Single Reference Frequency Loss for Multi-frequency Wavefield Representation using Physics-Informed Neural Networks

Xinquan Huang
KAUST
xinquan.huang@kaust.edu.sa

Tariq Alkhalifah
KAUST
tariq.alkhalifah@kaust.edu.sa

Abstract

Physics-informed neural networks (PINNs) can offer approximate multidimensional functional solutions to the Helmholtz equation that are flexible, require low memory, and have 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 multi-frequency wavefields by adding frequency to the NN multidimensional function, as the variation of the wavefield 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 stationary reducing the learning burden on the NN function. We demonstrate the effectiveness of this modified loss function on a layered model.

1 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, like in ultrasound medical imaging, ground penetrating radar, and seismic full waveform inversion. Consequently, an accurate and efficient multi-frequency 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 wave equation in elastic or anisotropic media can considerably add to the computational burden. The recently developed physics-informed neural network (PINN) for solving for the Helmholtz equation showed considerable potential in modeling because of its flexibility, low memory requirement, and no limitations are imposed on the shape of the solution space. However, it is hard to train, admitting less than optimal solutions for practical size neural network models. A major challenge is that when we increase the dimension of the input (like 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 based loss function to train a neural network (NN) for multi-frequency wavefield representation using PINN. The reference frequency allows us to effectively mitigate the change in 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

NeurIPS 2021 AI for Science Workshop.

apply our method using the frequency-domain scattered wave equation to predict multi-frequency wavefields. Compared to the traditional PINN with multi-frequency loss, our approach yields more accurate and efficient wavefield solutions.

2 Related Work

Physics-informed neural networks play a vital role in surrogate modeling with potential applications in many fields [9, 13, 11, 3, 5]. For wavefields, when the solution domain is large or the frequency is high, the complexity of the solution requires a large-size neural network, which is hard to train. 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.* [14] made use of Fourier input feature to solve the optimization using trigonometric functions. Liu *et al.* [7] proposed multi-scale deep neural networks with a PINN loss to improve the convergence. Recently, domain decomposition has gained attention [6, 4, 8] in which we divide the problem into many sub-domains. For inverse problems, we also need multi-frequency solutions and all these methods have not addressed this need. Alkhalifah *et al.* [1] demonstrated the multi-dimensional wavefield solutions potential of PINNs, but the accuracy over the range of frequencies was not good.

3 Single Reference Frequency Loss

3.1 Helmholtz equation for scattered wavefield

In this section, we briefly revisit several key concepts of the frequency-domain scattered wavefield representation using PINN. 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 $\mathbf{s} = (s_x, s_z)$ and ∇ is the gradient operator, $\mathbf{U}_0$ is the background wavefield, $\delta \mathbf{U} = \mathbf{U} - \mathbf{U}_0$ [2] 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 [10]:

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

where $H_0^{(2)}$ is the zero-order Hankel function of the second kind. To find a neural network representation $\Phi(\theta, \mathbf{x})$ satisfying the physical constraint (PINN), where θ represents the model parameters and $\mathbf{x} = (x, z, s_x, \omega)$, we use the physical multi-frequency loss function defined as:

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

3.2 Relationship between frequency upscaling and spatial rescaling

From equation 1, we observe that when we double the frequency, the first and the third term will quadruple. To keep equation 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. So when we double the frequency, we rescale the spatial coordinates by a half. For simplicity we use three frequency-domain wavefields (Figure 1) to demonstrate the process. We immediately arrive to the conclusion that the wavefield by the frequency upscaling (8Hz wavefield) and the wavefield by spatial rescaling (4Hz wavefield after spatial rescaling) share similar wavenumber content satisfying the Helmholtz equation, which means we just need one frequency here to describe two wavefields with different frequencies.

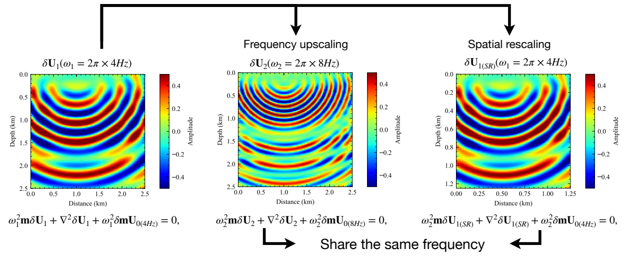

Figure 1: The original 4Hz wavefield based on a simple layered model (left), the wavefield after frequency upscaling (double the frequency), which is 8Hz wavefield (middle), and the wavefield after spatial rescaling (reducing the spatial scale for each coordinate by half), which looks like the 4Hz wavefield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

3.3 Dynamic frequency weighting derivation

The relationship between a frequency upscaling and spatial rescaling can be established to maintain a stationary wavenumber of the wavefield over the frequency range. However, we do so by literally rescaling the spatial coordinates as a function of frequency will admit nonuniform space dimension range for the various frequencies, in which the higher frequency component space dimension will be larger than the low frequency one. This dimension will non uniformity add complexity in the implementation of PINN.

