






# Spherical-Harmonic-Domain Feedforward Active Noise Control Using Sparse Decomposition of Reference Signals from Distributed Sensor Arrays

Yu Maeno , Member, IEEE, Yuki Mitsufuji , Member, IEEE, Prasanga N. Samarasinghe , Member, IEEE, Naoki Murata , Member, IEEE, and Thushara D. Abhayapala , Senior Member, IEEE

**Abstract**—Active acoustic noise attenuation over a sizable space is a challenging problem in signal processing. The noise attenuation performance of feedforward active noise control (ANC) relies on the preciseness of a reference signal of a primary noise field. To capture the precise reference signal for controlling a sizable space, a large number of reference microphones are required, which reduces system viability. In this study, we exploit an efficient representation of the reference signal in spherical harmonic (SH) domain by utilizing the inherent sparseness of the noise field. The main contributions of this work are as follows. (1) A general reference microphone geometry can be used. The implementation difficulty in the array structure, which is recognized as the common issue of SH-domain signal processing, e.g., use of a fully surrounding spherical array, is reduced by using the fields translation based on the addition theorem. (2) The accuracy of low-frequency signal decomposition is improved. The low accuracy of low-frequency signal decomposition in compressive sensing (CS), which is commonly reported in the literature, is improved by applying signal representation in SH domain. (3) System robustness is increased. The robustness of the system is increased by considering a noise source spatial distribution of both the interior and exterior sound fields, which is not possible in the case of a general signal representation in SH domain. Experimental results indicate that the noise attenuation performance of our proposed method exceeds that of existing solutions. The flexibility of the array structure is also increased, which leads to a more feasible practical system setup.

**Index Terms**—Active noise control, adaptive algorithm, spherical harmonic, compressive sensing, feedforward.

Manuscript received April 5, 2019; revised September 8, 2019 and December 7, 2019; accepted December 15, 2019. Date of publication December 27, 2019; date of current version January 22, 2020. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Simon Doclo. (Corresponding author: Yu Maeno.)

Y. Maeno is with the R&D Center, Sony Corporation, Tokyo, Japan (e-mail: yu.maeno@sony.com).

Y. Mitsufuji is with the R&D Center, Sony Corporation, Tokyo, Japan (e-mail: yuhki.mitsufuji@sony.com).

P. N. Samarasinghe is with the Research School of Engineering, Australian National University, Canberra, Australian Capital Territory, 2601 Australia (e-mail: prasanga.samarasinghe@anu.edu.au).

N. Murata is with the R&D Center, Sony Corporation, Tokyo, Japan (e-mail: naoki.murata@sony.com).

T. D. Abhayapala is with the Research School of Engineering, Australian National University, Canberra, Australian Capital Territory, 2601 Australia (e-mail: thushara.abhayapala@anu.edu.au).

Digital Object Identifier 10.1109/TASLP.2019.2962684

## I. INTRODUCTION

ACTIVE noise control is a technique by which a noise field is attenuated by generating an antinoise field using secondary sources. The antinoise field is produced by driving signals using secondary sources, which are generated by applying adaptive linear filters to reference signals captured by reference microphones. Error microphones capture residual errors, which can be fed back in an adaptive algorithm, e.g., the least mean squares (LMS) algorithm to update the adaptive filters. Such an approach is well known as the filtered-x LMS (FxLMS) algorithm and its other variants [1]–[4].

Spatial ANC aims to attenuate the noise field over a sizable space using multiple microphones and secondary sources. Multiple-channel FxLMS-based methods in the temporal frequency domain [5], [6] were proposed in practice. A major drawback of this method is that an entire region of interest (ROI) has to be covered with densely placed error microphones to produce a large quiet region, because the method tries to minimize the residual error only at the error microphone positions. Thus, we refer to this method as the multiple-point method. In contrast, harmonic-domain-based ANC methods, which decompose signals into harmonic basis functions, e.g., spherical harmonic (SH) functions, were proposed [7]–[11]. The advantages of this approach are that the sound field can be controlled in terms of a region rather than points and a fast convergence can be achieved by uniformly placing the microphones and secondary sources on an applicable surface, e.g., a sphere. Although this constraint of the array geometry gives stable and efficient signal representation in SH domain, spherical arrays in a 3D space with a number of microphones and loudspeakers are physically difficult to construct in practice. The reference microphone array (RMA) particularly occupies a large space since it is placed in the outermost position, which reduces the viability of a system. Instead of using a spherical array, multiple circular arrays were used to first transform signals into azimuth harmonic coefficients [12], and then the SH coefficients are calculated numerically to perform ANC in SH domain [13]. In this approach, the difficulty in constructing the array structure has been reduced but still the ROI has to be surrounded by the arrays.

To alleviate this problem in general spatial sound field processing task, several approaches take into account a practical sound field aspect. In most practical situations [14], [15], the

underlying sound field is due to a small number of underlying sound sources. In this context, compressive sensing (CS) has been successfully applied to sound field reproduction, sound field separation, and sound localization [16]–[20]. Koyama and co-workers [16], [17] showed that the spatial aliasing [21] originating from spatial sampling can be avoided in sound field reproduction. Takida *et al.* [18] showed that the exterior and interior sound fields corresponding to the direct source and reverberant component, respectively, can be precisely separated. CS has been applied to spatial ANC as well. Zhang *et al.* [22] showed that the number of secondary sources can be reduced. We previously introduced the CS approach to reduce the number of reference microphones and showed that noise can be attenuated by using a number of microphones that is smaller than a number required by the spatial Nyquist theory [23]. One limitation of this approach is that the reproduction accuracy at low frequencies is limited since CS is based on an assumption, that the columns of a sensing matrix are mutually incoherent [24]. However, in general, they are highly coherent at low frequencies [25]–[27]. To overcome this problem, CS has been applied to SH coefficients [28], [29].

In this paper, we focus on reference signal reconstruction and exploit its efficient representation by utilizing the inherent spatial sparseness of the noise field. As mentioned above, the RMA occupies a large space and it also limits the system performance owing to the spatial aliasing; therefore, we investigate sparse recovery of reference signals to improve both the system viability and performance. The flexibility of the RMA structure and placement is achieved by combining the sparse recovery with the fields translation [30] based on the addition theorem [31]. The sensing matrix is transformed into an orthonormal space; thus, the reconstruction accuracy at low frequencies is improved. Furthermore, a compact representation of signals based on CS improves the performance at frequencies higher than the spatial Nyquist frequency. Another fundamental problem of the sound field representation based on harmonic expansion is that both outgoing and incoming sound fields have to be considered separately [32]. This is because the harmonic expansion is based on a complete set of solutions to the homogeneous Helmholtz equation, which does not allow any sound sources to exist inside a closed region, when an incoming sound field is considered within the region. In a practical setup, however, the existence of sound sources inside the ROI is easily assumed. Regarding this problem, the proposed method describes the sound field in a sparsely represented manner and estimates the spatial noise source distribution. On the basis of the distribution, we can define target noise sources, which are expected to be attenuated, and nontarget noise sources, which are expected to be ignored. The nontarget sources, e.g., sources existing inside the ROI, can be removed from the ANC processing, which leads to high system robustness.

The main contributions of this work are as follows: (1) Generalization of the SH-domain reference signal representation utilizing the fields translation, which can be applied to an arbitrary RMA structure; (2) Improvement of aforementioned low-frequency accuracy in CS by transforming signals into the SH domain; (3) Increasing the system robustness by

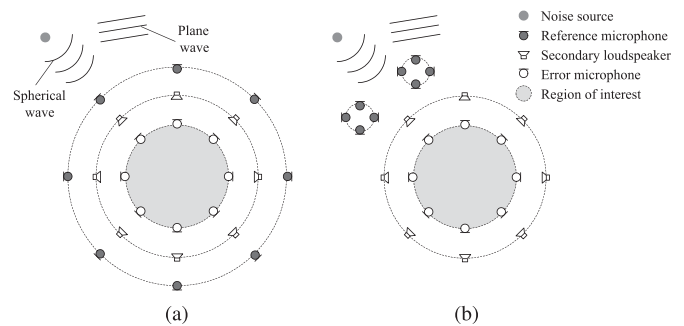


Fig. 1. System model of the interior sound field constructed using different RMA structures: (a) ROI fully surrounded by RMA; (b) Compactly distributed RMAs.

compensating the spatial noise source distribution of both the interior and exterior sound fields. To sum up, our proposed method gives flexibility in reference microphone placement, which improves system viability. Furthermore, system robustness is increased for the aforementioned problem associated with the composite noise field of outgoing and incoming sound fields.

## II. PRELIMINARIES AND PROBLEM STATEMENT

An interior noise field attenuation in a sizable 3D space is usually achieved by surrounding an ROI with a number of sensors and loudspeakers. One of the most ideal configurations for performing SH-domain ANC is to place spherical arrays of error microphones, secondary loudspeakers, and reference microphones around a global origin, i.e., the center of the ROI to fully surround the ROI, as shown in Fig. 1(a). The RMA has to be placed in the outermost location to satisfy the causality constraint since the reference signals of a noise field have to be captured before the noise wave front arrives at the ROI. Therefore, the RMA becomes relatively large, which may be difficult to construct in practice. Furthermore, a large number of microphones are required to fill the gap between the microphones to avoid spatial aliasing. The basic idea in dealing with these issues is based on the assumption of a practical situation, that is, the noise source positions are known and the noise sources are sparsely distributed. In such a scenario, sufficient noise attenuation may be achieved by placing the reference microphones closer to the possible noise source direction(s), as shown in Fig. 1(b). Hereafter, the center points of distributed RMAs are referred to as local origins.

Now, let us consider an arbitrary noise field  $S(\mathbf{r}, k)$  in the free field at position  $\mathbf{r}$  within the ROI,  $\Omega$ . Since we assume that the sound sources exist inside the region,  $S(\mathbf{r}, k)$  satisfies the inhomogeneous Helmholtz equation [33], [34]:

$$(\nabla^2 + k^2) S(\mathbf{r}, k) = -\rho_{\text{int}}^{(\text{ps})}(\mathbf{r}, k), \quad (1)$$

where  $k = 2\pi f/c$  is the wave number,  $f$  is the temporal frequency,  $c$  is the speed of sound, and  $\rho_{\text{int}}^{(\text{ps})}(\mathbf{r}, k)$  is the distribution of point sources inside the region  $\Omega$ . Hereafter, we omit  $k$  for notational simplicity. By approximating  $\rho_{\text{int}}^{(\text{ps})}(\mathbf{r})$  as a linear combination of  $L_{\text{ps,int}}$  point sources,  $S(\mathbf{r})$  can be represented

as a sum of the particular and homogeneous solutions [17], [35]:

$$S(\mathbf{r}) = S_P(\mathbf{r}) + S_H(\mathbf{r}). \quad (2)$$

The particular solution can be described as a convolution of  $\rho_{\text{int}}^{(\text{ps})}(\mathbf{r})$  and the acoustic transfer function inside  $\Omega$ :

$$S_P(\mathbf{r}) = \int_{\mathbf{y} \in \Omega} \rho_{\text{int}}^{(\text{ps})}(\mathbf{y}) G(\mathbf{r}|\mathbf{y}) d\mathbf{y} \approx \sum_{\ell=1}^{L_{\text{ps,int}}} \gamma_{\ell,\text{int}}^{(\text{ps})} G(\mathbf{r}|\mathbf{y}_{\ell,\text{int}}), \quad (3)$$

where  $G(\mathbf{r}|\mathbf{y})$  is the free-field Green's function between  $\mathbf{r}$  and  $\mathbf{y}$ ,  $L_{\text{ps,int}}$  is the number of the point sources inside  $\Omega$ ,  $\gamma_{\ell,\text{int}}^{(\text{ps})}$  is the complex amplitude of the  $\ell^{\text{th}}$  source inside  $\Omega$ , and  $\mathbf{y}_{\ell,\text{int}}$  is the source position inside  $\Omega$ . The homogeneous solution can be described as a superposition of plane waves, which is referred to the Herglotz wave function as

