

Improved impedance inversion by deep learning and iterated graph Laplacian

Davide Bianchi*

Florian Boßmann[†]Wenlong Wang[†]Mingming Liu[†]

Abstract. Deep learning techniques have shown significant potential in many applications through recent years. The achieved results often outperform traditional techniques. However, the quality of a neural network highly depends on the used training data. Noisy, insufficient, or biased training data leads to suboptimal results.

We present a hybrid method that combines deep learning with iterated graph Laplacian and show its application in acoustic impedance inversion which is a routine procedure in seismic explorations. A neural network is used to obtain a first approximation of the underlying acoustic impedance and construct a graph Laplacian matrix from this approximation. Afterwards, we use a Tikhonov-like variational method to solve the impedance inversion problem where the regularizer is based on the constructed graph Laplacian. The obtained solution can be shown to be more accurate and stable with respect to noise than the initial guess obtained by the neural network. This process can be iterated several times, each time constructing a new graph Laplacian matrix from the most recent reconstruction. The method converges after only a few iterations returning a much more accurate reconstruction.

We demonstrate the potential of our method on two different datasets and under various levels of noise. We use two different neural networks that have been introduced in previous works. The experiments show that our approach improves the reconstruction quality in the presence of noise.

Keywords: Impedance inversion; graph Laplacian; iterated Tikhonov; Deep Learning;

*School of Mathematics (Zhuhai), Sun Yat-sen University, Zhuhai, 518055, China.

Email: bianchid@mail.sysu.edu.cn

[†]Department of Mathematics, Harbin Institute of Technology, Harbin, 150001, China.

Email: f.bossmann@hit.edu.cn

Email: wenlong.wang@hit.edu.cn

1 Introduction

In seismic acoustic impedance inversion, subsurface properties and structures have to be reconstructed from seismic profiles. Mathematically, a subsurface impedance $\mathbf{x} \in X$ has to be recovered from a possibly noisy seismic profile $\mathbf{y} \in Y$, i.e., we have to find $\mathbf{x}$ such that

$$\mathbf{y}^\delta = F(\mathbf{x}) + \boldsymbol{\eta}, \quad \|\boldsymbol{\eta}\| \leq \delta \quad (1.1)$$

where $F : X \rightarrow Y$ is the forward operator. The inverse problem (1.1) is ill-posed in many cases which means that even a small amount of noise $\delta \ll 1$ can lead to extreme approximation errors. This ill-posedness originates from a non-linear operator F , heterogeneity in the profile $\mathbf{x}$, or an insufficient amount of data available.

A classical approach to overcome the ill-posedness is to regularize the inverse problem. Many of these methods fall under the broader category of Tikhonov-like variational methods which assume the form

$$\mathbf{x}_\alpha := \operatorname{argmin}_{\mathbf{x} \in X} \left\{ \mathcal{D}(F(\mathbf{x}); \mathbf{y}^\delta) + \alpha \mathcal{R}(\mathbf{x}) \right\}, \quad (1.2)$$

where $\mathcal{D}(\cdot; \cdot)$ is a pseudo-distance that quantifies the fidelity of the reconstruction, $\mathcal{R}(\cdot)$ serves as regularization term, and $\alpha > 0$ balances the trade-off between data fidelity and the regularization effect. A standard choice is

$$\mathcal{D}(F(\mathbf{x}); \mathbf{y}^\delta) = \|F(\mathbf{x}) - \mathbf{y}^\delta\|_2^2, \quad \mathcal{R}(\mathbf{x}) = \|\mathbf{x}\|_1.$$

As references see [17, 36].

The regularization operator $\mathcal{R}$ is of crucial importance. If we have access to a priori information on the ground-truth solution, we can tailor a specific $\mathcal{R}$ to incorporate such information and guide the overall regularization towards a narrower subset of approximate solutions which present the features we aim to recover.

Another typical example in image processing is given by $\mathcal{R}(\mathbf{x}) = \|L\mathbf{x}\|_q^q$, where L is a linear differential operator, such as the first or second derivative along each axis, and $\|\cdot\|_q$ is the Euclidean q -norm, $q \geq 1$, see [21, 23, 34]. The main reason is that 2D images are nearly piecewise-constant signals. Therefore, differential operators can help to preserve edges and identify regions with similar color intensity by minimizing the functional $\|L\mathbf{x}\|_q^q$.

Digital images, which consist of pixels arranged on a 2D grid, naturally possess a graph structure. For this reason, in recent years, graph-based differential operators Δ have been introduced to replace standard Euclidean differential operators L in various image processing tasks such as denoising [19], image deblurring [1, 8, 10, 14, 42], Computed Tomography (CT) [31], and other applications [5, 18, 35]. Graph operators showed a general good performance, due to the fact that they can more effectively model the complex structures and textures present in images. Unlike traditional Euclidean differential operators that primarily consider the spatial proximity of pixels, graph-based operators can take into account the intensity similarity between pixels as well, allowing for a deeper information extrapolation from image data.

Besides these classical methods, machine learning and deep learning approaches have gained interest over recent years. They have successfully been applied in seismic interpolation [37], full waveform inversion [43], impedance inversion [16], and seismic interpretation [39]. See also [27,

40] for an overview. In many of these cases neural networks generate the best outcome by far. A great advantage of such techniques is that the non-linearity in (1.1) is directly learned from the data. This can also be used to construct a forward operator without requiring a detailed and complicated physical model [3]. However, training a neural network requires a suitable amount and quality of data. A poorly trained network can result in strong artifacts and errors in the reconstruction, especially when dealing with ill-posed inverse problems [4, 15]. This can happen whenever there is insufficient data, biased data, noise, inaccurate labeling, or a shift in the dataset between training and deployment [22, 25, 41]. This makes data acquisition a complicated and expensive process.

To overcome said problem hybrid methods that combine neural networks with classical approaches have been studied lately [6, 29]. In [11, 12], a new ‘two-step’ method called **graphLa+Ψ** was proposed. The idea is to get a first approximate reconstruction of the ground-truth solution via a generic reconstructor operator $\Psi: Y \rightarrow X$, applied to the observed data $\mathbf{y}^\delta$, and then solve (1.2) replacing $\mathcal{R}(\mathbf{x})$ with $\|\Delta_\Psi \mathbf{x}\|_1$, where Δ_Ψ is a graph Laplacian built from $\Psi(\mathbf{y}^\delta)$. The first approximate reconstruction $\Psi(\mathbf{y}^\delta)$ helps identifying the pixel connections in the ground-truth signal by both their spatial proximity and intensity similarity. This approach greatly improves the quality of the approximate solution $\Psi(\mathbf{y}^\delta)$ for any given initial reconstructor Ψ .

