Sound Field Reconstruction Using Physics-Informed Boundary Integral Networks

Stefano Damiano, Toon van Waterschoot
Dept. of Electrical Engineering (ESAT-STADIUS), KU Leuven, Leuven, Belgium
{stefano.damiano, toon.vanwaterschoot}@esat.kuleuven.be

***Abstract*—Sound field reconstruction refers to the problem of estimating the acoustic pressure field over an arbitrary region of space, using only a limited set of measurements. Physics-informed neural networks have been adopted to solve the problem by incorporating in the training loss function the governing partial differential equation, either the Helmholtz or the wave equation. In this work, we introduce a boundary integral network for sound field reconstruction. Relying on the Kirchhoff-Helmholtz boundary integral equation to model the sound field in a given region of space, we employ a shallow neural network to retrieve the pressure distribution on the boundary of the considered domain, enabling to accurately retrieve the acoustic pressure inside of it. Assuming the positions of measurement microphones are known, we train the model by minimizing the mean squared error between the estimated and measured pressure at those locations. Experimental results indicate that the proposed model outperforms existing physics-informed data-driven techniques.**

***Index Terms*—sound field reconstruction, boundary integral networks, boundary integral equation, physics-informed neural networks, Helmholtz equation**

I. Introduction

Immersive audio technologies aiming to create a realistic auditory experience in virtual and augmented reality demand a precise spatial representation of acoustic environments, in order to enhance the user engagement [1]. Obtaining direct measurements in a large space is a demanding process [2], that can rarely be pursued in practice. As a solution, sound field reconstruction aims at estimating the acoustic pressure in areas where direct measurements are not available, given a limited set of microphone recordings.

The problem has been addressed using model-based and data-driven methods. Model-based solutions either rely on a parametric representation of the acoustic scene [3] or describe the sound field using closed-form solutions to the wave equation [4], including plane waves [5], spherical waves [6], and equivalent sources [7]–[9]. These models, while robust and interpretable, are based on assumptions and specific propagation models, that hinder their capacity to capture the complexity of the sound field over a considerably large area inside a room. Data-driven techniques have thus been explored to learn representations of the sound field from extensive datasets. Deep neural networks have proven effective in learning implicit representations of the acoustic field [10], [11], achieving good performance in sound field reconstruction [12], [13]. Moreover, dictionary learning techniques have been investigated to represent the audio signal as a combination of functions (called dictionary atoms) learned from data [14]. As a drawback, both types of data-driven methods necessitate substantial data for training and lack interpretability.

To bridge the gap between model-based and data-driven methods, physics-informed machine learning approaches have been proposed, exploiting prior knowledge of the physical laws that govern sound propagation in the learning process [15]. Physics-informed neural networks (PINNs) integrate in the training objective the underlying partial differential equation (PDE), either the wave equation, in the time domain [16]–[19], or the Helmholtz equation, in the frequency domain [20]–[22]. This is achieved by introducing in the loss function a term that penalizes the deviation of the estimated sound field from the underlying PDE, evaluated at a dense set of collocation points in the target area. As a result, these networks are trained using solely a sparse set of observations. Similarly, in [23] a zero-shot physics-informed dictionary learning technique, not requiring pre-training data, has been introduced.

Recently, physics-informed boundary integral networks (PIBI-Nets) have been proposed to solve PDEs [24]–[27]. These models are derived from the boundary element method (BEM) [28]: rather than directly exploiting PDEs, in fact, the associated *boundary integral equation* (BIE) is used to guide the training. It has been shown that PIBI-Nets outperform PINNs in the solution of Laplace and Poisson problems [27].

In this work, we investigate PIBI-Nets for sound field reconstruction. According to the Green's theorem, the sound field in a closed area can be expressed, using the Kirchhoff-Helmholtz BIE, as a function of the pressure density on its boundary [28], [29]. While the BEM relies on the solution of a linear system, the proposed method, inspired by [27], employs a multi-layer perceptron (MLP) to learn the pressure density at a discrete set of integration points on the boundary of the reconstruction region. The boundary density is then used to estimate, by means of the Kirchhoff-Helmholtz BIE, the sound field inside the region.

