Complex-Valued Physics-Informed Neural Network for Near-Field Acoustic Holography

Xinmeng Luan, Marco Olivieri, Mirco Pezzoli, Fabio Antonacci, Augusto Sarti
Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano
Piazza Leonardo Da Vinci 32, 20133 Milan, Italy
xinmeng.luan@mail.polimi.it, marco1.olivieri@polimi.it,
mirco.pezzoli@polimi.it, fabio.antonacci@polimi.it, augusto.sarti@polimi.it

***Abstract*—We present a novel approach to Near-field Acoustic Holography (NAH) with the introduction of the Complex-Valued Kirchhoff-Helmholtz Convolutional Neural Network (CV-KHCNN). Our study focuses on analyzing Complex-Valued Neural Networks (CVNNs) in the application of NAH scenario. We compare the performance between CV-KHCNN and its equivalent Real-Valued Neural Networks (RVNNs). Moreover, different complex activation functions are evaluated for CV-KHCNN. The results emphasize the effectiveness of CVNNs in tackling NAH challenges and highlight the suitability of Cardioid as the activation function for CVNNs. This discovery underscores the promising contributions of CVNNs to the field of NAH. T-distributed Stochastic Neighbor Embedding (t-SNE) is further adopted to visualize the features of the embedding layer. The results show that even without prior knowledge of the vibrations, CV-KHCNN demonstrates the capability to distinguish between different boundary conditions (BCs) and mode shapes.**

***Index Terms*—near-field acoustic holography, complex-valued neural networks, physics-informed neural network**

I. Introduction

Near-field Acoustic Holography (NAH) is a non-contact method used to analyze vibrating structures [1]. It aims at reconstructing surface velocity fields by measuring nearby microphone array data, allowing the identification of vibrating regions. This technique is especially beneficial for delicate or lightweight objects like musical instruments, as it avoids the potential damage caused by accelerometers and the impact of increasing the mass of the vibrating surface. The challenge with NAH lies in the inversion of Kirchhoff-Helmholtz (KH) integral, which is highly ill-conditioned [2], [3].

Recently, following the success of Deep Neural Networks (DNNs) in solving several acoustic problems [4]–[7], DNNs have been employed for NAH. A UNet-based [8] Convolutional Neural Network (CNN) for planar NAH solver has been proposed in [9], [10]. Later, inspired by the paradigm of physics-informed neural networks [11], the authors proposed to incorporate prior knowledge of the KH integral into the DNN model leading to a physics-informed CNN [12] known as Kirchhoff-Helmholtz Convolutional Neural Network (KHCNN). While previous CNNs [9], [10] estimate the magnitude of the velocity field, KHCNN provides an estimate of the real and imaginary part of the vibrational field using a two-channel Real-Valued (RV) CNN. Exploiting the KH integral, the output of the network is thus propagated to the microphones reconstructing the measured sound field. The KHCNN has been analyzed in [13] employing a Grad-CAM-inspired approach for visual explanation of the network. The results aid in understanding the decision-making process of KHCNN by examining mode shapes symmetry and edge behavior. A 3D CNN based Stacked Autoencoder approach for NAH (CSA-NAH) has been proposed for planar [14], [15] and cylindrical [16] NAH. A physics-constrained approach is instead employed in [17], where the coefficients of the equivalent sources [18] are estimated using machine learning.

Complex-Valued Neural Networks (CVNNs) is expressly designed for processing complex numbers [19], [20]. CVNNs appear to be the natural choice for describing data that are inherently Complex-Valued (CV), rather than resorting to two inputs (real and imaginary parts) or concatenating the real and imaginary parts into a longer input vector. They offer advantages due to the unique arithmetic of complex numbers compared to their RV counterparts, such as complex multiplication and Wirtinger calculus [21]. Therefore, CVNNs provide a more constrained system for processing CV data. CVNNs have found successful application in different fields, including sound field estimation [22]. However the complex arithmetic operations involved in CVNNs results in a higher number of trainable parameters. Achieving the comparison between CVNNs and RVNNs requires constructing an equivalent RVNN architecture that aligns with the trainable parameters of the CVNN [23], [24]. This approach eliminates bias stemming from the amount of trainable parameters discrepancies, facilitating a fair performance assessment.

