PROCEEDINGS of the
23rd International Congress on Acoustics
9 to 13 September 2019 in Aachen, Germany

[Pressure-matching-based 2D sound field synthesis with equivalent source array]

Izumi TSUNOKUNI[(1)], Kakeru KUROKAWA[(2)], Yusuke IKEDA[(3)]

[(1)]Tokyo Denki University, Japan, 19fmi23@ms.dendai.ac.jp
[(2)]Tokyo Denki University, Japan, 19fmi13@ms.dendai.ac.jp
[(3)]Tokyo Denki University, Japan, yusuke.ikeda@mail.dendai.ac.jp

Abstract
Pressure-matching(PM) method is one of well-known effective methods for physical sound field synthesis using a massive loudspeaker array. In the PM method, all transfer functions between loudspeakers and matching points have to be measured so as to control the sound pressure at each corresponding matching point. Thus, as the number of loudspeakers and matching points increases, it becomes more difficult to measure transfer functions using a microphone. In addition, it is difficult to achieve high accurate placement of a microphone without a robot or microphone array. In this paper, we proposed an efficient method of local sound field synthesis with pressure matching. This method does not require large-scale measurements as it estimates the transfer functions between loudspeakers and dense matching points from transfer functions between the loudspeakers and a small number of points. By sparse optimization, each set of equivalent sources of a loudspeaker is selected from a dictionary of mono-pole sources placed near each loudspeaker based on a small number of transfer functions.The driving functions of loudspeakers can be derived from transfer functions between numerous virtual matching points and all equivalent sources. To evaluate our proposed method, two-dimensional simulation experiments are conducted. Keywords: Sparse modeling, Impulse response, Transfer function

1 INTRODUCTION

Pressure Matching(PM) is one of sound field synthesis techniques to physically reproduce a sound field by matching a sound pressure at each matching point[1]. Generally, numerous matching points with higher density are required to control sound field at higher frequency. As the number of matching points increases, it takes more time and effort to obtain all transfer functions. In most of physical sound field synthesis(SFS) methods, transfer functions between loudspeakers and control points are also required, especially when compensating the characteristics of loudspeakers and the sound reflections in the room.

To measure the sound field, measurement methods with a moving microphone have been proposed[2, 3]. To move a microphone accurately, a robot or a moving instruments is required. In most cases, the size of instruments to move a microphone is not small enough to ignore the sound reflections by itself. In addition, moving instruments must be sufficiently silent to measure the transfer functions. In [4], a method of interpolation for room impulse responses in a whole room has been proposed under the assumption of sparsity in the early part of reverberation. However, the large number and complex arrangement of microphones are required. In most cases, transfer functions in a whole room is not required to control sound field in a specified area.

In this paper, we propose local sound field control method with an estimation for the direct-sound transfer functions between loudspeakers and arbitrary points in a local area. Assuming sparsity of the sound source, the wave field can be represented by superposition of the sparse point sources[5, 6]. In our proposed method, sparse equivalent sources and its weight coefficients are derived to estimate the transfer functions between a loudspeaker and a specified area from impulse responses measured at a small number of points. In this paper, estimated transfer function are utilized for PM method to synthesis sound field in a local area. In simulation experiment, we evaluate the errors of the estimated transfer functions in the specified area. Then, we evaluate reproduction errors of PM using the estimated transfer functions by 2D simulation.

2 PROPOSED METHOD

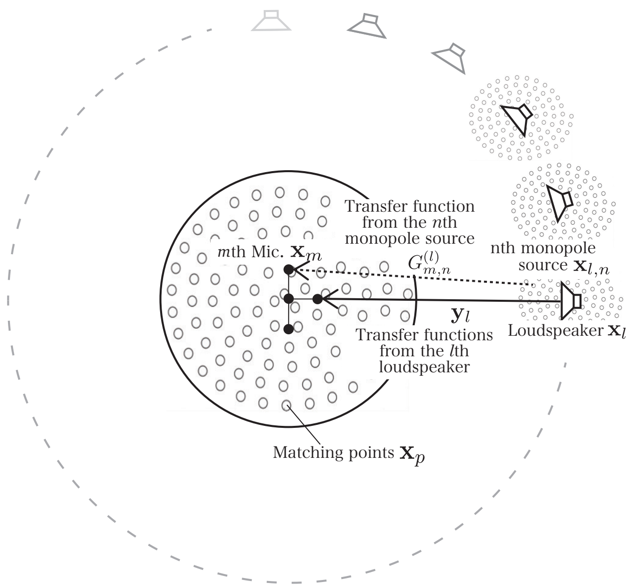


Figure 1. Concept of proposed reproduction method with equivalent sources. $\mathbf{y}_l$ is a transfer function vector from l-th loudspeaker to measurement points.