$$S_H(\mathbf{r}) = \int_{\hat{\mathbf{y}} \in S^2} \rho^{(\text{pw})}(\hat{\mathbf{y}}) e^{ik\hat{\mathbf{y}} \cdot \mathbf{r}} d\hat{\mathbf{y}}, \quad (4)$$

where  $S^2$  is the unitary sphere,  $\rho^{(\text{pw})}(\hat{\mathbf{y}})$  is the distribution of the plane waves, and  $\hat{\mathbf{y}}$  is a unit vector towards the direction of the plane wave. We also consider the approximation of the solution by discretizing the direction of plane waves. A sound field consisting of far-field sources can be decomposed into a sparse set of weights in the plane-wave domain. By contrast, a sound field consisting of near-field sources can be expressed by the superposition of few point sources. The combination of these two types of sound source enables a sparse signal representation of diverse sound fields. Therefore, we introduce the model of  $S_H$  as a sum of the plane waves and point sources:

$$S_H(\mathbf{r}) \approx \sum_{\ell=1}^{L_{\text{pw}}} \gamma_{\ell}^{(\text{pw})} e^{ik\hat{\mathbf{y}}_{\ell} \cdot \mathbf{r}} + \sum_{\ell=1}^{L_{\text{ps,ext}}} \gamma_{\ell,\text{ext}}^{(\text{ps})} G(\mathbf{r}|\mathbf{y}_{\ell,\text{ext}}), \quad (5)$$

where  $\gamma_{\ell}^{(\text{pw})}$  and  $\gamma_{\ell,\text{ext}}^{(\text{ps})}$  are the complex amplitudes of the plane waves and point sources, respectively, which exist outside  $\Omega$ ,  $L_{\text{pw}}$  and  $L_{\text{ps,ext}}$  are the number of the plane waves and point sources, respectively, and  $\mathbf{y}_{\ell,\text{ext}}$  is the source position outside  $\Omega$ . From (3) and (5), we describe the solution to (1) as

$$S(\mathbf{r}) \approx \sum_{\ell=1}^{L_{\text{pw}}} \gamma_{\ell}^{(\text{pw})} e^{ik\hat{\mathbf{y}}_{\ell} \cdot \mathbf{r}} + \sum_{\ell=1}^{L_{\text{ps,ext}}} \gamma_{\ell,\text{ext}}^{(\text{ps})} G(\mathbf{r}|\mathbf{y}_{\ell,\text{ext}}) + \sum_{\ell=1}^{L_{\text{ps,int}}} \gamma_{\ell,\text{int}}^{(\text{ps})} G(\mathbf{r}|\mathbf{y}_{\ell,\text{int}}). \quad (6)$$

Now, we assume that  $L_{\text{mic}}$  microphones are placed inside  $\Omega$  to capture the noise field. We discretize the integration to a sum of  $L = L_{\text{pw}} + L_{\text{ps,ext}} + L_{\text{ps,int}}$  defined grid points as shown in (6), and we derive the total noise field captured at the microphone position  $\mathbf{S} = [S(\mathbf{r}_1), \dots, S(\mathbf{r}_{L_{\text{mic}}})]^T$  as a sum of finite point sources and plane waves:

$$\mathbf{S} = \mathbf{H}\boldsymbol{\gamma}, \quad (7)$$

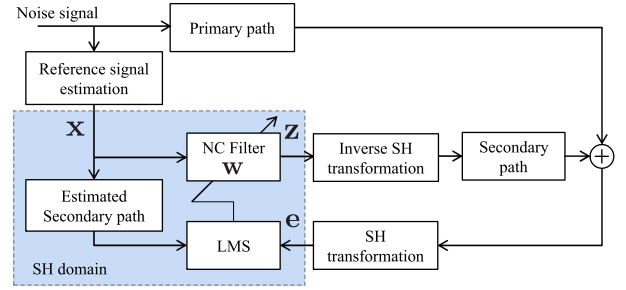


Fig. 2. Block diagram of the proposed method.

where  $\mathbf{H}$  is the sensing matrix consisting of the basis functions of the point sources and plane waves, and  $\boldsymbol{\gamma} \in \mathbb{C}^{L \times 1}$  is the distribution of the point source and plane wave. In sparse decomposition, (7) is set to be an underdetermined problem and the matrix  $\mathbf{H}$  is the so-called overcomplete dictionary, which usually has a dense resolution of the spatial discretization. The noise distribution  $\boldsymbol{\gamma}$  describes the total primary noise field. The problem model in this work is similar to that in [36]. The difference is the definition of the sensing matrix. In [36], a sensing matrix is defined as an acoustic transfer function in the spatial (temporal frequency) domain. In contrast, a sensing matrix is defined in the SH domain in this work. This gives the advantage to the system when a small microphone array is used to capture the sound field in the low frequency. Further discussions of the difference in the sensing matrix can be found in Sect. IV-B and VI-B.

The goal of this study is to precisely estimate the reference signals of the primary noise field from the measurements of RMAs. In the array configuration shown in Fig. 1(b), extrapolation of the noise field is required since the RMAs are placed at local origins, which have certain distances to the global origin. In the literature, at least the same number of microphones are required in local origins as in the global origin to reproduce the same sound field in theory [37], [38]. In fact, a much larger number of microphones are required for the precise reproduction of the ROI when the distances between the local origins and the global origin are large. Instead of increasing the number of microphones, we use an SH-domain representation of the sensing matrix and measurements in the CS framework to efficiently decompose the noise field into a sparse set of basis functions.

Fig. 2 shows a block diagram of the proposed method, which basically follows the standard FxLMS algorithm. Acoustic feedback from the secondary sources to the reference microphones is a common problem in ANC tasks and several solutions are proposed in the literature [39], [40]. Although, these methods can also be applied to our system, we ignore the feedback path in this work for the sake of simplicity. Our approach focuses on the reference signal estimation block by utilizing the sparsity of the noise field. The flow of the estimation process is shown in detail in Fig. 3. The measurements are first transformed into SH coefficients and then decomposed into a sparse set of basis functions of the sound field, i.e., point sources and plane waves, in the decomposition stage, followed by the reconstruction stage. Finally, the reconstructed reference signal is input to the SH-domain FxLMS algorithm.

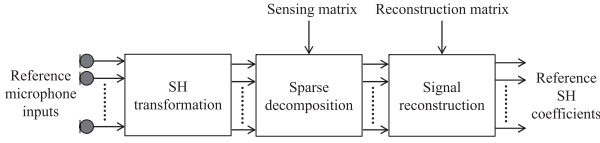


Fig. 3. Block diagram of the reference signal estimation part.

### III. SPHERICAL-HARMONIC-DOMAIN ANC

#### A. Spherical Harmonic Representation

As illustrated in Fig. 1, an arbitrary interior noise field is considered. The noise field can be decomposed into SH coefficients, which can be derived as the solution to the wave equation in the spherical coordinate:

$$S_H(\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \alpha_{nm} j_n(kr) Y_{nm}(\theta, \phi), \quad (8)$$

where  $\alpha_{nm}$  is the interior SH coefficient and  $j_n(\cdot)$  is the  $n^{\text{th}}$ -order spherical Bessel function,

$$Y_{nm}(\theta, \phi) = (-1)^m \sqrt{\frac{2n+1}{4\pi}} \sqrt{\frac{(n-m)!}{(n+m)!}} P_{nm}(\cos \theta) e^{im\phi} \quad (9)$$

is the SH of order  $n$  and degree  $m$ ,  $P_{nm}(\cos \theta)$  is the associated Legendre function, and  $\theta$  and  $\phi$  are the colatitude and azimuth angle, respectively.

The antinoise field generated by a secondary loudspeaker array (SLA) can be described as

$$S_c(\mathbf{r}) = \sum_{\ell=1}^{L_s} d_{\ell} G(\mathbf{r}|\mathbf{r}_{\ell}), \quad (10)$$

where  $L_s$  is the number of loudspeakers and  $d_{\ell}$  is the driving signal of the  $\ell^{\text{th}}$  loudspeaker.

The free-field Green's function,  $G(\mathbf{r}|\mathbf{r}_{\ell})$ , can be parameterized as

$$G(\mathbf{r}|\mathbf{r}_{\ell}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n C_{nm,\ell} j_n(kr) Y_{nm}(\theta, \phi), \quad (11)$$

where  $C_{nm,\ell}$  is the SH coefficient of the transfer function from the  $\ell^{\text{th}}$  loudspeaker to the ROI. By substituting (11) to (10), we can represent the antinoise field as

$$S_c(\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{\sum_{\ell=1}^{L_s} d_{\ell} C_{nm,\ell}}_{z_{nm}} j_n(kr) Y_{nm}(\theta, \phi). \quad (12)$$

The coefficient  $C_{nm,\ell}$  can be further decomposed as

$$C_{nm,\ell} = g_{nm} Y_{nm}^*(\theta_{\ell}, \phi_{\ell}), \quad (13)$$

where  $g_{nm}$  is the SH coefficient of the transfer function. From (12) and (13),  $z_{nm}$  can be written as

$$z_{nm} = g_{nm} \underbrace{\sum_{\ell=1}^{L_s} d_{\ell} Y_{nm}^*(\theta_{\ell}, \phi_{\ell})}_{\xi_{nm}}, \quad (14)$$

where  $\xi_{nm}$  is the SH coefficient of the driving signals.

#### B. Spherical-Harmonic-Domain Adaptive Filter Update

Noise field attenuation is considered in SH domain. An error microphone array (EMA) is placed at the coordinate origin to surround the ROI. A residual noise field captured by the EMA can be described as a sum of the primary noise field given in (8) and the antinoise field given in (12):

$$\begin{aligned} e_{nm} &= \alpha_{nm} + z_{nm} \\ &= \alpha_{nm} + g_{nm} \xi_{nm}. \end{aligned} \quad (15)$$

As depicted in Fig. 2, the FxLMS algorithm generates an antinoise signal by applying a linear filter to the reference signal; thus, the total residual noise field can be written as

$$e_{nm} = \alpha_{nm} + g_{nm} x_{nm} w_{nm}, \quad (16)$$

where  $x_{nm}$  and  $w_{nm}$  are the SH coefficients of the reference signal and the adaptive filter, respectively. By concatenating all available coefficients for each order  $n$  and degree  $m$ , we can write (16) in a matrix form:

$$\mathbf{e} = \boldsymbol{\alpha} + \mathbf{g}\mathbf{x}\mathbf{w}, \quad (17)$$

where

$$\mathbf{g} = \text{diag}(g_{00}, \dots, g_{NN}) \in \mathbb{C}^{(N+1)^2 \times (N+1)^2},$$

$$\mathbf{x} = \text{diag}(x_{00}, \dots, x_{NN}) \in \mathbb{C}^{(N+1)^2 \times (N+1)^2},$$

and

$$\mathbf{w} = [w_{00}, \dots, w_{NN}]^T \in \mathbb{C}^{(N+1)^2 \times 1}.$$

The notation  $\text{diag}(\cdot)$  indicates the diagonal matrix whose diagonal elements consist of the vector of argument. The truncation order  $N$  is usually determined by adopting the criterion  $N = \lceil ekR/2 \rceil$  [41] or  $N = \lceil kR \rceil$  [42], where  $R$  is the radius of the ROI.

A cost function is set as the sum of the squared residual signals [10]:

$$\mathcal{J}(k, \iota) = \sum_{n=0}^N \sum_{m=-n}^n |e_{nm}(k, \iota)|^2 = \mathbf{e}^H(k, \iota) \mathbf{e}(k, \iota), \quad (18)$$

where  $\iota$  is the iteration index. A gradient of the cost function  $\mathcal{J}(k, \iota)$  can be calculated as

$$\frac{\partial \mathcal{J}(k, \iota)}{\partial \mathbf{w}^*(k, \iota)} = [\mathbf{g}(k) \mathbf{x}(k, \iota)]^H \mathbf{e}(k, \iota), \quad (19)$$

so that we can update the filter coefficient as

$$\mathbf{w}(k, \iota + 1) = \mathbf{w}(k, \iota) - \mu [\mathbf{g}(k) \mathbf{x}(k, \iota)]^H \mathbf{e}(k, \iota), \quad (20)$$

where  $\mu$  is the step size.

