Spatial Active Noise Control Based on Spatial Frequency Analysis Using Practical Microphone-Loudspeaker Array Structures

March 2022

Yu Maeno

Spatial Active Noise Control Based on Spatial Frequency Analysis Using Practical Microphone-Loudspeaker Array Structures

Graduate School of Science and Technology
Degree Programs in Systems and Information Engineering
University of Tsukuba

March 2022

Yu Maeno

Contents

List of Figures

Acronyms

ANC Active Noise Control

LMS Least Mean Squares

FXLMS Filtered-X Least Mean Squares

FXNLMS Filtered-X Normalized Least Mean Squares

WFS Wave Field Synthesis

HOA Higher Order Ambisonics

MADI Multichannel Audio Digital Interface

ROI Region of Interest

SH Spherical Harmonic

CH Circular Harmonic

CS Compressive Sensing

FFT Fast Fourier Transform

IFFT Inverse Fast Fourier Transform

EMA Error Microphone Array

SLA Secondary Loudspeaker Array

RMA Reference Microphone Array

BPD Basis Pursuit Denoising

IRLS Iteratively Reweighted Least Squares

CHFF Circular-Harmonic-domain Feedforward

SDR Signal-to-Distortion Ratio

DOA Direction of Arrival

HOM Higher-Order Microphone

RNR Regional Noise Reduction

Chapter 1

Introduction

1.1 Background

The progress of industrial technology is making our daily life more and more convenient. Meanwhile, we are facing environmental disasters caused by it. One of the most common environmental disasters is noise pollution. An innovation to solve this issue is expected in the transportation industry, for example, to create a quiet space inside and outside of aircraft, trains, and automobiles. There are primarily two approaches to dealing with this issue. First is the passive noise control shown in Figure 1.1a, which uses sound-absorbing materials to prevent unwanted noise from propagating through the air [1, 2]. This method is effective for noise attenuation at high frequencies. However, due to the long acoustic wavelength of low-frequency noise, a massive sound absorber is required to attenuate the noise. Another approach is the active noise control (ANC) shown in Figure 1.1b, which uses actuators or loudspeakers to cancel out the noise by producing an antinoise sound. When compared to the passive approach, which is ineffective or expensive, the ANC approach can efficiently attenuate noise at low frequencies. As noise-canceling earphones and headphones become more popular in the consumer market, several products incorporating the ANC feature are becoming more common.

ANC is a technique by which a noise signal is attenuated by generating an antinoise signal from the actuators. The simplest single-channel ANC system is depicted in Figure 1.2. The antinoise signal is produced by driving the actuator, whose signal is generated by applying the linear filter to the reference signal captured by the reference sensor. The error sensor detects residual error signals, which can then be fed back into an adaptive algorithm, such as the least mean squares (LMS) algorithm, to adaptively update the filter. The filtered-X least mean squares (FXLMS) algorithm and its variants [3–7] are well known examples of such an approach. Furthermore, the system's multiple-channel extension has been developed to control multiple points with multiple sensors and actuators. There are some challenges in the multiple-channel ANC, which will be mentioned in Section 1.5.

In the context of sound controlling in a large space, studies on the research field called *sound field representation* are intimately involved. The sound field repre-

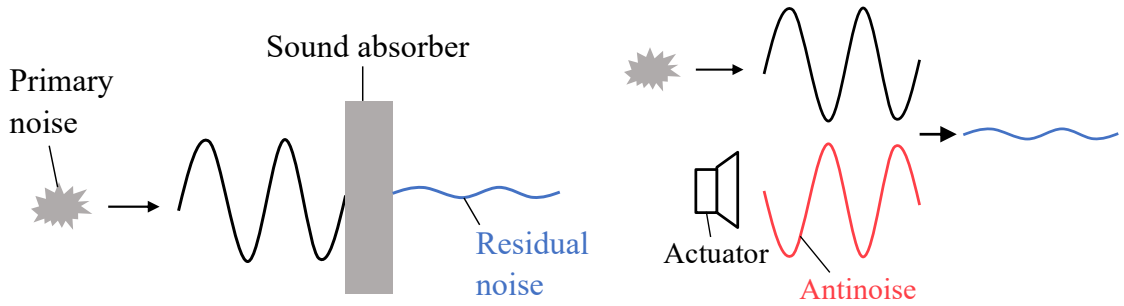


Figure 1.1: ANC approaches: (a) Passive control. (b) Active control.

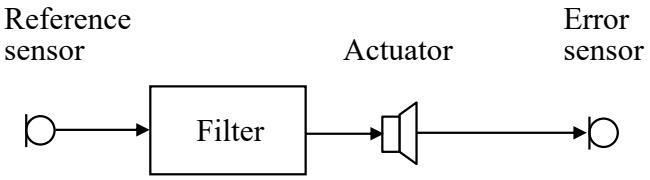


Figure 1.2: Single-input-single-output system.

sentation theory is based on the physical modeling of the sound field propagation characteristics. The well known sound field synthesis approaches are the *wave field synthesis (WFS)* [8] and *higher order ambisonics (HOA)* [9]. The theory itself has been studied for a few decades using computer simulation; however, real-world implementation has been hampered by hardware availability. The implementation based on these methods assumes that many audio channels are independently and concurrently driven. Figure 1.3 illustrates the sound field reproduction using multiple loudspeakers, whose amplitude and phase are individually tuned to reproduce the desired sound field. In recent years, there has been progressing in audio hardware, which has broken through this difficulty. Several audio interfaces with digital I/O formats such as the *multichannel audio digital interface (MADI)* [10] and *Dante* are commercially available. This situation changes accelerated the research and development in the research field.

1.2 Objectives

As mentioned in Section 1.1, the ANC and sound field synthesis techniques have a high affinity in the context of controlling a sound field in a large space. The combination of these two research fields has the potential to achieve noise field attenuation across a large area. Meanwhile, the addition of sound field synthesis techniques increases the system's complexity. This is because a sufficient number

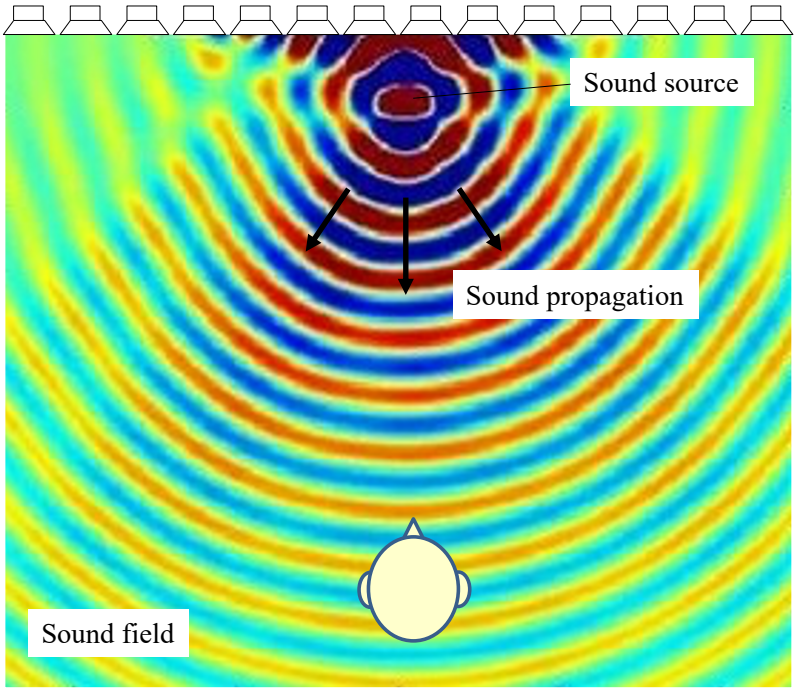


Figure 1.3: Illustration of sound field reproduction.

of sensors and actuators have to be introduced to accurately analyze and reproduce the sound field based on the theory. The goal of this thesis is to develop spatial ANC system leveraging the sound field synthesis techniques. The system feasibility is focused on, and several approaches to addressing the difficulty are investigated.

1.3 Applications

Spatial ANC systems are expected to achieve noise attenuation without wearing any earmuffs and headphones. The systems have the potential to be used in a wide range of situations, including public spaces, private spaces for a small number of people, and personal spaces. Listed below are some examples of applications:

Transportation. Including aircraft, trains, buses, and so on. Engine noise, road noise, and traffic noise are the target noise to be attenuated. Passengers can stay in a comfortable space and can hear announcements easily. Furthermore, the passengers no need to amplify the volume of the listening entertainment content while traveling.

Automotive. Including cars, vans, trucks, and so on. Engine noise, road noise, traffic noise, and wind noise are the target noise to be attenuated. Passengers can enjoy in-car entertainment such as listening to music and watching movies. There is no need to shout to each other to have a conversation with other passengers.

Room. Including a room at home, meeting room, telephone booth, and so on. The target noises to be reduced are construction noise, traffic noise, and speaking noise. One can be free of construction and traffic noise, allowing them to focus on their work.

1.4 Contributions

In this thesis, it is aimed to achieve the improvement of the noise attenuation performance as well as the improvement of the system feasibility in spatial ANC system. The summary of the contributions is listed as follows:

1. Introduction of the compressive sensing (CS) approach to spatial ANC.

   The CS approach is successfully adopted in spatial ANC system to improve the noise attenuation performance. The spatial Nyquist theorem's performance limitation has been overcome. To improve the system's noise attenuation, we define several types of sensing matrices in the transformed domain.

2. Development of the feasible microphone-loudspeaker array structure.

   Instead of using equally sampled spherical arrays, we propose a feasible array structure such as the multiple circular arrays, which are easy to install in practice. Furthermore, for spatial ANC, the concept of distributed sensor arrays with sensor placement flexibility is proposed.

3. Investigation of the novel adaptive algorithm.

   A novel weighted LMS algorithm for the feasible array structure is proposed. The proposed method tries to prioritize the suppressible noise signal by defining weights based on direction of arrival. Based on the secondary source geometry, the weights can be defined intuitively.

1.5 Related works

1.5.1 Multiple-channel ANC

Spatial ANC aims to attenuate the noise field over a sizable space using multiple microphones and secondary sources. In practice, multiple-channel FXLMS-based methods [11, 12] were proposed. The brief system structure is depicted in Figure 1.4. The adaptive filter's slow convergence speed when the reference signals are highly correlated is a significant disadvantage of this method. To address this issue, frequency-domain algorithms have been proposed [13–18].

Another drawback of this approach 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.

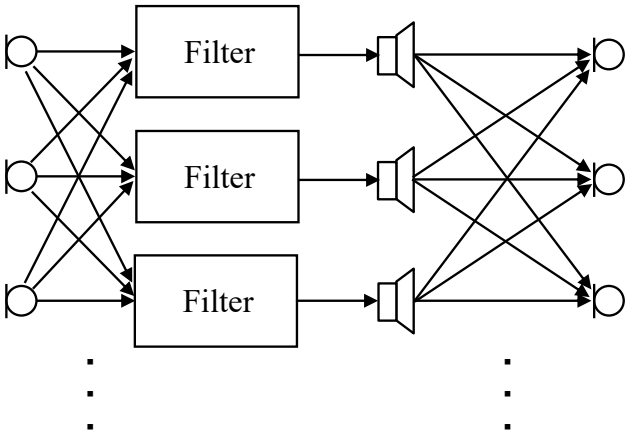


Figure 1.4: Multiple-input-multiple-output system.

1.5.2 Harmonic-domain ANC

In contrast to the multiple-point method, harmonic-domain-based ANC methods, which decompose signals into harmonic basis functions, e.g., spherical harmonic (SH) functions, were proposed [19–24]. The advantages of this approach are that the sound field can be controlled in terms of a region rather than points, and fast convergence can be achieved by uniformly placing the microphones and secondary sources on a suitable surface, such as a sphere. Although this array geometry constraint results in stable and efficient signal representation in the SH-domain, spherical arrays in 3D space with a large number of microphones and loudspeakers are physically difficult to build in practice. The reference microphone array (RMA) takes up a lot of room because it is at the very edge of the system, reducing its viability.

1.5.3 Sound field reproduction

Instead of using a spherical array, multiple circular arrays were used to first transform signals into azimuth harmonic coefficients, and then the SH coefficients are calculated numerically to reproduce a 3D sound field [25]. The difficulty in constructing the array structure has been reduced in this approach, but the ROI must still be surrounded by circular arrays.

In most practical situations [26, 27], the underlying sound field is due to a small number of underlying sound sources. In this context, CS has been successfully applied to sound field reproduction, sound field separation, and sound localization [28–32]. Koyama and co-workers [28, 29] showed that spatial aliasing [33, 34] originating from spatial sampling can be avoided in sound field reproduction. Takida et al. [30] demonstrated that the exterior and interior sound fields, which correspond to the direct source and reverberant component, can be precisely separated. CS has also been used in spatial ANC. Zhang et al. [35] demonstrated that the number of secondary sources can be reduced.

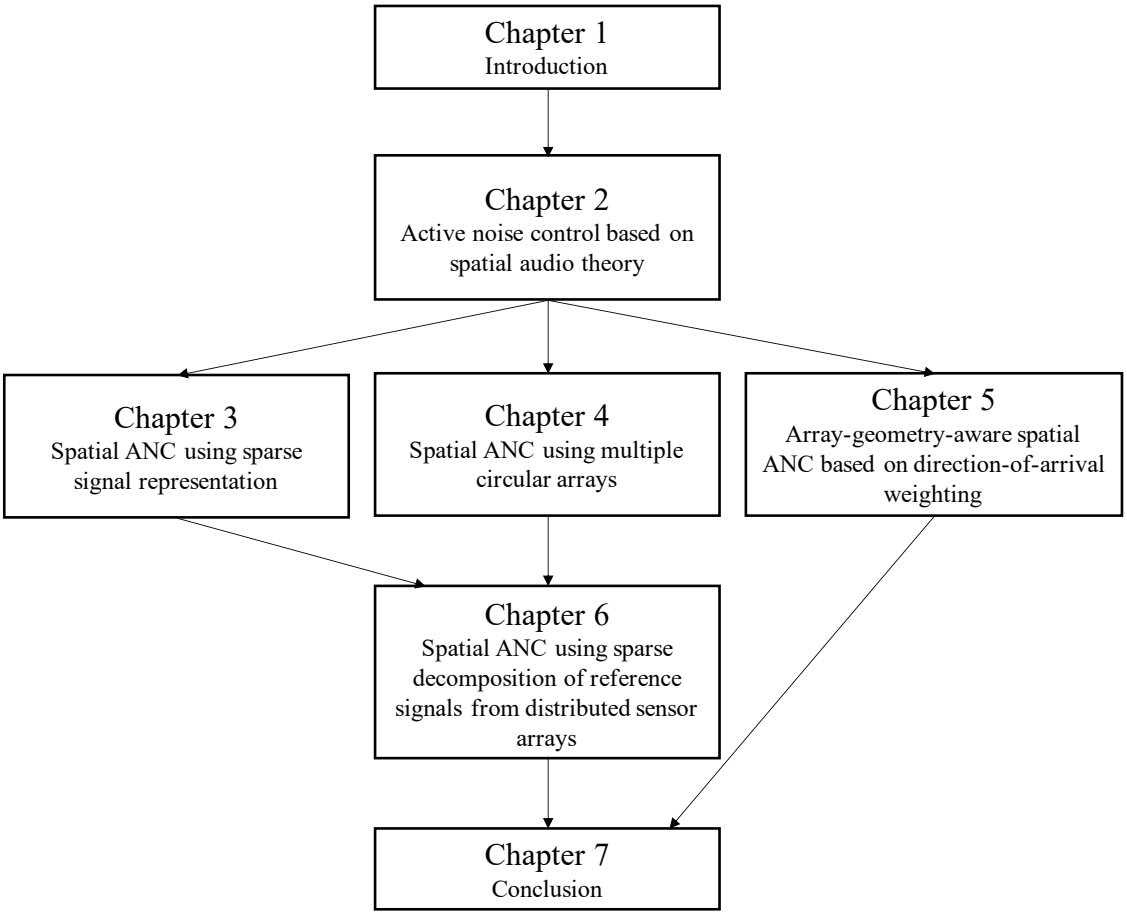


Figure 1.5: Structure of this thesis.

1.6 Structure of this thesis

The rest of the thesis is organized as in the flow chart in Figure 1.5. In Chapter 2, the basic theory of the sound field representation and ANC is reviewed. In Chapter 3, the sparse signal representation in the spatial ANC task is investigated to improve the system performance while reducing the system complexity. The multiple circular array structure is introduced in Chapter 4 to improve system viability. The weighting based on direction of arrival is investigated in Chapter 5. The sparse signal representation and distributed sensor array structure are combined in Chapter 6. Finally, Chapter 7 brings the thesis to a close.

Chapter 2

Active noise control based on spatial audio theory

Sound field analysis and reproduction based on a physical model of the sound propagation has the potential to efficiently control the sound field in a sizable region. The adaptive filtering approach is introduced to control the ever-changing sound field by combining it with the sound field representation theory. Spatial ANC refers to the task of specifically attenuating the unwanted noise field in a large region.

In this chapter, the background theory related to the spatial ANC is reviewed. Besides, the problem of the existing approaches is stated.

2.1 Sound field representation

Let $s(\mathbf{r}, t)$ be an acoustical pressure in the volume of interest. In the source-free region, the sound field satisfies the time-domain acoustic wave equation [36]:

$$\nabla^2 s(\mathbf{r}, t) - \frac{1}{c^2} \frac{\partial^2 s(\mathbf{r}, t)}{\partial t^2} = 0, \tag{2.1}$$

where ∇^2 is the Laplacian, $\mathbf{r}$ is the position vector, t is the time index, and c is the speed of sound. To analyze the steady-state, we can apply the temporal Fourier transform to Eq. (2.1), which leads to the homogeneous Helmholtz equation [36]:

$$\left(\nabla^2 + k^2\right) S(\mathbf{r}, k) = 0, \tag{2.2}$$

where $k = 2\pi f / c$ is the wave number, f is the temporal frequency. Hereafter, wave number k is omitted for notational simplicity.

The general solution of the Helmholtz equation, Eq. (2.2), is known to be plane waves and harmonic functions [36]. One of the solutions to the homogeneous Helmholtz equation is a superposition of plane waves, which is known as Herglotz wave function [37]:

$$S(\mathbf{r}) = \int_{\hat{\mathbf{r}} \in \mathbb{S}^2} \rho^{(\mathrm{pw})}(\hat{\mathbf{r}}) e^{ik\hat{\mathbf{r}} \cdot \mathbf{r}} d\hat{\mathbf{r}}, \tag{2.3}$$

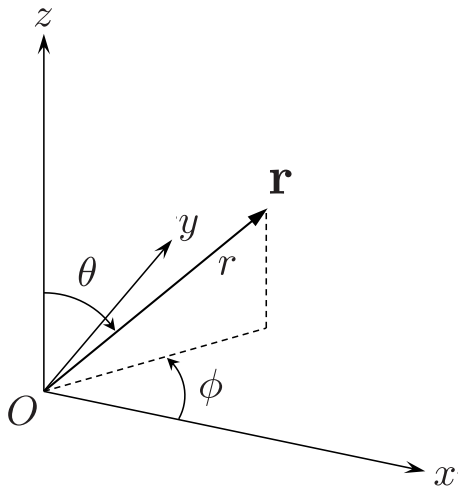


Figure 2.1: The coordinate system used in this study.

where $\mathbb{S}^2$ is the unitary sphere, $\rho^{(\mathrm{pw})}(\hat{\mathbf{r}})$ is the distribution of the plane waves, and $\hat{\mathbf{r}}$ is a unit vector toward the direction of the plane wave. Figure 2.1 depicts the coordinate system used in this study. Angle θ is termed azimuth and ϕ is termed colatitude.

Another solution is obtained in spherical coordinates by utilizing the separation of variables [38], which leads to:

$$S(\mathbf{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \alpha_{\nu,\mu} j_\nu(kr) Y_{\nu,\mu}(\hat{\mathbf{r}}), \tag{2.4}$$

where $\alpha_{\nu,\mu}$ is the harmonic coefficient of the sound field and $j_\nu(kr)$ is the spherical Bessel function.

$$Y_{\nu,\mu}(\hat{\mathbf{r}}) = (-1)^\mu \sqrt{\frac{2\nu+1}{4\pi}} \sqrt{\frac{(\nu-\mu)!}{(\nu+\mu)!}} P_{\nu,\mu}(\cos\theta) e^{i\mu\phi} \tag{2.5}$$

is the SH function of order ν and degree μ, $P_{\nu,\mu}(\cos\theta)$ is the associated Legendre function, and θ and φ are the colatitude and azimuth angle, respectively. Eq. (2.4) represents the interior sound field in the absence of a sound source within the ROI. This condition, known as an interior problem, is depicted in Figure 2.2. Refer to Chapter 6 for the form for describing the exterior field shown in Figure 2.3.

2.2 Active noise control based on adaptive filtering

ANC is a technique, which can actively attenuate the time-varying noise signal by measuring the signal with sensors and driving the antinoise signal from transducers. One of the simplest ANC systems consists of a single reference sensor, single

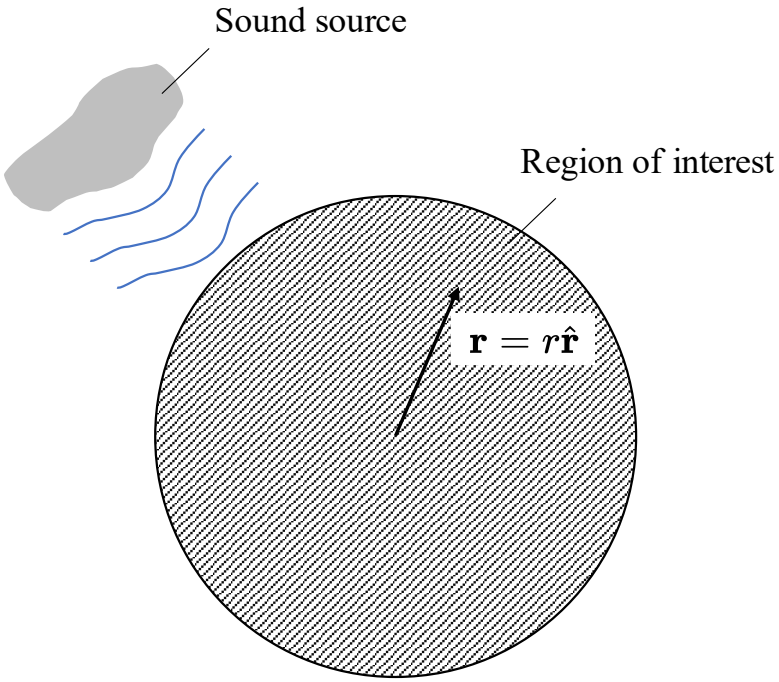


Figure 2.2: Interior problem where all the sound sources exist outside the ROI.

secondary source, and single error sensor. This is referred to as a single-channel feedforward ANC system, in which the reference signal captured by the reference sensor is used to generate the antinoise signal from the secondary source. The residual signal is measured by the error sensor in order to monitor the system's performance. There are variety of methods applied to several acoustic problems based on this principle [4, 5].

