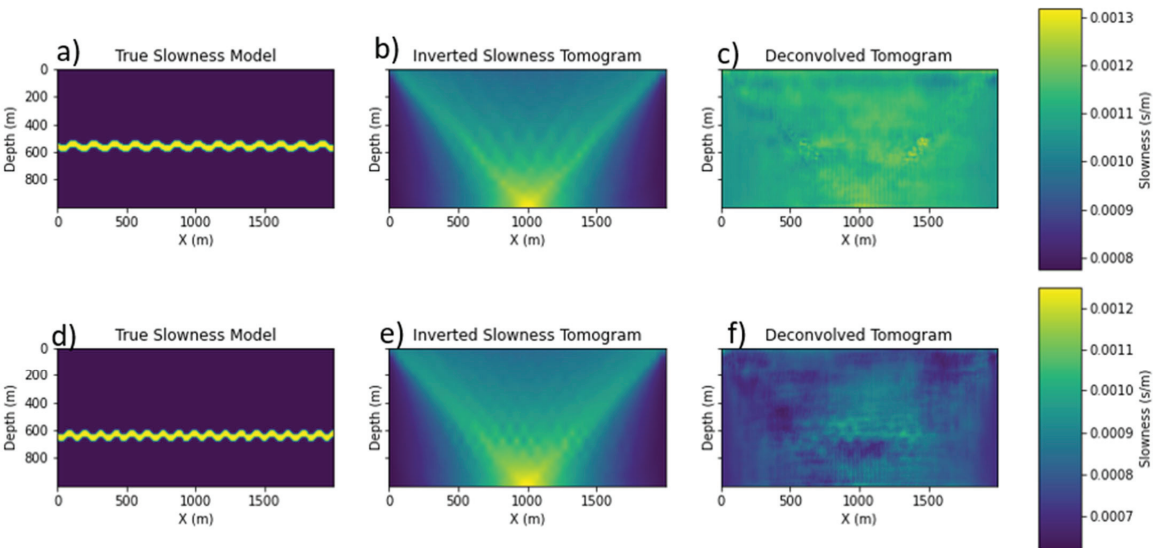


Figure 4: Same as the previous figure except the input slowness models consists of oscillatory low-velocity layers at different depths.