forum acusticum 2023

CIRCUMVENT SPHERICAL BESSEL FUNCTION NULLS FOR OPEN SPHERE MICROPHONE ARRAYS WITH PHYSICS INFORMED NEURAL NETWORK

Fei Ma   **Thushara D. Abhayapala**   **Prasanga N. Samarasinghe**
College of Engineering, Computing & Cybernetics, Australian National University, Australia

ABSTRACT

Open sphere microphone arrays (OSMAs) are simple to design and do not introduce scattering fields, and thus can be advantageous than other arrays for implementing spatial acoustic algorithms under spherical model decomposition. However, an OSMA suffers from spherical Bessel function nulls which make it hard to obtain some sound field coefficients at certain frequencies. This paper proposes to assist an OSMA for sound field analysis with physics informed neural network (PINN). A PINN models the measurement of an OSMA and predicts the sound field on another sphere whose radius is different from that of the OSMA. Thanks to the fact that spherical Bessel function nulls vary with radius, the sound field coefficients which are hard to obtain based on the OSMA measurement directly can be obtained based on the prediction. Simulations confirm the effectiveness of this approach and compare it with the rigid sphere approach.

Keywords: *Microphone array signal processing, physics informed neural network, spherical harmonics.*

1. INTRODUCTION

The products of the spherical harmonics (SHs) and the spherical Bessel functions (or the spherical Hankel functions) form the spherical nodes [1], a complete and orthogonal function set for the Helmholtz equation, the governing parti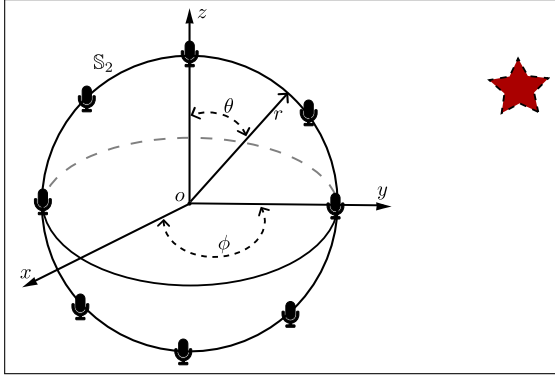al differential equation (PDE) of acoustic wave propagation. The SH decomposition of a sound field (the angular dependent SHs, the radial dependent spherical Bessel functions, and the frequency dependent sound field coefficients) greatly facilitates its analysis and manipulation [1–3]. Thus, spherical modal decomposition has become popular in many diverse spatial acoustic applications, such as spatial active noise control [4–6], beamforming [3,7,8], and direction of arrival estimation [9–11].

Due to their simplicity, open sphere microphone arrays (OSMAs) are intuitively chosen for implementing the spherical modal decomposition [7]. However, the spherical Bessel function nulls make it hard to obtain some order of the sound field coefficients at certain frequencies with an OSMA. We can mitigate this problem through arranging microphones on a rigid sphere [12], inside a spherical shell, or using vector sensors on an open sphere [3,13,14]. However, those approaches will unavoidably introduce scattering fields, request more microphones, and significantly increase the cost, respectively.

In this paper, we propose to circumvent the problem of spherical Bessel function nulls for an OSMA with the help of physics informed neural network (PINN) [15–17], a neural work which incorporates physical knowledge into its architecture and training. We model the measurement of an OSMA with a PINN, and then use it to predict the sound field on another sphere whose radius is different from that of the OSMA. Thanks to the fact that the spherical Bessel function nulls vary with radius, we can obtain the sound field coefficients which are difficult to obtain with the OSMA measurement based on the predicted sound field. The effectiveness of this approach is confirmed by simulations and compared with the rigid sphere approach.

Figure 1. A microphone array on an open sphere and some sound sources ⋆.