2.2.1 Time-domain filtered-X least mean squares algorithm

The adaptive filtering approach has been widely used in ANC applications to control the time-varying noise field. By updating the filter coefficients using the LMS algorithm, the adaptive filter minimizes the residual signal monitored by the error microphone. It is well understood that the presence of a secondary path generally leads to instability [39]. Because of the presence of the secondary path, the error signal and the reference signal are misaligned. To solve this problem, the FXLMS algorithm has been proposed [40–42]. The block diagram of the FXLMS algorithm is depicted in Figure 2.4. The estimated secondary path is placed in the reference signal path to compensate the secondary path.

As shown in Figure 2.4, the residual error signal at the error microphone position can be written as a sum of the primary noise signal and the antinoise signal:

$$e(n) = d(n) - s(n) * \left[\boldsymbol{w}^{\mathsf{T}}(n)\boldsymbol{x}(n)\right], \tag{2.6}$$

where n is the time index, $s(n)$ is the impulse response of the secondary path, $*$ is the linear convolution operator, and $(\cdot)^{\mathsf{T}}$ is the transpose operator. $\boldsymbol{w}(n)$ and $\boldsymbol{x}(n)$

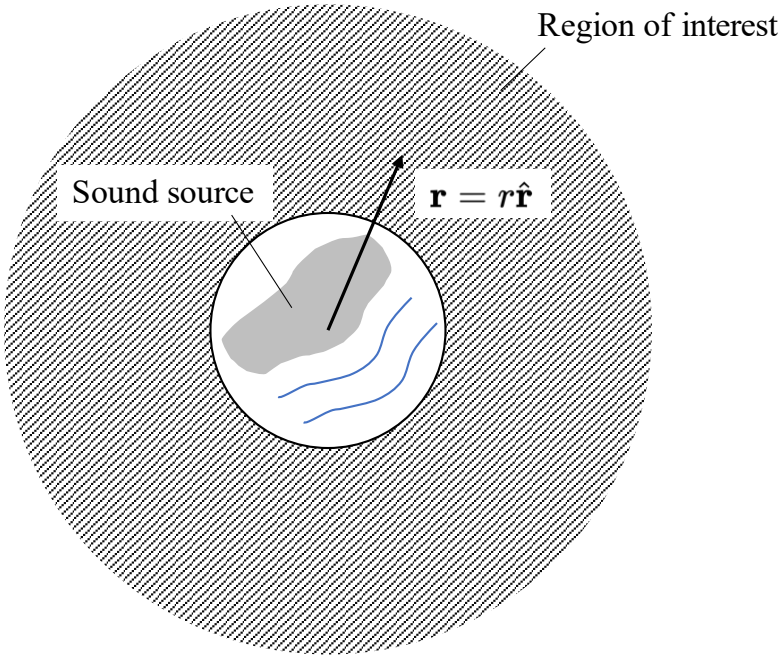


Figure 2.3: Exterior problem where all the sound sources exist inside the ROI.

are the coefficient vectors of the adaptive filter and reference signal, respectively. The cost function is expressed as the expectation of the squared residual signals:

$$\mathcal{J}(n) = E\left[e^2(n)\right]. \tag{2.7}$$

The adaptive filter minimizes the instantaneous error

$$\hat{\mathcal{J}}(n) = e^2(n) \tag{2.8}$$

in each step using the steepest descent algorithm. The adaptive update of the filter coefficients is performed by stepping toward the negative gradient direction of the cost function with step size u

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \frac{u}{2}\nabla\hat{\mathcal{J}}(n) \tag{2.9}$$

$$= \boldsymbol{w}(n) + u\boldsymbol{x}'(n)e(n), \tag{2.10}$$

where

$$\boldsymbol{x}'(n) = \hat{s}(n) * \boldsymbol{x}(n). \tag{2.11}$$

$\hat{s}(n)$ is the impulse response of the estimated secondary path, which can be measured in advance or modeled using online modeling methods [43, 44] in practical applications.

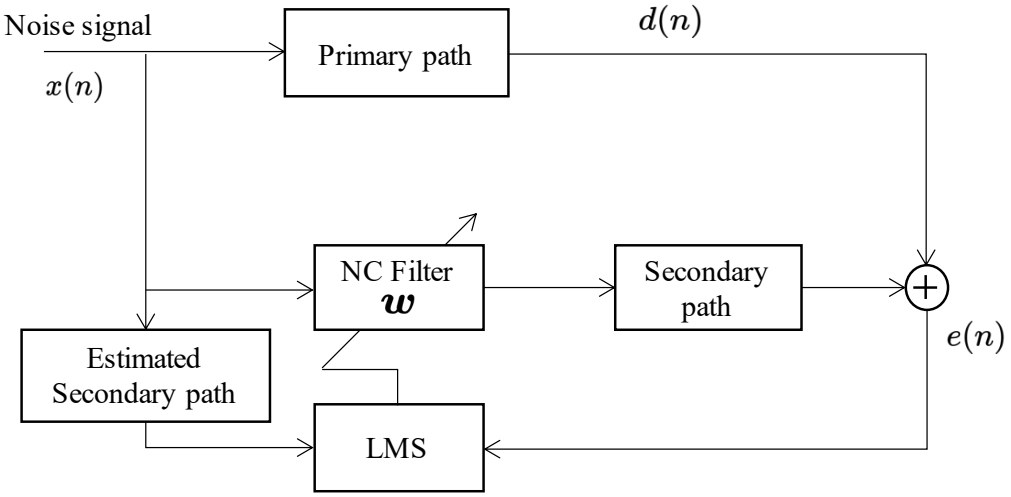


Figure 2.4: Block diagram of the time-domain FXLMS algorithm.

2.2.2 Temporal-frequency-domain FXLMS

The time-domain FXLMS system [45,46] has been shown effectiveness in many ANC applications. However, its high computational complexity and slow convergence of adaptive processing limit the practical implementation.

To overcome these drawbacks, temporal-frequency-domain FXLMS algorithms were developed. The temporal-frequency-domain FXLMS algorithm's block diagram is shown in Figure 2.5. The M-point signal block $\boldsymbol{x}(\bar{n})$ is formed by accumulating the reference signal $x(n)$ in the buffer, where $\bar{n}$ is the block index. The signal vector $\boldsymbol{x}(\bar{n})$ is transformed into the temporal frequency domain by using the fast fourier transform (FFT):

$$\mathbf{X}(\bar{n}) = \mathrm{FFT}[\boldsymbol{x}(\bar{n})] = [X_0(\bar{n}), X_1(\bar{n}), \ldots, X_{M-1}(\bar{n})]^{\mathsf{T}}, \tag{2.12}$$

where M is the length of the block.

The residual error signal at the error microphone can be obtained by using the inverse fast fourier transform (IFFT):

$$e(n) = d(n) - s(n) * \mathrm{IFFT}\left[\mathbf{W}(\bar{n}) \circ \mathbf{X}(\bar{n})\right]. \tag{2.13}$$

where $\mathbf{W}(\bar{n})$ is the temporal-frequency-domain adaptive filter and $\circ$ is the element-wise product operator. The residual error signal is also accumulated in the buffer and transformed into the temporal frequency domain every M sample:

$$\mathbf{E}(\bar{n}) = \mathrm{FFT}[\boldsymbol{e}(\bar{n})] = [E_0(\bar{n}), E_1(\bar{n}), \ldots, E_{M-1}(\bar{n})]^{\mathsf{T}}, \tag{2.14}$$

The adaptive filter update can be performed in each frequency bin based on the complex LMS algorithm:

$$W_m(\bar{n}+1) = W_m(\bar{n}) + \mu_m X'^{*}_m(\bar{n}) E_m(\bar{n}), \tag{2.15}$$

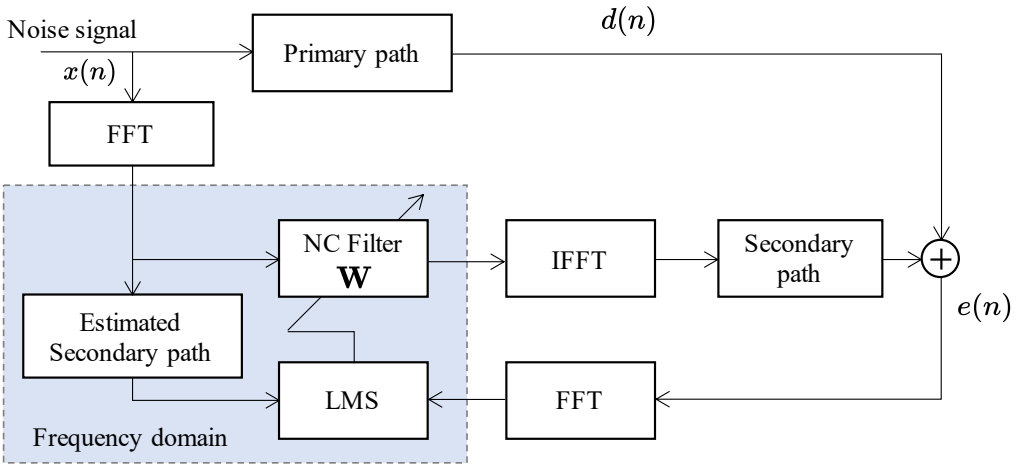


Figure 2.5: Block diagram of the temporal-frequency-domain FXLMS algorithm.

where m is the frequency bin index and $(\cdot)^*$ is the complex conjugate operator. The filtered reference signal $X'_m(\bar{n})$ can be defined as

$$X'_m(\bar{n}) = \hat{S}_m(\bar{n})X_m(\bar{n}), \tag{2.16}$$

where $\hat{S}_m(\bar{n})$ is the estimated secondary path in the temporal frequency domain.

2.2.3 Spherical-harmonic-domain FXLMS

The adaptive algorithm in the temporal frequency domain has been successfully applied to several practical applications including a multiple-channel system to enlarge the quiet region [47–50]. However, increasing the number of channels to control a large region remains a challenge. To address this issue, a novel approach based on spatial frequency analysis has been developed. This method employs the spatial audio theory described in Section 2.1 to analyze and control the large sound field. The signals are transformed into harmonic coefficients using orthogonal basis function such as SH function. The transformed domain is also known as the SH-domain, mode domain, or wave domain. The adaptive algorithm in the SH-domain is described in this section.

Let the ROI be a spherical region surrounded by an error microphone array (EMA), a secondary loudspeaker array (SLA), and a RMA. The antinoise field in the temporal frequency domain generated by the SLA can be described as

$$S_c(\mathbf{r}) = \sum_{\ell=1}^{L_\mathrm{s}} d_\ell G(\mathbf{r}|\mathbf{r}_\ell), \tag{2.17}$$

where L_s is the number of loudspeakers and d_ℓ is the driving signal of the ℓ-th loudspeaker. The free field Green's function, $G(\mathbf{r}|\mathbf{r}_\ell)$, can be parameterized as

$$G(\mathbf{r}|\mathbf{r}_\ell) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} C_{\nu,\mu,\ell} j_\nu(kr) Y_{\nu,\mu}(\hat{\mathbf{r}}), \tag{2.18}$$

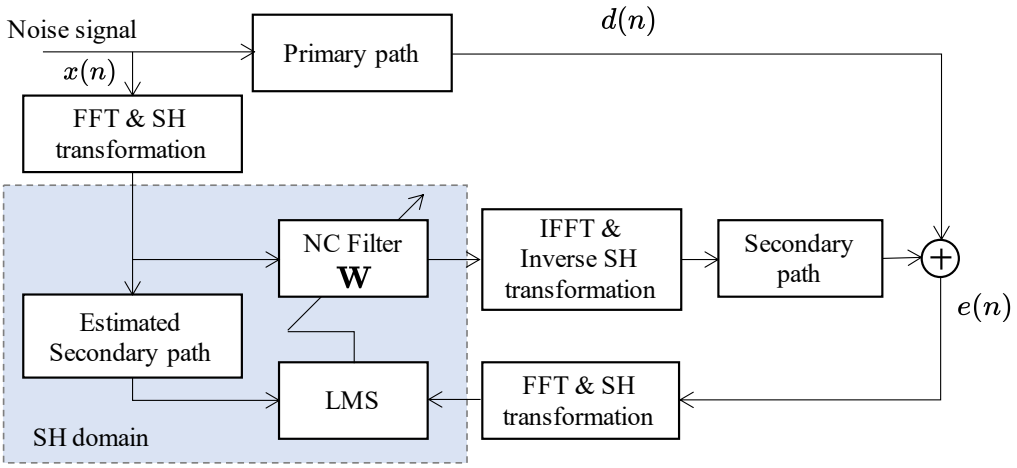


Figure 2.6: Block diagram of the SH-domain FXLMS algorithm.

where $C_{\nu,\mu,\ell}$ is the SH coefficient of the acoustic transfer function from the ℓ-th loudspeaker to the ROI. By substituting (2.18) to (2.17), the antinoise field can be represented as

$$S_c(\mathbf{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \underbrace{\sum_{\ell=1}^{L_\mathrm{s}} d_\ell C_{\nu,\mu,\ell}}_{z_{\nu,\mu}} j_\nu(kr) Y_{\nu,\mu}(\hat{\mathbf{r}}). \tag{2.19}$$

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

$$C_{\nu,\mu,\ell} = g_{\nu,\mu} Y^*_{\nu,\mu}(\hat{\mathbf{r}}_\ell), \tag{2.20}$$

where $g_{\nu,\mu}$ is the SH coefficient of the acoustic transfer function. From Eq. (2.19) and Eq. (2.20), $z_{\nu,\mu}$ can be written as

$$z_{\nu,\mu} = g_{\nu,\mu} \underbrace{\sum_{\ell=1}^{L_\mathrm{s}} d_\ell Y^*_{\nu,\mu}(\hat{\mathbf{r}}_\ell)}_{\xi_{\nu,\mu}}, \tag{2.21}$$

where $\xi_{\nu,\mu}$ is the SH coefficient of the driving signals.

Now, the noise field attenuation is considered in SH domain. The residual noise field captured by the EMA can be described as a sum of the primary noise field given in Eq. (2.4) and the antinoise field given in Eq. (2.19):

$$\begin{aligned} e_{\nu,\mu} &= \alpha_{\nu,\mu} + z_{\nu,\mu} \\ &= \alpha_{\nu,\mu} + g_{\nu,\mu}\xi_{\nu,\mu}. \end{aligned} \tag{2.22}$$

As depicted in Figure 2.6, the FXLMS algorithm generates the antinoise signal by applying a linear filter to the reference signal; thus, the total residual noise field can be written as

$$e_{\nu,\mu} = \alpha_{\nu,\mu} + g_{\nu,\mu} x_{\nu,\mu} w_{\nu,\mu}, \tag{2.23}$$

where $x_{\nu,\mu}$ and $w_{\nu,\mu}$ are the SH coefficients of the reference signal and the adaptive filter, respectively. By concatenating all available coefficients for each order ν and degree μ, we can write Eq. (2.23) in a matrix form:

$$\mathbf{e} = \boldsymbol{\alpha} + \mathbf{gxw}, \tag{2.24}$$

where

$$\mathbf{g} = \mathrm{diag}(g_{0,0}, \ldots, g_{N,N}) \in \mathbb{C}^{(N+1)^2 \times (N+1)^2},$$

$$\mathbf{x} = \mathrm{diag}(x_{0,0}, \ldots, x_{N,N}) \in \mathbb{C}^{(N+1)^2 \times (N+1)^2},$$

and

$$\mathbf{w} = [w_{0,0}, \ldots, w_{N,N}]^{\mathsf{T}} \in \mathbb{C}^{(N+1)^2 \times 1}.$$

The notation $\mathrm{diag}(\cdot)$ indicates the diagonal matrix whose diagonal elements consist of the vector of argument. The truncation order N is usually determined by adopting the criterion $N = \lceil ekR/2 \rceil$ [51] or $N = \lceil kR \rceil$ [52], where R is the radius of the ROI.

A cost function is set as the sum of the squared residual signals [23]:

$$\mathcal{J}(\iota) = \sum_{\nu=0}^{N} \sum_{\mu=-\nu}^{\nu} |e_{\nu,\mu}(\iota)|^2 = \mathbf{e}(\iota)^{\mathsf{H}} \mathbf{e}(\iota), \tag{2.25}$$

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

$$\frac{\partial \mathcal{J}(\iota)}{\partial \mathbf{w}^*(\iota)} = \mathbf{x}'(\iota)^{\mathsf{H}} \mathbf{e}(\iota), \tag{2.26}$$

so that we can update the filter coefficients as

$$\mathbf{w}(\iota + 1) = \mathbf{w}(\iota) - u\mathbf{x}'(\iota)^{\mathsf{H}} \mathbf{e}(\iota). \tag{2.27}$$

The filtered reference coefficients can be defined as

$$\mathbf{x}'(\iota) = \hat{\mathbf{g}}\mathbf{x}(\iota), \tag{2.28}$$

where $\hat{\mathbf{g}}$ is the coefficients of the estimated secondary path.

2.3 Problems of existing approaches

The harmonic-domain approaches in spatial ANC leveraging the spatial frequency analysis [23] have improved the size of the controllable region. This is because the signal representation, for example, in the SH-domain, allows for the description of a large continuous region of the sound field rather than just a few points. It is

well known that signal decomposition into the SH-domain leads to faster adaptive processing convergence and lower computational complexity due to the signal independence characteristics in the SH-domain [21].

These approaches have been shown certain breakthroughs in the spatial ANC tasks, however, there is another shortcoming that needs to be addressed to realize practical systems. In reality, the sound field must be spatially sampled before it can be processed from the captured signal by the microphones or reproduced by driving discretized loudspeakers. For accurate signal processing, this spatial sampling theoretically necessitates a spatially uniform sampling. Spatially uniform sampling is a difficult problem in three dimensions [53]. Furthermore, installing the microphone and loudspeaker array physically is usually difficult because it has a huge impact on the system's feasibility.

The spherical microphone array and loudspeaker array have been used for the study in the sound field capturing [54], sound field synthesizing [55], and binaural processing [56, 57]. [55, 57] show the hardware implementations. One of the difficulties in these studies is reducing the complexity of the array structure for practical application.

2.4 Summary

The basic theory related to the spatial ANC was reviewed. The sound field representation theory gives an analytical solution to analyze and process the sound field in a large space. Adaptive filtering approaches, such as FXLMS algorithms, can be applied to signals that are constantly changing. Although the combination of these two approaches has demonstrated progress in spatial ANC tasks, the challenge of the array structure's complexity remains for practical use. Starting from the next chapter, several solutions to improve the system feasibility of spatial ANC are discussed.

Chapter 3

Spatial active noise control using sparse signal representation

Spatial ANC systems aim to attenuate undesired noise over a spatial region by generating an antinoise sound field over the region using secondary sources. Such a system is viewed as an extension from the well-studied single-channel ANC [3–5] to a multiple-channel system [3, 6]. Controlling a sound field over a large area, on the other hand, necessitates a large number of reference and error microphones to capture the reference noise signals and residual error signals, respectively. In addition, a large number of loudspeakers are required to generate the antinoise. The requirement for a large number of microphones limits the practicality of using spatial ANC systems. Nevertheless, in most practical applications [26, 27], the underlying noise field is caused by a small number of underlying noise sources. The goal of this chapter is to use a sparse signal representation to reduce the number of microphones required by spatial ANC systems.

In practice, multiple-channel ANC in the frequency domain [11, 12] is widely used to attenuate the noise field at multiple points where the error microphones are placed. The requirement to uniformly place many error microphones inside the control region is one disadvantage of this approach to creating a large quiet zone. In contrast, spatial sound field representation techniques, such as WFS [8] and HOA [58, 59] are promising techniques to control not only multiple points but the entire space of interest. These techniques have been applied to ANC, and it has been demonstrated that harmonic-domain ANC can achieve noise attenuation over a large space while requiring less computational complexity [19, 21, 22]. However, such systems still require a large number of microphones and loudspeakers to reproduce the sound field properly, thus limiting the feasibility of practical implementations. The system suffers from artifacts due to spatial aliasing [60] due to violating the spatial Nyquist sampling criterion when there are not a theoretically sufficient number of sensors.

One way to alleviate the aliasing restriction is by expressing the noise field using only a small number of basis functions from an over-complete dictionary. This approach is well known as CS [61], which can provide an accurate solution for

underdetermined problems where the sparse characteristic of the underlying signal field is usually used to solve the problem [62].

In this chapter, a harmonic-domain ANC system is proposed that can perform noise attenuation in large space beyond the spatial Nyquist frequency using fewer reference microphones than required by the Nyquist theory. A superposition of weighted plane waves impinging from the far-field is used to describe the reference noise field. The reference harmonic coefficients are reconstructed using estimated plane wave weights calculated using various types of CS methods, which are expected to include fewer artifacts caused by spatial aliasing. The proposed methods are evaluated and compared in terms of reference-noise-field reproduction accuracy and noise attenuation level. The results show that the proposed method is capable of overcoming the spatial Nyquist frequency limitation. Furthermore, while properly controlling the noise field, the number of reference microphones can be reduced.

Nicolas [31] proposed CS sound field reproduction based on plane wave decomposition by introducing an HOA-based constraint. The loudspeaker weights, however, were calculated using amplitude-panning, which cannot be used directly for harmonic-domain processing. Jihui [35, 63] proposed the sparse complex FXLMS algorithm, despite using the CS approach to calculate sparse loudspeaker weights, so that spatial aliasing artifacts caused by the spatial sampling of the microphone array could not be avoided.

3.1 Problem statement

Let us assume an arbitrary noise field $S(\mathbf{r})$, where all the noise sources are outside of the ROI in 2D space. The geometrical setup with the microphone arrays and loudspeaker array is illustrated in Figure 3.1 The noise field can be represented as a decomposition of harmonic coefficients:

$$S(\mathbf{r}) \quad = \quad \sum_{\mu=-\infty}^{\infty} \beta_\mu J_\mu(kr) e^{i\mu\phi}, \tag{3.1}$$

where β_μ is the μth-order circular harmonic (CH) coefficient and $J_\mu(\cdot)$ is the μth-order Bessel function. The coefficient β_μ can be calculated by exploiting the orthogonality of exponential functions:

$$\beta_\mu = \frac{1}{J_\mu(kr)} \int S(\mathbf{r}) e^{-i\mu\phi} d\phi. \tag{3.2}$$

Our objective is to derive an algorithm to acquire precise reference CH coefficients of a noise field, which are then used to perform ANC instead of using direct coefficient extraction Eq. (3.2). Because the noise field is measured at multiple discrete points in practice, the calculation Eq. (3.2) contains spatial aliasing artifacts. Even with an insufficient number of microphones, the noise field is expected to be attenuated at higher frequencies beyond the spatial Nyquist frequency.