To reduce the cost of sound field measurement for SFS, we will obtain the transfer functions at an arbitrary point in the local region from the measurement at a small number of points. Especially, the direct sound including frequency characteristics and directivity of loudspeaker is important to control sound field. Now, consider modeling the direct sound component of the transfer functions from a loudspeaker with equivalent sources. Note that the direct sound includes reflections from a loudspeaker by itself, since it is difficult to divide them.

Figure 1 shows the concept of our proposed reproduction method with equivalent sources. Firstly, it is assumed that the approximate positions of the loudspeakers $\mathbf{x}_l$ relative to the microphones are known.

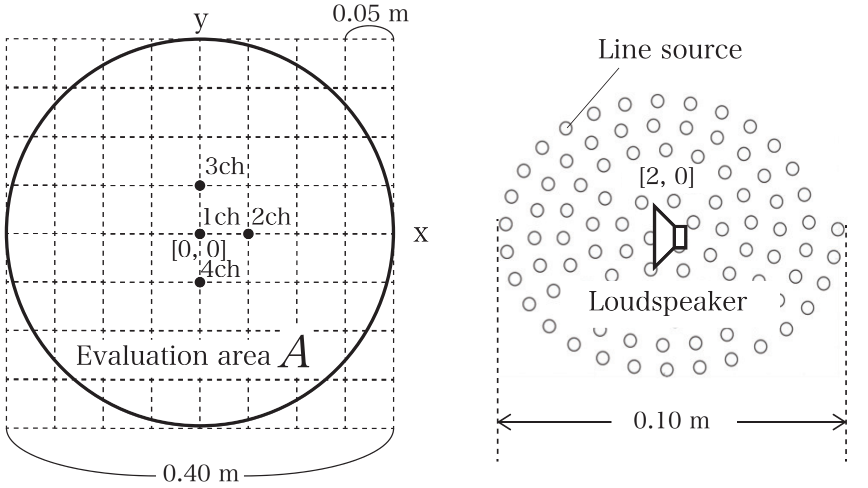
Consider that measured transfer functions are represented by the superposition of multiple point sources $\mathbf{x}_{l,n}$, which are placed in the area surrounding the expected position of a loudspeaker. In the frequency domain, the transfer function vector $\mathbf{y}(\in \mathbb{C}^{M\times 1})$ at the M measurement points is given by

$$\mathbf{y}_l = \mathbf{G}_l \mathbf{w}_l \ \ (l = 1,2,\ldots,L). \tag{1}$$

$\mathbf{G}_l = [G^{(l)}{}_{m,n}](\in \mathbb{C}^{M\times N})$ denotes a Green's function matrix between the M measurement points and N point sources surrounding the position of lth loudspeaker, and $\mathbf{w}_l(\in \mathbb{C}^{N\times 1})$ denotes the weight coefficient vector for N point sources. Since each group of point sources includes just one loudspeaker, sparsity of the point sources can be assumed. Thus, the weight coefficient vector $\mathbf{w}_l$ can be derived by solving the following optimization problem:

$$\underset{\mathbf{w}_l}{\text{minimize}} \quad \frac{1}{2}||\mathbf{y}_l - \mathbf{G}_l\mathbf{w}_l||_2^2 + \lambda||\mathbf{w}_l||_1, \tag{2}$$

where λ and $||\cdot||_1$ are the penalty parameter and L_1-norm, respectively. When the elements of weight coefficient $\mathbf{w}_l$ is sufficiently small, the value is considered to be zero. Subsequently, the coefficient is compensated by using the transfer function at a single measurement point. If the transfer function of direct sound from the loudspeaker is well-modeled by the non-zero weight coefficient w_s and the corresponding point sources $\mathbf{x}_s$, the



(a) Arrangement of the microphone

(b) Position of a loudspekaer and line sources

Figure 2. Arrangement in single loudspeaker simulation. A loudspeaker is positioned at $[2,0]$. The dictionary of line sources in the area with a radius of 0.10 m is placed along with the loudspeaker. The number of line sources is 300.

