DOI: 10.61782/fa.2025.0601

FORUM ACUSTICUM
EURONOISE 2025

NEURAL NETWORK SOLVERS FOR BINAURAL ENCODING WITH PERCEPTION-BASED LOSSES

O. Berebi[1*] **Z. Ben-Hur**[2] **D. L. Alon**[2] **B. Rafaely**[1]

[1] School of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Israel.
[2] Reality Labs Research, Meta, Redmod, WA, USA.

ABSTRACT

Binaural reproduction is essential for immersive spatial audio in applications such as virtual and augmented reality (VR/AR). Achieving high-quality spatial audio requires accurately modeling perceptual cues, often leading to non-convex optimization tasks. Recent examples include magnitude least squares (MagLS) with interaural level-difference (ILD) denoted as iMagLS optimization for first-order Ambisonics or binaural signal matching (BSM). Traditional numerical solvers for these tasks can be computationally expensive and time-consuming. This paper introduces a neural network-based optimizer for signal-independent binaural rendering in non-convex optimization tasks. The proposed network is trained with a perceptually motivated loss function incorporating mean-squared error (MSE), magnitude error, and ILD matching, providing a faster and potentially more perceptually accurate alternative. We compare the neural network approach against conventional gradient-based methods, like Quasi-Newton methods, in terms of computational efficiency and binaural signal accuracy. Preliminary results demonstrate competitive performance while significantly reducing computational overhead, showing promise for optimizing other perception-based losses without analytical or efficient iterative solutions.

Keywords: *Spatial Audio, Binaural Reproduction, Neural Network, Perception Loss.*

**Corresponding author*: *berebio@post.bgu.ac.il.*

***Copyright**: ©2025 This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.*

1. INTRODUCTION

Binaural reproduction is essential for creating immersive spatial audio experiences in applications such as virtual and augmented reality (VR/AR) [1]. High-quality binaural rendering relies on accurately modeling perceptual cues, which often leads to non-convex optimization problems. These optimization tasks arise in various binaural reproduction methods, such as spherical arrays using the Ambisonics format or arbitrary microphone arrays where binaural signal matching (BSM) is used to render the binaural signals [2, 3].

A widely used approach for perceptual modeling is magnitude least squares (MagLS) optimization, which improves spectral accuracy by minimizing magnitude errors between the reference and the reproduced binaural signals [3, 4]. In the MagLS case, the optimization problem can be efficiently solved with a simple one-shot frequency-recursive approach for the directional phase [2, 5]. However, recent methods have sought to incorporate additional perceptual cues into the optimization process beyond magnitude matching. One such example is interaural level difference (ILD) magnitude least squares (iMagLS), which simultaneously optimizes magnitude and ILD cues [6–8]. While these approaches improve binaural rendering quality, their optimization typically relies on iterative numerical solvers, such as Quasi-Newton methods, which can be computationally expensive and time-consuming.

In this paper, we propose a neural network-based optimizer for computing signal-independent binaural reproduction coefficients. The proposed approach enables the optimization of complex perceptually inspired loss functions, such as the iMagLS loss, providing a computationally efficient alternative to traditional numerical solvers.

We evaluate the proposed method performance against conventional optimization methods across three loss functions: the Least Squares (LS) loss, the MagLS loss, and the iMagLS loss. Results demonstrate that the proposed method achieves competitive accuracy while significantly reducing computational overhead in the iMagLS task, making it a promising solution for other perceptually based losses.

2. PROBLEM FORMULATION

This section provides the mathematical background for signal-independent binaural reproduction using a microphone array. Given an array with M microphones measuring the sound pressure, binaural signals—representing the acoustic pressure at the listener's left and right ears—can be estimated using the BSM formulation [3]:

$$\hat{p}^{l,r}(f) = [\mathbf{c}^{l,r}(f)]^H \mathbf{x}(f) \tag{1}$$

where f refer to frequency, $\mathbf{x}(f) \in \mathbb{C}^{M\times 1}$ is the array measurement vector, and $\mathbf{c}^{l,r}(f) \in \mathbb{C}^{M\times 1}$ are the BSM coefficients encoding both head-related transfer function (HRTF) and array transfer function (ATF) information. In signal-independent rendering, these coefficients remain constant for any input $\mathbf{x}(f)$.

