

Deep learning decomposition for null and active space estimation for thin-bed reflectivity inversion

Kristian Torres*, Mauricio D. Sacchi, University of Alberta

SUMMARY

We investigate deep decomposition learning for solving ill-posed seismic inverse problems. We pay particular attention to the solution of one ubiquitous problem in seismic exploration: the recovery of a full-band reflectivity from band-limited seismic traces. The proposed approach combines classical regularization theory with a learned deep decomposition framework. The method extends the popular learned post-processing approach by learning how to improve an initial reconstruction with estimated missing components from the null space of the forward operator, which in our case, are the missing frequency components of the reflectivity. Numerical experiments show that the proposed method naturally enforces a high-resolution prediction consistent with the low-resolution input seismic traces. We also compare the proposed technique with a classical thin-bed reflectivity estimation method on a real data set.

INTRODUCTION

This paper considers the computation of an approximate solution for seismic inverse problems of the form

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

where $\mathbf{L} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ represents a linear forward operator that maps an earth model or unknown signal $\mathbf{m} \in \mathbb{R}^n$ to a data vector $\mathbf{d} \in \mathbb{R}^m$, and ε denotes an unknown data error (i.e., the noise). The operator, \mathbf{L} , contains the seismic wavelet properly arranged into a Toeplitz matrix to represent convolution via simple matrix-times-vector multiplication. The sought vector of model parameters \mathbf{m} is the reflectivity, and \mathbf{d} contains the seismic traces. We are interested in recovering full-band reflectivity series, but bear in mind that the framework presented in this paper applies to other inverse problems associated with seismic data processing and imaging.

Geophysical inverse problems are ill-posed due to the presence of a non-trivial null space, which in our case, is a consequence of the lack of low and high frequencies of the seismic wavelet. Given the existence of a non-trivial null space, many solutions fit the acquired data equally well, and a direct inversion $\mathbf{m} = \mathbf{L}^{-1}\mathbf{d}_\varepsilon$ to obtain the sought-after model is impossible.

Regularization techniques are generally implemented in traditional model-based methods to reduce null space ambiguity by imposing prior information that permits to find a plausible solution. Many methods have been proposed for high-resolution, sparse-spike or thin-bed reflectivity inversion, including techniques that use l_1 sparsity constraints to estimate full-band reflectivity estimators (Taylor et al., 1979; Oldenburg et al., 1983; Debeye and Van Riel, 1990; Sacchi, 1997; Chopra et al., 2006; Zhang and Castagna, 2011).

Recently, several seismic applications have explored supervised techniques using Convolutional Neural Networks (CNNs), either as direct learned inversion (end-to-end) approaches (Araya-Polo et al., 2018; Mandelli et al., 2019), as learned iterative schemes (Torres and Sacchi, 2021), or as learned post-processing of an initial reconstruction (Kaur et al., 2020). Even though these methods have demonstrated remarkable empirical success, many supervised approaches still lack a data consistency restraint to enforce that the predicted model matches the acquired data, a necessary condition for a reliable solution to the inverse problem. Hence, the results might look realistic, but there is no way to assess their accuracy. Unsupervised approaches (Zhang et al., 2021; Kong et al., 2022) address this issue by design, but they often amplify the expensive iterative nature of traditional methods when the trainable weights are not correctly initialized.

Regularization via null space networks (Schwab et al., 2019) has been introduced as an alternative to the learned post-processing method to account for data consistency. By computing a projection onto the null space of the forward operator after the last weight layer of a residual architecture, it is possible to train a neural network to learn the missing components of the initial reconstruction given by the pseudoinverse $\mathbf{L}^\dagger : \mathbb{R}^m \rightarrow \mathbb{R}^n$ (or an approximation to it). As a generalization to null space networks, Chen et al. (2021) introduced the concept of deep decomposition learning, which attaches a complementary network to act as a denoiser on the range of the pseudoinverse. Based on this idea, we investigate the extension of deep null space learning on the inversion of reflectivity. Specifically, we extend the deep decomposition approach by using the truncated singular value decomposition (TSVD) as an initial regularized reconstruction for approximating the low-frequency components of the model. In a second step, we trained two neural networks to recover the missing parts of the model and the "inverted" noise, respectively.