While PINNs require a dense set of collocation points over the reconstruction region, the proposed solution relies on integration points only on the boundary: the dimensionality and

This project has received funding from KU Leuven internal funds C3/23/056 and from FWO Research Project G0A0424N. The resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation - Flanders (FWO) and the Flemish Government.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on January 05,2026 at 13:27:29 UTC from IEEE Xplore. Restrictions apply.

complexity of the problem are thus reduced, facilitating training. Experimental results indicate that the proposed method outperforms both the physics-informed dictionary learning approach [23] and a traditional PINN, is more robust to a decreasing number of available measurements, and requires fewer integration points compared to the collocation points required by the PINN. Finally, the proposed method does not entail prior knowledge on the source position, room geometry or boundary conditions, and produces an interpretable solution.

II. Problem Formulation and Background

A. Problem formulation

Let us consider a reverberant room containing one or more acoustic sources, and a two-dimensional, source-free region of space $\Omega \in \mathbb{R}^2$ lying on a horizontal plane inside the room. We can express by $p(\mathbf{r}, \omega) \in \mathbb{C}$ the complex-valued acoustic pressure in the frequency domain, at position $\mathbf{r} \in \Omega$ and angular frequency $\omega = 2\pi f$, where f is the temporal frequency. The frequency-domain sound field satisfies the homogeneous Helmholtz equation [29]

$$(\nabla^2 + k^2) p(\mathbf{r}, \omega) = 0\,, \tag{1}$$

where ∇^2 denotes the Laplace operator and $k = \omega/c$, with c being the speed of sound in air. We sample the acoustic pressure in Ω using M microphones, whose positions are indicated by $\{\mathbf{r}_m\}_{m=1}^M$. The sound field measured at the mth microphone can be expressed as

$$s(\mathbf{r}_m, \omega) = p(\mathbf{r}_m, \omega) + e_m\,, \tag{2}$$

where e_m denotes additive measurement noise. We can collect the M microphone measurements in the vector $\mathbf{s} \in \mathbb{C}^M$,

$$\mathbf{s} = [s(\mathbf{r}_1, \omega), s(\mathbf{r}_2, \omega), \ldots, s(\mathbf{r}_M, \omega)]^\top\,. \tag{3}$$

Our goal is to retrieve the pressure field at arbitrary positions within Ω, given the available measurements $\mathbf{s}$. For the sake of readability, we omit the dependency on the frequency ω in the rest of the paper.

B. Background

Traditional data-driven methods for sound field reconstruction rely on learning from data a function $f(\mathbf{r}, \boldsymbol{\theta})$, parametrized by the vector $\boldsymbol{\theta}$, that maps spatial coordinates in Ω to the corresponding pressure value. This function can be parametrized using a neural network trained to minimize a data-fidelity loss function, expressing how much the estimated pressure at the positions of the available microphones $\{\mathbf{r}_m\}_{m=1}^M$ deviates from the actual measurements $\mathbf{s}$. The problem can be framed as

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \mathcal{L}_{\text{data}}\left(f(\{\mathbf{r}_m\}_{m=1}^M, \boldsymbol{\theta}), \mathbf{s}\right)\,, \tag{4}$$

where the loss function $\mathcal{L}_{\text{data}}$ usually takes the form of a mean squared error

$$\mathcal{L}_{\text{data}}\left(f(\{\mathbf{r}_m\}_{m=1}^M, \boldsymbol{\theta}), \mathbf{s}\right) = \sum_{m=1}^{M} \left|s(\mathbf{r}_m) - f(\mathbf{r}_m, \boldsymbol{\theta})\right|^2\,. \tag{5}$$

To enable f to effectively generalize to positions where microphones are not available (i.e., to prevent overfitting), regularization techniques are adopted in the literature. To encourage f to produce a physically meaningful sound field, PINNs introduce an additional term in the loss function, penalizing solutions that deviate from the underlying PDE. This PDE loss term can be obtained from (1) as

$$\mathcal{L}_{\text{PDE}} = \sum_{n=1}^{N} \left|(\nabla^2 + k^2) f(\mathbf{r}_n, \boldsymbol{\theta})\right|^2\,, \tag{6}$$

where the PDE is evaluated at a set of N points, usually called collocation points, that can be arbitrarily picked within Ω, independently from the M sampling points. The loss function used to train the neural network becomes

$$\mathcal{L}_{\text{PINN}} = \mathcal{L}_{\text{data}} + \lambda \mathcal{L}_{\text{PDE}}\,, \tag{7}$$

where the positive constant λ represents a weighting factor between the two terms, constituting a hyperparameter of the model. The advantage of PINNs is that the measurements $\mathbf{s}$ constitute the only data used to train the network, significantly reducing the need for labeled data, that represents a major drawback of traditional data-driven techniques.