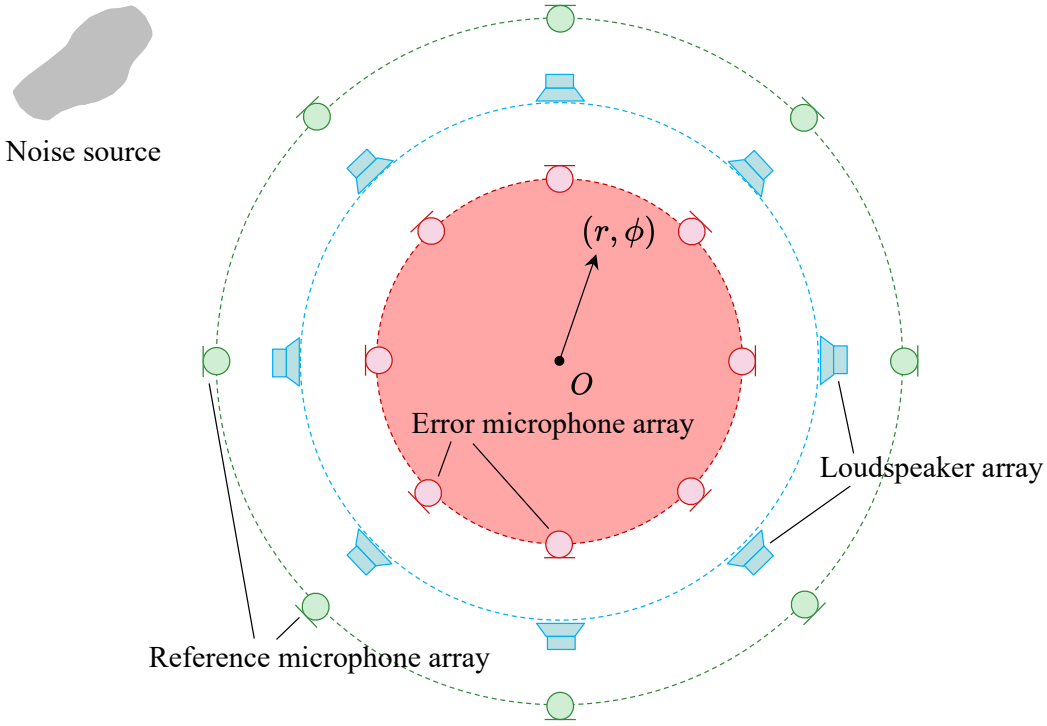


Figure 3.1: Geometric setup of feedforward ANC system. The colored circle indicates the ROI.

3.2 Sparse weight estimation of plane wave

In this section, we derive the CH coefficient reconstruction based on plane wave decomposition assuming that the noise field is constructed by a superposition of weighted plane waves. A sound field in a specific region is known to be approximated by a superposition of a few plane waves [64, 65]. The weights of the plane waves can be accurately estimated using the CS approach, which can be used to derive a sparse solution from a few measurements [66].

As an alternative to Eq. (3.1), the solution to the homogeneous Helmholtz equation can be derived by utilizing a plane wave function [65], which is also referred to as the Herglotz wave function:

$$S(\mathbf{r}) = \int_{\hat{\mathbf{r}} \in \mathbb{S}} \rho^{(\mathrm{pw})}(\hat{\mathbf{r}}) e^{ik\hat{\mathbf{r}}\cdot\mathbf{r}} d\hat{\mathbf{r}}, \tag{3.3}$$

where $\rho^{(\mathrm{pw})}(\hat{\mathbf{r}})$ is the distribution of plane wave and $\hat{\mathbf{r}} \equiv (1, \phi_\ell)$ is a unit vector toward the direction of the plane wave. Figure 3.2 illustrates this sound field expression also termed plane wave composition. The circular expansion of the incident noise field due to a unit-magnitude plane wave is given by [65]

$$e^{ik\hat{\mathbf{r}}_\ell\cdot\mathbf{r}} = \sum_{\mu=-\infty}^{\infty} i^\mu e^{-i\mu\phi_\ell} J_\mu(kr) e^{i\mu\phi}. \tag{3.4}$$

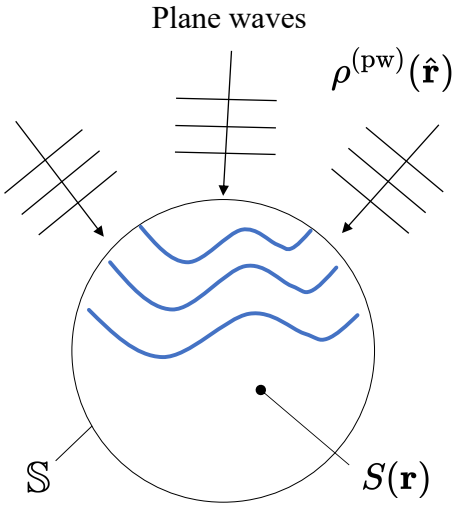


Figure 3.2: Plane wave composition in 2D field.

From Eq. (3.3) and Eq. (3.4), we have

$$S(\mathbf{r}) = \sum_{\mu=-\infty}^{\infty} i^{\mu} J_{\mu}(kr) e^{i\mu\phi} \int_{\hat{\mathbf{r}}\in\mathbb{S}} \rho^{(\mathrm{pw})}(\hat{\mathbf{r}}) e^{-i\mu\phi_{\ell}} d\hat{\mathbf{r}} \tag{3.5}$$

$$= \sum_{\mu=-\infty}^{\infty} \underbrace{i^{\mu} \int_{\hat{\mathbf{r}}\in\mathbb{S}} \rho^{(\mathrm{pw})}(\hat{\mathbf{r}}) e^{-i\mu\phi_{\ell}} d\hat{\mathbf{r}}}_{\beta_{\mu}} J_{\mu}(kr) e^{i\mu\phi}. \tag{3.6}$$

Equation Eq. (3.6) shows that the CH coefficient β_{μ} can be reconstructed using the plane wave distribution $\rho^{(\mathrm{pw})}(\hat{\mathbf{r}})$ to reproduce the incident noise field.

The plane wave composition Eq. (3.3) can be approximated by a limited number of plane waves [64]:

$$S(\mathbf{r}) \approx \sum_{\ell=1}^{L} \gamma_{\ell} e^{ik\hat{\mathbf{r}}\cdot\mathbf{r}}. \tag{3.7}$$

We represent Eq. (3.7) in a matrix form:

$$\boldsymbol{s} = \boldsymbol{E}\boldsymbol{\gamma}, \tag{3.8}$$

where $\boldsymbol{s} \in \mathbb{C}^{Q}$ is a vector of the observed signals, $\boldsymbol{\gamma} \in \mathbb{C}^{L}$ is an vector of plane wave weights, and $\boldsymbol{E} \in \mathbb{C}^{Q\times L}$ is a matrix given by

$$\boldsymbol{E} = \begin{bmatrix} e^{ik\hat{\mathbf{r}}_1\cdot\mathbf{r}_1} & \cdots & e^{ik\hat{\mathbf{r}}_L\cdot\mathbf{r}_1} \\ \vdots & \ddots & \vdots \\ e^{ik\hat{\mathbf{r}}_1\cdot\mathbf{r}_Q} & \cdots & e^{ik\hat{\mathbf{r}}_L\cdot\mathbf{r}_Q} \end{bmatrix}. \tag{3.9}$$

We assume that the composition of plane waves is sparse with the underdetermined condition ($Q < L$) in Eq. (3.8).

3.2.1 ℓ_1-norm constrained minimization

One of the most widely used techniques to estimate a sparse solution is a convex relaxation:

$$\underset{\boldsymbol{\gamma}}{\text{minimize}} \, \|\boldsymbol{\gamma}\|_p^p, \text{ s.t. } \boldsymbol{s} = \boldsymbol{E}\boldsymbol{\gamma}. \tag{3.10}$$

Applying basis pursuit denoising (BPD) [67], Eq. (3.10) can be written in an unconstrained form as:

$$\mathcal{J} = \underset{\boldsymbol{\gamma}}{\text{minimize}} \, \frac{1}{2}\|\boldsymbol{s} - \boldsymbol{E}\boldsymbol{\gamma}\|_2^2 + \lambda_1\|\boldsymbol{\gamma}\|_1. \tag{3.11}$$

where λ_1 controls the strength of the sparsity constraint. By utilizing the gradient descent algorithm, we can update the weights $\boldsymbol{\gamma}$ as

$$\boldsymbol{\gamma}(\iota+1) = \boldsymbol{\gamma}(\iota) - \frac{\eta_{\text{pw}}}{2}\nabla\mathcal{J}(\iota), \tag{3.12}$$

where ι is the iteration index and η_{pw} is the step size. A gradient of the cost function $\mathcal{J}(\iota)$ can be calculated as [35, 63]

$$\nabla\mathcal{J}(\iota) = 2\boldsymbol{E}^{\mathsf{H}}\boldsymbol{v}(\iota) + \lambda_1\Big(\text{sgn}\big\{\Re\left[\boldsymbol{\gamma}(\iota)\right]\big\} + i\text{sgn}\big\{\Im\left[(\boldsymbol{\gamma}(\iota)\right]\big\}\Big), \tag{3.13}$$

where $(\cdot)^{\mathsf{H}}$ denotes the Hermitian-transpose operation, $\text{sgn}(\cdot)$ denotes the sign function, and $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts of the argument. Finally, substituting Eq. (3.13) into Eq. (3.12), the plane-wave weights can be calculated as:

$$\boldsymbol{\gamma}(\iota+1) = \boldsymbol{\gamma}(\iota) - \eta_{\text{pw}}\boldsymbol{E}^{\mathsf{H}}\boldsymbol{v}(\iota) - \frac{1}{2}\eta_{\text{pw}}\lambda_1\Big(\text{sgn}\big\{\Re\left[\boldsymbol{\gamma}(\iota)\right]\big\} + i\text{sgn}\big\{\Im\left[(\boldsymbol{\gamma}(\iota)\right]\big\}\Big). \tag{3.14}$$

3.2.2 Iteratively reweighted least squares

It is known that by introducing the ℓ_p-norm and solving the following optimization problem, where $0 < p < 1$, one can provide a sparse solution with much fewer measurements than $p = 1$ [68, 69]. Although the optimization problem Eq. (3.10) is nonconvex, we can adopt the iteratively reweighted least squares (IRLS) to solve it iteratively by replacing the ℓ_p term with a weighted ℓ_2 norm [67, 69]. The IRLS is known to converge relatively quickly in terms of iterations; however, the algorithm includes matrix inversion in every iteration, which may result in a high computational cost. To address this issue, there is a quick implementation of the IRLS algorithm [70].

3.3 Circular-harmonic-domain active noise control

3.3.1 Order truncation

Let us consider a series truncation in sound field decomposition Eq. (3.1):

$$S(\mathbf{r}) \approx \sum_{\mu=-N}^{N} \beta_\mu J_\mu(kr)e^{i\mu\phi}. \tag{3.15}$$

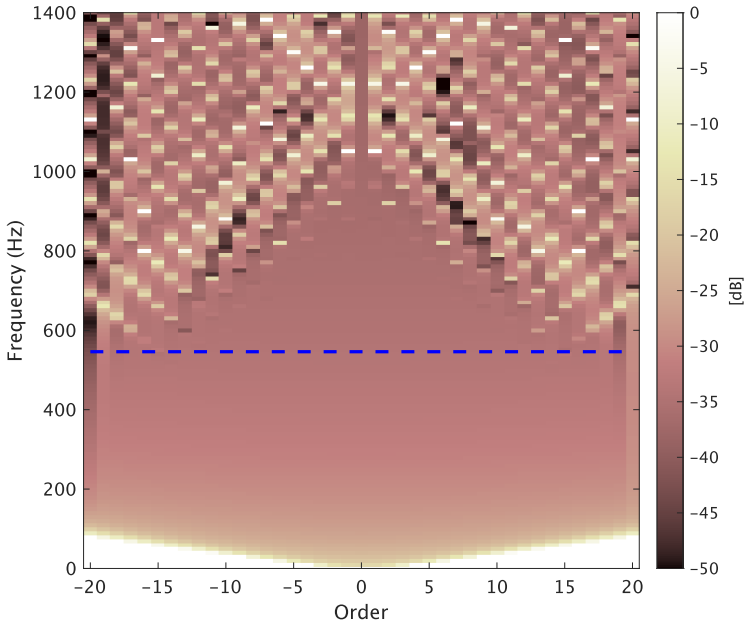


Figure 3.3: Extracted CH coefficients of the cylindrical wave using microphone outputs. The horizontal broken line indicates the spatial Nyquist frequency.

To determine the truncation order N, $N = \lceil ekR/2 \rceil$ [51] or $N = \lceil kR \rceil$ [52], where R is the radius of the ROI, are known to be the adequate criteria.

Now, we discretize the sound field by placing finite microphones on the surface of the ROI:

$$\beta_\mu \approx \frac{1}{QJ_\mu(kr)} \sum_{q=1}^{Q} S(\mathbf{r}_q) e^{-i\mu\phi_q}, \tag{3.16}$$

where Q is the number of microphones and ϕ_q is the azimuth angle of the q-th microphone. We need at least $Q \geq 2N+1$ microphones to avoid spatial aliasing [71]. The CH coefficients of a noise field captured by a 41-element circular microphone array with a radius of 2 m are shown in Figure 3.3. $N = 20$ is the maximum order of the CH coefficients. Look for spatial aliasing artifacts in higher-order signals above 546 Hz, which is the spatial Nyquist frequency for this condition. The high amplitude of low frequency corresponds to an evanescent wave [36] of the cylindrical source. This evanescent component only exists near the source, and the amplitude decays exponentially [60]. As a result, we assume that the impact on an ANC system's performance within the control region is minimal.

3.3.2 Adaptive filter update

We describe a CH-domain adaptive filtering algorithm based on the FXLMS algorithm [41]. Consider two circular microphone arrays placed in a free field, as depicted in Figure 3.1. The outer circle is a RMA captures reference signals, and the inner

circle is an EMA, measuring the residual signals. A circular SLA is positioned between the two microphone arrays to drive the antinoise signals.

Adopting the circular-harmonic-domain feedforward (CHFF) FXLMS algorithm, a residual SH coefficient at step ι can be written as:

$$e_\mu(\iota) = \beta_\mu + w_\mu(\iota)x'_\mu(\iota), \tag{3.17}$$

where $w_\mu(\iota)$ is the μth-order SH-domain weight of the adaptive filter, $x'_\mu(\iota) = g_\mu x_\mu(\iota)$ is the μth-order filtered reference SH coefficient. Here, g_μ is the μth-order SH-domain coefficient of the acoustic transfer function of the loudspeaker, and $x_\mu(\iota)$ is the μth-order reference SH coefficient. In the 2D free field, the coefficient of the acoustic transfer function represented by a circular wave can be described as:

$$g_\mu = -\frac{i}{4}H_\mu^{(2)}(kR_s), \tag{3.18}$$

where $H_\mu^{(2)}(\cdot)$ is the Hankel function of second kind of order μ, and R_s is the radius of the loudspeaker array. The SH-domain reference coefficient can be calculated as:

$$x_\mu(\iota) = i^\mu \sum_{\ell=1}^{L} \gamma_\ell(\iota) e^{-i\mu\phi_\ell}, \tag{3.19}$$

using $\gamma_\ell(\iota)$ reconstructed from the reference signals by solving the optimization problem mentioned in the previous section. Finally, the filtered-X normalized least mean squares (FXNLMS) [72] algorithm updates the weight as:

$$w_\mu(\iota+1) = w_\mu(\iota) - \eta_w \frac{e_\mu(\iota)x'^*_\mu(\iota)}{x'^*_\mu(\iota)x'_\mu(\iota)}, \tag{3.20}$$

where $(\cdot)^*$ denotes the complex conjugate operation and η_w is the step size.

3.4 Validation

3.4.1 Experimental conditions

Methods

We conducted simulation studies to evaluate and compare the accuracy of the reconstructed reference SH coefficients and noise attenuation performance among the following methods:

(i) "CHFF" corresponding to the conventional method using SH decomposition based on Eq. (3.2) and SH-domain filter update based on Eq. (3.20) as a baseline;

(ii) "Proposed (L1)" corresponding to the method using ℓ_1-norm constrained plane-wave-weight estimation based on Eq. (3.14) for the reconstruction of the reference coefficients;

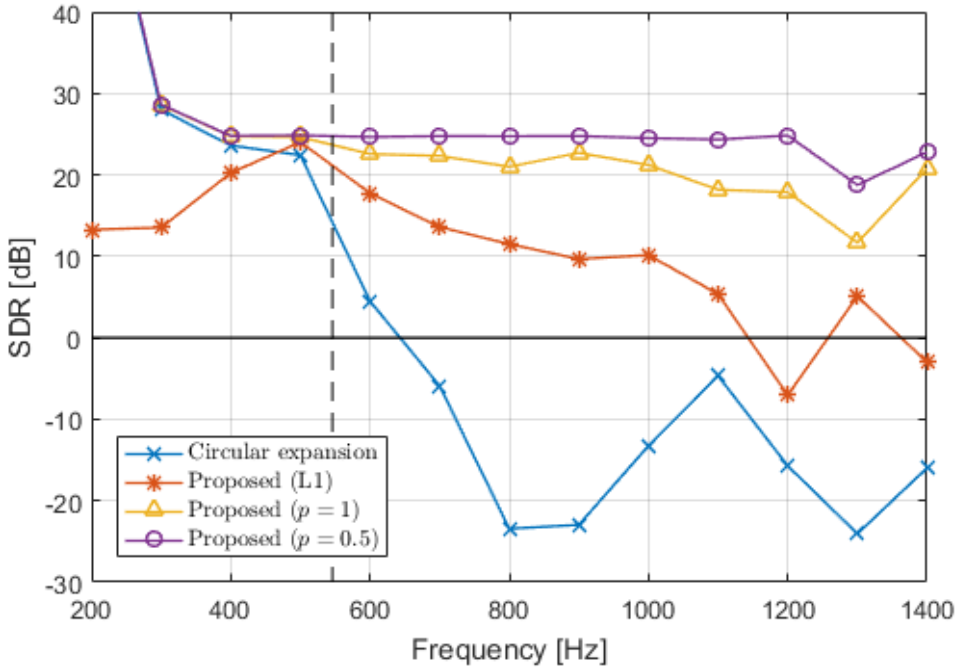


Figure 3.4: SDR of reference sound field at several frequency bins. The vertical broken line indicates the spatial Nyquist frequency corresponding to the reference microphone array.

(iii) "Proposed ($p = 1$)" and (iv) "Proposed ($p = 0.5$)" corresponding to the methods using IRLS methods for the plane-wave-weight estimation based on Eq. (3.11).

Geometrical setup

The ANC system consists of RMA, SLA, and EMA surrounding the ROI in a 2D free field as shown in Figure 3.1. The radii of the EMA and RMA are 1 m and 2 m, respectively. Each array of microphones is made up of 41 omnidirectional microphones. The SLA is made up of 41 monopole loudspeakers with a radius of 1.5 m. The harmonic coefficients have a maximum order of $N = 20$. For these simulation conditions, the spatial Nyquist frequency corresponding to the RMA is 546 Hz, and to the EMA is 1092 Hz. We use $L = 128$ plane waves to construct the over-complete matrix $\boldsymbol{E}$.

Due to spatial aliasing, there will be more artifacts in the calculation of the reference signal than the error signal because the RMA is placed at the outermost. The methods described above were applied to the reference signal to reduce artifacts and reconstruct accurate reference signals.

Noise signal

A single point source of white noise was placed outside of the RMA as the primary noise source.

Evaluation metric

The signal-to-distortion ratio (SDR) [28] was calculated inside the ROI described as the colored region in Figure 3.1 to evaluate the reconstruction accuracy of the reference signals, which affects the noise attenuation performance of feedforward ANC.

Next, an average noise level inside the ROI was calculated to evaluate the noise attenuation performance. Due to the spatial aliasing artifacts in the reference coefficients, it is expected that performance differences will be visible above the spatial Nyquist frequency.

3.4.2 Accuracy of reconstructed reference signal

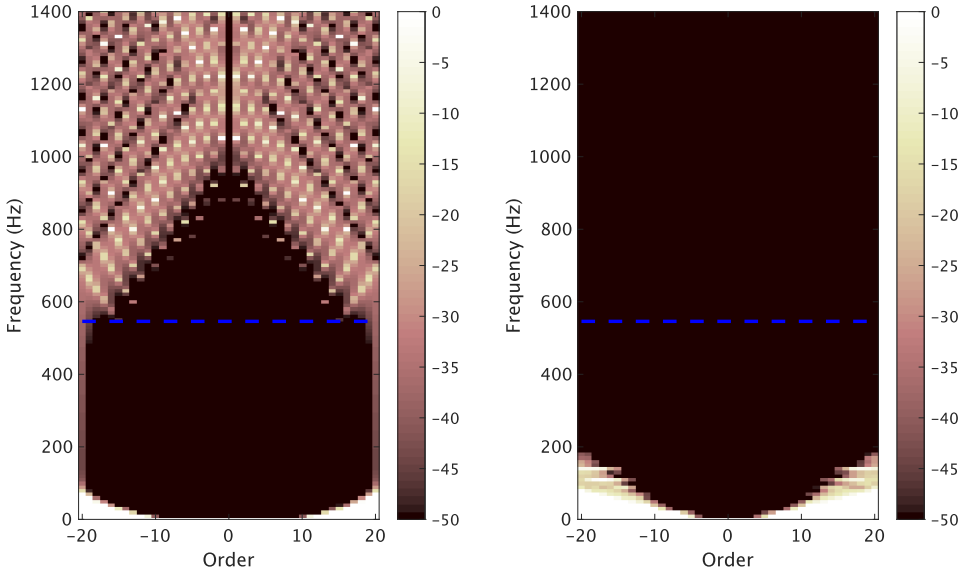
The SDR among four different methods at the frequency range from 200 Hz to 1400 Hz is shown in Figure 3.4. The vertical broken line is the spatial Nyquist frequency in this condition. The circular expansion result corresponds to the SH decomposition based on Eq. (3.16), which is a step in the "CHFF" system. The SDR in the circular expansion method degrades noticeably above the spatial Nyquist frequency. The other three proposed methods, on the other hand, maintain high SDR values at higher frequencies.

The reconstruction error of the SH coefficient of the reference signal was calculated and shown in Figure 3.5. Higher errors are appearing from higher-order components above spatial Nyquist frequency in "CHFF." The striped pattern represents spatial aliasing artifacts. As previously stated, higher errors in the evanescent wave region have only a minor impact on the ANC system.

Our preliminary examination showed that the convergence speed varies among each proposed method. For example, "Proposed (L1)" needs around 100 iterations to converge, whereas, "Proposed ($p = 0.5$)" needs only around 10 iterations. We heuristically determined the number of iterations to be sufficient for convergence.

3.4.3 Noise attenuation performance

The noise attenuation performance among the methods was evaluated while performing the adaptive ANC algorithm. Figure 3.6 depicts the level of noise attenuation after 50 iterations at various frequencies. The vertical broken lines represent the spatial Nyquist frequencies for the reference and error microphone arrays, respectively. Below the Nyquist frequency of the RMA, the "CHFF" and the proposed method gave almost the same results except "Proposed (L1)," which gave an even worse performance. We found that this was due to the strength of the constraint that was controlled by the parameter λ_1. The ℓ_1 constraint was too strong to calculate accurate reference SH coefficients for lower frequencies below the spatial Nyquist frequency. We confirmed that the result improved after manually tuning the parameter for each frequency; however, we use a fixed parameter in this paper because the issue is beyond the scope of this thesis, and we are not tuning any parameters for other methods.



