Single Reference Frequency Loss for Multifrequency Wavefield Representation Using Physics-Informed Neural Networks

Xinquan Huang and Tariq Alkhalifah

***Abstract*— Physics-informed neural networks (PINNs) can offer approximate multidimensional functional solutions to the Helmholtz equation that is flexible, requires low memory, and has no limitations on the shape of the solution space. However, the neural network (NN) training can be costly, and the cost dramatically increases as we train for multifrequency wavefields by adding frequency as an additional input to the NN multidimensional function. In this case, the often large variation of the wavefield features (specifically wavelength) with frequency adds more complexity to the NN training. Thus, we propose a new loss function for the NN multidimensional input training that allows us to seamlessly include frequency as a dimension. We specifically utilize the linear relation between frequency and wavenumber (the wavefield space representation) to incorporate a reference frequency scaling to the loss function. As a result, the effective wavenumber of the wavefield solution as a function of frequency remains almost stationary, which reduces the learning burden on the NN function. We demonstrate the effectiveness of this modified loss function on a layered model.**

***Index Terms*— Multifrequency wavefield, partial differential equation, physics-informed neural network (NN) (PINN), single reference frequency loss.**

I. Introduction

FREQUENCY-DOMAIN wave equation modeling, based on the Helmholtz equation, is quite common and of great importance in modeling many physical phenomena, e.g., electromagnetic and seismic wave propagation. However, in inverting for the subsurface properties, we have to solve the Helmholtz equation for many frequencies to recover fine-scale details, such as those in ultrasound medical imaging, ground-penetrating radar, and seismic full waveform inversion. Consequently, an accurate and efficient multifrequency solution is extremely important in many scientific and industrial applications. However, when the size of the subsurface model is large and the frequency is high, the computational cost of classical methods, such as finite difference, finite element, and spectral methods, is high. Besides, the complexity of the wave equation in elastic or anisotropic media can considerably add to the computational burden. The recently developed physics-informed neural network (PINN) for solving the Helmholtz equation showed considerable potential in modeling because of its flexibility and low memory requirement, and no limitations are imposed on the shape of the solution space [1]. However, it is hard to train, admitting less than optimal solutions for practical size neural network (NN) models [2]. A major challenge is that, when we increase the dimension of the input (such as including frequency), the complexity of the wavefield increases, yielding poor convergence of PINN. Reducing the complexity of the PINN optimization problem is an important objective.

Here, we propose to use a reference frequency modified Helmholtz equation-based loss function to train an NN for multifrequency wavefield representation using PINN. The reference frequency allows us to effectively mitigate the change in the spatial wavenumber over frequency by adapting the spatial scale to frequency, thus reducing the complexity of the wavefield as if it was representing a single frequency. This is important, as the PINN convergence depends highly on the complexity of the wavefield. We apply our method using the frequency-domain scattered wave equation to predict multifrequency wavefields. Compared to the traditional PINN with multifrequency loss, our approach yields more accurate and efficient wavefield solutions.

II. Related Work

PINN plays a vital role in surrogate modeling with potential applications in many fields [3]–[7]. For wavefields, when the solution domain is large or the frequency is high, the complexity of the solution requires a large-size NN, which is hard to train [8], [9]. To address the limitations of PINN, e.g., convergence, especially its convergence in scenarios with a large solution domain and high frequency, Wang *et al.* [10] made use of the Fourier input feature to solve the optimization using trigonometric functions. Liu *et al.* [11] proposed multiscale deep NNs with a PINN loss to improve the convergence. Recently, domain decomposition has gained attention [2], [12], [13] in which we divide the problem into many subdomains. For inverse problems, we also need multifrequency solutions, and all these methods have not addressed this need. Alkhalifah *et al.* [14] demonstrated the multidimensional wavefield solutions potential of PINNs, but the accuracy over the range of frequencies was not good.