III. Proposed Method

Boundary integral networks aim at solving the PDE exploiting the boundary integral equation. Let us denote the boundary of Ω as $\partial\Omega$ and define the boundary density function $h(\mathbf{y})$, representing the pressure at some points $\mathbf{y}$ lying on $\partial\Omega$. An explicit integral representation of the solution of the homogeneous Helmholtz equation can then be written as [29]

$$p(\mathbf{r}) = \int_{\partial\Omega} \left[G(\mathbf{y}, \mathbf{r}) \frac{\partial h}{\partial \mathbf{n}_\mathbf{y}}(\mathbf{y}) - \frac{\partial G}{\partial \mathbf{n}_\mathbf{y}}(\mathbf{y}, \mathbf{r}) h(\mathbf{y})\right] d\sigma_\mathbf{y}\,, \tag{8}$$

where $\mathbf{n}_\mathbf{y}$ denotes the normal to $\partial\Omega$ at point $\mathbf{y}$ and $G(\mathbf{y}, \mathbf{r})$ is the fundamental solution of the Helmholtz equation. This solution is represented by the Green's function, that defines the transfer function between a source located at position $\mathbf{y}$ and a receiver at position $\mathbf{r}$, and takes the form [29]

$$G(\mathbf{y}, \mathbf{r}) = \begin{cases} -\frac{j}{4} H_0^{(1)}(k\|\mathbf{r} - \mathbf{y}\|_2) & \text{2D case} \\ \frac{1}{4\pi\|\mathbf{r}-\mathbf{y}\|_2} e^{jk\|\mathbf{r}-\mathbf{y}\|_2} & \text{3D case}\,. \end{cases} \tag{9}$$

Here, j denotes the imaginary unit, $\|\cdot\|_2$ is the ℓ_2 norm and $H_0^{(1)}$ is the zeroth-order Hankel function of the first kind. Note that (8) relates the sound field at arbitrary points *inside* Ω to the pressure at boundary points on $\partial\Omega$. Given that Green's functions and their derivatives are available in closed-form, and that $\partial\Omega$ is known (in fact, Ω is arbitrarily defined), inspired by [27] we adopt a neural network to estimate the unknown boundary pressure distribution $h(\mathbf{y})$. The network follows the same architecture presented in [27], with 2 fully-connected layers of 64 neurons each and a single-neuron output layer. The network takes as input the spatial coordinate $\mathbf{r}$ of a position within Ω and outputs the value $h(\mathbf{y})$, then used to compute the estimated $\hat{p}(\mathbf{r})$ using (8). As the pressure is complex-valued, its real and imaginary parts are treated

Authorized licensed use limited to: KU Leuven University Library. Downloaded on January 05,2026 at 13:27:29 UTC from IEEE Xplore. Restrictions apply.

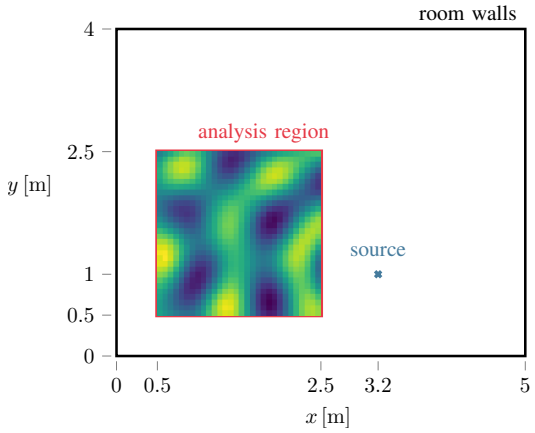


Fig. 1. Geometry of the simulated room considered in the experiments.