In this work we apply the **graphLa+Ψ** method iteratively to the seismic impedance inversion equation (1.1), using as first reconstructor two Deep Neural Networks (DNNs) specifically trained on seismic impedance datasets [3, 30]. Both DNNs learn the non-linear seismic impedance forward operator (1.1) directly from the given data. As seismic impedance measurements only have slight non-linearities, we use a linear approximation in the second step to simplify (1.2). Our iterative method then reads,

$$\begin{cases} \mathbf{x}_n = \underset{\mathbf{x} \in X}{\operatorname{argmin}} \left\{ \frac{1}{2} \|K\mathbf{x} - \mathbf{y}^\delta\|_2^2 + \alpha_n \|\Delta_{\mathbf{x}_{n-1}} \mathbf{x}\|_1 \right\} & \text{for } n \geq 1, \\ \mathbf{x}_0^\delta := \Psi(\mathbf{y}^\delta), \end{cases}$$

where $\Psi: Y \rightarrow X$ is a trained DNN and K is the linear approximation of F , i.e., $K\mathbf{x} \approx F(\mathbf{x})$. The idea is that at each step n , having at our disposal a better approximation $\mathbf{x}_{n-1}$ of the ground-truth, we can update the weights of the graph Laplacian Δ taking into account this new information. The updated weights bring a sharper insight on the pixel connections in the ground-truth signal, driving then the overall reconstruction into a closer neighborhood of the ground truth $\mathbf{x}_{\text{gt}}$.

The choice of a DNN as the initial reconstructor Ψ is based on the numerical evidence presented in [11, 12] and the regularization properties of the **graphLa+Ψ** method. DNNs outperformed any other traditional regularization methods when used as the initial reconstructor Ψ . Additionally, despite the well-known instability of DNNs in direct methods, the **graphLa+Ψ** variational approach has been shown to be regularizing and stable, even when integrated with a DNN. Specifically, it overcomes the typical hallucinations issue of a DNN when processing input data $\mathbf{y}^\delta$ that falls outside of the DNN’s training domain.

The manuscript is organized as follows. In the Section 2 we give the basic notation and theory required for this work. We also introduce the standard **graphLa+Ψ** method and shortly summarize its theory. In Section 3 we present our iterated **graphLa+Ψ** algorithm, followed by a discussion on the used initial DNN in Section 4. Finally, in Section 5 numerical results for

different scenarios are shown. A discussion of the results and outlook on future problems can be found in the conclusion in Section 6.

2 The graph Laplacian and the graphLa+ Ψ method

We first introduce some preliminaries on graph theory and theoretical results about the graphLa+ Ψ regularization method.

2.1 Images and graphs

For a modern introduction to graph theory we invite the interested reader to consult [26]. For simplicity, in this work we consider as a graph any pair $G = (P, w)$ where P is a finite set, called *node set*, and $w: P \times P \rightarrow [0, \infty)$ is a symmetric function, called *edge-weight function*. The set of the edges of a graph is given by $E := \{(p, q) \in P \times P \mid w(p, q) \neq 0\}$. Two nodes p and q are connected if $(p, q) \in E$, and we write $p \sim q$. The intensity of the connection is given by $w(p, q)$.

For any function $\mathbf{x}: P \rightarrow \mathbb{R}$ we can compute the graph Laplacian $\Delta \mathbf{x}: P \rightarrow \mathbb{R}$, defined by the action

$$\Delta \mathbf{x}(p) := \sum_{q \sim p} w(p, q)(\mathbf{x}(p) - \mathbf{x}(q)). \quad (2.1)$$

2.1.1 Induced graph

Observe that to define a graph we only need a node set P and an edge-weight function w . Let us see now how to generate a graph from an image.

Since any image is made by the union of several pixels $p \in P$ disposed on a grid, it is then natural to identify the pixels as ordered pairs $p = (i_p, j_p)$ with $i_p = 1, \dots, n$ and $j_p = 1, \dots, m$, and where n and m indicates the total number of pixels along the horizontal and vertical axis, respectively. So, we can set the node set as

$$P = \{p \mid p = (i_p, j_p), i_p = 1, \dots, n, j_p = 1, \dots, m\}.$$

For the sake of simplicity, consider now a gray-scale image, which is given by the light intensities of its pixels. That is, a gray-scale image can be represented by a function

$$\mathbf{x}: P \rightarrow [0, 1],$$

where 0 means black and 1 means white. A very popular choice to make the connection w of two pixels depending on both their spatial proximity and light intensity is given by

$$w_{\mathbf{x}}(p, q) = \mathbb{1}_{(0, R]}(\text{dist}(p, q))g_{\mathbf{x}}(p, q), \quad (2.2)$$

where $\text{dist}(\cdot, \cdot)$ is a (pseudo) distance on P and

$$g_{\mathbf{x}}(p, q) := e^{-\frac{|\mathbf{x}(p) - \mathbf{x}(q)|^2}{\sigma}}. \quad (2.3)$$

The function $\mathbb{1}_{(0, R]}$ is the indicator function of the interval $(0, R]$ and $R > 0$ is a parameter of control which tells the maximum proximity distance allowed for two pixels to be neighbors. If

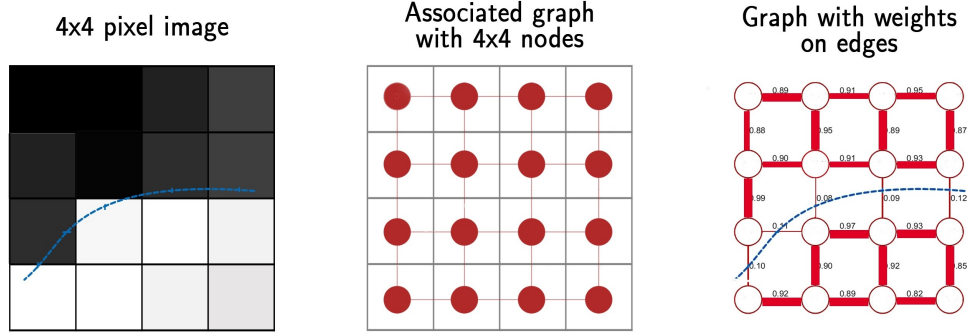


Figure 1: A simple example of constructing a graph from an image. Left: the original grayscale image composed of 4x4 pixels. Center: a node is associated with each pixel, and an initial set of edges is established by comparing the spatial distance between pixels. In this case, two pixels p, q are connected if $\text{dist}_1(p, q) = 1$. Right: following (2.2), the edges are then weighted based on the difference in intensities between two spatially connected pixels. The greater the intensity similarity between two spatially connected pixels, the higher the value of the edge weight connecting them. This is effectively represented by the thickness of the edges.

the distance between pixels p and q satisfies $0 < \text{dist}(p, q) \leq R$, then p and q are connected by an edge with magnitude $g_{\mathbf{x}}(p, q)$, i.e., $w_{\mathbf{x}}(p, q) = g_{\mathbf{x}}(p, q)$. The second control parameter, $\sigma > 0$, determines how sharply the edge weights vary with the difference in pixel light intensities. Common choices for the distance function $\text{dist}(\cdot, \cdot)$ are

$$\begin{aligned} \text{dist}(p, q) &= \text{dist}_1(p, q) := |i_p - i_q| + |j_p - j_q|, \\ \text{dist}(p, q) &= \text{dist}_{\infty}(p, q) := \max\{|i_p - i_q|; |j_p - j_q|\}. \end{aligned}$$

See Figure 1 for a simple example of building a graph from an image.