III. Single Reference Frequency Loss

A. Helmholtz Equation for Scattered Wavefield

In this section, we briefly revisit several key concepts of the frequency-domain scattered wavefield representation

Manuscript received March 3, 2022; revised April 22, 2022; accepted May 13, 2022. Date of publication May 23, 2022; date of current version June 2, 2022. This work was supported by the King Abdullah University of Science and Technology (KAUST). *(Corresponding author: Xinquan Huang.)*

The authors are with the Physical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955, Saudi Arabia (e-mail: xinquan.huang@kaust.edu.sa; tariq.alkhalifah@kaust.edu.sa).

Digital Object Identifier 10.1109/LGRS.2022.3176867

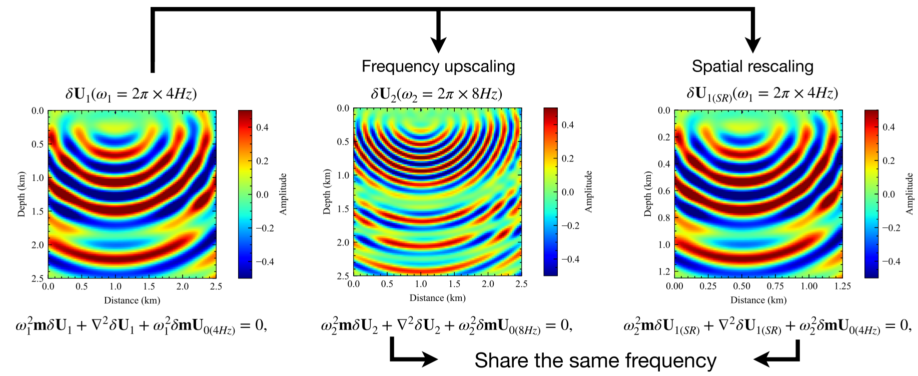

Fig. 1. Original 4-Hz wavefield based on a simple layered model (left), the wavefield after frequency upscaling (double the frequency), which is an 8-Hz wavefield (middle), and the wavefield after spatial rescaling (SR, reducing the spatial scale for each coordinate by half), which looks like the 4-Hz wavefield but only changes the scale (right). The Helmholtz equations based on wavefields of the middle and the right ones share the same frequency.

using PINN. In this letter, we focus on the forward modeling problem for the case of absorbing boundary conditions, and the velocity is known. To reduce the spatial samples needed for training the 2-D frequency-domain acoustic wave equation (Helmholtz equation) and mitigate the source singularity of the frequency-domain wavefield, we use the scattered wavefield instead, given by

$$\omega^2\mathbf{m}\delta\mathbf{U} + \nabla^2\delta\mathbf{U} + \omega^2\delta\mathbf{m}\mathbf{U}_0 = 0 \tag{1}$$

where $\mathbf{m}$ is the squared slowness, ω is the angular frequency, $\mathbf{U}$ is the frequency-domain wavefield as a function of (x, z) due to the source term $\mathbf{s} = (s_x, s_z)$, ∇ is the gradient operator, $\mathbf{U}_0$ is the background wavefield, $\delta\mathbf{U} = \mathbf{U} - \mathbf{U}_0$ [15] is the scattered wavefield, and $\delta\mathbf{m} = \mathbf{m} - \mathbf{m}_0$ is the squared slowness perturbation. Considering the background velocity to be constant, $\mathbf{U}_0$ can be directly calculated using the constant background squared slowness $\mathbf{m}_0$ with an analytical relation [16]

$$\mathbf{U}_0(x, z) = \frac{i}{4}H_0^{(2)}\left(\omega\sqrt{m_0\{(x - s_x)^2 + (z - s_z)^2\}}\right) \tag{2}$$

where $H_0^{(2)}$ is the zero-order Hankel function of the second kind. With this background wavefield, we consider the implicit setting that seismic waves propagate in infinite space. We, also, consider sources located at the same depth as it is often practiced in surface seismic exploration settings [17]. To find an NN representation $\Phi(\theta, \mathbf{x})$ satisfying the physical constraint (PINN), where θ represents the NN parameters and the NN's input $\mathbf{x} = (x, z, s_x, \omega)$, we use the physical multifrequency loss function, defined as

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}\left|\omega^2\mathbf{m}^i\Phi(\theta, \mathbf{x}^i) + \nabla^2\Phi(\theta, \mathbf{x}^i) + \omega^2\delta\mathbf{m}^i U_0^i\right|_2^2 \tag{3}$$

where N is the number of training samples.