The SH coefficient  $e_{nm}$  can be calculated by measuring the sound field on a boundary of a sphere:

$$e_{nm} = \frac{1}{j_n(kr_q)} \int_0^{2\pi} \int_0^{\pi} S(\mathbf{r}) Y_{nm}^*(\theta, \phi) \sin \theta d\theta d\phi, \quad (21)$$



where  $r_q$  is the radius of the EMA. Note that we use the orthonormality of the SH function:

$$\int_0^{2\pi} \int_0^\pi Y_{nm}(\theta, \phi) Y_{n'm'}^*(\theta, \phi) \sin \theta d\theta d\phi = \delta_{nn'} \delta_{mm'}, \quad (22)$$

where  $\delta_{nn'}$  denotes the Kronecker delta function. In (21), the spherical Bessel function in the denominator causes a numerical instability, which is known as the Bessel zero problem or forbidden frequency problem [30]. To avoid this problem, a rigid baffle array is widely used in practice. Instead of using the spherical Bessel function in (21), a different radial function for a rigid baffle array,  $b_n(kr_q)$ , is introduced:

$$b_n(kr_q) = j_n(kr_q) - \frac{j_n'(kr_q)}{h_n^{(1)'}(kr_q)} h_n^{(1)}(kr_q), \quad (23)$$

where  $h_n^{(1)}(\cdot)$  is the first kind of the  $n^{\text{th}}$  order spherical Hankel function, and  $h_n^{(1)'}(\cdot)$  and  $j_n'(\cdot)$  are the derivative of  $h_n^{(1)}(\cdot)$  and  $j_n(\cdot)$ , respectively. The coefficient  $e_{nm}$  can be calculated as

$$e_{nm} = \frac{1}{b_n(kr_q)} \int_0^{2\pi} \int_0^\pi S(\mathbf{r}) Y_{nm}^*(\theta, \phi) \sin \theta d\theta d\phi. \quad (24)$$

Alternatively, one can use either a dual-radius open spherical array [43], multiple rigid spherical arrays [44], or multiple flat arrays [45] to avoid the Bessel zero problem.

In practice, the SH coefficient can be calculated from the error microphone signal,  $S_e(\mathbf{r}_a)$ , by discretizing the sound field on a sphere based on the quadrature method described in [30]:

$$e_{nm} \approx \frac{1}{b_n(kr_q)} \sum_{a=1}^A \Lambda_a S_e(\mathbf{r}_a) Y_{nm}^*(\theta_a, \phi_a), \quad (25)$$

where  $A$  is the total number of the sampling points and  $\Lambda_a$  is the sampling weights.

In the SH-domain FxLMS algorithm, the reference SH coefficients  $\mathbf{x}$  are used to generate the SH coefficients of the antinoise noise field as  $\mathbf{z} = \mathbf{g}\mathbf{x}\mathbf{w}$ , as well as to update the adaptive filter  $\mathbf{w}$  using (20). Therefore, the precise estimation of the reference SH coefficients is important for both the noise attenuation and the filter convergence performance. The estimation method of the reference SH coefficients is described in Sect. IV and V.

#### IV. SPARSE DECOMPOSITION OF NOISE FIELD

##### A. CS Applied to Spherical-Harmonic-Domain Signals

We derive an SH-domain sensing matrix and measurement signals in the CS framework to efficiently decompose a noise field into a sparse set of basis functions. We here assume all noise sources exist outside the ROI. An arbitrary noise field can be expressed as a sum of plane waves and point sources, which exist outside the ROI, as described in (5). An SH-domain representation can be derived by applying the SH expansion of a plane wave [46],

$$e^{ik\hat{\mathbf{y}}\cdot\mathbf{r}} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{4\pi i^{-n} Y_{nm}^*(\theta_s, \phi_s)}_{\psi_{nm}^{(\text{pw})}(\hat{\mathbf{y}})} j_n(kr) Y_{nm}(\theta, \phi), \quad (26)$$

and that of a point source [32],

$$G(\mathbf{r}|\mathbf{y}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \underbrace{-ik h_n^{(1)}(kr_s) Y_{nm}^*(\theta_s, \phi_s)}_{\psi_{nm}^{(\text{ps})}(\mathbf{y})} \times j_n(kr) Y_{nm}(\theta, \phi), \quad (27)$$

where  $(r_s, \theta_s, \phi_s)$  is the source position and  $\psi_{nm}^{(\text{pw})}(\hat{\mathbf{y}})$  and  $\psi_{nm}^{(\text{ps})}(\mathbf{y})$  are the SH coefficients of the plane wave and point source, respectively. From (5), (8), (26), and (27), an SH-domain representation of the noise field can be described as

$$\alpha_{nm} \approx \sum_{\ell=1}^{L_{\text{pw}}} \gamma_{\ell}^{(\text{pw})} \psi_{nm, \ell}^{(\text{pw})} + \sum_{\ell=1}^{L_{\text{ps}}} \gamma_{\ell}^{(\text{ps})} \psi_{nm, \ell}^{(\text{ps})}, \quad (28)$$

where  $L_{\text{pw}}$  and  $L_{\text{ps}}$  are the numbers of the grid directions and points of the plane wave and point source, respectively. The matrix form of (28) can be derived by truncating the order of SH expansion in a sufficient order, which can be written as

$$\boldsymbol{\alpha} = \boldsymbol{\Psi} \boldsymbol{\gamma}, \quad (29)$$

where

$$\boldsymbol{\Psi} = [\boldsymbol{\Psi}^{(\text{pw})} \boldsymbol{\Psi}^{(\text{ps})}] = \begin{bmatrix} \psi_{00,1}^{(\text{pw})} & \cdots & \psi_{00,L_{\text{pw}}}^{(\text{pw})} & \psi_{00,1}^{(\text{ps})} & \cdots & \psi_{00,L_{\text{ps}}}^{(\text{ps})} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \psi_{NN,1}^{(\text{pw})} & \cdots & \psi_{NN,L_{\text{pw}}}^{(\text{pw})} & \psi_{NN,1}^{(\text{ps})} & \cdots & \psi_{NN,L_{\text{ps}}}^{(\text{ps})} \end{bmatrix} \quad (30)$$

and

$$\boldsymbol{\gamma} = [\gamma_1^{(\text{pw})}, \dots, \gamma_{L_{\text{pw}}}^{(\text{pw})}, \gamma_1^{(\text{ps})}, \dots, \gamma_{L_{\text{ps}}}^{(\text{ps})}]^T. \quad (31)$$

One way of finding the optimal solution of (29) is to assume the spatial sparsity of  $\boldsymbol{\gamma}$  and solve the following minimization problem:

$$\underset{\boldsymbol{\gamma}}{\text{minimize}} \|\boldsymbol{\alpha} - \boldsymbol{\Psi} \boldsymbol{\gamma}\|_2^2 + \lambda \|\boldsymbol{\gamma}\|_p^p, \quad (32)$$

where the operator  $\|\cdot\|_p$  ( $0 < p \leq 1$ ) indicates the  $\ell_p$ -norm. The first term of (32) represents the approximation error and the second term is introduced to induce a sparse solution [47], where the parameter  $\lambda$  balances both terms. There are several algorithms to solve the problem [48]–[51]. Although the sensing matrix is described in the SH domain, the sparsity constraint is still imposed on the spatial source distribution  $\boldsymbol{\gamma}$ .

##### B. Coherence of Sensing Matrix

The performance of the CS in (32) highly depends on the coherence of the sensing matrix. One way to measure the coherence is to calculate the Gram matrix  $\Gamma$  [26]:

$$\Gamma_{ij} = \frac{\|\mathbf{h}_i^H \mathbf{h}_j\|_2}{\|\mathbf{h}_i\|_2 \|\mathbf{h}_j\|_2}, \quad (33)$$

where each  $(i, j)$  element in  $\Gamma$  is the normalized inner product of the columns  $\mathbf{h}_i$  and  $\mathbf{h}_j$  of the sensing matrix  $\mathbf{H}$ .

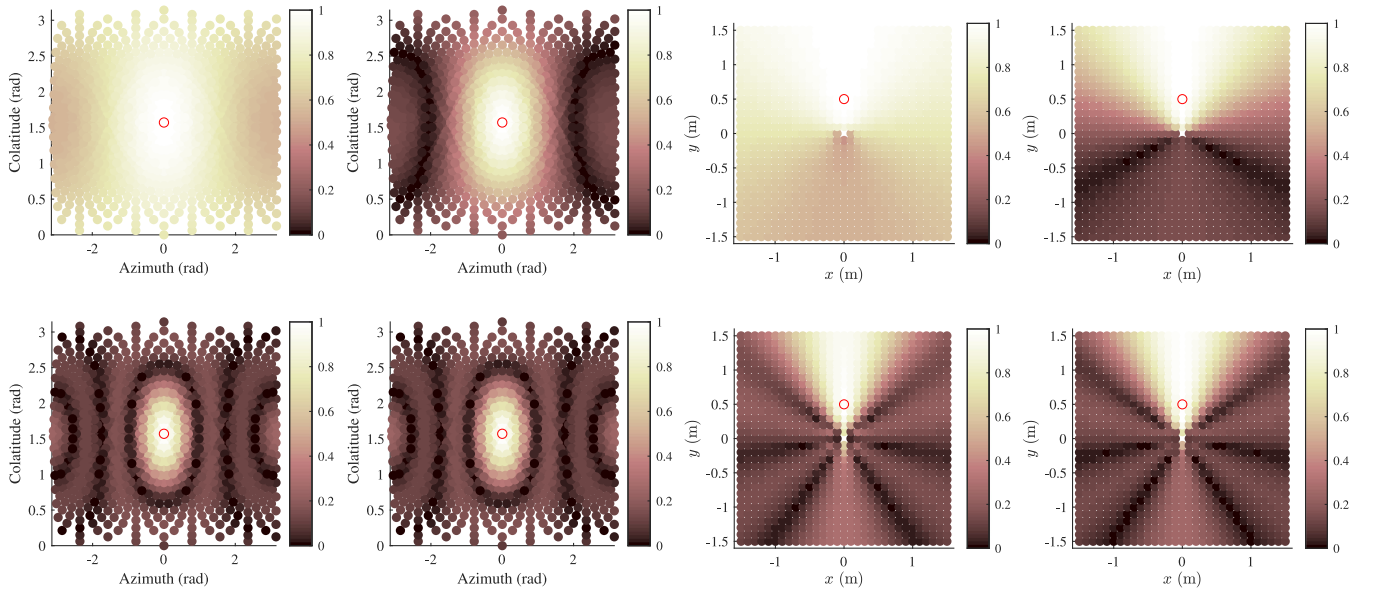


Fig. 4. (a) (b) Column coherence between a given direction  $(\theta_1, \phi_1) = (\pi/2, 0)$ , which is shown as a red solid circle, to other directions of the plane wave sensing matrix: (a) 500 Hz, (b) 1000 Hz. (c) (d) Column coherence between a given point  $(x_1, y_1) = (0, 0.5)$ , which is shown as a red solid circle, to other points of the point source sensing matrix: (c) 500 Hz, (d) 1000 Hz. Each dot corresponds to a single plane wave direction and a point source position. The first row is the result of the spatial-domain sensing matrix and the second row is that of SH-domain sensing matrix.

