Physics-Informed Neural Networks for Estimation of Scattered Sound Fields with Boundary Condition

Ryosuke Onizawa*, Gen Sato*, Izumi Tsunokuni* and Yusuke Ikeda*
* Tokyo Denki University, Tokyo
E-mail: {23fmi07, 24udc03, 21udc02}@ms.dendai.ac.jp, yusuke.ikeda@mail.dendai.ac.jp

***Abstract*—Sound field measurement helps us to understand the sound field, which is difficult to understand only by hearing. Recently, to reduce the measurement points, methods for estimating sound fields using physics-informed neural networks (PINNs) based on the wave equation or the Helmholtz equation have attracted much attention. In this study, we propose the sound field estimation method for the scattered sound field with PINNs by introducing the boundary condition into the loss function as additional physical laws. From the simulation experiments, the proposed method can improve the estimation accuracies in three-dimensional sound field with a scattering object with the acoustic impedance.**

I. Introduction

In an ordinary room, the sound propagates through reflections, diffraction, and scattering , depending on the materials and boundary shapes of the objects such as walls and furniture. Analyzing such a complex sound field is useful for a variety of applications, including room acoustics design and acoustic rendering in virtual reality.

Especially, the acoustic scattering on walls and objects complicates the sound field and is difficult to be modeled. In recent years, many modeling methods have been proposed to estimate the scattered sound fields. Based on the concept of compressive sensing [1], [2], the actual sound field can be modeled only by a small number of microphones. In [3], a scattered sound field from a rigid sphere is modeled using a sparse equivalent source method [4] with a small number of microphones. Using the superposition of point sources as a physical model of sound propagation [5], the scattered sound field from a rigid sphere can be estimated at lower frequencies.

On the other hands, recently, the estimation methods using deep learning have been proposed [6], [7]. In [8], the sound pressure level in the scattered sound field was estimated using a convolutional neural network (CNN) from the geometry of scattering objects. However, the typical data-driven methods do not guarantee that the physical laws of the outputs will hold, since the models are trained only on the dataset.

By contrast, a physics-informed neural network (PINN) has been proposed in 2019 [9]. A notable feature of PINN is that the output follows the laws of physics by introducing governing equations in the loss function. In acoustics, PINN has been applied in studies to estimate the sound field [10]–[13]. In these studies, the wave equation or Helmholtz equation is used as the governing equation for acoustics. Furthermore, the scattered sound field was estimated around a rigid sphere by using the wave equation and Neumann boundary condition using 32 microphones [14]. However, the problem of the case where the scattering object is not rigid, i,e., where the acoustic impedance of the object is not infinite, has not been investigated.

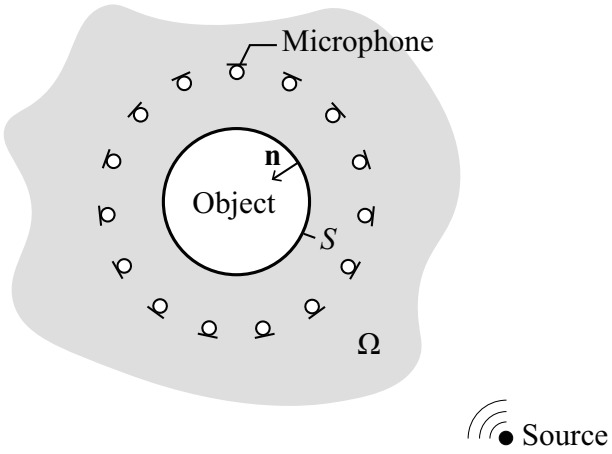

Fig. 1. Concept of estimation for scattered sound field Ω around a scattering object from an enclosing microphone array. The incident sound field is propagated from a sound source outside the microphone array. The boundary S of the scattering object has the acoustic impedance. $\mathbf{n}$ is normal vector on the boundary S.

In this study, based on the PINN, we propose an estimation method of the scattered sound field around a object with the acoustic impedance. The proposed method introduces the Helmholtz equation and the boundary condition into the loss function. Notably, in this study, the acoustic impedance of each boundary element is assumed to change smoothly with respect to adjacent boundary elements. Then, we introduce a loss function such that the differences between the acoustic impedance of the neighbor boundary elements become zero. Thus, we will model and estimate the scattered sound field from a object with an arbitrary acoustic impedance.