In this paper, we transform RV KHCNN (RV-KHCNN) [12] into its CV counterpart, thus defining the Complex-Valued Kirchhoff-Helmholtz Convolutional Neural Network (CV-KHCNN). Our main goal is to implement CVNNs into the physics-inspired model KHCNN for NAH scenario and compare the performance of CV-KHCNN with its RV-KHCNN version. Moreover, different complex activation functions are evaluated for CV-KHCNN. Subsequently, in order to gain deeper insights into the hidden behaviors of the network, we utilize t-distributed Stochastic Neighbor Embedding (t-SNE) to visualize the features of the embedding layer. The

This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on “Telecommunications of the Future” (PE00000001 - program “RESTART”).

analysis on the learned features highlights that CV-KHCNN has the ability to discern different BCs and mode shapes even without prior knowledge. In general, the results show an improved performance of the CV-KHCNN with respect to its RV counterpart underlying the effectiveness of CV models for solving NAH.

II. Background on NAH

A. Kirchhoff-Helmholtz Integral

The exterior pressure field radiated from a vibrating structure can be characterized by the well-known KH integral [25]. Let $\mathcal{S}$ denote the vibrating surface with points $\mathbf{s}$, and $\mathbf{r}$ denote the exterior measured point. The pressure at $\mathbf{r}$ can be solved using the KH integral as [25]

$$\begin{aligned} p(\mathbf{r},\omega) = &\int_{\mathcal{S}} p(\mathbf{s},\omega)\frac{\partial}{\partial \mathbf{n}} g_\omega(\mathbf{r},\mathbf{s}) d\mathbf{s} \\ &- j\omega\rho_0 \int_{\mathcal{S}} v_n(\mathbf{s},\omega) g_\omega(\mathbf{r},\mathbf{s}) d\mathbf{s}, \end{aligned} \tag{1}$$

where $\mathbf{n}$ is the outward normal direction unit vector, $p(\cdot,\omega)$ and $v_{\mathrm{n}}(\cdot,\omega)$ are the pressure and the normal velocity field, respectively, ω is the angular frequency and $\rho_0 \approx 1.225\,\mathrm{kg\,m^{-3}}$ is the air mass density at $20\,^{\circ}\mathrm{C}$. $g_\omega(\mathbf{r},\mathbf{s})$ is the free-field Green's function from $\mathbf{s}$ to $\mathbf{r}$, written as [1]

$$g_\omega(\mathbf{r},\mathbf{s}) = \frac{1}{4\pi}\frac{e^{-j\frac{\omega}{c}||\mathbf{r}-\mathbf{s}||}}{||\mathbf{r}-\mathbf{s}||}, \tag{2}$$

with c the sound speed in air and j the imaginary unit.

B. NAH Problem Formulation

Using the KH integral, we can calculate the pressure emitted by a vibrating source based on the pressure and normal velocity fields at the surface of the object. However, NAH aims at solving the inverse problem, i.e. computing $v_n(\mathbf{s},\omega)$ starting from $p(\mathbf{r},\omega)$ acquired by a microphone array, which is a highly ill-posed problem [1].

In this study, we specifically focus on planar NAH. In close proximity of the vibrating surface, we designate the positions $\mathbf{r}_m$, $m \in [1, M]$ of M measurement points on a plane $\mathcal{H}$, which we refer to as the hologram plane. Additionally, we consider a sampled normal velocity field evaluated at points $\mathbf{s}_n$, $n \in [1, N]$ lying on the surface plane $\mathcal{S}$. The normal distance between plane $\mathcal{H}$ and $\mathcal{S}$ is denoted as $z_{\mathcal{H}}$. The inverse problem can be represented by

$$\hat{v}_n(\mathbf{s},\omega)\Big|_{\mathbf{s}\in\mathcal{S}} \approx \Gamma^{-1}\left[p(\mathbf{r},\omega)\right]\Big|_{\mathbf{r}\in\mathcal{H}}, \tag{3}$$

where Γ is a discrete estimator that approximates the pressure field on the hologram plane given the normal velocity field on the surface plane.

