SFC-L1: Sound Field Control With Least Absolute Deviation Regression

Takuma Okamoto

National Institute of Information and Communications Technology, Japan
okamoto@nict.go.jp

***Abstract*—Sound field control using loudspeaker arrays is an important acoustic and audio signal processing applications. In sound field control, least squares (LS) regression based on pressure matching or mode-matching is typically introduced to derive the driving signals of loudspeakers as a closed-form solution. The LS regression is a maximum-likelihood estimation, in which the error is assumed to be Gaussian distribution. Compared with the LS regression, the least absolute deviation (LAD) regression, in which the error is assumed to be Laplace distribution, is robust against outliers. In pressure matching-based sound field methods, outliers appear at higher frequencies according to the spatial Nyquist frequency. To improve the control accuracy for pressure matching-based methods at high frequencies, this paper proposes SFC-L1, pressure matching-based sound field control method with LAD regression instead of LS regression. In the proposed method, the LAD regression combined with L1 regularization is solved with gradient method simply implemented on PyTorch. The results of computer simulations demonstrate that the proposed LAD-based methods can improve the sound field control accuracy at high frequencies compared with the conventional LS-based methods. Additionally, PyTorch-based implementation, Torch-SFC, is open-sourced for accelerating sound field control research.**

1. INTRODUCTION

Sound field control with multiple loudspeakers is an important acoustic and audio signal processing applications. Wave field synthesis [1]–[3], higher-order Ambisonics (HOA) [4]–[6], and spectral division method [7], [8] can analytically calculate the driving functions of loudspeakers. However, these methods can be used for linear, planar, circular and spherical arrays of loudspeakers.

Compared to these analytical methods, pressure matching [9]–[12] and mode-matching [4], [13]–[17] are widely used for arbitrary distributions of microphones and loudspeakers. In these methods, least squares (LS) regression is typically introduced to derive the driving signals of loudspeakers as a closed-form solution. The LS regression is a maximum-likelihood estimation, in which the error is assumed to be Gaussian distribution. Compared with the LS regression, the least absolute deviation (LAD) regression, in which the error is assumed to be Laplace distribution, is robust against outliers [18]. In pressure matching-based sound field methods, outliers appear at higher frequencies according to the spatial Nyquist frequency.

To improve the control accuracy for pressure matching-based methods at high frequencies, this paper proposes SFC-L1, pressure matching-based sound field control method with LAD regression instead of LS regression. In the proposed method, the LAD regression combined with L1 regularization is solved with gradient method simply implemented on PyTorch. The results of computer simulations demonstrate that the proposed LAD-based methods can improve the sound field control accuracy at high frequencies compared with the conventional LS-based methods.

Additionally, PyTorch-based implementation used in computer simulations conducted in Sec. 4, Torch-SFC, is open-sourced for accelerating sound field control research.[1]

This study was partly supported by JSPS KAKENHI Grant Number JP23K11177.

[1] https://www.okamotocamera.com/torch-sfc.html

2. CONVENTIONAL LEAST SQUARES-BASED PRESSURE MATCHING METHODS

2.1. Simple pressure matching

The sound pressure at $\boldsymbol{r}$, synthesized by L loudspeakers, is given by

$$p_{\mathrm{syn}}(\boldsymbol{r}) = \sum_{l=1}^{L} d_l(\boldsymbol{r}_l) G\left(\boldsymbol{r}|\boldsymbol{r}_l\right) = \boldsymbol{g}(\boldsymbol{r})^{\top}\boldsymbol{d}, \tag{1}$$

where $\boldsymbol{d} = [d_1(\boldsymbol{r}_1), \cdots, d_L(\boldsymbol{r}_L)]^{\top}$ and $\boldsymbol{g}(\boldsymbol{r}) = [G\left(\boldsymbol{r}|\boldsymbol{r}_1\right), \cdots, G\left(\boldsymbol{r}|\boldsymbol{r}_L\right)]^{\top}$ are the vectors of the driving signals of the loudspeakers and transfer functions from $\boldsymbol{r}_l$ to $\boldsymbol{r}$, respectively.

