DFSNet FOR EIGENMODE SIMULATION OF RF CAVITY

A. H. Shali*, M. Fukuda, T. Yorita, H. Kanda, T. H. Chong, H. Zhao, S. Matsui,
K. Watanabe, T. Imura, S. Ishihata, N. Itakura
Research Center for Nuclear Physics Osaka University, Osaka, Japan

Abstract

Optimization of the quantity of interest of an RF cavity, such as the frequency and the quality factor, requires calculation from many samples with different geometric configuration with each of them requiring finite element method simulation. Geometric modification of smaller components requires finer mesh thus a very high computational cost requirement. In this research, we propose a physics informed neural network model to solve eigenvalue problems for the lowest frequency mode of a fully three dimensional RF cavity. Five loss functions are used, related to the differential equation (L_{in}), boundary condition (L_{bc}), non-trivial solution (L_{norm}), divergence free constraint (L_{div}) and mode forcing constraint (L_{m}). The input for the neural networks are the sampled spatial position, while the output is the value of the RF magnetic and electric field at the corresponding spatial position. The sampled positions are chosen to be related to the nodes of a generated unstructured mesh of the cavity model. The result shows that the neural network model could give an accurate prediction of the frequency and the field for the lowest mode.

INTRODUCTION

In recent years, there has been significant progress in neural networks across various domains, including medical imaging, language processing, and image recognition [1]. A particularly interesting application in science and engineering is the use of neural networks to solve partial differential equations (PDEs), known as physics-informed neural networks (PINNs) [2]. These networks incorporate extra loss functions that are related to the physical equations, in addition to the usual data-based loss functions. Additionally, it is possible to train these networks without any data, a method sometimes referred to as data-free surrogate networks (DFSNet) [3].

PINNs have some properties that is not shared with the more traditional numerical techniques used to solve partial differential equation. PINNs are not plagued with curse of dimensionality, there is no error accumulation [4], transfer learning possibility [5], and inverse design problem can be treated systematically [6]. Not to mention the parallelizability of neural network, which means that huge-sized problem can be solved concurrently with a large, parallel setting.

In this paper, a DFSNet scheme for electromagnetic RF eigenmode is proposed. The model is a simple fully connected network with loss functions relevant to the electromagnetic equations inside a perfectly conducting resonant cavity. The model is quite similar with network proposed to solve Schrödinger equation for one dimension [7] or two dimension [8] with infinite potential well setting, with the main difference being the loss functions and how the eigenvalue is calculated.

Transfer learning and the less impact from curse of dimensionality is the main motivation of the proposal of this scheme, which might be useful for cavity optimization problem. Several authors have published cavity optimization scheme using evolutionary algorithm with conventional finite element as the evaluation function [9]. Each of function evaluation will require simulation from scratch, which might be excessive when the variation is small [10, 11].

METHODS

We propose a fully connected neural network trained without external data to calculate the electromagnetic excitation inside a perfectly conducting resonant cavity. Three neurons representing position are used as the input, and six fields which correspond to three electric field components and three magnetic field components as the output. The network are trained to minimize Helmholtz equations related to the Maxwell equatios, along with the divergence free equation as the following [12]

$$\left(\nabla^2 + \mu\varepsilon\omega^2\right)\psi_i = 0 \qquad i = 1, ..., 6 \tag{1}$$

$$\nabla \cdot \vec{E} = 0 \tag{2}$$

$$\nabla \cdot \vec{H} = 0 \tag{3}$$

with ψ_i represents the component of electric or magnetic field. Additionally, the network also need to satisfy the boundary conditions given by