III. Kirchhoff-Helmholtz-based CNN for NAH

A. CV-KHCNN Architecture

An RV physics-inspired UNet-based model, RV-KHCNN, has already been introduced for addressing the reconstruction challenges in NAH [12]. Building upon the framework of RV-KHCNN, we incorporate the state-of-the-art CVNNs architecture, resulting in the enhanced model, CV-KHCNN.

The backbone of the model is the modified UNet [8]. CV-KHCNN adopts a Multi-Task Learning (MTL) approach, featuring a shared trunk encoder and bottleneck, as well as two separate decoders, followed by the KH propagation model, which has been demonstrated in Sec. II. Note that the frequencies f, NAH configurations (sampled points positions $\mathbf{r}_m$, $\mathbf{s}_n$, and normal distances between planes $z_{\mathcal{H}}$) are all known in the KH propagation model. The main architecture of CV-KHCNN remains aligned with that of RV-KHCNN, while using complex activation functions, convolution, transposed convolution, max pooling, and batch normalization [20]. Refer to Fig. 1 for an illustration of the architecture.

We adopt the technique in [24] to ensure competitive number of trainable parameters for CVNN and RVNN. The transformation ratio for RV-KHCNN is computed to be $r = 1.4162$, with the channel transformation as: 4 to 6, 8 to 11, 16 to 23, 32 to 45, 48 to 68, and 64 to 90. In such case, the trainable parameters for CV-KHCNN and RV-KHCNN are 362,220 and 359,747, respectively.

B. Input/Output data

The network input is the pressure field acquired at the hologram plane $\mathbf{P}_{\mathcal{H}}(\omega)$ with a resolution of $M = 8 \times 8$, corresponding to the samples on the hologram plane. On the other hand, the two outputs of the backbone coming from two decoders are the higher resolution pressure $\hat{\mathbf{P}}_{\mathcal{S}}(\omega)$ and velocity $\hat{\mathbf{V}}(\omega)$ on the surface plane in $N = 16 \times 64$, respectively.

The proposed approach is evaluated using the open dataset of simulated aluminum rectangular plates for NAH [12]. This dataset includes different boundary conditions (BCs), such as simply supported, clamped, and free edges. Note that the surface plane is located $z_{\mathcal{H}} = 3\,\mathrm{cm}$ away from the hologram plane, and the eigenfrequencies of the plates are in the range $f \in [0, 2]\,\mathrm{kHz}$. Moreover, the dataset pre-processing procedure keeps the same as [12].

C. Activation functions

Various activation functions have been proposed in the literature for CVNNs. The most renowned ones are the variations of ReLU, such as $\mathbb{C}$ReLU, zReLU [26] and modReLU [27]. Virtue et. al. [28] propose the Cardioid activation function. Later, the adaptive Cardioid function (denoted as A-Cardioid) was proposed by introducing a trainable bias [29]. The mathematical expression of the activation functions are summarized in Table I.

The encoding of phase information has varying performance among different complex activation functions. Both zReLU and modReLU divide the complex plane only into two regions. For zReLU, phase is preserved only in the first quadrant, while modReLU preserves phase only outside the circle centered at the origin [20]. $\mathbb{C}$ReLU retains the phase of the first quadrant, maps the phase of the second quadrant to $\pi/2$, projects the phase of the fourth quadrant to 0, and neglects the phase of

