JOINT SOURCE AND SENSOR PLACEMENT FOR SOUND FIELD CONTROL BASED ON EMPIRICAL INTERPOLATION METHOD

Shoichi Koyama[1,2]*, Gilles Chardon*[3]*, and Laurent Daudet*[2]

[1]The University of Tokyo, Graduate School of Information Science and Technology,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
[2]Institut Langevin, ESPCI, Université Paris Diderot, CNRS UMR 7587,
1 rue Jussieu, Paris 75005, France
[3]L2S, CentraleSupélec, CNRS, Université Paris-sud, Université Paris-Saclay,
3 rue Joliot-Curie, Gif-sur-Yvette 91192, France

ABSTRACT

This study proposes a principled method to jointly determine the placement of acoustic sources (loudspeakers) and sensors (control points/microphones) in sound field control. The goal of this setup is to efficiently produce a sound field using multiple loudspeakers, approximately matching a target sound field over a region of interest. Therefore, the loudspeaker and control-point placement problem can be seen as the problem of finding interpolating functions (associated with individual loudspeaker sound fields) and sampling points (corresponding to control points or microphones) to approximate the target sound field in the given domain. We here solve this problem using the empirical interpolation method, originally developed for the numerical analysis of partial differential equations. The proposed method enables a joint determination of loudspeaker and control-point placement, from a large set of candidate locations, independently of the desired sound field. Numerical simulation results indicate that accurate and stable sound field control can be achieved by the proposed method, with significantly better results than with random and regular placements.

Index Terms— source and sensor placement, sound field control, sound field reproduction, interpolation, magic points

1. INTRODUCTION

Sound field control is aimed at synthesizing a desired sound field inside a region of interest. It can be applied to various settings including high fidelity audio systems and noise cancellation systems. A typical strategy for the sound field control problem is the use of a distributed loudspeaker array, allowing the control of sound pressures at multiple discrete positions inside the region [1–3]. In general, the inverse of the given transfer function matrix between loudspeakers and control points is calculated in a (regularized) least-square-error sense. The positions of the loudspeakers and control points (microphones) must be carefully chosen because they have a great effect not only on the control accuracy but also on the stability of the inverse filter. For example, when the control points of the sound pressures are arranged on the boundary of the enclosed space, it is known that the sound field inside this space cannot be uniquely determined, which leads to significantly unstable inverse filters [4]. The estimated inverse filter can also be sensible to perturbations in the transfer functions.

In the context of sound field reproduction, a similar problem has been addressed by modeling a sound field in a continuous system [5–8]. The positions of the loudspeakers and microphones are generally determined by regularly discretizing the continuous space. The above-mentioned non-uniqueness problem is typically avoided by using microphones mounted on an acoustically-rigid object or directional microphones [7, 9–11]. Although this regular placement will perform well when the array has a simple shape, such as sphere, plane, line, and circle, that is not obvious for more complicated geometries. Besides, their performance depends on the acoustic characteristics of the sound field to be controlled.

As discussed above, in the sound field control and reproduction problems, how to place the loudspeakers and control points (or microphones) is still an open problem, especially for an arbitrary shape of the control region. Several attempts have been made to optimize the loudspeaker placement [12,13]. However, even though the placement of loudspeakers and control points both have a significant impact on performance, to the best of our knowledge their joint placement has not yet been investigated. Furthermore, current algorithms for loudspeaker placement are optimized on a single target sound field, or by averaging over a limited set of target sound fields - this hinders their practical applicability. Independently, the microphone placement problem has been addressed in the context of sensor array design [14, 15].

Here, we propose a method to jointly determine these positions from candidate positions independently of the desired sound field. We consider the loudspeaker and control point placement problem as a problem of finding interpolating functions and associated sampling points for approximating these transfer functions. This interpretation motivates us to apply a method called *empirical interpolation method (EIM)* [16, 17], which was originally proposed in the field of numerical analysis of partial differential equations. Numerical simulations are conducted to evaluate the proposed method, that is compared to regular and random sampling schemes in a two-dimensional (2D) sound field.