Figure 2. The spherical Bessel function $j_u(2\pi f r_a/s)$ as a function of frequency, $r_a = 0.05$ m, $s = 343$ m/s, $u = 0, 1, 2, 3, 4$.
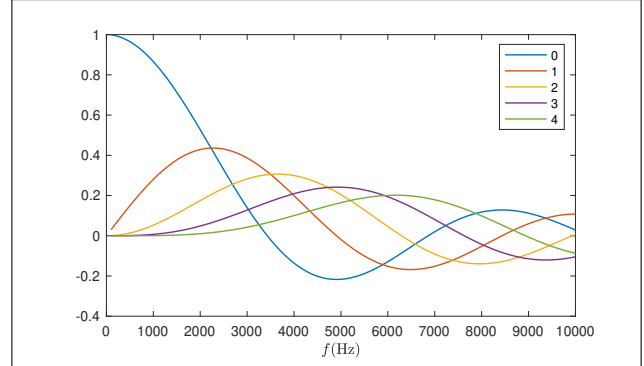
2. PROBLEM FORMULATION

We consider the set up shown in Fig. 1, where there are Q omni-directional pressure microphones on an open sphere S_2 of radius r_a and some sound sources. The Cartesian coordinates and the spherical coordinates of a point with respect to an origin are denoted as O as (x, y, z) and (r, θ, ϕ), respectively [3]. One would like to reconstruct the sound field around the sphere or locate the sound sources based on the OSMA measurement.

The tasks could be approached with SH decomposition. We decompose the sound pressure at microphone position $\{(r_a, \theta_q, \phi_q)\}_{q=1}^{Q}$ onto SHs as [1]

$$
\begin{aligned}
P(\omega, r_a, \theta_q, \phi_q) &\approx \sum_{u=0}^{U} \sum_{v=-u}^{u} \mathsf{P}_{u,v}(\omega, r_a) Y_{u,v}(\theta_q, \phi_q) \\
&= \sum_{u=0}^{U} \sum_{v=-u}^{u} \mathsf{K}_{u,v}(\omega) j_u(\omega r_a/s) \\
&\quad \times Y_{u,v}(\theta_q, \phi_q),
\end{aligned}
\tag{1}
$$

where $\omega = 2\pi f$ is the angular frequency (f is the frequency), s is the speed of sound propagation, $U = \lceil 2\pi f r_a/s \rceil$ is the up-order of the SHs that are needed to represent the sound pressure accurately [18] ($\lceil \cdot \rceil$ is the ceiling operation),. $\mathsf{P}_{u,v}(\omega, r_a)$ are the pressure field coefficients, $\mathsf{K}_{u,v}(\omega)$ are the sound field coefficients [1], $j_u(\cdot)$ is the spherical Bessel function of the first kind of order u, $Y_{u,v}(\theta, \phi)$ is the SH of order u and degree v [1] at is evaluated at (θ, ϕ).

The sound field coefficients $\mathsf{K}_{u,v}(\omega)$ characterize the sound sources and allow us to reconstruct the sound field

or to locate the sound sources [3]. To obtain the sound field coefficients, we first estimate the pressure field coefficients through [3]

$$
\hat{\mathsf{P}}_{u,v}(\omega, r_a) = \sum_{q=1}^{Q} P(\omega, r_a, \theta_q, \phi_q) Y_{u,v}(\theta_q, \phi_q) \gamma_q, \tag{2}
$$

where $\{\gamma_q\}_{q=1}^{Q}$ are the sampling weights, and then estimate the sound field coefficients through

$$
\hat{\mathsf{K}}_{u,v}(\omega) = \hat{\mathsf{P}}_{u,v}(\omega, r_a)/j_u(\omega r_a/s). \tag{3}
$$

The problem with (3) is the spherical Bessel function $j_u(\cdot)$ nulls [3,13,14]. Figure 2 presents $j_u(2\pi f r_a/s)$ with $r_a = 0.05$ m, $s = 343$ m/s, $u = 0, 1, 2, 3, 4$. We can see that $j_0(2\pi f r_a/s) = 0$ for $f = 3430$ Hz, and $j_1(2\pi f r_a/s) = 0$ for $f = 4905$ Hz. This makes it difficult to estimate the sound field coefficients of order 0, $\mathsf{K}_{0,0}(\omega)$, and order 1, $\mathsf{K}_{1,v}(\omega)$, at frequency 3430 Hz and 4905 Hz with an OSMA array of radius $r_a = 0.05$ m.