$$\vec{n} \cdot \vec{H} = 0 \tag{4}$$

$$\vec{n} \times \vec{E} = 0 \tag{5}$$

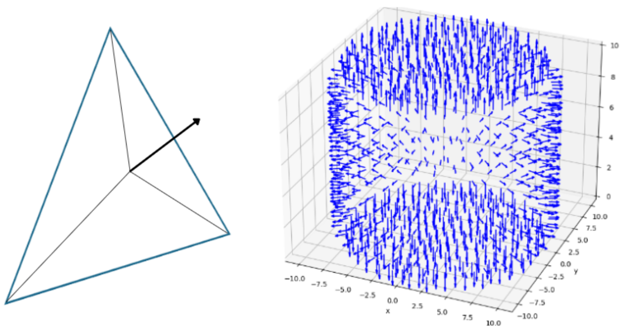
where $\vec{n}$ is the normal vector perpendicular to the surface of the resonant cavity at the sampled boundary point. Equation (1) − (5) will still be satisfied when $\vec{E} = \vec{H} = 0$, which is unwanted. To avoid obtaining trivial solution, a loss function related to the energy normalization is used

$$\int \frac{\varepsilon|\vec{E}|^2}{2} dV + \int \frac{\mu|\vec{H}|^2}{2} dV = 1 \tag{6}$$

One cavity is different with each other based on the geometry of the cavity. In our case the difference is represented by the normal vector $\vec{n}$ sampled on the boundary points. Thus, PINN for resonant cavity cannot be made completely mesh free [4], unlike other PINN implementations. The surface

* ahsahafi@rcnp.osaka-u.ac.jp

mesh is obtained using Gmsh scheme [13] with triangular shape surface mesh, while the sampling for interior points are done with uniform spacing. The normal vector is then calculated using the cross product of two sides of the triangular cell, and then the normal vector is positioned in the middle of the triangle. The normal vector used to calculate boundary condition loss function is shown in Fig. 1.



a. b.

Figure 1: a. Illustration of how normal vector is related to triangular mesh. b. Sampled normal vectors for a cylindrical shape surface.

The total loss function is a linear combination of individual loss functions related to the minimized equations, which can be written as

$$L = w_{\text{in}}L_{\text{in}} + w_{\text{div}}L_{\text{div}} + w_{\text{norm}}L_{\text{norm}} + w_{\text{bc}}L_{\text{bc}} + w_{\text{m}}L_{\text{m}} \tag{7}$$

with w_i's are the weights and L_i's are the losses. The subscript *in*, *div*, *norm*, *bc*, and *m* correspond to Helmholtz equation, divergence free relation, normalization requirement, boundary condition, and mode forcing loss respectively. In general, the weights is not set equal to each other. Here, the optimized composition of loss weights are found through trial and errors.

The mode forcing loss function is used to enforce the model to learn a certain excitation mode. For instance, if transverse magnetic (TM) mode is desired, $H_z = 0$ can be used. Likewise, if transverse electric (TE) mode is desired, additional loss function E_z can also be used. This will help the model to converge more easily to the desired mode, especially when the mode is not the lowest frequency mode for the given cavity.

The minimization of Helmholtz equation error requires the eigenvalue $\lambda = -\mu\varepsilon\omega^2$ to be known. In this paper, the eigenvalue is calculated by choosing one of Helmholtz equation of the field that we suspect is non-trivial, then multiplying the equation with the said field, and integrating the equation with respect to volume. The scheme is similar to the calculation method energy expectation value for Schrödinger [14]. The equation can also be written in discrete form as follows

$$\lambda = \frac{\int \psi_i \nabla^2 \psi_i dV}{\int \psi_i^2 dV} = \frac{\sum_j \psi_{ij} \nabla^2 \psi_{ij} dV_j}{\sum_j \psi_{ij}^2 dV_j} \tag{8}$$

where the index j indicates the node volume related to the sampled interior points. For uniform sampling, the value of dV_j will be the same for all sampled interior points, thus for that case dV_j can be dropped.

Training Procedure

The fully connected network proposed here consists of five hidden layer with forty neurons in each hidden layer, illustrated in Fig. 2. All of the hidden layers use sine as the activation function. The training done for 25000 epochs using a single batch for all cases. Adam optimizer is used to minimize the loss function with the initial learning rate is set to be 10^{-3} and then changed to 10^{-4} after 10000 epochs.