In sound field control based on LS regression, the sound field inside the target region V is optimized by minimizing the objective function:

$$\underset{\boldsymbol{d}}{\text{minimize}}\ \mathcal{J} = \int_{\boldsymbol{r}\in V} \left|\boldsymbol{g}(\boldsymbol{r})^{\top}\boldsymbol{d} - p_{\mathrm{des}}(\boldsymbol{r})\right|^2 \mathrm{d}\boldsymbol{r}, \tag{2}$$

where $p_{\mathrm{des}}(\boldsymbol{r})$ is the desired sound pressure at $\boldsymbol{r}$. In pressure matching-based method, N control points are located inside V, and the cost function $\mathcal{J}$ in (2) is approximated as the error between the synthesized and desired pressures at the control points. The optimization problem of pressure matching method with LS regression and L2 regularization[2] is described as:

$$\underset{\boldsymbol{d}}{\text{minimize}}\ ||\boldsymbol{G}\boldsymbol{d} - \boldsymbol{p}_{\mathrm{des}}||^2 + \lambda_{\mathrm{PM}}||\boldsymbol{d}||_2, \tag{3}$$

where $\boldsymbol{G} = [\boldsymbol{g}(\boldsymbol{r}_1), \cdots, \boldsymbol{g}(\boldsymbol{r}_N)]^{\top}$, $\boldsymbol{p}_{\mathrm{des}} = [p_{\mathrm{des}}(\boldsymbol{r}_1), \cdots, p_{\mathrm{des}}(\boldsymbol{r}_N)]^{\top}$, and λ_{PM} is a weighting parameter, respectively. Then, the driving signals of pressure matching method with LS regression is solved as a closed-form:

$$\boldsymbol{d}_{\mathrm{PM}} = \left(\boldsymbol{G}^{\mathsf{H}}\boldsymbol{G} + \lambda_{\mathrm{PM}}\boldsymbol{I}\right)^{-1}\boldsymbol{G}^{\mathsf{H}}\boldsymbol{p}_{\mathrm{des}}, \tag{4}$$

where $\boldsymbol{I}$ the identity matrix.

2.2. Weighted pressure matching

Although (4) only considers the sound pressures on the control points, the sound pressures throughout the target region V can be considered by weighted pressure matching [15]. In weighted pressure matching, sound pressures without the control points can be estimated by kernel interpolation of sound field [19]. The optimization problem of weighted pressure matching method with LS regression is defined as:

$$\underset{\boldsymbol{d}}{\text{minimize}}\ (\boldsymbol{G}\boldsymbol{d} - \boldsymbol{p}_{\mathrm{des}})^{\mathsf{H}}\boldsymbol{W}_{\mathrm{PM}}(\boldsymbol{G}\boldsymbol{d} - \boldsymbol{p}_{\mathrm{des}}) + \lambda_{\mathrm{WPM}}||\boldsymbol{d}||_2, \tag{5}$$

where λ_{WPM} is also a weighting parameter. $\boldsymbol{W}_{\mathrm{PM}}$ is defined as:

$$\boldsymbol{W}_{\mathrm{PM}} = \left((\boldsymbol{\Psi} + \lambda\boldsymbol{I})^{-1}\right)^{\mathsf{H}} \int_{\boldsymbol{r}\in V} \boldsymbol{\kappa}(\boldsymbol{r})^{*}\boldsymbol{\kappa}(\boldsymbol{r})^{\top} d\boldsymbol{r} (\boldsymbol{\Psi} + \lambda\boldsymbol{I})^{-1}, \tag{6}$$

$$(\boldsymbol{\Psi})_{n,n'} = j_0(k||\boldsymbol{r}_n - \boldsymbol{r}_{n'}||), \tag{7}$$

$$\boldsymbol{\kappa}(\boldsymbol{r}) = \left[j_0(k||\boldsymbol{r} - \boldsymbol{r}_1||), \cdots, j_0(k||\boldsymbol{r} - \boldsymbol{r}_N||)\right]^{\top}, \tag{8}$$

[2] Lasso-based pressure matching introduces L1 regularization in (3) for obtaining sparse solution [11], [12].