To implement this stationary wavenumber concept, we introduce a frequency weighting approach which up-weights 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:

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

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

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

where ω_{ref} is the reference angular frequency. In practice, we implement equation 5 by utilizing a computational graph (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: model 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: model parameters θ

4 Experiments

In this section, we show the effectiveness of our proposed loss function for the scattered wave equation, and compare the accuracy of wavefield and the corresponding velocity. We compare the performances of conventional PINN using the multi-frequency loss function and PINN utilizing our single reference frequency loss function.

Based on a simple layered model covering an area of $2.5 \times 2.5 km^2$, we generate 1280000 random samples from our four dimensional wavefield, with f, the frequency, ranging from 3.0 Hz to 8.0 Hz, along with δm for squared slowness perturbation, m_0 for background squared slowness at these points. The depth of sources s_z is set to 0.025 km. The background wavefield is calculated analytically for a background velocity of 1.5km/s. The reference frequency here is 8.0 Hz. The basic network architecture for both methods is a Multi-Layer Perception with three hidden layers, as well as positional encoding [5]. 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 40000. The initial learning rate is chosen to be $1e{-}3$ and it is gradually decreased to $5e{-}5$. We have trained our model with these settings for 15000 epochs. To evaluate the results, we solve the Helmholtz equation numerically for specific frequencies and source locations and use the solution as reference, and these solutions are provided in the $2.5 \times 2.5 km^2$ area using 100 samples in both the x and z directions. For Figure 2, the source is located at a depth of 0.025 km and lateral distance of 1.0 km. Figure 2 shows

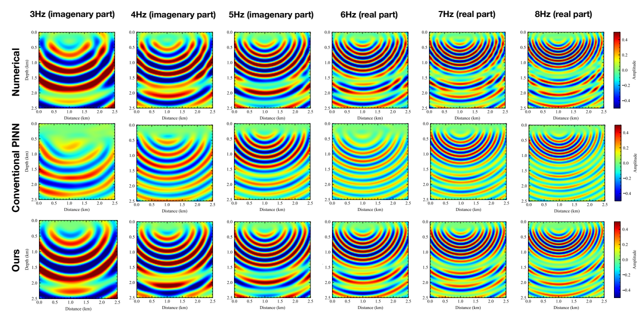

Figure 2: Predicted multi-frequency wavefields using numerical solutions (on the top, considered ground truth), conventional PINN (in the middle), and the PINN with our method (on the bottom).

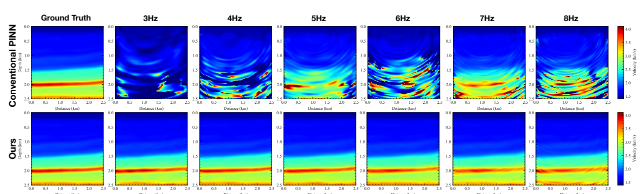

Figure 3: The estimated velocities from the multi-frequency wavefields.

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 the PINN is expected to predict. We also calculated the velocity models from the PINN predictions using equation 1[12], as shown in Figure 3. We can observe that the PINN with our proposed loss function reconstructs the details of the velocity model much better than the conventional method.

5 Conclusion

We proposed using a reference frequency based loss function to train the NN for multi-frequency wavefield representation and we demonstrated that this approach admitted superior performance and wavefield accuracy. 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 multi-frequency Helmholtz equation even in a large and complex model. The method can be generalized to other physical problems, which may include a multi-frequency or multi-scale component.

References

[1] Tariq Alkhalifah, Chao Song, and Xinquan Huang. "High-dimensional wavefield solutions based on neural network functions". In: *SEG Technical Program Expanded Abstracts*. Society of Exploration Geophysicists, 2021, pp. 2440–2444. DOI: `https://doi.org/10.1190/segam2021-3584030.1`.