In this paper, we aim to circumvent the problem of spherical Bessel function nulls for an OSMA.

3. PINN ASSISTED OSMA

In this section, we propose a PINN method to assist an OSMA for sound field analysis. To simplify the calculation of the Laplacian, we express acoustic quantities in Cartesian coordinates.

The key idea is to exploit the fact that spherical Bessel function nulls vary with radius. The spherical Bessel

function $j_u(\cdot)$ is a function of both frequency f and radius r, and thus that $j_u(2\pi f r_b/s) \neq 0$ if $r_b \neq r_a$ and $j_u(2\pi f r_a/s) = 0$. Thus that we can obtain the sound field coefficients which are difficult to obtain with an OSMA of radius r_a based on the sound field on another sphere of radius r_b. An OSMA of radius r_a can not measure the sound field on another sphere of radius r_b directly, but we can build a PINN [15–17] to predict the sound field on the other sphere based on the measurement of the OSMA.

We build up a L layer N node (on each layer) full connected feedforward neural network [15] whose inputs are Cartesian coordinates (x, y, z) and output is the sound field estimation $\hat{P}_{\mathrm{PI}}(\omega, x, y, z)$, and update the trainable parameters of the network by minimizing the following cost function

$$L = \underbrace{\frac{1}{Q}\sum_{q=1}^{Q} \|P(\omega, x_q, y_q, z_q) - \hat{P}_{\mathrm{PI}}(\omega, x_q, y_q, z_q)\|_2^2}_{L_{\mathrm{data}}}$$
$$+ \underbrace{\frac{1}{A}\sum_{a=1}^{A} \|\frac{\nabla \hat{P}_{\mathrm{PI}}(\omega, x_a, y_a, z_a)}{(w/s)^2} + \hat{P}_{\mathrm{PI}}(\omega, x_a, y_a, z_a)\|_2^2}_{L_{\mathrm{PDE}}},$$
$$(4)$$

where $\|\cdot\|_2$ is the 2-norm, $\nabla \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplacian. The data loss L_{data} makes the network output to approximate the OSMA measurement $\{P(\omega, x_q, y_q, z_q)\}_{q=1}^{Q}$ where $(x_q, y_q, z_q)_{q=1}^{Q}$ correspond to $(r_a, \theta_q, \phi_q)_{q=1}^{Q}$. The PDE loss L_{PDE} informs the network output to conform with the Helmholtz equation on the measurement sphere of radius r_a, where $\{(x_a, y_a, z_a)\}_{a=1}^{A}$ are uniformly arranged sampling points on the sphere.

To obtain the sound field coefficients, we first train the PINN and use it to estimate the pressure $\hat{P}_{\mathrm{PI}}(\omega, x_d, y_d, z_d)$ (which are equal to $\hat{P}_{\mathrm{PI}}(\omega, r_b, \theta_d, \phi_d)$) on a sphere of radius r_b. Next we estimate the pressure field coefficients $\hat{\mathsf{P}}_{u,v}(\omega, r_b)$

$$\hat{\mathsf{P}}_{u,v}(\omega, r_b) = \sum_{d=1}^{D} \hat{P}_{\mathrm{PI}}(\omega, r_b, \theta_d, \phi_d) Y_{u,v}(\theta_d, \phi_d)\gamma_d,$$
$$(5)$$

where $\{\gamma_d\}_{d=1}^{D}$ are the sampling weights [3]. We further estimate the sound field coefficients through

$$\hat{\mathsf{K}}_{u,v}(\omega) = \hat{\mathsf{P}}_{u,v}(\omega, r_b)/j_u(\omega r_b/s). \qquad (6)$$

rcl In summary, for spatial acoustics with an OSMA, we can estimate the sound field coefficients through (3) when $j_u(kr_a) \neq 0$ and through (4), (5), (6) when $j_u(kr_a) = 0$. In this way, the problem of spherical Bessel function nulls is circumvented.

Note that the spherical Bessel function nulls is a problem under the spherical modal decomposition, but it is not a problem with the PINN. This is the fundamental fact that make the PINN assisted OSMA sound field analysis possible.