METHOD

We can decompose the domain of the forward operator into two sub-spaces: the measurement space and the null space. Accordingly, we might think of any model \mathbf{m} in the domain of \mathbf{L} as being made up of two unique orthogonal vectors,

$$\mathbf{m} = \mathbf{m}_R + \mathbf{m}_N = P_R(\mathbf{m}) + P_N(\mathbf{m}), \quad (2)$$

such that \mathbf{m}_R lies in the range of the pseudoinverse \mathbf{L}^\dagger which is also the active space solution, and \mathbf{m}_N lies in the null space. By definition, these two components satisfy, respectively,

$$\mathbf{m}_R = \mathbf{L}^\dagger \mathbf{d}_\varepsilon = \mathbf{L}^\dagger \mathbf{L}\mathbf{m} + \mathbf{L}^\dagger \varepsilon, \quad (3)$$

and

$$\mathbf{L}\mathbf{m}_N = 0. \quad (4)$$

Deep decomposition

Based on this fragmentation of the model, we can express the ideal reconstruction as

$$\mathbf{m}^* = \mathbf{L}^\dagger \mathbf{d}_\varepsilon - \mathbf{L}^\dagger \varepsilon + \mathbf{m}_N. \quad (5)$$

In other words, the solution is expressed in terms of a unique minimum norm least-squares solution ($\mathbf{L}^\dagger \mathbf{d}_\varepsilon$) minus the "inverted" noise plus the null space vector.

As denoted in equation 2, the model components can be obtained from two orthogonal projections, P_R and P_N , defined as

$$P_R = \mathbf{L}^\dagger \mathbf{L}, \quad (6)$$

and

$$P_N = \mathbf{I} - \mathbf{L}^\dagger \mathbf{L}, \quad (7)$$

where $\mathbf{I} \in \mathbb{R}^n$ is the identity operator. Using a physics-engaged approach promoted by the application of the above-mentioned orthogonal projections, deep decomposition learning attempts to solve equation 5 with a trained estimator $\Lambda : \mathbb{R}^m \rightarrow \mathbb{R}^n$ defined as

$$\begin{aligned} \Lambda(\mathbf{d}_\varepsilon; \theta_1, \theta_2) &= \mathbf{L}^\dagger \mathbf{d}_\varepsilon + P_R \circ \mathbf{F}_{\theta_1} \circ \mathbf{L}^\dagger \mathbf{d}_\varepsilon \\ &\quad + P_N \circ \mathbf{N}_{\theta_2} \circ (\mathbf{L}^\dagger \mathbf{d}_\varepsilon + P_R \circ \mathbf{F}_{\theta_1} \circ \mathbf{L}^\dagger \mathbf{d}_\varepsilon), \end{aligned} \quad (8)$$

where \mathbf{F}_{θ_1} and \mathbf{N}_{θ_2} are two trainable neural networks. Compared to equation 5, it is clear that the second term in the right-hand side of equation 8 tries to estimate the negative "inverted" noise by projecting the output of the network \mathbf{F}_{θ_1} onto the range of the pseudoinverse. Likewise, the third term in equation 8 tries to estimate the null space component from the denoised input $\mathbf{L}^\dagger \mathbf{d}_\varepsilon + P_R \circ \mathbf{F}_{\theta_1} \circ \mathbf{L}^\dagger \mathbf{d}_\varepsilon \approx \mathbf{L}^\dagger \mathbf{d}_\varepsilon - \mathbf{L}^\dagger \varepsilon$. When Λ lacks the explicit denoising element ($P_R \circ \mathbf{F}_{\theta_1} \circ \mathbf{L}^\dagger \mathbf{d}_\varepsilon = 0$), the estimator turns into a standard null space regularization network of the form

$$\begin{aligned} \Lambda(\mathbf{d}_\varepsilon; \theta) &= \mathbf{L}^\dagger \mathbf{d}_\varepsilon + P_N \circ \mathbf{N}_\theta \circ \mathbf{L}^\dagger \mathbf{d}_\varepsilon \\ &= (\mathbf{I} + P_N \circ \mathbf{N}_\theta)(\mathbf{L}^\dagger \mathbf{d}_\varepsilon), \end{aligned} \quad (9)$$

where the data consistency property, $\mathbf{L}\Lambda(\mathbf{d}_\varepsilon; \theta) = \mathbf{d}_\varepsilon$ is preserved.