The rest of this paper is organized as follows. Section II describes the problem statement and proposed method. Section III-A describes the simulation conditions. Section III-B describes the results of three simulation experiments. Finally, we conclude this work in section IV.

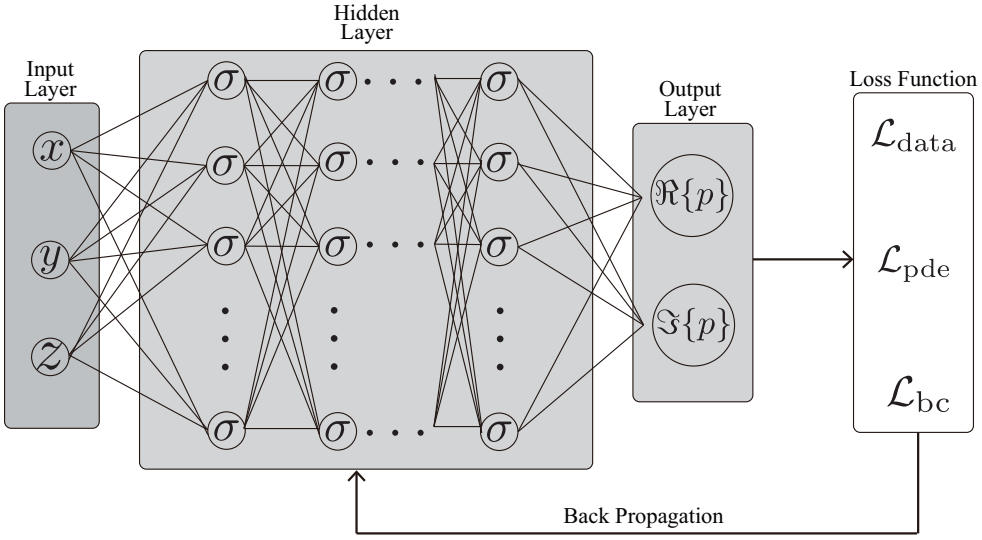

Fig. 2. Network architecture of proposed method. Input is three-dimensional coordinates, σ is the activation function, $\Re\{p\}$ and $\Im\{p\}$ are real and imaginary parts of complex sound pressure, respectively.

II. Proposed Method

A. Problem Statement

Fig. 1 shows the concept of the problem of scattered sound field estimation. This study estimates the three-dimensional scattered sound field Ω around a scattering object with spatially-smooth absorption coefficients. We assume that the sound sources are located outside the region Ω and outside the microphone array. That is, since the sound field in Ω is homogeneous, the Helmholtz equation [15] holds as follows,

$$k^2 p(\mathbf{x}) + \nabla^2 p(\mathbf{x}) = 0 \quad (\mathbf{x} \in \Omega) \tag{1}$$

where k is wave number, ∇ is the Laplacian operator and $p(\mathbf{x})$ is the sound pressure $p(\in \mathbb{C})$ at the position $\mathbf{x}(\in \mathbb{R}^3)$. The N measurement points at position $\mathbf{x}_n (n = 1, 2, ..., N)$ are positioned around the scattering object in the region Ω.

A acoustic impedance is defined at the boundary of the scattering object in Ω and is represented as follows.

$$Z(\mathbf{x}) = \frac{\hat{p}(\mathbf{x})}{v(\mathbf{x})} \quad (\mathbf{x} \in S) \tag{2}$$

where v is normal particle velocity to the boundary S, $\hat{p}$ is sound pressure on the boundary. In this study, the acoustic impedance $Z(\mathbf{x})$ on S is assumed to be approximately uniform.

B. Proposed Method

Raissi *et al.* proposed a PINN, which introduces the governing equations into the loss function [9]. This study uses the Helmholtz equation as the governing equation to obtain an output that follows the laws of sound propagation. In addition, the proposed method introduces the boundary conditions into the loss function.

The proposed method employs the SIREN architecture [16] to perform higher-order derivatives of the loss function. Fig. 2 shows the network architecture of the proposed method. Using multilayer-perceptron (MLP), the function $\phi(\cdot)$ and the estimated sound field $\hat{p}(\mathbf{x})$ is represented as follows:

$$\hat{p}(\mathbf{x}) = \phi(\mathbf{w}, \mathbf{x}) = (f_L \circ f_{L-1}, ..., \circ f_2 \circ f_1) \tag{3}$$

where $\mathbf{x} = (x, y, z)$ is the coordinates as input and $\mathbf{w}$ is the weights. f_l is the function of the l-th layer of MLP. Then, the Helmholtz equation in (1) must be hold on the estimated sound pressures $p(\mathbf{x})$. In other words, the model should be trained so that the outputs satisfy the laws of acoustic physics.