(a) (b)

Figure 3.5: CH coefficients error between original and reconstructed noise field. The horizontal broken line indicates the spatial Nyquist frequency: (a) "CHFF." (b) "Proposed ($p = 0.5$)."

For higher frequencies, the attenuation level of "CHFF" degrades due to spatial aliasing; however, the proposed method can still attenuate the noise. All methods were out of control above the Nyquist frequency of the EMA, and no attenuation could be achieved.

3.4.4 Reduction of the number of reference microphones

When the noise field can be expressed by a sparse set of plane wave weights, the number of reference microphones can be reduced to calculate the reference SH coefficients. Figure 3.7 shows the average noise level after 50 iterations of ANC. Recalling the Nyquist frequency of the RMA, it is natural that the noise attenuation performance starts degrading above 546 Hz for $Q = 41$ and 1092 Hz for $Q = 81$ in "CHFF." On the other hand, despite a reduction in the number of reference microphones, the proposed method performs well over a wide range of frequencies. The ANC system begins to diverge above 1092 Hz for both methods, which is the Nyquist frequency of the EMA.

3.5 Summary

In this chapter, the author focused on the reference signal reconstruction on feedforward ANC system and explored the improvement of the noise attenuation performance by adopting sparse signal representation. As a result of the comparison

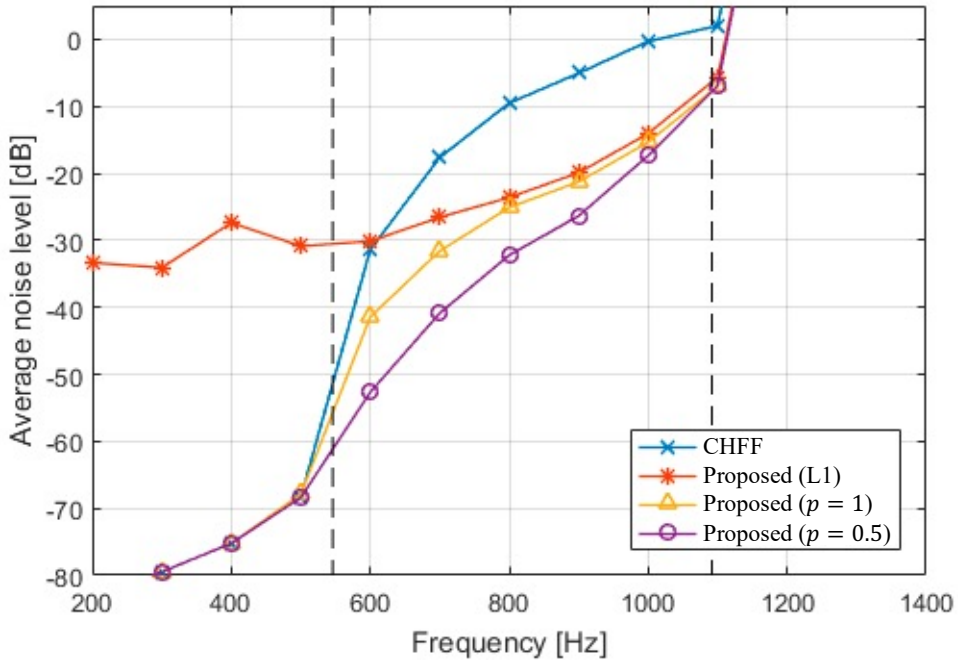
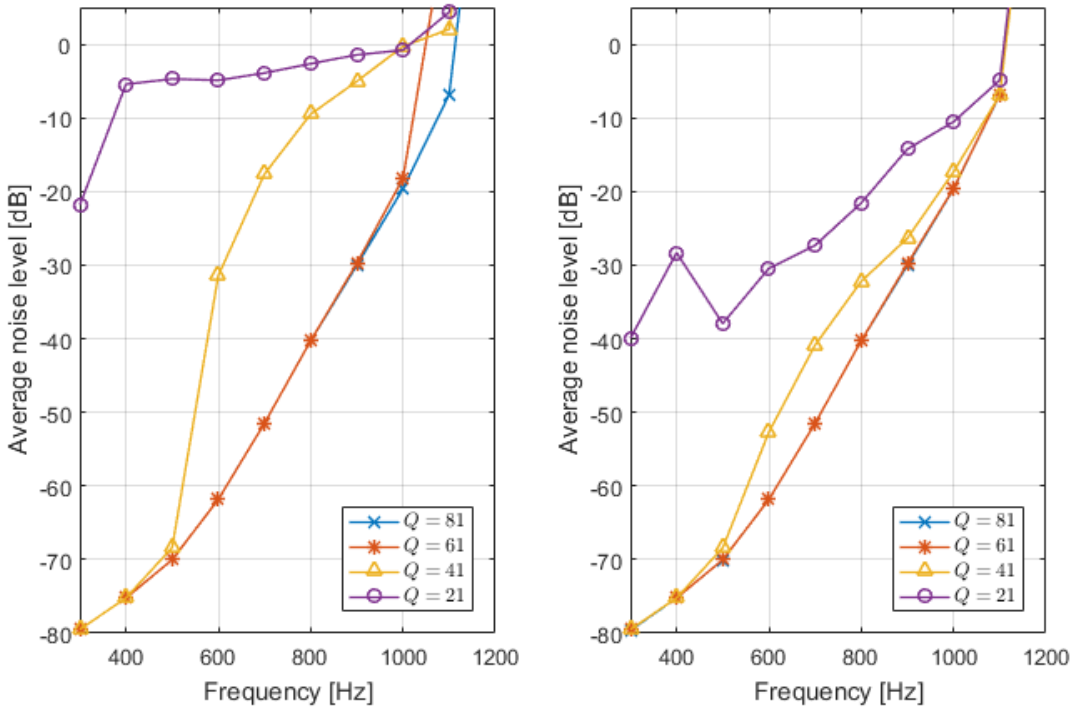


Figure 3.6: Average noise level at several frequency bins after 50 iterations. The vertical broken lines indicate the spatial Nyquist frequencies corresponding to the reference (left) and the error (right) microphone array.

among several methods based on CS, the IRLS approach with $p = 0.5$ performs well in terms of both convergence speed and noise attenuation performance.

Future work will be directed toward applying the proposed approach to the error signal to further improve the system performance. Furthermore, investigating the descent number of iterations in the sparse weights reconstruction is critical for practical implementation. The goal of adaptive filtering-based ANC is to track the changing noise field. When the noise field is stationary, however, frequent adaptation is not required. Similarly, the required number of iterations in sparse weight recovery would be adjusted by monitoring the stationarity of the noise field.



(a) (b)

Figure 3.7: Average noise level among various setups for the number of reference microphones: (a) "CHFF." (b) "Proposed ($p = 0.5$)."

Chapter 4

Spatial active noise control using multiple circular arrays

ANC is a technique, which attenuates an undesired primary noise field by reproducing an antinoise field, having the opposite phase and the same amplitude of the primary noise field using secondary sources. Several attempts have been made in practical applications to cancel the noise field in automobiles [26,27] and aircraft [73], which requires noise attenuation over a large region. Multiple-channel ANC methods [3, 6, 74] have sparked increased interest in controlling the sound field over a large space for this purpose. However, the ANC system requires many microphones uniformly placed to achieve sufficient attenuation of the noise field. Also, it needs many loudspeakers to generate an appropriate antinoise field, both of which result in high implementation costs.

Multiple-channel ANC methods adopting the FXLMS algorithm in temporal-frequency-domain was successful in many practical applications [5,11,12]. Nonetheless, they suffer from the aforementioned disadvantage of requiring a large number of uniformly distributed error microphones to generate a large quiet zone within the ROI. Because the temporal-frequency-domain representation of a sound field does not implicitly describe the underlying sound field aspects, the system can only attenuate the noise field near the error microphone positions using the so-called multiple points method [75].

In contrast, spatial sound field reproduction techniques such as WFS [8] and HOA [58, 59] are decent approaches to mitigate this problem. To reproduce an arbitrary sound field, these methods control acoustic pressure only on a planar or spherical surface rather than the entire region. As a result, the harmonic-domain sound field representation has been applied to ANC methods, resulting in a larger quiet zone and faster convergence speed with even lower computational complexity [19, 21–23]. However, the array geometry for 3D space is constrained to a regularly sampled spherical array in order not to violate the spatial Nyquist theory [33], which makes it impractical in a real environment.

Taking into account the feasibility of the ANC system, several methods have been proposed to reduce the number of microphones and loudspeakers by introduc-

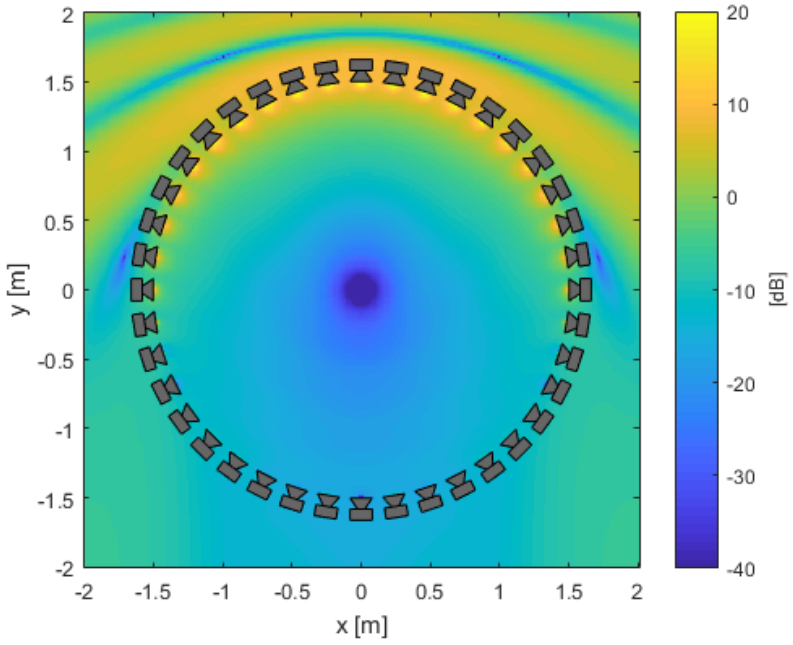


Figure 4.1: Sound field reproduction error caused by dimensional mismatch. The figure is plotted using [76, 77].

ing a sparseness constraint [31, 35, 63, 78], which is known as the CS approach [61]. Although the array structure can be flexible, it is preferable to have a regularly sampled array without any prior knowledge of the noise field, such as the source's direction. Another approach that is commonly used in most practical setups is to use 3D point sources as secondary sources for a 2D sound field reproduction, known as 2.5-dimensional synthesis [79, 80]. The number of microphones and loudspeakers is reduced by using a 2D array structure in this approach. Because the 2D reproduction problem assumes a height invariant sound field model, line sources are a good choice of secondary sources. Actual cabinet loudspeakers, on the other hand, typically exhibit point source characteristics. This dimensional mismatch prevents a perfect reproduction of the desired sound field due to a difference of amplitude decay between the 2D line source and the 3D point source. Perfect reproduction can be achieved only at a given reference point or line [81] as depicted in Figure 4.1, which directly affects the performance of ANC.

In this chapter, the author proposes a simple microphone and loudspeaker array structure for 3D space attenuation in SH-domain ANC, which improves upon the feasibility of spatial Nyquist based arrays. Multiple circular microphone arrays are used to capture the 3D SH coefficients of the noise field, and multiple circular loudspeaker arrays are used to generate the antinoise field. When compared to the conventional SH-domain ANC system, which uses an impractical spherical array, we show that our proposed method provides comparable noise attenuation performance with a practical array structure.

4.1 Problem statement

In this section, the author reviews the theory of capturing and reproducing a sound field using multiple circular microphone and loudspeaker arrays based on the SH decomposition of the sound field proposed in [25, 82].

Let us consider a noise field generated by noise sources outside the ROI in 3D space. We can represent the incident noise field at an arbitrary point $\mathbf{r} \equiv (r, \theta, \phi)$:

$$S(\mathbf{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \alpha_{\nu,\mu} j_\nu(kr) \mathcal{P}_{\nu,\mu}(\cos\theta) E_\mu(\phi), \tag{4.1}$$

where ν and m are integers representing order and degree respectively, $\alpha_{\nu,\mu}$ are the SH coefficients, $j_\nu(\cdot)$ is the νth-order spherical Bessel function, $E_\mu(\phi) = (1/\sqrt{2\pi})e^{i\mu\phi}$ is the normalized exponential function and

$$\mathcal{P}_{\nu,|\mu|}(\cos\theta_q) = \sqrt{\frac{2\nu+1}{2}} \sqrt{\frac{(\nu-|\mu|)!}{(\nu+|\mu|)!}} P_{\nu,|\mu|}(\cos\theta_q) \tag{4.2}$$

is the normalized associated Legendre function. Hereafter, the series expansion in Eq. (4.1) is truncated to $N = \lceil ekR/2 \rceil$ [51] or $N = \lceil kR \rceil$ [52], where R is the radius of the ROI.

4.2 Microphone and loudspeaker arrangements

4.2.1 Multiple circular microphone arrays

The noise field on a circle with a radius r_q and a colatitude θ_q is given by:

$$S(\mathbf{r}_q) \approx \sum_{\nu=0}^{N} \sum_{\mu=-\nu}^{n} \alpha_{\nu,\mu} j_\nu(kr_q) \mathcal{P}_{\nu,|\mu|}(\cos\theta_q) E_\mu(\phi), \tag{4.3}$$

where $\mathbf{r}_q \equiv (r_q, \theta_q, \phi)$. We reformulate Eq. (4.3) by exchanging the order of summations to reveal the Fourier series expansion coefficients, which leads to:

$$S(\mathbf{r}_q) \approx \sum_{\mu=-N}^{N} E_\mu(\phi) \underbrace{\sum_{\nu=|\mu|}^{N} \alpha_{\nu,\mu} j_\nu(kr_q) \mathcal{P}_{\nu,|\mu|}(\cos\theta_q)}_{a_\mu^{(q)}}. \tag{4.4}$$

The coefficient $a_\mu^{(q)}$ is expressed as azimuth harmonic coefficient in [25]. Consider a circular microphone array positioned on the circle. Using orthogonal exponential functions, the azimuth harmonic coefficient can be calculated from microphone outputs:

$$a_\mu^{(q)} \approx \frac{2\pi}{P_q} \sum_{p=1}^{P_q} S(\mathbf{r}_{q,p}) E_{-\mu}(\phi_{q,p}), \tag{4.5}$$

where P_q is the number of microphones of q-th array and $\mathbf{r}_{q,p} \equiv (r_q, \theta_q, \phi_{q,p})$ is the position of the p-th microphone. Note that instead of taking an integral in Eq. (4.5), a summation is taken, which is an approximation due to a spatial discretization.

From Eq. (4.4), the relation between the SH coefficients and azimuth harmonic coefficients is written as:

$$a_\mu^{(q)} = \sum_{\nu=|\mu|}^{N} \alpha_{\nu,\mu} j_\nu(kr_q) \mathcal{P}_{\nu,|\mu|}(\cos\theta_q). \tag{4.6}$$

Taking into account Q circular microphone arrays, a matrix formulation can be constructed:

$$\mathbf{J}_\mu \boldsymbol{\alpha}_\mu = \mathbf{A}_\mu, \tag{4.7}$$

where

$$\mathbf{J}_\mu = \begin{bmatrix} j_{|\mu|}(kr_1)\mathcal{P}_{|\mu|,|\mu|}(\cos\theta_1) & \cdots & j_N(kr_1)\mathcal{P}_{N,|\mu|}(\cos\theta_1) \\ \vdots & \ddots & \vdots \\ j_{|\mu|}(kr_Q)\mathcal{P}_{|\mu|,|\mu|}(\cos\theta_Q) & \cdots & j_N(kr_Q)\mathcal{P}_{N,|\mu|}(\cos\theta_Q) \end{bmatrix}, \tag{4.8}$$

$\boldsymbol{\alpha}_\mu = [\alpha_{|\mu|,\mu}, \alpha_{(|\mu|+1),\mu}, \ldots, \alpha_{N,\mu}]^\mathsf{T}$, and $\mathbf{A}_\mu = [a_\mu^{(1)}, \ldots, a_\mu^{(Q)}]^\mathsf{T}$.

In order to get a stable solution of Eq. (4.7), we need to set up an over-determined system, which requires more than $N - |\mu|$ arrays. Furthermore, a high condition number in Eq. (4.8) must be avoided by designing an array geometry to avoid containing small values of spherical Bessel functions and normalized associated Legendre functions. [25] contains a more detailed discussion of the design of microphone array geometry.

In the x-y plane, odd coefficients have no contribution to the sound field [25]. By setting the colatitude $\theta_q = \pi/2$ for every array, Eq. (4.8) can be reduced for $\nu = |\mu|, |\mu| + 2, \ldots, N$, which leads to a reduction of the number of microphone arrays.

4.2.2 Multiple circular loudspeaker arrays

A desired antinoise field can be written in a same manner as Eq. (4.3):

$$S_{\text{des}}(\mathbf{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \gamma_{\nu,\mu} j_\nu(kr) \mathcal{P}_{\nu,\mu}(\cos\theta) E_\mu(\phi). \tag{4.9}$$

Suppose the ℓ-th circular loudspeaker aperture is producing an antinoise field:

$$S(\mathbf{r}) = \int_0^{2\pi} \rho_\ell(\varphi) \frac{1}{4\pi} \frac{e^{ik||\mathbf{r}_\ell - \mathbf{r}||}}{||\mathbf{r}_\ell - \mathbf{r}||} d\varphi, \tag{4.10}$$

where $\mathbf{r}_\ell \equiv (r_\ell, \theta_\ell, \varphi)$ is the position on the circular aperture and $\rho_\ell(\varphi)$ is the ℓth circular continuous loudspeaker aperture function. Spherical expansion of the incident

sound field due to a unit-magnitude point source is given by

$$\frac{1}{4\pi}\frac{e^{ik||\mathbf{r}_\ell-\mathbf{r}||}}{||\mathbf{r}_\ell-\mathbf{r}||}=\sum_{\nu=0}^{\infty}\sum_{\mu=-\nu}^{\nu}-ikh_\nu^{(2)}(kr_\ell)\mathcal{P}_{\nu,\mu}(\cos\theta_\ell) \times E_{-\mu}(\varphi)j_\nu(kr)\mathcal{P}_{\nu,\mu}(\cos\theta)E_\mu(\phi), \tag{4.11}$$

where $h_\nu^{(2)}(\cdot)$ is the νth-order spherical Hankel function of second kind. Substituting Eq. (4.11) into Eq. (4.10) leads to:

$$S(\mathbf{r})=\sum_{\nu=0}^{\infty}\sum_{\mu=-\nu}^{\nu}-ik\beta_\mu^{(\ell)}h_\nu^{(2)}(kr_\ell)\mathcal{P}_{\nu,\mu}(\cos\theta_\ell) \times j_\nu(kr)\mathcal{P}_{\nu,\mu}(\cos\theta)E_\mu(\phi), \tag{4.12}$$

where

$$\beta_\mu^{(\ell)}=\int_0^{2\pi}\rho_\ell(\varphi)E_{-\mu}(\varphi)d\varphi \tag{4.13}$$

is the Fourier series expansion coefficient of the circular aperture function. Equating Eq. (4.9) and Eq. (4.12) while taking into account L circular loudspeaker apertures, we have

$$\gamma_{\nu,\mu}=\sum_{\ell=1}^{L}-ik\beta_\mu^{(\ell)}h_\nu^{(2)}(kr_\ell)\mathcal{P}_{\nu,\mu}(\cos\theta_\ell). \tag{4.14}$$

We now have simultaneous equations:

$$\mathbf{\Gamma}_\mu=\mathbf{H}_\mu\mathbf{B}_\mu, \tag{4.15}$$

where

$$\mathbf{H}_\mu=-ik\times\begin{bmatrix}h_{|\mu|}^{(2)}(kr_1)\mathcal{P}_{|\mu|,|\mu|}(\cos\theta_1) & \cdots & h_{|\mu|}^{(2)}(kr_L)\mathcal{P}_{|\mu|,|\mu|}(\cos\theta_L)\\ \vdots & \ddots & \vdots\\ h_N^{(2)}(kr_1)\mathcal{P}_{N,|\mu|}(\cos\theta_1) & \cdots & h_N^{(2)}(kr_L)\mathcal{P}_{N,|\mu|}(\cos\theta_L)\end{bmatrix}, \tag{4.16}$$

$\mathbf{\Gamma}_\mu=[\gamma_{|\mu|,\mu},\gamma_{(|\mu|+1),\mu},\ldots,\gamma_{N,\mu}]^\mathsf{T}$, and $\mathbf{B}_\mu=[\beta_\mu^{(1)},\ldots,\beta_\mu^{(L)}]^\mathsf{T}$. The SH-domain coefficients of the loudspeaker weight can be calculated by solving Eq. (4.15) in a least-square sense. Finally, the loudspeaker weight of the v-th loudspeaker at the ℓ-th circular array is given by

$$\tilde{\rho}_\ell(\phi_v)=\frac{2\pi}{V_\ell}\sum_{\mu=-N}^{N}\beta_\mu^{(\ell)}E_\mu(\phi_v), \tag{4.17}$$

where V_ℓ is the number of loudspeakers of ℓ-th array.

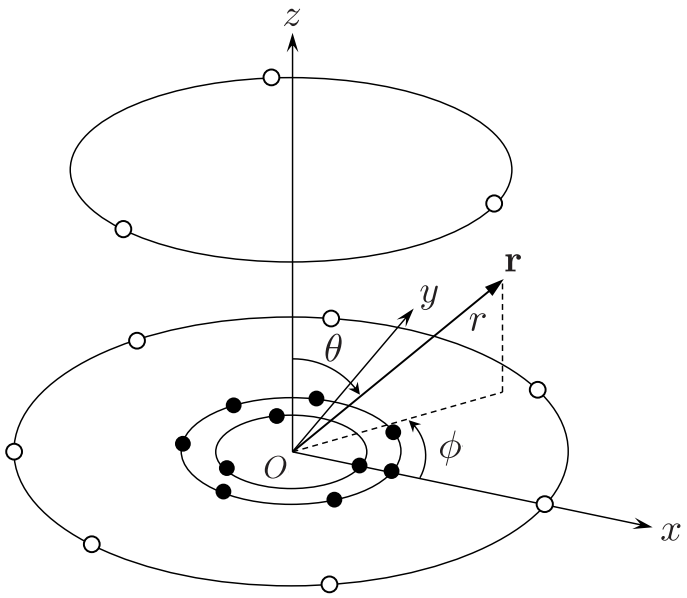


Figure 4.2: Multiple circular microphone and loudspeaker array structure. The microphones are marked as • and the loudspeakers are marked as ∘.

4.3 Spherical-harmonic-domain active noise control

In this section, an SH-domain adaptive filtering algorithm utilizing azimuth harmonic representation is described.