B. Relationship Between Frequency Upscaling and Spatial Rescaling

From (1), we observe that, when we double the frequency, the first and third terms will quadruple. To keep (1) stationary, we need the second term (the Laplacian operator acting on $\delta\mathbf{U}$) to quadruple as well. To do so, we rescale the spatial axes to maintain the effective wavenumber. Thus, when we double the frequency, we rescale the spatial coordinates by half. For simplicity, we use three frequency-domain wavefields (see Fig. 1) to demonstrate the process. We immediately arrive at the conclusion that the wavefield by the frequency upscaling (8-Hz wavefield) and the wavefield by spatial rescaling (4-Hz wavefield after spatial rescaling) share similar wavenumber content satisfying the Helmholtz equation, which means that we just need one frequency here to describe two wavefields with different frequencies.

C. Dynamic Frequency Weighting Derivation

The relationship between frequency upscaling and spatial rescaling can be used to maintain a stationary wavenumber of the wavefield over the frequency range. However, the way to achieve this by literally rescaling the input spatial coordinates as a function of frequency will admit a nonuniform space dimension range for the various frequencies, in which the higher frequency component space dimension will be larger than the low-frequency one (see Fig. 2). This dimension will nonuniformly add complexity to the implementation of PINN.

To implement this stationary wavenumber concept, we introduce a frequency weighting approach, which upweights the derivation of low-frequency wavefields. The frequency weighting is dynamically determined by the ratio of current frequency of the training sample to the reference frequency, and as a result, the gradient of the scattered wavefield is given by

$$\text{gradient}(\delta\mathbf{U}) = \frac{\partial\delta\mathbf{U}}{\partial(\alpha\mathbf{x})} \tag{4}$$

where α is the scaling factor, equal to the ratio of the current frequency to a reference frequency. Inserting (4) into (1), we have the new loss function

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}\left|\omega_{\text{ref}}^2\mathbf{m}^i\Phi(\theta, \mathbf{x}^i) + \frac{\partial^2\Phi(\theta, \mathbf{x}^i)}{\partial^2(\alpha\mathbf{x}^i)} + \omega_{\text{ref}}^2\delta\mathbf{m}^i U_0^i\right|_2^2 \tag{5}$$

where ω_{ref} is the reference angular frequency and N is the number of training samples. In practice, we implement (5) by utilizing a computational graph (as shown in Algorithm 1).

Algorithm 1: Training With the Single Reference Frequency Loss Function

Draw N points $\{\mathbf{x}^i\}_{i=1}^N$ sampled from the 4-D model region, and no boundary points.
Initiate: NN parameters θ
for *each epoch* **do**
  Compute the $\mathbf{x}_{ref} = (x, z, s_x) \times \frac{\omega}{\omega_{ref}}$
  Then the input of the NN: $\mathbf{x} = (\mathbf{x}_{ref} \times \frac{\omega_{ref}}{\omega}, \omega)$
  normalize $\mathbf{x}$ and feed them into the network to get $\Phi(\theta, \mathbf{x})$
  automatic differentiation: $\frac{\partial^2 \Phi(\theta, \mathbf{x})}{\partial^2 \mathbf{x}_{ref}}$
  calculate the loss function of equation 5
  Update: NN parameters θ

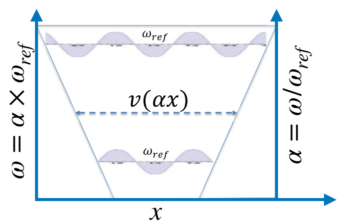

Fig. 2. Schematic plot describing how the reference frequency is utilized to make the wavefield nearly stationary over frequency (similar wavenumbers). Thus, the velocity model is stretched in the transformed space domain so that the wavefield will be accurate in the none transformed domain. As frequency increases, α increases.