Eventually, by all the above considerations, given an image $\mathbf{x}: P \rightarrow [0, 1]$ we can then define its associated graph $G = (P, w_{\mathbf{x}})$, and consequently the graph Laplacian on G , $\Delta = \Delta_{\mathbf{x}}$, as per Equation (2.1). Let us observe that $\mathbf{z} \mapsto \|\Delta_{\mathbf{x}} \mathbf{z}\|_1$ is a pseudo-metric. In some sense, it is telling us how close we are to $\mathbf{x}$. See Figure 2, where we plot $|\Delta_{\mathbf{x}} \mathbf{x}|$.

2.2 The standard graphLa+ Ψ method: convergence and stability results

For the reader convenience, we report here a convergent result about the **graphLa+ Ψ** method which makes it a regularizing method. It was proven in [12]. We will use the notation introduced in Section 2.1

By a reconstruction method Ψ , we more broadly mean a family of operators of the form $\{\Psi_{\Theta} : Y \rightarrow X \mid \Theta \in \mathbb{R}^k\}$, which we call reconstructors. The parameters Θ may depend on δ and $\mathbf{y}^{\delta}$.

For instance, any variational method as (1.2) qualifies as a reconstructor, with parameter $\Theta = \alpha \in (0, +\infty)$. In this case, some specific parameter choice rules $\alpha = \alpha(\delta, \mathbf{y}^{\delta})$ can make the pair (Ψ_{α}, α) a regularization method, as per [17, Definition 3.1].

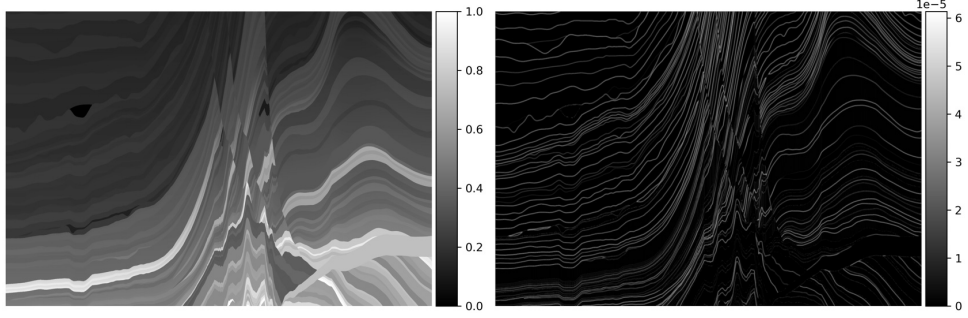


Figure 2: Left: The original image $\mathbf{x}$ in grayscale. Right: $|\Delta_{\mathbf{x}}\mathbf{x}|$, where $\Delta_{\mathbf{x}}$ is computed using (2.2). As it can be seen, $|\Delta_{\mathbf{x}}\mathbf{x}| \approx \mathbf{0}$.

Another example of interest is when Ψ_{Θ} is a trained DNN. Here, Θ denotes the parameters of the DNN, as discussed in Section 4. Even if DNNs can be theoretically trained to take into account different kind of noise levels δ and noise distributions, in practice this is impossible due the paramount computational time it would require. Therefore, DNNs are typically trained on just a few fixed level of noise intensities or noise distributions. So, in this case, Θ does not depend on δ nor on $\mathbf{y}^{\delta}$, that is, $\Theta(\delta, \mathbf{y}^{\delta}) \equiv \hat{\Theta}$ for a fixed $\hat{\Theta}$.

Chosen then a reconstructor Ψ_{Θ} and defining $\Psi_{\Theta}^{\delta} := \Psi_{\Theta}(\mathbf{y}^{\delta})$, the approximated solution $\mathbf{x}_{\alpha}$ given by the **graphLa+Ψ** method reads

$$\mathbf{x}_{\alpha} \in \operatorname{argmin}_{\mathbf{x} \in X} \left\{ \frac{1}{2} \|K\mathbf{x} - \mathbf{y}^{\delta}\|_2^2 + \alpha \|\Delta_{\Psi_{\Theta}^{\delta}} \mathbf{x}\|_1 \right\}. \quad (2.4)$$

Let $\mathbf{x}_0 \in X$ and $\Theta = \Theta(\delta, \mathbf{y}^{\delta})$ be such that

$$\mathbf{x}_0 = \lim_{\delta \rightarrow 0} \Psi_{\Theta(\delta, \mathbf{y}^{\delta})}(\mathbf{y}^{\delta}).$$

Remark 2.1. In the case of a DNN, since it is a (Lipschitz) continuous map for any fixed Θ , then $\mathbf{x}_0 = \Psi_{\Theta}(\mathbf{y})$. That is, $\mathbf{x}_0$ is the corresponding output of the DNN with input the noiseless data $\mathbf{y}$.

We have the following definition of solution.

Definition 2.2. We call $\mathbf{x}$ a graph-minimizing solution with respect to $\mathbf{x}_0$, if

$$\begin{aligned} & \text{(i) } K\mathbf{x} = \mathbf{y}, \\ & \text{(ii) } \|\Delta_{\mathbf{x}_0} \mathbf{x}\|_1 = \min\{\|\Delta_{\mathbf{x}_0} \mathbf{x}\|_1 \mid \mathbf{x} \in X, K\mathbf{x} = \mathbf{y}\}. \end{aligned} \quad (2.5)$$

The next theorem provide a convergence result. For more details, see [12, Section 3].

Theorem 2.3. Fix a sequence $\{\delta_k\}$ and $\alpha: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be such that

$$\|\mathbf{y}^{\delta_k} - \mathbf{y}\|_2 \leq \delta_k, \quad \lim_{k \rightarrow \infty} \delta_k = 0, \quad \lim_{\delta_k \rightarrow 0} \alpha(\delta_k) = 0, \quad \lim_{\delta_k \rightarrow 0} \frac{\delta_k^2}{\alpha(\delta_k)} = 0.$$

Then every sequence $\{\mathbf{x}_k\}$ of elements that minimize the functional (2.4), with δ_k and $\Theta(\delta_k, \mathbf{y}^{\delta_k})$, has a convergent subsequence. The limit $\mathbf{x}$ of the convergent subsequence $\{\mathbf{x}_{k'}\}$ is a graph-minimizing solution with respect to $\mathbf{x}_0$. If $\mathbf{x}$ is unique, then $\mathbf{x}_k \rightarrow \mathbf{x}$.