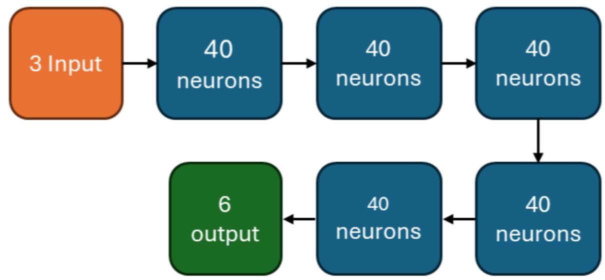
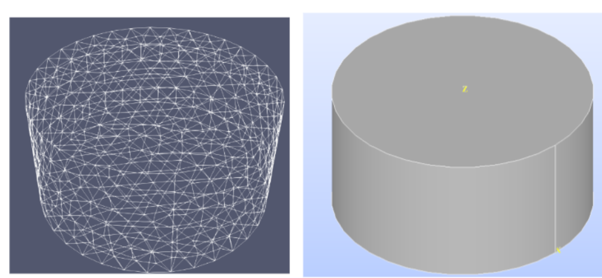


Figure 2: Fully connected network used in this paper.

A simple cylindrical cavity shown in Fig. 3 is used as the test case. The height and the radius of the cylinder is both equal to 10 (unitless). For that particular cavity, the eigenvalue obtained using analytical calculation for TM_{010} and TE_{111} are 0.0578 and 0.1325. 1088 points are sampled at the boundary following Gmsh algorithm, while 2652 interior points are sampled in uniform manner. For simplicity, both of permittivity and permeability of free space is set equal to one.



a. b.

Figure 3: a. Model of cavity with a. Mesh view. b Geometry view.

Two modes are considered here, the first one is TM_{010} and the second one is TE_{111}. As mentioned previously, to enforce TM and TE case, additional mode forcing loss function can be used. Here, in addition of setting $H_z = 0$ and $E_z = 0$ to make it even easier for the network to learn the field pattern, additional $E_x = E_y = 0$ and $H_x = E_y = 0$ is used for TM and TE case respectively.

Adaptive weighting scheme is adopted in this paper, where the loss function weights changes after a certain amount of epochs has passed. The scheme is used to make it easier for certain characteristic of solution to occur first. For example, in this case the most important loss function is the boundary condition loss (because it distinguish one cavity with another) so the loss weight for boundary condition should be made high. However, setting it high from the beginning might make it hard for the system to have a low normalization error. In essence, the model often prioritize itself to satisfy the boundary condition only, often resulting in trivial solution. Thus, initially the boundary condition loss weight is made very small so that the electric and magnetic field is large enough. The Helmholtz loss weight is also made small initially since the calculation will not be accurate as long as the boundary is not fixed yet. The initial loss weights and the latter changes are shown in Table 1 and Table 2 respectively

Table 1: Initial Loss Weights

No.	Loss Type	Weights
1	Helmholtz equation	20
2	Divergence free equation	10000
3	Normalization constraint	0.01
4	Boundary condition	10
5	Mode forcing	1000

Table 2: Modified Loss Weights

No.	Loss Type	Weights	Epochs
1	Helmholtz equation	200	3000
2	Boundary condition	1000	6000

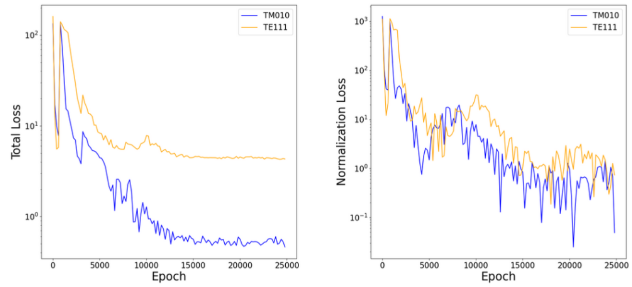
All of the procedure shown in this paper is carried using Research Center of Nuclear Physics' workstation PC. The workstation uses Intel core-i9 14900KF processor and NVIDIA RTX 4070 Super with 12 GB of VRAM. The code is written using pytorch, and is trained completely using the GPU.