separately and then combined. To compute the integral, a set of N_{int} integration points are selected on $\partial\Omega$ and (8) is discretized similarly to [27] to compute the pressure within Ω. To train the neural network, the sound field at the positions $\{\mathbf{r}_m\}_{m=1}^{M}$ of the M microphones is estimated using PIBI-Net, and the deviation from the measured pressure is minimized using a mean squared loss function of the form

$$\mathcal{L}_{\text{PIBI}} = \sum_{m=1}^{M} \left| s(\mathbf{r}_m) - \hat{p}(\mathbf{r}_m) \right|^2 , \tag{10}$$

where $\hat{p}(\mathbf{r}_m)$ indicates the pressure estimated using PIBI-Net as in (8). Note that, differently from the PINN loss (7), $\mathcal{L}_{\text{PIBI}}$ has a unique term and does not contain any hyperparameter: thus, PIBI-Net has one hyperparameter less than PINN.

The boundary integral network comes with two main advantages over PINNs. First, as the predicted sound field is computed using (8), it always satisfies the Helmholtz equation: thus, the proposed model is interpretable and produces physically meaningful solutions. Note also that the neural network output $h(\mathbf{y})$ is constrained to be in the solution space of (8) which, as discussed in [24], reduces the complexity of the problem and leads to an improved convergence. Second, while PINNs necessitate collocation points across the entire domain Ω, PIBI-Nets require integration points only on its boundary $\partial\Omega$, which reduces by 1 the dimensionality of the problem. Consequently, as will be shown in Sec. IV, a PIBI-Net achieves better performance with a lower M and with fewer integration points than the collocation points needed by a PINN.

IV. Evaluation

To assess the performance of the proposed method, we implement it in Python[1] and run an experimental campaign on simulated data. Using pyroomacoustics [30], we create a 2D rectangular room with size $[5, 4]\,\text{m}$, having $T_{60} = 0.4\,\text{s}$ and containing a single, omnidirectional source located at position $[3.2, 1]\,\text{m}$. We choose our analysis region Ω to be a square with $2\,\text{m}$ side, with its bottom right corner at position $[0.5, 0.5]\,\text{m}$. In this area, we define a grid of 30×30 equally spaced points, simulate room impulse responses (RIRs) between the source and these points, with a sampling frequency $f_s = 16\,\text{kHz}$, and transform them to the frequency domain using the fast Fourier transform (FFT) with 1024 frequency bins. The simulated scenario is illustrated in Fig. 1. Note that this 2D geometry corresponds to a shoebox room having totally absorbing ceiling and floor, with the source and all microphones located on the same horizontal plane.

To evaluate the proposed method, we use two different baselines: the physics-informed dictionary learning method [23] (PIDL) and a physics-informed neural network (PINN) having the same architecture of the proposed method (PIBI), described in Sec. III. The PIDL can be directly applied to the grid defined in the reconstruction area, and its hyperparameters are set as in [23]. We implement the PINN and PIBI networks in Pytorch [31] and train them for $5\,000$ steps using the Adam optimizer, a learning rate of 0.001, and the loss functions defined in (7), with manually tuned $\lambda = 0.001$, and (10), respectively. Both architectures have $8\,577$ parameters.

As we consider a source-free reconstruction region, we adopt the homogeneous Helmholtz PDE (1) for PINN, and its boundary integral representation (8) for PIBI. When training the PINN, the partial derivatives of the PDE are computed using automatic differentiation, available in Pytorch.

We adopt two different evaluation metrics: the normalized mean squared error (NMSE) in dB scale

$$\text{NMSE} = 10 \log_{10} \left(\frac{\|\mathbf{p} - \hat{\mathbf{p}}\|_2^2}{\|\mathbf{p}\|_2^2} \right) , \tag{11}$$

where $\mathbf{p}$ contains the measured pressure and $\hat{\mathbf{p}}$ the predicted pressure at the 30×30 grid points, and the normalized cross-correlation (NCC)

$$\text{NCC} = \frac{|\hat{\mathbf{p}}^{\text{H}} \mathbf{p}|}{\|\hat{\mathbf{p}}\|_2^2 \|\mathbf{p}\|_2^2} , \tag{12}$$

where $(\cdot)^{\text{H}}$ indicates the Hermitian transpose. The NCC measures the similarity between the predicted and ground truth pressure, expressed by a value in the range $[-1, 1]$, where 1, 0, and -1 correspond to perfect similarity, no correlation, and perfect anti-correlation, respectively [23].