Another approach to signal-independent binaural rendering is Ambisonics [2]. If a spherical microphone array captures the sound pressure, the array measurements $\mathbf{x}(f)$ can be transformed via the spherical harmonics transform (SHT) into an order-N Ambisonics signal, denoted as $\mathbf{a}_{nm}(f) \in \mathbb{C}^{(N+1)^2\times 1}$. The Ambisonics order depends on the number of microphones in the array [2]. These Ambisonics signals are then rendered into binaural signals using order-N HRTF coefficients, $\mathbf{h}^{l,r}_{nm}(f) \in \mathbb{C}^{(N+1)^2\times 1}$ [9]:

$$\hat{p}^{l,r}(f) = [\tilde{\mathbf{a}}_{nm}(f)]^* \mathbf{h}^{l,r}_{nm}(f) \tag{2}$$

where the modified Ambisonics signal is $\tilde{\mathbf{a}}_{nm} = (-1)^m [\mathbf{a}_{n(-m)}]^*$.

In both approaches, the sound pressure information is encoded either in the array signal $\mathbf{x}(f)$ for the BSM case or in the SH signal $\mathbf{a}_{nm}(f)$ in the Ambisonics case. In the signal-independent reproduction task, these signals are treated as the input and are usually not affected by the reproduction algorithm. The part that depends on the algorithm consists of the coefficients $\mathbf{h}^{l,r}_{nm}$ or $\mathbf{c}^{l,r}$ in the Ambisonics and BSM case, respectively. These coefficients can be viewed as a low-order representation of the array and HRTFs. Calculating the coefficients can be done by evaluating their HRTF estimation ability. The HRTF can be estimated from the coefficients as [2, 3]:

$$\hat{\mathbf{h}}^{l,r}(\Omega_k, f) = \begin{cases} \mathbf{Y}(\Omega_k)\mathbf{h}^{l,r}_{nm}(f) & \text{, Ambisonics,} \\ \mathbf{V}^T(\Omega_k, f)\left[\mathbf{c}^{l,r}(f)\right]^* & \text{, BSM.} \end{cases} \tag{3}$$

Here, $\mathbf{Y}(\Omega_k) \in \mathbb{C}^{K\times(N+1)^2}$ is the spherical harmonics transformation matrix [10], and $\mathbf{V}(\Omega_k, f) \in \mathbb{C}^{M\times K}$ is the measured or simulated ATF matrix for a dense set of K directions $\{\Omega_k = (\theta_k, \phi_k)\}_{k=1}^K$. Finally, the calculation of the coefficients can be framed as an optimization problem. Let $\mathbf{z}^{l,r}(f)$ represent these coefficients, i.e.,

$$\mathbf{z}^{l,r}(f) = \begin{cases} \mathbf{h}^{l,r}_{nm}(f) & \text{, Ambisonics,} \\ \left[\mathbf{c}^{l,r}(f)\right]^* & \text{, BSM.} \end{cases} \tag{4}$$

Then, the optimization problem is formulated as:

$$\hat{\mathbf{z}}^{l,r}(f) = \underset{\mathbf{z}^{l,r}(f)}{\arg\min}\, \mathcal{L}\left(\mathbf{h}^{l,r}_{\text{ref}}(\Omega_k, f), \hat{\mathbf{h}}^{l,r}(\Omega_k, f)\right), \tag{5}$$

where $\mathcal{L}(\cdot,\cdot)$ is a loss function measuring the dissimilarity between $\mathbf{h}^{l,r}_{\text{ref}}(\Omega_k, f) \in \mathbb{C}^{K\times 1}$, the measured or simulated HRTF from a dense set of K directions $\{\Omega_k = (\theta_k, \phi_k)\}_{k=1}^K$, and the estimated HRTF, $\hat{\mathbf{h}}^{l,r}(\Omega_k, f) \in \mathbb{C}^{K\times 1}$ of Eq. (3).