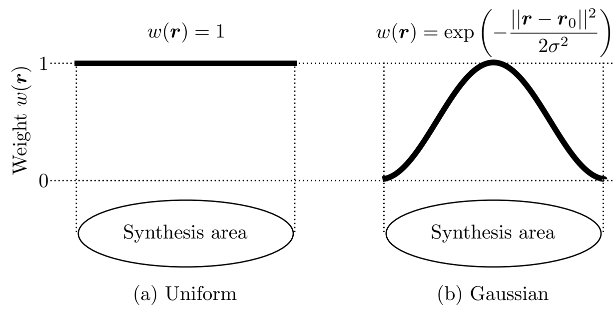

Fig. 1: Spatial weighting functions where $\boldsymbol{r}_0$ is the center of the synthesis area.

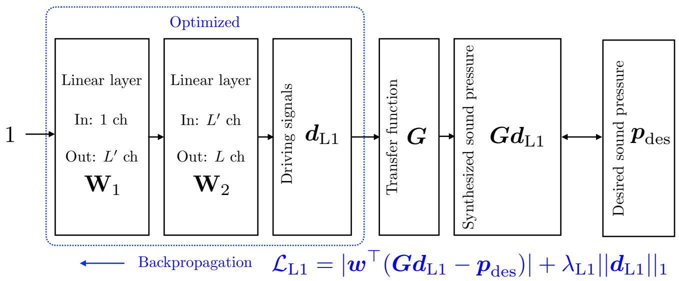

Fig. 2: Proposed method to obtain driving signals for least absolute deviation regression-based sound field control with backpropagation.

where k and j_0 are the wavenumber and 0-th order spherical Bessel function [20], respectively. λ is also a weighting parameter. Then, the driving signals of weighted pressure matching method with LS regression is also solved as a closed-form:

$$\boldsymbol{d}_{\mathrm{WPM}} = \left(\boldsymbol{G}^{\mathsf{H}}\boldsymbol{W}_{\mathrm{PM}}\boldsymbol{G} + \lambda_{\mathrm{PM}}\boldsymbol{I}\right)^{-1}\boldsymbol{G}^{\mathsf{H}}\boldsymbol{W}_{\mathrm{PM}}\boldsymbol{p}_{\mathrm{des}}. \tag{9}$$

By introducing (9) instead of (4), the control accuracy can be improved at higher frequencies. Weighted pressure matching can be regarded as a special case of weighted mode matching with spherical harmonic spectra estimated with infinite-dimensional harmonic analysis [21].

3. PROPOSED LEAST ABSOLUTE DEVIATION-BASED PRESSURE MATCHING METHOD: SFC-L1

3.1. Pressure matching with least absolute deviation

In the proposed method, the cost function $\mathcal{J}$ in (2) is replaced with LAD regression. Additionally, the spatial weighting function $w(\boldsymbol{r})$ used in weighted mode matching [14], such as uniform or Gaussian (Fig. 1),[3] is also introduced in the cost function. Then, the cost function of the proposed method is defined as:

$$\underset{\boldsymbol{d}}{\text{minimize}}\ \mathcal{J} = \int_{\boldsymbol{r}\in V}\left|w(\boldsymbol{r})\left(\boldsymbol{g}(\boldsymbol{r})^{\top}\boldsymbol{d} - p_{\mathrm{des}}(\boldsymbol{r})\right)\right|\,\mathrm{d}\boldsymbol{r}. \tag{10}$$

Compared with (2), (10) cannot be directly solved because it is not a closed-form. Some methods to solve the LAD regression have been investigated, such as simplex-based methods including Barrodale-Roberts algorithm [22], iteratively re-weighted least squares [23], Wesolowsky's direct descent method [24], Li-Arce's maximum likelihood approach [25], and recursive reduction of dimensionality approach [26].

[3]The spatial weighting function $w(\boldsymbol{r})$ can also be directly applied to pressure matching and weighted pressure matching.