Consider a residual noise field, which is a sum of the noise field described by Eq. (4.1) and the antinoise field described by Eq. (4.9). SH coefficients of the residual sound field can be written as:

$$r_{\nu,\mu} = \alpha_{\nu,\mu} + \gamma_{\nu,\mu}. \tag{4.18}$$

By concatenating all available coefficients for each degree μ, we can write Eq. (4.18) in a matrix form:

$$\mathbf{R}_\mu = \boldsymbol{\alpha}_\mu + \boldsymbol{\Gamma}_\mu, \tag{4.19}$$

where $\mathbf{R}_\mu = [r_{|\mu|,\mu}, r_{(|\mu|+1),\mu}, \ldots, r_{N,\mu}]^\mathsf{T}$. We set a cost function as a sum of the squared residual signal:

$$\mathcal{J}_\mu = \sum_{\nu=|\mu|}^{N} |r_{\nu,\mu}|^2 = \mathbf{R}_\mu^\mathsf{H}\mathbf{R}_\mu, \tag{4.20}$$

where $(\cdot)^\mathsf{H}$ denotes the Hermitian-transpose operation. A gradient of the cost function $\mathcal{J}_\mu$ can be calculated as

$$\frac{\partial \mathcal{J}_\mu}{\partial \mathbf{B}_\mu} = \mathbf{H}_\mu^\mathsf{H}(\boldsymbol{\alpha}_\mu + \mathbf{H}_\mu\mathbf{B}_\mu) = \mathbf{H}_\mu^\mathsf{H}\mathbf{R}_\mu. \tag{4.21}$$

By utilizing the steepest descent algorithm, we can update the coefficients $\mathbf{B}_\mu$ as

$$\mathbf{B}_\mu(\iota+1) = \mathbf{B}_\mu(\iota) - \eta_\mu \mathbf{H}_\mu^{\mathsf{H}} \mathbf{R}_\mu, \tag{4.22}$$

where ι is the iteration index and η_μ is the step size. Finally, the canceling loudspeaker weight of each loudspeaker can be calculated by Eq. (4.17).

To calculate the residual signal $\mathbf{R}_\mu$, first, calculate the azimuth harmonic coefficients from microphone outputs by Eq. (4.5), then solve the inverse problem of Eq. (4.7), i.e., $\boldsymbol{\alpha}_\mu = \mathbf{J}_\mu^+ \mathbf{A}_\mu$. The use of azimuth harmonic representation has the advantage of allowing us to design a flexible microphone array geometry to calculate SH coefficients while taking into account the underlying structure of the wave propagation [25]. It is important to note that once the array geometry is fixed, $\mathbf{J}_\mu^+$ is the static matrix, which is appropriate for an adaptive process because there is no matrix inversion with each iteration.

4.4 Validation

4.4.1 Experimental conditions

Methods

The noise attenuation performance is evaluated among the following methods:

(i) "SHD" corresponding to the conventional SH-domain adaptive algorithm, which utilizes spherical array geometry, as a reference method;

(ii) "FD" corresponding to the conventional temporal-frequency-domain adaptive algorithm, also termed multi-point method;

(iii) "AHD" corresponding to the proposed method.

Geometrical setup

"SHD" requires a spherical microphone and loudspeaker array, whose microphones and loudspeakers are equally distributed on the surface of a sphere. To avoid spatial aliasing, arrays must have at least $(N+1)^2$ microphones and loudspeakers [33]. We used 20 microphones and loudspeakers to create a third order system. The microphones and loudspeakers were placed at the vertices of the dodecahedron to avoid orthonormality error [83], with radii 0.2 m and 1.0 m.

Although we consider sound propagation in 3D space, we limited the ROI in the evaluation to the x-y plane because the height of the human head varies little in many practical applications, such as sitting in a vehicle seat. As mentioned before, we only need the even-order coefficients to control the x-y plane. Table 4.1 shows the SH coefficients up to the third order and Table 4.2 shows only the even-order coefficients. For each degree μ in Table 4.2, there are at most two coefficients, thus we need at least two circular microphone arrays to calculate the coefficients.

Table 4.1: All SH coefficients.

ν \ μ	-3	-2	-1	0	1	2	3
0				$\alpha_{0,0}$			
1			$\alpha_{1,(-1)}$	$\alpha_{1,0}$	$\alpha_{1,1}$		
2		$\alpha_{2,(-2)}$	$\alpha_{2,(-1)}$	$\alpha_{2,0}$	$\alpha_{2,1}$	$\alpha_{2,2}$	
3	$\alpha_{3,(-3)}$	$\alpha_{3,(-2)}$	$\alpha_{3,(-1)}$	$\alpha_{3,0}$	$\alpha_{3,1}$	$\alpha_{3,2}$	$\alpha_{3,3}$

Table 4.2: Even-order SH coefficients.

ν \ μ	-3	-2	-1	0	1	2	3
0				$\alpha_{0,0}$			
1			$\alpha_{1,(-1)}$		$\alpha_{1,1}$		
2		$\alpha_{2,(-2)}$		$\alpha_{2,0}$		$\alpha_{2,2}$	
3	$\alpha_{3,(-3)}$		$\alpha_{3,(-1)}$		$\alpha_{3,1}$		$\alpha_{3,3}$

When we look at $\mu = 1$, for example, two even-order coefficients are present. Seven microphones are placed in the first array, which needs at least $2N + 1$ microphone to capture the third order SH coefficient $\alpha_{3,1}$. The coefficient $\alpha_{1,1}$ can be captured by placing three microphones as the second array. Figure 4.2 depicts the simulated array structure of "FD" and "AHD."

The loudspeaker array geometry, like the microphone array geometry, was based on the result of our preliminary experiment. Two loudspeaker arrays are used to control the sound field on the x-y plane. Seven loudspeakers are placed at $z = 0$ m with a radius $r_1 = 1$ m as a first array, and three loudspeakers at $z = 0.8$ m with a radius $r_2 = 0.8$ m as a second array.

Both free field and reverberant field were simulated. A rectangular room of 6 m $\times$ 5 m $\times$ 4 m in size with a wall reflection coefficient of 0.6 was simulated as the reverberant field. We set the coordinate origin at $(x, y, z) = (-1, -0.5, -0.5)$ with respect to the center of the room.

Noise signal

We placed a single noise source outside of the microphone and loudspeaker arrays. As the primary noise source, a point source of tone signal with an operation frequency of 400 Hz was used. Each error microphone output receives a signal-to-noise ratio of 60 dB white Gaussian noise. Note that a signal power is calculated as an average of the initial error microphone outputs and is used as a reference value to simulate the power of the measurement noise.

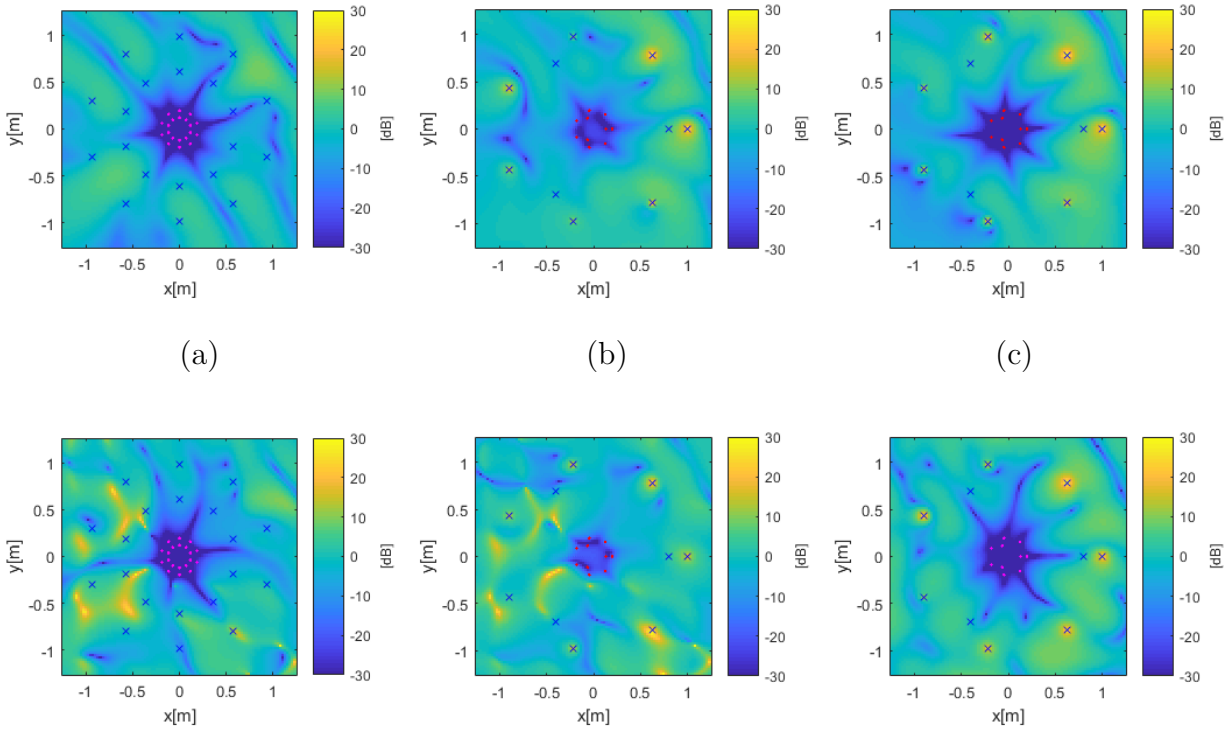


Figure 4.3: Residual noise fields in a free field (upper row) and reverberant field (lower row) after 50 steps of ANC processing: (a)(d) "SHD." (b)(e) "FD." (c)(f) "AHD." Symbols "x" and "." indicate the loudspeaker and microphone positions, respectively. Note that all microphones and loudspeakers are plotted by mapping to the x-y plane.

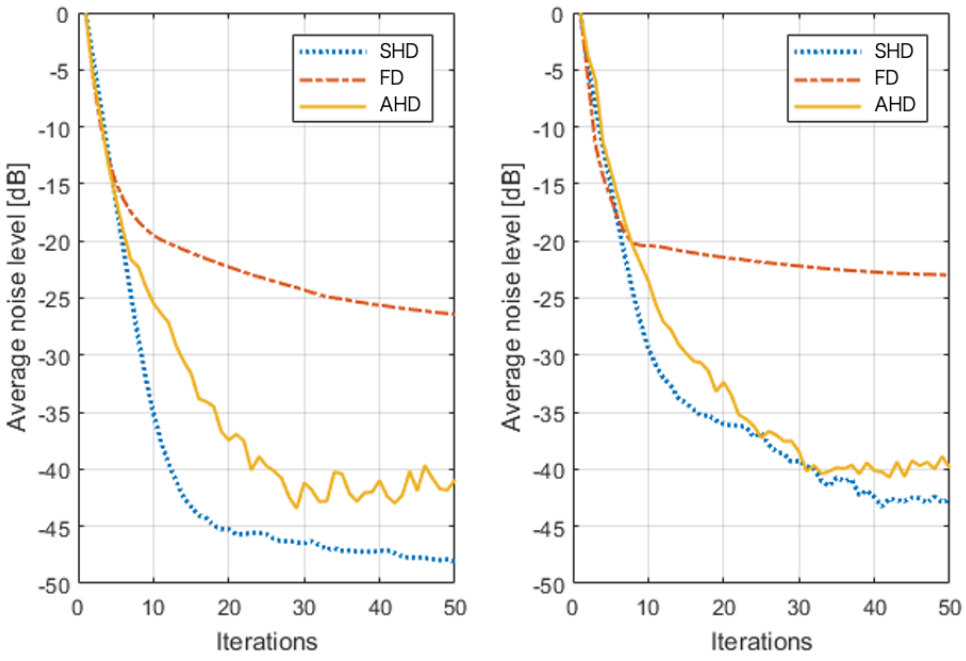
Evaluation metric

To compare the noise attenuation performance of each method, the average noise level inside the ROI was calculated. The ROI was a circular region with a radius of 0.2 m on the x-y plane inside the inner microphone array.

Since "SHD" used the spherical arrays, which is an ideal configuration to control all the SH coefficients up to a certain order, we expected that "SHD" gives the best performance. "AHD" used fewer microphones and loudspeakers to control only the even-order SH coefficients to limit the ROI to the x-y plane. "AHD" is expected to give better performance than "FD," which is the conventional method with the same array structure. Besides, we evaluated how close the performance of "AHD" with fewer components and the simple array structure is to that of "SHD."

4.4.2 Noise attenuation performance

We evaluated the noise attenuation performance of each method. Figure 4.3 demonstrates the amplitude difference of the residual noise field from the original noise field



(a) (b)

Figure 4.4: Average noise level inside the region of interest: (a) Free field. (b) Reverberant field.

after 50 steps of the adaptive processing. As we can see, "AHD" provides a comparable quiet zone to "SHD." Because it does not implicitly describe the underlying aspects of the sound field, "FD" can only attenuate noise near the microphone position. Figure 4.4 shows the average noise level inside the ROI, i.e., inside the circle of radius 0.2 m. Both the free field and reverberant field results show a similar tendency among the methods. "SHD" achieves the fastest convergence speed as expected because it employs ideal spherical arrays with enough microphones and loudspeakers to control the entire region in the SH-domain and performs the adaptive filtering order independently. The convergence speed of "AHD" is much faster compared to "FD," on the other hand, it is slightly slower compared to "SHD." However, the number of microphones and loudspeakers is less and the array geometry is quite simple since there are only a few circular arrays, which is applicable for practical implementation. "AHD" seems to be more sensitive to the measurement noise compared to the other methods. One possible explanation is that "AHD" has to solve two minimization problems, i.e., one is for solving Eq. (4.7) to calculate the error SH coefficients and another one is for solving Eq. (4.15) to calculate the SH coefficients of the loudspeaker outputs, hence estimating errors are accumulated.

4.5 Summary

In this chapter, we investigated the SH-domain ANC algorithm using multiple circular arrays in 3D space. Normally, noise attenuation over a large area necessitates uniformly distributed many microphones and loudspeakers, which limits the system's practicability. To alleviate this problem, we simplified the array structure by introducing the azimuth harmonic representation of the signal. The adaptive update of the ANC system was derived in the azimuth harmonic-domain.

In the simulation, the ANC system controlling the sound field up to a third order was constructed using two circular microphone arrays as well as two circular loudspeaker arrays. This simpler array structure is appropriate for use in practical applications. Using the appropriate spherical array, the proposed method demonstrated comparable noise attenuation performance to the reference SH-domain ANC method. Although the simulation results demonstrated a significant improvement in the specific condition, the arrays must be carefully designed for other situations, such as a different frequency band, truncation order, and ROI size. One of the future works would be a study on a systematic algorithm for determining array geometry.

The results shown in this chapter indicate that a non-spherical array structure has the potential of attenuating the noise in a sizable 3D space. The research presented in this chapter provided some insights into the use of multiple groups of microphones and loudspeakers to reduce the feasibility of the array structure while maintaining noise attenuation performance, which is the motivation for this thesis.

Chapter 5

Array-geometry-aware spatial active noise control based on direction of arrival weighting

ANC over a sizable space is a challenging problem in acoustic signal processing. An arbitrary noise field is captured and attenuated using multiple sensors and secondary sources. This method is thought to be a progression from single-channel ANC [3–5] to multiple-channel ANC [3,6]. One of the most widely used techniques for creating a quiet region over a space is a multiple-channel system using the FXLMS algorithm [11,12]. This system attempts to minimize the residual error only in the vicinity of the error sensors. Thus, a drawback of the system is that the error sensors must be distributed densely throughout the ROI to achieve a large quiet region.

In the context of controlling a sound field over a space, spatial sound field representation in the harmonic-domain [58,59], also known as mode decomposition, is a reasonable approach. Harmonic-domain ANC methods [19,21–24,84–86] have a larger quiet region and a faster convergence speed than conventional multiple-point pressure control methods, despite having a lower computational cost. However, one disadvantage of most of these approaches [19,21–24] is that the sensors and secondary sources must be uniformly placed on a suitable surface, such as a sphere, to achieve optimal performance. Although this constraint of the array geometry gives a stable and efficient signal representation in the harmonic-domain, spherical arrays in a 3D space with numerous sensors and secondary sources are physically difficult to set up in practice for the following reasons: (1) the uniform sampling of a spherical surface is difficult; (2) the physical array structure becomes complex; (3) sensors and secondary sources placed in the lower hemisphere are difficult to install.

Taking into account the feasibility of the ANC system, several attempts have been made to reduce the number of sensors and secondary sources by introducing a sparseness constraint [35,63,78,87]. Another approach is that instead of using a spherical array, multiple circular arrays are used to first transform signals into azimuth harmonic coefficients [25] and then the SH coefficients are calculated in a least-squares sense to perform ANC in the SH-domain as introduced in Chap-

ter 4 [88]. The noise attenuation capability, on the other hand, is determined by the array geometry. As a result, the geometry must be carefully designed based on the results of the harmonic-domain analysis [25,82] in order to precisely capture and generate the primary noise and antinoise fields.

In this chapter, we propose a direction of arrival (DOA) weighting method that attempts to prioritize the controllable DOA of the sound field determined by the secondary source geometry. We present array geometries that employ multiple horizontal rings for both sensors and secondary sources, which are more practical than spherical array geometries. We consider them to be reasonable geometries because horizontal control of the sound field is critical for human perception due to the height of ears. The multiple-horizontal-ring arrays have a high resolution of the sensors and secondary sources on the azimuth angle; hence, we assume that the DOA nearly parallel to the horizontal plane is the controllable DOA. Although a noise field arriving from the same height as the arrays can be effectively captured and controlled, due to the sensor and secondary source geometries, it is difficult to attenuate a noise field arriving from other directions, i.e., above and below the arrays. The power of the DOA distribution is defined based on the DOA weighting and then used to formulate the cost function that will be minimized by driving the secondary sources. Controllable harmonic orders dependent on array geometry are prioritized without using prior knowledge of the primary noise field by appropriately designing the DOA weighting. A numerical simulation in three-dimensional space is used to evaluate the proposed algorithm.

5.1 Problem statement

The objective of spatial ANC is to attenuate a primary noise field in a target region by using multiple sensors and secondary sources. We use M_{e} error sensors, M_{r} reference sensors, and L secondary sources. The error sensors are assumed to be in a spherical region Ω that includes the target region, while the secondary source and reference sensor arrays are outside Ω. In the practical setup, for example, multiple-horizontal-ring geometry is used for each array, as shown in Figure 5.1. Our goal is to achieve spatial ANC with these practical sensors and the secondary source geometry by calculating the optimal driving signals of the secondary sources, denoted by $\mathbf{d} = [d_1, \ldots, d_L]^{\mathsf{T}}$ with the transpose operator $(\cdot)^{\mathsf{T}}$, from signals received by error and reference sensors.

The origin is set at the center of Ω and the position vector is denoted as $\mathbf{r} = r\hat{\mathbf{r}}$ in spherical coordinates, where $\hat{\mathbf{r}}$ is a unit vector in 3D space. The sound pressure of wave number k at position $\mathbf{r}$, denoted as $S(\mathbf{r}, k)$, is approximated in the SH-domain as

$$S(\mathbf{r}, k) = \sum_{\nu=0}^{N} \sum_{\mu=-\nu}^{\nu} \alpha_{\nu,\mu}(k) j_\nu(kr) Y_{\nu,\mu}(\hat{\mathbf{r}}), \tag{5.1}$$

where $Y_{\nu,\mu}(\cdot)$ is the (ν, μ)th-order SH function and $j_\nu(\cdot)$ is the νth-order spherical Bessel function. The truncation order N is usually determined by adopting the

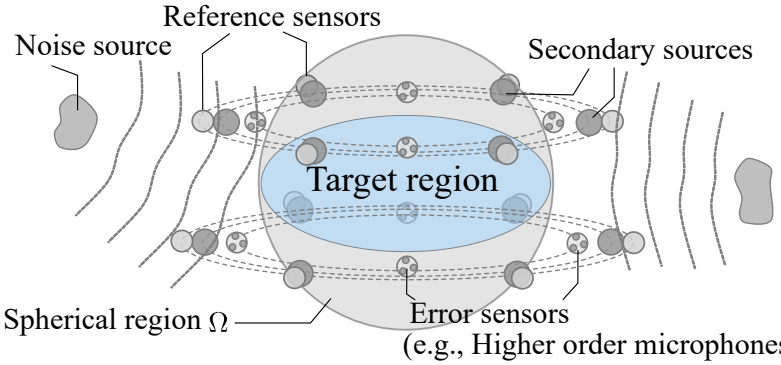


Figure 5.1: Example of the geometric setup of feedforward spatial ANC system.

criterion $N = \lceil ekR/2 \rceil$ [51] or $N = \lceil kR \rceil$ [52], where R is the radius of the target region. We hereafter omit k for notational simplicity. The SH coefficients $\alpha_{\nu,\mu}$ up to the truncation number N, $\boldsymbol{\alpha} := [\alpha_{0,0}, \ldots, \alpha_{N,N}]^\mathsf{T} \in \mathbb{C}^{(N+1)^2}$, can be obtained using the observed signals from the error sensors distributed inside or around Ω, as described in the next paragraph. As error signals, we use the global SH coefficients $\boldsymbol{\alpha}$, whose expansion origin is the coordinate origin.

The global SH coefficients $\boldsymbol{\alpha}$ can be estimated by several methods, such as uniformly placing error microphones on a spherical surface or distributing higher-order microphone (HOM) as the error sensors [71]. We return to the latter method, which is based on the translation operator [89], because it allows for greater flexibility in the placement of the error microphones. HOMs are typically spherical/circular microphone arrays that are intended to capture higher-order SH coefficients. We assume that the local SH coefficient series up to the V-th order with the expansion origin $\mathbf{r}_q$, denoted as $\boldsymbol{\alpha}_{\mathrm{local}}^{(q)} \in \mathbb{C}^{(V+1)^2}$, are obtained from the q-th HOM at $\mathbf{r}_q$. By using the translation operator, which relates the SH coefficients for different expansion origins, the global and local coefficients are approximately represented as

$$\begin{bmatrix} \vdots \\ \boldsymbol{\alpha}_{\mathrm{local}}^{(q)} \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \mathbf{T}^{(q)} \\ \vdots \end{bmatrix} \boldsymbol{\alpha}, \tag{5.2}$$

where $\mathbf{T}^{(\cdot)} \in \mathbb{C}^{(V+1)^2 \times (N+1)^2}$ is the translation matrix. By denoting the error SH coefficients as $\boldsymbol{\alpha}_{\mathrm{local}} := [\boldsymbol{\alpha}_{\mathrm{local}}^{(1)\mathsf{T}}, \ldots, \boldsymbol{\alpha}_{\mathrm{local}}^{(Q)\mathsf{T}}]^\mathsf{T}$, Eq. (5.2) can be represented as $\boldsymbol{\alpha}_{\mathrm{local}} = \boldsymbol{\tau}\boldsymbol{\alpha}$ with the concatenated matrix $\boldsymbol{\tau} := [\mathbf{T}^{(1)\mathsf{T}}, \ldots, \mathbf{T}^{(Q)\mathsf{T}}]^\mathsf{T} \in \mathbb{C}^{Q(V+1)^2 \times (N+1)^2}$; therefore, the global SH coefficients $\boldsymbol{\alpha}$ can be obtained as $\boldsymbol{\alpha} = \boldsymbol{\tau}^+ \boldsymbol{\alpha}_{\mathrm{local}}$, where $(\cdot)^+$ denotes the Moore–Penrose pseudoinverse matrix.