3. CURRENT METHODS

The solution $\hat{\mathbf{z}}^{l,r}(f)$ of Eq. (5) depends on the choice of loss function $\mathcal{L}(\cdot,\cdot)$, which significantly affects the spectral and spatial quality of the resulting binaural audio [2, 3, 6–8]. This section reviews commonly used loss functions and their corresponding solutions.

3.1 Least Squares

A straightforward choice for $\mathcal{L}(\cdot,\cdot)$ is the LS loss, defined as

$$\mathcal{L}_{\text{LS}}\left(\mathbf{h}^{l,r}_{\text{ref}}, \hat{\mathbf{h}}^{l,r}\right) = \|\mathbf{h}^{l,r}_{\text{ref}} - \hat{\mathbf{h}}^{l,r}\|_2^2. \tag{6}$$

This measures the complex error between the estimated and reference vectors, ensuring accurate reconstruction in both magnitude and phase. The LS loss results in a convex problem with a closed-form solution [2, 3]:

$$\hat{\mathbf{z}}^{l,r}_{\text{LS}}(f) = \begin{cases} \mathbf{Y}^\dagger \mathbf{h}^{l,r}_{\text{ref}} & \text{, Ambisonics} \\ \left[\mathbf{V}^T\right]^\dagger \mathbf{h}^{l,r}_{\text{ref}} & \text{, BSM} \end{cases} \tag{7}$$

where $(\cdot)^{\dagger}$ denotes the pseudo-inverse. However, in practical applications with a limited number of microphones, this solution is accurate only at lower frequencies. For example, in Ambisonics, frequencies above a cutoff value given by $f_c = \frac{cN}{2\pi R}$, where c is the speed of sound and $R \approx 8.75$ cm represents the average head radius, exhibit phase reconstruction errors that increase with frequency. This leads to spatial degradation and an audible frequency roll-off [2, 11].

3.2 Magnitude Least Squares

The spectral degradation of the LS solution can be mitigated by modifying $\mathcal{L}(\cdot,\cdot)$. The MagLS loss prioritizes magnitude accuracy over phase above a cut-off frequency, f_c [4, 5]:

$$\mathcal{L}_{\text{MLS}}\left(\mathbf{h}_{\text{ref}}^{l,r}, \hat{\mathbf{h}}^{l,r}\right) = \| \, |\mathbf{h}_{\text{ref}}^{l,r}| - |\hat{\mathbf{h}}^{l,r}| \, \|_2^2. \tag{8}$$

Unlike the LS loss, MagLS is non-convex and lacks a closed-form solution. As proposed in [2, 3, 5], an iterative approach can be used, where the reconstructed phase from the previous frequency is combined with the optimal HRTF magnitude of the current frequency:

$$\hat{\mathbf{z}}_{\text{MLS}}^{l,r}(f_k) = \begin{cases} \mathbf{Y}^{\dagger}\left[|\mathbf{h}_{\text{ref}}^{l,r}(f_k)| e^{i\mathbf{\Phi}^{l,r}(f_k)}\right], & \text{Ambisonics} \\ \left[\mathbf{V}^{T}\right]^{\dagger}\left[|\mathbf{h}_{\text{ref}}^{l,r}(f_k)| e^{i\mathbf{\Phi}^{l,r}(f_k)}\right], & \text{BSM} \end{cases} \tag{9}$$

where the reconstructed phase is given by:

$$\mathbf{\Phi}^{l,r}(f_k) = \begin{cases} \angle\left[\mathbf{Y}\mathbf{h}_{nm}^{l,r}(f_{k-1})\right] & \text{, Ambisonics} \\ \angle\left[\mathbf{V}^{T}(f_{k-1})\mathbf{c}^{l,r}(f_{k-1})\right] & \text{, BSM} \end{cases} \tag{10}$$

with f_k representing the current frequency bin and f_{k-1} the previous one. Below f_c, the LS solution from Eq. (7) is used. While MagLS improves magnitude accuracy and reduces frequency roll-off [2,3], its spatial quality remains suboptimal in applications with a limited number of microphones [6–8].