2. PROBLEM STATEMENT

We assume that the sound field inside a region of interest Ω is controlled by L loudspeakers [18]. The synthesized sound pressure $u_{\mathrm{syn}}(\mathbf{r}, \omega)$ of the frequency ω at the position $\mathbf{r}$ is represented by a linear combination of transfer functions of the loudspeakers.

$$u_{\mathrm{syn}}(\mathbf{r}, \omega) = \sum_{l=1}^{L} d_l(\omega) g_l(\mathbf{r}, \omega), \tag{1}$$

where $d_l(\omega)$ and $g_l(\mathbf{r},\omega)$ are the driving signal and the transfer function (i.e., Green's function) at $\mathbf{r}$ of the lth loudspeaker, respectively. Hereafter, ω is omitted for notational simplicity. By denoting the desired pressure field as $u_{\mathrm{des}}(\mathbf{r})$ ($\mathbf{r} \in \Omega$), the objective function $\mathcal{J}$ for the sound field control problem can be described as

$$\mathcal{J} = \int_{\mathbf{r}\in\Omega} \left| \sum_{l=1}^{L} d_l g_l(\mathbf{r}) - u_{\mathrm{des}}(\mathbf{r}) \right|^2 d\mathbf{r}. \tag{2}$$

Therefore, it is necessary to estimate d_l that minimizes $\mathcal{J}$.

However, this minimization problem is difficult to solve directly because (2) includes an integral with respect to $\mathbf{r}$: it is difficult to measure or estimate $g_l(\mathbf{r})$ in the entire Ω for continuous $\mathbf{r}$. A typical strategy for solving it is to discretize $\mathbf{r}$ inside Ω, which leads to the following linear equation:

$$\mathbf{u}^{\mathrm{des}} = \mathbf{G}\mathbf{d}, \tag{3}$$

where $\mathbf{u}^{\mathrm{des}} \in \mathbb{C}^M$ is the vector of the desired pressures at M discrete positions inside Ω, $\mathbf{G} \in \mathbb{C}^{M\times L}$ is the matrix of the transfer functions, and $\mathbf{d} \in \mathbb{C}^L$ is the vector of the driving signals. Then, $\mathbf{d}$ can be obtained by

$$\hat{\mathbf{d}} = \mathbf{G}^{\dagger}\mathbf{u}^{\mathrm{des}}, \tag{4}$$

where $(\cdot)^{\dagger}$ represents Moore-Penrose pseudoinverse. Since the calculation of the inverse of $\mathbf{G}$ becomes frequently unstable, it is usually necessary to regularize (4). For instance, the use of Tikhonov regularization is represented as

$$\hat{\mathbf{d}} = \left(\mathbf{G}^{\mathsf{H}}\mathbf{G} + \lambda\mathbf{I}\right)^{-1}\mathbf{G}^{\mathsf{H}}\mathbf{u}^{\mathrm{des}}, \tag{5}$$

where λ is the regularization parameter.

The question is how to determine the discrete positions and number of loudspeakers and control points (microphones). Since measuring the transfer functions by using microphones is necessary to obtain $\mathbf{G}$, it is preferable that the number of sampling positions is as small as possible. The number of available loudspeakers is also limited in general. Besides, excessively dense placement of the microphones and loudspeakers leads to extremely unstable inverse filters. Their positions also have a great effect on the control accuracy and filter stability. For example, the sampling only on the boundary of Ω should be avoided because of the non-uniqueness problem.

3. SOURCE AND SENSOR PLACEMENT USING MAGIC POINTS

We consider EIM for the above-mentioned problem. EIM was first presented in [16] in the context of numerical analysis of partial differential equations based on the reduced basis method [19]. In [17], EIM was discussed as a general interpolation procedure.

3.1. Empirical interpolation method in general case

Given a functional space $\mathcal{V}$ of large or infinite dimension defined on a domain Ω and an order Q, the EIM selects Q interpolating functions and Q sampling points in Ω (so-called *magic points*). An approximation of a function is then obtained by matching a linear combination of the interpolating functions with values of the function at the magic points, i.e., by solving a linear system of equations. Interpolating functions and sampling points are chosen such that this system remains stable when Q increases.

Following [17], the first interpolation function h_1 is chosen as $v \in \mathcal{V}$ with maximum L_∞-norm in Ω, and the first sampling point x_1 is chosen as the point where $|h_1|$ attains its bounds. The subsequent interpolating functions and sampling points are iteratively identified. Given a set of Q interpolating functions h_q and sampling points x_q, the next interpolating function h_{Q+1} and the next sampling point x_{Q+1} are found by the following greedy algorithm:

1. The interpolation $I_Q(v)$ for any $v \in \mathcal{V}$ is computed using h_q and x_q identified so far. This interpolation is obtained by solving the following equation:

$$I_Q(v) = \sum_{q=1}^{Q} c_q h_q, \tag{6}$$

   where c_q is the solutions of

$$v(x_q) = \sum_{q'=1}^{Q} c_{q'} h_{q'}(x_q), \tag{7}$$

   for $1 \le q \le Q$.
2. v that maximizes the L_∞-norm of the error between v and its interpolation $I_Q(v)$ is taken as h_{Q+1}.
3. The point of the maximal absolute value of the error between $v(x)$ and its interpolation $I_Q(v)$ is taken as x_{Q+1}.