In this situation, $S(\mathbf{r})$ inside Ω can be expanded using the Herglotz wave function as

$$S(\mathbf{r}) = \int_{\hat{\mathbf{k}} \in \mathbb{S}^2} p(\hat{\mathbf{k}}) e^{ik\langle \hat{\mathbf{k}}, \mathbf{r} \rangle} d\hat{\mathbf{k}}, \tag{5.3}$$

where the unit vector $\hat{\mathbf{k}}$ denotes the DOA of the plane wave, $\langle \cdot, \cdot \rangle$ is the Euclidean inner product, and the right-hand side of Eq. (5.3) is integral on the unit sphere $\mathbb{S}^2$.

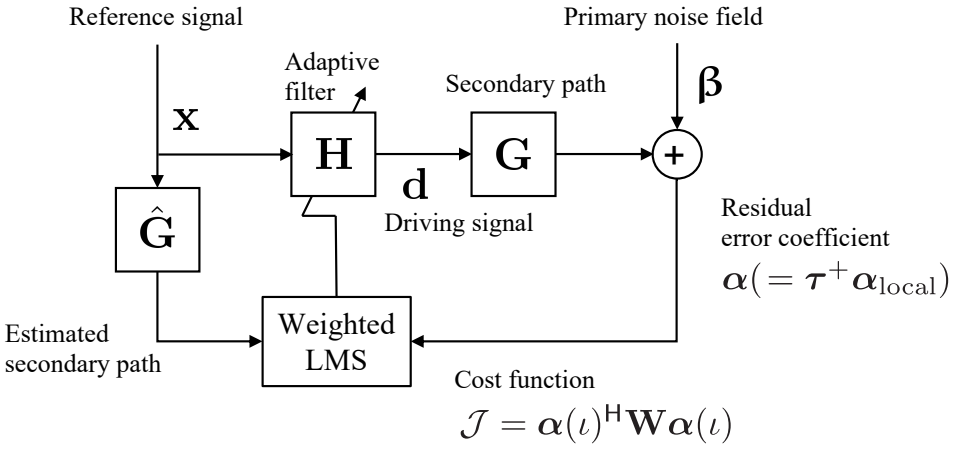


Figure 5.2: Block diagram of the proposed algorithm.

The DOA distribution of the plane wave $p(\hat{\mathbf{k}})$ can be expanded in the SH-domain as

$$p(\hat{\mathbf{k}}) = \sum_{\nu=0}^{N}\sum_{\mu=-\nu}^{\nu} \rho_{\nu,\mu} Y_{\nu,\mu}(\hat{\mathbf{k}}) \tag{5.4}$$

$$=: \sum_{\nu=0}^{N}\sum_{\mu=-\nu}^{\nu} p_{\nu,\mu}(\hat{\mathbf{k}}). \tag{5.5}$$

From Eq. (5.3) and Eq. (5.5), $S(\mathbf{r})$ can be represented by the plane wave decomposition in the SH-domain as

$$S(\mathbf{r}) = \sum_{\nu=0}^{N}\sum_{\mu=-\nu}^{\nu} \int_{\hat{\mathbf{k}}\in\mathbb{S}^2} p_{\nu,\mu}(\hat{\mathbf{k}}) e^{ik\langle\hat{\mathbf{k}},\mathbf{r}\rangle} d\hat{\mathbf{k}}. \tag{5.6}$$

Note that the coefficients $\rho_{\nu,\mu}$ are related to the SH coefficients $\alpha_{\nu,\mu}$ as described in [89]:

$$\alpha_{\nu,\mu} = 4\pi i^{\nu} \rho_{\nu,\mu}. \tag{5.7}$$

From Eq. (5.7), the DOA distribution $p_{\nu,\mu}(\hat{\mathbf{k}})$ can be interpreted as the sound field element of direction $\hat{\mathbf{k}}$ derived from the SH coefficients $\alpha_{\nu,\mu}$ and represented as

$$p_{\nu,\mu}(\hat{\mathbf{k}}) = (4\pi i^{\nu})^{-1} \alpha_{\nu,\mu} Y_{\nu,\mu}(\hat{\mathbf{k}}). \tag{5.8}$$

5.2 Direction of arrival weighting

We derive an adaptive ANC algorithm with the DOA weighting based on the FXLMS algorithm depicted in Figure 5.2. In the proposed algorithm, the driving signals $\mathbf{d}$ are

obtained by the linear filtering of the observed signals from the reference sensors, denoted as $\mathbf{x} = [x_1, \ldots, x_{M_\mathrm{r}}]^\mathsf{T}$. By denoting the truncated SH coefficients of the primary noise field as $\boldsymbol{\beta} := [\beta_{0,0}, \ldots, \beta_{N,N}]^\mathsf{T} \in \mathbb{C}^{(N+1)^2}$, the global SH coefficients $\boldsymbol{\alpha}$ are represented as

$$\boldsymbol{\alpha} = \boldsymbol{\beta} + \mathbf{G}\mathbf{d} = \boldsymbol{\beta} + \mathbf{G}\mathbf{H}\mathbf{x}, \tag{5.9}$$

where $\mathbf{G} \in \mathbb{C}^{(N+1)^2 \times L}$ is the acoustic transfer function from the secondary sources to the error sensors in the SH-domain and $\mathbf{H} \in \mathbb{C}^{L \times M_\mathrm{r}}$ is the filter coefficient matrix.

We define the power of the DOA distribution derived from the (ν, μ)-th SH coefficient based on the plane wave decomposition in the SH-domain Eq. (5.6) as

$$\Pi_{\nu,\mu} = \int_{\hat{\mathbf{k}} \in \mathbb{S}^2} |p_{\nu,\mu}(\hat{\mathbf{k}})|^2 d\hat{\mathbf{k}}. \tag{5.10}$$

In the configuration of the secondary sources described in Section 5.1, the sound field component from a specific DOA, such as a higher elevation direction, cannot be well controlled by the sources. We define the cost of the power of the DOA distribution to attenuate the controllable sound field component in the proposed algorithm by taking the DOA weight into account. Therefore, we define the weighted power of the DOA distribution for the (ν, μ)-th order as

$$\Pi^w_{\nu,\mu} = \int_{\hat{\mathbf{k}} \in \mathbb{S}^2} w(\hat{\mathbf{k}}) |p_{\nu,\mu}(\hat{\mathbf{k}})|^2 d\hat{\mathbf{k}}, \tag{5.11}$$

where $w(\hat{\mathbf{k}})$ is the non-negative DOA weighting function. By substituting Eq. (5.8) into Eq. (5.11), the reformulation of Eq. (5.11) in the SH-domain is obtained as

$$\Pi^w_{\nu,\mu} = \alpha^*_{\nu,\mu} \alpha_{\nu,\mu} \tilde{w}_{\nu,\mu}, \tag{5.12}$$

where

$$\tilde{w}_{\nu,\mu} = \frac{1}{(4\pi)^2} \int_{\hat{\mathbf{k}} \in \mathbb{S}^2} w(\hat{\mathbf{k}}) Y_{\nu,\mu}(\hat{\mathbf{k}})^* Y_{\nu,\mu}(\hat{\mathbf{k}}) d\hat{\mathbf{k}}. \tag{5.13}$$

The proposed adaptive algorithm minimizes the instantaneous summation of the weighted power of the DOA distribution, which is

$$\mathcal{J} = \sum_{\nu=0}^{N} \sum_{\mu=-\nu}^{\nu} \Pi^w_{\nu,\mu}(\iota), \tag{5.14}$$

where ι represents the iteration index. We design $w(\hat{\mathbf{k}})$ according to the controllable sound field component dependent on the secondary source geometry so that $w(\hat{\mathbf{k}})$ becomes large for the controllable DOA $\hat{\mathbf{k}}$. From Eq. (5.12), Eq. (5.14) can be formulated using the SH coefficients $\boldsymbol{\alpha}$ as

$$\mathcal{J} = \boldsymbol{\alpha}(\iota)^\mathsf{H} \mathbf{W} \boldsymbol{\alpha}(\iota), \tag{5.15}$$

where $\mathbf{W} \in \mathbb{C}^{(N+1)^2 \times (N+1)^2}$ is a diagonal matrix and $(\mathbf{W})^{\nu,\mu}_{\nu,\mu} = \tilde{w}_{\nu,\mu}$. The update rule for the filter coefficients $\mathbf{H}$ is derived by the LMS algorithm as

$$\mathbf{H}(\iota+1) = \mathbf{H}(\iota) - \eta \frac{\partial}{\partial \mathbf{H}^*} \mathcal{J} \tag{5.16}$$

$$= \mathbf{H}(\iota) - \eta \hat{\mathbf{G}}^{\mathsf{H}} \mathbf{W} \boldsymbol{\alpha}(\iota) \mathbf{x}(\iota)^{\mathsf{H}}, \tag{5.17}$$

where η represents the step size of the adaptation and $\hat{\mathbf{G}}$ is the estimated secondary path. Note that $\mathbf{W}$ is independent of both k and $\mathbf{H}(\iota)$ and can be calculated in advance.

The proposed algorithm is closely related to the method of minimizing the squared error of SH coefficients ($\boldsymbol{\alpha}^{\mathsf{H}}\boldsymbol{\alpha}$), which corresponds to the theoretical extension of [23] to the feedforward FXLMS algorithm. We hereafter call this method the mode-matching-based method (MM). When $w(\hat{\mathbf{k}})$ is designed as a uniform function in the proposed method, $\mathbf{W}$ becomes the identity matrix and $\mathcal{J} = \boldsymbol{\alpha}^{\mathsf{H}}\boldsymbol{\alpha}$; therefore, "MM" is regarded as a special case of the proposed method.

5.3 Validation

5.3.1 Experimental conditions

Methods

We compared the noise attenuation performance of the proposed method with "MM."

(i) "MM" corresponding to the conventional mode-matching-based algorithm, which can be considered as the uniformly weighted case of the proposed method;

(ii) "Proposed" corresponding to the proposed method utilizing the DOA weighting. The details of the definition of the weighting function are described below.

Geometrical setup

The geometrical setup of the arrays is illustrated in Figure 5.3. Instead of using a fully surrounding spherical array geometry for the sensor and secondary source arrays, multiple-horizontal-ring arrays were introduced, which can be considered as a practical geometry. We considered two circular loudspeaker arrays placed at the elevations of $\theta_1^{(\mathrm{sec})} = 0$ and $\theta_2^{(\mathrm{sec})} = \pi/6$. The loudspeaker arrays consisting of 16 loudspeakers per ring were used to produce an antinoise sound field. The distance to each loudspeaker array was 0.4 m from the coordinate origin. The two horizontal ring-shaped RMAs, each with 16 microphones, were positioned at the same elevation as the loudspeaker arrays, 0.5 m from the coordinate origin. The residual noise field was recorded by the error microphones' HOMs. Each HOM consisted of four microphones uniformly placed on a sphere with a radius of 0.1 m, which captures up to first-order local SH coefficients. Eight HOMs were placed 0.2 m from the coordinate origin at the vertices of a cube. The simulations were carried out in a three-dimensional free field.

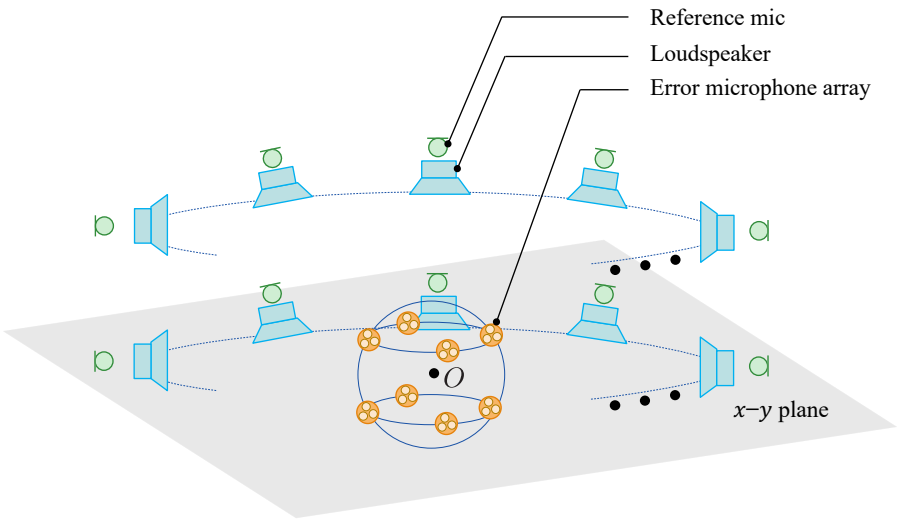


Figure 5.3: Geometrical setup of the arrays.

Noise signal

We placed a point source of white noise as the primary noise source. To evaluate the noise attenuation performance dependent on the direction of the primary source, two different heights of the primary source were simulated. Each error and the reference microphone output were subjected to white Gaussian noise with a signal-to-noise ratio of 40 dB.

Evaluation metric

The noise attenuation performance was evaluated by defining the regional noise reduction (RNR) inside the region as

$$\mathrm{RNR}(n) = 10 \log_{10} \frac{\sum_i |S^{(0)}(\mathbf{r}_i)|^2}{\sum_i |S^{(\iota)}(\mathbf{r}_i)|^2}, \tag{5.18}$$

where $S^{(\iota)}(\mathbf{r}_i)$ is the residual error at the i-th grid point in the evaluation region at the ι-th iteration and $S^{(0)}(\mathbf{r}_i)$ is the initial noise field at the i-th point. We defined the two evaluation regions S_1 and S_2 shown in Figure 5.4 as horizontal circular regions at heights of $z = 0$ m and $z = 0.05$ m, respectively. Each circular region was inscribed in a sphere with a radius of 0.2 m placed at the coordinate origin. There were 1876 and 1780 uniformly sampled grid points in S_1 and S_2, respectively.

5.3.2 Analysis of weighting matrix

We defined a binary weighting function as a simple case according to the loudspeaker array geometry, which only depends on the elevation θ, as

$$w(\hat{\mathbf{k}}) = \begin{cases} 1 & (\theta_1^{(w)} < \theta < \theta_2^{(w)}) \\ 0 & (\text{otherwise}). \end{cases} \tag{5.19}$$

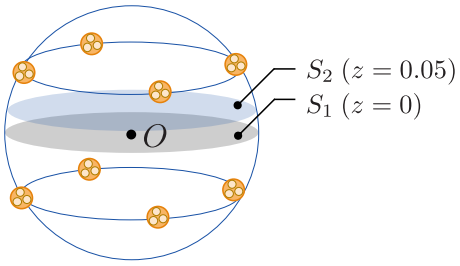


Figure 5.4: Definition of the evaluation regions.

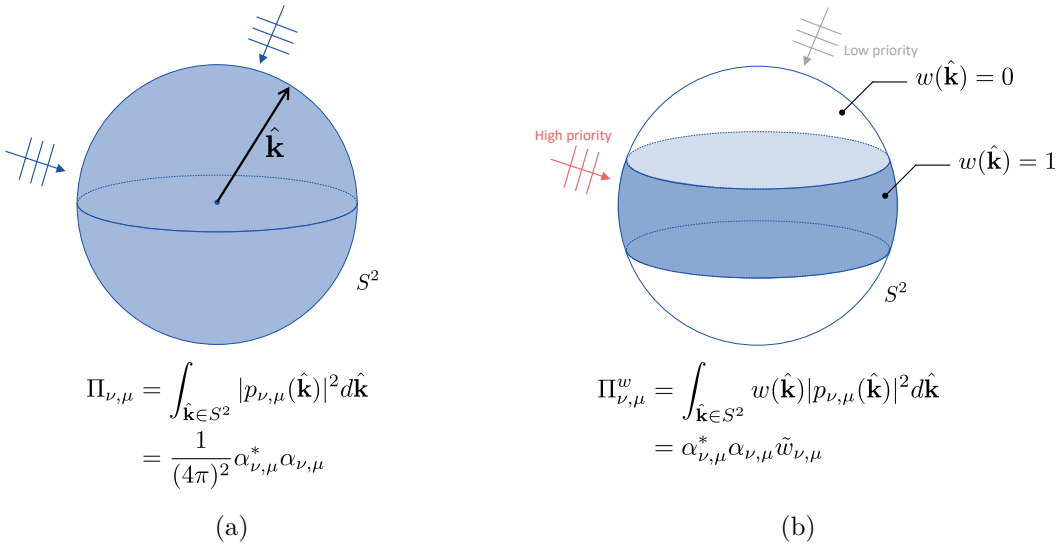


Figure 5.5: Comparison of weighting functions: (a) Uniform weighting. (b) Binary weighting.

Since the multiple horizontal rings of loudspeakers have a high resolution on the azimuth angle, the DOA nearly parallel to the x-y plane should be prioritized as the controllable direction. Therefore, we defined the weighting function with $\theta_1^{(w)} = 0$ and $\theta_2^{(w)} = \pi/12$. This weighting prioritizes the DOA nearly parallel to the x-y plane. The weighted power of the DOA distribution using the uniform weighting function and binary weighting function is illustrated in Figure 5.5.

Once the DOA weighting function is defined, the SH-domain weighting coefficient can be calculated by Eq. (5.13). We used the Gauss–Kronrod quadrature formula [90] to approximate the integral calculation because there is no analytical solution for the defined weighting function. Figure 5.6 depicts the power of the weighting coefficient for each order and degree. From the figure, we can see that the even orders, which make the main contribution to the sound field close to the x-y plane, have large weights. Moreover, the weights with $\nu = \mu$ have a larger value than the other even-order weights.

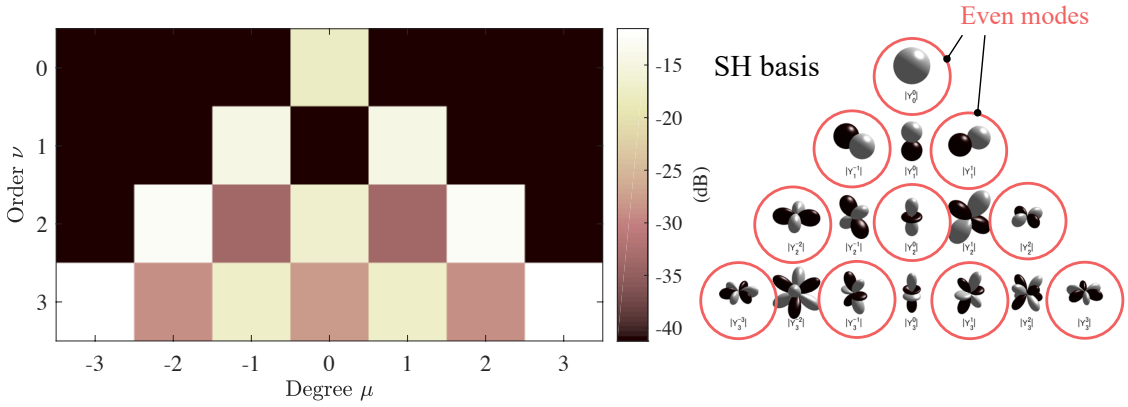


Figure 5.6: SH-domain weighting coefficients $\tilde{w}_{\nu,\mu}$.

5.3.3 Noise attenuation performance

We compared the proposed method with "MM" under the same array geometries as described in Section 5.3.1. Note that the truncation order was determined by $N = \lceil kr_{\text{tar}} \rceil$, where $r_{\text{tar}} = 0.2$ m is the radius of the target region.

The noise attenuation performance for the frequency range of 100 Hz to 1000 Hz is shown in Figure 5.7. The result was obtained by calculating the average of 10 random sequence patterns of noise. The primary noise source was placed at $(x, y, z) = (0.0, 3.0, 0.2)$ m. The primary source was placed near the x-y plane, as illustrated in Figure 5.8a. At all frequencies, we can see that the proposed method outperforms "MM." In S_1 and S_2, "MM" performed nearly identically, whereas the proposed method significantly reduced noise in S_1, indicating that the weighted cost function was appropriate for the array geometry.

Figure 5.9 shows the convergence performance at 500 Hz. Both "MM" and the proposed method showed similar convergence performance characteristics, whereas the proposed method gave a higher RNR than "MM." This implies that the DOA-weighted cost function Eq. (5.15), as written, is appropriate for the array geometry. The power distribution of the residual noise field on the x-y plane after 500 iterations of adaptive processing at 500 Hz is shown in Figure 5.10. The RNR in S_1 for "MM" was 11.75 dB, while the proposed method had an RNR of 17.52 dB.

5.3.4 Noise attenuation performance of higher source position

We moved the primary noise source to a higher position $(x, y, z) = (0.0, 1.0, 3.0)$ m as illustrated in Figure 5.8b to evaluate the robustness of the methods. Figure 5.11 shows the result. Although the proposed method's performance degraded rapidly above 600 Hz, we can see that the overall tendency was the same as when the noise source was close to the x-y plane. Figure 5.12 shows the power distribution of the residual noise field on the x-y plane ($z = 0$) and the higher plane ($z = 0.05$). We can

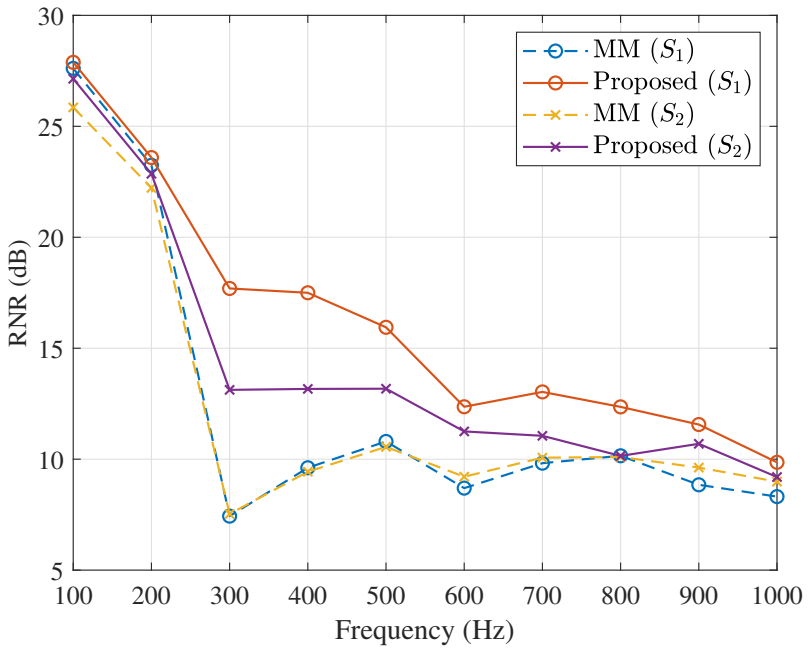


Figure 5.7: Regional noise reduction inside the target region for various frequencies.