We compare the performance of PIDL, PINN and PIBI for varying frequency, number of microphones M, and number of collocation/integration points. We perform all experiments using a 10-fold procedure, where we randomly select different measurement microphones in each fold, and report the average and standard deviation of the results. In all experiments, the real and imaginary parts are reconstructed separately and combined to obtain the total reconstructed pressure. Following [27], we first choose $N_{\text{coll}} = 200$ collocation points in the reconstruction region for PINN, randomly selected at each iteration, and $N_{\text{int}} = 200$ equally spaced integration points on the boundary for PIBI. We randomly pick $M = 50$ microphones from the defined grid and analyze the reconstruction performance across frequencies ranging from $50\,\text{Hz}$ to $1050\,\text{Hz}$. In Fig. 2a and Fig. 3a we report, respectively,

[1]Code available: `https://github.com/steDamiano/pibi-sfr`

Authorized licensed use limited to: KU Leuven University Library. Downloaded on January 05,2026 at 13:27:29 UTC from IEEE Xplore. Restrictions apply.

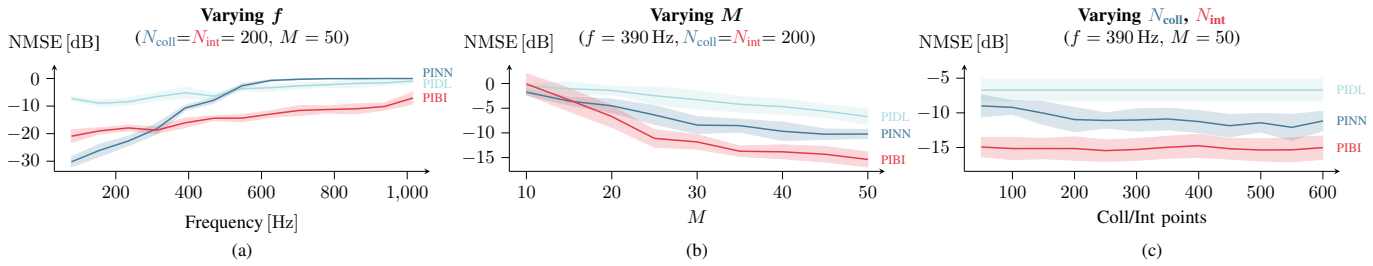


Fig. 2. Mean (solid line) and standard deviation (shaded region) of the NMSE scores, obtained with the baselines PIDL and PINN and the proposed method PIBI when varying: (a) the frequency; (b) the number of measurement microphones M; (c) the number of collocation (PINN) or integration (PIBI) points.

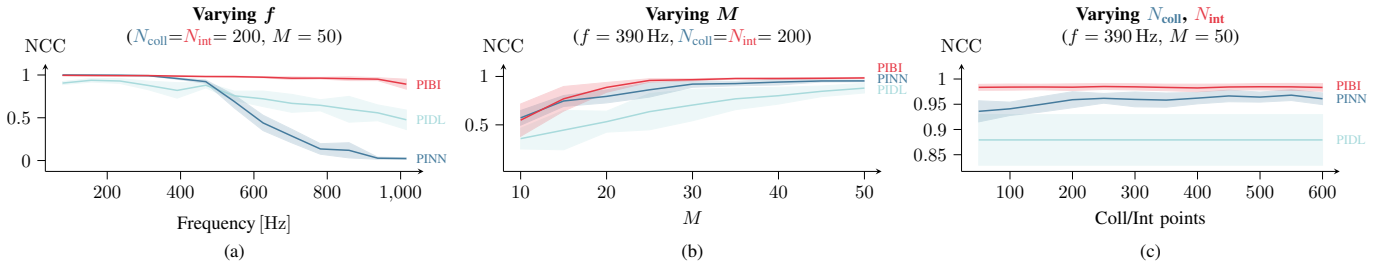


Fig. 3. Mean (solid line) and standard deviation (shaded region) of the NCC scores, obtained with the baselines PIDL and PINN and the proposed method PIBI when varying: (a) the frequency; (b) the number of measurement microphones M; (c) the number of collocation (PINN) or integration (PIBI) points.