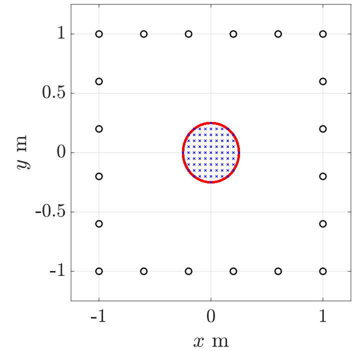

Fig. 3: Aarrangements of 20 loudspeakers (black circles) and 81 control points (blue cross marks). Red circle is control target region V.

3.2. Least absolute deviation regression with PyTorch

Although the previous methods to solve the LAD regression have been investigated, (10) can also be easily solved by gradient method with backpropagation [27], which can be simply implemented on a Python-based deep learning toolkit, PyTorch [28]. PyTorch can directly treat complex-valued optimization. In the proposed method, two linear layers with complex weight matrices $\mathbf{W}_1^{\in(1\times L')}$ and $\mathbf{W}_2^{\in(L'\times L)}$ without biases are introduced to obtain the proposed driving signals $\boldsymbol{d}_{\mathrm{L1}}$.[4] Then, the loss function to solve (10) is defined as:

$$\mathcal{L}_{\mathrm{L1}} = |\boldsymbol{w}^{\top}(\boldsymbol{G}\boldsymbol{d}_{\mathrm{L1}} - \boldsymbol{p}_{\mathrm{des}})| + \lambda_{\mathrm{L1}}||\boldsymbol{d}_{\mathrm{L1}}||_1, \tag{11}$$

where $\boldsymbol{w} = [w(\boldsymbol{r}_1), \cdots, w(\boldsymbol{r}_N)]^{\top}$, and λ_{L1} is also a weighting parameter. Compared to (3) and (5) with L2 regularization, (11) introduces L1 regularization rather than L2 regularization. This is because L1 regularization can obtain more stable solution at low frequencies according to the results of preliminary computer simulations. By updating the complex weight matrices $\mathbf{W}_1^{\in(1\times L')}$ and $\mathbf{W}_2^{\in(L'\times L)}$, $\boldsymbol{d}_{\mathrm{L1}}$ can be optimized (Fig. 2). By introducing the proposed method with LAD regression, improving the control accuracy at higher frequencies can be expected.

By replacing the loss function in (11), other sound field control methods can also be easily solved, such as Lasso-LS-based pressure matching [11], [12], and multiple sound spot synthesis [29]–[33]. Investigating sound field control with backpropagation for other tasks is future work.

4. COMPUTER SIMULATIONS

4.1. Simulation conditions

Computer simulations were performed to evaluate the proposed approach and compare it with the conventional LS-based methods. All the simulations were conducted on Python and PyTorch. In all the simulations, a 3D free-field was assumed and the speed of sound c was 343.36 m/s. Then, the transfer function between a loudspeaker at $\boldsymbol{r}_l$ and a control point $\boldsymbol{r}_n$ is three-dimensional free-field Green's function given as [20]:

$$G_{\mathrm{3D}}(\boldsymbol{r}_l, \boldsymbol{r}_n) = \frac{e^{jk|\boldsymbol{r}_l - \boldsymbol{r}_n|}}{4\pi|\boldsymbol{r}_l - \boldsymbol{r}_n|}. \tag{12}$$

In the simulations, the primary sound field was three-dimentional sound field propagated from a point source located at $x = 4$ m, $y = 5$ m, $z = 0$ m. Then, 2.5-dimensional sound field control [5], [7] with a square loudspeaker array on the horizontal plane is considered. 20

[4]Using two layers can realize higher control accuracy at low frequencies than using one layer according to the results of preliminary simulations.

