A novel Fourier feature physics-informed neural networks based on the boundary element method for solving scattering of SH wave induced by complex topography

Yongxin Wu[1], Haochen Yuan[1], Yufeng Gao[1], Shangchuan Yang[2], Yue Hou[3]

[1] Key Laboratory of Ministry of Education for Geomechanics and Embankment Engineering, College of Civil and Transportation Engineering,. Hohai University, Nanjing 210098, China

[2] Key Laboratory of High-speed Railway Engineering of Ministry of Education, Southwest Jiaotong University, Chengdu, 610031, China.

[3] Department of Civil Engineering, Faculty of Science and Engineering, Swansea University, Swansea, UK.

Correspondence:

Yufeng Gao, Key Laboratory of Ministry of Education for Geomechanics and Embankment Engineering, College of Civil and Transportation Engineering, Hohai University, Nanjing 210098, China. E-mail: yfgao66@163.com

E-mail addresses:

yongxin@hhu.edu.cn (Yongxin Wu); haochen_yuan@hhu.edu.cn (Haochen Yuan); shangchuan.yang@swjtu.edu.cn (Shangchuan Yang); yue.hou@swansea.ac.uk (Yue Hou)

A novel Fourier feature physics-informed neural networks based on the boundary element method for solving scattering of SH wave induced by complex topography

ABSTRACT

The topographic effects of seismic waves often amplify seismic hazards. Thus, solving the scattering phenomena of seismic waves due to complex topography is a significant challenge in seismology and earthquake engineering. To address this problem efficiently and accurately, we propose a novel method, termed FF PINNs-BEM, which integrates Fourier feature physics-informed neural networks (FF PINNs) with the boundary element method (BEM) to solve parameterized frequency-domain wave equations. The accuracy of the proposed method is verified by three representative topographies, i.e., a semi-cylindrical canyon, a semi-cylindrical mountain and a V-shaped canyon, for which analytical solutions are available. Furthermore, we simulate scattered wavefields parameterized by the incidence angle and dimensionless frequency. Transfer learning, leveraging a pre-trained model, is employed to accelerate the neural network convergence. Experimental results demonstrate that the proposed method achieves high accuracy and efficiency while exhibiting robust convergence and generalization capabilities.

KEYWORDS: PINNs, BEM, ground motion, SH waves, topographic effect

1 INTRODUCTION

Earthquakes cause significant economic losses and structural damage. Empirical studies of earthquake damage show that topographic effects severly influence damage severity.[1-3] Thus, the amplification and attenuation of seismic waves caused by local irregular topography have been extensively studied in earthquake engineering.

To reveal the topographic effects on the scattering of seismic waves, the analytical solution is primarily based on the wave function expansion method (WFEM). Early studies introduced wave function series solutions for classical semi-cylindrical and semi-elliptical canyons, serving mainly as benchmarks for validating other numerical methods.[4,5] Subsequent research developed wave function series solutions for SH-wave scattering in various concave terrains in the half-plane. However, these studies focus exclusively on terrains characterized by circular arcs and elliptical shapes.[6-8] A significant advancement was achieved when Tsaur et al. [9,10] introduced an analytical solution for symmetric V-shaped canyons using the region-matching technique, which better simulated real-world scenarios. Based on this technique, analytical solutions of non-symmetrical V-shaped and U-shaped canyons were later proposed.[11,12] By directly solving the underlying mathematical model, the WFEM provides insights into the physical nature of the wave scattering process. However, these analytical methods often require simplifying irregular topography, limiting their application to realistic scenarios.

In recent decades, numerical simulation methods have facilitated the study of more complex topographies. Commonly used approaches include the finite difference method (FDM),[13,14] finite element method (FEM),[15,16] spectral element method (SEM)[17-19], and boundary element method (BEM).[20-23] However, each method exhibits unique advantages and limitations stemming from the different ways they handle the wave equation and its boundary conditions. For instance, FDM, FEM, and SEM can accurately simulate complicated geological conditions but require artificial absorbing boundaries to constrain the problem within a finite computational domain.[24] These absorbing boundaries, however, do not strictly satisfy Sommerfeld's radiation condition, which may compromise accuracy and stability. In contrast, the BEM inherently satisfies the required radiation conditions, yet its dependence on the calculation of Green’s functions renders it computationally expensive. Consequently, it is of practical importance to develop novel approaches that address these limitations of the aforementioned methods.

More recently, data-driven machine learning techniques have been widely used in earthquake engineering. [25-29] However, the lack of physical interpretability and the scarcity of accessible data limit its applicability. With advances in automatic differentiation techniques, the integration of physical constraints into deep learning has become an effective approach to enhance the generality and interpretability of machine learning models.[30,31] Notable approaches include the deep Galerkin method (DGM),[32] the deep Ritz method (DRM),[33] physics-informed neural networks (PINNs),[34] and the deep energy method (DEM).[35] Among these, PINNs impose the governing equations and their initial and boundary conditions as soft constraints by integrating them into the loss function. This conceptually simple and highly flexible framework has been successfully adopted for solving partial differential equations (PDEs) in diverse fields. In geophysics and seismology, PINNs have been employed to solve the Eikonal equations[36-38] and Maxwell equations[39]. Additionally, due to PINNs’ unified framework for addressing both forward and inverse PDE problems, they have been extensively used to solve the forward and inverse problems of wave equations in complex media, in both the time[40-43] and frequency domains.[44-46]

To simulate the frequency-domain scattered wavefields induced by complex media, Song et al.[44,45] proposed a PINN-based framework for solving single-frequency Helmholtz equations. In practice, using this PINNs to solve frequency-domain wave equations requires training a separate network for each frequency. To overcome this limitation, Alkhalifah et al.[47] proposed feeding frequency to the neural network, thereby enabling the generation of multi-frequency wavefields using a single model. Nevertheless, multi-frequency wavefields inherently exhibit multi-scale characteristics. Neural networks often struggle to capture these features effectively and tend to learn predominantly low-frequency components caused by their spectral bias.[48,49] To alleviate this limitation, various input encoding strategies have been proposed. For example, multi-resolution hash encoding has recently been introduced into PINNs to accelerate convergence by enriching the input representation with local and multi-scale information.[50] Although this approach primarily aims at improving training efficiency, its multi-resolution nature implicitly enhances the network’s capability to represent high-frequency components. To more explicitly enhance the representation of high-frequency wavefields, Tancik et al.[51] proposed the Fourier feature (FF) neural network, which improves the network’s ability to capture high-frequency details by applying Fourier mapping to low-dimensional coordinates. Song et al.[52] employed FF PINNs to generate multi-frequency scattered wavefields induced by complex velocity media for arbitrary source positions. Despite these advances, the potential of FF PINNs for simulating frequency-domain scattered wavefields induced by complex topographies, especially over multiple frequencies and varying incidence angles, remains underexplored.

Moreover, the direct application of PINNs to seismic wave scattering problems in semi-infinite domains remains challenging. In particular, traction-free boundary conditions associated with free surfaces cannot be enforced exactly within the vanilla PINN framework. To address this issue, several enhanced PINN formulations have been proposed. For example, physics-constrained neural networks (PCNN)[42] achieve a hard embedding of traction-free boundary conditions by images method; self-adaptive PINNs[43] improve the enforcement of boundary conditions by dynamically adjusting the weights of different loss terms during training. Despite these advances, all the aforementioned PINNs-based methods for solving wave equations rely on artificial absorbing boundaries to truncate wave propagation in semi-infinite domains. Common approaches include the absorbing boundary condition (ABC)[41] and the perfectly matched layer (PML)[44,45]. The ABC applies a soft constraint at the truncated boundary, which is incorporated into the boundary loss. The PML method reformulates the wave equations by introducing spatially dependent coefficients related to the source function into the differential terms. These coefficients are then incorporated into the equation loss to enforce the wave propagation truncation. However, these soft constraints increase the model complexity and can compromise convergence stability.

To address the above challenges, we propose a hybrid framework that integrates FF PINNs with BEM (FF PINN–BEM). By reformulating the learning task in terms of a virtual force defined on the physical boundary, the proposed method eliminates the need for artificial absorbing boundaries and enforces traction-free boundary conditions in a physically consistent manner. By leveraging Fourier features to mitigate the spectral bias of neural networks, the method in this study enhances the prediction of high-frequency wavefields.

The paper proceeds as follows. Section 2 provides an introduction to the BEM and the network architecture. We further discuss the spectral bias inherent in traditional neural networks and explain how Fourier feature mapping alleviates this limitation to improve the prediction of high-frequency functions. Section 3 employs benchmark analytical solutions to validate the computational precision of the novel methodology in modeling SH wave scattering across three topographic configurations: semi-cylindrical canyons, semi-cylindrical mountains, and V-shaped canyons. Section 4 evaluates and discusses performance characteristics of the presented approach in generating surface displacements of a V-shaped canyon for multiple frequencies and angles of incidence. Additionally, the method's applicability is demonstrated in a more complex coupled terrain. Finally, the conclusion section provides a summary of the work presented in this study, highlighting its key contributions.

2 METHODOLOGY

2.1 Problem setup

In this work, the two-dimensional SH wavefield in an elastic half-space with complex topography is modeled, as shown in Figure 1. The elastic half-space is assumed to be isotropic and homogeneous with its shear modulus μ. The medium is excited by a unit-amplitude plane SH wave impinging at an incidence angle α. The Cartesian coordinate system (x, y, z) is employed with the positive x–axis pointing to the right, the positive y–axis pointing downward and positive z-axis pointing out of the plane.

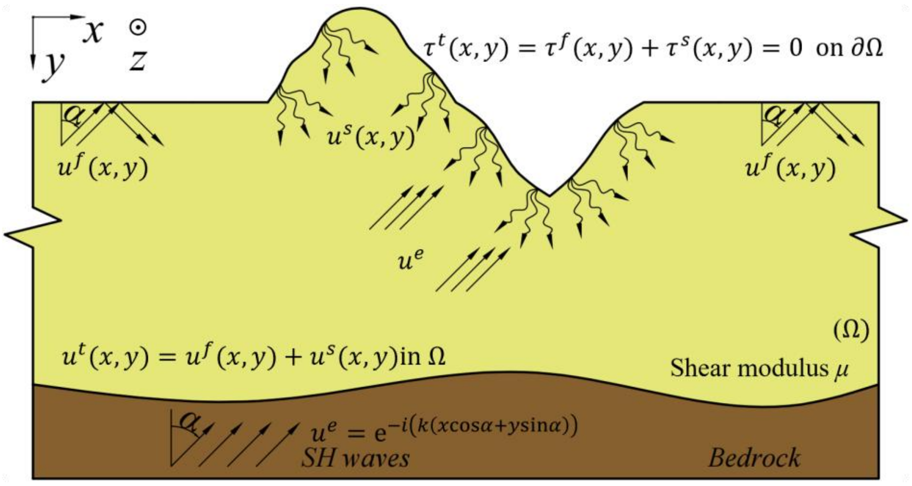


FIGURE 1 2D model depicting the scattering of SH waves.

The total displacement is oriented in the z-direction, denoted by $u^t(x, y)$ and governed by the Helmholtz

Equation within the half-space Ω:

$$\frac{\partial^2 u^t(x,y)}{\partial x^2}+\frac{\partial^2 u^t(x,y)}{\partial y^2}+k^2 u^t(x,y)=0\ , \tag{1}$$

where k is the shear wavenumber.

Assuming a unit-amplitude plane SH wave incident on the half-space at angle α, the unit plane wavefield u^e can be expressed as:

$$u^e(x,y)=\exp - i(k(x\cos\alpha + y\sin\alpha))\,. \tag{2}$$

The total traction $\tau^t(x, y)$ should satisfy the traction-free boundary condition on the boundary ∂Ω (i.e., the ground surface):

$$\tau^t(x,y)=0\,. \tag{3}$$

For convenience, the displacement amplitude $|U|$ and the dimensionless frequency η are defined as follows:

$$|U|=\sqrt{u_r{}^2+u_i{}^2}\ , \tag{4}$$

$$\eta=\frac{2d}{\lambda}=\frac{kd}{\pi}\ , \tag{5}$$

where λ is the wavelength; d is the canyon depth (or the mountain height); u_r and u_i are the real and imaginary components of the complex expression of the wavefield, respectively.