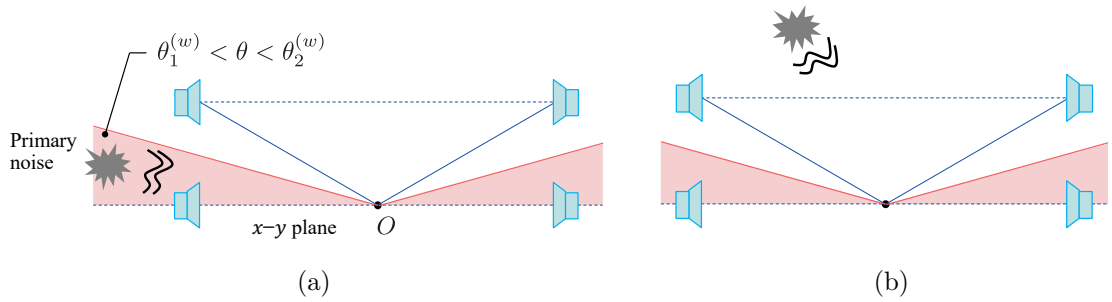


Figure 5.8: Positional relationship between the primary source and loudspeaker array: (a) Primary source placed close to the x-y plane. (b) Primary source placed at the higher position.

see that on the $z = 0$ plane, the residual noise of "Proposed" was lower than "MM." By contrast, the residual noise on the $z = 0.05$ plane remained for both methods. This indicates that the noise signal impinging from the higher elevation angle was difficult to attenuate using the array structure used in the simulation. Although the $z = 0$ plane was also affected by this difficult situation in "MM," "Proposed" managed to attenuate the plane to some extent.

Regarding the practical environment, sound reflections from the ceiling and floor may affect the performance. Placing sound-absorbing materials on the ceiling and floor, for example, could reduce the noise field impinging from directions where the noise field is difficult to attenuate. As a result, performance degradation at high

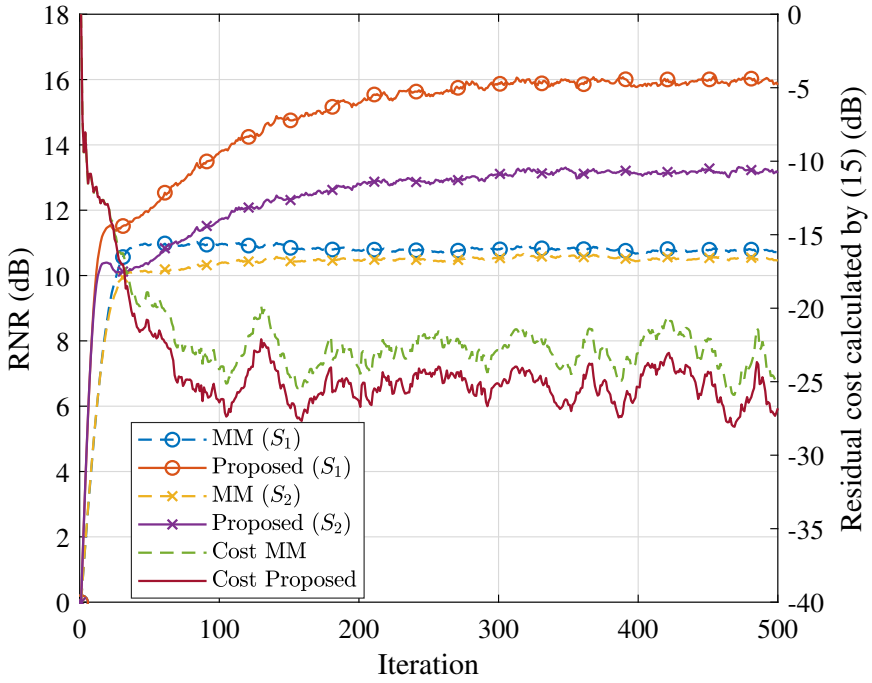


Figure 5.9: Regional noise reduction at 500 Hz for each iteration index.

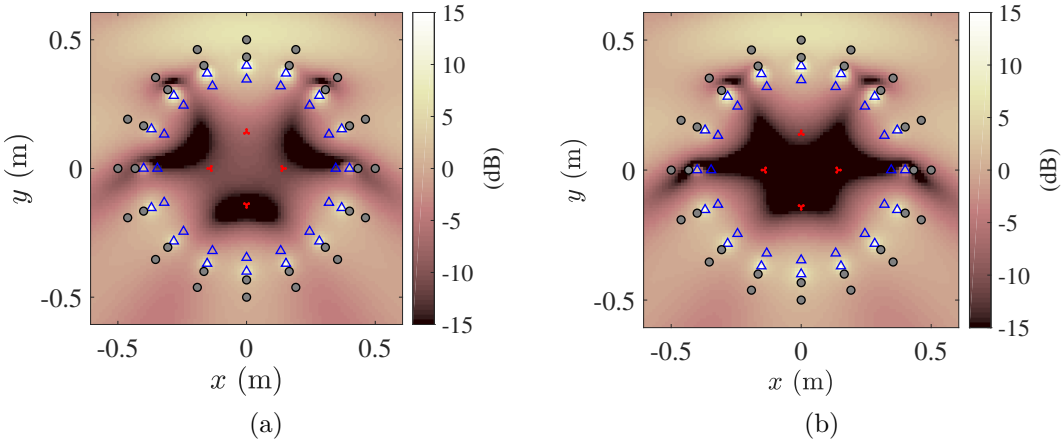


Figure 5.10: Residual noise field after 500 iterations of FXLMS algorithm: (a) "MM." (b) "Proposed." The reference microphones, secondary loudspeakers, and error microphones are represented by gray dots, blue triangles, and red dots, respectively. It should be noted that all microphones and loudspeakers are projected to the x-y plane.

frequencies would be reduced.

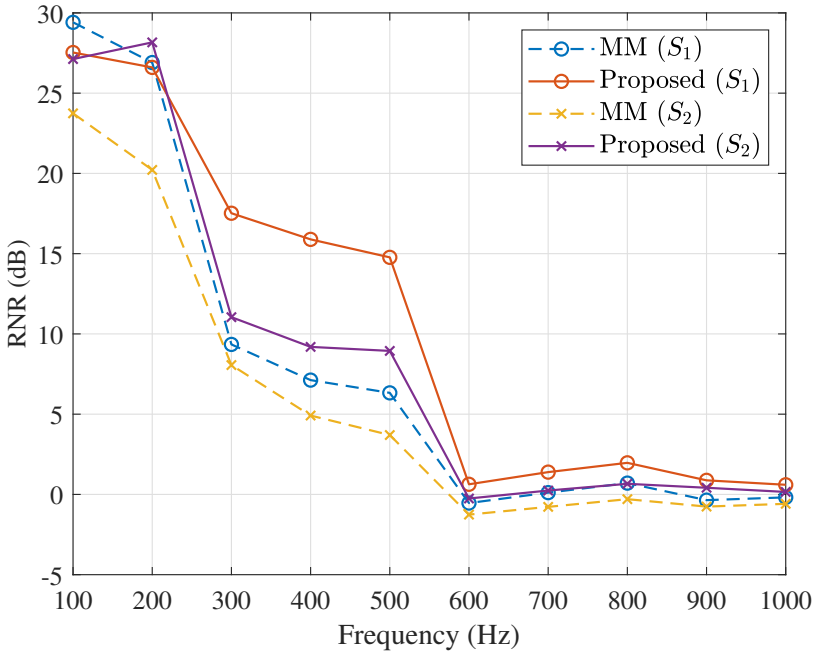


Figure 5.11: Regional noise reduction inside the target region for various frequencies.

5.4 Summary

In this chapter, we proposed the SH-domain ANC algorithm utilizing the DOA weighting, which controls the noise field attenuation in terms of direction. Traditionally, the direction of noise propagation was treated uniformly. To put it another way, noise components impinging from any direction were designed to be attenuated. As a result, spherical arrays of microphones and loudspeakers with uniform distribution were required. However, this limits the system's practicability. To tackle the problem, we investigated the approach to give the DOA weighting in the LMS procedure according to the array geometry, which leads the system to a reasonable solution.

In the simulation, the proposed method outperformed the conventional "MM"-based algorithm using a nonuniform array geometry, for which it can be considered difficult to control coefficients of all orders. Because of the lack of microphones and loudspeakers in the vertical direction, the noise from the higher position was difficult to attenuate by the circular arrays. Although the conventional method degrades significantly in this condition, the proposed method maintains its attenuation performance. This can be considered that the LMS update was not too much affected by the vertical components of the noise field, which will be an unstable solution seeking due to less number of microphones and loudspeakers.

The DOA weighting function can be intuitively defined based on the secondary source array geometry without any prior information of the primary noise field. One of the future projects will be to investigate the reasonable weighting function.

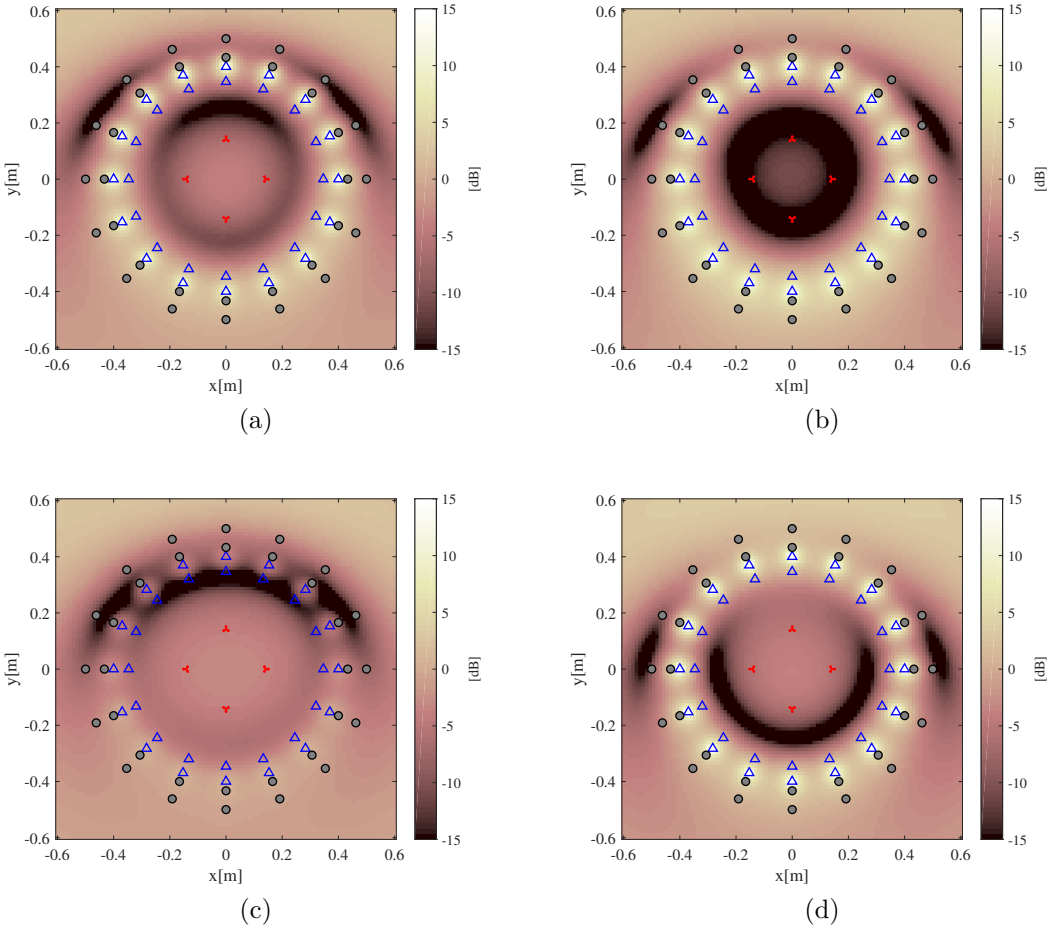


Figure 5.12: Residual noise field after 500 iterations of FXLMS algorithm: (a) "MM" ($z = 0$). (b) "Proposed" ($z = 0$). (c) "MM" ($z = 0.05$). (d) "Proposed" ($z = 0.05$).

A closed-form function can also be defined as the weighting function, allowing the numerical integration to be replaced by the analytical one, resulting in less numerical error.

The study in this chapter showed the validity of the use of non-spherical arrays. Compared to Chapter 4, the array geometry is not restricted to the circular shape, therefore there is more degree of freedom in designing the array structure.

Chapter 6

Spatial active noise control using sparse decomposition of reference signals from distributed sensor arrays

We previously introduced the CS approach to reduce the number of reference microphones and showed that noise can be attenuated by using several microphones that is smaller than a number required by the spatial Nyquist theory [78]. Because CS is based on the assumption that the columns of a sensing matrix are mutually incoherent, the reproduction accuracy at low frequencies is limited [91]. However, at low frequencies, they are generally highly coherent [66, 92, 93]. To overcome this problem, CS has been applied to SH coefficients [94, 95].

In this chapter, we focus on reference signal reconstruction and exploit its efficient representation by utilizing the inherent spatial sparseness of the noise field. As previously stated, RMA takes up a lot of space and limits system performance due to spatial aliasing; thus, we investigate sparse recovery of reference signals to improve both system viability and performance. The flexibility of the RMA structure and placement is achieved by combining the sparse recovery with the translation of the field [89] based on the addition theorem [96]. The sensing matrix is transformed into orthonormal space, which improves low frequency reconstruction accuracy. Furthermore, using CS to represent signals compactly improves performance at frequencies higher than the spatial Nyquist frequency. Another fundamental issue with harmonic expansion-based sound field representation is that outgoing and incoming sound fields must be considered separately [36]. This is due to the harmonic expansion being based on a complete set of solutions to the homogeneous Helmholtz equation, which does not allow any sound sources to existing inside a closed region when an incoming sound field is considered. 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. We can define target noise sources as

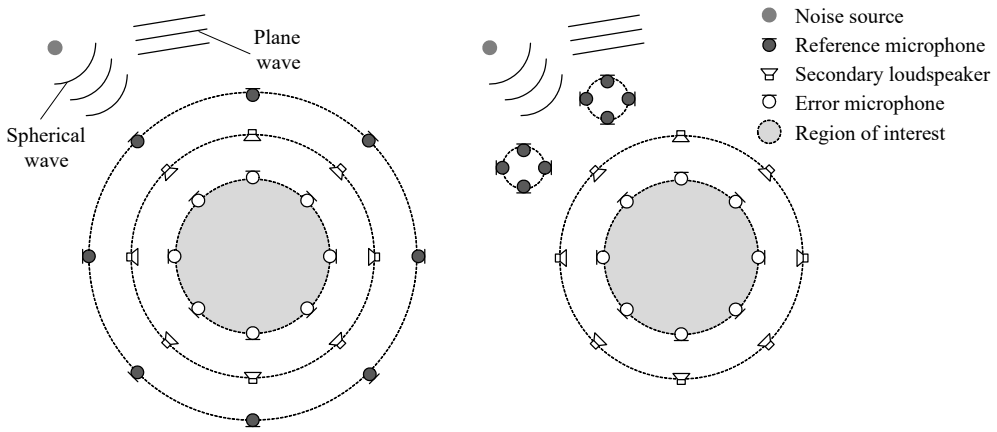


Figure 6.1: System model of the interior sound field constructed using different RMA structures: (a) ROI fully surrounded by RMA. (b) Compactly distributed RMAs.

those that are expected to be attenuated and nontarget noise sources as those that are expected to be ignored based on the distribution. Nontarget sources, such as those found within the ROI, can be eliminated from the ANC processing, resulting in increased system robustness.

The main contributions in this chapter are as follows: (1) Generalization of the SH-domain reference signal representation utilizing the translation of the field, 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 compensating the spatial noise source distribution of both the interior and exterior sound fields. To summarize, our proposed method allows for greater flexibility in reference microphone placement, which increases system viability. Furthermore, for the aforementioned problem associated with the composite noise field of outgoing and incoming sound fields, system robustness is increased.

6.1 Problem statement

An interior noise field attenuation in a sizable 3D space is usually achieved by surrounding an ROI with a large number of sensors and loudspeakers. As shown in Figure 6.1a, one of the best 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 surround the ROI. Because the reference signals of a noise field must be captured before the noise wave front arrives at the ROI, the RMA must be placed in the farthest location to satisfy the causality constraint. Therefore, the RMA becomes relatively large, which may be difficult to construct in practice. Furthermore, to avoid spatial aliasing, a large

number of microphones are required to fill the gap between the microphones. The basic idea in dealing with these issues is based on the assumption of a practical situation, namely that the positions of the noise sources 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 Figure 6.1b. Hereafter, the center points of distributed RMAs are referred to as local origins.

Now, let us consider an arbitrary noise field $S(\mathbf{r})$ in the free field at position $\mathbf{r}$ within the ROI, Ω. In contrast to the homogeneous Helmholtz equation Eq. (2.2), $S(\mathbf{r})$ satisfies the inhomogeneous Helmholtz equation [97, 98] when we assume that the sound sources exist inside the region:

$$(\nabla^2 + k^2)S(\mathbf{r}) = -\rho_{\mathrm{int}}^{(\mathrm{ps})}(\mathbf{r}), \tag{6.1}$$

where $\rho_{\mathrm{int}}^{(\mathrm{ps})}(\mathbf{r})$ is the distribution of point sources inside the region Ω. By approximating $\rho_{\mathrm{int}}^{(\mathrm{ps})}(\mathbf{r})$ as a linear combination of $L_{\mathrm{ps,int}}$ point sources, $S(\mathbf{r})$ can be represented as a sum of the particular and homogeneous solutions [29, 99]:

$$S(\mathbf{r}) = S_{\mathrm{P}}(\mathbf{r}) + S_{\mathrm{H}}(\mathbf{r}). \tag{6.2}$$

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

$$S_{\mathrm{P}}(\mathbf{r}) = \int_{\mathbf{r}'\in\Omega} \rho_{\mathrm{int}}^{(\mathrm{ps})}(\mathbf{r}')G(\mathbf{r}|\mathbf{r}')d\mathbf{r}' \approx \sum_{\ell=1}^{L_{\mathrm{ps,int}}} \gamma_{\ell,\mathrm{int}}^{(\mathrm{ps})} G(\mathbf{r}|\mathbf{r}'_{\ell,\mathrm{int}}), \tag{6.3}$$

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

$$S_{\mathrm{H}}(\mathbf{r}) = \int_{\hat{\mathbf{r}}\in\mathbb{S}^2} \rho^{(\mathrm{pw})}(\hat{\mathbf{r}})e^{ik\hat{\mathbf{r}}\cdot\mathbf{r}}d\hat{\mathbf{r}}, \tag{6.4}$$

where $\mathbb{S}^2$ is the unitary sphere, $\rho^{(\mathrm{pw})}(\hat{\mathbf{r}})$ is the distribution of the plane waves, and $\hat{\mathbf{r}}$ is a unit vector toward the direction of the plane wave. We also consider the solution's approximation by discretizing the direction of plane waves. In the plane wave domain, a sound field composed of far-field sources can be decomposed into a sparse set of weights. A sound field composed of near-field sources, on the other hand, can be expressed by the superposition of a few point sources. The combination of these two types of sound sources enables a sparse signal representation of diverse sound fields. Therefore, we introduce the model of $S_{\mathrm{H}}(\mathbf{r})$ as a sum of the plane waves and point sources:

$$S_{\mathrm{H}}(\mathbf{r}) \approx \sum_{\ell=1}^{L_{\mathrm{pw}}} \gamma_{\ell}^{(\mathrm{pw})} e^{ik\hat{\mathbf{r}}_{\ell}\cdot\mathbf{r}} + \sum_{\ell=1}^{L_{\mathrm{ps,ext}}} \gamma_{\ell,\mathrm{ext}}^{(\mathrm{ps})} G(\mathbf{r}|\mathbf{r}'_{\ell,\mathrm{ext}}), \tag{6.5}$$

where $\gamma_\ell^{(\mathrm{pw})}$ and $\gamma_{\ell,\mathrm{ext}}^{(\mathrm{ps})}$ are the complex amplitudes of the plane waves and point sources, respectively, which exist outside Ω, L_{pw} and $L_{\mathrm{ps,ext}}$ is the number of the plane waves and point sources, respectively, and $\mathbf{r}'_{\ell,\mathrm{ext}}$ is the source position outside Ω. From Eqs. (6.3) and (6.5), we describe the solution to Eq. (6.1) as

$$S(\mathbf{r}) \approx \sum_{\ell=1}^{L_{\mathrm{pw}}} \gamma_\ell^{(\mathrm{pw})} e^{ik\hat{\mathbf{r}}_\ell \cdot \mathbf{r}} + \sum_{\ell=1}^{L_{\mathrm{ps,ext}}} \gamma_{\ell,\mathrm{ext}}^{(\mathrm{ps})} G(\mathbf{r}|\mathbf{r}'_{\ell,\mathrm{ext}}) + \sum_{\ell=1}^{L_{\mathrm{ps,int}}} \gamma_{\ell,\mathrm{int}}^{(\mathrm{ps})} G(\mathbf{r}|\mathbf{r}'_{\ell,\mathrm{int}}). \tag{6.6}$$

Now, we assume that L_{mic} microphones are placed inside Ω to capture the noise field. We discretize the integration into a sum of $L = L_{\mathrm{pw}} + L_{\mathrm{ps,ext}} + L_{\mathrm{ps,int}}$ defined grid points as shown in Eq. (6.6), and we derive the total noise field captured at the microphone position $\mathbf{S} = [S(\mathbf{r}_1), \ldots, S(\mathbf{r}_{L_{\mathrm{mic}}})]^{\mathsf{T}}$ as a sum of finite point sources and plane waves:

$$\mathbf{S} = \mathbf{H}\boldsymbol{\gamma}, \tag{6.7}$$

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, Eq. (6.7) is set to be an underdetermined problem and the matrix $\mathbf{H}$ is the so-called over-complete dictionary, which usually has a dense resolution of the spatial discretization. The total primary noise field is described by the noise distribution $\boldsymbol{\gamma}$. This work's problem model is similar to that of [100]. The distinction is in the definition of the sensing matrix. An acoustic transfer function in the spatial (temporal frequency) domain is defined as a sensing matrix in [100]. In this work, however, a sensing matrix is defined in the SH-domain. This gives an 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 Sections 6.3.2 and 6.5.3.

The goal of this study is to precisely estimate the reference signals of the primary noise field from the measurements of RMAs. Extrapolation of the noise field is required in the array configuration shown in Figure 6.1b because the RMAs are placed at local origins that are at different distances from the global origin. To reproduce the same sound field in theory, at least the same number of microphones are required in local origins as in global origins, according to the literature [101,102]. 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 decompose the noise field into a sparse set of basis functions using an SH-domain representation of the sensing matrix and measurements in the CS framework.