3 The iterated graphLa+ Ψ method

Iterative variants of (1.2) are very popular and have a long story. Iterated Tikhonov methods can often converge to approximated solutions of higher quality in many applications. For example, we refer the readers to [7, 9, 13, 20, 24].

In this work we propose an iterative version of (2.4) which reads

$$\begin{cases} \mathbf{x}_n = \underset{\mathbf{x} \in X}{\operatorname{argmin}} \left\{ \frac{1}{2} \|K\mathbf{x} - \mathbf{y}^\delta\|_2^2 + \alpha_n \|\Delta_{\mathbf{x}_{n-1}}\mathbf{x}\|_1 \right\} & \text{for } n \geq 1, \\ \mathbf{x}_0^\delta := \Psi(\mathbf{y}^\delta), \end{cases} \quad (3.1)$$

where $\Psi \equiv \Psi_\Theta$ is a fixed trained DNN. That is, at each step n we build a new graph Laplacian $\Delta_{\mathbf{x}_{n-1}}$ on the previous approximated solution $\mathbf{x}_{n-1}$. For $n = 1$ we recover exactly (2.4), with $\mathbf{x}_0^\delta = \Psi_\Theta^\delta$.

The fundamental intuition regarding this choice is that if $\mathbf{x}_{n-1}$ is a better approximation of $\mathbf{x}_{\text{gt}}$ than $\mathbf{x}_{n-2}$, then $\|\Delta_{\mathbf{x}_{n-1}}\mathbf{x}_{\text{gt}}\|_1 \leq \|\Delta_{\mathbf{x}_{n-2}}\mathbf{x}_{\text{gt}}\|_1$, making $\mathbf{x} \mapsto \|\Delta_{\mathbf{x}_{n-1}}\mathbf{x}\|_1$ a better regularization term.

We employ a Majorization–Minimization technique along with a Generalized Krylov Subspace approach to address the optimization problem of (3.1) at each step, reducing its complexity and computational cost. For an in-depth explanation of the algorithm, we refer to [28].

The convergence analysis of (3.1) necessitates a comprehensive investigation which is beyond the scope of the present manuscript. It will be addressed in a future work.

4 Initialization: Neural Network architecture, setup and training

For the seismic impedance experiments conducted here, we use two different neural networks to obtain an initial approximate solution $\mathbf{x}_0^\delta := \Psi(\mathbf{y}^\delta)$.

The first network was introduced by Alfarraj and AlRegib [3]. They use a forward and an inverse model that are trained simultaneously using the loss function

$$L(\Theta_1, \Theta_2) = \frac{0.2}{N_p} \left\| \mathbf{x}_{\text{gt}} - \Psi_{\Theta_1}(\mathbf{y}^\delta) \right\|_{2,\Omega}^2 + \frac{1}{N_s} \left\| \mathbf{y}^\delta - \Psi_{\Theta_2}(\Psi_{\Theta_1}(\mathbf{y}^\delta)) \right\|_2^2. \quad (4.1)$$

Here Ψ_{Θ_1} is the inverse model and Ψ_{Θ_2} the forward model. N_s is the number of seismic traces in the given data. Furthermore, we have given a ground truth $\mathbf{x}_{\text{gt}}$ on a subset Ω of the given traces which is usually obtained from a certain number N_p of well log samples. The loss function (2.2) calculates the weighted sum of the Mean Squared Errors on the impedance data and seismic traces. As forward model Ψ_{Θ_2} a simple 4-layer CNN is used. For the inverse model a more complicated structure consisting of four submodules is used: A sequence modeling build from GRU layers, a local pattern analysis based on convolutional layers, an upsampling step using deconvolution layers, and a regression module using a GRU and a linear layer. A schematic of the network is shown in Figure 3. For a more detailed description we refer to the original work [3] as well as to the Python code published by the authors [2].

As we will see in the numerical experiments, the above network can be prone to noise on the input data. Hence, some of the authors proposed a new approach where the the inverse model is replaced by a more stable design while the forward model and loss function are kept the

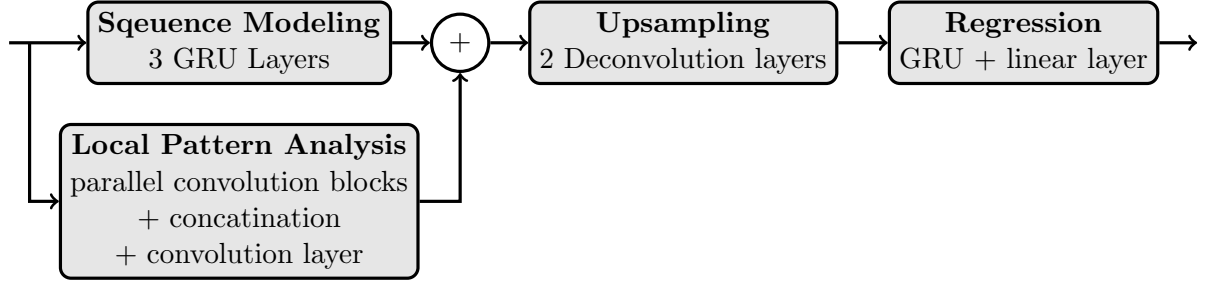


Figure 3: Schematic of the network architecture proposed by Alfarraj and AlRegib [3].

same. The new inverse model uses four submodules that are applied in serial: super resolution build from convolution layers and a pixel shuffle, time domain graph convolutional attention (TDGC-Attention), deconvolution, and a fully connected layers module. The schematics of the network architecture are shown in Figure 4, and more details can be found in [30].

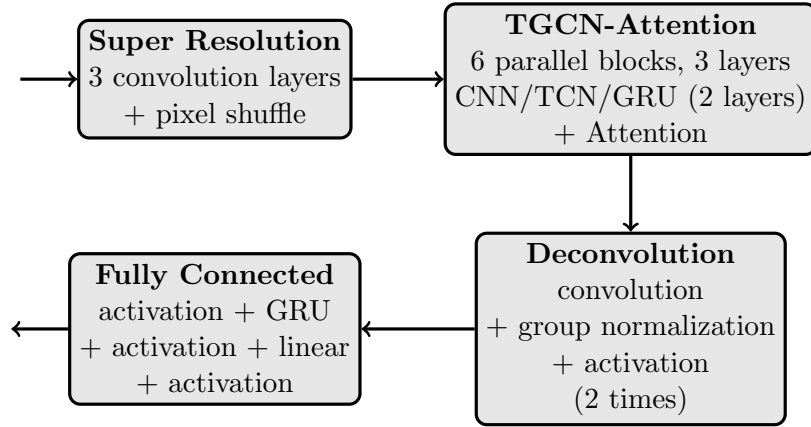


Figure 4: Network architecture for a more stable inverse model under noise proposed in [30].