4. SIMULATION

In this section, we use a sound field reconstruction task to demonstrate the performance of the PINN assisted OSMA and compare it with the rigid sphere approach.

We consider the setup shown in Fig. 1. There is a radius $r_a = 0.05$ m OSMA with 36 uniformly arranged omni-directional pressure microphones on it. There is a sound source located at $(0.5 \text{ m}, 0.5 \text{ m}, 0.75 \text{ m})$. The sound source generates a unit amplitude sinusoidal signal at $f = 3430$ Hz. In the case, the up-order of SHs needed to represent the sound field is $U = \lceil 2\pi \times 3430 \times 0.05/343 \rceil = 4$ [18]. The transfer functions between the sound source and the microphones are simulated based on the Green's function [1]. The aim is to reconstruct the sound field on a smaller sphere of radius $r_c = 0.04$ m.

Three approaches for sound field reconstruction are considered. The first is the OSMA approach based on the spherical modal decomposition. For this approach, we estimate the sound field coefficients $\{\hat{\mathsf{K}}_{u,v}(\omega)\}_{u=1}^{4}$ through (3) and reconstruct the sound field on the smaller sphere by

$$\hat{P}_{\mathrm{SH}}(\omega, r_c, \theta, \phi) \approx \sum_{u=1}^{U}\sum_{v=-u}^{u} \hat{\mathsf{K}}_{u,v}(\omega) j_u(\omega r_c/s) Y_{u,v}(\theta, \phi),$$
$$(7)$$

because $\hat{\mathsf{K}}_{0,0}(\omega)$ is not obtainable.

The second one is the PINN assisted OSMA method. For this method, we build up a $L = 3$ layer and $N = 3$ node PINN, with the activation function being $\tanh$, and initialize the trainable parameters with the Xavier initialization [19]. PINN is trained for 10^8 epochs with a learning rate of 10^{-5} using the ADAM optimizer. The data loss L_{data} is evaluated with respect to the 36 microphone measurements, and the PDE losses L_{PDE} with respect to the Cartesian coordinates of 500 uniformly arranged sampling
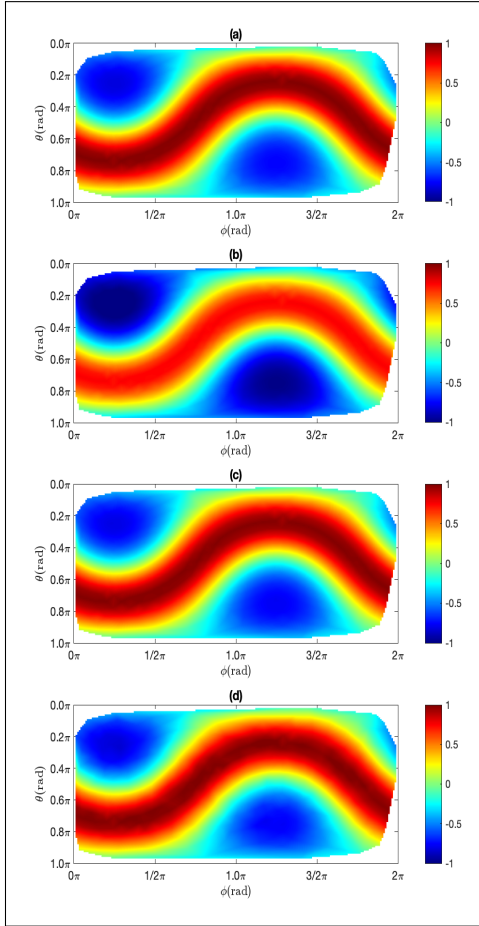
Figure 3. Sound field at $f = 3430$ Hz : (a) the ground truth, (b) OSMA reconstruction, (c) PINN assisted OSMA reconstruction, and (d) the rigid sphere reconstruction.
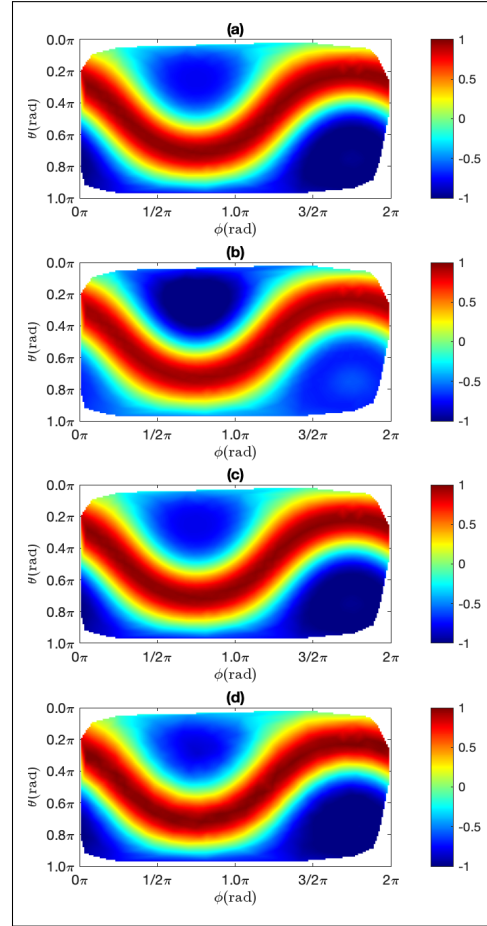
Figure 4. Sound field at $f = 4905$ Hz : (a) the ground truth, (b) OSMA reconstruction, (c) PINN assisted OSMA reconstruction, and (d) the rigid sphere reconstruction.