As shown in (1), the Helmholtz equation has the second derivative of p. ReLU, one of the typical activation functions, is not appropriate for calculating the Helmholtz equation because it becomes zero after second-order differentiation. To obtain the derivatives by the automatic differentiation, it is required to choose the activation functions that are two-time differentiable. Thus, in this study, we use the sine function as an activation function. The MLP with sine activation function is known as the SIREN architecture [16]. Other examples include the use of hyperbolic tangent functions for activation functions [14].

By using the sine function, the l-th layer function f_l in (3) is as follows:

$$f_l(\mathbf{x}_l) = \sin\left(\omega(\mathbf{x}_l \mathbf{w}_l + \mathbf{b}_l)\right) \tag{4}$$

where $\mathbf{x}_l$ is the input of l-th layer, $\mathbf{w}_l$ is the weight parameter, $\mathbf{b}_l$ is bias of l-th layer and ω is hyperparameter.

To determine the weight parameter $\mathbf{w}$, the PINN introduces a law of physics or a governing equation into the loss function L. By also using the complex sound pressures measured at the microphone positions, the loss function $\mathcal{L}$ is defined by

$$\mathcal{L} = \mathcal{L}_{\mathrm{data}} + \lambda_1 \mathcal{L}_{\mathrm{pde}} + \lambda_2 \mathcal{L}_{\mathrm{bc}} \tag{5}$$

$$\mathcal{L}_{\mathrm{data}} = \frac{1}{N} \sum_{n=1}^{N} \left| \hat{p}(\mathbf{x}_n) - p(\mathbf{x}_n) \right|^2 \ (\mathbf{x}_n \in \Omega) \tag{6}$$

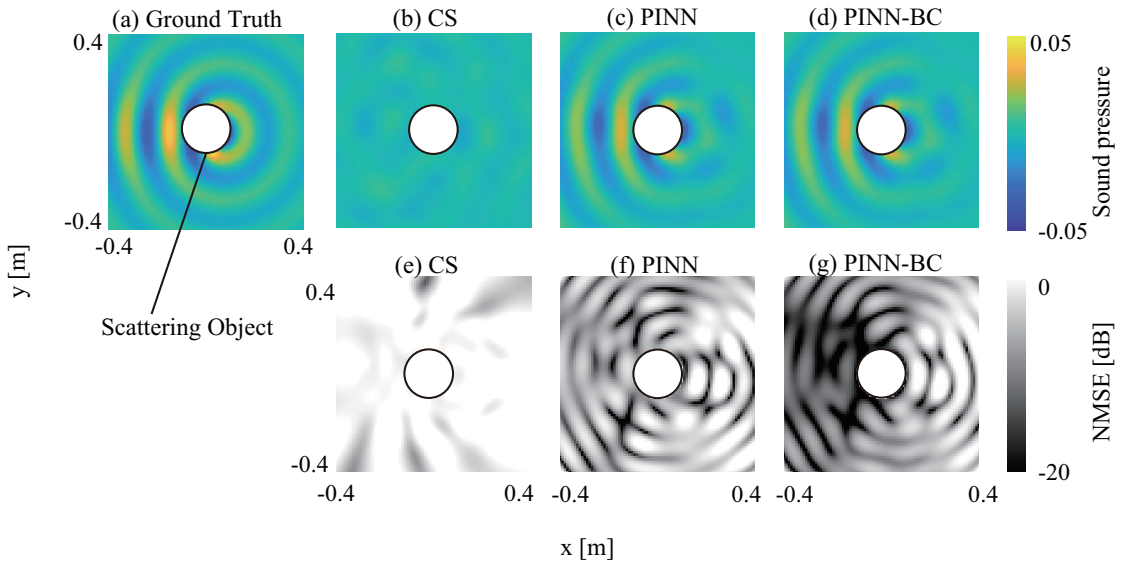

Fig. 3. Comparison of scattered sound fields and NMSE distributions ($z = 0$) at frequency is 2 kHz. (a) Ground Truth, (b) Estimated sound field by CS. (c) Estimated sound field by PINN (without $\mathcal{L}_{\mathrm{bc}}$), (d) Estimated sound field by the proposed method, PINN-BC (with the loss of boundary condition $\mathcal{L}_{\mathrm{bc}}$), (e)-(g) are NMSE distributions by CS, PINN and PINN-BC, respectively. The number of microphone for the estimation was 32. The circle at the center indicates the scattering object.