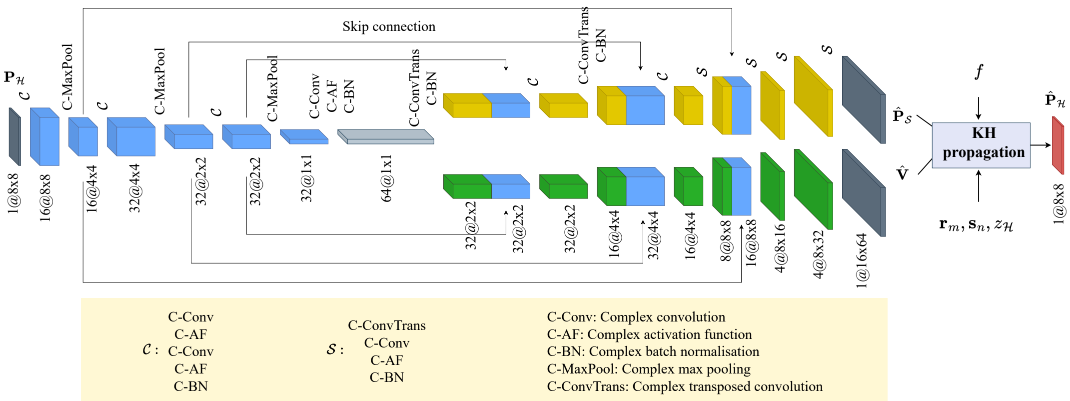


Fig. 1: The architecture of CV-KHCNN. $\mathbf{P}_{\mathcal{H}}$ is fed in the encoder (in blue), through a bottleneck (in light gray), followed by two decoders with output $\hat{\mathbf{P}}_{\mathcal{S}}$ (in yellow) and $\hat{\mathbf{V}}$ (in green), respectively. The network outputs then are fed into the KH propagation model to get the reconstruction $\hat{\mathbf{P}}_{\mathcal{H}}$, with known frequency and NAH configuration.

modReLU	$f(z) = \mathrm{ReLU}(\lvert z\rvert + b)e^{i\theta_z}$
zReLU	$f(z) = \begin{cases} z & \text{if} \quad \theta_z \in [0, \pi/2] \\ 0 & \text{otherwise} \end{cases}$
$\mathbb{C}$ReLU	$f(z) = \mathrm{ReLU}(\Re(z)) + i\mathrm{ReLU}(\Im(z))$
Cardioid	$f(z) = \frac{1}{2}z(1 + \cos(\theta_z))$
A-Cardioid	$f(z) = \frac{1}{2}z(1 + \cos(\theta_z + \theta_b))$

* $z = \lvert z\rvert\, e^{i\theta_z}$ and b is the trainable bias.

TABLE I: Mathematical expressions of complex activation functions in CVNN.

	$\hat{\mathbf{V}}$		$\hat{\mathbf{P}}_{\mathcal{H}}$	
	NMSE	**NCC**	**NMSE**	**NCC**
RV-KHCNN	-17.46	99.23%	-23.72	99.83%
modReLU	-13.55	98.28%	-24.27	99.87%
$\mathbb{C}$ReLU	-17.92	99.32%	-23.46	99.83%
Cardioid	**-18.99**	**99.48%**	**-26.30**	**99.91%**
A-Cardioid	-18.89	99.47%	-25.99	99.90%

*Values marked in bold are the best performances.

TABLE II: The mean value of NMSE and NCC of $\hat{\mathbf{V}}$ and $\hat{\mathbf{P}}_{\mathcal{H}}$ for different models. The first row represents RV-KHCNN, while the subsequent rows correspond to CV-KHCNN with different activation functions.

the third quadrant [20]. On the other hand, for the Cardioid activation function, the phase projection follows a cardioid shape with a fixed orientation toward the real axis [28]. The A-Cardioid can rotate the orientation through the bias term, which may be adapted for the data phase characteristic.

D. Training Procedure

We use *complexPyTorch* [30] to implement CVNNs in this study. The network is trained using the Adam optimizer with an initial learning rate of 0.01. We decrease the learning rate by a factor of 0.1 with a patience of 5 until it reached a minimum of 0.0009 through learning plateau. Moreover, we adopt the early stopping regularization technique to prevent over-fitting. The training stops after 20 epochs during which no improvement in the validation loss was observed. Moreover, the dataset is splitted in 80% for the training set, 10% for the validation set and 10% for the test set.

Several models are adopted with different complex activation functions: $\mathbb{C}$ReLU, zReLU, modReLU, Cardioid, and A-Cardioid. The same configuration is used for all the models to ensure the reproducibility.