3.3 ILD Magnitude Least Squares

A recent modification of the MagLS loss is the iMagLS loss [6,7], which incorporates ILD error for improved perceptual accuracy:

$$\mathcal{L}_{\text{iMLS}}\left(\mathbf{h}_{\text{ref}}^{l,r}, \hat{\mathbf{h}}^{l,r}\right) = \mathcal{L}_{\text{MLS}} + \lambda \mathcal{L}_{\text{ILD}} \tag{11}$$

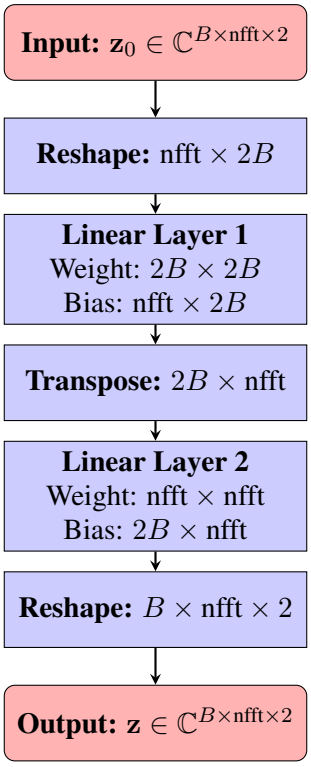


Figure 1. Neural network architecture with two linear layers. The input is reshaped, processed through transformations, and then reshaped back to its original dimensions. nfft refer to the number of positive frequency bins, B is the number of coefficients and 2 refer to the left and right ears. The linear layers are complex right-hand multiplication with bias, which are the network parameters.

where $\lambda \in \mathbb{R}$ is a weighting factor, and the ILD error is defined as

$$\mathcal{L}_{\text{ILD}} = \|ILD_{\text{ref}}(\Omega_{\text{az}}) - \hat{ILD}(\Omega_{\text{az}})\|_2^2. \tag{12}$$

The ILD is computed from a frequency-averaged, Gammatone-filtered version of the HRTF along the horizontal plane, denoted as Ω_{az} [12]. The iMagLS loss presents the most challenging optimization problem among the three, as it jointly minimizes MagLS loss and ILD error. Reducing ILD error has been shown to improve localization accuracy along the horizontal plane [13]. Due to the complexity of the joint optimization, gradient-based numerical solvers are typically

required. In this study, the Quasi-Newton Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [14] is used to minimize Eq. (11), with the LS solution from Eq. (7) serving as the initial estimate.

4. PROPOSED METHOD

The proposed method offers alternative network-based solvers for the perceptually motivated losses used for binaural rendering. Optimization of Eq. (5) with Eq. (3) requires finding the coefficients $\mathbf{h}_{nm}^{l,r}$ or $\left[\mathbf{c}^{l,r}\right]^*$, given $\mathbf{Y}$ or $\mathbf{V}^T$. Let $\mathbf{z}_0(f)$ denote an initial estimate of these coefficients. For brevity, we omit the left and right superscripts, though this notation applies to both the left and right ear BSM and Ambisonics coefficients.

A deep neural network (DNN) is employed to refine this initial estimate and produce optimized coefficients, expressed as:

$$\mathbf{z} = f_{\boldsymbol{\theta}}\left(\mathbf{z}_0\right) \tag{13}$$

where $f_{\boldsymbol{\theta}}(\cdot)$ represents the forward pass of the DNN, parameterized by $\boldsymbol{\theta}$. An example for $f_{\boldsymbol{\theta}}(\cdot)$ is presented in Fig. 1. The network parameters are optimized by minimizing the loss function in Eq. (5), leading to:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \mathcal{L}\left(\mathbf{h}_{\text{ref}}, f_{\boldsymbol{\theta}}\left(\mathbf{z}_0\right)\right) \tag{14}$$

Once the optimal network parameters $\hat{\boldsymbol{\theta}}$ are obtained, the refined coefficients are computed using Eq. (13) with the initial estimate $\mathbf{z}_0$.