RESULTS

Training result as a function of epoch is presented in Fig. 4a. The simulation took about 32.4 minutes and 33.11 minutes for TM_{010} and TE_{111} mode respectively.

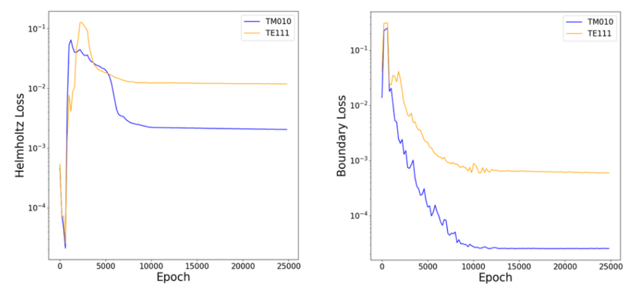
Figure 4a shows that the total loss could be minimized by the neural network model, and that the total loss almost plateaued after 15000 iterations (slight fluctuation in the graph is caused by the logarithmic scaling). From Fig. 4b it is clear that the normalization loss is significantly larger than other losses, which is the main motivation of lowering the value of normalization loss weight. It is evident from Fig 5a and 5b that it is harder for the network to minimize the Helmholtz and boundary loss function of TE_{111} mode compared to TM_{010}. It is likely that it is easier for the network to converge to the mode with lower eigenvalue.

There are two big bumps in the total loss, at epoch 600 and 3000, which is the point where the boundary condition



a. b.

Figure 4: a. Loss vs epoch for a. Total loss. b. Normalization loss.



a. b.

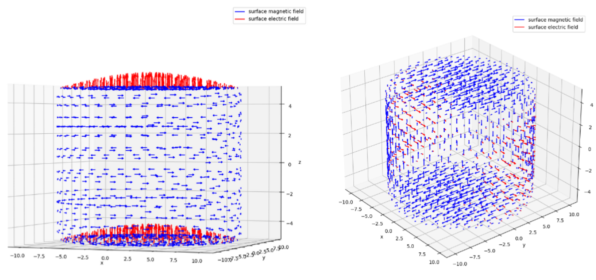
Figure 5: a. Loss vs epoch for a. Helmholtz loss. b. Boundary loss.

weight and Helmholtz equation weight is increased. The bump implies that before the weight is increased, the network simply ignore the significance of boundary condition and the Helmholtz equation. After the boundary loss is prioritized by the increase of weight, the network need to give compromise to other losses, which will affect the loss value. It is proven from the result that the emergence of non-trivial solution can be avoided by using the adaptive scheme.

As can be seen from the normalization loss graph, it is harder for the network in TE_{111} case to minimize normalization loss after the boundary condition weight is increased, compared to TM_{010}. The normalization loss stayed high for about 1000 epochs after the boundary weight is increased. It is likely that the network tend to converge to the lowest frequency mode first, which is the TM_{010} mode for this cavity.

Figure 4a shows that the total loss function start to plateau after 10000 epochs. The result is not that dissimilar with other published physics informed neural network algorithm that has been published [3, 5, 15]. An author has proposed a scheme to accelerate the convergence by using gradient enhanced loss function [16]. The procedure is not done here because in doing so will increase the number of loss function in a model with already a lot of loss functions.

The neural network was able to predict the general distribution of the electromagnetic field (the analytical calculation can be seen on [12]). The electromagnetic field distribution at the surface of the cylindrical cavity for TM_{010} and TE_{111} are shown on Fig. 6a and 6b respectively. The high bound-



a. b.

Figure 6: Surface electric field (red) and magnetic field (blue) for a. TM_{010}. b. TE_{111} mode.

ary condition loss enforce the electric field to be always perpendicular to the surface while the magnetic field is always parallel to the surface. The radial distribution of axial electric field E_z also resembles zeroth order Bessel function of the first kind. Such distribution is the result of Helmholtz equation loss function. However, the loss weight cannot be made too large, as it will make the network to prioritize the Helmholtz equation over the boundary condition, resulting in an incorrect distribution (and in turn, incorrect eigenvalue).