Both networks have been trained using $N_p = 20$ evenly spaced traces for which we assume the acoustic impedance ground truth is known from well logs. The networks are initialized with random parameters and updated using gradient descent on the loss function (4.1). The networks are designed to be trained on-the-spot, i.e., the training data is the data used in the according experiments in the next section.

Furthermore, the linear operator K required for the iterated **graphLa**+ Ψ algorithm is learned as a one-layer network from the forward model Ψ_{Θ_2} . The non-linear effects in acoustic impedance are minor and thus the non-linear forward model can be approximated by a linear operator quite accurately.

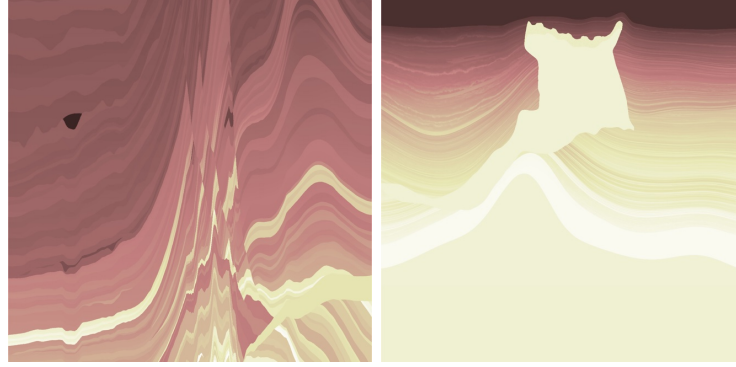


Figure 5: acoustic impedance data used in experiments: Marmousi2 and SEAM model.

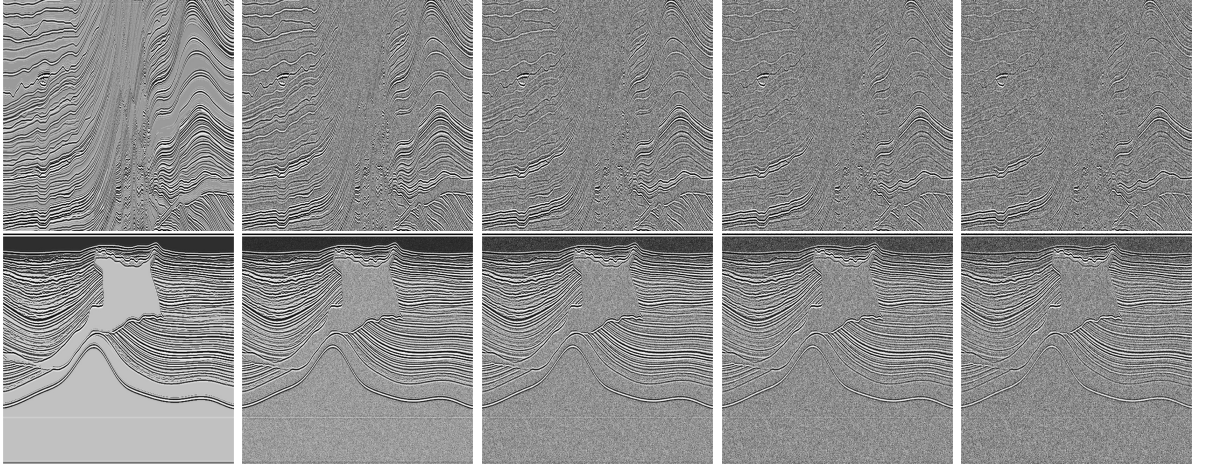


Figure 6: Simulated seismic profile with different noise levels. (A non-linear grayscale colormap was used to achieve a better contrast.)

5 Numerical Experiments

5.1 Initial step

In the following experiments we use the two different acoustic impedance data shown in Figure 5. The first dataset is taken from the Marmousi2 model [32], the second one is the SEAM model [33]. The corresponding seismic profile is shown in Figure 6 where an increasing level of noise is added up to a PSNR value of 30 (right column). Note that a non-linear colormap was chosen to improve the contrast of the image.

To obtain a first approximate reconstruction we then use the Alfarraj and AlRegib network (AA-Net) [3] and the Liu network (L-Net) [30] to reconstruct the impedance profile from the data shown in Figure 6. The results are shown in Figure 7 (AA-Net) and Figure 8 (L-Net). As can be seen, AA-Net is more prone to noise on the data. In both cases the reconstruction error increases with the noise level. In a next step we will use iterated $\text{graphLa}+\Psi$ to remove those artifacts and increase the overall reconstruction quality.

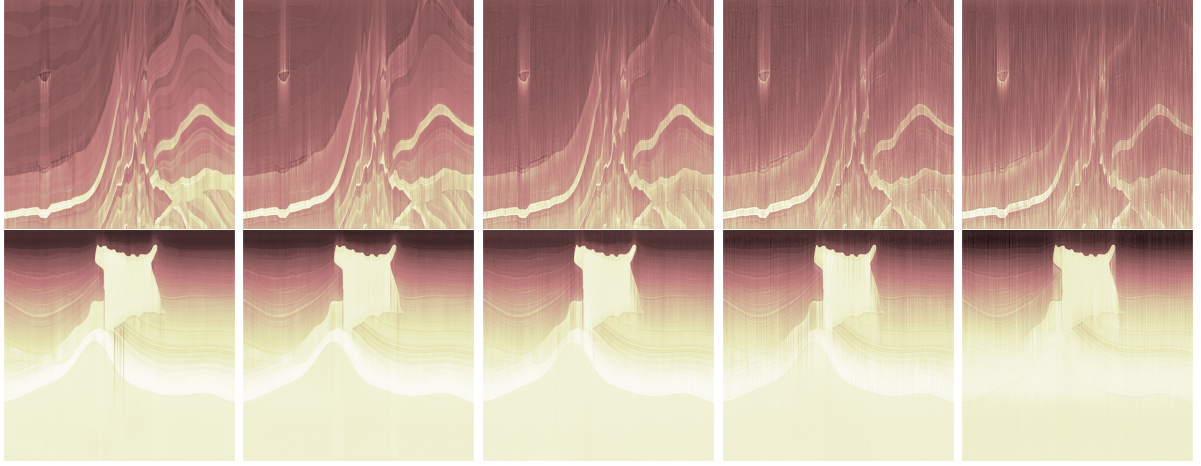


Figure 7: Reconstructed impedance profile using AA-Net.