points on the sphere of radius $r_a = 0.05$ m. We first estimate the sound field coefficients $\{\hat{K}_{u,v}(\omega)\}_{u=1}^{4}$ through (3), $\hat{K}_{0,0}(\omega)$ through (4), (5) with (6) with $r_b = 0.048$ m, and next reconstruct the sound field on the smaller sphere similar as (1) but with $\hat{K}_{0,0}(\omega)j_u(\omega r_c/s)Y_{0,0}(\theta,\phi)$ included.

The third one is the rigid sphere approach. The OSMA in Fig. 1 is replaced with a rigid sphere of the same radius, and the rest of simulation setting is the same. we
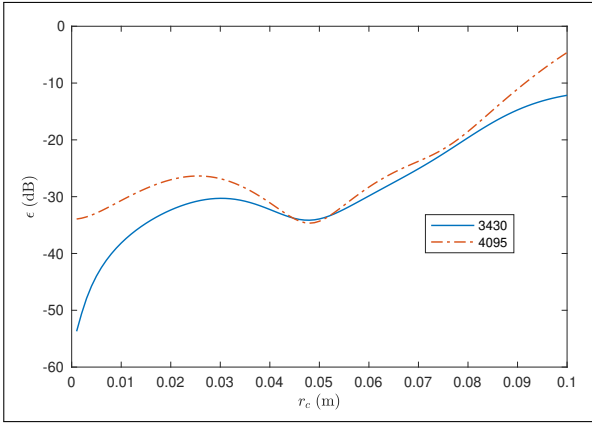
reconstruct the sound field on the sphere of radius r_c as

$$P(\omega, r_c, \theta, \phi) \approx \sum_{u=0}^{U} G_u(\omega, r_c, r_a)$$
$$\times \sum_{v=-u}^{u} \hat{P}_{u,v}(\omega, r_a) Y_{u,v}(\theta, \phi), \tag{8}$$

where the pressure field coefficients $\hat{P}_{u,v}(\omega, r_a)$ are obtained similar to (2), $G_u(\omega, r_c, r_a)$ is the radial transla-
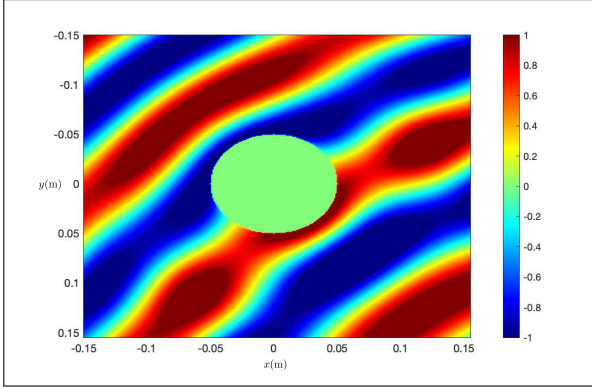
Figure 5. The scattering field around a rigid sphere at 3430 Hz.