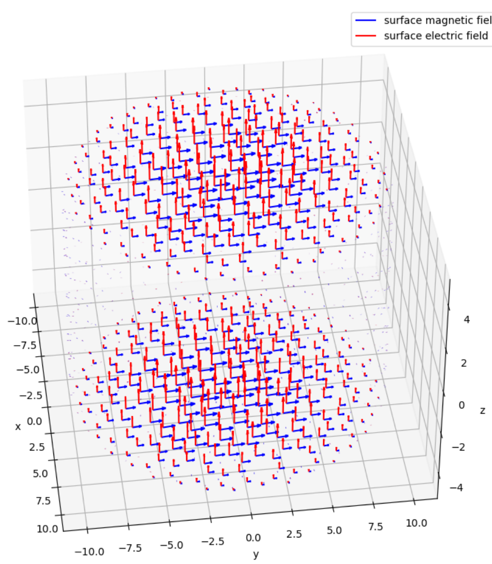


Figure 7: Surface electric field (red) and magnetic field (blue) for TM_{010} mode with no zero divergence constraint.

The divergence loss weight was set very high from the beginning to emphasize the importance the importance of the constraint. Without it, the field will find a configuration that still satisfy the boundary and Helmholtz equation constraint, but might be unphysical. Figure 7 shows when the divergence constraint is ignored. The figure shows that the magnetic field could behaves as if it is originated from a source, which is obviously not the case.

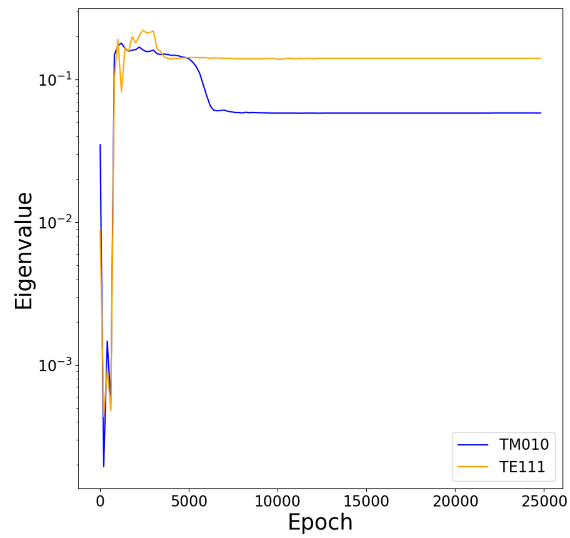


Figure 8: Eigenvalue vs epoch for TM_{010} and TE_{111} mode.

Figure 8 shows the evolution of eigenvalue with respect to epoch. At the early stage of training, the eigenvalue is much smaller than the actual value. During this period, the electric and magnetic field are both non-trivial, since the normalization error is quite small. Therefore, the significant discrepancy stems from the fact that boundary condtion is not obeyed at the early stages. When the boundary condition loss weight is made significant, suddenly the eigenvalue start to gain value comparable to the actual eigenvalue.

The eigenvalue for TM_{010} and TE_{111} mode are 0.0584 and 0.141 respectively. The error for each computed eigenvalue with respect to the analytical value is 0.98% and 6.9% respectively. It is likely that transverse electric case has a larger error than transverse magnetic case because the network converge more easily to the ground state instead one of the state with higher frequency above.

For this simple cavity, the error is significantly larger than conventional finite element methods. Some authors have pointed out about the possibility of improving the PINNs by modifying the network (at the very least reducing convergence time), such as by using symmetrical or antisymmetric network or by using Fourier features [17]. In this case however, it can be seen that the eigenvalue plateaued after about 6000 epochs, which means that the network already consider the computed result as the best value. It would seem that another method to compute the eigenvalue is more suited to this problem. There has been proposal on the use of one input weight at the network as the eigenvalue [15].