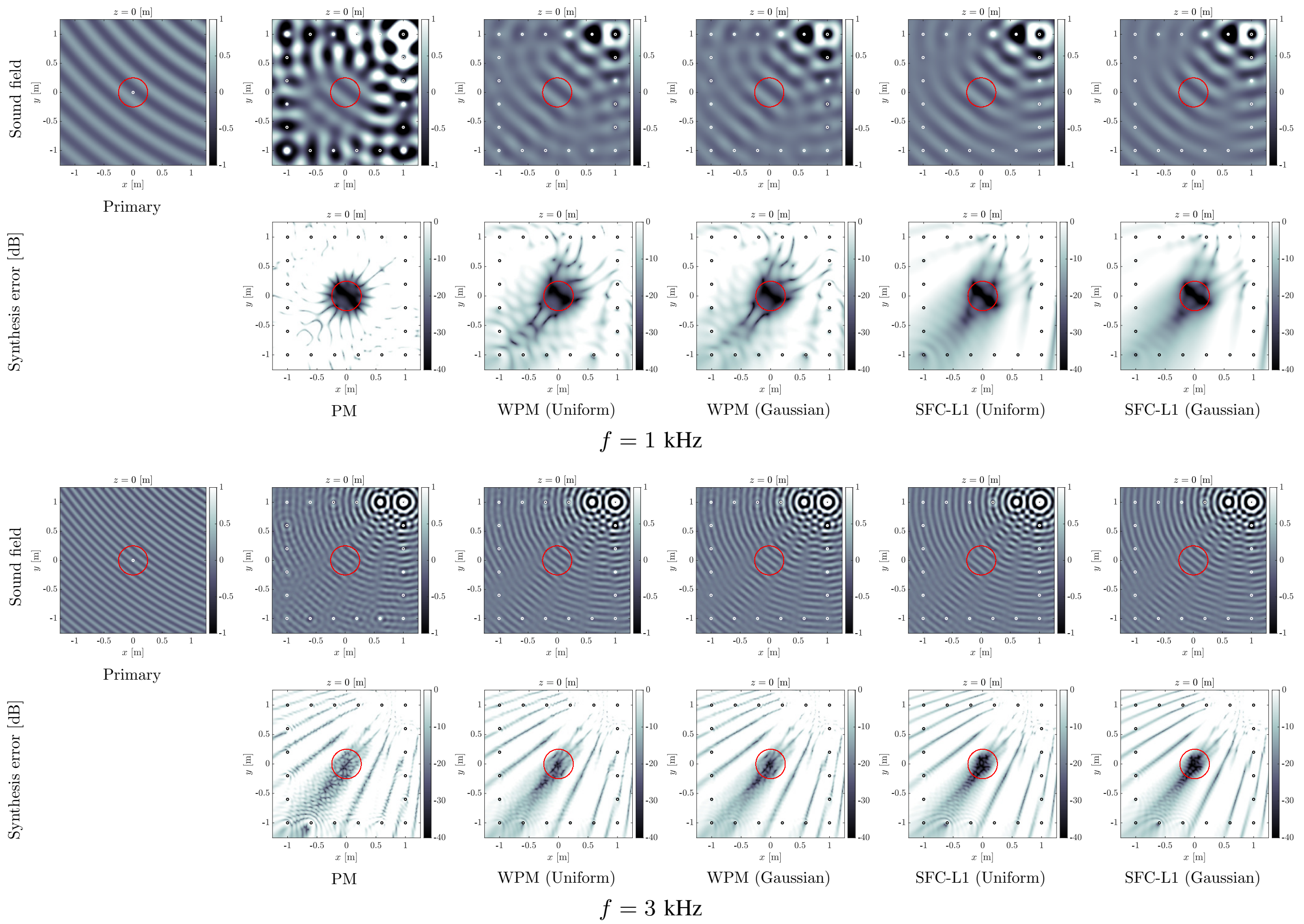

Fig. 4: Results of original and synthesized sound field (upper) and synthesis error (lower) for $f = 1$ kHz and 3 kHz, respectively. PM and WPM are conventional pressure matching and weighted pressure matching with least squares regression. SFC-L1 is proposed pressure matching with least absolute deviation regression. Uniform and Gaussian are spatial weighting windows. White and red circles are loudspeakers and control regions.

loudspeakers and 81 control points (= microphone array) were located on the horizontal plane ($z = 0$). The arrangements of loudspeakers and control points are shown in Figure 3. The control target region V is inside a circle with a radius of 0.5 m.