IV. Results

A. Metrics Description

The performance of the proposed network is assessed by two metrics: the Normalized Mean Square Error (NMSE) (defined in dB) and the Normalized Cross Correlation (NCC). They are expressed by

$$\mathrm{NMSE}(\hat{\mathbf{x}}, \mathbf{x}) = 10\log_{10}\left(\frac{\mathbf{e}^H \cdot \mathbf{e}}{\mathbf{x}^H \cdot \mathbf{x}}\right), \tag{4}$$

and

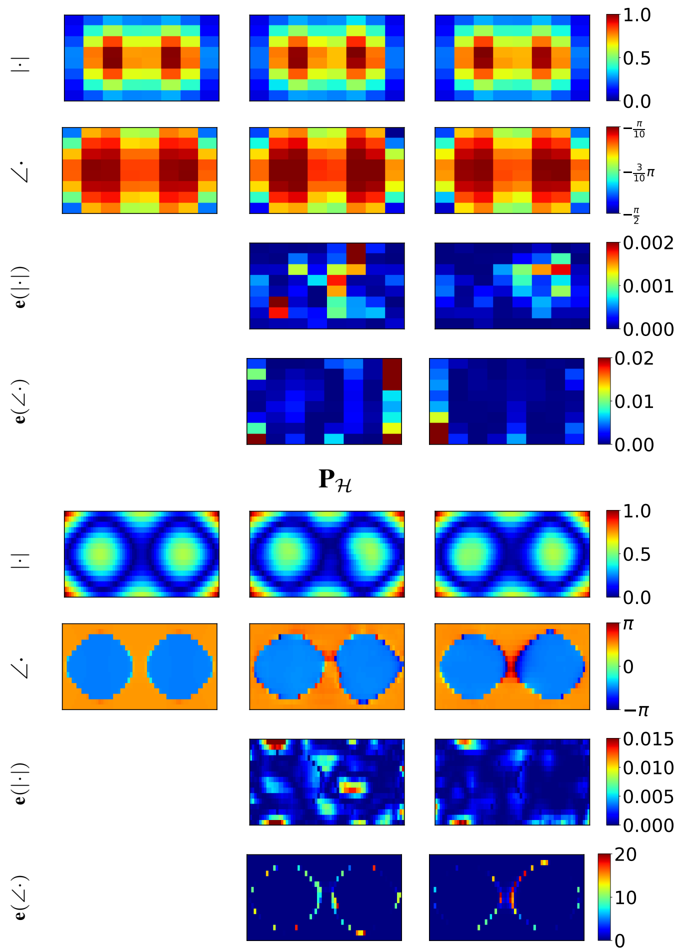
$$\mathrm{NCC}(\hat{\mathbf{x}}, \mathbf{x}) = \frac{\left\lvert \hat{\mathbf{x}}^H \cdot \mathbf{x} \right\rvert}{\lvert\lvert \hat{\mathbf{x}} \rvert\rvert_2 \cdot \lvert\lvert \mathbf{x} \rvert\rvert_2}, \tag{5}$$

where $\mathbf{x}$ are the ground truth data, $\hat{\mathbf{x}}$ are the predicted data, and $\mathbf{e} = \hat{\mathbf{x}} - \mathbf{x}$ denote the error. Additionally, the metrics are computed with a column-vector representation of the data, and NCC reaches 1 when the two quantities match perfectly. Note that both the metrics are for complex numbers and ^{H} is the Hermitian transpose operator.

B. Results and Comparisons

For the sake of simplicity, CV-KHCNN with different activation functions are denoted by their activation function names. By applying the early stopping strategy, RV-KHCNN, modReLU, $\mathbb{C}$ReLU, Cardioid and A-Cardioid get no further improvements after 22, 79, 30, 131 and 58 epochs, respectively. Instead, zReLU fails during training after several epochs.

Both the reconstructed $\hat{\mathbf{V}}$ and $\hat{\mathbf{P}}_{\mathcal{H}}$ are evaluated by NMSE and NCC. Table II presents the mean value of NMSE and NCC for $\hat{\mathbf{V}}$ and $\hat{\mathbf{P}}_{\mathcal{H}}$ across the different models. Both Cardioid and A-Cardioid show better results than RV-KHCNN in terms of NMSE and NCC for both $\hat{\mathbf{P}}_{\mathcal{H}}$ and $\hat{\mathbf{V}}$, while $\mathbb{C}$ReLU only predicts better $\hat{\mathbf{V}}$ than RV-KHCNN. On the other hand, modReLU performs worse in predicting $\hat{\mathbf{V}}$. Intuitively, these results