We use TSVD for computing a regularized approximation \mathbf{L}_k^\dagger to the exact pseudoinverse, where the regularization parameter is given by the number of singular values $k > 0$ and is chosen such that $\mathbf{m}_{\text{TSVD}}^* = \mathbf{L}_k^\dagger \mathbf{d}_\varepsilon$ has the minimal l_2 difference to the true model. The solution obtained by TSVD is the classical regularized least-square solution that will yield a reflectivity of spectral properties similar to the original data and not significant gain in bandwidth. By only recovering signal components corresponding to sufficiently large singular values, TSVD produces a stable solution that prevents small singular values of \mathbf{L} to amplify noise. The absent components (frequencies in the null space of the operator) have an unreliable model-to-data mapping and, therefore, will be recovered in the learning stage.

Finally, the overall procedure aims to jointly seek

$$\begin{aligned} \arg \min_{\theta_1, \theta_2} \frac{1}{N} \sum_{i=1}^N & \| \mathbf{m}^i - \Lambda(\mathbf{d}_\varepsilon^i; \theta_1, \theta_2) \|_2^2 + \\ & \lambda_1 \sum_{i=1}^N \| \mathbf{L} \mathbf{F}_{\theta_1}(\mathbf{L}_k^\dagger \mathbf{d}_\varepsilon^i) - \varepsilon^i \|_2^2 + \lambda_2 \| \theta_2 \|_2^2, \end{aligned} \quad (10)$$

where the first term carries out the supervised training on a synthetic dataset $\mathcal{D} = \{(\mathbf{m}^i, \mathbf{d}_\varepsilon^i)\}_{i=1}^N$ using the MSE loss, the second term prevents the denoising component from breaking the data consistency property, and the third term provides the null space estimator with robustness to small perturbations via weight regularization (Schwab et al., 2019). In this work, we define $\mathbf{F}_{\theta_1} = \sigma_4 \circ \mathbf{W}_4 \circ \dots \circ \sigma_1 \mathbf{W}_1$ as a four-layered CNN denoising network where σ_i represents the ReLU function, and \mathbf{W}_i is a matrix with trainable weights. We set \mathbf{N}_{θ_2} as the original U-net architecture following Ronneberger et al. (2015).

Reflectivity Inversion

We evaluate the learned regularization method described in the previous section on synthetic and real data. Setting the trade-off parameters $\lambda_1 = 10e^{-6}$ and $\lambda_2 = 10e^{-8}$ yields stable solutions for the two experiments consider in this paper. In both experiments, we incorporate different realizations of additive Gaussian noise to the clean data such that the signal-to-noise ratio is equal to 20%. In all cases, equation 10 was minimized using stochastic gradient descent with 400 epochs and a learning rate of 0.001.

Relying on the convolutional model of the seismic trace, we attempt to recover a full-band reflectivity sequence from a zero-offset trace. The forward operator represents a band-limited wavelet matrix expressing time-invariant convolution in this example. To train the estimator, we randomly generate 5000 reflectivity samples and obtain the corresponding data by convolving a 60 Hz Ricker wavelet and adding the random noise. Figure 1 shows the results for a test reflectivity model different from all training samples but generated using the same random procedure. For a quantitative evaluation of the results, we calculate the reconstruction accuracy (dB) = $10 \times \log_{10} \| \mathbf{m} \|_2^2 / \| \mathbf{m} - \mathbf{m}^* \|_2^2$, where \mathbf{m} and \mathbf{m}^* are the true and inverted reflectivity models, respectively. Figure 1b shows the initial TSVD solution, where some of the most prominent reflections are partially recovered, but a residual band-limited component masks the non-resolvable events. Moreover, oscillations are visible in the result due to the truncated expansion. For comparison, Figure 1c displays a sparse-spike deconvolution result obtained with 100 iterations of the FISTA solver. Even though the inversion with sparsity promotion enables a full-band solution, the noisy data impedes successfully retrieving the correct amplitude and positioning of some events. Figure 1d shows the result obtained with deep null space regularization, where we conclude that, both visually and quantitatively, the learned estimator inversion produces higher quality results compared to the two previous techniques. Finally, Figures 1e and 1f show the amplitude spectrum of the true and the estimated result and the amplitude spectrum of the solution's range and null space components. Both figures show that deep null space regularization approximately recovers the missing frequency components of the original signal.