λ_{PM} in (4), λ in (6), λ_{WPM} in (9) and λ_{L1} in (11) were set to $\sigma_{\max}(\boldsymbol{G}^{\mathsf{H}}\boldsymbol{G}) \times 10^{-5}$, $\sigma_{\max}(\boldsymbol{\Sigma}) \times 10^{-5}$, $\sigma_{\max}(\boldsymbol{G}^{\mathsf{H}}\boldsymbol{W}_{\mathrm{PM}}\boldsymbol{G}) \times 10^{-5}$, and 10^{-5}, respectively. $\sigma_{\max}(\cdot)$ denotes the maximum eigenvalue. For weighted pressure matching and proposed SFC-L1, spatial weighting functions (uniform and Gaussian with $\sigma = 0.2$) in Fig. 1 were introduced. In weighted pressure matching, $\int_{\boldsymbol{r}\in V} \boldsymbol{\kappa}(\boldsymbol{r})^{*}\boldsymbol{\kappa}(\boldsymbol{r})^{\top} d\boldsymbol{r}$ in (6) was numerically calculated where $\boldsymbol{r}$ was discretized at 0.01 m intervals. j_0 in (7) and (8) was calculated using SciPy (https://scipy.org). In the proposed SFC-L1, Adam optimizer [34] with a learning rate of 10^{-2} was introduced to minimize $\mathcal{L}_{\mathrm{L1}}$ in (11). L' and L in $\mathbf{W}_1^{\in(1\times L')}$ and $\mathbf{W}_2^{\in(L'\times L)}$ were 20. Although the number of parameter updates for calculating (11) was 4,000, it can be calculate in about 0.6 seconds by using Apple MacBook Air M2 2023.

As the evaluation criteria, the synthesis error at $\boldsymbol{r}$ was introduced:

$$E(\boldsymbol{r}) = 10\log_{10}\frac{|p_{\mathrm{des}}(\boldsymbol{r}) - p_{\mathrm{syn}}(\boldsymbol{r})|^2}{|p_{\mathrm{des}}(\boldsymbol{r})|^2}. \tag{13}$$

The L2 norm of the driving signals $||\boldsymbol{d}||_2$ was also measured to evaluate the stability. These criteria were evaluated up to $f = 3$ kHz.

4.2. Results

Figure 4 shows the results of synthesized sound field and synthesis error for $f = 1$ kHz and 3 kHz, respectively. Additionally, Figure 5 shows the results of the averaged synthesis error inside the control region V. Furthermore, Figure 6 shows the results of the L2 norm of the driving signals. The results of Figures 4 and 5 indicated that the proposed methods with uniform and Gaussian weights outperform the conventional pressure matching and weighted pressure matching for $f > 1.5$ kHz in terms of the synthesis accuracy. Additionally, the results of Fig. 6 suggested the stability of the proposed method. Consequently, the effectiveness of the proposed SFC-L1 with LAD regression solved by backprppagation is validated for high frequencies.

5. CONCLUSION

To improve the control accuracy for pressure matching-based methods at high frequencies, this paper proposed SFC-L1, pressure matching-based sound field control methods with LAD regression instead of LS regression. In the proposed method, the LAD regression combined with L1 regularization is solved with gradient method simply implemented on PyTorch. The results of computer simulations showed that the proposed LAD-based methods can improve the sound field control accuracy at high frequencies compared with the conventional LS-based methods. PyTorch-based implementation, Torch-SFC, is open-sourced for accelerating sound field control research.