Sensing matrices for the plane waves can be described as

$$\mathbf{E} = \begin{bmatrix} e^{ik\hat{\mathbf{y}}_1 \cdot \mathbf{r}_1} & \cdots & e^{ik\hat{\mathbf{y}}_{L_{pw}} \cdot \mathbf{r}_1} \\ \vdots & \ddots & \vdots \\ e^{ik\hat{\mathbf{y}}_1 \cdot \mathbf{r}_{L_{mic}}} & \cdots & e^{ik\hat{\mathbf{y}}_{L_{pw}} \cdot \mathbf{r}_{L_{mic}}} \end{bmatrix} \quad (34)$$

in the spatial domain and  $\boldsymbol{\psi}^{(pw)}$  in the SH domain, respectively. Likewise, sensing matrices for the point sources can be described as

$$\mathbf{G} = \begin{bmatrix} G(\mathbf{r}_1 | \mathbf{y}_1) & \cdots & G(\mathbf{r}_1 | \mathbf{y}_{L_{ps}}) \\ \vdots & \ddots & \vdots \\ G(\mathbf{r}_{L_{mic}} | \mathbf{y}_1) & \cdots & G(\mathbf{r}_{L_{mic}} | \mathbf{y}_{L_{ps}}) \end{bmatrix} \quad (35)$$

in the spatial domain and  $\boldsymbol{\psi}^{(ps)}$  in the SH domain, respectively. By plotting a column of the Gram matrix  $\Gamma_{i1}$ , the direction- and position-dependent coherence can be visualized for the sensing matrices  $\mathbf{E}$ ,  $\boldsymbol{\psi}^{(pw)}$ ,  $\mathbf{G}$ , and  $\boldsymbol{\psi}^{(ps)}$ . In Fig. 4(a) and Fig. 4(b),  $\Gamma_{i1}$  indicates the correlation between one direction  $(\theta_1, \phi_1) = (\pi/2, 0)$  and all other possible directions  $(\theta_i, \phi_i)$ , while in Fig. 4(c) and Fig. 4(d),  $\Gamma_{i1}$  indicates the correlation between one point  $(x_1, y_1) = (0.5, 0)$  and all other possible directions  $(x_i, y_i)$ . Note that only the grid points in the  $x$ - $y$  plane is considered for the better visibility. 974 grid directions of plane waves are defined based on the Lebedev quadrature scheme [52]. 960 grid points are uniformly placed in a  $3 \times 3$  m region for point sources. The measurement points are uniformly sampled on a surface of a sphere with a radius of 0.1 m. The locations of the points correspond to 20 vertexes of a dodecahedron. The figure indicates that the source distributions of closely located 2 point sources as well as 2 plane waves are difficult to be estimated

separately by using the spatial-domain sensing matrix especially at low frequencies since most columns of the matrix are highly correlated. In contrast, the SH-domain sensing matrix shows a high correlation only at relatively close directions and positions at around  $(\theta_1, \phi_1)$  and  $(x_1, y_1)$ . Also note that the SH-domain sensing matrix based on the plane wave is frequency-invariant.

### C. Translation of Interior Spherical Harmonic Coefficients

One straightforward way to capture the interior SH coefficient  $\alpha_{nm}$  of the primary noise field in (8) from observations of the noise field  $S(\mathbf{r})$  is to place an RMA with respect to the coordinate origin. In this setup, the ROI is surrounded by an RMA as shown in Fig. 1(a). However, distributing the numbers of reference microphones so as to sample the surface of the region is difficult to implement. Alternatively, we can apply the concept of distributed higher-order microphones (HOMs), i.e., typically a spherical or circular microphone array, which is designed to capture the higher-order harmonics, to capture the noise field of a large space with distributed arrays [44]. Instead of uniformly distributing many microphones to enclose the entire ROI, we can reduce the total number of units by clustering the microphones. Fig. 5(a) illustrates the array structure and the defined grid points. We define the interior field as the inside region, whose boundary is a sphere of radius  $R_q$ , which is shown in white in Fig. 5(a). Note that we choose the outermost local origin  $O_q$  to define the boundary. Each RMA should be able to extract SH coefficients up to a certain order. One can use either spherical arrays or any other 3D array structure, which is mentioned in Sect. III-B, with sufficiently large aperture to capture SH coefficients. In this method, the translation relationship of SH coefficients among different expansion origins is utilized to estimate the global SH coefficients by placing HOMs with respect to the local origin

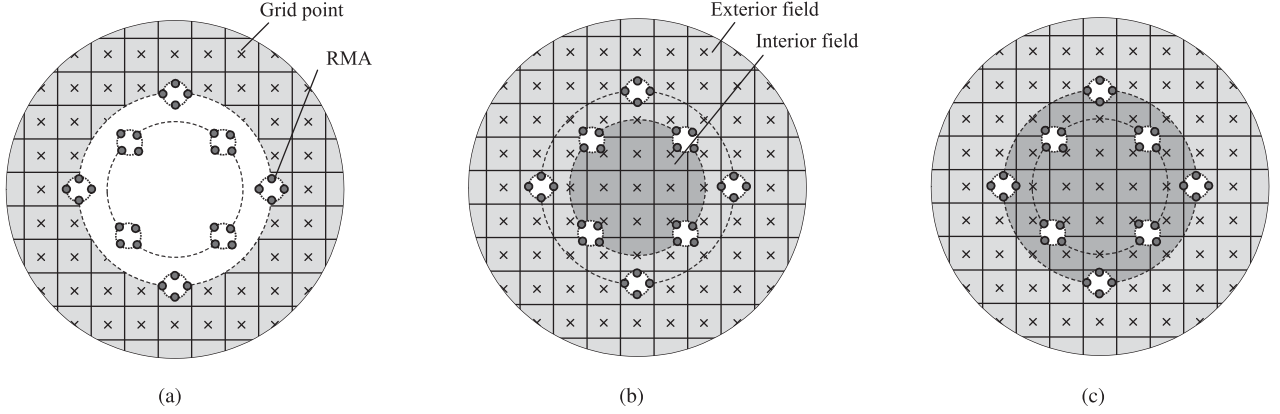


Fig. 5. Array structure and grid point configuration: (a) Grid points of exterior field; (b) Grid points of both interior and exterior sound fields for the inner RMAs. The light and dark regions correspond to the exterior and interior fields, respectively. (c) Grid points of both interior and exterior sound fields for the outer RMAs.

$O_q$ :

$$S_H^{(q)}(\mathbf{r}) = \sum_{v=0}^{\infty} \sum_{u=-v}^v \alpha_{vu}^{(q)} j_v(kr') Y_{vu}(\theta, \phi'), \quad (36)$$

where

$$\alpha_{vu}^{(q)} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \alpha_{nm} T_{nv}^{mu}(\mathbf{r}_q). \quad (37)$$

$T_{nv}^{mu}(\mathbf{r}_q)$  is the translation coupling coefficient [31], [46] between the local and global interior SH coefficient, where  $\mathbf{r}_q = (R_q, \theta_q, \phi_q)$  is the position of the  $q^{\text{th}}$  local origin. Note that we assume the noise sources do not exist inside each HOM. Taking into account all  $Q$  HOMs, we can construct the following matrix formulation:

$$\boldsymbol{\alpha}^{(O')} = \boldsymbol{\tau} \boldsymbol{\Psi} \boldsymbol{\gamma}, \quad (38)$$

where  $\boldsymbol{\alpha}^{(O')} = [\boldsymbol{\alpha}^{(1)T}, \dots, \boldsymbol{\alpha}^{(Q)T}]^T \in \mathbb{C}^{(V+1)^2 Q}$  and  $\boldsymbol{\tau} = [\mathbf{T}^{(1)T}, \dots, \mathbf{T}^{(Q)T}]^T \in \mathbb{C}^{(V+1)^2 Q \times (N+1)^2}$ .  $\boldsymbol{\alpha}^{(O')}$  and  $\boldsymbol{\tau}$  consist of  $Q$  vectors and matrices, whose  $q^{\text{th}}$  block elements are  $\boldsymbol{\alpha}^{(q)} = [\alpha_{00}^{(q)}, \dots, \alpha_{NN}^{(q)}]^T$  and

$$\mathbf{T}^{(q)} = \begin{bmatrix} T_{00}^{00}(\mathbf{r}_q) & \cdots & T_{N0}^{N0}(\mathbf{r}_q) \\ \vdots & \ddots & \vdots \\ T_{0V}^{0V}(\mathbf{r}_q) & \cdots & T_{NV}^{NV}(\mathbf{r}_q) \end{bmatrix} \in \mathbb{C}^{(V+1)^2 \times (N+1)^2}. \quad (39)$$

The truncation order of the global coefficient  $N$  is determined so as to make the sensing matrix,  $\boldsymbol{\tau} \boldsymbol{\Psi}$ , a full-rank matrix. This can be achieved by choosing a sufficiently large order, which satisfies  $(N+1)^2 \geq (V+1)^2 Q$ . As shown in Fig. 5(a), the grid points of the point source are defined to be in the exterior field, which is outside of the outermost-positioned RMAs. The local SH coefficient  $\alpha_{uv}^{(q)}$  can be calculated by measuring the sound pressure on a boundary of a sphere with respect to

the local origin  $O_q$ :

$$\alpha_{uv}^{(q)} \approx \frac{1}{b_n(k\bar{r}_q)} \sum_{a=1}^A \Lambda_a S_r^{(q)}(\mathbf{r}_a) Y_{nm}^*(\theta_a, \phi_a), \quad (40)$$

where  $\bar{r}_q$  is the radius of the  $q^{\text{th}}$  RMA and  $S_r^{(q)}(\mathbf{r}_a)$  is the reference signal of the  $q^{\text{th}}$  RMA.

Finally, the spatial noise distribution  $\boldsymbol{\gamma}$  can be estimated by solving the following minimization problem:

$$\underset{\boldsymbol{\gamma}}{\text{minimize}} \|\boldsymbol{\alpha}^{(O')} - \boldsymbol{\tau} \boldsymbol{\Psi} \boldsymbol{\gamma}\|_2^2 + \lambda \|\boldsymbol{\gamma}\|_p^p. \quad (41)$$

#### D. Translation of Interior and Exterior SH Coefficients

A typical setup of SH-domain spatial ANC assumes noise sources propagating from the exterior field. Therefore, the interior problem of a sound field is usually considered. In the interior problem, it is assumed that the ROI is free from any sound sources. When this assumption is violated, (i.e., sound sources exist in the region), estimation of the sound field fails. To increase the robustness of the ANC system, we develop a way to consider both an interior sound field and exterior sound field. The considered array structure and the grid points are illustrated in Figs. 5(b) and 5(c). Note that there are multiple definitions of the boundary of the interior, and the exterior field depends on which local origin to focus on.

In analogy to the interior sound field representation, the exterior sound field can be written as

$$S_P(\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \check{\beta}_{nm} h_n^{(1)}(kr) Y_{nm}(\theta, \phi), \quad (42)$$

where  $\check{\beta}_{nm}$  is the SH coefficient of the exterior field, and  $h_n^{(1)}(\cdot)$  is the first kind of the  $n^{\text{th}}$ -order spherical Hankel function. From (8) and (42), a composite noise field at a boundary of the interior and exterior sound field can be written as

$$S(\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n [\alpha_{nm} j_n(kr) Y_{nm}(\theta, \phi) + \check{\beta}_{nm} h_n^{(1)}(kr) Y_{nm}(\theta, \phi)]. \quad (43)$$

The translation relationship for the interior field in (36), as well as for the exterior field, can be applied to (43), which leads to

$$S(\mathbf{r}) = \sum_{v=0}^{\infty} \sum_{u=-v}^v \underbrace{[\alpha_{vu}^{(q)} + \beta_{vu}^{(q)}]}_{\zeta_{vu}^{(q)}} j_v(kr') Y_{vu}(\theta, \phi'), \quad (44)$$

where

$$\beta_{vu}^{(q)} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \check{\beta}_{nm} \check{T}_{nv}^{mu}(\mathbf{r}_q). \quad (45)$$

$\check{T}_{nv}^{mu}(\mathbf{r}_q)$  is the translation coupling coefficient [31], [46] between the local interior SH coefficient  $\beta_{vu}^{(q)}$  originating from the interior sources and the global exterior SH coefficient  $\check{\beta}_{nm}$ .