2.2 The PINN

PINN provides a unified framework for solving both forward and inverse PDE problems and can be applied to equations of the following form:

$$\mathrm{N}[u]=(\boldsymbol{x},p)=0\,, \tag{6}$$

$$\mathrm{B}[u]=(\boldsymbol{x},p)=0\,, \tag{7}$$

where $\mathcal{N}$ and $\boldsymbol{\mathcal{B}}$ are nonlinear differential and boundary condition operators, respectively. The latent solution $u(\boldsymbol{x}, \boldsymbol{p})$ can be approximated by a deep neural network(DNN) $u\ (\boldsymbol{x}, p, \boldsymbol{\theta})$ with parameters $\boldsymbol{\theta}$. The loss function for the network is defined as:

$$\mathrm{L}\left(u,\boldsymbol{\theta}\right)|=\lambda_p \mathrm{L}_p\left(\boldsymbol{\theta}\right)+\lambda_b \mathrm{L}_b\left(\boldsymbol{\theta}\right), \tag{8}$$

where $\boldsymbol{\mathcal{L}}_p(\boldsymbol{\theta})$ and $\boldsymbol{\mathcal{L}}_b(\boldsymbol{\theta})$ are the PDE residual and boundary condition loss terms, respectively, λ_p, λ_b are the weights for the PDE residual and boundary loss, respectively. The DNN takes the coordinate $\boldsymbol{x}$ as input and outputs the corresponding solution value at that location. The partial derivatives of $u(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{\theta})$ with respect to the input points can be conveniently computed using automatic differentiation to construct the loss function. The neural network is trained by minimizing the loss function $\boldsymbol{\mathcal{L}}(u;\ \boldsymbol{\theta})$ through the gradient descent method to obtain the network parameter $\boldsymbol{\theta}$ that minimizes the loss function.

However, vanilla PINN methods for solving the wave equation with free boundary conditions and complex geometries may yield inaccurate approximations. Moreover, because PINNs operate on finite computational domains, artificial absorbing boundaries must be introduced to truncate the half-space, which may incur additional computational costs.

2.3 The boundary elements method(BEM)

To avoid introducing artificial absorbing boundaries and to enable the network to stably enforce free boundary conditions, we reformulate the learning problem such that the neural network predicts a virtual force field defined on the physical boundary. Through the BEM, this virtual force field uniquely determines the displacement field throughout the half-space, thereby circumventing the need to directly learn the wave field.

The total wavefield displacement u^t is the superposition of the free field displacement u^f and the scattered field displacement u^s:

$$u^t(x,y) = u^f(x,y) + u^s(x,y) . \tag{9}$$

Given the incident wave defined in Eq. (2), the free wavefield displacement u^f can be expressed as

$$u^f = \exp(-ik(x\cos(\alpha) + y\sin(\alpha))) + \exp(-ik(x\cos(\alpha) - y\sin(\alpha))) . \tag{10}$$

Similarly, the traction-free boundary condition on the ground surface can be written as

$$\tau^t(x,y) = \tau^f(x,y) + \tau^s(x,y) = 0 , \tag{11}$$

where τ^f is the free field traction, and τ^s is the scattered field traction. Specifically, with the incident wave from Eq. (2), the free field traction is given by

$$\begin{aligned}\tau^f(x,y) = &-i\mu k n_x \cos\alpha(\exp(-ik(x\cos\alpha + y\sin\alpha)) + \exp(-ik(x\cos\alpha - y\sin\alpha))) \\ &-i\mu k n_y \sin\alpha(\exp(-ik(x\cos\alpha + y\sin\alpha)) - \exp(-ik(x\cos\alpha - y\sin\alpha))) ,\end{aligned} \tag{12}$$

where (n_x, n_y) is the unit normal vector at point (x, y).

Based on the BEM, the anti-plane steady-state problem studied herein is reformulated by ignoring the body force components. Consequently, the scattered field displacement and traction can be expressed in integral form as[53]

$$u^s(x,y) = \int_S G_{zz}((x,y),(x_0,y_0))\phi_z(x_0,y_0)dS , \tag{13}$$

$$\tau^s(x,y) = \int_S T_{zz}((x,y),(x_0,y_0))\phi_z(x_0,y_0)dS , \tag{14}$$

where $G_{zz}((x, y), (x_0, y_0))$ and $T_{zz}((x, y), (x_0, y_0))$ are the displacement and traction Green's function, respectively; ϕ_z is the distributed force of each boundary element (i.e., virtual force density). (x, y) and (x_0, y_0) denote field point and source point, respectively. Note that in this work, the source points and field points coincide at the same spatial locations.

The expression of half-space displacement Green's function is given by[54]

$$G_{zz}((x,y),(x_0,y_0)) = -\frac{i}{4\mu} H_0^{(2)}(kr) , \tag{15}$$

where $r = \sqrt{(x-x_0)^2 + (y-y_0)^2}$, and $H_0^{(2)}$ is the Hankel function of the second kind of order 0.

The half-space traction Green's function is expressed as

$$T_{zz}((x,y),(x_0,y_0)) = -\frac{i}{4r} H_1^{(2)}(kr)(\gamma_x n_x + \gamma_y n_y) , \tag{16}$$

where $H_1^{(2)}$ is the Hankel function of the second kind of order 1, n_x and n_y are the direction cosines of the boundary element and $\gamma_x = (x - x_0)/r, \gamma_y = (y - y_0)/r$.

Assuming the virtual force density ϕ_z is uniform over each boundary element, the surface is discretized into boundary elements, leading to a system of linear equations. To meet the accuracy requirements, the x-direction computational domain should be extended beyond the topography by eight times the wavelength and at least 10 boundary elements per wavelength are required. Consequently, the total wavefield displacement and the traction-free boundary condition in Eqs. (9) and (11) can be written as

$$u^t(x,y)_n = u^f(x,y)_n + \sum_{l=1}^{N_s} \phi_z((x_0,y_0)_l) \int_{(x_0,y_0)_l-\frac{\Delta S}{2}}^{(x_0,y_0)_l+\frac{\Delta S}{2}} G_{zz}((x,y)_n,(x_0,y_0)_l)dS_{(x_0,y_0)}, (n = 1,2,3...,N_r) \quad (17)$$

$$\tau^t(x,y)_n = \tau^f(x,y)_n + \sum_{l=1}^{N_s} \phi_z(x_0,y_0)_l \int_{(x_0,y_0)_l-\frac{\Delta S}{2}}^{(x_0,y_0)_l+\frac{\Delta S}{2}} T_{zz}((x,y)_n,(x_0,y_0)_l)dS_{(x_0,y_0)} = 0, (n = 1,2,3...,N_r) \quad (18)$$

where N_r and N_s are the numbers of field points and source points, respectively.

The integrals in Eqs. (17) and (18) are evaluated using a three-point Gaussian quadrature formula, except when the source point and the field point coincide (i.e., $x=x_0$ and $y=y_0$). In these cases, analytical expressions are obtained from an ascending series of Bessel functions[54]

$$\int_{(x_0,y_0)_l-\frac{\Delta S}{2}}^{(x_0,y_0)_l+\frac{\Delta S}{2}} G_{zz}((x,y)_n,(x_0,y_0)_l)dS_{(x_0,y_0)} = -\frac{i\Delta S}{4\mu}[1+i\frac{2}{\pi}(0.4228-\log(\frac{k\Delta S}{4}))], \quad (19)$$

$$\int_{(x_0,y_0)_l-\frac{\Delta S}{2}}^{(x_0,y_0)_l+\frac{\Delta S}{2}} T_{zz}((x,y)_n,(x_0,y_0)_l)dS_{(x_0,y_0)} = \frac{1}{2}, \quad (20)$$

where i denotes the imaginary unit.

In summary, the governing equation and boundary conditions originally represented by Eqs. (1) and (3) are reformulated into Eqs. (17) and (18). The unknowns in Eqs. (17) and (18) are virtual force density ϕ_z at the source points on the boundary. Once Eq. (13) is enforced and the unknowns are solved, the total wavefield is generated using Eq. (17). Note that the wavefield solution in Eq. (17) is a linear combination of fundamental solutions. Since the fundamental solutions inherently satisfy the Sommerfeld radiation condition, the wavefield solution naturally satisfies this condition without the need for an artificial absorbing boundary.

2.4 The implementation of FF PINN-BEM

To approximate the virtual force density $\phi(x, y, \psi)$, we construct a deep feedforward neural network shown in Figure 2 with N+1 layers, comprising the input layer, N−1 hidden layers, the output layer. Nonlinear activation functions are employed only in the hidden layers, while the output layer remains linear. For each hidden layer l, the output (ϕ^l) is given by:

$$\phi^l = \varpi(\mathbf{W}^l \phi^{l-1} + \mathbf{b}^l), \quad (21)$$

where ϖ denotes a nonlinear activation function. The loss function is designed to quantify the deviation between the model's predicted results and the reference, guiding the optimizer to update the weights **W** and biases **b** iteratively toward an optimal set of parameters. For all experiments in this paper, we employ a NVIDIA GeForce RTX 3060 12 GB GPU to train the networks.

In this work, the Gaussian Error Linear Unit (GELU) activation function is employed due to its smooth and continuously differentiable nature, which improves gradient stability when computing spatial derivatives in PINN[55]. The network parameters are initialized using the Xavier initialization scheme[56]. To reduce overfitting and stabilize the training process, L2 regularization is incorporated by penalizing excessively large network weights. Based on BEM, our network is designed not to output the displacement solution $u(x, y, \psi; \boldsymbol{\theta})$ directly. Instead, it outputs the real and imaginary components of the virtual force $\phi(x, y, \psi; \boldsymbol{\theta})$. Thus, the network is trained by enforcing Eq. (18) solely on the boundary. Moreover, both the physical parameters ψ and the spatial coordinates are fed to the network, ensuring that the virtual force $\phi(x, y, \psi; \boldsymbol{\theta})$ is a continuous function of the spatial coordinates and the physical parameters ψ. The loss function is given by:

$$\mathrm{L}(\theta)=\frac{1}{N_r}\sum_{n=1}^{N_r}\left|\,\tau_n^f(x,y,\psi)+\sum_{l=1}^{N_s}\hat{\phi}(x_0,y_0,\boldsymbol{\theta})_l\int_{(x_0,y_0)_l-\frac{\Delta S}{2}}^{(x_0,y_0)_l+\frac{\Delta S}{2}}T_{zz}((x,y)_n,(x_0,y_0)_l)dS_{(x_0,y_0)}\,\right|^2 , \tag{22}$$

where $\hat{\phi}(x_0, y_0, \boldsymbol{\theta})_l$ is the predicted virtual force in complex value;

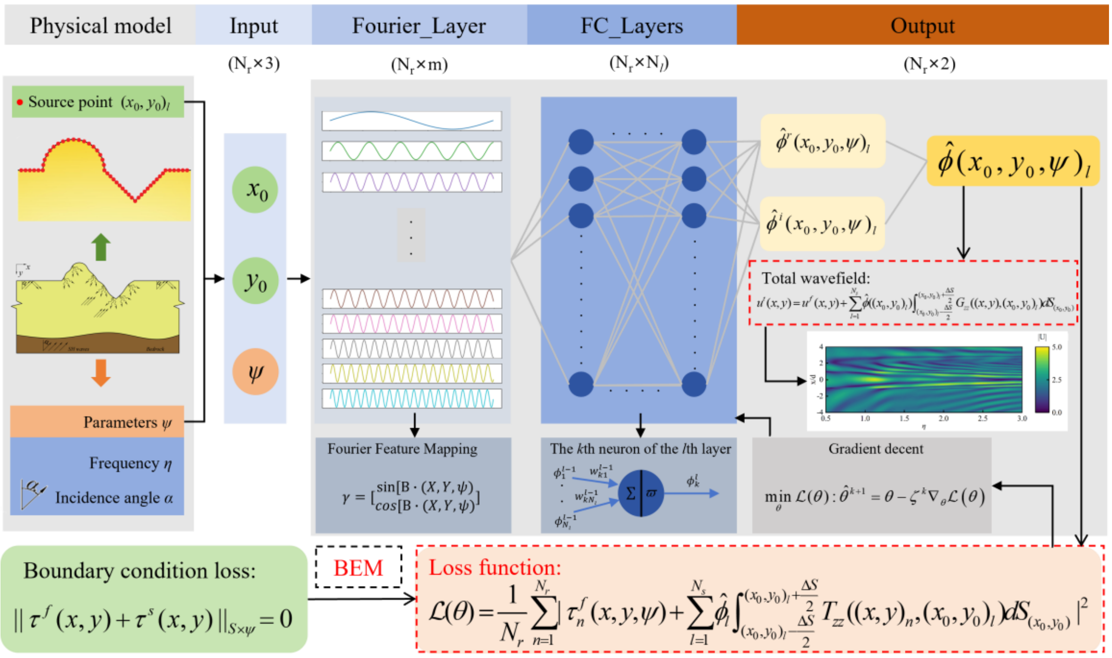


FIGURE 2 FF PINNs-BEM architecture diagram for simulating the scattering of SH wave

According to neural tangent kernel (NTK) theory[46], an invariant NTK exists during neural networks training, and analysis of its eigenvalue spectra reveals a spectral bias: coordinate-based multilayer perceptron (MLP) corresponds to a rapid decay of the NTK eigenvalues, which limits the network's capacity to simulate high-frequency functions.