Field data example

We also adopted the trained network to estimate full-band reflectivity from a real dataset. The network was trained with 2D pseudo-random reflectivity models similar to Figure 2. Results are portrayed in Figure 3, where for completeness, we

Deep decomposition

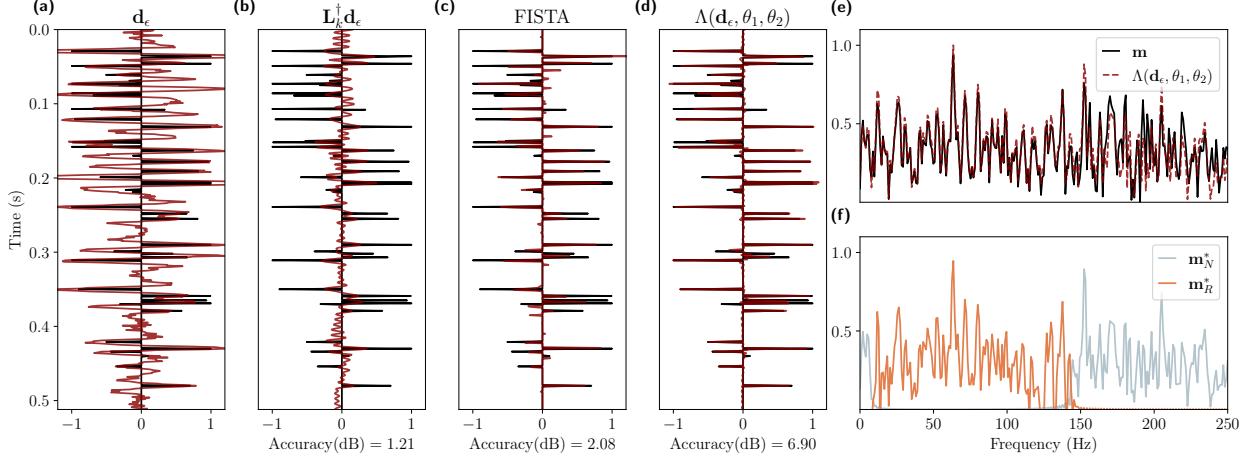


Figure 1: Deconvolution example. True reflectivity \mathbf{m} superposed with (a) the noisy trace \mathbf{d}_ϵ , (b) the TSVD result, (c) the FISTA result, and (d) the deep null space regularization. Figure (e) compares the amplitude spectrum of the true reflectivity model and the deep null space regularization solution, and Figure (f) shows the amplitude spectrum of the individual range (orange) and null space (gray) retrieved components.

also compare the proposed method to the output of a thin-bed reflectivity inversion (Chopra et al., 2006).

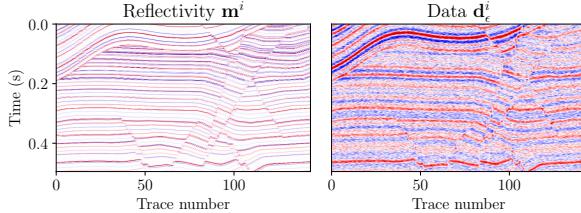


Figure 2: One realization of reflectivity model and data used for training the network to perform 2D thin-bed reflectivity inversion.

Figure 4 compares the average power spectral density of the inverted reflectivity and the data. Both the proposed method and thin-bed reflectivity inversion yield broadband solutions. Absence of sonic logs in the area has precluded the incorporation of additional information into the network training process.