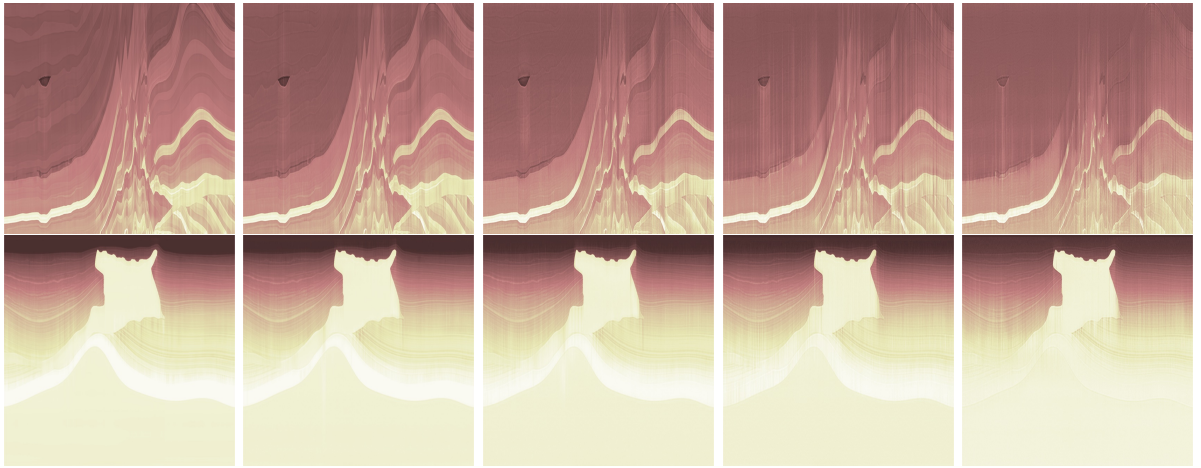


Figure 8: Reconstructed impedance profile using L-Net.

5.2 Parameter setting for iterated $\text{graphLa}+\Psi$

In our experiments we achieved good reconstruction results for a maximum proximity distance of $R \geq 5$ in Equation (2.2). The parameter should be chosen larger the more noise is present in the data. However, with increasing R the graph Laplacian matrix gets more and more dense which leads to an increased time and space complexity. The experiments shown in the following are all performed using $R = 7$.

We set $\sigma = 0.001$ in all presented experiments as this value always resulted in close to optimal results compared to other choices. Iterated $\text{graphLa}+\Psi$ can be slightly optimized further by fine tuning the weight σ . We recommend a value $\sigma \in [0.1, 0.001]$ where the parameter should be smaller for higher noise levels.

Last the balancing parameter α in Equation (1.2) was simply set to $\alpha = 1$ as it does not influence the reconstruction as much as the previously discussed parameters and thus was not worth optimizing.

5.3 Iterated $\text{graphLa}+\Psi$

To measure the reconstruction quality of the proposed method we use the structural similarity index (SSIM index) [38]. This index compares the similarity of two images by calculating

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (5.1)$$

where μ_x , μ_y is the mean value, σ_x , σ_y is the variance, and σ_{xy} is the covariance. Small constants c_1 and c_2 are used for stabilization. Equation (5.1) is calculated over all 10×10 pixel sized windows and the mean value is returned. The SSIM index returns a value in $[-1, 1]$ where 1 is a perfect match between both images.

In a first experiment we investigate on the number of required iterations. Therefore iterated $\text{graphLa}+\Psi$ was applied to the DNN reconstructions shown in Figures 7 and 8. We performed a total of 10 iterations for each case. The resulting SSIM value compared to the original ground truth for each iteration is shown in Figure 9. Note that the first SSIM index used for each plot is the starting SSIM achieved by the DNN initialization.

We observe that the SSIM index can drastically increase using our new method. However, the results depend on the noise level. For low noise levels the SSIM index reaches its optimum after only a few iterations and afterwards decreases slightly before converging. For noiseless data the results are even slightly worse than the original DNN reconstruction. We assume this is due to the linear approximation K of the forward operator used in (3.1). While using a linear operator simplifies the algorithm, it is a slight oversimplification compared to the used DNN forward operator Ψ_{Θ_2} .

Another interesting observation is that the choice of the initial DNN has a much higher impact for the Marmousi model than for the SEAM model.

In Figure 10 and 11 the reconstructions of iterated $\text{graphLa}+\Psi$ are shown for an initial guess using AA-Net and L-Net respectively. Based on the previous experiment, we use a different number of iterations $n = 1, 3, 5, 10, 10$ depending on the noise level. It can be seen that the amount of noise is drastically reduced and we are able to reconstruct more details.

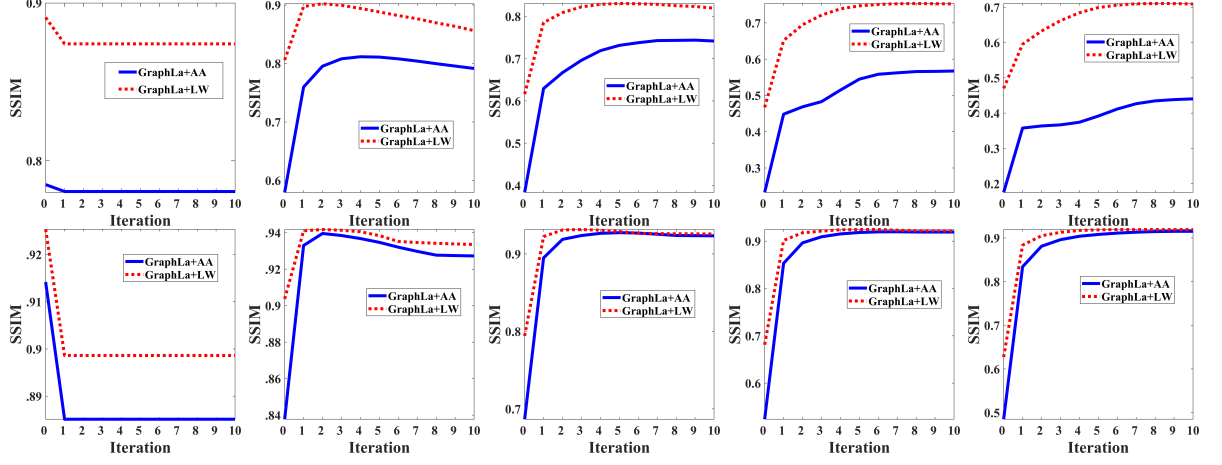


Figure 9: SSIM value of reconstruction using iterated $\text{graphLa}+\Psi$ on Marmousi (top) and SEAM (bottom) model for increasing levels of noise (left to right). As initialization AA-Net (blue) and L-Net (red,dotted) were used.