transfer function at any points in the area A is estimated as follows,

$$\hat{G}(\mathbf{x}_s, \mathbf{x}, \omega) = -\frac{1}{4\pi} \sum_{s=1}^{S} w_s H_0^{(2)}(k|\mathbf{x} - \mathbf{x}_s|) \ \ (\mathbf{x} \in A), \tag{3}$$

where k denotes the wave number, $H_0^{(2)}$ denotes the Hankel function of 0-th order and second kind. S is the number of sparse point sources, which has a non-zero corresponding coefficient. In the following section,2D simulation was conducted by using 2D Green's function as shown in Eq.(**??**).

In PM method, a sound pressures $P(\mathbf{x}_m, \omega)$ by line sources at $\mathbf{x}_l$ in 2D situation is represented with driving function of l-th loudspeaker $D(\mathbf{x}_l, \omega)$ and a transfer function $G(\mathbf{x}_l, \mathbf{x}_p, \omega)$ from l-th loudspeaker $\mathbf{x}_l$ to matching points $\mathbf{x}_p$ as following[1],

$$P(\mathbf{x}_p, \omega) = \sum_{l=1}^{L} D(\mathbf{x}_l, \omega) G(\mathbf{x}_l, \mathbf{x}_p, \omega), \tag{4}$$

$$G(\mathbf{x}_l, \mathbf{x}_p, \omega) = H_0^{(2)}(k|\mathbf{x}_l - \mathbf{x}_p|) \tag{5}$$

. By using the modeled transfer function $\hat{G}$ instead of the transfer function G, Eq.(4) is changed as follows,

$$P(\mathbf{x}_p, \omega) = \sum_{l=1}^{L} D(\mathbf{x}_l, \omega) \hat{G}(\mathbf{x}_l, \mathbf{x}_p, \omega), \tag{6}$$

The desired sound field $P(\mathbf{x}_s, \mathbf{x}_p, \omega)$ from a line source $\mathbf{x}_{\mathrm{src}}$ is given by

$$P(\mathbf{x}_{\mathrm{src}}, \mathbf{x}_p, \omega) = H_0^{(2)}(k|\mathbf{x}_l - \mathbf{x}_p|). \tag{7}$$

The vector of solution of driving function vector $\mathbf{d} = [D(\mathbf{x}_l)] \in \mathbb{C}^{L \times 1}$ can be derived by least squares method as following,

$$\mathbf{d} = [\hat{\mathbf{G}}^{\mathrm{H}} \hat{\mathbf{G}} + \rho \mathbf{I}]^{-1} \hat{\mathbf{G}}^{\mathrm{H}} \mathbf{p} \tag{8}$$

where ρ is a regularization parameter, $\hat{\mathbf{G}}$ describes the matrix of transfer functions between equivalent sources and matching points. $\mathbf{I}$ is identity matrix. $[\cdot]^{\mathrm{H}}$ is complex conjugate transposition.

The number of matching points can be freely increased with higher dense to obtain driving functions. In addition, this method is able to apply to sound field synthesis techniques matching the sound pressure at the points such as BoSC.

3 SIMULATION EXPERIMENTS

3.1 Conditions

We conducted simulation of a sound field in a 2D free-field condition. Firstly, the errors of estimated and desired transfer functions were compared. Secondly, the accuracy of synthesized sound field by estimated transfer function and the desired Green's function was compared. The directivity of the loudspeaker is unidirectional i.e., toward the origin, which is given by $(1+\cos\theta)/2$ where θ is angle between the directions of the loudspeaker and measurement position.

To model each loudspeaker, 300 line sources were placed in a circle with a radius of 0.05 m. The penalty parameter is $\lambda = 10^{-4}$ for sparse optimization.

To evaluate reproduction accuracy, the average of reproduction error in the area A with a radius of 0.20 m was defined by

$$\mathrm{Error}(\omega) = 10\log_{10}\frac{\sum_{i=1}^{I}|P_{\mathrm{d}}(\mathbf{x}_i,\omega)-P_{\mathrm{s}}(\mathbf{x}_i,\omega)|^2}{\sum_{i=1}^{I}|P_{\mathrm{d}}(\mathbf{x}_i,\omega)|^2} \tag{9}$$

where P_{d} and P_{s} denote the desired complex sound pressure and the estimated complex sound pressure, respectively. I is the number of evaluation points inside a circular area A.