To address this issue, Tancik et al.[48] introduced Fourier feature mapping, transforming the NTK into a static (shift-invariant) kernel and adjusts its spectrum through hyperparameter tuning. This method maps the low-dimensional standard coordinates into the high-dimensional feature space, from which these features are then passed into the network. Let $\boldsymbol{v} = \{x, y, \psi\}$ represent the inputs. The Fourier feature mapping $\gamma(\boldsymbol{v})$ and the first hidden layer are written as:

$$\gamma(\boldsymbol{v})=\left[\cos(\mathbf{B}\boldsymbol{v}),\sin(\mathbf{B}\boldsymbol{v})\right]^T , \tag{23}$$

$$\phi^1=\varpi(\mathbf{W}^1\gamma(\boldsymbol{v})+\mathbf{b}^1) , \tag{24}$$

where the Fourier basis frequency $\mathbf{B}\in\mathbb{R}^{d\times m}$ is sampled from a normal distribution $\mathcal{N}(0,\sigma^2)$; m is the number of basis functions, d is the dimension of the input coordinate and σ is the standard deviation of the normal distribution. $\mathbf{B}$ can be kept constant or allowed to be trainable. In this study, $\mathbf{B}$ is set with fixed parameters.

The scale of the Fourier basis frequency distribution and the resulting NTK feature space are determined by σ and m directly. These hyperparameters are usually determined by hyperparameter search or based on a priori frequency information. Notably, the optimal network performance can only be achieved when σ is fine-tuned within a specific range.

2.5 Transfer learning

To improve computational efficiency in multi-frequency simulations, we formulate the learning problem as a frequency-dependent transfer learning task. Specifically, the computational domain of the dimensionless frequency η is partitioned into several adjacent subdomains. Each subdomain corresponds to a distinct learning task because the virtual traction field depends on the frequency parameter.

2.5.1 Definition of Source and Target Tasks

Let the dimensionless frequency range $\mathrm{D}_\eta = [\eta_{\min}, \eta_{\max}]$ be partitioned in to K adjacent subdomains:

$$\mathrm{D}_\eta = \bigcup_{k=1}^{K} \mathrm{D}_\eta^{(k)}, \ \mathrm{D}_\eta^{(k)} = [\eta_{\min}^{(k)}, \eta_{\max}^{(k)}]. \tag{25}$$

For each subdomain $\mathrm{D}_\eta^{(k)}$, we define a learning task

$$\mathrm{T}^{(k)} = \{(x, y; \eta), \phi(x, y, \eta)\}, \tag{26}$$

where the objective is to approximate the virtual force $\phi(x, y; \eta)$ over the boundary $\partial\Omega$ using a neural network.

Given three adjacent subdomains $\mathrm{D}_\eta^{(k)}$, $\mathrm{D}_\eta^{(k-1)}$ and $\mathrm{D}_\eta^{(k+1)}$ we define the source task T_s and target task T_t as follows:

$$\mathrm{T}_s = \mathrm{T}^{(k)}, \tag{27}$$

$$\mathrm{T}_t = \mathrm{T}^{(k\pm1)}. \tag{28}$$

The tasks differ only in the dimensionless frequency η; all other physical settings (topography, incident angle) remain unchanged.

2.5.2 Transfer Strategy

Let $\boldsymbol{\theta}_s$ be the optimized network parameters obtained from solving the source task:

$$\boldsymbol{\theta}_s = \arg\min_{\boldsymbol{\theta}} \mathrm{L}\left(\boldsymbol{\theta}; \mathrm{T}_s\right). \tag{29}$$

Instead of random initialization, the target-task training is initialized as:

$$\boldsymbol{\theta}_t^{(0)} = \boldsymbol{\theta}_s, \tag{30}$$

and then optimized on T_t:

$$\boldsymbol{\theta}_t = \arg\min_{\boldsymbol{\theta}} \mathrm{L}\left(\boldsymbol{\theta}; \mathrm{T}_t\right). \tag{31}$$

Since wave fields at nearby frequencies share similar spatial patterns and satisfy the same boundary integral formulation, the pretrained model from T_s contains useful physical priors for T_t. Therefore, the network parameters (weights and biases) learned in T_s are used to initialize the network in T_t, instead of training from scratch. This design effectively reduces optimization difficulty, accelerates convergence, and minimizes the computational cost of simulating wavefields across a broad frequency spectrum.

2.6 Evaluation mertics

Two mertics are adopted to assess the performance of the proposed approach. First, the coefficient of determination (R^2) is employed to quantify the model's fitting capacity, that is

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (u_i^{ref} - \hat{u}_i)^2}{\sum_{i=1}^{n} (u_i^{ref} - \bar{u}^{ref})^2}, \tag{32}$$

where $\hat{u}_i$ is the predicted value, u_i^{ref} is the true value, and $\bar{u}^{ref}$ is the average value of the true values. An R^2 value closer to 1 indicates that the model better fits the data, effectively reflecting the predictive ability and reliability of the model.

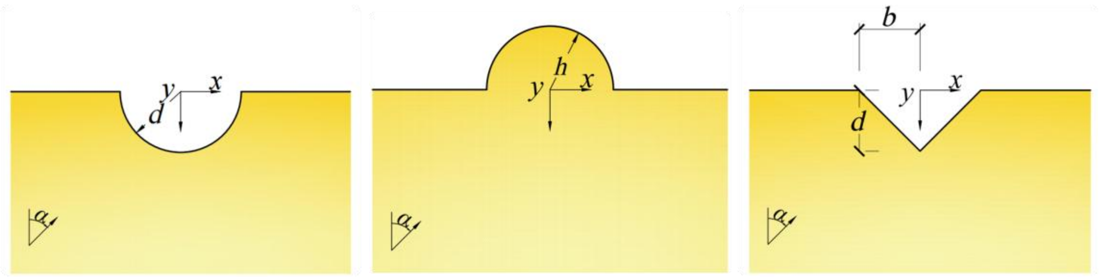
Second, the accuracy of predictions $u(x, y, \psi; \boldsymbol{\theta})$ for both plain PINNs (i.e., PINNs without the Fourier feature layer) and FF PINNs are quantified in terms of the relative error ε, defined as

$$\varepsilon = \frac{1}{N} \sum_{i=1}^{N} \frac{|(\hat{u}_i - u_i^{ref})|}{u_i^{ref}}, \tag{33}$$

where $\hat{u}_i$ is the prediction result u_i^{ref} denotes the true displacement as obtained from analytical solutions and BEM. N denotes the number of sampling points for testing.

3 SOLUTION VERIFICATION

In this section, the model predictions are validated against exact solutions corresponding to SH-wave scattering caused by three distinct topographies: a semi-cylindrical canyon (Trifunac[4]), a semi-cylindrical mountain (Yuan and Men[8]) and a V-shaped canyon (Tsaur and Chang[9]), as illustrated in Figure 3 with model parameters listed in Table 1. To simulate real-world conditions with varying incident angles, the results for four incident angles (0°, 30°, 60°, 90°) and the dimensionless frequency $\eta = 2.0$ for the above topographies are evaluated. In the proposed framework, collocation points are sampled exclusively on the physical boundary. Therefore, the number of collocation points is identical to the number of boundary elements used for discretization, which is summarized in Table 2 for all numerical cases. The network consists of six hidden layers, each containing 150 neurons. A hybrid optimization strategy is adopted, in which the network is first trained using the AdamW optimizer with a learning rate of 1e-3 for 40,000 epochs, followed by 10,000 optimization steps using the L-BFGS optimizer with a learning rate of 1e-1 to further improve convergence stability. The Fourier feature mapping employs 64 basis functions with a standard deviation of 1.0. For comparison, additional results obtained using AdamW-only optimization are provided in Appendix A.



(a) Semi-cylindrical canyon (b) Semi-cylindrical mountain (c) V-shaped canyon

FIGURE 3 Three representative topographies.

Figures 4(a)-(d), Figures 5(a)-(d) and Figures 6(a)-(d) compare the surface displacement amplitude of semi-cylindrical canyon, semi-cylindrical mountain and symmetric V-shaped canyon predicted by the FF PINN with the analytical solutions for four incident angles, respectively. Table 2 shows the relative errors (ε) and the coefficient of determination (R^2) for these predictions, demonstrating that the FF PINN results closely match the analytical solution by (Trifunac, 1973[4]; Yuan and Men, 1992[8]; Tsaur and Chang, 2008[9])

TABLE 1 Model parameters of three topographies

Topography	Topographic size(m)	x-direction computational domain(m)	Element number	
			flat surface	terrain surface
Semi-cylindrical canyon	$d = 1{,}000$	$[-8\lambda-d, 8\lambda+d]$	240	100
Semi-cylindrical mountain	$h = 1{,}000$	$[-8\lambda-h, 8\lambda+h]$	240	100
Symmetric V-shaped canyon	$d = 1{,}000$, $b = 1{,}000$	$[-8\lambda-b, 8\lambda+b]$	240	100

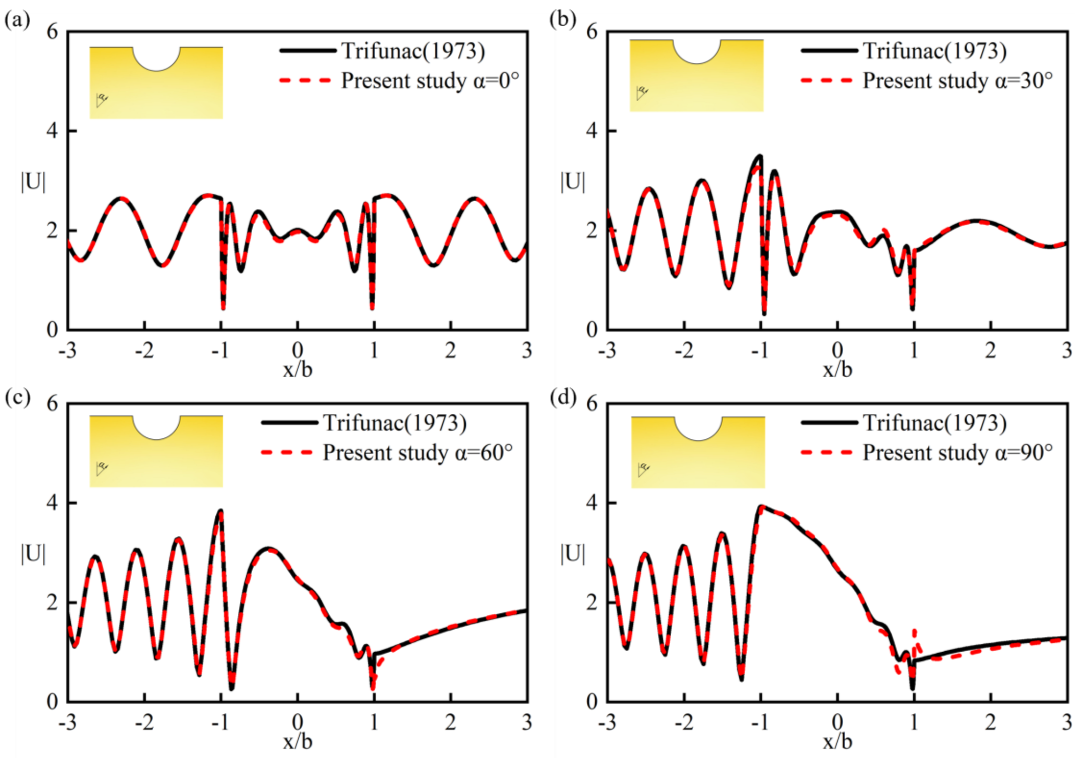


FIGURE 4 Comparisons of surface displacement amplitudes for semi-cylindrical canyon: (a) *α*=0°; (b) *α*=30°; (c) *α*=60°; (d) *α*=90°.