CONCLUSIONS

In this work, we investigated the use of regularizing networks via the deep decomposition approach to estimate full-band reflectivity from band-limited seismic data. We illustrate how learned null space regularization adds reasonable estimates from the null space, improving classic regularization solutions. Additionally, we combined the deep decomposition learning method with TSVD, which helps produce clean inputs for the efficient training of the null space network. An important research direction is to identify ways of efficiently integrating deep null space regularization with bigger problems where the direct computation of the orthogonal projections is prohibitively

expensive.

ACKNOWLEDGEMENTS

The research leading to this paper was supported by the Signal Analysis and Imaging Group (SAIG) sponsors at the University of Alberta. We thank ARCIS-TGS for providing the field data example. We also thank Satinder Chopra for providing constructive feedback about our work.

Deep decomposition

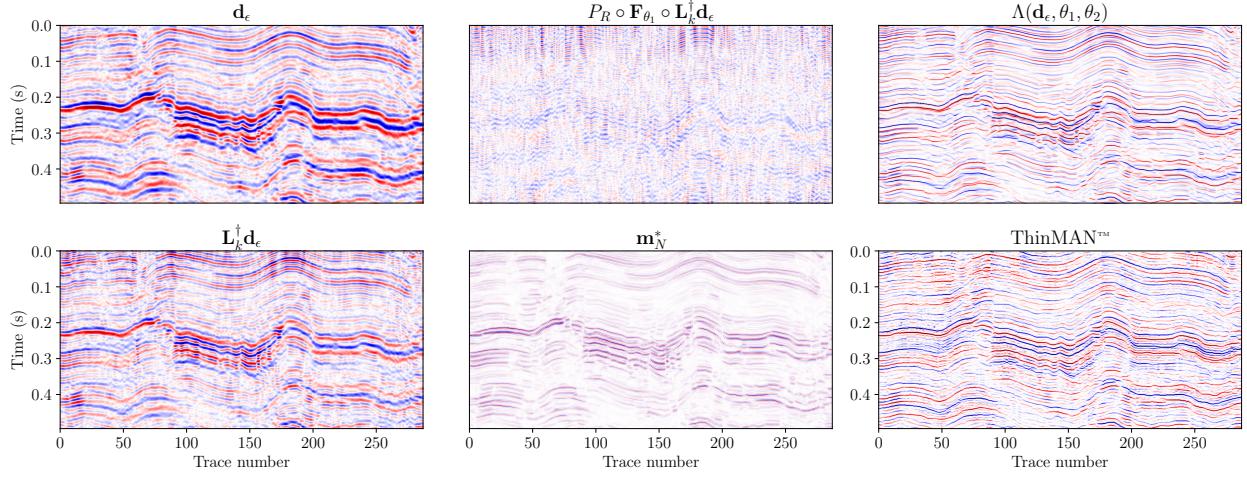


Figure 3: First column: Field data example portraying the input band-limited data \mathbf{d}_ϵ and the band-limited solution obtained via the truncated SVD $\mathbf{L}_k^\dagger \mathbf{d}_\epsilon$. Central column: Noise leaking into solution through the pseudo-inverse and discovered null space component (\mathbf{m}_N^*). Third column inverted reflectivity by the proposed method and classical thin-bed reflectivity inversion via ThinMAN.