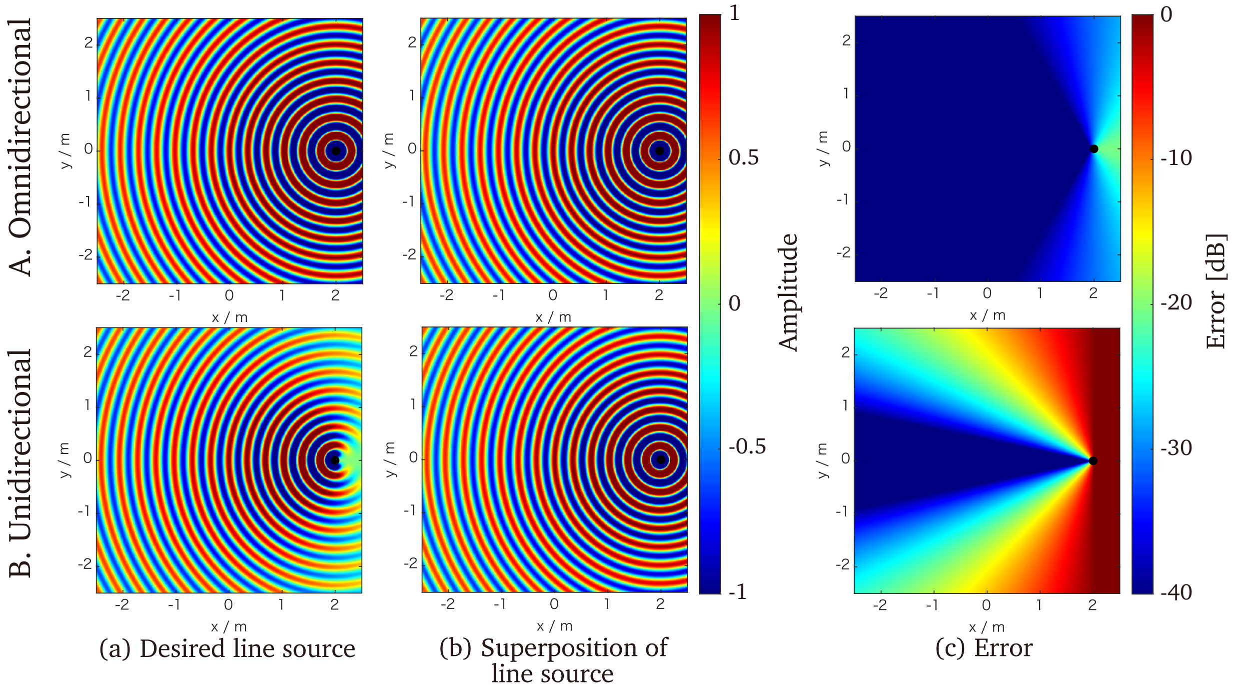


Figure 3. The wave front represented by sparse line source, $f = 1000$ Hz. In the upper figures, original source is an omnidirectional line source in horizontal plane. In the lower figures, original source is a unidirectional line source.

Table 1. The errors of estimated transfer function and desired transfer function

Frequency [kHz]	1	2	3	4	5	6	7	8
Error [dB]	-54.5	-56.5	-57.8	-59.9	-59.7	-55.6	-58.6	-58.4
The number of line source	19	20	18	8	9	7	6	7

3.2 Estimation of transfer function

The microphones were placed in the interval of 0.05 m at around the origin, and a loudspeaker was placed at $[2, 0]$. Fig.2 shows an arrangement of instruments for simulation experiment. The number of line sources as the dictionary is 300. Table 1 shows the errors of estimated transfer function of direct sound using four microphones. The errors were very small at about -55–-60 dB for 1–8 kHz. Thus, direct sound in the area around the origin were accurately modeled by superposition of sparse line sources.

Figure3 shows the wave front and the errors of estimated transfer function. As shown in Fig.3, omnidirectional line source can be accurately modeled. In contrary, when the unidirectional line source was modeled, the size of area that transfer functions can be accurately estimated becomes smaller. Thus, the desired sound source is closer to line source (or point source), the accuracy improves. Thus, when the source has a directivity, the accuracy of estimated transfer function is lower.