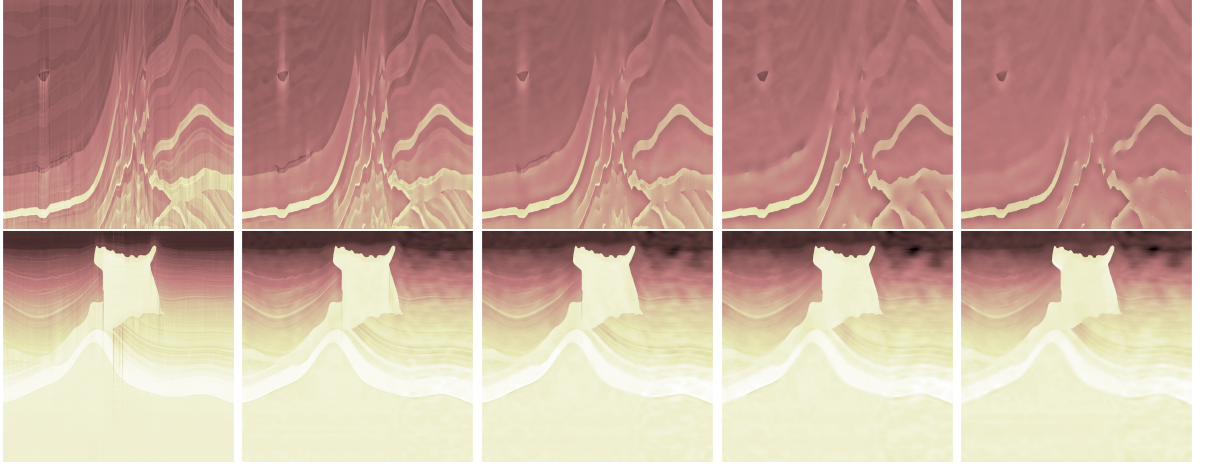


Figure 10: Reconstructed impedance profile using AA-Net initialization and $\text{graphLa}+\Psi$ for increasing noise levels and $n = 1, 3, 5, 10, 10$ iterations (left to right).

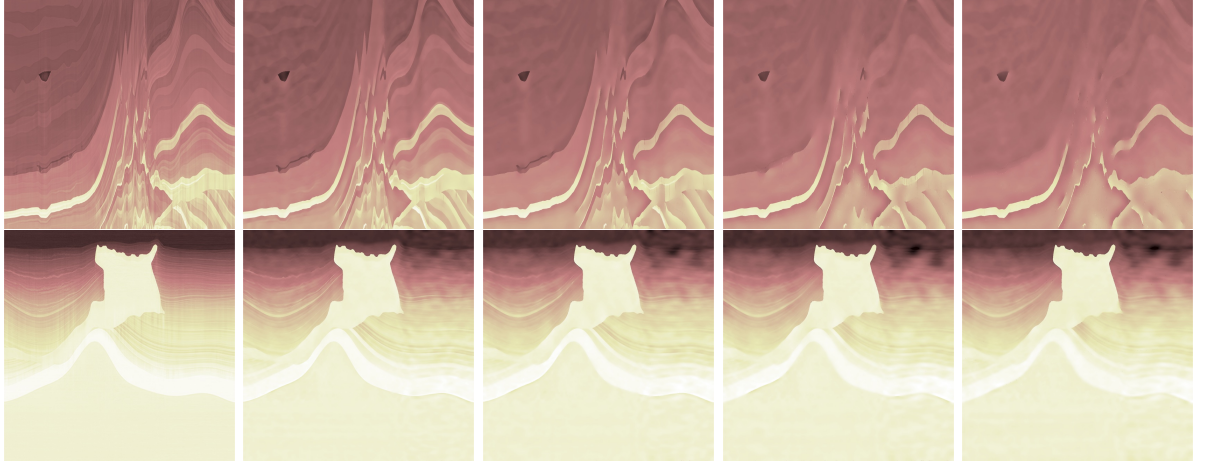


Figure 11: Reconstructed impedance profile using L-Net initialization and **graphLa+ Ψ** for increasing noise levels and $n = 1, 3, 5, 10, 10$ iterations (left to right).

6 Conclusion

In this work we introduced an iterative version of a Tikhonov-like regularization method which combines both deep learning and graph operators, for solving an ill-posed impedance inversion problem. We showcased different numerical experiments with different level of noise and state-of-the-art DNN architectures designed for the acoustic impedance inversion problem.

Despite the instabilities of the DNNs in presence of noise, the iterated graph Laplacian is able to provide stable final reconstructions of much higher quality after a few iterations. The influence of noise is greatly reduced and more details of the impedance profile are recovered. The method maintains stability under many scenarios without the need for significant parameter tuning.

Future directions: To improve results even further, the linear forward operator can be replaced with the more accurate non-linear DNN. However, this requires a more extensive theory and a sophisticated numerical algorithm. Another interesting question is if the proposed method can be used to reduce the number of required ground truth traces which are hard to obtain in practise.

References

- [1] S. Aleotti, A. Buccini, and M. Donatelli. “Fractional graph Laplacian for image reconstruction”. In: *Applied Numerical Mathematics* (2023).
- [2] M. Alfarraj and G. AlRegib. *Python Code: Semi-supervised learning for acoustic impedance inversion*. <https://github.com/olivesgatech/Semi-supervised-Learning-for-Acoustic-Impedance-Inversion>. 2019.
- [3] M. Alfarraj and G. AlRegib. “Semi-supervised learning for acoustic impedance inversion”. In: *SEG Technical Program Expanded Abstracts 2019*. Society of Exploration Geophysicists, 2019, pp. 2298–2302.