[2] Tariq Alkhalifah, Chao Song, and Umair bin Waheed. "Machine learned Green's functions that approximately satisfy the wave equation". In: *SEG Technical Program Expanded Abstracts 2020*. Society of Exploration Geophysicists, 2020, pp. 2638–2642. ISBN: 2020342146. DOI: `10.1190/segam2020-3421468.1`. URL: `https://library.seg.org/doi/10.1190/segam2020-3421468.1`.

[3] Shengze Cai et al. "Flow over an espresso cup: inferring 3-D velocity and pressure fields from tomographic background oriented Schlieren via physics-informed neural networks". In: *Journal of Fluid Mechanics* 915 (2021), A102. ISSN: 0022-1120. DOI: `10.1017/jfm.2021.135`. URL: `https://www.cambridge.org/core/product/identifier/S002211202100135X/type/journal_article`.

[4] Alexander Heinlein et al. "Combining machine learning and domain decomposition methods for the solution of partial differential equations—A review". In: *GAMM-Mitteilungen* 44.1 (2021). ISSN: 0936-7195. DOI: `10.1002/gamm.202100001`. URL: `https://onlinelibrary.wiley.com/doi/10.1002/gamm.202100001`.

[5] Xinquan Huang, Tariq Alkhalifah, and Chao Song. "A modified physics-informed neural network with positional encoding". In: *SEG Technical Program Expanded Abstracts*. Society of Exploration Geophysicists, 2021, pp. 2480–2484. DOI: `https://doi.org/10.1190/segam2021-3584127.1`.

[6] Ameya D. Jagtap and George Em Karniadakis. "Extended Physics-Informed Neural Networks (XPINNs): A Generalized Space-Time Domain Decomposition Based Deep Learning Framework for Nonlinear Partial Differential Equations". In: *Communications in Computational Physics* 28.5 (2020), pp. 2002–2041. ISSN: 1815-2406. DOI: `10.4208/cicp.OA-2020-0164`. URL: `http://global-sci.org/intro/article_detail/cicp/18403.html`.

[7] Ziqi Liu, Wei Cai, and Zhi-Qin John Xu. "Multi-Scale Deep Neural Network (MscaleDNN) for Solving Poisson-Boltzmann Equation in Complex Domains". In: *Communications in Computational Physics* 28.5 (2020), pp. 1970–2001. ISSN: 1815-2406. DOI: `10.4208/cicp.OA-2020-0179`.

[8] Ben Moseley, Andrew Markham, and Tarje Nissen-Meyer. "Finite Basis Physics-Informed Neural Networks (FBPINNs): a scalable domain decomposition approach for solving differential equations". In: (2021). arXiv: `2107.07871`. URL: `http://arxiv.org/abs/2107.07871`.

[9] Maziar Raissi, Alireza Yazdani, and George Em Karniadakis. "Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations". In: *Science* 367.6481 (2020), pp. 1026–1030. ISSN: 0036-8075. DOI: `10.1126/science.aaw4741`. URL: `https://www.sciencemag.org/lookup/doi/10.1126/science.aaw4741`.

[10] P G Richards and K. Aki. *Quantitative seismology: theory and methods*. Vol. 859. Freeman New York, 1980. ISBN: 9781891389634.

[11] Francisco Sahli Costabal et al. "Physics-Informed Neural Networks for Cardiac Activation Mapping". In: *Frontiers in Physics* 8 (2020). ISSN: 2296-424X. DOI: `10.3389/fphy.2020.00042`. URL: `https://www.frontiersin.org/article/10.3389/fphy.2020.00042/full`.

[12] Chao Song, Tariq Alkhalifah, and Umair Bin Waheed. "Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks". In: *Geophysical Journal International* 225.2 (2021), pp. 846–859. ISSN: 0956-540X. DOI: `10.1093/gji/ggab010`. URL: `https://academic.oup.com/gji/article/225/2/846/6081098`.

[13] Luning Sun et al. "Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data". In: *Computer Methods in Applied Mechanics and Engineering* 361 (2020), p. 112732. ISSN: 00457825. DOI: `10.1016/j.cma.2019.112732`. URL: `https://linkinghub.elsevier.com/retrieve/pii/S004578251930622X`.

[14] Sifan Wang, Hanwen Wang, and Paris Perdikaris. "On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale PDEs with physics-informed neural networks". In: (2020). DOI: 10.1016/j.cma.2021.113938. arXiv: 2012.10047. URL: http://arxiv.org/abs/2012.10047%20http://dx.doi.org/10.1016/j.cma.2021.113938.