As mentioned earlier, the simulation took about half an hour to finish. For simple cavity, this might seem to be too lengthy, since commercial software on eigenmode simulation only requires simulation on order of seconds for such simple cavity (it might took a longer time if the mesh is refined, but for simple cavity, the numerical accuracy will not increase significantly). The benefit of PINNs would be when larger sampling size (corresponding to fine mesh) is required, as the simulation duration will not scale exponentially with respect to sample size. In addition, transfer learning might also greatly reduce the required simulation duration. Initial network can be trained using a simple cavity as the base model. For a more complex cavity with larger sampling size, the base model can be used for starting point of the simulation. These assertions will require further tests.

CONCLUSION

A physics informed neural network model to solve RF eigenmode problem on a resonant cavity have been proposed and tested. The network could give a relatively good prediction on the lowest transverse magnetic and transverse electric mode. A further research on an improved network and eigenvalue calculation scheme is necessary to reduce training duration and improve accuracy. This scheme might be useful for cavity optimization problem with the help of transfer learning and its ability to handle large data samples.

REFERENCES

[1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[2] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational physics*, vol. 378, pp. 686–707, 2019.

[3] X. Chen, R. Chen, Q. Wan, R. Xu, and J. Liu, "An improved data-free surrogate model for solving partial differential equations using deep neural networks," *Scientific reports*, vol. 11, no. 1, p. 19 507, 2021.

[4] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, no. 6, pp. 422–440, 2021.

[5] S. Markidis, "The old and the new: Can physics-informed deep-learning replace traditional linear solvers?" *Frontiers in big Data*, vol. 4, p. 669 097, 2021.

[6] L. Lu, R. Pestourie, W. Yao, Z. Wang, F. Verdugo, and S. G. Johnson, "Physics-informed neural networks with hard constraints for inverse design," *SIAM Journal on Scientific Computing*, vol. 43, no. 6, B1105–B1132, 2021.

[7] H. Jin, M. Mattheakis, and P. Protopapas, "Unsupervised neural networks for quantum eigenvalue problems," *arXiv preprint arXiv:2010.05075*, 2020.

[8] E. G. Holliday, J. F. Lindner, and W. L. Ditto, "Solving quantum billiard eigenvalue problems with physics-informed machine learning," *AIP Advances*, vol. 13, no. 8, 2023.

[9] S. Smith, M. Southerby, S. Setiniyaz, R. Apsimon, and G. Burt, "Multiobjective optimization and pareto front visualization techniques applied to normal conducting rf accelerating structures," *Physical Review Accelerators and Beams*, vol. 25, no. 6, p. 062 002, 2022.

[10] S. Shin and J. Chai, "Optimization of the rf cavity of the medical purpose electron linac by using genetic algorithm," *Generations*, vol. 9200, p. 9250, 2014.

[11] M. Kranjčević, A. Adelmann, P. Arbenz, A. Citterio, and L. Stingelin, "Multi-objective shape optimization of radio frequency cavities using an evolutionary algorithm," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 920, pp. 106–114, 2019.

[12] J. D. Jackson, *Classical electrodynamics*. John Wiley & Sons, 2021.

[13] C. Geuzaine and J.-F. Remacle, "Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities," *International journal for numerical methods in engineering*, vol. 79, no. 11, pp. 1309–1331, 2009.

[14] H. Li, Q. Zhai, and J. Z. Chen, "Neural-network-based multistate solver for a static schrödinger equation," *Physical review A*, vol. 103, no. 3, p. 032 405, 2021.

[15] H. Jin, M. Mattheakis, and P. Protopapas, "Physics-informed neural networks for quantum eigenvalue problems," in *2022 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2022, pp. 1–8.

[16] J. Yu, L. Lu, X. Meng, and G. E. Karniadakis, "Gradient-enhanced physics-informed neural networks for forward and inverse pde problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 393, p. 114 823, 2022.

[17] 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," *Computer Methods in Applied Mechanics and Engineering*, vol. 384, p. 113 938, 2021.