- [4] V. Antun, F. Renna, C. Poon, B. Adcock, and A. C. Hansen. “On instabilities of deep learning in image reconstruction and the potential costs of AI”. In: *Proceedings of the National Academy of Sciences*. Vol. 117. 48. 2020, pp. 30088–30095.
- [5] P. Arias, V. Caselles, and G. Sapiro. “A variational framework for non-local image inpainting”. In: *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*. Springer. 2009, pp. 345–358.
- [6] S. Arridge, P. Maass, O. Öktem, and C.-B. Schönlieb. “Solving inverse problems using data-driven models”. In: *Acta Numerica* 28 (2019), pp. 1–174.
- [7] M. Bachmayr and M. Burger. “Iterative total variation schemes for nonlinear inverse problems”. In: *Inverse Problems* 25.10 (2009), p. 105004.
- [8] D. Bianchi, A. Buccini, M. Donatelli, and E. Randazzo. “Graph Laplacian for image deblurring”. In: *ETNA* 55 (2022), pp. 169–186.
- [9] D. Bianchi, A. Buccini, M. Donatelli, and S. Serra-Capizzano. “Iterated fractional Tikhonov regularization”. In: *Inverse Problems* 31.5 (2015), p. 055005.
- [10] D. Bianchi and M. Donatelli. “Graph approximation and generalized Tikhonov regularization for signal deblurring”. In: *2021 21st International Conference on Computational Science and Its Applications (ICCSA)*. IEEE. 2021, pp. 93–100.
- [11] D. Bianchi, M. Donatelli, D. Evangelista, W. Li, and E. L. Piccolomini. “Graph Laplacian and Neural Networks for Inverse Problems in Imaging: GraphLaNet”. In: *International Conference on Scale Space and Variational Methods in Computer Vision*. 2023, pp. 175–186.
- [12] D. Bianchi, D. Evangelista, S. Aleotti, M. Donatelli, E. L. Piccolomini, and W. Li. *A data-dependent regularization method based on the graph Laplacian*. Preprint arXiv:2312.16936. 2023.
- [13] D. Bianchi, G. Lai, and W. Li. “Uniformly convex neural networks and non-stationary iterated network Tikhonov (iNETT) method”. In: *Inverse Problems* 39.5 (2023), p. 055002.
- [14] A. Buccini and M. Donatelli. “Graph Laplacian in $\ell^2 - \ell^q$ regularization for image reconstruction”. In: *2021 21st International Conference on Computational Science and Its Applications (ICCSA)*. IEEE. 2021, pp. 29–38.
- [15] M. J. Colbrook, V. Antun, and A. C. Hansen. “The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale’s 18th problem”. In: *Proceedings of the National Academy of Sciences*. Vol. 119. 12. 2022, e2107151119.
- [16] V. Das, A. Pollack, U. Wollner, and T. Mukerji. “Convolutional neural network for seismic impedance inversion”. In: *Geophysics* 84.6 (2019), R869–R880.
- [17] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*. Springer Science & Business Media, 1996.
- [18] G. Gilboa and S. Osher. “Nonlocal linear image regularization and supervised segmentation”. In: *Multiscale Modeling & Simulation* 6.2 (2007), pp. 595–630.

- [19] G. Gilboa and S. Osher. “Nonlocal operators with applications to image processing”. In: *Multiscale Modeling & Simulation* 7.3 (2009), pp. 1005–1028.
- [20] M. Hanke and C. W. Groetsch. “Nonstationary iterated Tikhonov regularization”. In: *Journal of Optimization Theory and Applications* 98 (1998), pp. 37–53.
- [21] P. C. Hansen, J. G. Nagy, and D. P. O’leary. *Deblurring images: matrices, spectra, and filtering*. SIAM, 2006.
- [22] F. Huot, B. Biondi, and G. Beroza. “Jump-starting neural network training for seismic problems”. In: *SEG International Exposition and Annual Meeting*. SEG. 2018, SEG–2018.
- [23] A. K. Jain. *Fundamentals of digital image processing*. Prentice-Hall, Inc., 1989.
- [24] Q. Jin and M. Zhong. “Nonstationary iterated Tikhonov regularization in Banach spaces with uniformly convex penalty terms”. In: *Numerische Mathematik* 127 (2014), pp. 485–513.
- [25] S. Karimpouli and P. Tahmasebi. “Physics informed machine learning: Seismic wave equation”. In: *Geoscience Frontiers* 11.6 (2020), pp. 1993–2001.
- [26] M. Keller, D. Lenz, and R. K. Wojciechowski. *Graphs and Discrete Dirichlet Spaces*. Grundlehren der mathematischen Wissenschaften. Springer, Cham, 2021.
- [27] F. Khosro Anjom, F. Vaccarino, and L. V. Socco. “Machine learning for seismic exploration: Where are we and how far are we from the holy grail?” In: *Geophysics* 89.1 (2024), WA157–WA178.
- [28] A. Lanza, S. Morigi, L. Reichel, and F. Sgallari. “A generalized Krylov subspace method for ℓ_p - ℓ_q minimization”. In: *SIAM Journal on Scientific Computing* 37.5 (2015), S30–S50.
- [29] H. Li, J. Schwab, S. Antholzer, and M. Haltmeier. “NETT: Solving inverse problems with deep neural networks”. In: *Inverse Problems* 36.6 (2020), p. 065005.
- [30] M. Liu, J. Ma, W. Wang, and F. Boßmann. *Semi-supervised Impedance Inversion Based on Super Resolution and Attention*. Personal draft, preprint available on arXiv soon. 2024.
- [31] Y. Lou, X. Zhang, S. Osher, and A. Bertozzi. “Image recovery via nonlocal operators”. In: *Journal of Scientific Computing* 42.2 (2010), pp. 185–197.
- [32] G. Martin, K. Marfurt, and S. Larsen. “An updated model for the investigation of avo in structurally complex area”. In: *Society of Exploration Geophysics*. doi 10.1.1817083 (2002).
- [33] A. Mustafa, M. Alfarraj, and G. AlRegib. “Spatiotemporal modeling of seismic images for acoustic impedance estimation”. In: *SEG International Exposition and Annual Meeting*. SEG. 2020, D041S101R005.
- [34] M. K. Ng, R. H. Chan, and W.-C. Tang. “A fast algorithm for deblurring models with Neumann boundary conditions”. In: *SIAM Journal on Scientific Computing* 21.3 (1999), pp. 851–866.
- [35] G. Peyré, S. Bougleux, and L. Cohen. “Non-local regularization of inverse problems”. In: *Computer Vision – ECCV 2008*. Ed. by D. Forsyth, P. Torr, and A. Zisserman. Springer Berlin Heidelberg, 2008, pp. 57–68.

- [36] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational methods in imaging*. Springer, 2009.
- [37] B. Wang, N. Zhang, W. Lu, and J. Wang. “Deep-learning-based seismic data interpolation: A preliminary result”. In: *Geophysics* 84.1 (2019), pp. V11–V20.
- [38] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. “Image quality assessment: from error visibility to structural similarity”. In: *IEEE transactions on image processing* 13.4 (2004), pp. 600–612.
- [39] T. Wrona, I. Pan, R. E. Bell, R. L. Gawthorpe, H. Fossen, and S. Brune. “3D seismic interpretation with deep learning: A brief introduction”. In: *The Leading Edge* 40.7 (2021), pp. 524–532.
- [40] S. Yu and J. Ma. “Deep learning for geophysics: Current and future trends”. In: *Reviews of Geophysics* 59.3 (2021), e2021RG000742.
- [41] J. Zhang, J. Li, X. Chen, Y. Li, G. Huang, and Y. Chen. “Robust deep learning seismic inversion with a priori initial model constraint”. In: *Geophysical Journal International* 225.3 (2021), pp. 2001–2019.
- [42] X. Zhang, M. Burger, X. Bresson, and S. Osher. “Bregmanized nonlocal regularization for deconvolution and sparse reconstruction”. In: *SIAM Journal on Imaging Sciences* 3.3 (2010), pp. 253–276.
- [43] Z. Zhang and T. Alkhalifah. “Regularized elastic full-waveform inversion using deep learning”. In: *Advances in subsurface data analytics*. Elsevier, 2022, pp. 219–250.