A key distinction between this approach and traditional gradient-based optimization methods is that instead of directly optimizing the coefficients, we optimize the parameters of the neural network. This enables efficient differentiation through the DNN, avoiding reliance on numerical differentiation schemes [15, 16]. This is particularly advantageous when the loss function is complex and does not have an analytical derivative.

5. EXPERIMENTAL EVALUATION

A comparative analysis is conducted between the proposed method and state-of-the-art solvers.

5.1 Setup

The analysis is performed using the measured Cologne HRTF database for the Neumann KU100 dummy head as the reference HRTF [17]. First-order ($N = 1$) Ambisonics rendering is used as the estimate, with four coefficients per frequency bin. A simulated ATF of a semi-circular array with six microphones is used for the BSM estimation. The microphones were positioned on a 12 cm radius rigid sphere, with an elevation of 90° and azimuth angles of $[\pm 64°, \pm 38°, \pm 13°]$ in the front-facing direction. The loss terms are evaluated on a high-order Lebedev grid consisting of 2702 directions with a sampling rate of 48kHz and a 512 bin FFT. All methods were executed on an Apple M2 MAX laptop, utilizing only its CPU cores. The DNN solver was implemented in PyTorch, while all other methods were executed in MATLAB 2023a.

Table 1. Computation Time

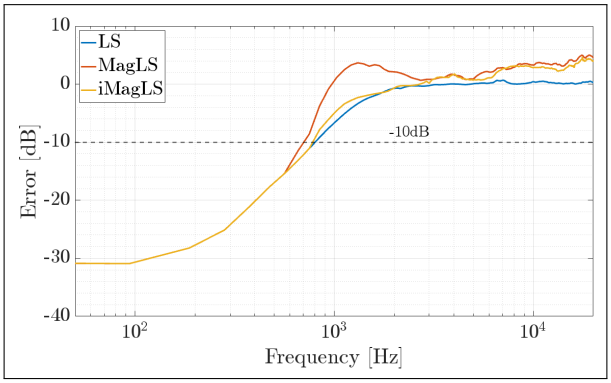
Loss	Benchmark	DNN solver
LS (Ambisonics)	**28.4 (ms)**	1.3 (min)
MagLS (Ambisonics)	**80.7 (ms)**	0.8 (min)
iMagLS (Ambisonics)	3.1 (h)	**1.4 (min)**
LS (BSM)	**42.5 (ms)**	3.8 (min)
MagLS (BSM)	**1.8 (sec)**	2.1 (min)
iMagLS (BSM)	5.4 (h)	**4.2 (min)**

Table 2. Normalized relative difference between benchmark solutions and DNN solutions

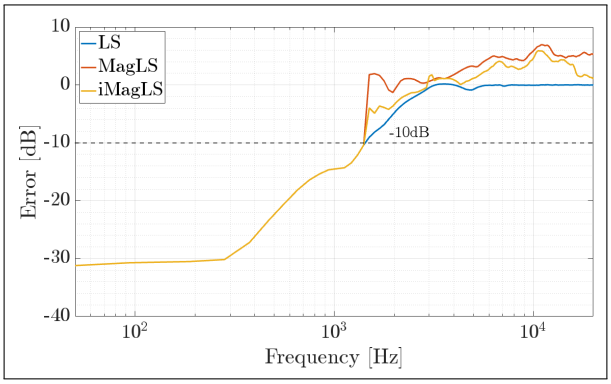
Loss	$\text{Mean}_{(\text{STD})}$[dB]	(Min,Max)[dB]
LS (Ambi)	$-246.7_{5.2}$	$(-256, -236)$
MagLS (Ambi)	$-60.5_{4.2}$	$(-71.3, -49.7)$
iMagLS (Ambi)	$-20.2_{1.7}$	$(-22.4, -16.8)$
LS (BSM)	$-122.5_{7.3}$	$(-133, -114)$
MagLS (BSM)	$-57.2_{2.9}$	$(-65.0, -46.2)$
iMagLS (BSM)	$-22.8_{1.9}$	$(-27.4, -18.1)$