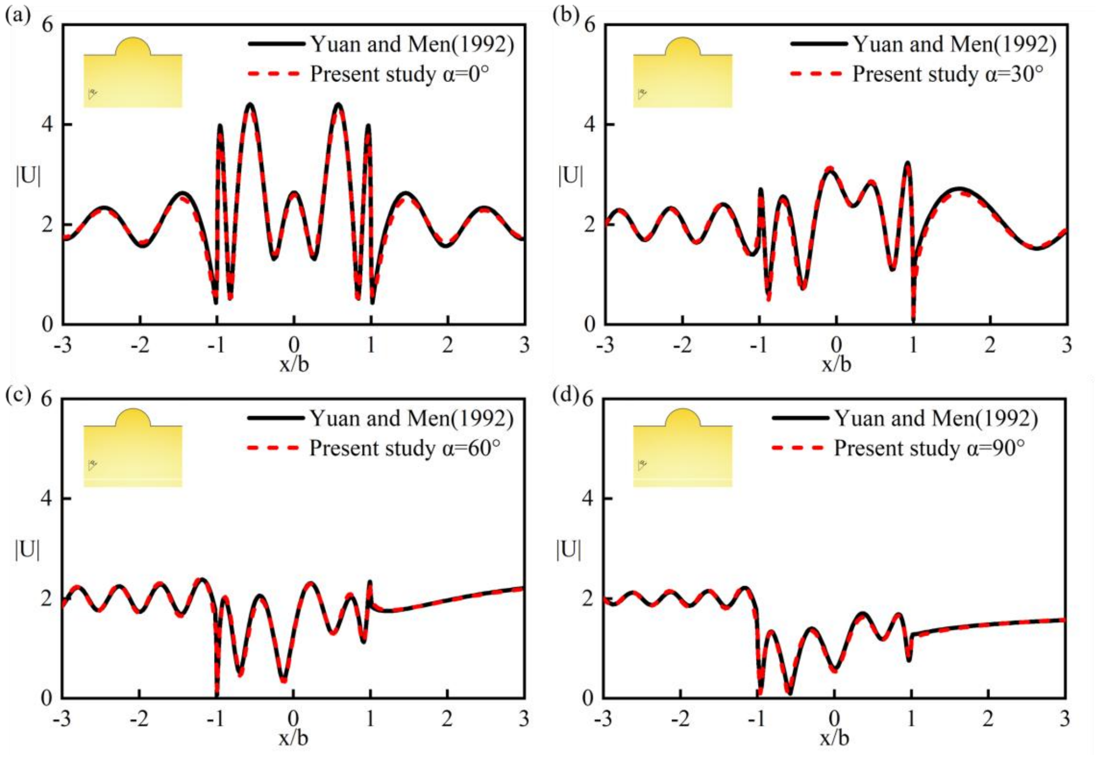


FIGURE 5 Comparisons of surface displacement amplitudes for semi-cylindrical mountain: (a) *α* = 0°; (b) *α* = 30°; (c) *α* = 60°; (d) *α* = 90°.

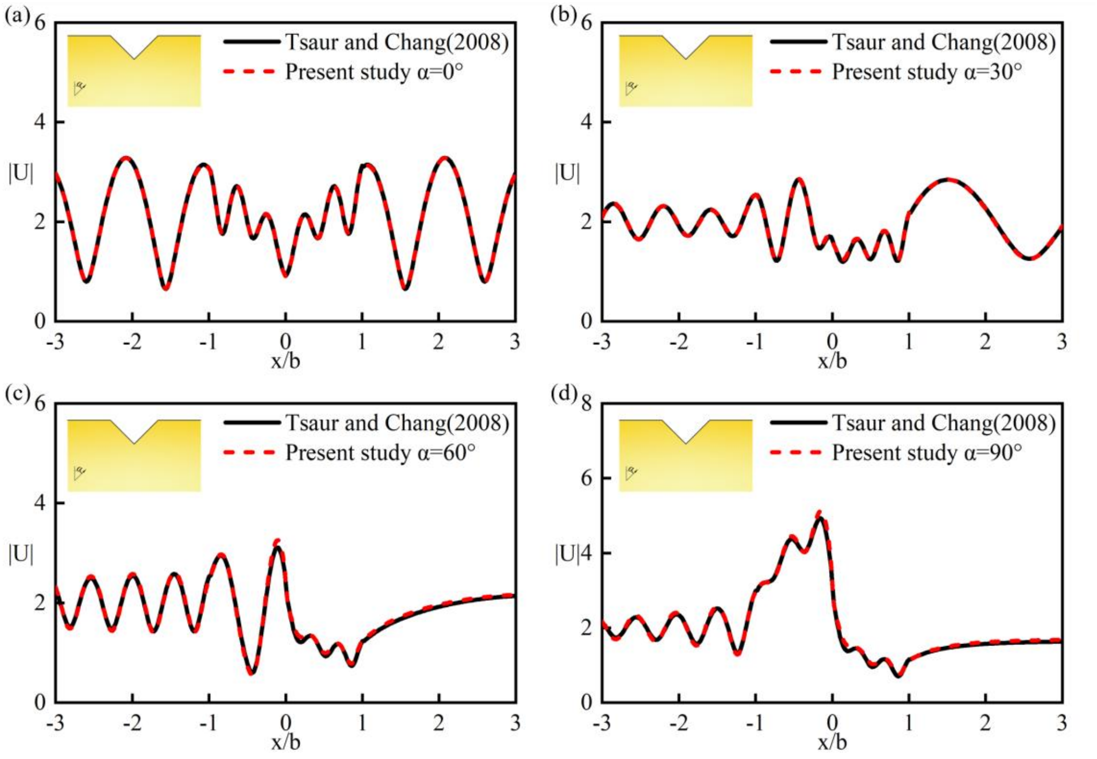


FIGURE 6 Comparisons of surface displacement amplitudes for symmetrical V-shaped canyon: (a) *α* =

0°; (b) α = 30°; (c) α = 60°; (d) α = 90°.

TABLE 2 Relative errors and coefficients of determination of predicted surface displacement amplitudes for three topographies.

Topography	α	ε-FF PINN	R^2-FF PINN
Semi-cylindrical canyon	0°	5.85×10^{-3}	0.997
	30°	1.12×10^{-2}	0.993
	60°	1.57×10^{-2}	0.994
	90°	2.61×10^{-2}	0.994
Semi-cylindrical mountain	0°	1.80×10^{-2}	0.991
	30°	1.60×10^{-2}	0.984
	60°	1.25×10^{-2}	0.992
	90°	1.96×10^{-2}	0.993
Symmetrical V-shaped canyon	0°	2.51×10^{-3}	0.999
	30°	5.59×10^{-3}	0.999
	60°	2.15×10^{-2}	0.994
	90°	2.27×10^{-2}	0.997

4. NUMERICAL RESULTS AND DISCUSSIONS

4.1 Parameterized surface displacement simulation

In this study, we simulate surface displacement for multiple frequencies and angles of the incident wave. Each input sample fed into the neural network comprises three dimensions, given by $\boldsymbol{v} = \{x_i, y_i, \psi_i\}$, where i is the sample index, ψ_i denotes the parameter of interest (i.e., incident angle α and dimensionless frequency η). The symmetric V-shaped canyon shown in Figure 3 is adopted here as a test of the influence of topography due to its more complex scattering behavior relative to the other two terrains. To accurately simulate parameterized surface displacement amplitudes, both the plain PINN and the FF PINN in this section contain eight hidden layers with 200 neurons per layer. The network was trained using a hybrid optimization strategy, consisting of 50,000 epochs of AdamW optimization followed by 100 epochs of the L-BFGS optimizer. To ensure adequate representation capacity, the number of Fourier basis functions m is set to 128. Note that the plain PINN and the FF-PINN have comparable model sizes. Although the Fourier feature mapping increases the input dimensionality of the FF-PINN and consequently leads to a modest increase in the number of trainable parameters in the first layer, the overall number of parameters remains of the same order of magnitude for both networks, since the network depth and the number of neurons per hidden layer are kept identical.

4.1.1 Incident angle parameterized surface displacement simulation

For the incident angle parameterization, the incident angle of the plane wave is treated as a network input over the domain [0°, 45°], and to reduce computational cost we sample this domain at 5° increments. Furthermore, three cases of η=5.0, 6.0, and 7.0 are considered to evaluate the networks' performance in solving the angle-parameterized equations.

Since sampling collocation points only at specific angles may lead to overfitting, an L2 regularization term is introduced to enhance the generalization ability of the neural network. To further investigate its impact on generalization performance, an additional parametric study is conducted and reported in Appendix B. However, the plain PINN suffers from spectral bias: it tends to generate low-frequency solutions, globally resulting in a poorer generalization ability when the dimensionless frequency increases. Figures 7 shows the results of plain PINN and FF PINN for simulating the V-shaped canyon surface displacement amplitude. We define $\delta_p(|U|)$ and $\delta_f(|U|)$ to represent the pointwise absolute errors in surface displacement amplitudes predicted by the plain PINN and FF PINN relative to the BEM, respectively. It shows that the increase of dimensionless frequency leads to the deterioration of Plain PINN performance and the error is concentrated in the high-frequency cases ($\eta = 6.0$, and 7.0). To solve this problem, we adjusted the standard deviation σ of the Fourier features according to the dimensionless frequency to enhance the network's capacity to generate high-frequency wavefields.

Figures 8(a)-(c) show the training loss curves for various values of σ at $\eta = 5.0$, 6.0 and 7.0. The results illustrate the training loss of the plain PINN decreases at a slow rate and the final value is one order of magnitude larger than that of FF PINN, demonstrating that the ability of FF PINN to accurately simulate surface displacements under multi-angle incidence is better than plain PINN.

Note that Song et al[45] showed that optimal results can be achieved by tuning the standard deviation so that the maximum value of the Fourier basis frequency $\mathbf{B}$ corresponds to the maximum wave number $k_{\max}$ in the equation. However, because our network outputs the virtual load (an intermediate quantity) rather than directly outputting the wavefield displacement, the optimal standard deviation still needs to be determined by hyperparameter sweeping. Figure 9 illustrates the results of this hyperparameter σ sweeping for $\eta = 5.0$, 6.0, and 7.0. The scatter plots exhibit a U-shaped trend as σ varies, indicating that both excessively large and excessively small σ values adversely affect the model's accuracy.

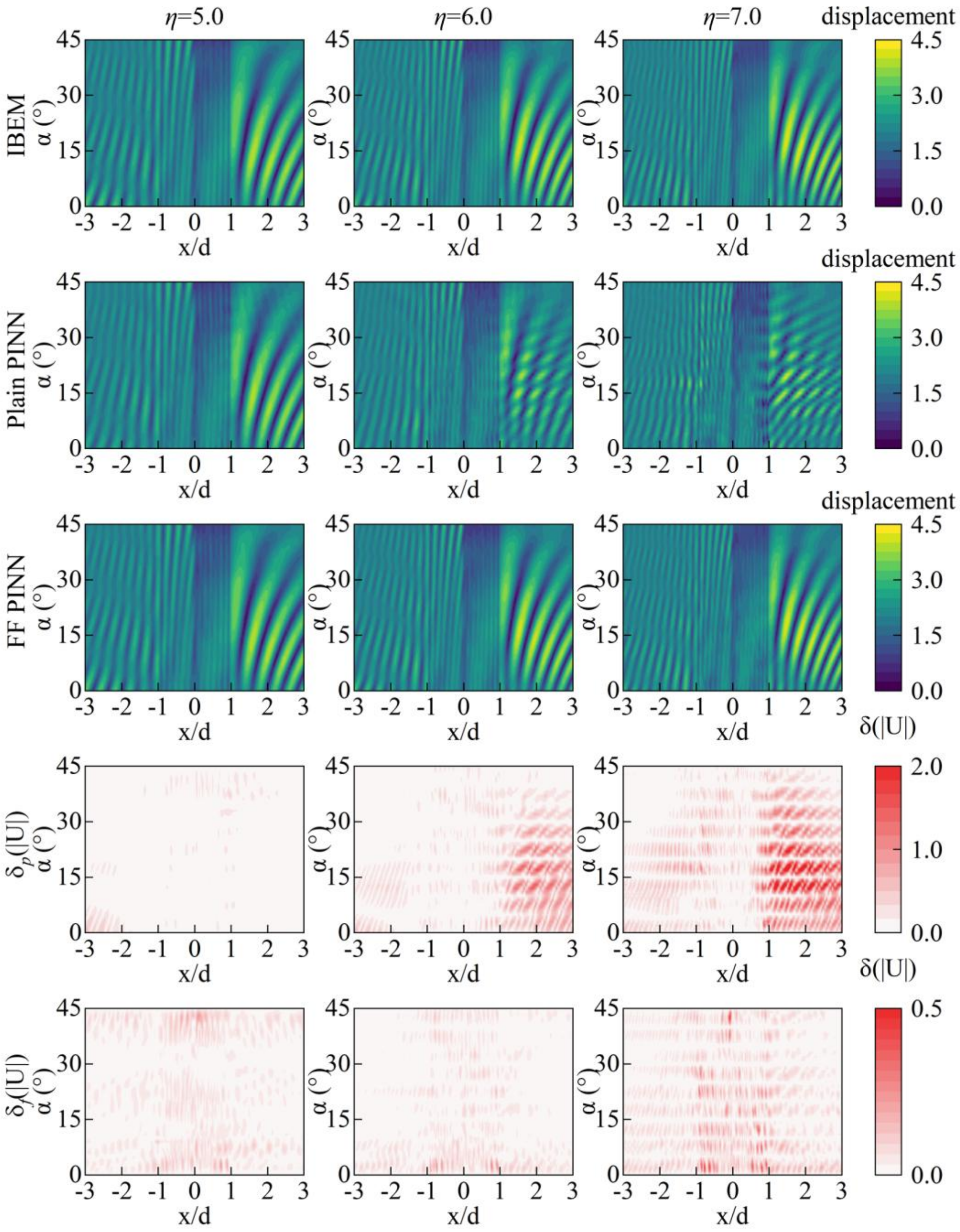


FIGURE 7 Comparison of surface displacement amplitudes for a symmetric V-shaped canyon predicted by plain PINNs and FF-PINNs, with BEM solutions as reference, under various incidence angles.