Fig. 2 demonstrates the scaling of the velocity model along the x-axis. A similar scaling is applied to the z-axis. Thus, α provides the proper velocity $v(\alpha x, \alpha z)$ transformation, and we can use (5) directly with setting ω_{ref} for all frequencies. This is easy to implement using NN functions as the training sample points are often chosen randomly. Thus, we only need to train using random values from within the inverted trapezoid region in Fig. 2. Thus, depending on the frequency (or α), the range of samples involved in the training varies. Note that, for a uniform distribution of samples, low frequencies will get far fewer samples in the training than high frequencies, which is natural for a proper representation of the complexity of the wavefield.

IV. Experiments

In this section, to show the effectiveness of our proposed loss function for the scattered wave equation, we test the performance of conventional PINN using the multifrequency loss function and PINN utilizing our single reference frequency loss function on a simple layered model. Then, we compare the accuracy of the predicted wavefields by computing the corresponding velocities obtained via inserting the predicted wavefields into (1).

Based on a simple layered model (see Fig. 3) extracted from the Marmousi model covering an area of 2.5×2.5 km^2, we generate 1 280 000 random samples for our 4-D wavefield space, with the frequency, ranging from 3.0 to 8.0 Hz, along with δm for squared slowness perturbation and m_0 for background squared slowness at these points. The depth of

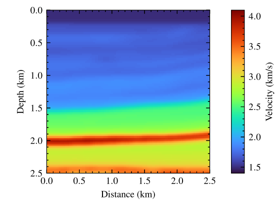

Fig. 3. Simple layered velocity model extracted from the Marmousi model.

sources s_z is set to 0.025 km. The background wavefield is calculated analytically for a background velocity of 1.5 km/s. The reference frequency here is 8.0 Hz. The basic network architecture for both methods is a multilayer perception with three hidden layers, as well as positional encoding [7]. The inputs include $\{x, z, s_x, \omega\}$, and the hidden layers are of the size $\{512, 512, 512\}$ from shallow to deep.

We use an Adam optimizer to train our networks. During training, we set the batch size to 40 000. The initial learning rate is chosen to be $1e{-}3$, and it is gradually decreased to $5e{-}5$. We have trained (optimized) our model with these settings for 15 000 epochs. The training costs of our proposed method and vanilla PINN are almost the same. To evaluate the results, we numerically solve the Helmholtz equation for specific frequencies and source locations, and use the solution as a reference; these solutions are provided in the 2.5×2.5 km^2 area using 100 samples in both the x- and z-directions. For Figs. 4 and 5, the source is located at a depth of 0.025 km and the lateral position of 1.0 km. Figs. 4 and 5 show the real and imaginary parts of the predicted wavefield for various frequencies. It is obvious that, with one more input dimension, the representation of NNs for the wavefield becomes harder to obtain via conventional training. On the other hand, our proposed loss function provided reasonable results considering the larger (four) dimensional space, and we obtain a much more accurate amplitude and phase representation of the wavefield. We also calculated the velocity models from the PINN predictions using (1) [18], as shown in Fig. 6. We can observe that the PINN with our proposed loss function reconstructs the details of the velocity model much better than the vanilla method.