5.2 Methodology

The evaluation considers three loss functions, defined in Eqs. (6), (8), and (11). Each loss function is paired with its corresponding state-of-the-art solver, while the proposed method serves as an alternative solver. The HRTF estimates are computed using Eq. (3), both the Ambisonics and BSM coefficients are obtained based on the loss functions as described in Secs. 3 and 4. The LS solu-

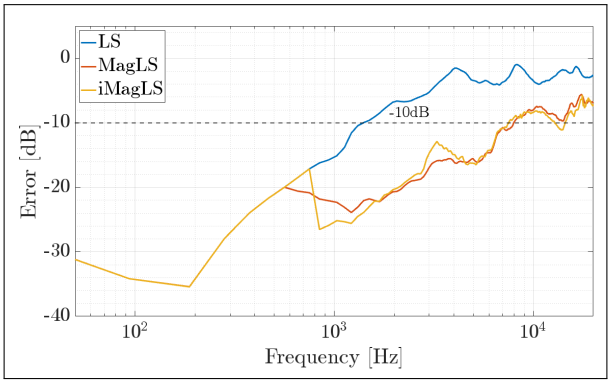


(a) Ambisonics

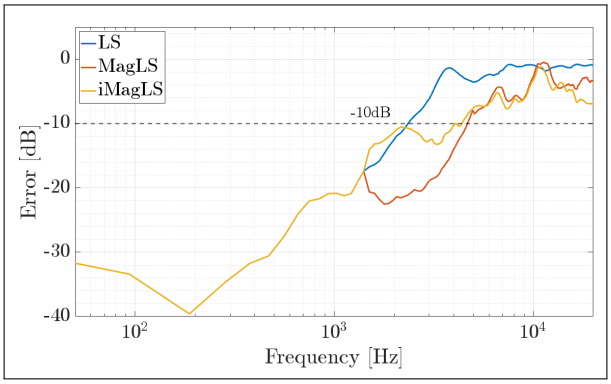


(b) BSM

Figure 2. Normalized MSE as a function of frequency, averaged across a dense Lebedev grid of 2702 directions. The Benchmark solvers are used in the LS and MagLS case, where the proposed DNN solver is used in the iMagLS task.



(a) Ambisonics



(b) BSM

Figure 3. Magnitude errors as a function of frequency, averaged across a dense Lebedev grid of 2702 directions. The Benchmark solvers are used in the LS and MagLS case, where the proposed DNN solver is used in the iMagLS task.

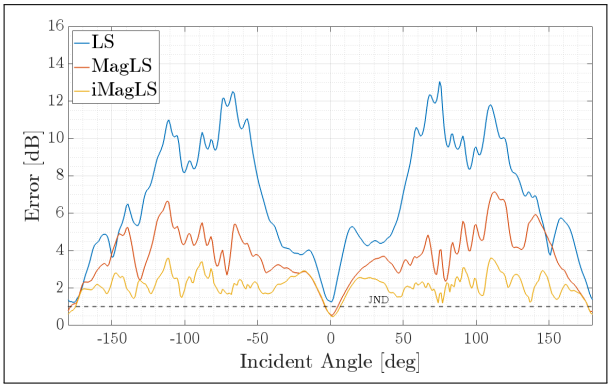
tion of Eq. (7) is used as the initial estimate for both the Quasi-Newton solver and the proposed DNN solver in the iMagLS task. The performance of the proposed DNN solver is also evaluated on the MagLS and LS loss task, and its initial estimates are the LS and a zeros vector, respectively. Both the Quasi-Newton and DNN solvers were stopped when their loss value improved by less than 0.001%. The neural network in the proposed method was optimized using the Adam optimizer with a learning rate of 0.0005 and a weighting parameter of $\lambda = 2$.