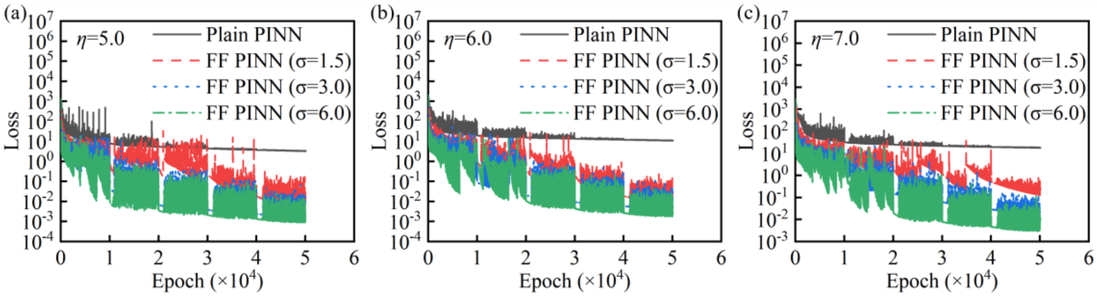


FIGURE 8 Training loss curves for different network configurations: (a) η = 5.0; (b) η = 6.0; (c) η = 7.0.

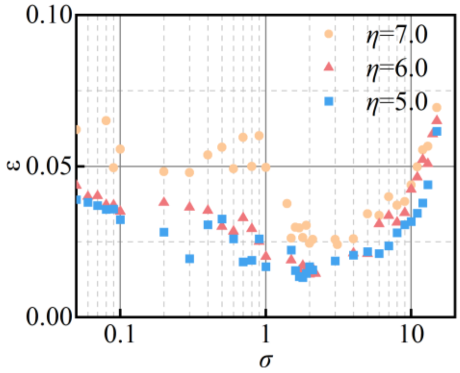


FIGURE 9 Relative error ε versus σ hyperparameter sweeps with Fourier features mapping sampled from a Gaussian distribution.

4.1.2 Dimensionless frequency parameterized surface displacement simulation

To perform the time-domain analysis, the frequency-domain wavefield is first computed at discrete frequency points and treated as the transfer function of the linear system. The time-domain incident wave is transformed into the frequency domain, and the corresponding wavefield response is obtained by multiplying the transfer function with the Fourier spectrum of the incident wave at each frequency. The time-domain wavefield is then reconstructed by applying an inverse Fourier transform.[57] In contrast, the conventional BEM for solving the frequency domain wavefield must calculate the integral of the Green's function at that frequency, which is computationally demanding.

To reduce the computational cost, the proposed method generates collocation points at only a few dimensionless frequency points during training. In other words, both the solution of the wavefield over the entire x-direction computational domain and the integration of the traction Green's function (see Eq. 9) are computed only for a few selected frequencies. This strategy avoids repeated integration across all frequencies.

The method is tested on the dimensionless frequency computational domain [4.0, 4.5] with an incident wave angle of 0°. The training collocation points are chosen at the endpoints and midpoint of this domain, i.e., at η = 4.0, 4.25, and 4.5. To satisfy the global accuracy requirement, the x-direction computational domain is $S = [-8\lambda_{\max}-b, 8\lambda_{\max}+b]$ m, where $\lambda_{\max}$ is the wavelength corresponding to η = 4.0. On the flat surfaces on both sides of the canyon, 240 boundary elements are assigned to each side, while 400 boundary elements are assigned along the V-shaped canyon surface. The method's accuracy is evaluated at 11 points from η = 4.0 to 4.5 in 0.05 increments, since over 10 frequency samples are typically required for accurate inverse Fourier transformation.

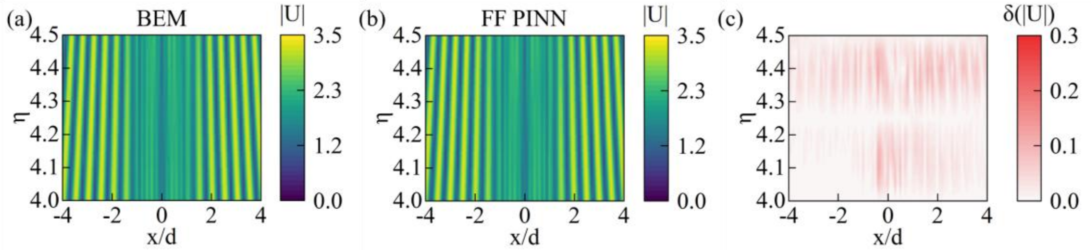


FIGURE 10 Comparisons of symmetric V-shaped canyon surface displacement amplitudes between BEM and FF PINNs for different dimensionless frequency.

Figure 10 compares the symmetric V-shaped canyon surface displacement amplitudes, as simulated by the FF PINNs, with the BEM reference over the specified frequency domain. The results demonstrate that the network generalizes well throughout the entire frequency computational domain, achieving an average relative error ε of 1.26% and the coefficient of determination R^2 is 0.997 over the test points. In the same case, BEM requires 9.9 hours for the surface displacement calculation, while the proposed FF PINNs take 7.2 hours. Moreover, Table 3 compares the two methods with varying numbers of dimensionless frequency calculation points. With additional calculation points, the efficiency advantage of the FF PINN becomes even more pronounced. For instance, when the number of frequency calculation points reaches 49, the BEM computation requires 38.8 hours, whereas the FF PINN completes the calculation in 20.0 hours, approximately halving the computation time of the BEM.

TABLE 3 Time cost assessment for FF PINN versus BEM under different numbers of dimensionless frequencies (N_η).

Method	Time (hours)		
	$N_\eta = 11$	$N_\eta = 25$	$N_\eta = 49$
FF PINN[a]	7.2	12.1	20.1
BEM	9.9	19.8	38.8

[a] Time includes network training, wavefield prediction, and Green's function computation.

The tests confirm that FF PINN can accurately simulate surface displacement amplitudes induced by irregular topography for multiple frequencies. However, the trained model generalizes well only within the dimensionless frequency range used for training, and cases outside the training range require retraining. To mitigate this, transfer learning is employed. By using a model pre-trained on a specific dimensionless frequency domain as an initialization, the network can rapidly adapt to a related frequency range because it already contains the virtual load features.

This transfer learning approach was validated by employing the trained network from $\eta = [4.0, 4.5]$ as a pre-trained model for new domains $\eta = [3.0, 3.5]$ and $\eta = [5.0, 5.5]$. Figure 11 shows the loss function curves of the network trained with the pre-trained model and trained from scratch for both cases. For the majority of the training process, the loss values of transfer learning remain consistently lower than those of scratch training throughout the training process. The model trained from scratch requires more than 50,000 epochs with the AdamW optimizer to achieve the same level of performance as the pre-trained model, which converges in only 20,000 epochs, indicating it is more efficient than scratch training. By adopting a transfer learning approach, the time cost of network training was reduced by half. Figure 12 shows a comparison between the reference displacement of the BEM and the result of the FF PINN optimized by the AdamW optimizer for 20,000 epochs. It demonstrates that the predicted results do not differ significantly from the BEM results. These show that transfer learning can also accurately predict surface displacement amplitude after training with fewer epochs than scratch training.

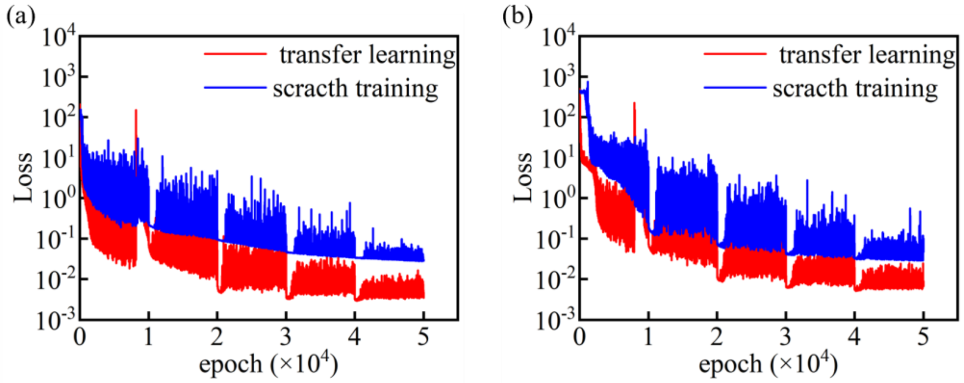


FIGURE 11 Loss function curves for networks trained with a pre-trained model for η = [4.0, 4.5] versus from scratch: (a) η = [3.0, 3.5]; (b) η = [5.0, 5.5].

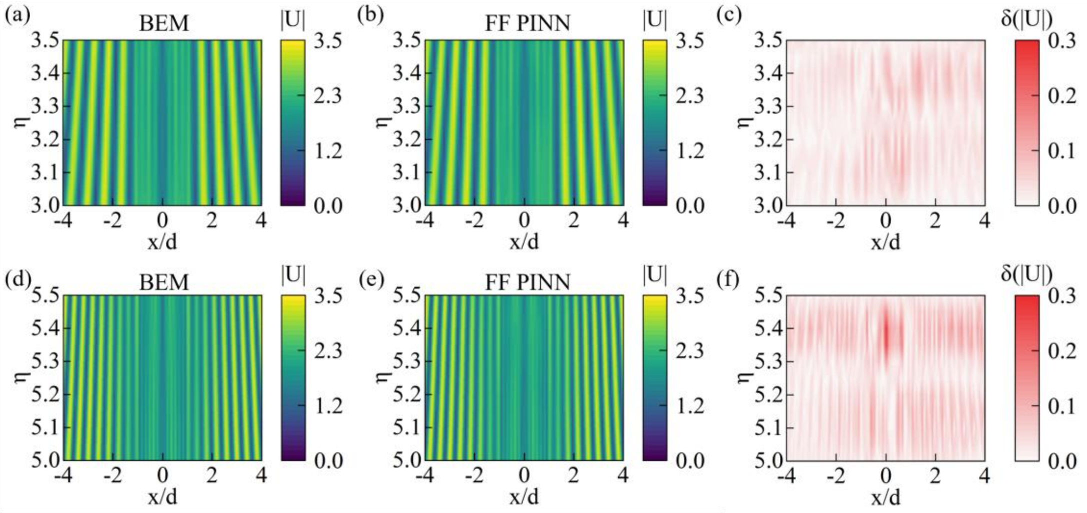


FIGURE 12 Comparison of surface displacement amplitudes for the symmetric V-shaped canyon between BEM and FF PINNs trained with a pre-trained model over two frequency domains: (a) η = [3.0, 3.5]; (b) η = [5.0, 5.5].

4.2 A coupled terrain of a semi-cylindrical mountain and a V-shaped canyon example

The model shown in Figure 1 is simplified to a terrain consisting of a semi-cylindrical mountain coupled with a V-shaped canyon terrain (see Figure 13). The model parameters are set as follows: canyon depth d = 1,000 m, canyon half-width b = 1,000 m, and mountain height h = 1,000 m. The frequency η ranges from 0.01 to 3.0, and is divided into six subdomains of width 0.5, namely [0.01, 0.5], [0.5, 1.0], [1.0, 1.5], [1.5, 2.0], [2.0, 2.5], and [2.5, 3.0]. The ith x-direction computational domain is $S_i = [-8\lambda_{\max,i}-h, 8\lambda_{\max,i}+h+2b]$ m, where $\lambda_{\max,i}$ represents the maximum wavelength in ith subdomain. Note that $\lambda_{\max,i}$ in subdomain [0.01, 0.5] is the wavelength corresponding to η = 0.25. In each subdomain, 11 frequency points are selected to generate the training collocation points. For spatial discretization, 240 boundary elements are assigned to each of the flat surfaces on either side of the terrain. Additionally, the V-shaped canyon surface is discretized using $200/\lambda_{\min}$ boundary elements, and the semi-cylindrical mountain surface is discretized with $32/\lambda_{\min}$ boundary elements, where $\lambda_{\min}$ represents the minimum wavelength value in each subdomain. To simulate the terrain's surface displacement, a network with 8 hidden layers, each containing 200 neurons, is optimized over 50,000 training epochs using AdamW with a learning rate of 1e-3. Subsequently, the trained network is further refined using the L-BFGS optimizer for 100 epochs. The number of Fourier basis functions $m = 256$ with a standard deviation $\sigma = 3.0$.