(a) Ground truth (b) RV-KHCNN (c) CV-KHCNN

$\mathbf{V}$

Fig. 2: An example of rectangular plate for free BC at 734.92 Hz. The upper and lower collection corresponds to $\mathbf{P}_{\mathcal{H}}$ and $\mathbf{V}$, respectively. In each collection, the first two rows represent the module and the phase, respectively. The last two rows represent the squared difference between the ground truth and the reconstruction. The first column (a) represents the ground truth, followed by (b) RV-KHCNN and (c) CV-KHCNN respectively.

aligns with our expectations in the terms of the way different complex activation functions handles phase information, as discussed in Sec. II. The outstanding performance of Cardioid is particularly noteworthy. Once again, the experimental results reinforce the belief that Cardioid is a strong choice for CVNNs activation functions [28]. Cardioid places more emphasis on the phase information, which is suitable for dealing with complex numbers. Regarding the A-Cardioid, introducing a trainable bias does not enhance the performance in our test. However, the difference is minimal, with the NMSE and NCC results being nearly identical.

A visualization example is shown in Fig. 2. Based on the metric results, we only compare the best-performing CV-KHCNN model, which uses the Cardioid activation function, with RV-KHCNN and ground truth. The reconstructions $\hat{\mathbf{P}}_{\mathcal{H}}$ and $\hat{\mathbf{V}}$, including their magnitude, phase, and errors, are displayed. The squared difference error $\mathbf{e} = (\hat{\mathbf{x}} - \mathbf{x})^2$ is visualized in the last two rows. Note that to emphasize the differences, a consistent normalized range for the colorbar of the same type of data is carefully selected to facilitate comparison. By directly comparing the magnitude and phase images, it is evident that CV-KHCNN reconstructs $\hat{\mathbf{V}}$ more accurately, especially the phase. The patterns reconstructed by CV-KHCNN closely resemble the ground truth. Observing the error images can provide a more intuitive understanding of the differences.

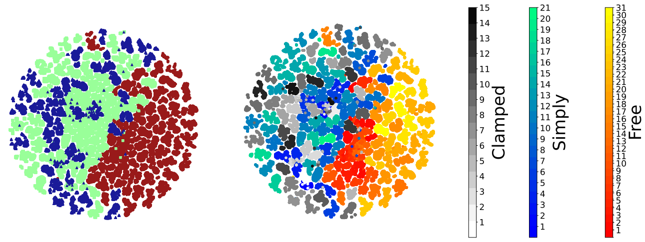
C. *t-SNE Visualization of the bottleneck*

Understanding the behavior of the network is a topic worthy of attention. Previously, an analysis of KHCNN has been done in [13]. Observing CV-KHCNN's ability to reconstruct fields with different BCs, we anticipate discovering informative features encoded within the network's representations. Therefore, we adopt the t-distributed Stochastic Neighbor Embedding (t-SNE) to visualize the underlying behavior of the CV-KHCNN.

T-SNE is a statistical method for visualizing high-dimensional data by reducing the tendency for data-points to cluster densely in a 2D or 3D map [31]. We utilize t-SNE technique on the output features of the CV-KHCNN embedding layer, i.e., the bottleneck. The dataset is labeled with both BCs and mode shapes. This allows us to assess whether the shared trunk has encoded information pertaining to the classification of distinct BCs and mode shapes without informing the network about the boundary conditions.

Figure 3 shows the 2D t-SNE results of CV-KHCNN embedding layer. The dataset is labeled by different BCs in Fig. 3a. Free and clamped BCs features are placed all together in a partition of the space since their mode shapes are completely different. Conversely, simply supported BC features are mixed with clamped BC features, which is reasonable since their mode shapes are similar. It is worth noticing that the data is accumulated into numerous small clusters. To further examine whether each small cluster is related to the mode shape, the t-SNE in Fig. 3b is labeled by different mode numbers with respect to the BCs. To distinguish the mode number for different BCs, three different colorbars are utilized according to the BCs. The visualization reveals that most small clusters correspond to the same mode number for identical BCs. Larger mode numbers cluster more tightly in one region, whereas smaller mode shapes tend to be relatively mixed together. This may be due to the fact that higher mode shapes have more unique patterns. In practice, lower frequencies tend to exhibit greater similarity to each other, with their representations concentrated towards the center of t-SNE space. As frequencies increase, mode shapes become more "complex" and different each other, causing them to spread towards the edges of t-SNE space.