Transforming the multifrequency wavefield into time-domain records could help us understand better what parts of the wavefield we managed to predict accurately in time. Using inverse Fourier transform, we obtain time-domain snapshots for the numerical implementation, the conventional PINN, and our proposed PINN of the multifrequency wavefield representation. Fig. 7 shows the time-domain snapshot at 1.0 s transformed from 3- to 8-Hz wavefields with a frequency interval of 0.2 Hz. We observe that the time-snap shot obtained by the model trained with a single reference frequency loss function shows more agreement with the numerical result compared to the vanilla multifrequency loss function, specifically the key reflection corresponding to the high-velocity perturbation at 2.0 km.

V. Discussion

From the above experiments and analysis, we found that, though PINN provides a general framework to represent

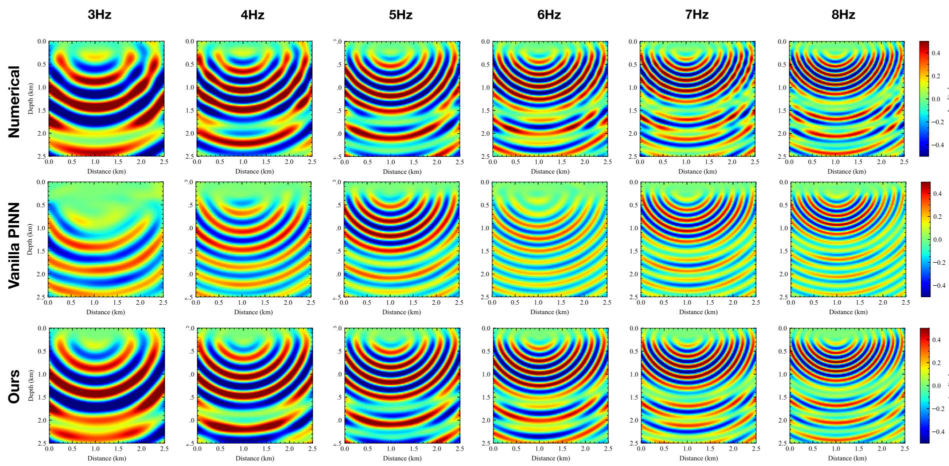

Fig. 4. Real parts of predicted multifrequency wavefields using numerical solutions (on the top, considered ground truth), vanilla PINN (in the middle), and the PINN with our method (on the bottom).

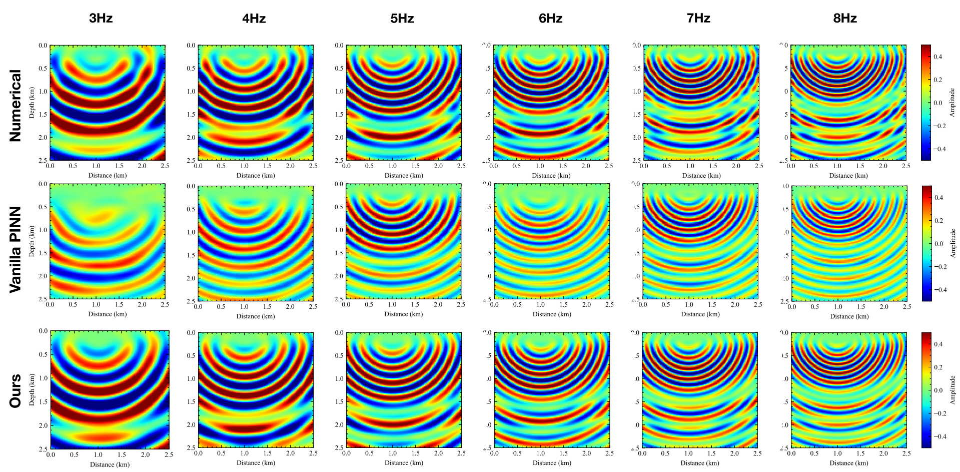

Fig. 5. Imaginary parts of predicted multifrequency wavefields using numerical solutions (on the top, considered ground truth), vanilla PINN (in the middle), and the PINN with our method (on the bottom).