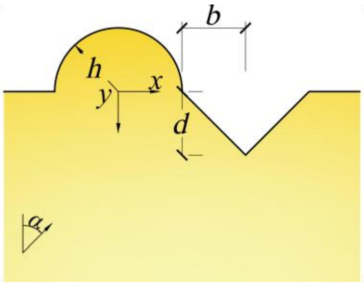
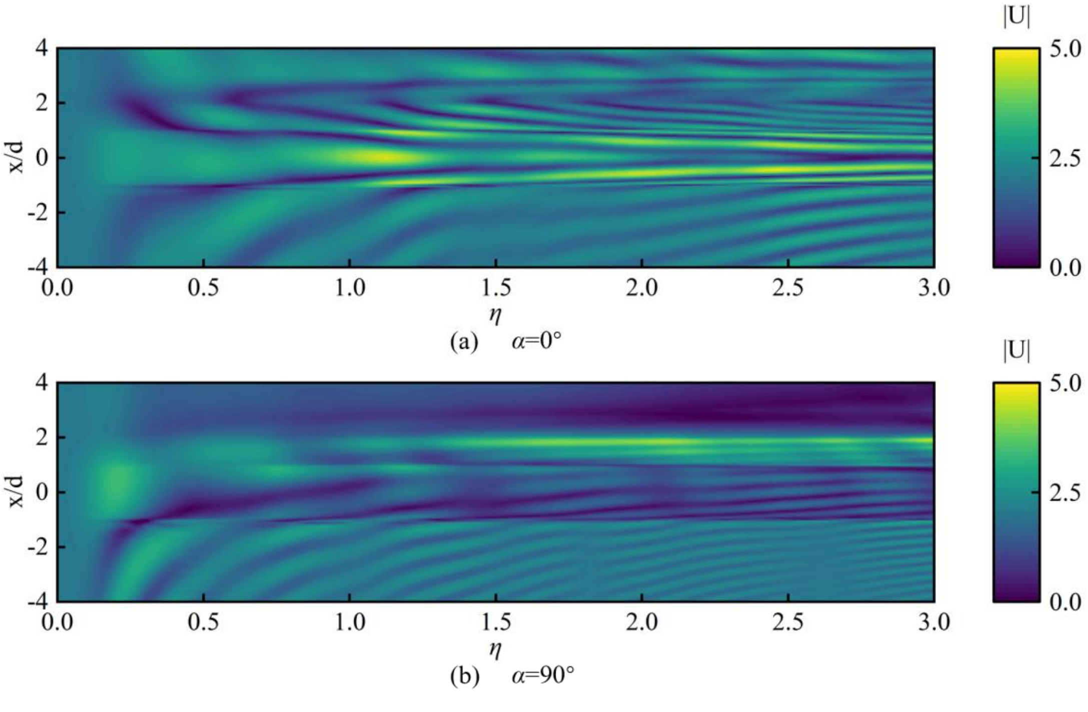


FIGURE 13 Coupled terrain of semi-cylindrical mountain and V-shaped canyon

The trained networks can efficiently generate surface displacements at any frequency in the computational domain. Figures 14(a) and (b) present the ground surface displacement amplitude spectra for two incidence angles over the frequency range [0.01, 3.0]. Specifically, for an incidence angle of α = 0º, the peak displacement amplitude is concentrated near the mountain top (i.e., $x/d = 0$), while for an incidence angle of α = 90º, the peak displacement amplitude is distributed near the bottom of the V-shaped canyon (i.e., $x/d = 2$).



(a) α=0°

(b) α=90°

FIGURE 14 Ground surface displacement amplitude spectra for two incidence angles: (a) $\alpha = 0°$; (b) $\alpha = 90°$.

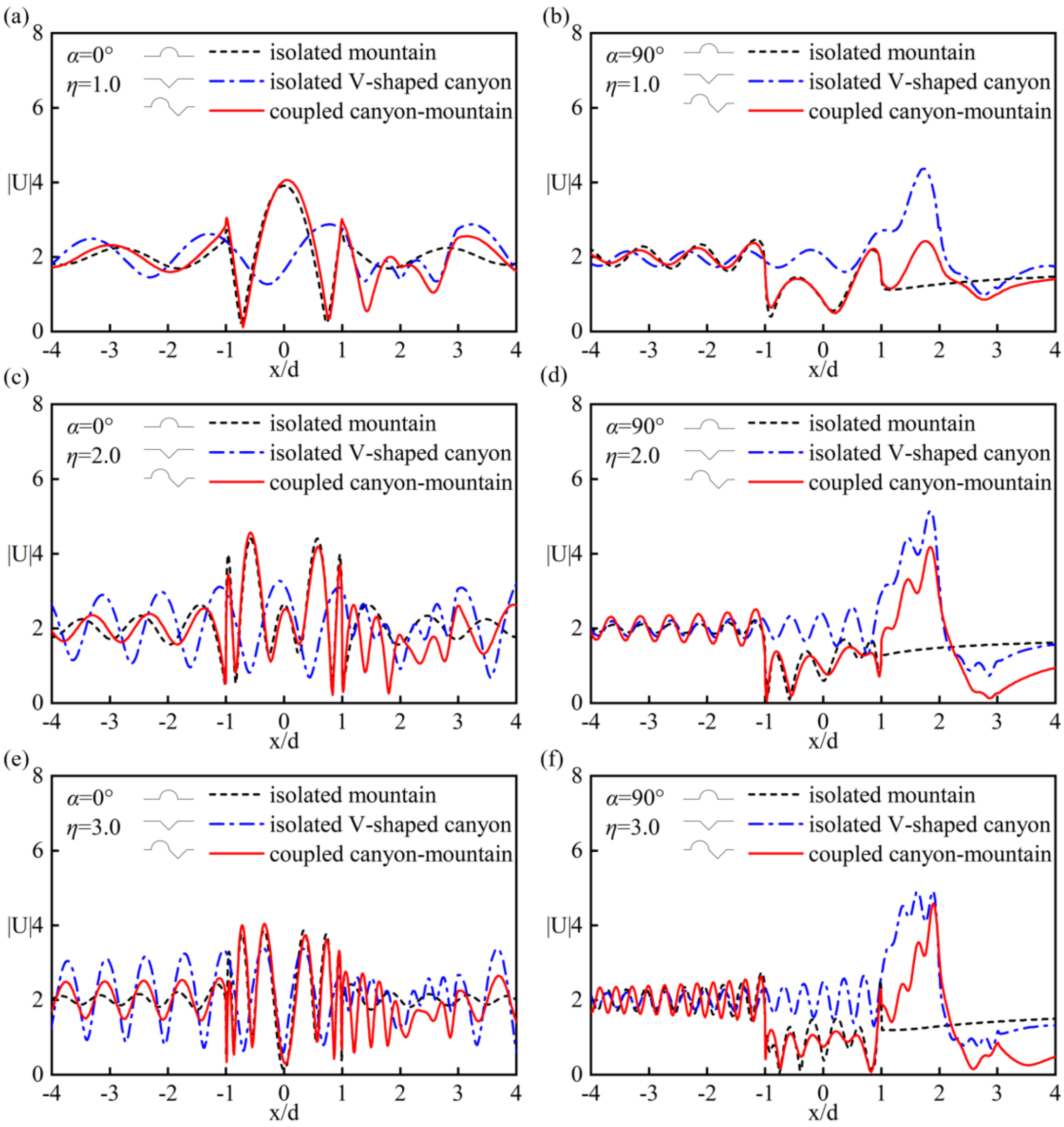


FIGURE 15 Displacement amplitude around the surface of three terrain scenarios for various *η* (1.0 and 2.0) and *α* (0° and 90°).

Figures 15(a)-(f) show the displacement amplitude of the ground surface for three terrain scenarios (i.e., isolated semi-cylindrical mountain, isolated V-shaped canyon, and the coupled semi-cylindrical mountain–V-shaped canyon terrains), for three dimensionless frequencies ($\eta = 1.0$, 2.0 and 3.0) and two incidence angles ($\alpha = 0°$ and 90°). Note that the coordinate origin is at the center of the mountain and the bottom of the canyon corresponds to x/d=2. It can be observed that the coupled semi-cylindrical mountain–V-shaped canyon terrain has a significant effect on seismic wave propagation.

Specifically, as shown in Figures 15(a), 15(c), and 15(e), at vertical incidence ($\alpha = 0°$) the displacement pattern on the mountain region and its adjacent flat surface with $x/d < 1$ for the coupled terrain is similar to that observed for the isolated mountain. However, for the canyon side immediately adjacent to the mountain ($1 < x/d < 2$), the coupled terrain exhibits a more intricate seismic response compared to the two isolated scenarios. At horizontal incidence ($\alpha = 90°$, Figures 15(b), 15(d), and 15(f)), the mountain exerts a barrier effect on the canyon, reducing the surface displacement in the canyon portion. For example, at $\eta = 1.0$ and $\alpha = 90°$, the dimensionless peak displacement amplitude in the isolated canyon is 4.36, whereas it drops to 2.42 in the coupled terrain, presenting a reduction of approximately 44.5%.

Overall, the coupling effect is most pronounced on the canyon side adjacent to the mountain ($1 < x/d < 2$).

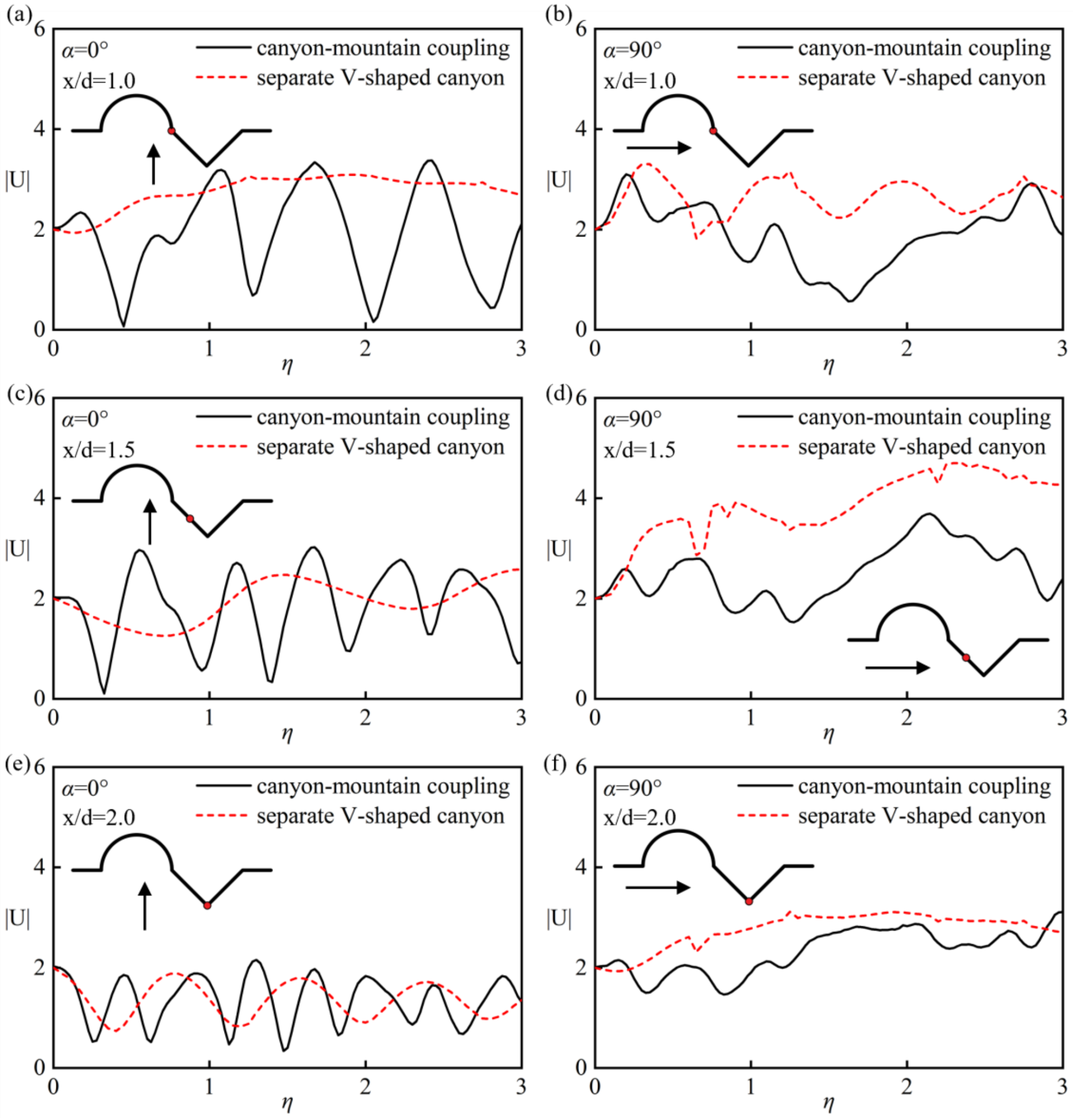


FIGURE 16 Displacement amplitude versus dimensionless frequency at three positions on the canyon side adjacent to the mountain.