Finally, taking into account all  $Q$  HOMs and applying the translation relationships in (37) and (45), we derive the following matrix formulation:

$$\zeta^{(O')} = \begin{bmatrix} \tau^{(1)} & & & \\ & \tau^{(2)} & & \\ & & \ddots & \\ & & & \tau^{(Q)} \end{bmatrix} \begin{bmatrix} \psi^{(1)} \\ \psi^{(2)} \\ \vdots \\ \psi^{(Q)} \end{bmatrix} \gamma, \quad (46)$$

where  $\zeta^{(O')} = [\zeta^{(1)T}, \dots, \zeta^{(Q)T}]^T \in \mathbb{C}^{(V+1)^2 Q}$ ,  $\tau^{(q)} = [\mathbf{T}^{(q)}, \check{\mathbf{T}}^{(q)}]$ , and

$$\psi^{(q)} = \begin{bmatrix} \Psi^{(q)} & O \\ O & \check{\Psi}^{(q)} \end{bmatrix} \in \mathbb{C}^{2(N+1)^2 \times L}.$$

$\check{\Psi}^{(q)}$  corresponds to the  $q^{\text{th}}$  sensing matrix consisting of exterior point source basis functions and  $L$  is the total number of the grid directions and points correspond to both the interior and exterior sound fields. Similar to (39),  $\check{\mathbf{T}}^{(q)}$  can be written as

$$\check{\mathbf{T}}^{(q)} = \begin{bmatrix} \check{T}_{00}^{00}(\mathbf{r}_q) & \cdots & \check{T}_{N0}^{N0}(\mathbf{r}_q) \\ \vdots & \ddots & \vdots \\ \check{T}_{0V}^{0V}(\mathbf{r}_q) & \cdots & \check{T}_{NV}^{NV}(\mathbf{r}_q) \end{bmatrix} \in \mathbb{C}^{(V+1)^2 \times (N+1)^2}. \quad (47)$$

In contrast to (38), we define different sensing matrices  $\psi^{(q)}$  for every  $Q$  local origin since the distance between the global origin and the local origin defines the boundary between the interior and exterior sound fields, as illustrated in Figs. 5(b) and 5(c). A sparse set of weights can be estimated by solving the following minimization problem:

$$\underset{\gamma}{\text{minimize}} \|\zeta^{(O')} - \tau \psi \gamma\|_2^2 + \lambda \|\gamma\|_p^p. \quad (48)$$

If all the RMAs are placed at the same radius from the global origin, all sensing matrices  $\psi^{(q)}$  become equivalent, which we denote as  $\psi^{(0)}$ . Therefore, the matrix formulation can be

simplified as

$$\zeta^{(O')} = \begin{bmatrix} \tau^{(1)} \\ \tau^{(2)} \\ \vdots \\ \tau^{(Q)} \end{bmatrix} \psi^{(0)} \gamma. \quad (49)$$

## V. RECONSTRUCTION OF REFERENCE SH COEFFICIENT

In this section, we describe the reference signal reconstruction stage of our proposed method shown in Fig. 3. The reference SH coefficients can be reconstructed from the estimated spatial noise distribution  $\hat{\gamma}$  and a reconstruction matrix. The reference signal can be reconstructed in several ways by defining the appropriate reconstruction matrix. We describe several definitions of the reconstruction matrix, which is required to perform the SH-domain FxLMS algorithm afterwards.

First, let us assume that only incoming noise sources exist in the exterior field, which is the same configuration discussed in IV-C. Note that the grid points of the point source are defined to be outside of the outermost-positioned RMA, as shown in Fig. 5(a). Since the global interior SH coefficients of the reference signals are required, we define a reconstruction matrix as  $\mathbf{\Pi} = \check{\Psi}$ , which is the order-truncated sensing matrix in the global SH domain.  $\tilde{N}$  is the required order in the following ANC processing. Thus, the global SH coefficients of the reference signals can be calculated as

$$\mathbf{x} = \mathbf{\Pi} \hat{\gamma}. \quad (50)$$

The proposed algorithm is summarized in Algorithm 1. We use M-FOCUSS algorithm [50] to estimate  $\hat{\gamma}$ . As shown in Algorithm 1, M-FOCUSS is nested in the adaptive ANC processing, hence this structure is computationally expensive in general. However, we confirmed that by initializing  $\bar{\gamma}$  with  $\hat{\gamma}$ , which is the last value in the previous iteration, only few iterations were needed to satisfy the termination condition of M-FOCUSS. This is because the spatial sparsity does not change when the noise sources have the same distribution. Therefore, the computational complexity can be reduced if the distribution of the noise sources does not change rapidly.

Next, we consider the general setup shown in Figs. 5(b) and 5(c), where both the interior and exterior noise sources exist. Target grid points, from which the noise source is expected to be attenuated, correspond to the exterior noise sources, whereas nontarget grid points correspond to the interior noise sources. We define the reconstruction matrix as  $\mathbf{\Pi} = [\check{\Psi}, O] \in \mathbb{C}^{(\tilde{N}+1)^2 \times L}$ , where  $\mathbf{\Pi}$  is now a reconstruction matrix consisting only of the interior field basis functions  $\check{\Psi}$ . The global SH coefficients of the reference signals can be reconstructed using (50). As a result, nontarget noise sources corresponding to the interior noise sources can be removed from the reconstructed reference SH coefficients.

Now, we show a practical definition of the target grid point  $\mathbf{y}_{\text{tar}}$  based on the distance between the local origins and the global origin. When a grid point is closer to the global origin



**Algorithm 1:** Proposed Algorithm.

---

```

i = 1; Initialize  $\mathbf{w}[i] \leftarrow \mathbf{0}$ ;
while i reaches last frame number do
   $\boldsymbol{\alpha}[i] = [\alpha_{00}^{(1)}[i], \dots, \alpha_{NN}^{(Q)}[i]]^T$ ;
  Update  $\boldsymbol{\alpha}[i]$  by calculating (40);
  Update  $\mathbf{e}[i]$  by calculating (25);
  j = 1; Initialize  $\bar{\boldsymbol{\gamma}}[j] \leftarrow \hat{\boldsymbol{\gamma}}[i - 1]$ ;
  // M-FOCUSS
  while stop = 0 and j is smaller
    than predefined maximum value do
     $W[j] \leftarrow \text{diag}(\bar{\boldsymbol{\gamma}}[j]^{(1-p/2)})$ ;
     $Q[j] \leftarrow$ 
     $[(\boldsymbol{\Psi}W[j])^H \boldsymbol{\Psi}W[j] + \lambda I]^{-1} (\boldsymbol{\Psi}W[j])^H \boldsymbol{\alpha}[i]$ ;
     $\bar{\boldsymbol{\gamma}}[j] \leftarrow W[j]Q[j]$ ;
    if Termination condition is satisfied then
       $\hat{\boldsymbol{\gamma}}[i] \leftarrow \bar{\boldsymbol{\gamma}}[j]$ ;
      stop = 1;
    end
    j  $\leftarrow j + 1$ ;
  end
   $\mathbf{x}[i] \leftarrow \Pi \hat{\boldsymbol{\gamma}}[i]$ ;
   $\mathbf{z}[i] \leftarrow \mathbf{g}\mathbf{x}[i]\mathbf{w}[i]$ ;
   $\mathbf{w}[i + 1] \leftarrow \mathbf{w}[i] - \mu(\mathbf{g}\mathbf{x}[i])^H \mathbf{e}[i]$ ;
  i  $\leftarrow i + 1$ ;
end

```

---

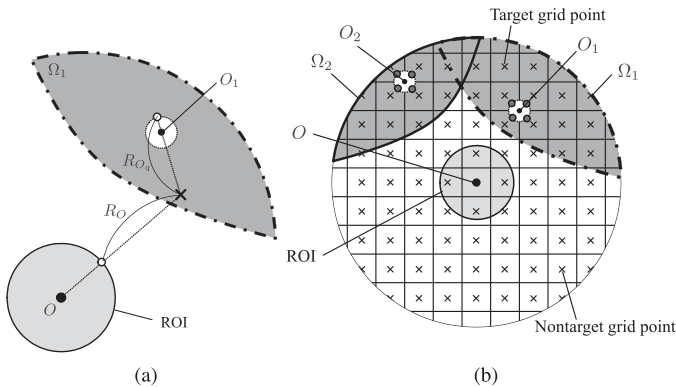


Fig. 6. Definition example of target and nontarget grid points: (a) Illustration of a boundary of the target region; (b) Overall configuration.

than to any local origins, where an RMA exists, a noise wavefront propagates to the ROI before it is captured by the RMA; thus, noise attenuation cannot be achieved in a feedforward structure because of the causality constraint. Let us consider the distance  $R_{O_q}$ , which is from a grid point to the farthest reference microphone of  $q^{\text{th}}$  RMA and  $R_O$ , which is from a grid point to a nearest point of ROI. The 2 distances are shown in Fig. 6(a). The noise wavefront has to be captured by all reference microphones of at least one RMA before it arrives at ROI. Therefore, we define the target grid point as a point that satisfies  $R_{O_q} < R_O$ . We denote  $\mathbf{y}_{\text{tar}}^{(q)} \in \Omega_q$  as the target grid point for the  $q^{\text{th}}$  local origin, where  $\Omega_q$  is the target region defined for  $O_q$ . Finally, the target grid

point can be defined as a point inside a union of each region  $\Omega_q$  as  $\mathbf{y}_{\text{tar}} \in \{\Omega_1 \cup \dots \cup \Omega_Q\}$ . The overall model is described in Fig. 6(b). We can then define the reconstruction matrix by using only the interior basis functions corresponding to the grid point  $\mathbf{y}_{\text{tar}}$ .

The target and nontarget grid points can be arbitrarily defined in the reconstruction stage. Prior information of the target and nontarget regions can be used to define the reconstruction matrix. Therefore, the sound sources from the nontarget region are passed through and only the target noise sources propagating from the target region are attenuated.

## VI. VALIDATIONS

We conducted simulation studies in a 3D free field to evaluate the noise attenuation performance. An RMA, SLA, and EMA are arranged such that they surround the ROI, which is the typical setup of spatial ANC. The RMA captures the reference noise field, whereas the EMA captures and feeds back the residual noise field to update the adaptive filter. The signal-to-noise ratio of 40 dB white Gaussian noise is added to each error and reference microphone output. The SLA produces the antinoise field to attenuate the undesired noise field. We assume an open array structure for the microphone arrays. On the other hand, a rigid baffle array is widely used for an actual implementation of a microphone array. In that case, reflections among HOMs may be taken into account, which are analytically derived in the literature [53], to improve the accuracy of the sound field estimation.

We considered the following array configurations:

1) *Concentric Array Structure*: A typical array setup of spatial ANC is a concentric array configuration, as shown in Fig. 1(a). In this configuration, RMA, SLA, and EMA are placed at the global origin. Microphones and loudspeakers are placed on a surface of a sphere, which is sampled at 20 vertices of a dodecahedron to uniformly sample the sphere. A 20 ch microphone array and SLA can control a sound field up to the third order in theory. Since SH-domain representation relies on an expansion origin, a straightforward formulation can be considered with the common origin. The radii of the RMA, SLA, and EMA are 0.6 m, 0.4 m, and 0.2 m, respectively.

2) *Nonconcentric Array Structure*: A generalized array setup has a nonconcentric array configuration, as shown in Fig. 1(b). This configuration gives more flexibility of RMA placement. For example, an HOM can be placed at a local origin. Moreover, multiple HOMs can be placed at multiple local origins as RMAs, which do not necessarily surround the ROI. In this configuration, we assume that the SLA and EMA are placed at the global origin; in contrast, the RMAs are placed at local origins.