5.3 Results

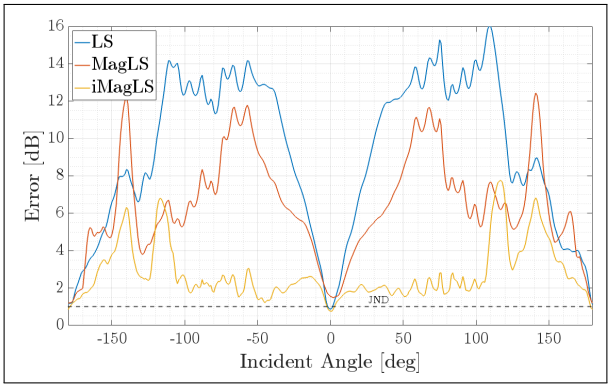
The proposed method in Sec. 4 is compared with the benchmark methods in Sec. 3 across the three loss functions in terms of accuracy and computational efficiency.

5.3.1 Computational Time

The execution times for the evaluated methods are presented in Table 1. For the LS and MagLS loss tasks, the benchmark methods execute significantly faster, completing in tens of milliseconds, whereas the DNN solver requires approximately one minute. However, for the more complex iMagLS task, the proposed method significantly outperforms the benchmark, completing in a few minutes compared to several hours for the Quasi-Newton solver.



(a) Ambisonics



(b) BSM

Figure 4. ILD error averaged across 1.5-20 kHz as a function of incident direction on the horizontal plane. The Benchmark solvers are used in the LS and MagLS case, where the proposed DNN solver is used in the iMagLS task.

5.3.2 Accuracy

To evaluate accuracy, the estimates were assessed using MSE, magnitude error, and ILD error, as defined in Eqs. (6), (8), and (12). The LS and MagLS loss performances were evaluated using the benchmark solutions from Sec. 3, while the iMagLS loss performance was assessed using the proposed DNN solver from Sec. 4. As shown in Figs. 2, 3, the optimized LS and MagLS solutions achieve the best performance for their respective tasks, as expected. However, the DNN solver shows only a minimal compromise in magnitude and complex error compared to the benchmark solutions. Figure 4 presents the ILD errors for the different solutions. The iMagLS approach outperforms the other methods, reducing ILD error across all directions and approaching the Just Noticeable Difference (JND) threshold of 1 dB [18].

Finally, evaluating the similarity between the proposed DNN-based approach and the benchmark can provide insight into the overall differences between the two methods. Table 2 summarizes these results, where the relative difference is averaged over directions, frequencies, and ears. The results in the table suggest that, across all tasks, the differences between the DNN solution and the benchmark solution are minimal, indicating similar audible performance.

5.4 Discussion

The experimental evaluation reveals two key findings. First, conventional solvers deliver fast and accurate solutions for the LS and MagLS tasks, whereas the traditional gradient-based solver was notably slow for the iMagLS task, taking approximately several hours to converge, as seen in Table. 1. In contrast, the DNN solver completed the same task in just a few minutes, marking a substantial reduction in processing time. This efficiency could be particularly beneficial for applications, where the preprocessing of the HRTF must be computed individually for each user or through a parametric technique that relies on re-evaluating the low-order HRTF representation based on the measured acoustic scene. Second, as shown in Table 2, the DNN solver performed very similarly to the benchmark methods in the LS and MagLS tasks, as evidenced by the relatively low difference between the solutions. This suggests that it successfully reached the optimal local minima in these tasks. For the iMagLS task, as illustrated in Figs. 2 and 3, the DNN solver performed comparably to the benchmark LS and MagLS solutions in terms of magnitude and complex error performance. Furthermore, it achieved a noticeable reduction in ILD error, approaching the JND threshold of 1 dB. This shows that DNN provides high computational efficiency while maintaining high perceptual fidelity in complex loss tasks such as the iMagLS task.

6. CONCLUSIONS

This study introduced a DNN solver for signal-independent binaural rendering that leverages perceptually motivated loss functions. By reframing the coefficient optimization problem, the proposed method bypasses the traditional reliance on iterative, gradient-based numerical solvers. Instead, the DNN directly refines an initial estimate of the binaural reproduction optimization param-

eters, efficiently handling the non-convex nature of the iMagLS task.