Figures 16(a)-(f) show displacement amplitude versus η at three specific positions on the canyon side adjacent to the mountain (i.e., x/d=1.0, x/d=1.5 and x/d=2.0) for two incidence angles. At vertical incidence (α = 0°, Figures 16(a), 16(c), and 16(e)), the coupled terrain exhibits more pronounced oscillations, with additional peaks and higher amplitude values than those observed for the isolated V-shaped canyon. At horizontal incidence (α = 90°, Figures 16(b), 16(d), and 16(f)), the displacement amplitudes at the three positions on the coupled terrain are generally smaller than those in the isolated V-shaped canyon.

5 CONCLUSIONS

We introduced an efficient SH wave simulation method developed by combining fundamental principles of BEM and PINNs. The method was validated in complex topography problems in semi-infinite domains by comparing its results with analytical solutions. Our study demonstrated that the FF PINN -BEM approach effectively simulates SH wave scattering in complex terrain, combining the advantages of both techniques. The main features of the proposed method are presented as follows:

(1) The PINN method cannot directly solve the infinite or semi-infinite domain problems and must rely on an appropriate absorbing boundary. In contrast, the wavefield construction of the proposed method is straightforward: the virtual load outputs from the fully connected neural network can be directly derived from the boundary integral equation, enabling the calculation of displacement at any point within the semi-infinite space.

(2) By taking the incidence angle (α) and dimensionless frequency (η) as network inputs, the algorithm can output the virtual load continuously in these parameter dimensions. The virtual load predicted by the network to solve the displacements can avoid calculating the integration of the traction force Green's function and solving the linear equations, thereby significantly reducing computation time compared to the traditional BEM.

(3) The method uses a pre-trained model to initialize the network within a transfer learning framework. As the pre-trained model has already captured essential characteristics of ground motion, it enables faster convergence and significantly shortens the training duration.

Despite these advantages, several limitations of the present study should be acknowledged:

(1) The proposed framework is formulated for two-dimensional anti-plane SH-wave propagation in homogeneous media and is not directly applicable to wave equations with spatially varying material properties. Extending the method to variable-coefficient problems remains challenging and requires further investigation.

(2) The current formulation relies on a three-point Gaussian quadrature scheme for numerical integration, which imposes an inherent limitation on the achievable accuracy. More accurate integration strategies may further improve the precision of the proposed method.

(3) Although the present work focuses on forward seismic wavefield simulations using noise-free reference solutions, the robustness of the proposed framework in the presence of noisy observational data has not been explicitly examined. Incorporating noise-aware regularization or hybrid physics–data constraints is an important direction for future research, particularly for seismic inversion applications.

(4) Finally, while the concept of learning virtual boundary forces through physics-informed neural networks is not fundamentally restricted to two dimensions, extending the framework to three-dimensional wave propagation problems would require three-dimensional Green's functions and surface integral formulations, leading to increased computational and optimization challenges.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used ChatGPT4 in order to improve language. After using this tool/service, the authors reviewed and edited the content as needed and takes full responsibility for the content of the publication.

ACKNOWLEDGMENT

The authors are grateful for the research funding provided by the National Natural Science Foundation of China (Grant No. 42377140, 5227837), Key Research and Development Program of Sichuan Province

(Grant No. 2025YFHZ0300). The authors would like to thank the two anonymous reviewers for their constructive comments and suggestions, which significantly improved the quality of this paper.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY STATEMENT

The data and codes used in this paper will be publicly available on GitHub at https://github.com/yxwuhhu/FF_PINN-BEM, after the paper is published.

Appendix A. Comparison Between AdamW-only and Hybrid AdamW–L-BFGS Optimization

A.1 Experimental setup

To further investigate the influence of different optimization strategies on the training performance of the proposed FF PINN–BEM framework, two optimization schemes are considered and compared in this study.

The first scheme employs the AdamW optimizer alone for a total of 50,000 training epochs. The second scheme adopts a hybrid optimization strategy, in which the network is first trained using AdamW for 40,000 epochs to achieve stable initial convergence, followed by 10,000 optimization steps using the L-BFGS optimizer for fine-tuning.

For reproducibility, the configuration of the L-BFGS optimizer used in this study is reported below.

Training is performed in a full-batch manner, where each closure evaluates the complete physics-informed loss function.

The L-BFGS hyperparameters are:

- learning rate: lr = 0.1
- maximum inner iterations per step: max_iter = 50
- maximum function evaluations: default PyTorch setting (automatically determined from max_iter)
- history size: history_size = 50
- gradient tolerance: 1×10^{-8}
- parameter change tolerance: 1×10^{-12}
- line search: not used (default PyTorch behavior)

Except for the optimization strategy, all other training configurations—including network architecture, learning rate, collocation point distribution, and loss function formulation—are identical to those described in Section 3.

A.2 Convergence behavior comparison

Figures A1-3 illustrate the training loss histories obtained using AdamW-only and hybrid AdamW–L-BFGS optimization for the cases in Section3. During the first 40,000 epochs, the two strategies exhibit nearly identical convergence behavior, as the same optimizer and training settings are employed. After switching to L-BFGS, the optimization gradually approaches a near-stationary region of the loss landscape. In this regime, AdamW may exhibit oscillatory behavior. When L-BFGS is applied, the updates incorporate local curvature information, resulting in smaller and smoother parameter corrections.

Consequently, the loss variation in the L-BFGS phase may appear limited even though the optimization continues to progress toward convergence.

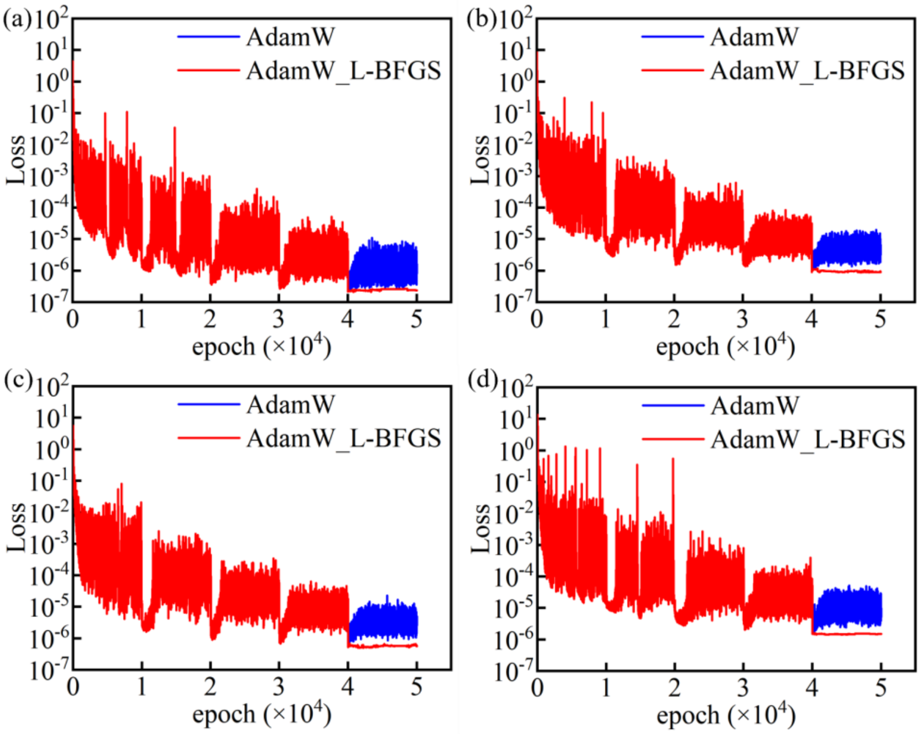


FIGURE A1 Training loss curves for the semi-cylindrical canyon under different incidence angles: (a) α=0°; (b) α=30°; (c) α=60°; (d) α=90°.

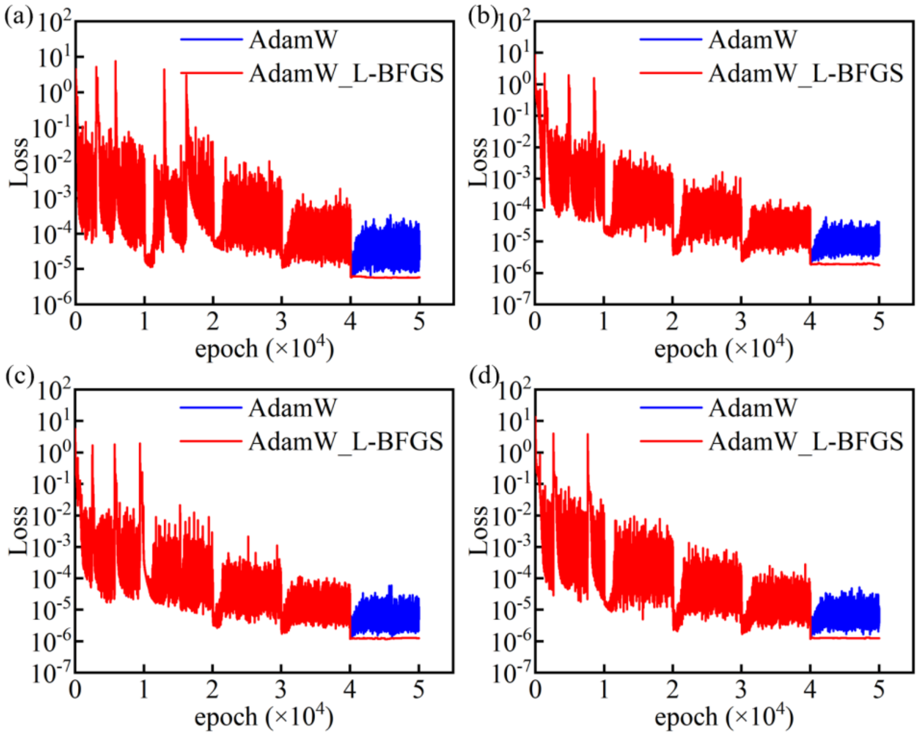


FIGURE A2 Training loss curves for the semi-cylindrical mountain under different incidence angles: (a) α=0°; (b) α=30°; (c) α=60°; (d) α=90°.

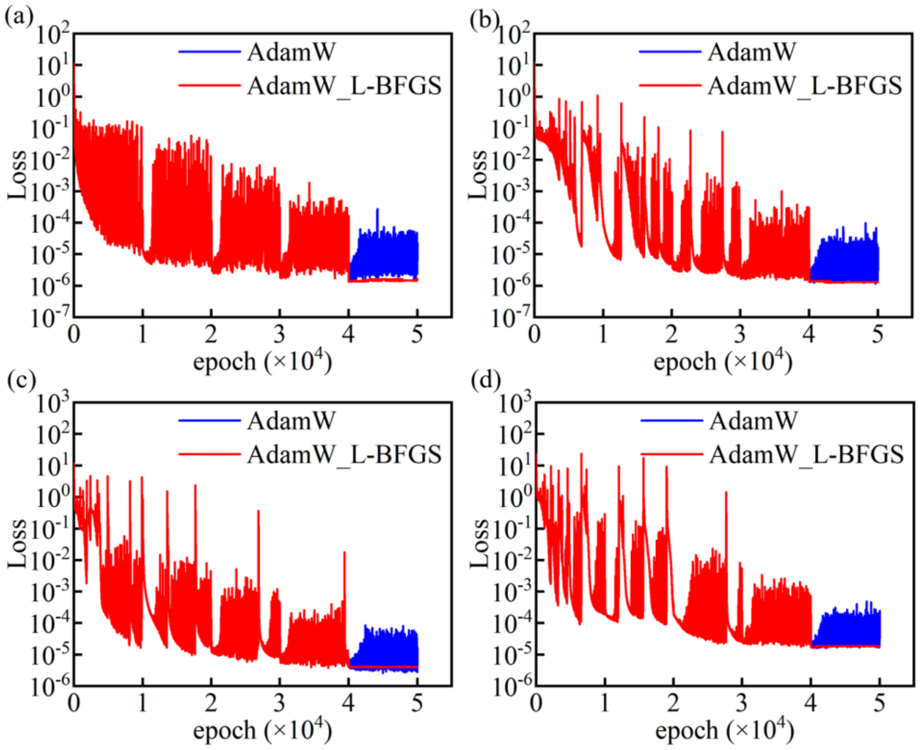


FIGURE A3 Training loss curves for the V-shaped canyon under different incidence angles: (a) α=0°; (b) α=30°; (c) α=60°; (d) α=90°.