TABLE I
SIMULATION CONDITION

Frequency [Hz]	2000
Absorption rate	0.5
Radius of scattering object [m]	0.1
Number of microphones	16×2, 25×2, 36×2
Radius of microphone array [m]	0.2, 0.25
Balance parameterλ_1, λ_2	10^{-4}, 1.0

$$\mathcal{L}_{\mathrm{pde}} = \frac{1}{M}\sum_{m=1}^{M}\left|k^2\hat{p}(\mathbf{x}'_m) + \nabla^2\hat{p}(\mathbf{x}'_m)\right|^2 \quad (\mathbf{x}'_m \in \Omega) \tag{7}$$

$$\mathcal{L}_{\mathrm{bc}} = \sum_{e\in E}\sum_{e'\in E'}\left|Z(\mathbf{x}''_e) - Z(\mathbf{x}''_{e'})\right|^2 \quad (\mathbf{x''}_e, \mathbf{x''}_{e'} \in S) \tag{8}$$

where $\|\cdot\|_2$ is ℓ_2 norm, $\hat{p}$ and p are estimated and measured sound pressures, respectively. λ_1, λ_2 are balance parameters. $\mathbf{x}_n$ is the microphone position, $\mathbf{x}'_m$ is three-dimensional coordinates selected randomly, $\mathbf{x''}_e$ is the point on the boundary of scattering object, $\mathbf{x''}_{e'}$ is a point neighbor to $\mathbf{x''}_e$.

$\mathcal{L}_{\mathrm{data}}$ is a loss function to calculate MSE between the estimated and measured sound pressures. $\mathcal{L}_{\mathrm{pde}}$ is a loss function to evaluate how well the estimated sound pressures satisfy the Helmholtz equation, as shown in (1). $\mathcal{L}_{\mathrm{bc}}$ is a loss function to evaluate the difference in acoustic impedance at the neighbor points. Thus, by learning the weight parameter $\mathbf{w}$ so that $\mathcal{L}_{\mathrm{pde}}$ and $\mathcal{L}_{\mathrm{bc}}$ become small, the output of the MLP will more closely follow the Helmholtz equation and the boundary condition.

III. Simulation Experiment

A. Simulation Condition

We conducted the simulation experiment to evaluate the accuracy of estimating the scattered sound field around a sphere with a radius of 0.1 m in three dimensions. For the ground truth, the analytical solution [17] was calculated with spherical harmonics. The center of the scattering sphere object is the origin of the coordinate system. For $\mathcal{L}_{\mathrm{data}}$, the equally-spaced spherical microphone array had two layers with radii of 0.2 m and 0.25 m. The microphone array was positioned at the origin. The numbers of microphones were $M = 32, 50, 72$. The sound speed was 343 m/s, and the source position is placed at $(x, y, z) = (2, 0, 0)$ m. We employ the Adam optimizer with a learning rate of 1.0×10^{-5} for 15000 epochs. The number of randomly selected positions M for $\mathcal{L}_{\mathrm{pde}}$ is 10000 at each epoch. The number of boundary elements for $\mathcal{L}_{\mathrm{bc}}$ is 400 with equal spacing on the surface S. The weight parameters λ_1 and λ_2 are 1.0×10^{-4} and 1.0, respectively. The other details of simulation conditions are shown in Table I.

To evaluate the accuracy of the scattered sound field, we defined the normalized mean squared error (NMSE) as

$$\mathrm{NMSE} = 10\log_{10}\frac{\int\left|\hat{p}(\mathbf{x}) - p(\mathbf{x})\right|^2 d\mathbf{x}}{\int\left|p(\mathbf{x})\right|^2 d\mathbf{x}}. \tag{9}$$

B. Result

First, we compare the estimation accuracies of the proposed method (PINN–BC) and the two conventional methods: equivalent source method (CS) [3], [5] and PINN with no boundary condition (PINN). In CS, the scattered sound field was modeled from the 32 microphones using the sparse equivalent source method.In simple PINN, the model was trained without using the loss of boundary condition $\mathcal{L}_{\mathrm{bc}}$ (8).