3.3 Synthesized sound field using estimated transfer function

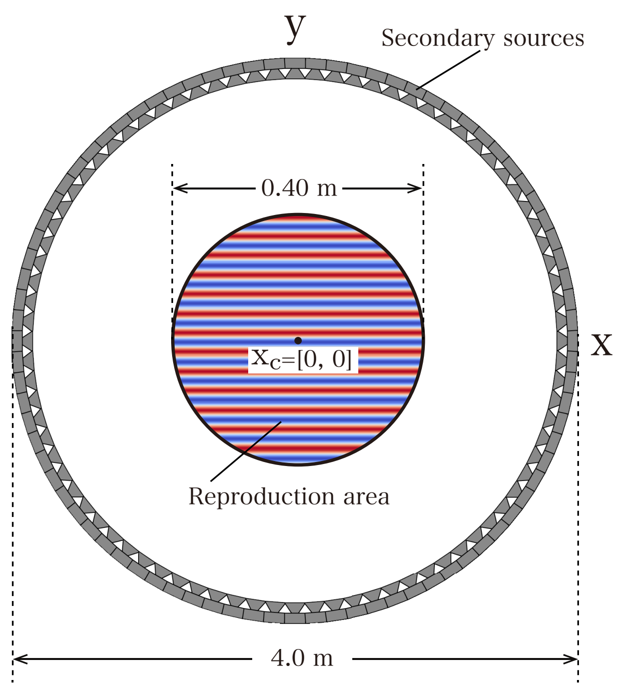


Figure 4. Condition of sound field synthesis by PM method. Sound field is synthesized by a circular array with radius of 2 m. The number of loudspeaker is 84.

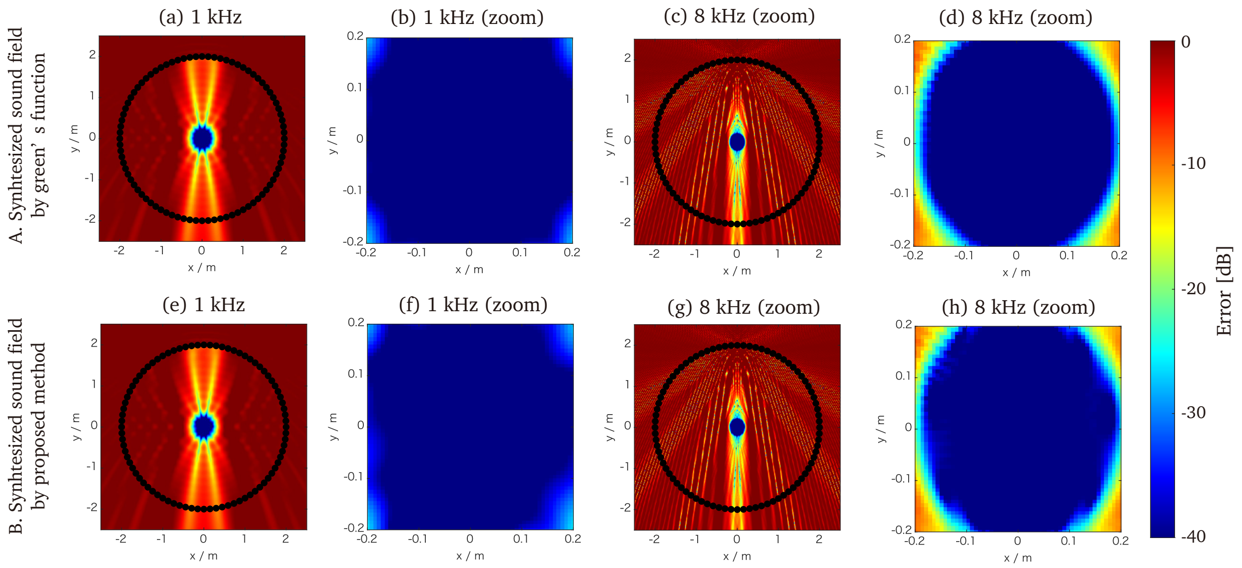


Figure 5. Errors of synthesized sound field : (a)–(d) sound field was synthesized by desired transfer function, and (e)–(h) sound field by estimated transfer function. Figs (a), (c), (e) and (g) show the errors of synthesized sound field. Figs (b), (d), (f) and (h) are expansion of left figure.