Similar qualitative convergence patterns are observed across the differentnumerical cases presented in Section 3, including various incidence angles and topographic configurations, indicating that this numerical behavior is not specific to a particular case.

It is further observed that after a limited number of L-BFGS epochs, the loss variation becomes very small. This reflects the diminishing improvement characteristic of quasi-Newton refinement once the optimization has entered a near-stationary regime. Therefore, in the remainder of this study, AdamW is used to approach the near-optimal region, followed by a short L-BFGS refinement stage to complete the training, avoiding unnecessary additional epochs while maintaining stable final convergence.

Appendix B. Effect of L2 regularization on generalization performance

To further investigate the role of L2 regularization in improving the generalization performance of the proposed FF-PINN-BEM framework, a comparative study with different regularization strengths was conducted. In this section, the network was trained using a hybrid optimization strategy, consisting of 50,000 epochs of AdamW optimization followed by 100 epochs of the L-BFGS optimizer. The standard deviation of the Fourier feature mapping is fixed at $\sigma = 2$. The L2 regularization coefficient was varied among {1.0, 0.1, 0.01, 0.001, 0}, where the case with coefficient 0 corresponds to training without L2 regularization. All problem parameters, network architecture, and training settings are identical to those used in Section 4.1.1 for the case of dimensionless frequency $\eta = 6.0$, except for the L2 regularization coefficient.

To assess generalization performance, the network is trained using sparsely sampled incidence angles at 5° intervals over the range of 0°–45°. After training, the surface displacement responses are predicted at every 1° within the same range, excluding the angles used for training. These unseen angles constitute the test set. The average relative error over the test set is used as the evaluation metric.

TABLE B1 Average relative errors on the test set for different L2 regularization coefficients (η = 6.0, σ = 2)

L2 coefficient	0.0	0.001	0.01	0.1	1.0
Test average relative error ε	4.14%	4.24%	3.48%	1.89%	8.87%

Table B1 summarizes the test-set average relative errors for different values of the L2 regularization coefficient. The results indicate that an appropriate level of L2 regularization effectively suppresses overfitting and improves generalization accuracy, whereas excessively large or vanishing regularization may deteriorate prediction performance.

REFERENCES

1. Trifunac M D, Hudson D E. Analysis of the Pacoima dam accelerogram—San Fernando, California, earthquake of 1971. *Bull Seismol Soc Am*. 1971;61(5):1393-1411.
2. Assimaki D, Kausel E, Gazetas G. Soil-dependent topographic effects: a case study from the 1999 Athens earthquake. *Earthq Spectra*. 2005;21(4):929-966.
3. Hough S E, Altidor J R, Anglade D, et al. Localized damage caused by topographic amplification during the 2010 M 7.0 Haiti earthquake. *Nat Geosc*. 2010;3(11):778-782.
4. Trifunac M D. Scattering of plane SH waves by a semi-cylindrical canyon. *Earthq Eng Struct Dyn*. 1973;1(3):267–281.
5. Wong H L, Trifunac M D. Scattering of plane SH waves by a semi-elliptical canyon. *Earthq Eng Struct Dyn*. 1974;3(2):157-169.
6. Lee V W. Scattering of plane SH-waves by a semi-parabolic cylindrical canyon in an elastic half-space. *Geophys J Int*. 1990;100(1):79-86.
7. Todorovska M I, Lee V W. Surface motion of shallow circular alluvial valleys for incident plane SH waves-analytical solution. *Soil Dyn Earthq Eng*. 1991;10(4):192-200.
8. Yuan X, Men F L. Scattering of plane SH waves by a semi-cylindrical hill. *Earthq Eng Struct Dyn*. 1992;21(12):1091-1098.
9. Tsaur D H, Chang K H. An analytical approach for the scattering of SH waves by a symmetrical V-shaped canyon: shallow case. *Geophys J Int*. 2008;174(1):255-264.
10. Tsaur D H, Chang K H, Hsu M S. An analytical approach for the scattering of SH waves by a symmetrical V-shaped canyon: deep case. *Geophys J Int*. 2010;183(3):1501-1511.
11. Zhang N, Gao Y, Cai Y, et al. Scattering of SH waves induced by a non-symmetrical V-shaped canyon. *Geophys J Int*. 2012;191(1):243-256.
12. Gao Y, Zhang N, Li D, et al. Effects of topographic amplification induced by a U-shaped canyon on seismic waves. *Bull Seismol Soc Am*. 2012;102(4):1748-1763.
13. Boore D M. A note on the effect of simple topography on seismic SH waves. *Bull Seismol Soc Am*. 1972;62(1):275-284.
14. Zhang W, Zhang Z, Chen X. Three-dimensional elastic wave numerical modelling in the presence of surface topography by a collocated-grid finite-difference method on curvilinear grids. *Geophys J Int*. 2012;190(1):358-378.
15. Smith W D. The application of finite element analysis to body wave propagation problems. G*eophys. J. R. Astron. Soc.* 1975;42(2):747–768.
16. Jahromi S G, Karkhaneh S. The plurality effect of topographical irregularities on site seismic response. *Earthq Eng Eng Vib*. 2019;18:521-534.

17. Komatitsch D, Vilotte J P. The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures. *Bull Seismol Soc Am*. 1998;88(2):368-392.
18. Wang G, Du C, Huang D, et al. Parametric models for 3D topographic amplification of ground motions considering subsurface soils. *Soil Dyn Earthq Eng*. 2018;115:41-54.
19. Wu M, Ba Z, Liang J. A procedure for 3D simulation of seismic wave propagation considering source - path - site effects: Theory, verification and application. *Earthq Eng Struct Dyn*. 2022;51(12):2925-2955.
20. Wong H L. Effect of surface topography on the diffraction of P, SV, and Rayleigh waves. *Bull Seismol Soc Am*. 1982;72(4):1167-1183.
21. Liu Z, Zhou T, Meng S, et al. 2-D FM-IBEM Simulation of broadband ground motions on near-fault mountain-valley coupling site. *Eng Anal Bound Elem*. 2022;145:224-241.
22. Sánchez-Sesma F J, Ramos-Martínez J, Campillo M. An indirect boundary element method applied to simulate the seismic response of alluvial valleys for incident P, S and Rayleigh waves. *Earthq Eng Struct Dyn*. 1993;22(4):279-295.
23. Ba Z, Yin X. Wave scattering of complex local site in a layered half-space by using a multidomain IBEM: incident plane SH waves. *Geophys J Int*. 2016;205(3):1382-1405.
24. Komatitsch D, Tromp J. A perfectly matched layer absorbing boundary condition for the second-order seismic wave equation. *Geophys J Int*. 2003;154(1):146-153.
25. Cheng Q, Niu C, Zhao H, et al. A cumulative absolute velocity prediction method based on surrounding strong motion records and deep learning. *Eng Appl Artif Intel*. 2025; 157: 111416.
26. Lu X, Xu Y, Tian Y, et al. A deep learning approach to rapid regional post-event seismic damage assessment using time-frequency distributions of ground motions. *Earthq Eng Struct Dyn*. 2021;50(6): 1612-1627.
27. Chen S, Liu X, Fu L, et al. Physics symbolic learner for discovering ground-motion models via NGA-West2 database. *Earthq Eng Struct Dyn*. 2024;53(1): 138-151.
28. Fei Y, Liao W, Zhao P, et al. Hybrid surrogate model combining physics and data for seismic drift estimation of shear-wall structures. *Earthq Eng Struct Dyn*. 2024;53(10): 3093-3112.
29. Huang X, Alkhalifah T. Learned frequency-domain scattered wavefield solutions using neural operators. *Geophys J Int*. 2025;241(3): 1467-1478.
30. Karniadakis G E, Kevrekidis I G, Lu L, et al. Physics-informed machine learning. *Nat Rev Phys*. 2021;3(6):422-440.
31. Baydin A G, Pearlmutter B A, Radul A A, et al. Automatic differentiation in machine learning: a survey. *J Mach Learn Res*. 2018;18(153):1-43.
32. Sirignano J, Spiliopoulos K. DGM: A deep learning algorithm for solving partial differential equations. *J Comput Phys*. 2018;375:1339-1364.
33. Yu B. The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems. *Commun Math Stat*. 2018;6(1):1-12.
34. Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys*. 2019;378:686-707.
35. Nguyen-Thanh V M, Anitescu C, Alajlan N, et al. Parametric deep energy approach for elasticity accounting for strain gradient effects. *Comput Method Appl M*. 2021;386:114096.
36. Smith J D, Azizzadenesheli K, Ross Z E. Eikonet: Solving the eikonal equation with deep neural networks. *IEEE Trans. Geosci.Remote Sens*. 2020;59(12):10685-10696.

37. U. B. Waheed, E Haghighat, T. Alkhalifah, et al. PINNeik: Eikonal solution using physics-informed neural networks. *Comput Geosci*. 2021;155:104833.
38. Taufik M H, Waheed U, Alkhalifah T A. A neural network based global traveltime function (GlobeNN). *Sci Rep*. 2023;13(1):7179.
39. Piao S, Gu H, Wang A, et al. A domain-adaptive physics-informed neural network for inverse problems of Maxwell's equations in heterogeneous media. *IEEE Antennas and Wireless Propagation Letters*. vol. 23, no. 10, pp. 2905-2909,Oct. 2024.
40. Rasht-Behesht M, Huber C, Shukla K, et al. Physics-informed neural networks (PINNs) for wave propagation and full waveform inversions. *J Geophys Res-Sol Ea*. 2022;127(5):e2021JB023120.
41. Ren P, Rao C, Chen S, et al. SeismicNet: Physics-informed neural networks for seismic wave modeling in semi-infinite domain. *Comput Phys Commun*. 2024;295:109010.
42. Ding Y, Chen S, Li X, et al. Physics-constrained neural networks for half-space seismic wave modeling. *Comput Geosci*. 2023;181:105477.
43. Ding Y, Chen S, Li X, et al. Self-adaptive physics-driven deep learning for seismic wave modeling in complex topography. *Eng Appl Artif Intel*. 2023;123:106425.
44. Song C, Alkhalifah T, Waheed U B. A versatile framework to solve the Helmholtz equation using physics-informed neural networks. *Geophys J Int*. 2022;228(3):1750-1762.
45. Song C, Alkhalifah T, Waheed U B. Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks. *Geophys J Int*. 2021;225(2):846-859.
46. Song C, Alkhalifah T A. Wavefield reconstruction inversion via physics-informed neural networks. *IEEE T Geosci Remote*. 2021;60:1-12.
47. Alkhalifah, T., Song, C. & Huang, X., 2021b. High-dimensional wavefield solutions based on neural network functions. in *Proceedings of the First International Meeting for Applied Geoscience & Energy*, Society of Exploration Geophysicists, pp. 2440–2444.
48. Wang S, Yu X, Perdikaris P. When and why PINNs fail to train: A neural tangent kernel perspective. *J Comput Phys*. 2022, 449: 110768.
49. Wang S, Teng Y, Perdikaris P. Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM J Sci Comput*. 2021, 43(5): A3055-A3081.
50. Huang X, Alkhalifah T. Efficient physics-informed neural networks using hash encoding. Journal of Computational Physics, 2024, 501: 112760.
51. Tancik M, Srinivasan P, Mildenhall B, et al. Fourier features let networks learn high frequency functions in low dimensional domains. *Adv Neural Inf Process Syst*. 2020;33:7537-7547.
52. Song C, Wang Y. Simulating seismic multifrequency wavefields with the Fourier feature physics-informed neural network. *Geophys J Int*. 2023;232(3):1503-1514.
53. Sánchez-Sesma F J, Rosenblueth E. Ground motion at canyons of arbitrary shape under incident SH waves. *Earthq Eng Struct Dyn*. 1979;7(5):441-450.
54. Sánchez-Sesma F J, Campillo M. Diffraction of P, SV, and Rayleigh waves by topographic features: a boundary integral formulation. *Bull Seismol Soc Am*. 1991, 81(6): 2234-2253.
55. Hendrycks D. Gaussian Error Linear Units (Gelus). arXiv preprint arXiv:1606.08415, 2016.
56. Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 2010: 249-256.
57. Huang X, Alkhalifah T A. Microseismic source imaging using physics-informed neural networks with hard constraints. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-11.