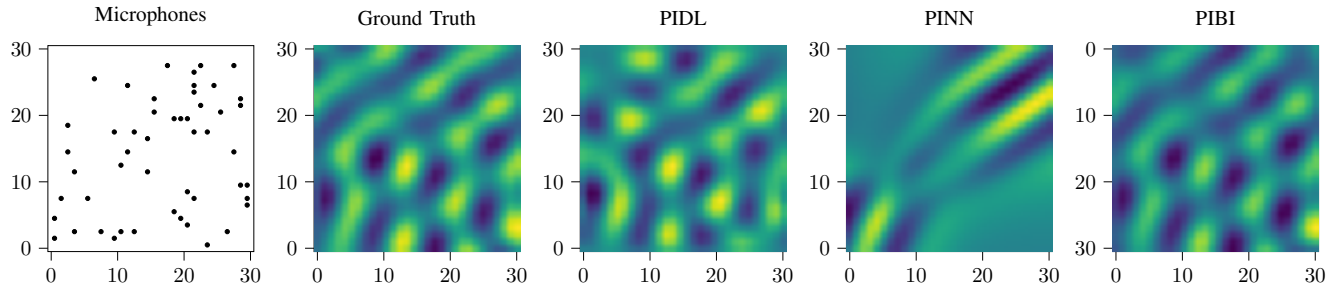


Fig. 4. Real part of the sound field at $f = 625$ Hz on the 30×30 point grid, obtained with the PIDL and PINN baselines and the proposed method PIBI, and compared to the ground truth. The problem is solved using $M = 50$ measurements with microphones randomly arranged as in the leftmost figure.

the NMSE and NCC scores for the three models. All models show a degradation of the performance as frequency increases: PIBI consistently outperforms PIDL at all frequencies on both metrics, and significantly outperforms PINN for frequencies higher than 350 Hz. In particular, the NCC score of PIBI is close to 1 (i.e., perfect similarity) across the entire range, whereas for PINN it gets close to zero (i.e., no correlation) for frequencies higher than 800 Hz. In Fig. 4, we show the real part of the sound field at 625 Hz, reconstructed in Ω by the three models and compared with the ground truth for $M = 50$.

We then fix the frequency to 390 Hz and set $N_{\text{coll}} = N_{\text{int}} = 200$, and progressively reduce the number of measurement microphones; we report the performance of the three models in Fig. 2b and Fig. 3b. All methods exhibit a degradation of the performance as the number of available measurements decreases. Nonetheless, PIBI is more robust than the two baselines to a decreasing number of available microphones, outperforming them on both metrics for $M > 15$ microphones. Note that PINN achieves its best NMSE with $M = 45$: to reach the same score, PIBI requires only 25 microphones instead. This indicates that the proposed PIBI method is more data-efficient than traditional PINNs. As discussed in Sec. III, whereas PINN relies on collocation points in the reconstruction region Ω, PIBI only requires integration points on the boundary: this results in a reduction of the dimensionality of the problem, facilitating the convergence of the network and increasing its robustness to a lower number of microphones.

We finally fix the frequency to 390 Hz and the number of microphones to $M = 50$, and vary the number of collocation and integration points for PINN and PIBI, respectively. We report the results in Fig. 2c and Fig. 3c. The performance of PIDL is also reported for reference (horizontal line): as PIDL does not require collocation or integration points, its performance is constant in this experiment. Both PINN and PIBI outperform PIDL for all considered cases. Moreover, the figures show that, for each considered number of collocation (integration) points, PIBI outperforms PINN on both metrics. Once again, PIBI relies on points lying on the boundary of

Authorized licensed use limited to: KU Leuven University Library. Downloaded on January 05,2026 at 13:27:29 UTC from IEEE Xplore. Restrictions apply.

Ω. Thus, for a given choice of $N_{\text{coll}} = N_{\text{int}}$, the density of integration points used in PIBI is higher than that of collocation points used in PINN, leading to a better performance. In other words, PIBI demands less integration points than PINN requires collocation points, resulting in a complexity gain.

These results indicate the potential of the proposed approach for sound field reconstruction. Imposing a physical constraint on the boundary of the target domain, rather than on points inside of it, helps to reduce the complexity of the problem and leads to improved performance and data-efficiency over PINNs. Moreover, the proposed approach has the advantage of yielding physically interpretable results that always represent solutions to the Kirchhoff-Helmholtz BIE.