This procedure is repeated until the error between any function of $\mathcal{V}$ and its interpolation is less than a predefined threshold value in L_2-norm, or the order Q reached to the predefined maximum value. This algorithm ensures that the linear system yielding the approximation of a given function remains sufficiently stable, which means that the condition number of the matrix to invert does not increase too fast.

3.2. Empirical interpolation method in source and sensor placement

Now, we apply EIM to sound field control problem. The candidate loudspeakers are assumed to be continuously distributed on a boundary of an enclosed space D including Ω and the boundary is denoted as ∂D. The functional space $\mathcal{V}$ is the set of the transfer functions of the candidate loudspeakers. The EIM will provide a finite set of loudspeakers that approximate transfer functions of an arbitrary loudspeaker with sufficient accuracy.

Suppose that the target space is free-field and the transfer functions are represented as monopole. The sound field inside D can be represented by the single layer potential as [20]

$$u(\mathbf{r}) = \int_{\mathbf{r}'\in\partial D} \varphi(\mathbf{r}')G(\mathbf{r}|\mathbf{r}')d\mathbf{r}' \qquad (\mathbf{r} \in D), \tag{8}$$

where φ is the density and G is the free-field Green's function. By approximating the Green's function on ∂D using the chosen loudspeakers of monopole, any sound field can be approximated by a finite sum of the Green's functions, i.e., by using a finite number of loudspeakers, on the basis of the principle of EIM. Also in a general reverberant case, by using the chosen set of loudspeakers, it is guaranteed that the transfer functions of the distribution of the candidate loudspeakers on ∂D can be approximated. The EIM also provides positions of the control points (microphones) inside Ω as sampling points to obtain the driving signals of the loudspeakers with a stable inverse filter.

In practice, the loudspeakers and sampling points are chosen from a predefined large discrete set of points on ∂D and inside Ω,

Algorithm 1 Proposed algorithm of source and sensor placement for single frequency

Input: A set of candidate points $\mathbf{y}_l$ ($l \in \{1, \ldots, L\}$) and $\mathbf{x}_m$ ($m \in \{1, \ldots, M\}$), a transfer function matrix $\mathbf{G} \in \mathbb{C}^{M \times L}$, and target error tolerance ϵ_{tol}.

Output: A set of loudspeaker and microphone indices.

Set $Q = 1$

while $\epsilon > \epsilon_{\mathrm{tol}}$ **do**

Select the loudspeaker index

$$l_Q = \underset{l=1,\ldots,L}{\arg\max} \left\| \mathbf{G}_{\cdot,l} - I_{Q-1}(\mathbf{G}_{\mathbf{m}_{Q-1},l}) \right\|_\infty$$

and the corresponding index of control point

$$m_Q = \underset{m=1,\ldots,M}{\arg\max} \left| \mathbf{G}_{m l_Q} - \left(I_{q-1} \left(\mathbf{G}_{\mathbf{m}_{Q-1},l_q} \right) \right)_m \right|.$$

Define the error by

$$\epsilon = \max_{l=1,\ldots,L} \left\| \mathbf{G}_{\cdot,l} - I_{Q-1} \left(\mathbf{G}_{\mathbf{m}_{Q-1},l} \right) \right\|_2.$$