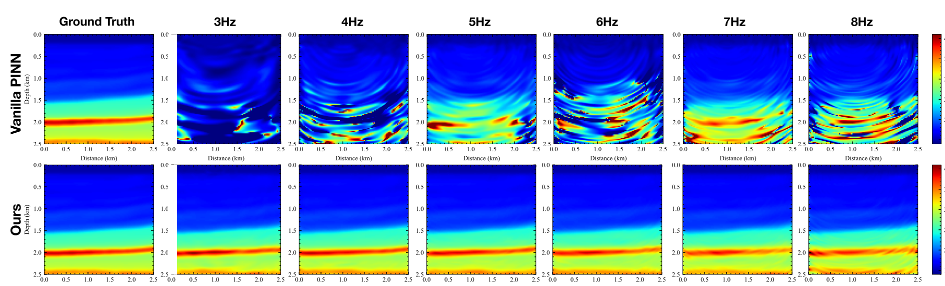

Fig. 6. Estimated velocities from the multifrequency wavefields.

functional solutions of PDE, these solutions tend to be smooth and prone to errors. This weakness often limits the potential of PINN for practical applications. Combining our understanding of the physical property underlying the PDEs, the proper architecture design, and the proper training of PINN, the accuracy of the solution could be improved at a reasonable cost. Here, we demonstrate how our single reference frequency loss function, taking into account the behavior of the wavefield with respect to frequency, can reduce the complexity of our solution space, which is important for high-dimensional wavefield representation by PINN. The benefits are achieved by leveraging an adaptive spatial scale that depends on the frequency and, thus, mitigates the change in the spatial wavenumber over frequency. As for the formulation of the loss function, it results in a natural weighting of the loss.

As for the selection of the reference frequency, in theory, it can be chosen as any value, and it is better to be in the frequency range. In our implementation, we found that using the upper bound of the frequency range is better.

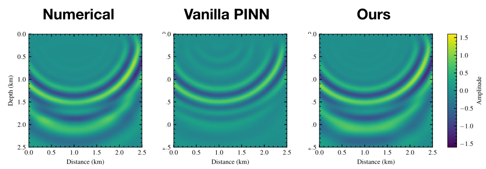

Fig. 7. Comparison of the time-domain wavefields at 1.0 s transformed from the wavefields ranging from 3 to 12 Hz with an interval of 0.2 Hz by numerical method (as a reference), vanilla PINN, and our proposed method.

As for the computational efficiency, although the total training time of the workflow may cost 5 h total (trained using a Quadro RTX 8000 GPU with 48 GB of memory), it can predict the wavefield value within the frequency range at an arbitrary space point (x, z) for any source location on the surface instantly, such as an extended Green's function. For the prediction of a monofrequency wavefield with 91 809 samples (101 $\times$ 101 resolution and nine sources), it only takes 3 s. In future work, we plan to study the positional encoding methods and the design of the PINN to make the training faster.

VI. Conclusion

We proposed using a reference frequency-based loss function to train the NN for multifrequency wavefield representation, and we demonstrated that this approach admitted superior performance and wavefield accuracy compared to the vanilla PINN. The reference frequency loss function implicitly embeds the relationship between frequency upscaling and spatial rescaling into the network, which makes the network easier to train and improves its representation. The method has the potential to solve the multifrequency Helmholtz equation even in a large and complex model. The method can be generalized to other physical problems, which may include a multifrequency or multiscale component.

Acknowledgment

The authors would like to thank KAUST for the support and the SWAG Group for the collaborative environment. They would also like to thank the resources of the Supercomputing Laboratory at KAUST.

References

[1] C. Song and T. A. Alkhalifah, "Wavefield reconstruction inversion via physics-informed neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–12, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9585726/

[2] B. Moseley, A. Markham, and T. Nissen-Meyer, "Finite basis physics-informed neural networks (FBPINNs): A scalable domain decomposition approach for solving differential equations," 2021, *arXiv:2107.07871*.