The proposed DNN-based approach provides a promising alternative to traditional gradient based optimization techniques in spatial audio processing for challenging perceptual losses. It achieves a substantial speedup in solving complex loss functions while consistently delivering high accuracy results compared to the benchmark solvers. Additionally, other perceptual losses can be considered, as the proposed solution does not rely on prior knowledge about the loss function itself. Future research will focus on integrating additional perceptual metrics, evaluating larger HRTF datasets, and exploring generalization across multiple HRTFs, given that the current formulation is tailored for a single HRTF. This work could pave the way for efficient, high-quality binaural reproduction systems that meet the demands of emerging immersive audio applications.

7. REFERENCES

[1] B. Rafaely, V. Tourbabin, E. Habets, Z. Ben-Hur, H. Lee, H. Gamper, L. Arbel, L. Birnie, T. Abhayapala, and P. Samarasinghe, "Spatial audio signal processing for binaural reproduction of recorded acoustic scenes–review and challenges," *Acta Acustica*, vol. 6, p. 47, 2022.

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

[3] L. Madmoni, Z. Ben-Hur, J. Donley, V. Tourbabin, and B. Rafaely, "Design and analysis of binaural signal matching with arbitrary microphone arrays and listener head rotations," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2025, no. 1, p. 11, 2025.

[4] C. Schörkhuber, M. Zaunschirm, and R. Höldrich, "Binaural rendering of ambisonic signals via magnitude least squares," in *Proceedings of the DAGA*, vol. 44, pp. 339–342, 2018.

[5] L. McCormack, N. Meyer-Kahlen, D. L. Alon, Z. Ben-Hur, S. V. A. Garí, and P. Robinson, "Six-degrees-of-freedom binaural reproduction of head-worn microphone array capture," *Journal of the Audio Engineering Society*, vol. 71, no. 10, pp. 638–649, 2023.

[6] O. Berebi, Z. Ben-Hur, D. L. Alon, and B. Rafaely, "Imagls: Interaural level difference with magnitude least-squares loss for optimized first-order head-related transfer function," in *10th Convention of the European Acoustics Association, EAA 2023*, European Acoustics Association, EAA, 2023.

[7] O. Berebi, Z. Ben-Hur, D. L. Alon, and B. Rafaely, "Feasibility of imagls-bsm-ild informed binaural signal matching with arbitrary microphone arrays," in *2024 18th International Workshop on Acoustic Signal Enhancement (IWAENC)*, pp. 429–433, IEEE, 2024.

[8] O. Berebi, Z. Ben-Hur, D. L. Alon, and B. Rafaely, "Bsm-imagls: Ild informed binaural signal matching for reproduction with head-mounted microphone arrays," *arXiv preprint arXiv:2501.18227*, 2025.

[9] B. Rafaely and A. Avni, "Interaural cross correlation in a sound field represented by spherical harmonics," *JASA*, vol. 127, no. 2, pp. 823–828, 2010.

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

[11] Z. Ben-Hur, D. L. Alon, R. Mehra, and B. Rafaely, "Efficient representation and sparse sampling of head-related transfer functions using phase-correction based on ear alignment," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 12, pp. 2249–2262, 2019.

[12] B. Xie, *Head-related transfer function and virtual auditory display*. J. Ross Publishing, 2013.

[13] B. F. Katz and R. Nicol, "Binaural spatial reproduction," in *Sensory Evaluation of Sound*, pp. 349–388, CRC Press, 2018.

[14] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms 1. general considerations," *IMA Journal of Applied Mathematics*, vol. 6, no. 1, pp. 76–90, 1970.

[15] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, "Automatic differentiation in machine learning: a survey," *Journal of machine learning research*, vol. 18, no. 153, pp. 1–43, 2018.

[16] 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.

[17] B. Bernschütz, "A spherical far field HRIR/HRTF compilation of the Neumann KU 100," in *Proceedings of the 40th Italian (AIA) annual conference on acoustics and the 39th German annual conference on acoustics (DAGA) conference on acoustics*, p. 29, AIA/DAGA, 2013.

[18] A. W. Mills, "Lateralization of high-frequency tones," *JASA*, vol. 32, no. 1, pp. 132–134, 1960.