and set $Q := Q + 1$

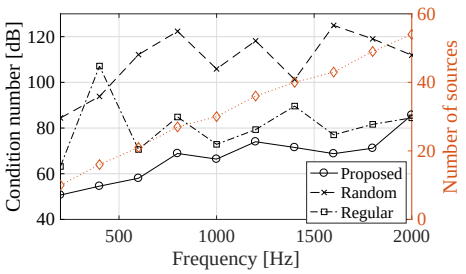
end while

respectively. The proposed algorithm is summarized in Algorithm 1. The discrete set of possible sampling positions inside Ω is defined as $\mathbf{x}_m$ ($m \in \{1, \ldots, M\}$), and candidate loudspeakers are located on a set of points $\mathbf{y}_l$ ($l \in \{1, \ldots, L\}$). This algorithm makes it possible to choose proper m and l for approximating the transfer function matrix $\mathbf{G}$ with the target error tolerance ϵ_{tol}. This transfer function matrix $\mathbf{G}$ will be calculated by a numerical acoustic simulation of the target space in advance. In Algorithm 1, the elements of matrices are represented by its subscript. For example, $\mathbf{G}_{\cdot,l}$ is the lth column and $\mathbf{G}_{\mathbf{m}_{Q-1},l}$ is the submatrix of $(\mathbf{m}_{Q-1}, l)$th elements of $\mathbf{G}$ where $\mathbf{m}_{Q-1} \in \mathbb{N}^{q-1}$ is the vector of selected indexes of m by the $Q-1$ iterations. It is also possible to stop the iteration when Q reached to the maximum available number of microphones and loudspeakers.

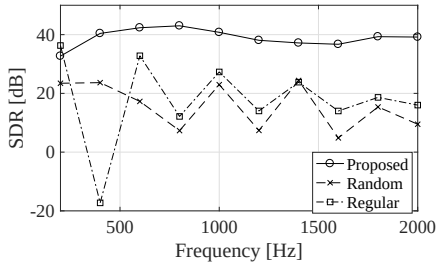
Algorithm 1 is applied to the transfer function matrix of a single frequency. The frequency of interest for the sound field control, however, can be broadband. In this case, the input of the algorithm becomes the third order tensor of the transfer functions including the dimension of frequency, i.e., $\mathbf{G} \in \mathbb{C}^{M \times L \times K}$, where K is the number of frequencies. The proposed method can be applied to the broadband case by evaluating the ℓ_∞-norm for the matrix $\mathbf{G}_{\cdot,l,\cdot}$ and vector $\mathbf{G}_{m,l_Q,\cdot}$ in line 1 and 2 of Algorithm 1, respectively.

4. EXPERIMENTS

Numerical simulations are conducted to evaluate the proposed method in a 2D sound field. We use FreeFem++ [21], a finite element method (FEM) solver, for acoustic simulation. We assumed a parallelogram room as depicted by the bold line in Fig. 2. The specific acoustic impedance ratio of each wall was set at 131.3 for all the frequencies, which corresponds to an absorption ratio of 0.03. Candidate loudspeaker locations are located along the boundary of the rectangular region in size of 2.4×2.8 m^2. The number of candidate loudspeakers is 256 and the rectangular boundary is discretized at regular intervals. The desired region is set as a rectangular region of 0.8×1.0 m^2, which is discretized at every 0.04 m to construct



(a) Condition number



(b) SDR

Fig. 1: Condition number and SDR with respect to frequency.

candidate control points. The candidate loudspeakers and control points are depicted by the dotted and dashed lines, respectively, in Fig. 2.

In the proposed method, the loudspeaker and control point locations are chosen from these candidate locations. We compared the proposed method with the random and regular placement. In the random placement, the same numbers of loudspeakers and control points as those of the proposed method are randomly chosen from the candidates. The rectangular boundaries of the loudspeaker candidate and the desired region are regularly discretized for the regular placement with the same number of loudspeakers and control points. We hereafter refer the proposed, random, and regular placements as **Proposed**, **Random**, and **Regular**, respectively.

We define signal-to-distortion ratio (SDR) for evaluation as

$$\mathrm{SDR}(\omega) = 10 \log_{10} \frac{\int_\Omega |u_{\mathrm{des}}(\mathbf{r}, \omega)|^2 d\mathbf{r}}{\int_\Omega |u_{\mathrm{syn}}(\mathbf{r}, \omega) - u_{\mathrm{des}}(\mathbf{r}, \omega)|^2 d\mathbf{r}}, \quad (9)$$

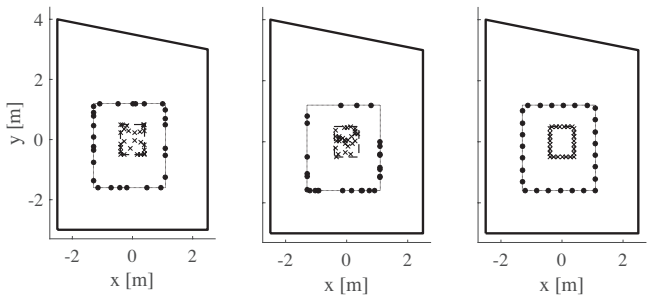
where $u_{\mathrm{des}}(\cdot)$ and $u_{\mathrm{syn}}(\cdot)$ are the desired and synthesized pressure fields, respectively. The desired sound field is a plane wave field and its arrival angle is varied from 0 to 350 deg at 10 deg intervals.