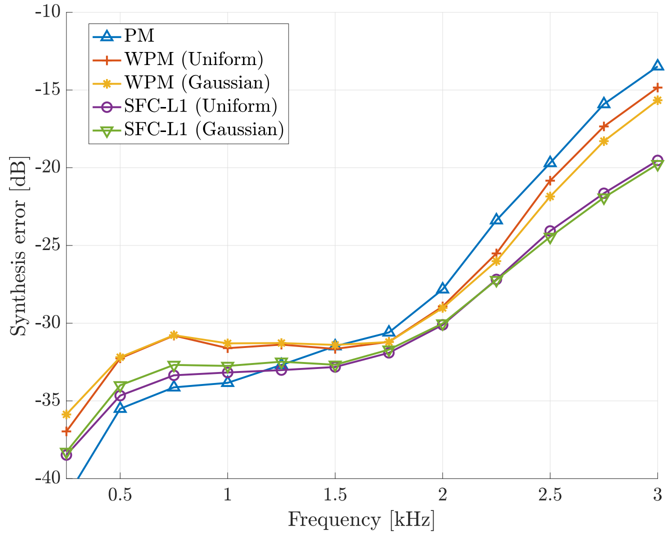

Fig. 5: Results of averaged synthesis error inside control region V. PM and WPM are conventional pressure matching and weighted pressure matching with least squares regression. SFC-L1 is proposed pressure matching with least absolute deviation regression. Uniform and Gaussian are spatial weighting windows.

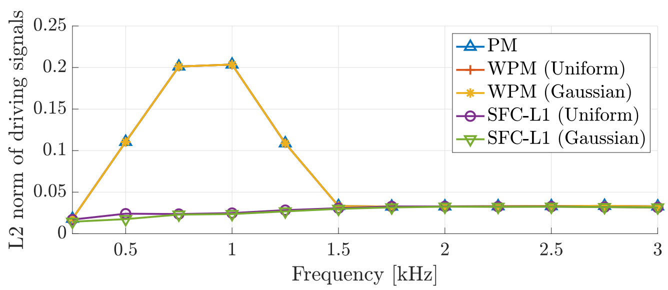

Fig. 6: Results of L2 norm of driving signals $||\boldsymbol{d}||_2$. PM and WPM are conventional pressure matching and weighted pressure matching with least squares regression. SFC-L1 is proposed pressure matching with least absolute deviation regression. Uniform and Gaussian are spatial weighting windows.

REFERENCES

[1] A. J. Berkhout, "A holographic approach to acoustic control," *J. Audio. Eng. Soc.*, vol. 36, no. 12, pp. 977–995, Dec. 1988.

[2] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764–2778, May 1993.

[3] S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," in *Proc. 124th Conv. Audio Eng. Soc.*, May 2008.

[4] J. Ahrens and S. Spors, "Analytical driving functions for higher order Ambisonics," in *Proc. ICASSP*, Mar. 2008, pp. 373–376.

[5] T. Okamoto, "Analytical approach to 2.5D sound field control using a circular double-layer array of fixed-directivity loudspeakers," in *Proc. ICASSP*, Mar. 2017, pp. 91–95.

[6] ——, "Horizontal 3D sound field recording and 2.5D synthesis with omni-directional circular arrays," in *Proc. ICASSP*, May 2019, pp. 960–964.

[7] J. Ahrens and S. Spors, "Sound field reproduction using planar and linear arrays of loudspeakers," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 8, pp. 2038–2050, Nov. 2010.

[8] T. Okamoto, "Angular spectrum decomposition-based 2.5D higher-order spherical harmonic sound field synthesis with a linear loudspeaker array," in *Proc. WASPAA*, Oct. 2017, pp. 180–184.

[9] O. Kirkeby and P. Nelson, "Reproduction of plane wave sound fields," *J. Acoust. Soc. Am.*, vol. 94, no. 5, pp. 2992–3000, Nov. 1993.

[10] T. Betlehem and T. D. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," *J. Acoust. Soc. Am.*, vol. 117, no. 4, pp. 2100–2111, Apr. 2005.

[11] G. N. Lilis, D. Angelosante, and G. B. Giannakis, "Sound field reproduction using the Lasso," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 8, pp. 1902–1912, Nov. 2010.