V. Conclusions

In this paper, we proposed a sound field reconstruction method based on a boundary-informed neural network. By describing the sound field using the Kirchhoff-Helmholtz boundary integral equation, we adopted a multi-layer perceptron to retrieve the boundary pressure distribution, then employed to reconstruct the sound field at arbitrary positions within the analysis region. Simulation results indicate that the proposed approach outperforms both physics-informed dictionary learning and traditional PINNs on a broad frequency range and is more robust to scarcity of available measurements. Moreover, it demands fewer integration points compared to the collocation points required by PINNs, leading to a reduced problem complexity. Future work will focus on the analysis of regions containing sources (i.e., where the inhomogeneous Helmholtz equation holds) and on the introduction of explicit boundary conditions in the model. Finally, tests on real data will be carried out.

References

[1] F. Zotter, *Ambisonics: A practical 3D audio theory for recording, studio production, sound reinforcement, and virtual reality*. Cham: Springer Nature, 2019.

[2] T. Dietzen, R. Ali, M. Taseska, and T. van Waterschoot, "MYRiAD: a multi-array room acoustic database," *EURASIP J. Audio Speech Music Process.*, vol. 2023, Apr. 2023.

[3] M. Pezzoli, F. Borra, F. Antonacci, S. Tubaro, and A. Sarti, "A Parametric Approach to Virtual Miking for Sources of Arbitrary Directivity," *IEEE/ACM Trans. Audio Speech Language Process.*, vol. 28, 2020.

[4] N. Ueno, S. Koyama, and H. Saruwatari, "Kernel ridge regression with constraint of Helmholtz equation for sound field interpolation," in *Proc. 2018 Int. Workshop Acoust. Signal Enhancement (IWAENC)*, 2018.

[5] W. Jin and W. B. Kleijn, "Theory and design of multizone soundfield reproduction using sparse methods," *IEEE/ACM Trans. Audio, Speech Language Process.*, vol. 23, no. 12, 2015.

[6] F. Borra, I. D. Gebru, and D. Markovic, "Soundfield Reconstruction in Reverberant Environments Using Higher-order Microphones and Impulse Response Measurements," in *Proc. 2019 IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, pp. 281–285, May 2019.

[7] N. Antonello, E. De Sena, M. Moonen, P. A. Naylor, and T. van Waterschoot, "Room Impulse Response Interpolation Using a Sparse Spatio-Temporal Representation of the Sound Field," *IEEE/ACM Trans. Audio Speech Language Process.*, vol. 25, Oct. 2017.

[8] S. Koyama and L. Daudet, "Sparse Representation of a Spatial Sound Field in a Reverberant Environment," *IEEE J. Selected Topics Signal Process.*, vol. 13, pp. 172–184, Mar. 2019.

[9] S. Damiano, F. Borra, A. Bernardini, F. Antonacci, and A. Sarti, "A Compressive Sensing Approach for the Reconstruction of the Soundfield Produced by Directive Sources in Reverberant Rooms," *IEEE/ACM Trans. Audio Speech Language Process.*, vol. 32, 2024.

[10] K. Su, M. Chen, and E. Shlizerman, "INRAS: Implicit Neural Representation for Audio Scenes," in *Adv. Neur. Inf. Process. Syst.*, (New Orleans, USA), 2022.

[11] A. Luo and Y. Du, "Learning Neural Acoustic Fields," in *Proc. 36th Int. Conf. Neur. Inf. Process. Syst. (NeurIPS)*, (New Orleans, USA), 2022.

[12] E. Fernandez-Grande, X. Karakonstantis, D. Caviedes-Nozal, and P. Gerstoft, "Generative models for sound field reconstruction," *J. Acoust. Soc. Amer.*, vol. 153, pp. 1179–1190, Feb. 2023.

[13] F. Miotello, L. Comanducci, M. Pezzoli, A. Bernardini, F. Antonacci, and A. Sarti, "Reconstruction of sound field through diffusion models," in *Proc. 2024 IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, pp. 1476–1480, 2024.