We evaluated and compared the noise attenuation level among the following methods: (i) BL-MIMO corresponds to the temporal-frequency-domain feedforward FxLMS algorithm [1]; (ii) BL-MD corresponds to the SH-domain feedforward FxLMS algorithm, which is a modification of the feedback version based on [10] using the concentric array structure; (iii) SDCS corresponds to the conventional SH-domain feedforward FxLMS algorithm utilizing CS, whose sensing matrix is represented in the

spatial domain [23]; (iv) MDCS corresponds to the SH-domain feedforward FxLMS algorithm utilizing CS by inputting SH coefficients for the reference signal reconstruction; (v) MDCS-C1 uses the same method as (ii) with only the difference in the use of the nonconcentric array structure, whose radius of RMA is 0.1 m placed at (0, 0.6, 0); thus, the array does not surround the entire ROI; (vi) MDCS-C2 uses 2 RMAs with a radius of 0.1 m, which consist of 12 microphones placed on vertexes of an icosahedron. The M-FOCUSS algorithm with  $p = 0.5$  is used to estimate the weight,  $\gamma$ . The filter update equation for each method was modified by its normalized version [54], where the update term was divided by the power of the filtered reference signals. The step size was set to 1 for all methods.

The truncation order for each RMA and EMA can be calculated as  $\tilde{N} = \lceil kR_{\text{mic}} \rceil$ , where  $R_{\text{mic}}$  is the radius of the array. According to this criterion, we obtain the spatial Nyquist frequencies of the RMA and EMA to be 273 Hz and 819 Hz, respectively, for the setup of BL-MD.

156 points of point sources and 110 directions of plane waves, in total 266 grid points and directions, are defined as a set of interior basis functions for the sensing matrix. We only define grid points on a horizontal plane since a rotational symmetry can be assumed. Note that the grid points are not defined inside the ROI for the interior basis functions. The truncation order  $N$  of the translation matrix described in (38) is heuristically determined, i.e.,  $N = 11$ , to satisfy  $(N + 1)^2 \geq (V + 1)^2 Q$ .

#### A. Noise Attenuation Performance in Concentric Array

Two monochromatic point sources of a white noise were used as primary noise sources. Note that uncorrelated noises were used as the primary source signals. The noise attenuation level was evaluated by defining the regional noise reduction (RNR) inside the region as

$$\epsilon_1(\iota) = 10 \log_{10} \frac{\sum_i |S_{(0)}(\mathbf{r}_i)|^2}{\sum_i |S_{(\iota)}(\mathbf{r}_i)|^2}, \quad (51)$$

where  $S_{(\iota)}(\mathbf{r}_i)$  is the residual error at the  $i^{\text{th}}$  point at the  $\iota^{\text{th}}$  iteration and  $S_{(0)}(\mathbf{r}_i)$  is the initial noise field at the  $i^{\text{th}}$  point. The evaluation region is defined as a sphere inside the radius of EMA. There are 3544 uniformly sampled points in the region.

The RNR for the frequency band from 100 Hz to 1000 Hz is shown in Fig. 7 and Fig. 8. An on-grid condition, where the primary sources are on the grid points (1, 2, 0) and (0, 5, 0), and an off-grid condition, where the sources are placed at non-grid points (0.8, 2.2, 0) and (0.3, 5, 0), are evaluated. The result obtained by calculating the average of 10 simulation runs. The proposed MDCS outperformed BL-MIMO and BL-MD in most frequency bins. It can be seen that the spatial aliasing of the reference signal affects the performance of BL-MD. On the other hand, MDCS can maintain its performance beyond the spatial Nyquist frequency of RMA. Fig. 9 shows the convergence performance and Fig. 10 shows the residual noise field on  $x$ - $y$  plane after 100 iterations of ANC at 500 Hz. A simple moving average is taken, where each average is calculated over 10 neighbor iterations to smooth the fluctuation of the result originating from the random amplitude of the primary noise

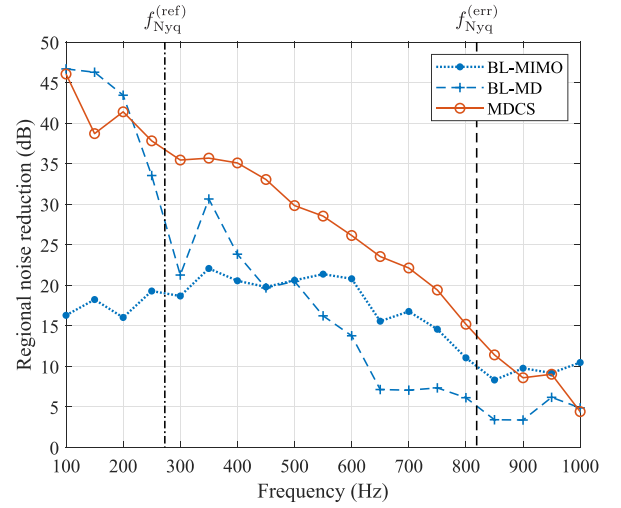


Fig. 7. Regional noise reduction with on-grid condition after 100 iterations of FxLMS algorithm for various frequencies.  $f_{\text{Nyq}}^{(\text{ref})}$  and  $f_{\text{Nyq}}^{(\text{err})}$  are the Nyquist frequencies of RMA and EMA, respectively.

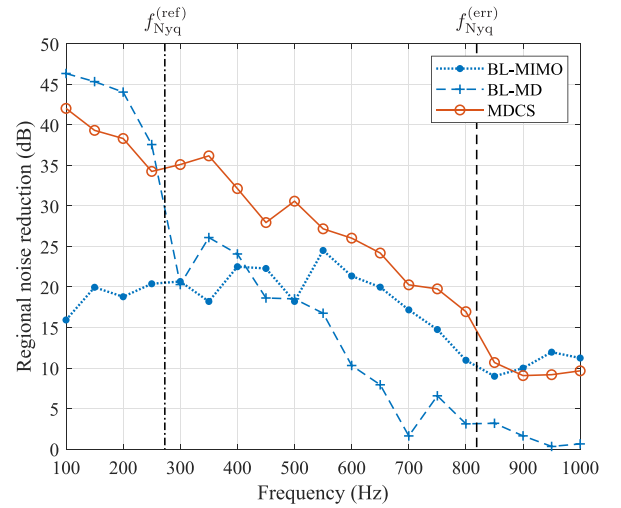


Fig. 8. Regional noise reduction with off-grid condition after 100 iterations of FxLMS algorithm for various frequencies.

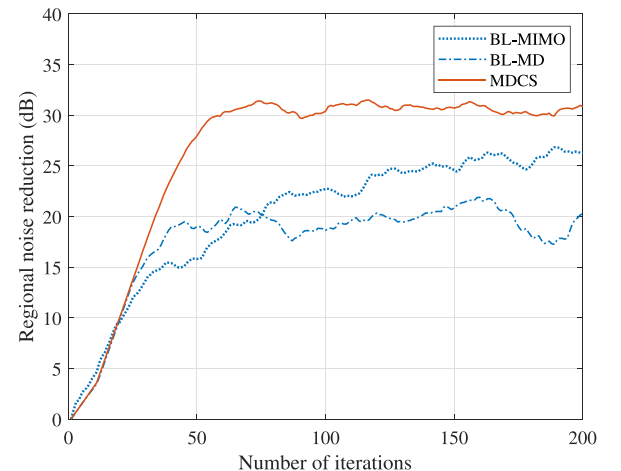


Fig. 9. Regional noise reduction for each iteration index.

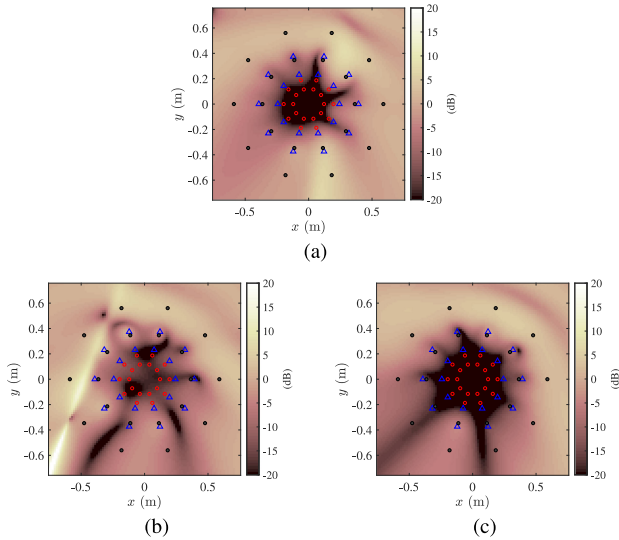


Fig. 10. Residual noise field after 100 iterations of FxLMS algorithm. (a) BL-MIMO. (b) BL-MD. (c) MDCS. Black dots, blue triangles, and red circle indicate the reference microphones, loudspeakers, and error microphones, respectively. Note that all microphones and loudspeakers are plotted by mapping to the  $x$ - $y$  plane.

source and measurement noise. From Fig. 9, we can see that the proposed method gives the highest RNR with a smaller number of iterations among the methods.

We compared the computational complexity of the filter update among BL-MIMO, BL-MD, and MDCS, which were  $\mathcal{O}(L_s(L_e + L_r))$ ,  $\mathcal{O}((N + 1)^2)$ , and  $\mathcal{O}(LL_r^2)$ , respectively.  $L_e$  and  $L_r$  are the number of the error and reference microphones, respectively.  $L_e$  and  $L_r$  are usually larger than  $(N + 1)^2$ ; thus BL-MD has the least computational complexity. Although MDCS requires iterative processing in M-FOCUSS, the number of iterations can be reduced as mentioned in Sect. V. Therefore, the computational order of MDCS can be regarded as  $L$  times higher than BL-MIMO.

### B. Evaluation of Different RMA Structures

The proposed method provides flexibility of the RMA structure and placement because RMAs can be located at local origins, as shown in Fig. 1(b). We compared the RNR among four setups, i.e., SDCS, MDCS, MDCS-C1, and MDCS-C2. One RMA was placed at  $(r, \theta, \phi) = (0.6, \pi/2, \pi/2)$  in SDCS, MDCS, and MDCS-C1, and 2 RMAs were placed at  $(0.6, \pi/2, \pi/4)$  and  $(0.6, \pi/2, 3\pi/4)$  in MDCS-C2.

The RNR is shown in Fig. 11. The proposed methods outperformed SDCS below 700 Hz. As shown in Fig. 4, the high coherence of the sensing matrix in SDCS deteriorates the accuracy of the sparse decomposition of the reference signals at low frequency. MDCS-C1 showed a comparable performance to MDCS above 600 Hz even when a compact RMA was used; however, at low frequencies, the performance degraded. This is due to the lack of spatial resolution of the captured reference signal [25], [26]. We can improve the performance at low frequencies by using distributed microphone arrays to sample the noise field with a higher spatial resolution. Fig. 11

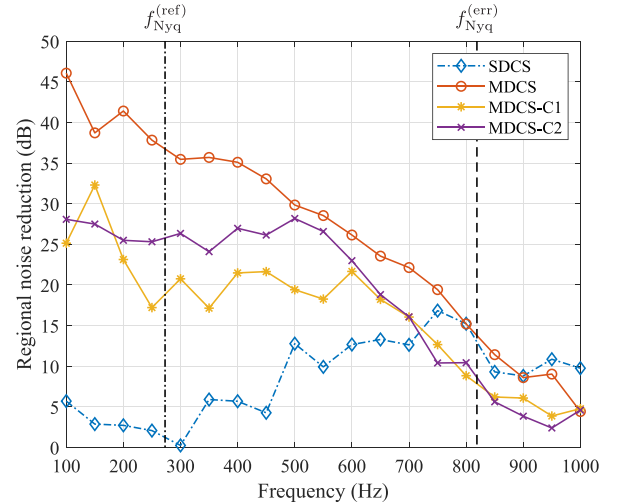


Fig. 11. Regional noise reduction after 100 iterations of FxLMS algorithm for various frequencies.

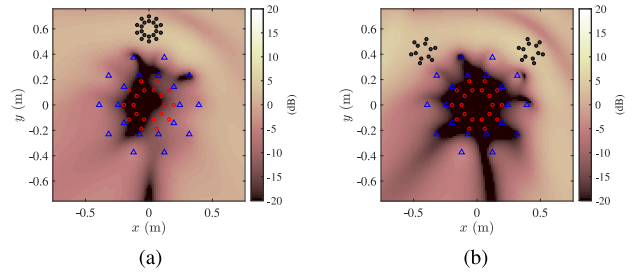


Fig. 12. Residual noise field after 100 iterations of FxLMS algorithm. (a) MDCS-C1. (b) MDCS-C2.

shows that by using 2 RMAs in MDCS-C2, we can improve the RNR at frequencies below 600 Hz compared with using 1 RMA in MDCS-C1, since the 2 RMAs can capture larger spatial information of the noise field. Fig. 12 shows the residual noise field on  $x$ - $y$  plane after 100 iterations of ANC at 500 Hz.