[12] N. Radmanesh and I. S. Burnett, "Generation of isolated wideband sound fields using a combined two-stage Lasso-LS algorithm," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 2, pp. 378–387, Feb. 2013.

[13] M. A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *J. Audio Eng. Soc.*, vol. 53, no. 11, pp. 1004–1025, Nov. 2005.

[14] N. Ueno, S. Koyama, and H. Saruwatari, "Three-dimensional sound field reproduction based on weighted mode-matching method," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 27, no. 12, pp. 1852–1867, Dec. 2019.

[15] S. Koyama, K. Kimura, and N. Ueno, "Weighted pressure and mode matching for sound field reproduction: Theoretical and experimental comparisons," *J. Audio Eng. Soc.*, vol. 71, no. 4, pp. 173–185, Apr. 2023.

[16] J. Zhang, W. Zhang, T. D. Abhayapala, and L. Zhang, "2.5D multizone reproduction using weighted mode matching: Performance analysis and experimental validation," *J. Acoust. Soc. Am.*, vol. 147, no. 3, pp. 1404–1417, Mar. 2020.

[17] T. Okamoto, "2D multizone sound field synthesis with interior-exterior Ambisonics," in *Proc. WASPAA*, Oct. 2021, pp. 276–280.

[18] T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," *Geosci. Model Dev.*, vol. 15, pp. 5481–5487, July 2022.

[19] N. Ueno, S. Koyama, and H. Saruwatari, "Kernel ridge regression with constraint of Helmholtz equation for sound field interpolation," in *Proc. IWAENC*, Sept. 2018, pp. 436–440.

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

[21] N. Ueno, S. Koyama, and H. Saruwatari, "Sound field recording using distributed microphones based on harmonic analysis of infinite order," *IEEE Signal Process. Lett.*, vol. 25, no. 1, pp. 135–139, Jan. 2018.

[22] I. Barrodale and F. D. K. Roberts, "An improved algorithm for discrete l_1 linear approximation," *SIAM J. Numer. Anal.*, vol. 10, no. 5, pp. 839–848, Oct. 1973.

[23] E. J. Schlossmacher, "An iterative technique for absolute deviations curve fitting," *J. Am. Stat. Assoc.*, vol. 68, no. 334, pp. 857–859, Dec. 1973.

[24] G. O. Wesolowsky, "A new descent algorithm for the least absolute value regression problem: A new descent algorithm for the least absolute value," *Commun. Stat. Simulat.*, vol. 10, no. 5, pp. 479–491, Oct. 1981.

[25] Y. Li and G. R. Arce, "Maximum likelihood approach to least absolute deviation regression," *EURASIP J. Adv. Signal Process.*, vol. 2004, no. 12, p. 948982, Sept. 2004.

[26] A. S. Kržić and D. Seršić, "L1 minimization using recursive reduction of dimensionality," *Signal Process.*, vol. 151, pp. 119–129, Oct. 2018.

[27] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, Oct. 1986.

[28] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "PyTorch: An imperative style, high-performance deep learning library," in *Proc. NeurIPS*, Dec. 2019, pp. 8024–8035.

[29] J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *J. Acoust. Soc. Am.*, vol. 111, no. 4, pp. 1695–1700, Apr. 2002.

[30] T. Okamoto, "Analytical methods of generating multiple sound zones for open and baffled circular loudspeaker arrays," in *Proc. WASPAA*, Oct. 2015.

[31] T. Okamoto and A. Sakaguchi, "Experimental validation of spatial Fourier transform-based multiple sound zone generation with a linear loudspeaker array," *J. Acoust. Soc. Am.*, vol. 141, no. 3, pp. 1769–1780, Mar. 2017.

[32] T. Lee, J. K. Nielsen, and M. G. Christensen, "Signal-adaptive and perceptually optimized sound zones with variable span trade-off filters," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 2412–2426, 2020.

[33] T. Okamoto and M. Kono, "Simultaneous speech translation integrated compact multiple sound spot synthesis system on a laptop carried out with a backpack," in *Proc. Interspeech*, Aug. 2025.

[34] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, May 2015.