Fig. 3 shows the comparison of scattered sound fields and NMSE distributions at 2 kHz by CS, simple PINN, and PINN–BC (proposed method) when 32 microphones were used. From Fig. 3 (a)–(d), it is evident that simple PINN and PINN–BC were able to estimate the scattered sound field

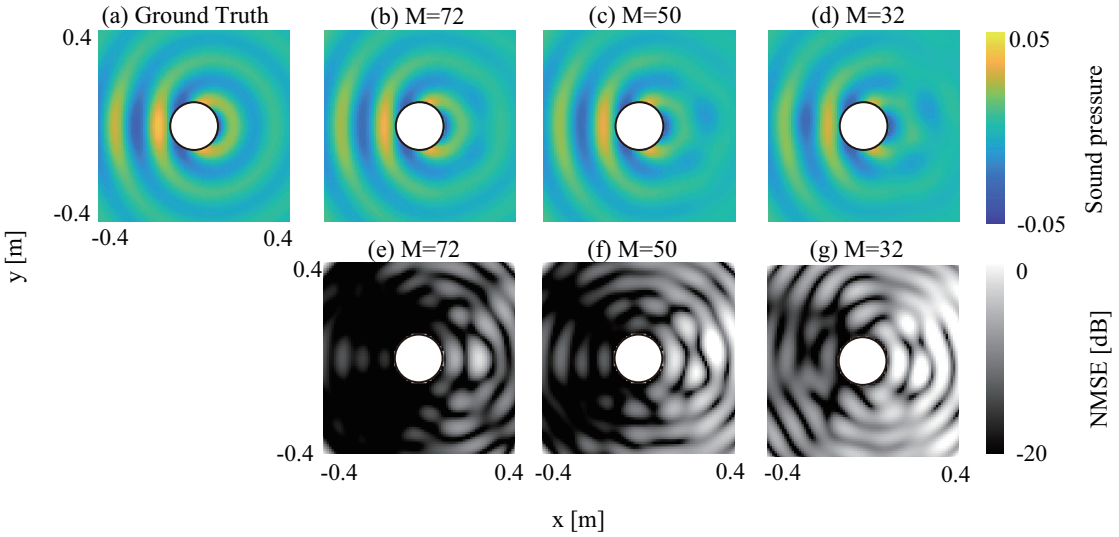

Fig. 4. Comparison of scattered sound fields and NMSE distributions ($z = 0$) by the number of microphones. (a) Ground Truth, (b) and (e) 72 microphones, (c) and (f) 50 microphones, (d) and (g) 32 microphones.

TABLE II
COMPARISON OF MEAN NMSES ON THE X-Y PLANE AND IN THREE-DIMENSIONAL EVALUATION POINTS

	CS	PINN	PINN–BC
Mean NMSE ($z = 0$)	0.2 dB	-5.1 dB	-8.3 dB
Mean NMSE (3D)	0.0 dB	-3.9 dB	-8.5 dB

significantly closer to the ground truth compared to CS. From Fig. 3 (e)–(g), PINN and PINN–BC significantly improved the estimation accuracies over the estimation region compared to CS. Furthermore, compared to simple PINN, PINN-BC has significantly improved the accuracy of estimating the scattered sound field, especially at locations opposite the sound source.

Table II shows the comparison of mean NMSEs on the x–y plane ($z = 0$) and in the three-dimensional evaluation points. For the $z = 0$ plane, the evaluation region was a square with 0.4 m on each side. For the 3D evaluation points, a spherical region between a radius of 0.1 m and 0.4 m was used as the evaluation region. On the $z = 0$ plane, the mean NMSEs of PINN and PINN–BC were improved by approximately 5.3 dB and 8.5 dB, respectively, compared to CS. Furthermore, in three-dimensional evaluation points, the estimation accuracy of PINN–BC was more stable than the other methods, with an estimation accuracy of -8.5 dB, approximately equivalent to the accuracy of the $z = 0$ plane. The proposed method achieves an improvement in accuracy of 4.6 dB NMSE compared to PINN and 8.5 dB NMSE compared to CS. Therefore, the proposed method improved the estimation accuracies of scattered sound fields by introducing the boundary condition into the loss function.

Second, we compare the estimation accuracies by the number of microphones in the proposed method. Fig. 4 shows the estimated sound fields and NMSE distributions with 32, 50, and 72 microphones. As shown in Figs. 4(a)–(d), it was found that the scattered sound field can be estimated with high accuracy regardless of the number of microphones. From Figs. 4 (e)–(g), as the number of microphones decreased, the estimation accuracies were significantly degraded over the entire region. However, it is clear that even with 32 microphones, the high estimation accuracy was maintained on the opposite side of the sound source. Based on these results, the number of microphones should be increased according to the desired estimation accuracy.