[3] M. Raissi, A. Yazdani, and G. E. Karniadakis, "Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations," *Science*, vol. 367, no. 6481, pp. 1026–1030, Feb. 2020, doi: 10.1126/science.aaw4741.

[4] L. Sun, H. Gao, S. Pan, and J.-X. Wang, "Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data," *Comput. Methods Appl. Mech. Eng.*, vol. 361, Apr. 2020, Art. no. 112732. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S004578251930622X

[5] F. S. Costabal, Y. Yang, P. Perdikaris, D. E. Hurtado, and E. Kuhl, "Physics-informed neural networks for cardiac activation mapping," *Frontiers Phys.*, vol. 8, p. 42, Feb. 2020. [Online]. Available: https://www.frontiersin.org/article/10.3389/fphy.2020.00042/full

[6] S. Cai, Z. Wang, F. Fuest, Y. J. Jeon, C. Gray, and G. E. Karniadakis, "Flow over an espresso cup: Inferring 3-D velocity and pressure fields from tomographic background oriented Schlieren via physics-informed neural networks," *J. Fluid Mech.*, vol. 915, p. A102, May 2021.

[7] X. Huang, T. Alkhalifah, and C. Song, "A modified physics-informed neural network with positional encoding," in *Proc. 1st Int. Meeting Appl. Geosci. Energy*. Houston, TX, USA: Soc. Explor. Geophysicists, 2021, pp. 2480–2484.

[8] N. Rahaman *et al.*, "On the spectral bias of neural networks," 2018, *arXiv:1806.08734*.

[9] Z.-Q. J. Xu, Y. Zhang, T. Luo, Y. Xiao, and Z. Ma, "Frequency principle: Fourier analysis sheds light on deep neural networks," 2019, *arXiv:1901.06523*.

[10] S. Wang, H. Wang, and P. Perdikaris, "On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale PDEs with physics-informed neural networks," 2020, *arXiv:2012.10047*.

[11] Z. Liu, "Multi-scale deep neural network (MscaleDNN) for solving Poisson-Boltzmann equation in complex domains," *Commun. Comput. Phys.*, vol. 28, no. 5, pp. 1970–2001, Jun. 2020.

[12] A. D. Jagtap and G. E. Karniadakis, "Extended physics-informed neural networks (XPINNs): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations," *Commun. Comput. Phys.*, vol. 28, no. 5, pp. 2002–2041, Jun. 2020.

[13] A. Heinlein, A. Klawonn, M. Lanser, and J. Weber, "Combining machine learning and domain decomposition methods for the solution of partial differential equations—A review," *GAMM-Mitteilungen*, vol. 44, no. 1, Mar. 2021, Art. no. e202100001. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/gamm.202100001

[14] T. Alkhalifah, C. Song, and X. Huang, "High-dimensional wavefield solutions based on neural network functions," in *Proc. 1st Int. Meeting Appl. Geosci. Energy*. Houston, TX, USA: Soc. Explor. Geophysicists, 2021, pp. 2440–2444.

[15] T. Alkhalifah, C. Song, and U. B. Waheed, "Machine learned Green's functions that approximately satisfy the wave equation," in *Proc. SEG Int. Expo. Annu. Meeting*. Houston, TX, USA: Soc. Explor. Geophysicists, 2020, pp. 2638–2642, doi: 10.1190/segam2020-3421468.1.

[16] P. G. Richards and K. Aki, *Quantitative Seismology: Theory and Methods*, vol. 859. New York, NY, USA: Freeman, 1980.

[17] O. Yilmaz, *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data*. Houston, TX, USA: Soc. Explor. Geophysicists, 2001.

[18] C. Song, T. Alkhalifah, and U. B. Waheed, "Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks," *Geophys. J. Int.*, vol. 225, no. 2, pp. 846–859, Feb. 2021. [Online]. Available: https://academic.oup.com/gji/article/225/2/846/6081098