### C. Interior Noise Compensation

A fundamental problem of SH-domain sound field analysis is that the interior and exterior sound fields cannot be represented at the same time. Since we focus on controlling the interior sound field, all sound sources must exist outside of the ROI. However, in practice, a sound source may easily be generated inside the region, e.g., a human voice. By the proposed method, we can estimate the sound source distribution to distinguish noise sources, which are generated inside the region and removed from the ANC processing.

We defined the interior region as a sphere, whose radius is the distance from the global origin to the center of the RMA. The array configuration is same as MDCS-C2 described in Sect. VI-B. We simulated an initial noise field consisting of a point source placed at  $(x, y, z) = (0.5, 1, 0)$ , which existed outside of the region. While performing the ANC processing, we added another point source placed at  $(0.5, 0, 0)$  inside the region at the 30<sup>th</sup> iteration and simulated a composite noise field consisting of

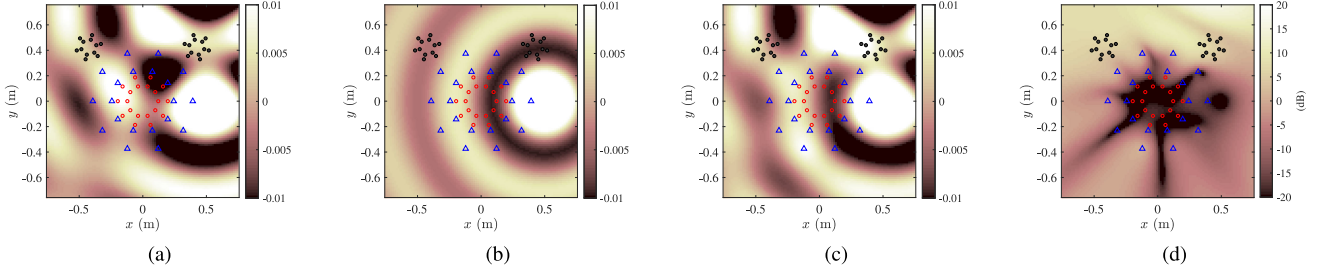


Fig. 13. Simulated sound fields of interior noise compensation: (a) Composite noise field; (b) Noise field of interior noise source; (c) Residual noise field after noise attenuation; (d) Error distribution based on (52).

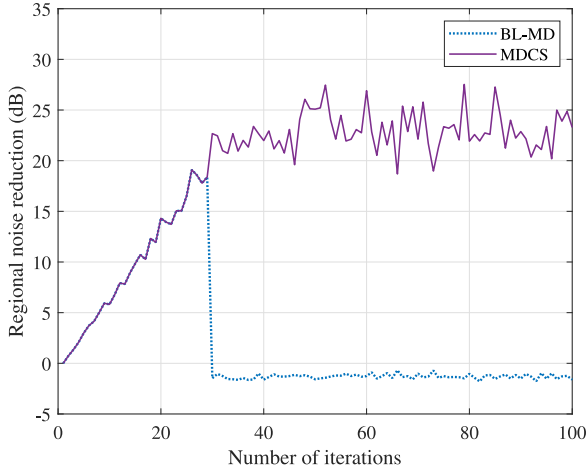


Fig. 14. Average performance of the incoming noise field attenuation. Interior noise source is added at the 30<sup>th</sup> iteration.

two point sources, as shown in Fig. 13(a). The frequency of the noise sources was 600 Hz. The noise field consisting of only the interior noise source is shown in Fig. 13(b). Only the exterior noise source is expected to be attenuated; thus, the interior noise field remains after the convergence of the adaptive processing. The resulting noise field is shown in Fig. 13(c), from which we can confirm that the noise field of the interior noise source is reproduced in the ROI. The noise attenuation performance of the incoming noise field was evaluated by calculating the RNR inside the region using (51) from the 1<sup>st</sup> iteration to the 29<sup>th</sup> iteration. After the 30<sup>th</sup> iteration, we defined another measure as

$$\epsilon_2(\ell) = 10 \log_{10} \frac{\sum_i |S_{(\ell)}(\mathbf{r}_i) - \check{S}_{(\ell)}(\mathbf{r}_i)|^2}{\sum_i |\check{S}_{(\ell)}(\mathbf{r}_i)|^2}, \quad (52)$$

where  $\check{S}_{(\ell)}(\mathbf{r}_i)$  is the outgoing noise field generated by the interior noise source, which is expected to be removed from the ANC processing in the proposed method. The final error distribution after convergence is shown in Fig. 13(d).

Fig. 14 shows the convergence performance for BL-MD and MDCS-C2. We can clearly see that BL-MD no longer attenuates the incoming noise field after the 30<sup>th</sup> iteration. In contrast, MDCS-C2 can detect the interior noise source, which we do not want to attenuate and thus remove it from the ANC processing; thus, there is no impact of the existence of the interior noise

source on the convergence performance. This separation of the target and nontarget sources could be also used to remove the acoustic feedback from the SLA to the RMAs.

The separation of a target region, whose noise sources are attenuated, and a nontarget region is not limited to the interior and the exterior regions. Consider a situation in which you want to listen to a sound, e.g., music from your front but you do not want to hear noise, e.g., traffic noise from behind. To separate noise sources into target noise sources, which are propagated from the target region, and nontarget noise sources, directional microphones or beamforming using HOMs may be utilized. However, narrow directivity is difficult to achieve at low frequencies thereby affecting the ANC performance owing to the capture of signals of nontarget noise sources. The proposed method can be applied to overcome this problem.

#### D. Actual Room Experiment

We conducted measurements in our audio laboratory to evaluate the noise attenuation performance in a real-world environment. Main differences from the ideal setup is that the acoustic transfer function contains reverberations and primary and secondary sources are not necessarily point sources. The Eigenmike was used as an RMA as well as an EMA and a 30 channel spherical loudspeaker array was used as an SLA. The radii of the RMA, SLA, and EMA are 0.042 m, 1.0 m, and 0.042 m, respectively. The EMA was placed inside the SLA at the center point. The RMA and primary loudspeaker, which generates the primary noise, were collinearly placed 1.7 m and 3.18 m away from the center point. SH coefficients of the secondary path can be calculated from the measured secondary path  $\mathbf{G}_{\text{act}} \in \mathbb{C}^{L_e \times L_s}$ . We write the matrix form of (25) as  $\mathbf{e} = \mathbf{T}_e \mathbf{S}_e$ , where  $\mathbf{T}_e \in \mathbb{C}^{(N+1)^2 \times L_e}$  is the SH transformation matrix. From (11) and (13), the SH coefficients of the secondary path can be derived in a matrix form as

$$\mathbf{g}_{\text{act}} = \mathbf{T}_e \mathbf{G}_{\text{act}} \mathbf{Y}, \quad (53)$$

where  $\mathbf{Y} \in \mathbb{C}^{L_s \times (N+1)^2}$  is the matrix consisting of SH functions. We confirmed that the diagonal elements of  $\mathbf{g}_{\text{act}}$  were dominant in our experimental condition. Hence, we extracted the diagonal elements and diagonalized the matrix similar to the related work [55].

The average noise attenuation level at the error microphones are shown in Fig. 15. The noise attenuation was achieved at the



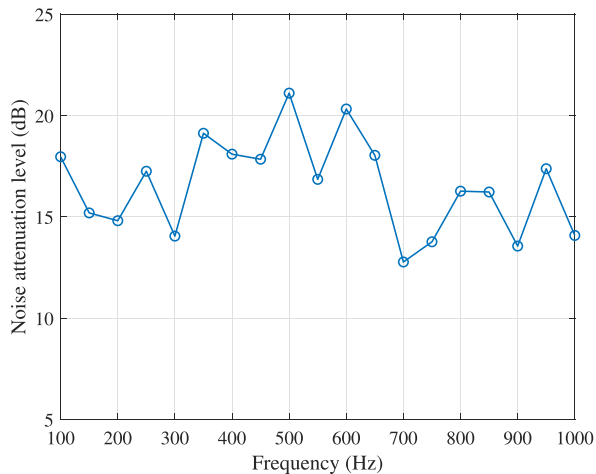


Fig. 15. Average noise attenuation level after 100 iterations of FxLMS algorithm for various frequencies.

frequency range of 100 to 1000 Hz. Unlike the results of the simulation in Sect. VI-A and VI-B, performance degradation at high frequencies cannot be seen. A possible reason is that the evaluation points are only at the error microphone positions in this experiment. The evaluation in terms of the sizable region in a real-world environment is a consideration for future study.

## VII. CONCLUSION

In this paper, we proposed the reference signal reconstruction algorithm for SH-domain feedforward ANC. We assumed that the primary noise field can be represented as a sum of a sparse set of point sources and plane waves and adopted the CS approach to decompose the noise field into them. The decomposition accuracy relies on the column correlation of the sensing matrix in CS; thus, we showed that the SH basis function is one of the efficient spaces to decompose the sound field. We evaluated the proposed method and showed that its noise attenuation performance exceeded that of the conventional method at any frequency above the Nyquist frequency of the reference microphone array. Another advantage exploited was that both the interior and exterior sound fields can be considered at once, which is not possible in the case of the general SH-domain signal representation. This property helps practical implementation since an unexpected noise source can be removed from the system. Furthermore, an arbitrary array structure can be used in our method, which also leads to the realization of a viable system.

## ACKNOWLEDGMENT

The authors would like to thank Ms. Jihui Zhang for collecting the multichannel recordings that are used in Section VI-D.

