CSIAM Trans. Appl. Math.
doi: 10.4208/csiam-am.SO-2024-0027

Vol. 7, No. 1, pp. 29-61
March 2026

A Deep Learning Approach for Solving the Inverse Problem of the Wave Equation

Xiong-Bin Yan[1], Keke Wu[2], Zhi-Qin John Xu[3,4,5]
and Zheng Ma[3,4,5,6,*]

[1] *School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, China.*
[2] *School of Mathematical Sciences and Suzhou Institute for Advanced Research, University of Science and Technology of China, Jiangsu 215217, China.*
[3] *School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai 200240, China.*
[4] *Qing Yuan Research Institute, Shanghai Jiao Tong University, Shanghai 200240, China.*
[5] *Institute of Natural Sciences, MOE-LSC, Shanghai Jiao Tong University, Shanghai 200240, China.*
[6] *CMA-Shanghai, Shanghai Jiao Tong University, Shanghai 200240, China.*

Received 24 June 2024; Accepted 13 April 2025

Abstract. Full-waveform inversion is a powerful geophysical imaging technique that infers high-resolution subsurface physical parameters by solving a non-convex optimization problem. However, due to limitations in observation, e.g. limited shots or receivers, and random noise, conventional inversion methods are confronted with numerous challenges, such as the local-minimum problem. In recent years, a substantial body of work has demonstrated that the integration of deep neural networks and partial differential equations for solving full-waveform inversion problems has shown promising performance. In this work, drawing inspiration from the expressive capacity of neural networks, we provide a new deep learning approach aimed at accurately reconstructing subsurface physical velocity parameters. This method is founded on a re-parametrization technique for Bayesian inference, achieved through a deep neural network with random weights. Notably, our proposed approach does not hinge upon the requirement of the labeled training dataset, rendering it exceedingly versatile and adaptable to diverse subsurface models. Furthermore, uncertainty analysis is effectively addressed through approximate Bayesian inference. Extensive experiments show that the proposed approach performs noticeably better than existing conventional inversion methods.

AMS subject classifications: 34K28, 35R30, 62M45

Key words: Full-waveform inversion, Bayesian inference, re-parametrization technique, deep neural network.

*Corresponding author. *Email addresses:* yanxb@lzu.edu.cn (X. Yan), wukekever@ustc.edu.cn (K. Wu), xuzhiqin@sjtu.edu.cn (Z. Xu), zhengma@sjtu.edu.cn (Z. Ma)

1 Introduction

In geophysics, seismic waveform inversion is frequently employed to obtain quantitative estimates of subsurface properties that can accurately predict the observed seismic data. The reconstruction of subsurface properties is a profoundly non-linear inverse problem. There are various techniques available to address the seismic inverse problem, such as velocity analysis of stacked seismic traces [2], migration-based travel-time approaches [6, 46], Born approximation [16, 26], and full-waveform inversion (FWI) [18, 33, 38]. Distinguished from other approaches, full-waveform inversion excels in the ability to deduce high-resolution subsurface structures. This is achieved through an iterative process that involves aligning observed and simulated seismograms while harnessing comprehensive wavefield information. This superiority stems from its capacity to fully exploit the informational richness of recorded seismic data, encompassing both amplitude and travel-time components.

In the context of full-waveform inversion, the typical approach for minimizing the objective function involves the use of gradient descent methods. These methods require the explicit computation of gradients pertaining to the cost function with respect to the velocity model. Nevertheless, the application of gradient descent methods to the full-waveform inversion objective function poses considerable challenges. This is primarily due to the presence of numerous local minima, a consequence of the inherent high nonlinearity and ill-posed nature of the problem [36]. Over the past few decades, researchers have developed numerous approaches aimed at enhancing the effectiveness of FWI. These approaches encompass the integration of multiple data components [4], the introduction of regularization terms [1, 44], the regularization by discretization [11, 14], and the adoption of novel objective functions [25, 28, 45].

Deep learning, primarily in the form of deep neural networks (DNNs), has gained significant attention in the fields of science and engineering. Its versatile applications span a wide range, including image classification/recognition [49], shape representation [42], and natural language processing [7]. Similarly, this method has also found extensive application in the solution of inverse problems and has demonstrated performance surpassing that of traditional methods. Examples include CT reconstruction [8, 37], image processing [5, 34], MRI reconstruction [17], and more.

A strong effort has been made in recent years to facilitate deep learning application in geophysics, including studies focused on fault detection [39], random noise attenuation [31], and so on. Some researchers have also attempted to design a DNN architecture that directly maps seismic data to subsurface models in a fully data-driven manner. Numerical experiments [9,19,40,43,48] have demonstrated that the inverse operator for full-waveform inversion can be acquired by training convolutional neural networks (CNNs) using datasets composed of wavefield and velocity model pairs. Nonetheless, it is important to note that the ill-posed and complex nature of seismic inversion poses challenges for training CNNs that are both sufficiently generalized and robust for inversion tasks.