IV. CONCLUSIONS

In this study, we proposed a method for estimating the scattered sound field from a small number of microphones using a physical model and deep learning. From the simulation experiments, the proposed method significantly improved the estimation accuracy by approximately 4.5 dB in NMSE when boundary conditions are introduced into the loss function as an additional physical model. From the other simulation experiments, it was found that the estimation accuracy significantly decreased with the number of microphones. However, even with a small number of microphones, it was found that it is possible to estimate a wide range of scattered sound fields.

In future work, we will study the estimation of scattered sound fields in the broad frequency band. Furthermore, we will estimate the scattered sound field around more complicated shapes of scattering objects. We also aim to evaluate the estimation accuracy in actual measurements and environments.

ACKNOWLEDGMENT

This work was partially supported by JSPS KAKENHI Grant Number 22K12099.

REFERENCES

[1] E. J. Candes and M. B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[2] P. Gerstoft, F. C. Mecklenbräuker, W. Seong, and M. Bianco, "Introduction to compressive sensing in acoustics," *The Journal of the Acoustical Society of America*, vol. 143, no. 6, pp. 3731–3736, Jun. 2018.

[3] R. Onizawa, I. Tsunokuni, and Y. Ikeda, "Visualization of scattered sound field by enclosing microphone array based on sparse equivalent source method," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, 2023, pp. 6771–6776.

[4] E. Fernandez-Grande, A. Xenaki, and P. Gerstoft, "A sparse equivalent source method for near-field acoustic holography," *The Journal of the Acoustical Society of America*, vol. 141, no. 1, pp. 532–542, Jan. 2017.

[5] I. Tsunokuni, K. Kurokawa, H. Matsuhashi, Y. Ikeda, and N. Osaka, "Spatial extrapolation of early room impulse responses in local area using sparse equivalent sources and image source method," *Applied Acoustics*, vol. 179, p. 108 027, 2021.

[6] F. Lluís, P. Martínez-Nuevo, M. B. Møller, and S. E. Shepstone, "Sound field reconstruction in rooms: Inpainting meets super-resolution," *The Journal of the Acoustical Society of America*, vol. 148, no. 2, pp. 649–659, 2020.

[7] E. Fernandez-Grande, X. Karakonstantis, D. Caviedes-Nozal, and P. Gerstoft, "Generative models for sound field reconstruction," *The Journal of the Acoustical Society of America*, vol. 153, no. 2, pp. 1179–1190, Feb. 2023.

[8] Z. Fan, V.Vineet, H. Gamper, and N. Raghuvanshi, "Fast acoustic scattering using convolutional neural networks," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 171–175.

[9] M. Raissi, P. Perdikaris, and G. 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.

[10] K. Shigemi, S. Koyama, T. Nakamura, and H. Saruwatari, "Physics-informed convolutional neural network with bicubic spline interpolation for sound field estimation," in *2022 International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2022, pp. 1–5.

[11] X. Karakonstantis and E. Fernandez-Grande, "Room impulse response reconstuction using physics-constrained neural networks," in *Proc. the 10th Convention of the European Acoustics Association Forum Acusticum 2023*, 2023, pp. 3181–3188.

[12] M. Pezzoli, F. Antonacci, and A. Sarti, "Implicit neural representation with physics-informed neural networks for the reconstruction of the early part of room impulse responses," in *Proc. the 10th Convention of the European Acoustics Association Forum Acusticum 2023*, 2023, pp. 2177–2184.

[13] I. Tsunokuni, G. Sato, Y. Ikeda, and Y. Oikawa, "Spatial extrapolation of early room impulse responses with noise-robust physics-informed neural network," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. advpub, 2024EAL2015, 2024.

[14] X. Chen, F. Ma, A. Bastine, P. Samarasinghe, and H. Sun, "Sound field estimation around a rigid sphere with physics-informed neural network," in *2023 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2023, pp. 1984–1989.

[15] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. 1999.

[16] V. Sitzmann, J. N. P. Martel, A. W. Bergman, D. B. Lindell, and G. Wetzstein, "Implicit neural representations with periodic activation functions," in *Proc. NeurIPS*, 2020.

[17] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Elsevier Science, 2005.