## REFERENCES

- [1] S. M. Kuo and D. R. Morgan, *Active Noise Control Systems: Algorithms and DSP Implementations*. New York: Wiley, 1995.
- [2] S. M. Kuo and D. R. Morgan, "Active noise control: A tutorial review," *Proc. IEEE*, vol. 87, no. 6, pp. 943–973, Jun. 1999.
- [3] Y. Kajikawa, W.-S. Gan, and S. M. Kuo, "Recent advances on active noise control: Open issues and innovative applications," *APSIPA Trans. Signal Inf. Process.*, vol. 1, pp. 1–21, Dec. 2012.
- [4] S. Elliott, I. Stothers, and P. Nelson, "A multiple error LMS algorithm and its application to the active control of sound and vibration," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 10, pp. 1423–1434, Oct. 1987.
- [5] S. C. Douglas, "Fast implementations of the filtered-X LMS and LMS algorithms for multichannel active noise control," *IEEE Speech Audio Process.*, vol. 7, no. 4, pp. 454–465, Jul. 1999.
- [6] M. Bouchard and S. Norcross, "Computational load reduction of fast convergence algorithms for multichannel active noise control," *Signal Process.*, vol. 83, no. 1, pp. 121–134, Jan. 2003.
- [7] D. R. Morgan, "An adaptive modal-based active control system," *J. Acoust. Soc. Amer.*, vol. 89, no. 1, pp. 248–256, Jan. 1991.
- [8] S. Spors and H. Buchner, "Efficient massive multichannel active noise control using wave-domain adaptive filtering," in *Proc. IEEE 3rd Int. Symp. Commun., Control Signal Process.*, St. Julians, Malta, Mar. 2008, pp. 1480–1485.
- [9] J. Zhang, W. Zhang, and T. D. Abhayapala, "Noise cancellation over spatial regions using adaptive wave domain processing," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, New Paltz, NY, USA, Oct. 2015, pp. 1–5.
- [10] J. Zhang, T. D. Abhayapala, W. Zhang, P. N. Samarasinghe, and S. Jiang, "Active noise control over space: A wave domain approach," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 4, pp. 774–786, Apr. 2018.
- [11] W. Zhang, C. Hofmann, M. Buerger, T. D. Abhayapala, and W. Kellermann, "Spatial noise-field control with online secondary path modeling: A wave-domain approach," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 12, pp. 2355–2370, Dec. 2018.
- [12] T. D. Abhayapala and A. Gupta, "Spherical harmonic analysis of wave-fields using multiple circular sensor arrays," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 6, pp. 1655–1666, Aug. 2010.
- [13] Y. Maeno, Y. Mitsufuji, P. N. Samarasinghe, and T. D. Abhayapala, "Mode-domain spatial active noise control using multiple circular arrays," in *Proc. Intl. Workshop Acoust. Signal Enhancement*, Tokyo, Japan, Sep. 2018, pp. 441–445.
- [14] H. Chen, P. Samarasinghe, T. D. Abhayapala, and W. Zhang, "Spatial noise cancellation inside cars: Performance analysis and experimental results," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, New Paltz, NY, USA, Oct. 2015, pp. 1–5.
- [15] P. N. Samarasinghe, W. Zhang, and T. D. Abhayapala, "Recent advances in active noise control inside automobile cabins: Toward quieter cars," *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 61–73, Nov. 2016.
- [16] S. Koyama, S. Shimauchi, and H. Ohmuro, "Sparse sound field representation in recording and reproduction for reducing spatial aliasing artifacts," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Florence, Italy, May 2014, pp. 4443–4447.
- [17] S. Koyama, N. Murata, and H. Saruwatari, "Sparse sound field decomposition for super-resolution in recording and reproduction," *J. Acoust. Soc. Amer.*, vol. 143, no. 6, pp. 3780–3795, Jun. 2018.
- [18] Y. Takida, S. Koyama, and H. Saruwatari, "Exterior and interior sound field separation using convex optimization: Comparison of signal models," in *Proc. 26th Eur. Signal Process. Conf.*, Rome, Italy, Sep. 2018, pp. 2549–2553.
- [19] N. Epain, C. Jin, and A. Van Schaik, "The application of compressive sampling to the analysis and synthesis of spatial sound fields," in *Proc. Audio Eng. Soc. 127th Conv.*, New York, NY, USA, Oct. 2009.
- [20] T. Tachikawa, K. Yatabe, and Y. Oikawa, "3D sound source localization based on coherence-adjusted monopole dictionary and modified convex clustering," *Appl. Acoust.*, vol. 139, pp. 267–281, Oct. 2018.
- [21] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 1, pp. 135–143, Jan. 2005.
- [22] J. Zhang, T. D. Abhayapala, P. N. Samarasinghe, W. Zhang, and S. Jiang, "Sparse complex FxLMS for active noise cancellation over spatial regions," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Shanghai, China, Mar. 2016, pp. 524–528.
- [23] Y. Maeno, Y. Mitsufuji, and T. D. Abhayapala, "Mode domain spatial active noise control using sparse signal representation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Calgary, AB, Canada, Apr. 2018, pp. 211–215.
- [24] J. A. Tropp, "Greedy is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

- [25] A. Xenaki, P. Gerstoft, and K. Mosegaard, "Compressive beamforming," *J. Acoust. Soc. Amer.*, vol. 136, no. 1, pp. 260–271, Jul. 2014.
- [26] E. J. Candes, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. Comput. Harmonic Anal.*, vol. 31, no. 1, pp. 59–73, Jul. 2011.
- [27] S. A. Verburg and E. Fernandez-Grande, "Reconstruction of the sound field in a room using compressive sensing," *J. Acoust. Soc. Amer.*, vol. 143, no. 6, pp. 3770–3779, Jun. 2018.
- [28] Y. Wang and K. Chen, "Compressive sensing based spherical harmonics decomposition of a low frequency sound field within a cylindrical cavity," *J. Acoust. Soc. Amer.*, vol. 141, no. 3, pp. 1812–1823, Mar. 2017.
- [29] Y. Wang and K. Chen, "Sparse plane wave decomposition of a low frequency sound field within a cylindrical cavity using spherical microphone arrays," *J. Sound Vib.*, vol. 431, pp. 150–162, Sep. 2018.
- [30] B. Rafaely, *Fundamentals of Spherical Array Processing*. New York: Springer, 2015.
- [31] P. N. Samarasinghe, T. D. Abhayapala, and M. A. Poletti, "3D spatial soundfield recording over large regions," in *Proc. Intl. Workshop Acoust. Signal Enhancement*, Aachen, Germany, Sep. 2012, pp. 1–4.
- [32] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. New York: Academic press, 1999.
- [33] F. Fazi, M. Noisternig, and O. Warusfel, "Representation of sound fields for audio recording and reproduction," in *Proc. Annu. Meeting Inst. Acoust.*, Nantes, France, Apr. 2012.
- [34] G. Chardon, T. Nowakowski, J. De Rosny, and L. Daudet, "A blind dereverberation method for narrowband source localization," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 5, pp. 815–824, Aug. 2015.
- [35] Y. Takida, S. Koyama, N. Ueno, and H. Saruwatari, "Gridless sound field decomposition based on reciprocity gap functional in spherical harmonic domain," in *Proc. IEEE Int. Sensor Array Multichannel Signal Process. Workshop (SAM)*, Sheffield, U.K., Jul. 2018, pp. 627–631.
- [36] S. Koyama and L. Daudet, "Sparse representation of a spatial sound field in a reverberant environment," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 1, pp. 172–184, Mar. 2019.
- [37] P. N. Samarasinghe, M. A. Poletti, S. A. Salehin, T. D. Abhayapala, and F. M. Fazi, "3D soundfield reproduction using higher order loudspeakers," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Vancouver, BC, Canada, May 2013, pp. 306–310.
- [38] P. Samarasinghe, T. Abhayapala, M. Poletti, and T. Betlehem, "An efficient parameterization of the room transfer function," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 12, pp. 2217–2227, Dec. 2015.
- [39] Y. Iwamatsu, K. Fujii, and M. Muneyasu, "Frequency domain method to estimate the coefficients of feedback control filter for active noise control systems," *Acoustical Sci. Technol.*, vol. 27, no. 5, pp. 264–269, 2006.
- [40] M. T. Akhtar, M. Abe, and M. Kawamata, "On active noise control systems with online acoustic feedback path modeling," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 2, pp. 593–600, Feb. 2007.
- [41] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2542–2556, Jun. 2007.
- [42] D. B. Ward and T. D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Speech Audio Process.*, vol. 9, no. 6, pp. 697–707, Sep. 2001.
- [43] I. Balmages and B. Rafaely, "Open-sphere designs for spherical microphone arrays," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 2, pp. 727–732, Feb. 2007.
- [44] P. Samarasinghe, T. Abhayapala, and M. Poletti, "Wavefield analysis over large areas using distributed higher order microphones," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 3, pp. 647–658, Mar. 2014.
- [45] H. Chen, T. D. Abhayapala, and W. Zhang, "Theory and design of compact hybrid microphone arrays on two-dimensional planes for three-dimensional soundfield analysis," *J. Acoust. Soc. Amer.*, vol. 138, no. 5, pp. 3081–3092, Nov. 2015.
- [46] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Elsevier, Oxford, U.K., 2005.
- [47] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Process. Lett.*, vol. 14, no. 10, pp. 707–710, Oct. 2007.
- [48] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statistical Soc., Series B*, vol. 58, no. 1, pp. 267–288, 1996.
- [49] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUS: A re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 600–616, Mar. 1997.
- [50] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.
- [51] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Las Vegas, NV, USA, Mar./Apr. 2008, pp. 3869–3872.
- [52] V. I. Lebedev and D. Laikov, "A quadrature formula for the sphere of the 131st algebraic order of accuracy," *Doklady Math.*, vol. 59, pp. 477–481, 1999.
- [53] P. A. Martin, *Multiple Scattering: Interaction of Time-Harmonic Waves with N Obstacles*. Cambridge Univ. Press, 2006.
- [54] S. Elliott, *Signal Processing for Active Control*. Elsevier, 2000.
- [55] N. Murata, J. Zhang, Y. Maeno, and Y. Mitsufuji, "Global and local mode-domain adaptive algorithms for spatial active noise control using higher-order sources," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Brighton, U.K., May 2019, pp. 526–530.



**Yu Maeno** received the B.E. degree in information technology, and the M.E. degree in information processing from the Tokyo Institute of Technology, Tokyo, Japan, in 2011 and 2013, respectively. He is currently an Engineer with Sony Corporation, Tokyo, Japan. From May 2017 to October 2018, he was a Visiting Researcher with the Applied Signal Processing Group, Australian National University, Canberra, ACT, Australia. His research interests include spatial sound recording, analysis, and reproduction and spatial noise cancellation.



**Yuki Mitsufuji** received the B.S. and M.S. degrees in information science from Keio University, Tokyo, Japan, in 2002 and 2004, respectively, where he is currently working toward the Ph.D. degree. He is a Deputy General Manager with Speech and Music Group, Sony Corporation, Tokyo, Japan. He led teams that developed the sound design for the PlayStation game title called "Gran Turismo Sport," and spatial audio solution called "Sonic Surf VR." In 2004, he joined Audio Technology Development Group, Sony Corporation. From 2011 to 2012, he was a Visiting

Researcher with Analysis/Synthesis Team, Institut de Recherche et Coordination Acoustique/Musique (IRCAM), Paris, France. He was involved in the 3DTV content search project sponsored by European Project FP7, in research collaboration with IRCAM. He is a Reviewer with ICASSP, INTERSPEECH, etc. Recently, he became a General Chair of Signal Separation Evaluation Campaign where his team had scored the best results for three consecutive years. He has numerous granted patents for audio signal processing.



**Prasanga N. Samarasinghe** received the B.E. degree (Hons.) from the University of Peradeniya, Peradeniya, Sri Lanka, in 2009, and the Ph.D. degree from the Australian National University (ANU), Canberra, ACT, Australia, in 2015. In 2019, she received a prestigious Fulbright Future Scholarship to visit the University of Maryland, College Park, MD, USA. She is currently a Future Engineering Research Leadership (FERL) Fellow and a Senior Lecturer with ANU. Her research interests include spatial sound recording, reproduction and analysis, room acoustics, spatial noise cancellation, and virtual acoustics.



**Naoki Murata** (S'16–M'17) received the B.E. degree in engineering and the M.S. degree in information science and technology from the University of Tokyo, Tokyo, Japan, in 2015, and 2017, respectively. He is currently an Engineer with Sony Corporation, Tokyo, Japan. His research interests include acoustic signal processing and numerical optimization. He was the recipient of the 12th Best Student Presentation Award of the Acoustical Society of Japan in 2016.



**Thushara D. Abhayapala** received the B.E. degree (hons.) in engineering in 1994 and the Ph.D. degree in telecommunications engineering in 1999, both from Australian National University (ANU), Canberra, ACT, Australia. He held a number of leadership positions including Deputy Dean of the College of Engineering and Computer Science (2015–2019), Head of the Research School of Engineering, ANU (2010–2014), and the Leader of the Wireless Signal Processing Program, National ICT Australia (NICTA) from 2005–2007. He worked with the industry for two years, before his doctoral study and has active collaboration with a number of companies. Among many contributions, he is one of the first researchers to use spherical harmonic based eigen-decomposition in microphone arrays and to propose the concept of spherical microphone arrays; novel contributions on the multizone sound field reproduction problem; was one of the first to show the fundamental limits of spatial sound field reproduction using arrays of loudspeakers and spherical harmonics. He has supervised 37 Ph.D. students and coauthored more than 280 peer-reviewed papers. His research interests are in the areas of spatial audio and acoustic signal processing, and multichannel signal processing. He was an Associate Editor for IEEE/ACM TRANSACTION ON AUDIO, SPEECH, AND LANGUAGE PROCESSING and was a Member of the Audio and Acoustic Signal Processing Technical Committee (2011–2016) of the IEEE Signal Processing Society. He is a fellow of Engineers Australia (IEAust).