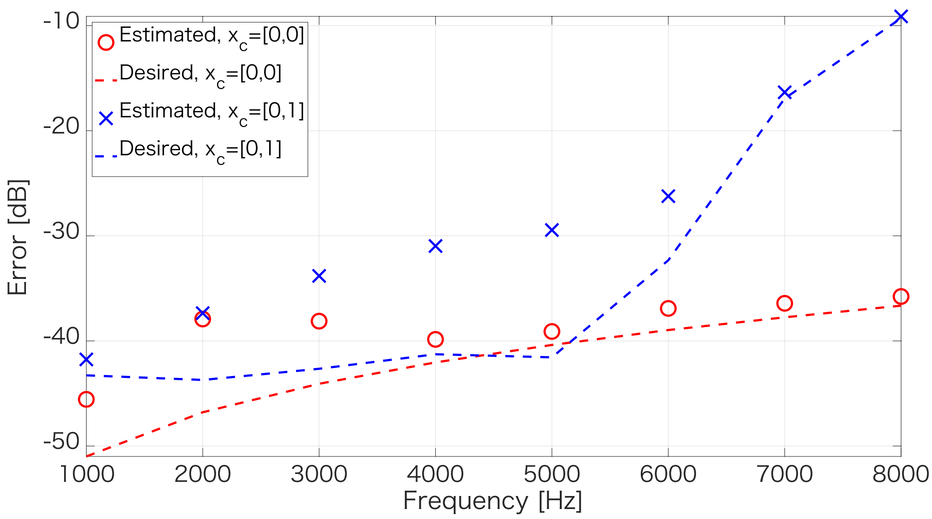


Figure 6. Comparison of errors with PM method: Red circle shows the errors of synthesized sound field by using modeled transfer functions when the center of reproduced area is at $x_c = [0, 0]$. Red dashed line shows the errors at $x_c = [0, 0]$ by using desired transfer functions. Blue cross and dashed line show the errors when the center of reproduction area is at $x_c = [0, 1]$.

Sound field was reproduced by PM method by using a circular loudspeaker array with an interval of 0.15 m. Sound field was controlled at 500 matching points. In this simulation, the plane wave propagating in direction negative of y-axis was synthesized in circle of radius 0.2 m (Fig.4). The error maps of synthesized sound field were illustrated in Fig.5. From the figure, the accuracy of synthesized sound field using modeled transfer function was extremely close to that using Green's function. Figure 6 shows the errors of synthesized sound field by modeled transfer function for 1–8 kHz. Comparing sound fields reproduced at $x_c = [0,0]$ by the proposed method and desired transfer functions, the difference is less than about 5 dB at 4 kHz or more. For 1–3 kHz, the differences are about 5–10 dB. When the center of reproduction area is at $x_c = [0,1]$, the errors becomes larger. However, this results was caused by lower accuracy of PM method.

4 CONCLUSIONS

In this paper, we proposed sound field synthesis method using estimated transfer function to reduce time and effort for implementation. The transfer functions between loudspeaker and any points in local area can be well-modeled by superposition of sparse line sources, which are calculated from a few measurement points. By our proposed method, the number of matching points to control the sound field can be arbitrarily increased. Thus, a small number of measurement points are only required, even when many matching points are required. However, when the sound source with directivity is modeled, the accuracy is lower than Omnidirectional sound source. In future works, we improve modeling a sound source with directivity.

ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Number 17KT0141.

REFERENCES

[1] Poletti. Mark A., "Improved methods for generating focused soures using circular arrays," *Audio Engineering Society Conventoin 133*, San Francisco, CA, USA, Oct.2012.

[2] Nara Hahn, and Sascha Spors, "Comparison of Continuous of Measurement Techniques for Spatial Room Impulse Response," *in Proc. of the EUSIPCO*, Nice, France, Aug. 2015.

[3] T. Ajdler, L. Sbaiz, and M. Vetterli, "Dynamic Measurement of Room Impulse Responses using a Moving Microphone," *J. Acoust. Soc. of America*, vol. 122, no. 3, pp. 1536–1645, 2007.

[4] R. Mignot, L. Daudet, and F. Ollivier, "Room Reverberation Reconstruction: Interpolation of the Early Part Using Compressed," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 11, pp. 2301–2312, Nov. 2013.

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

[6] N. P. Valdivia, and E. G. Williams, "Study of the comparison of the methods of equivalent sources and boundary element methods for near-field acoustic holography," *J. Acoust. Soc. of America*, vol. 120, no. 6, pp. 3694–3705, 2006.