Visualization of the embedding layer using t-SNE reveals that a 64-dimensional CV layer is capable of effectively encoding and distinguishing between different BCs and mode



(a) (b)

Fig. 3: T-SNE Visualization of the embedding layer features. The dataset is labeled by BCs in (a) (red for free, green for simply supported, and blue for clamped BCs), and different mode numbers with respect to the BCs in (b).

shapes. It also indicates that CV-KHCNN can infer different BCs and mode shapes without prior knowledge.

V. Conclusions

In this paper, the previous RV data-driven physics-inspired approach RV-KHCNN in NAH is transformed into its CV counterpart, i.e. CV-KHCNN. The performances of CV-KHCNN with various CV activation functions are compared with RV-KHCNN, revealing that CV-KHCNN with Cardioid activation function yields the best results. This result validates the effectiveness of CVNNs in solving the NAH problem, while also highlighting the suitability of Cardioid as the activation function for CVNNs. Furthermore, t-SNE analysis is conducted to visualize the features of the embedding layer. The results suggest that CV-KHCNN possesses the ability to differentiate between BCs and mode shapes, even without prior knowledge.

References

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

[2] E. Fernandez-Grande, A. Xenaki, and P. Gerstoft, "A sparse equivalent source method for near-field acoustic holography," *J. Acoust. Soc. Am.*, vol. 141, no. 1, pp. 532–542, 2017.

[3] G. Chardon, L. Daudet, A. Peillot, F. Ollivier, N. Bertin, and R. Gribonval, "Near-field acoustic holography using sparse regularization and compressive sampling principles," *J. Acoust. Soc. Am.*, vol. 132, no. 3, pp. 1521–1534, 2012.

[4] M. Olivieri, R. Malvermi, M. Pezzoli, M. Zanoni, S. Gonzalez, F. Antonacci, and A. Sarti, "Audio information retrieval and musical acoustics," *IEEE Instrum. Meas. Mag*, 2021.

[5] M. Acerbi, R. Malvermi, M. Pezzoli, F. Antonacci, A. Sarti, and R. Corradi, "Interpolation of irregularly sampled frequency response functions using convolutional neural networks," in *Int. Conf. Acoust. Speech Signal Process.*, IEEE, 2021.

[6] M. Olivieri, L. Comanducci, M. Pezzoli, D. Balsarri, L. Menescardi, M. Buccoli, S. Pecorino, A. Grosso, F. Antonacci, and A. Sarti, "Real-time multichannel speech separation and enhancement using a beamspace-domain-based lightweight cnn," in *Int. Conf. Acoust. Speech Signal Process.*, pp. 1–5, IEEE, 2023.

[7] M. Cobos, J. Ahrens, K. Kowalczyk, and A. Politis, "An overview of machine learning and other data-based methods for spatial audio capture, processing, and reproduction," *Eurasip J. Audio Speech Music Process.*, 2022.

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Med. Image Comput. Comput.-Assist. Interv.*, Springer, 2015.

[9] M. Olivieri, M. Pezzoli, R. Malvermi, F. Antonacci, and A. Sarti, "Near-field acoustic holography analysis with convolutional neural networks," in *INTER-NOISE NOISE-CON Congr. Conf. Proc.*, Institute of Noise Control Engineering, 2020.

[10] M. Olivieri, M. Pezzoli, F. Antonacci, and A. Sarti, "Near field acoustic holography on arbitrary shapes using convolutional neural network," in *Eur. Signal Process. Conf.*, IEEE, 2021.

[11] 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 *Forum Acusticum*, EAA, 2023.

[12] M. Olivieri, M. Pezzoli, F. Antonacci, and A. Sarti, "A physics-informed neural network approach for nearfield acoustic holography," *Sensors*, vol. 21, no. 23, p. 7834, 2021.