Figure 6. Sound field reconstruction error ϵ of the pure PINN method as a function of the reconstruction sphere radius r at 3430 Hz and 4095 Hz.

tor [3]

$$
G_u(\omega, r_c, r_a) = \\
\frac{h'_u(\omega r_a/s) j_u(\omega r_c/s)}{j_u(\omega r_a/s) h'_u(\omega r_a/s) - j'_u(\omega r_a/s) h_n(\omega r_a/s)}, \quad (9)
$$

$j_u(\cdot)$ and $h_u(\cdot)$ are the spherical Bessel function of the first kind and the spherical Hankel function of the second kind, respectively, and $j'_u(\cdot)$ and $h'_u(\cdot)$ are corresponding derivatives with respect to argument.

We denote the reconstruction error as

$$
\epsilon = \frac{\sum_{d=1}^{100} ||P(\omega, r_c, \theta_d, \phi_d) - \hat{P}(\omega, r_c, \theta_d, \phi_d)||_2^2}{\sum_{d=1}^{100} ||P(\omega, r_c, \theta_d, \phi_d)||_2^2}, \quad (10)
$$

where $P(\omega, r_c, \theta_d, \phi_d)$ and $\hat{P}(\omega, r_c, \theta_d, \phi_d)$ are the true pressure and its reconstruction at 100 uniformly selected sampling positions $(\theta_d, \phi_d)_{d=1}^{100}$.

Real part of the ground truth and its reconstructions by three methods is shown in Fig. 3. Comparing Fig. 3 (b) and (a), we can see that the sound field component $K_{0,0}(\omega) j_u(\omega r_c/s) Y_{0,0}(\theta, \phi)$ missing the OSMA approach is unable to accurately reconstruct the ground truth. Comparing Fig. 3 (c), (d) and (a), we can see that with the PINN assisted OSMA approach and the rigid sphere approach are able to accurately reconstruct the sound field. The reconstruction errors of the OSMA approach, the PINN assisted OSMA approach, and the rigid sphere approach are -8.5 dB, -28.4 dB and -29.3 dB, respectively. The simulation results of three approaches for reconstructing the imaginary part of the ground truth are similar to Fig. 3, and thus are not shown for brevity.

The simulations is repeated at $f = 4095$ Hz. We arrange the sound source at $(0.5 \text{ m}, -0.5 \text{ m}, -0.75 \text{ m})$ and the rest of simulation settings are the same as the $f = 3430$ Hz case. Real part of the ground truth and its reconstructions by three methods is shown in Fig. 4. The reconstruction errors of the OSMA approach, the PINN assisted OSMA approach, and the rigid sphere approach are -10.2 dB, -31.6 dB and -32.3 dB, respectively.

The PINN-assisted OSMA approach performs comparably to the rigid sphere approach. Nonetheless, the rigid sphere approach has a drawback: the scattering effect. In Fig. 5, we present the scattering field around the rigid sphere at 3430 Hz. For nearfield applications, where the rigid sphere is placed close to some object, the scattering field will further undergo multiple scatterings and is highly undesirable [20]. The PINN assisted OSMA approach, on the other hand, does not suffer from the scattering problems.

Figure 6 presents the sound field reconstruction error ϵ of a pure PINN method as a function of reconstruction sphere radius r_c at 3430 Hz and 4095 Hz. In this case, the sound field is reconstructed based on the PINN prediction directly and only, and does not go through the SH decomposition process. From Fig. 6, we can see that the reconstruction error ϵ is small when the reconstruction sphere radius is close to the array radius, or $r_c \approx r_a = 0.05$ m, and increases when they are not close $r_c \not\approx r_a = 0.05$ m. It is interesting that the reconstruction error ϵ decreases again when the reconstruction sphere radius is small, or $r_c < 0.025$ m. From a spherical modal decomposition point of view, this can be explained by the fact that when the reconstruction sphere radius is small less number of

10th Convention of the European Acoustics Association
Turin, Italy • 11th – 15th September 2023 • Politecnico di Torino
2173

SH coefficients are needed to describe the sound field.