[14] M. Hahmann, S. A. Verburg, and E. Fernandez-Grande, "Spatial reconstruction of sound fields using local and data-driven functions," *J. Acoust. Soc. Amer.*, vol. 150, pp. 4417–4428, Dec. 2021.

[15] S. Koyama, J. G. Ribeiro, T. Nakamura, N. Ueno, and M. Pezzoli, "Physics-informed machine learning for sound field estimation: Fundamentals, state of the art, and challenges," *IEEE Signal Process. Mag.*, vol. 41, no. 6, pp. 60–71, 2025.

[16] 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. 10th Conv. Europ. Acoust. Association Forum Acusticum 2023*, Jan. 2024.

[17] D. Sundström, S. Koyama, and A. Jakobsson, "Sound field estimation using deep kernel learning regularized by the wave equation," in *Proc. 2024 Int. Workshop Acoust. Signal Enhancement (IWAENC)*, 2024.

[18] M. Olivieri, X. Karakonstantis, M. Pezzoli, F. Antonacci, A. Sarti, and E. Fernandez-Grande, "Physics-informed neural network for volumetric sound field reconstruction of speech signals," *EURASIP J. Audio Speech Music Process.*, vol. 2024, no. 1, p. 42, 2024.

[19] X. Karakonstantis, D. Caviedes-Nozal, A. Richard, and E. Fernandez-Grande, "Room impulse response reconstruction with physics-informed deep learning," *J. Acoust. Soc. Amer.*, vol. 155, Feb. 2024.

[20] K. Shigemi, S. Koyama, T. Nakamura, and H. Saruwatari, "Physics-informed convolutional neural network with bicubic spline interpolation for sound field estimation," in *Proc. 2022 Int. Workshop Acoust. Signal Enhancement (IWAENC)*, 2022.

[21] J. G. Ribeiro, S. Koyama, R. Horiuchi, and H. Saruwatari, "Sound field estimation based on physics-constrained kernel interpolation adapted to environment," *IEEE/ACM Trans. Audio Speech Language Process.*, vol. 32, 2024.

[22] F. Ma, S. Zhao, and I. S. Burnett, "Sound field reconstruction using a compact acoustics-informed neural network," *J. Acoust. Soc. Amer.*, vol. 156, Sept. 2024.

[23] S. Damiano, F. Miotello, M. Pezzoli, A. Bernardini, F. Antonacci, A. Sarti, and T. van Waterschoot, "A zero-shot physics-informed dictionary learning approach for sound field reconstruction," in *Proc. 2025 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, (Hyderabad, India), Apr. 2025.

[24] G. Lin, P. Hu, F. Chen, X. Chen, J. Chen, J. Wang, and Z. Shi, "BINet: Learning to Solve Partial Differential Equations with Boundary Integral Networks," *arXiv:2110.00352*, Oct. 2021.

[25] G. Lin, F. Chen, P. Hu, X. Chen, J. Chen, J. Wang, and Z. Shi, "BI-GreenNet: Learning Green's Functions by Boundary Integral Network," *Commun. Math. Stat.*, vol. 11, Mar. 2023.

[26] J. Sun, Y. Liu, Y. Wang, Z. Yao, and X. Zheng, "BINN: A deep learning approach for computational mechanics problems based on boundary integral equations," *Comput. Methods Appl. Mech. Eng.*, vol. 410, 2023.

[27] M. Nagy-Huber and V. Roth, "Physics-informed boundary integral networks (PIBI-Nets): A data-driven approach for solving partial differential equations," *J. Comput. Sci.*, vol. 81, p. 102355, Sept. 2024.

[28] S. Kirkup, "The Boundary Element Method in Acoustics: A Survey," *Applied Sciences*, vol. 9, Apr. 2019.

[29] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustic Holography*. London: Academic Press, 1999.

[30] R. Scheibler, E. Bezzam, and I. Dokmanic, "Pyroomacoustics: A Python Package for Audio Room Simulation and Array Processing Algorithms," in *Proc. 2018 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, (Calgary, AB), Apr. 2018.

[31] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, and others, "Pytorch: An imperative style, high-performance deep learning library," in *Adv. Neural Inf. Process. Syst.*, vol. 32, 2019.

Authorized licensed use limited to: KU Leuven University Library. Downloaded on January 05,2026 at 13:27:29 UTC from IEEE Xplore. Restrictions apply.