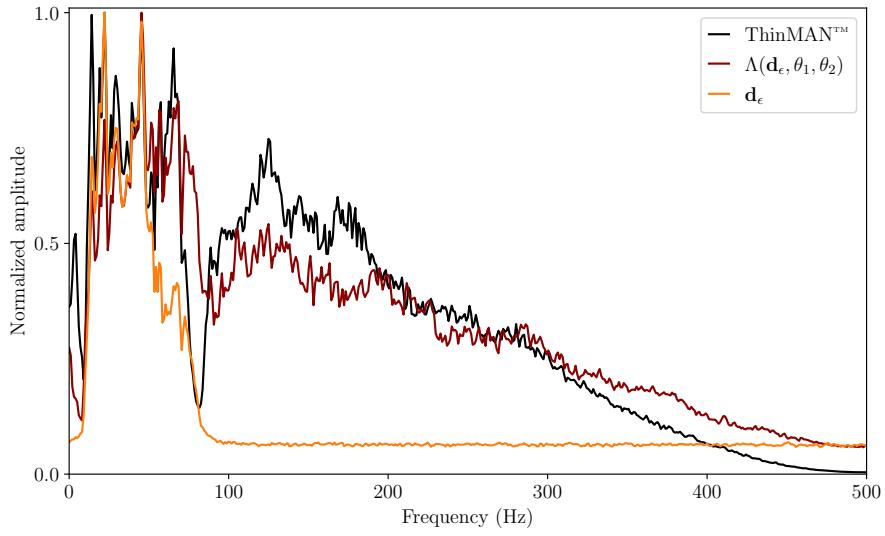


Figure 4: Normalized average power spectral density for the field data example. Deep decomposition learning solution $\Lambda(\mathbf{d}_\epsilon, \theta_1, \theta_2)$ for real data thin-bed reflectivity estimation. The resulting thin-bed reflectivity is compared with the ThinMAN solution (Chopra et al., 2006).

Deep decomposition

REFERENCES

- Araya-Polo, M., J. Jennings, A. Adler, and T. Dahlke, 2018, Deep-learning tomography: The Leading Edge, **37**, 58–66.
- Chen, D., J. Tachella, and M. E. Davies, 2021, Equivariant imaging: Learning beyond the range space: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 4379–4388.
- Chopra, S., J. Castagna, and O. Portniaguine, 2006, Thin-bed reflectivity inversion: SEG Technical Program Expanded Abstracts 2006, 2057–2061.
- Debeye, H. W. J., and P. Van Riel, 1990, L_p-norm deconvolution: Geophysical Prospecting, **38**, 381–403.
- Kaur, H., N. Pham, and S. Fomel, 2020, Improving the resolution of migrated images by approximating the inverse hessian using deep learning: GEOPHYSICS, **85**, WA173–WA183.
- Kong, F., F. Picetti, V. Lipari, P. Bestagini, X. Tang, and S. Tubaro, 2022, Deep prior-based unsupervised reconstruction of irregularly sampled seismic data: IEEE Geoscience and Remote Sensing Letters, **19**, 1–5.
- Mandelli, S., F. Borra, V. Lipari, P. Bestagini, A. Sarti, and S. Tubaro, 2019, Seismic data interpolation through convolutional autoencoder: 2018 SEG International Exposition and Annual Meeting, SEG 2018, 4101–4105.
- Oldenburg, D. W., T. Scheuer, and S. Levy, 1983, Recovery of the acoustic impedance from reflection seismograms: GEOPHYSICS, **48**, 1318–1337.
- Ronneberger, O., P. Fischer, and T. Brox, 2015, U-net: Convolutional networks for biomedical image segmentation: CoRR, [abs/1505.04597](https://arxiv.org/abs/1505.04597).
- Sacchi, M. D., 1997, Reweighting strategies in seismic deconvolution: Geophysical Journal International, **129**, 651–656.
- Schwab, J., S. Antholzer, and M. Haltmeier, 2019, Deep null space learning for inverse problems: Convergence analysis and rates: Inverse Problems, **35**.
- Taylor, H. L., S. C. Banks, and J. F. McCoy, 1979, Deconvolution with the l1 norm: Geophysics, **44**, 39–52.
- Torres, K., and M. Sacchi, 2021, Deep learning based least-squares reverse-time migration: First International Meeting for Applied Geoscience & Energy Expanded Abstracts, 2709–2713.
- Zhang, R., and J. Castagna, 2011, Seismic sparse-layer reflectivity inversion using basis pursuit decomposition: Geophysics, **76**, R147–R158.
- Zhang, W., J. Gao, X. Jiang, and W. Sun, 2021, Consistent least-squares reverse time migration using convolutional neural networks: IEEE Transactions on Geoscience and Remote Sensing, 1–18.