5. CONCLUSION

In this paper, we proposed to assist an OSMA for sound field analysis with the PINN. Under SH decomposition, the OSMA suffers from the spherical Bessel function nulls and is unable to obtain some orders of sound field coefficients at certain frequencies. We use a PINN to predict the sound field on a sphere whose radius is different from that of the OSMA, and obtain those order of sound field coefficients based on the PINN prediction. The performance of this approach is comparable with the rigid sphere approach and does not introduce the scattering field.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] E. G. Williams and J. A. Mann III, "Fourier acoustics: sound radiation and nearfield acoustical holography," 2000.

[2] R. A. Kennedy, D. B. Ward, and T. D. Abhayapala, "Nearfield beamforming using radial reciprocity," *IEEE Transactions on Signal Processing*, vol. 47, no. 1, pp. 33–40, 1999.

[3] B. Rafaely, *Fundamentals of spherical array processing*, vol. 8. Springer, 2015.

[4] J. Zhang, T. D. Abhayapala, W. Zhang, P. N. Samarasinghe, and S. Jiang, "Active noise control over space: A wave domain approach," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 4, pp. 774–786, 2018.

[5] F. Ma, W. Zhang, and T. D. Abhayapala, "Active control of outgoing noise fields in rooms," *The Journal of the Acoustical Society of America*, vol. 144, no. 3, pp. 1589–1599, 2018.

[6] F. Ma, W. Zhang, and T. D. Abhayapala, "Active control of outgoing broadband noise fields in rooms," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 529–539, 2020.

[7] T. D. Abhayapala and D. B. Ward, "Theory and design of high order sound field microphones using spherical microphone array," in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. II–1949–II–1952, 2002.

[8] G. Huang, J. Chen, and J. Benesty, "Insights into frequency-invariant beamforming with concentric circular microphone arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 12, pp. 2305–2318, 2018.

[9] A. H. Moore, C. Evers, and P. A. Naylor, "Direction of arrival estimation in the spherical harmonic domain using subspace pseudointensity vectors," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 1, pp. 178–192, 2016.

[10] S. Hafezi, A. H. Moore, and P. A. Naylor, "Augmented intensity vectors for direction of arrival estimation in the spherical harmonic domain," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 10, pp. 1956–1968, 2017.

[11] B. Jo and J.-W. Choi, "Parametric direction-of-arrival estimation with three recurrence relations of spherical harmonics," *The Journal of the Acoustical Society of America*, vol. 145, no. 1, pp. 480–488, 2019.

[12] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. II–1781–II–1784, 2002.

[13] F. Ma, W. Zhang, and T. D. Abhayapala, "Real-time separation of non-stationary sound fields on spheres," *The Journal of the Acoustical Society of America*, vol. 146, no. 1, pp. 11–21, 2019.

[14] G. Huang, J. Chen, and J. Benesty, "A flexible high directivity beamformer with spherical microphone arrays," *The Journal of the Acoustical Society of America*, vol. 143, no. 5, pp. 3024–3035, 2018.

[15] 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.

[16] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli, "Scientific machine learning through physics–informed neural networks: where we are and what's next," *Journal of Scientific Computing*, vol. 92, no. 3, p. 88, 2022.

[17] 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.

[18] T. D. Abhayapala and A. Gupta, "Higher order differential-integral microphone arrays," *The Journal of the Acoustical Society of America*, vol. 127, no. 5, pp. EL227–EL233, 2010.

[19] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pp. 249–256, JMLR Workshop and Conference Proceedings, 2010.

[20] D. L. Colton, R. Kress, and R. Kress, *Inverse acoustic and electromagnetic scattering theory*, vol. 93. Springer, 1998.