First, we compare the reproduction performance of the plane wave field in the single-frequency case. Next, a more practical situation of the broadband case is demonstrated.

4.1. Single-frequency case

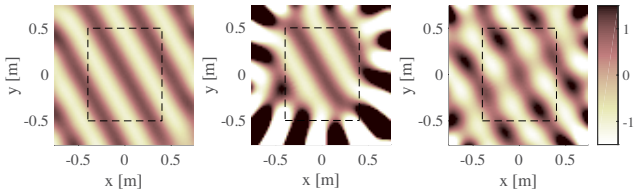
In the single frequency case, the numbers and locations of the loudspeakers and control points are determined at each frequency. We set ϵ_{tol} of Algorithm 1 as 1.0×10^{-2}. The driving signals of the loudspeakers are obtained by using (4) without regularization.

The condition number of $\mathbf{G}$ and SDR with respect to the frequency are shown in Fig. 1. The number of loudspeakers and control points is also plotted in Fig. 1a in red. Note that the placement of **Random** is randomly determined at each frequency. As expected,



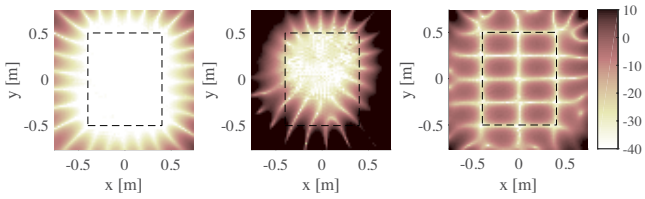
(a) Proposed (b) Random (c) Regular

Fig. 2: Selected positions of loudspeakers and microphones for 800 Hz. Black dots and cross marks represent loudspeakers and control points, respectively.



(a) Proposed (b) Random (c) Regular

Fig. 3: Reproduced sound pressure distributions at 800 Hz.



(a) Proposed (b) Random (c) Regular

Fig. 4: Normalized Squared error distributions at 800 Hz.

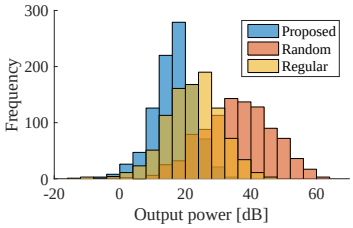


Fig. 5: Histogram of output power of loudspeakers at 800 Hz.

the number of loudspeakers and control points required for the reproduction increases as the frequency. Although the condition number of **Proposed** slightly increases as the frequency, it is the smallest compared to those of **Random** and **Regular** except 2000 Hz. The highest SDR is also achieved by **Proposed** as shown in Fig. 1b. The **Regular** curve features sharp dips in SDR and corresponding peaks in condition number at a number of frequencies. This is due to the non-uniqueness problem discussed above.

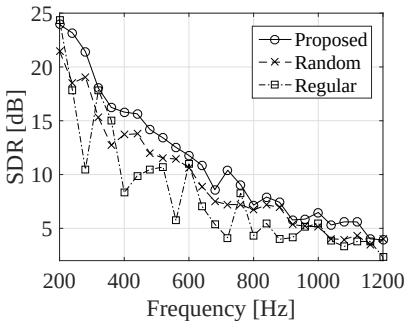


Fig. 6: SDR with respect to frequency for broadband case.

Figure 2 is the selected locations of loudspeakers and control points for 800 Hz, which are indicated by black dots and cross marks, respectively. Figures 3 and 4 are the reproduced pressure and squared error distributions, respectively, by using these placements. The region of high reproduction accuracy is limited in **Random** and the reproduction accuracy of **Regular** is poor. On the other hand, a high reproduction accuracy is achieved by **Proposed** over the entire desired region. To demonstrate the stability of the inverse filter, we also plot the histogram of the output power of the driving signals in Fig. 5. **Proposed** also results in a significantly lower average/max power of the driving signals.