[13] H. Kafri, M. Olivieri, F. Antonacci, M. Moradi, A. Sarti, and S. Gannot, "Grad-cam-inspired interpretation of nearfield acoustic holography using physics-informed explainable neural network," in *Int. Conf. Acoust. Speech Signal Process.*, IEEE, 2023.

[14] J. Wang, Z. Zhang, Y. Huang, Z. Li, and Q. Huang, "A 3d convolutional neural network based near-field acoustical holography method with sparse sampling rate on measuring surface," *Measurement*, vol. 177, p. 109297, 2021.

[15] J. Wang, Z. Zhang, Z. Li, and Q. Huang, "Research on joint training strategy for 3d convolutional neural network based near-field acoustical holography with optimized hyperparameters," *Measurement*, vol. 202, p. 111790, 2022.

[16] J. Wang, W. Zhang, Z. Zhang, and Y. Huang, "A cylindrical near-field acoustical holography method based on cylindrical translation window expansion and an autoencoder stacked with 3d-cnn layers," *Sensors*, vol. 23, no. 8, p. 4146, 2023.

[17] S. Chaitanya, S. Sriraman, S. Srinivasan, and K. Srinivasan, "Machine learning aided near-field acoustic holography based on equivalent source method," *J. Acoust. Soc. Am.*, vol. 153, no. 2, pp. 940–951, 2023.

[18] S. Lee, "The use of equivalent source method in computational acoustics," *J. Comput. Acoust.*, vol. 25, no. 01, p. 1630001, 2017.

[19] A. Hirose, *Complex-valued neural networks*, vol. 400. Springer Science & Business Media, 2012.

[20] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. J. Pal, "Deep complex networks (2017)," *arXiv preprint arXiv:1705.09792*, 2017.

[21] W. Wirtinger, "Zur formalen theorie der funktionen von mehr komplexen veränderlichen," *Math. Ann.*, vol. 97, no. 1, pp. 357–375, 1927.

[22] F. Ronchini, L. Comanducci, M. Pezzoli, F. Antonacci, and A. Sarti, "Room transfer function reconstruction using complex-valued neural networks and irregularly distributed microphones," *arXiv preprint arXiv:2402.04866*, 2024.

[23] N. Mönning and S. Manandhar, "Evaluation of complex-valued neural networks on real-valued classification tasks," *arXiv preprint arXiv:1811.12351*, 2018.

[24] J. A. Barrachina, C. Ren, C. Morisseau, G. Vieillard, and J.-P. Ovarlez, "Comparison between equivalent architectures of complex-valued and real-valued neural networks-application on polarimetric sar image segmentation," *J. Signal Process. Syst.*, vol. 95, no. 1, pp. 57–66, 2023.

[25] G. H. Koopmann, L. Song, and J. B. Fahnline, "A method for computing acoustic fields based on the principle of wave superposition," *J. Acoust. Soc. Am.*, vol. 86, no. 6, pp. 2433–2438, 1989.

[26] N. Guberman, "On complex valued convolutional neural networks," *arXiv preprint arXiv:1602.09046*, 2016.

[27] M. Arjovsky, A. Shah, and Y. Bengio, "Unitary evolution recurrent neural networks," in *Int. Conf. Mach. Learn.*, PMLR, 2016.

[28] P. Virtue, X. Y. Stella, and M. Lustig, "Better than real: Complex-valued neural nets for mri fingerprinting," in *Int. Conf. Image Process.*, pp. 3953–3957, IEEE, 2017.

[29] G. Daval-Frerot, X. Chen, S. Arberet, B. Mailhé, and M. Nadar, "Exploring complex-valued neural networks with trainable activation functions for magnetic resonance imaging," in *Int. Soc. Magn. Reson. Med.*, 2019.

[30] M. W. Matthès, Y. Bromberg, J. de Rosny, and S. M. Popoff, "Learning and avoiding disorder in multimode fibers," *Phys. Rev. X*, vol. 11, no. 2, p. 021060, 2021.

[31] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne.," *J. Mach. Learn. Res.*, vol. 9, no. 11, 2008.