4.2. Broadband case

In most practical use cases, the sound field control has to be broadband. The proposed method is applied using the transfer function calculated by FEM from 40 to 1200 Hz at intervals of 40 Hz. The threshold value ϵ_{tol} is again 1.0×10^{-2}. The number of loudspeaker and control-point locations chosen by **Proposed** is 33. The same number is used for **Random** and **Regular** methods. To obtain the driving signals, we added noise to the transfer functions so that the signal-to-noise ratio becomes 20 dB to simulate a mismatch in the transfer functions. In this noisy scenario, the regularization for calculating the inverse of the transfer functions is necessary; therefore, the driving signals are obtained by using (5). The regularization parameter λ was chosen in the range $[10^{-7}, 10^{0}]$, maximizing SDR.

Figure 6 is the SDR with respect to the frequency from 200 to 1200 Hz. As in the single-frequency case, the SDR of **Regular** sharply decreases at several frequencies. Above 200 Hz, **Proposed** consistently outperforms in terms of SDRs over the whole frequency range.

5. CONCLUSION

This study shows that there is a clear benefit in jointly placing loudspeakers and control points for the sound field control and reproduction problem. The so-called empirical interpolation method leads to a low-complexity procedure, providing optimal positions independently of the desired sound field. In the numerical experiments, this achieves significantly higher SDR and lower condition number than random and regular placements. Further experiments should confirm these findings, in a 3D setting, and with actual acoustic experiments.

6. ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number JP15H05312.

7. REFERENCES

[1] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 2, pp. 145–152, 1988.

[2] P. A. Nelson, "Active control of acoustic fields and the reproduction of sound," *J. Sound Vibr.*, vol. 177, no. 4, pp. 447–477, 1993.

[3] P.-A. Gauthier and A. Berry, "Sound-field reproduction in-room using optimal control techniques: Simulations in the frequency domain," *J. Acoust. Soc. Am.*, vol. 117, no. 2, pp. 662–678, 2005.

[4] F. M. Fazi and P. A. Nelson, "Nonuniqueness of the solution of the sound field reproduction problem with boundary pressure control," *Acta Acustica united with Acustica*, vol. 98, no. 1, pp. 1–14, 2012.

[5] 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, 1993.

[6] S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," in *Proc. 124th AES Conv.*, Amsterdam, Oct. 2008.

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

[8] S. Koyama, K. Furuya, Y. Hiwasaki, and Y. Haneda, "Analytical approach to wave field reconstruction filtering in spatio-temporal frequency domain," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 4, pp. 685–696, 2013.

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

[10] S. Koyama, K. Furuya, Y. Hiwasaki, Y. Haneda, and Y. Suzuki, "Wave field reconstruction filtering in cylindrical harmonic domain for with-height recording and reproduction," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 10, pp. 1546–1557, 2014.

[11] S. Koyama, K. Furuya, K. Wakayama, S. Shimauchi, and H. Saruwatari, "Analytical approach to transforming filter design for sound field recording and reproduction using circular arrays with a spherical baffle," *J. Acoust. Soc. Am.*, vol. 139, no. 3, pp. 1024–1036, 2016.

[12] F. Asano, Y. Suzuki, and D. C. Swanson, "Optimization of control source configuration in active control systems using gram–schmidt orthogonalization," *IEEE Trans. Speech Audio Process.*, vol. 7, no. 2, pp. 213–230.

[13] H. Khalilian, I. V. Bajić, and R. G. Vaughan, "Comparison of loudspeaker placement methods for sound field reproduction," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 8, pp. 1364–1379, 2016.

[14] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 13, no. 1, pp. 135–143, 2005.

[15] G. Chardon, W. Kreuzer, and M. Noisternig, "Design of spatial microphone arrays for sound field interpolation," vol. 9, no. 5, pp. 780–790, 2015.

[16] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera, "An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations," *C. R. Acad. Sci. Paris, Ser. I*, vol. 339, pp. 667–672, 2004.

[17] Y. Maday, N. C. Nguyen, A. T. Patera, and G. S. Pau, "A general, multipurpose interpolation procedure: the magic points," *Commun. Pure Appl. Anal.*, vol. 8, pp. 383–404, 2007.

[18] N. Ueno, S. Koyama, and H. Saruwatari, "Listening-area-informed sound field reproduction based on circular harmonic expansion," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, New Orleans, Mar. 2017, pp. 111–115.

[19] M. A. Grepl, Y. Maday, N. C. Nguyen, and A. T. Patera, "Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations," *ESAIM: M2AN*, vol. 41, no. 3, pp. 575–605, 2007.

[20] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer, 1998.

[21] F. Hecht, "New development in FreeFem++," *J. Numer. Math.*, vol. 20, no. 3-4, pp. 251–265, 2012.