Figure 6.2 shows a block diagram of the proposed method, which 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 [103,104]. Although, these methods can also be applied to

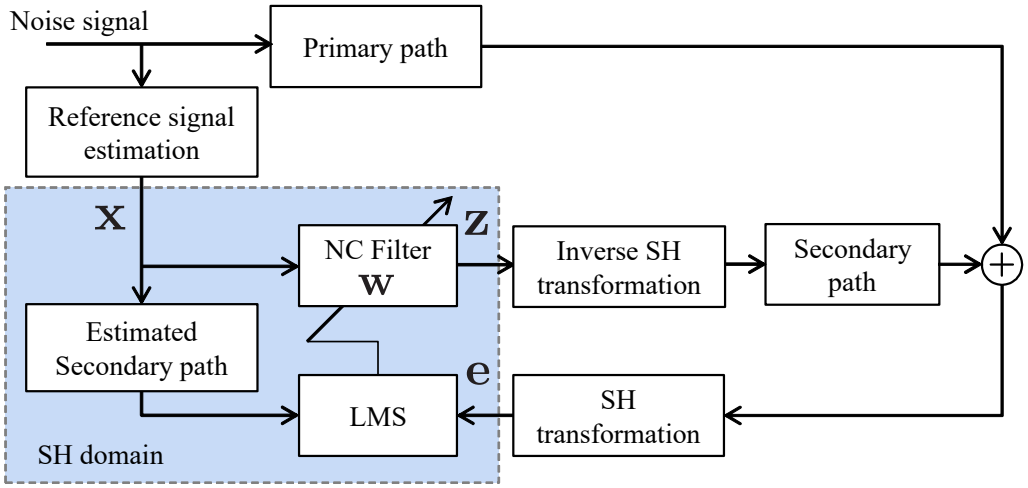


Figure 6.2: Block diagram of the proposed method.

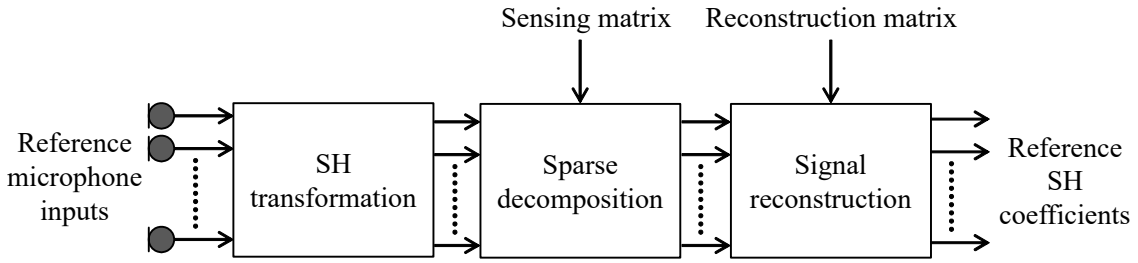


Figure 6.3: Block diagram of the reference signal estimation part.

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. Figure 6.3 depicts the flow of the estimation process in detail. In the decomposition stage, 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, followed by the reconstruction stage. Finally, the reconstructed reference signal is input to the SH-domain FXLMS algorithm.

6.2 Spherical-harmonic-domain ANC

The basic formulation of the sound field representation and FXLMS in SH-domain is described in Chapter 2. In this section, the detailed formulation of SH-domain representation utilizing the CS approach is explained.

6.2.1 Spherical-harmonic-domain adaptive filter update

As written in Eq. (2.25), the cost function of the ANC system is evaluated by calculating the squared residual signals in SH-domain. The SH coefficient e_{nm} can be calculated by measuring the sound field on a boundary of a sphere:

$$e_{\nu,\mu} = \frac{1}{j_\nu(kr_q)} \int_0^{2\pi} \int_0^{\pi} S(\mathbf{r}) Y^*_{\nu,\mu}(\theta,\phi) \sin\theta d\theta d\phi, \tag{6.8}$$

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

$$\int_0^{2\pi} \int_0^{\pi} Y_{\nu,\mu}(\theta,\phi) Y^*_{\nu',\mu'}(\theta,\phi) \sin\theta d\theta d\phi = \delta_{\nu,\nu'}\delta_{\mu,\mu'}, \tag{6.9}$$

where $\delta_{\nu,\nu'}$ denotes the Kronecker delta function. The numerical instability caused by the spherical Bessel function in the denominator in Eq. (6.8) is known as the Bessel zero problem or forbidden frequency problem [89]. A rigid baffle array is commonly used in practice to avoid this problem. Instead of using the spherical Bessel function in (6.8), a different radial function for a rigid baffle array, $b_\nu(kr_q)$, is introduced:

$$b_\nu(kr_q) = j_\nu(kr_q) - \frac{j'_\nu(kr_q)}{h^{(1)'}_\nu(kr_q)} h^{(1)}_\nu(kr_q), \tag{6.10}$$

where $h^{(1)}_\nu(\cdot)$ is the first kind of the ν-th order spherical Hankel function, and $h^{(1)'}_\nu(\cdot)$ and $j'_\nu(\cdot)$ are the derivative of $h^{(1)}_\nu(\cdot)$ and $j_\nu(\cdot)$, respectively. The coefficient $e_{\nu,\mu}$ can be calculated as

$$e_{\nu,\mu} = \frac{1}{b_\nu(kr_q)} \int_0^{2\pi} \int_0^{\pi} S(\mathbf{r}) Y^*_{\nu,\mu}(\theta,\phi) \sin\theta d\theta d\phi. \tag{6.11}$$

Alternatively, one can use either a dual-radius open spherical array [105], multiple rigid spherical arrays [71], or multiple flat arrays [106] 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 [89]:

$$e_{\nu,\mu} \approx \frac{1}{b_\nu(kr_q)} \sum_{a=1}^{A} \Lambda_a S_e(\mathbf{r}_a) Y^*_{\nu,\mu}(\theta_a,\phi_a), \tag{6.12}$$

where A is the total number of the sampling points and Λ_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 $\mathbf{z} = \mathbf{gxw}$, as well as to update the adaptive filter $\mathbf{w}$ using Eq. (2.27). As a result, accurate estimation of the reference SH coefficients is critical for both noise attenuation and filter convergence performance. Sections 6.3 and 6.4 describe the method for estimating the reference SH coefficients.

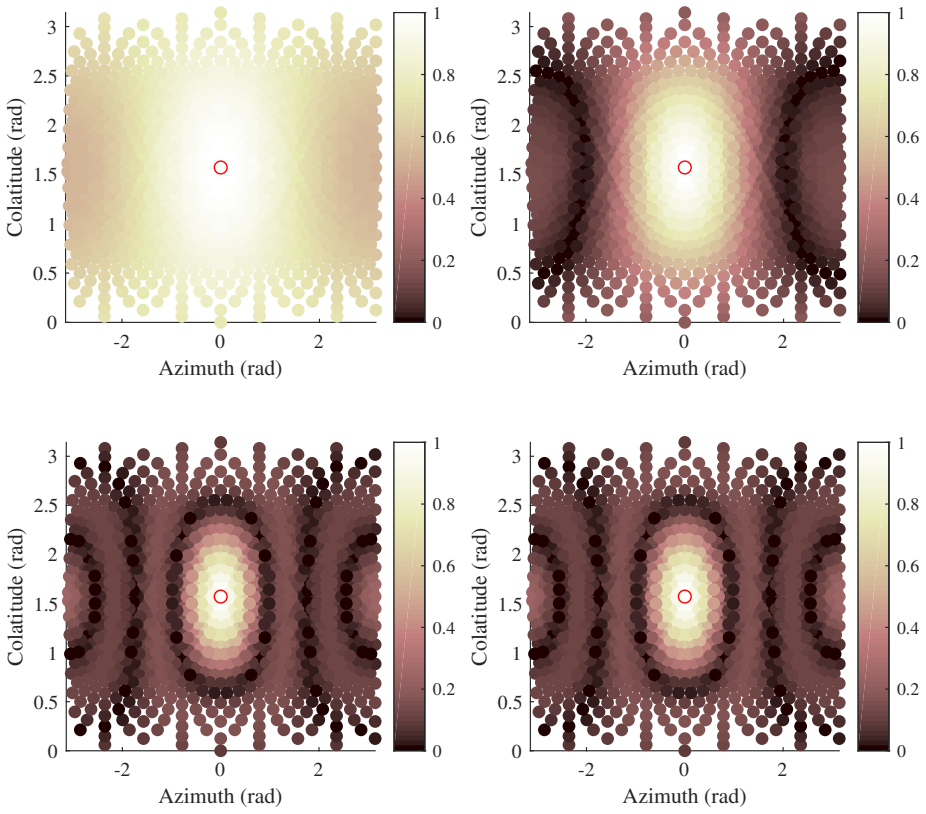


Figure 6.4: Column coherence between a given direction $(\theta_1, \phi_1) = (\pi/2, 0)$, which is shown as a red solid circle, to other directions of the plane wave sensing matrix: (a) 500 Hz. (b) 1000 Hz. Each dot represents a single plane wave direction as well as a point source position. The first row represents the result of the spatial-domain sensing matrix, and the second row represents the result of the SH-domain sensing matrix.

6.3 Sparse decomposition of noise field

6.3.1 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 assume that all noise sources exist outside of the ROI in this case. As described in (6.5), an arbitrary noise field can be expressed as a sum of plane waves and point sources that exist outside the ROI. An SH-domain representation can be derived by applying the SH expansion of a plane wave [38],

$$e^{ik\hat{\mathbf{r}}\cdot\mathbf{r}} = \sum_{\nu=0}^{\infty}\sum_{\mu=-\nu}^{\nu} \underbrace{4\pi i^{-\nu} Y_{\nu,\mu}^{*}(\theta_s, \phi_s)}_{\psi_{\nu,\mu}^{(\mathrm{pw})}(\hat{\mathbf{r}})} j_\nu(kr) Y_{\nu,\mu}(\theta, \phi), \tag{6.13}$$

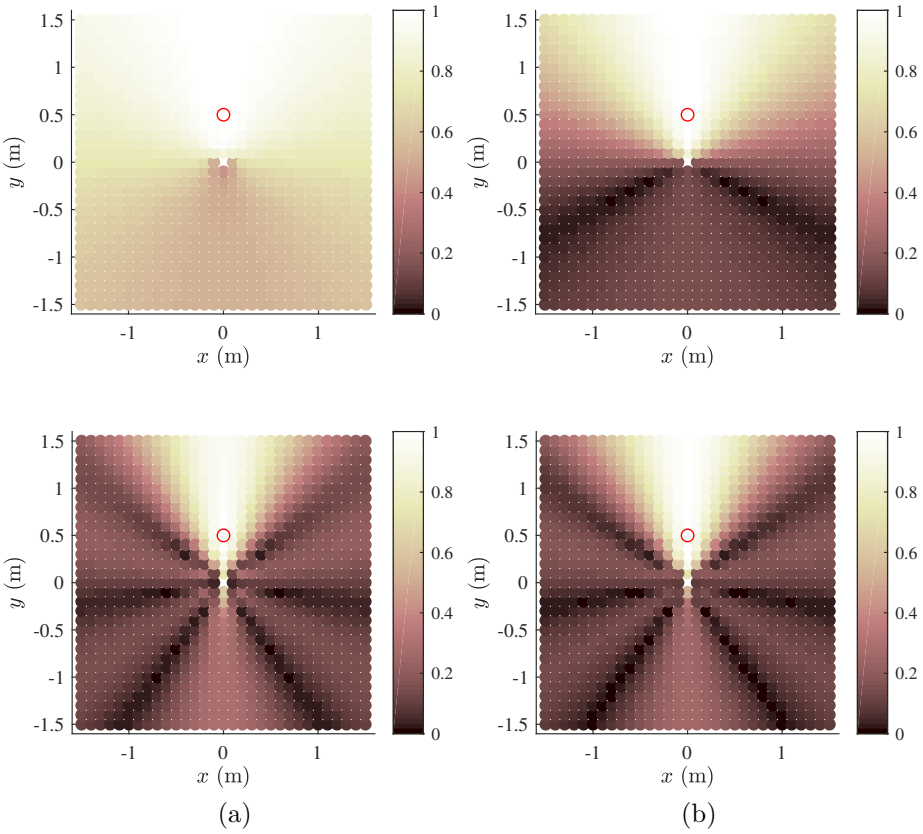


Figure 6.5: Column coherence between a given point $(x_1, y_1) = (0, 0.5)$, which is shown as a red solid circle, to other points of the point source sensing matrix: (a) 500 Hz. (b) 1000 Hz. Each dot represents a single plane wave direction as well as a point source position. The first row represents the result of the spatial-domain sensing matrix, while the second row represents the result of the SH-domain sensing matrix.

and that of a point source [36],

$$G(\mathbf{r}|\mathbf{r}') = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \underbrace{-ikh_\nu^{(1)}(kr_s)Y_{\nu,\mu}^*(\theta_s, \phi_s)}_{\psi_{\nu,\mu}^{(\mathrm{ps})}(\mathbf{r}')} \times j_n(kr)Y_{nm}(\theta, \phi), \tag{6.14}$$

where (r_s, θ_s, ϕ_s) is the source position and $\psi_{\nu,\mu}^{(\mathrm{pw})}(\hat{\mathbf{r}})$ and $\psi_{\nu,\mu}^{(\mathrm{ps})}(\mathbf{r}')$ is the SH coefficients of the plane wave and point source, respectively. From Eqs. (2.4), (6.5), (6.13), and (6.14), an SH-domain representation of the noise field can be described as

$$\alpha_{\nu,\mu} \approx \sum_{\ell=1}^{L_{\mathrm{pw}}} \gamma_\ell^{(\mathrm{pw})} \psi_{\nu,\mu,\ell}^{(\mathrm{pw})} + \sum_{\ell=1}^{L_{\mathrm{ps}}} \gamma_\ell^{(\mathrm{ps})} \psi_{\nu,\mu,\ell}^{(\mathrm{ps})}, \tag{6.15}$$

where L_{pw} and L_{ps} are the numbers of the grid directions and points of the plane wave and point source, respectively. The matrix form of Eq. (6.15) can be derived

by truncating the order of SH expansion in a sufficient order, which can be written as

$$\boldsymbol{\alpha} = \boldsymbol{\psi}\boldsymbol{\gamma}, \tag{6.16}$$

where

$$\begin{aligned} \boldsymbol{\psi} &= [\boldsymbol{\psi}^{(\mathrm{pw})}\boldsymbol{\psi}^{(\mathrm{ps})}] \\ &= \begin{bmatrix} \psi^{(\mathrm{pw})}_{0,0,1} & \cdots & \psi^{(\mathrm{pw})}_{0,0,L_{\mathrm{pw}}} & \psi^{(\mathrm{ps})}_{0,0,1} & \cdots & \psi^{(\mathrm{ps})}_{0,0,L_{\mathrm{ps}}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \psi^{(\mathrm{pw})}_{N,N,1} & \cdots & \psi^{(\mathrm{pw})}_{N,N,L_{\mathrm{pw}}} & \psi^{(\mathrm{ps})}_{N,N,1} & \cdots & \psi^{(\mathrm{ps})}_{N,N,L_{\mathrm{ps}}} \end{bmatrix} \end{aligned} \tag{6.17}$$

and

$$\boldsymbol{\gamma} = [\gamma^{(\mathrm{pw})}_1, \ldots, \gamma^{(\mathrm{pw})}_{L_{\mathrm{pw}}}, \gamma^{(\mathrm{ps})}_1, \ldots, \gamma^{(\mathrm{ps})}_{L_{\mathrm{ps}}}]^{\mathsf{T}}. \tag{6.18}$$

One way of finding the optimal solution of Eq. (6.16) is to assume the spatial sparsity of $\boldsymbol{\gamma}$ and solve the following minimization problem:

$$\underset{\boldsymbol{\gamma}}{\text{minimize}} \, \|\boldsymbol{\alpha} - \boldsymbol{\psi}\boldsymbol{\gamma}\|_2^2 + \lambda \|\boldsymbol{\gamma}\|_p^p, \tag{6.19}$$

where the operator $\|\cdot\|_p (0 < p \leq 1)$ indicates the ℓ_p-norm. The first term of Eq. (6.19) represents the approximation error and the second term is introduced to induce a sparse solution [68], where the parameter λ balances both terms. There are several algorithms for solving the problem [69, 107–109]. Although the sensing matrix is described in the SH-domain, the sparsity constraint is still imposed on the spatial source distribution $\boldsymbol{\gamma}$.

6.3.2 Coherence of sensing matrix

The performance of the CS in (6.19) highly depends on the coherence of the sensing matrix. One way to measure the coherence is to calculate the Gram matrix $\boldsymbol{\Gamma}$ [66]:

$$\boldsymbol{\Gamma}_{ij} = \frac{\|\mathbf{h}_i^{\mathsf{H}}\mathbf{h}_j\|_2}{\|\mathbf{h}_i\|_2\|\mathbf{h}_j\|_2}, \tag{6.20}$$

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

Sensing matrices for the plane waves can be described as

$$\mathbf{E} = \begin{bmatrix} e^{ik\hat{\mathbf{r}}_1 \cdot \mathbf{r}_1} & \cdots & e^{ik\hat{\mathbf{r}}_{L_{\mathrm{pw}}} \cdot \mathbf{r}_1} \\ \vdots & \ddots & \vdots \\ e^{ik\hat{\mathbf{r}}_1 \cdot \mathbf{r}_{L_{\mathrm{mic}}}} & \cdots & e^{ik\hat{\mathbf{r}}_{L_{\mathrm{pw}}} \cdot \mathbf{r}_{L_{\mathrm{mic}}}} \end{bmatrix} \tag{6.21}$$

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

$$\mathbf{G} = \begin{bmatrix} G(\mathbf{r}_1|\mathbf{r}'_1) & \cdots & G(\mathbf{r}_1|\mathbf{r}'_{L_{\mathrm{ps}}}) \\ \vdots & \ddots & \vdots \\ G(\mathbf{r}_{L_{\mathrm{mic}}}|\mathbf{r}'_1) & \cdots & G(\mathbf{r}_{L_{\mathrm{mic}}}|\mathbf{r}'_{L_{\mathrm{ps}}}) \end{bmatrix} \tag{6.22}$$

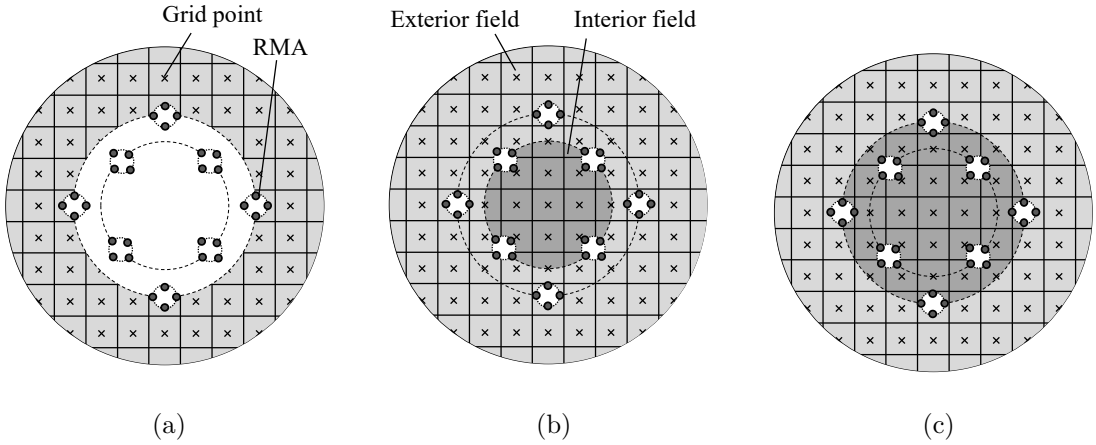


Figure 6.6: Array structure and grid point configuration: (a) Grid points of the exterior field. (b) Grid points of both interior and exterior sound fields for the inner RMAs. The light and dark regions correspond to the exterior and interior fields, respectively. (c) Grid points of both interior and exterior sound fields for the outer RMAs.

in the spatial-domain and $\boldsymbol{\psi}^{(\mathrm{ps})}$ in the SH-domain, respectively. By plotting a column of the Gram matrix $\boldsymbol{\Gamma}_{i1}$, the direction- and position-dependent coherence can be visualized for the sensing matrices $\mathbf{E}$, $\boldsymbol{\psi}^{(\mathrm{pw})}$, $\mathbf{G}$, and $\boldsymbol{\psi}^{(\mathrm{ps})}$. In Figures 6.4a and 6.4b, $\boldsymbol{\Gamma}_{i1}$ indicates the correlation between one direction $(\theta_1, \phi_1) = (\pi/2, 0)$ and all other possible directions (θ_i, ϕ_i), while in Figures 6.5a and 6.5b, $\boldsymbol{\Gamma}_{i1}$ 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 [110]. 960 grid points are uniformly placed in a 3×3 m region for point sources. The measurement points are sampled uniformly on the surface of a 0.1-m-radius sphere. The points' positions correspond to the vertexes of a dodecahedron's 20 faces. The figure shows that the source distributions of closely located two point sources, as well as two plane waves, are difficult to estimate separately using the spatial-domain sensing matrix, particularly at low frequencies, because the majority of the matrix's columns are highly correlated. In contrast, the SH-domain sensing matrix shows a high correlation only at relatively close directions and positions at around (θ_1, ϕ_1) and (x_1, y_1). Also, note that the SH-domain sensing matrix based on the plane wave is frequency invariant.

6.3.3 Translation of interior spherical harmonic coefficients

One straightforward way to capture the interior SH coefficient α_{nm} of the primary noise field in Eq. (2.4) from observations of the noise field $S(\mathbf{r})$ is to place an RMA to the coordinate origin. In this setup, the ROI is surrounded by an RMA as shown in Figure 6.1a. However, it is difficult to implement distribution of reference microphones to sample the region's surface. Alternatively, we can apply the concept

of distributed 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 [71]. By clustering the microphones, we can reduce the total number of units instead of uniformly distributing many microphones to enclose the entire ROI. Figure 6.6a depicts the array structure as well as 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 Figure 6.6a. 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. To capture SH coefficients, spherical arrays or any other 3D array structure mentioned in Section 6.2.1 with a sufficiently large aperture can be used. In this method, the translation relationship of SH coefficients among different expansion origins is utilized to estimate the global SH coefficients by placing HOMs to the local origin O_q:

$$S_{\mathrm{H}}^{(q)}(\mathbf{r}) = \sum_{v=0}^{\infty} \sum_{u=-v}^{v} \alpha_{v,u}^{(q)} j_v(kr') Y_{v,u}(\theta', \phi'), \tag{6.23}$$

where

$$\alpha_{v,u}^{(q)} = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \alpha_{\nu,\mu} T_{\nu,v}^{\mu,u}(\mathbf{r}_q). \tag{6.24}$$

$T_{\nu,v}^{\mu,u}(\mathbf{r}_q)$ is the translation coupling coefficient [38, 96] between the local and global interior SH coefficient, where $\mathbf{r}_q = (R_q, \theta_q, \phi_q)$ is the position of the q-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:

$$\boldsymbol{\alpha}_{\mathrm{local}} = \boldsymbol{\tau}\boldsymbol{\psi}\boldsymbol{\gamma}, \tag{6.25}$$

where

$$\boldsymbol{\alpha}_{\mathrm{local}} = [\boldsymbol{\alpha}_{\mathrm{local}}^{(1)\mathsf{T}}, \ldots, \boldsymbol{\alpha}_{\mathrm{local}}^{(Q)\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{Q(V+1)^2} \tag{6.26}$$

and

$$\boldsymbol{\tau} = [\mathbf{T}^{(1)\mathsf{T}}, \ldots, \mathbf{T}^{(Q)\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{Q(V+1)^2 \times (N+1)^2}. \tag{6.27}$$

$\boldsymbol{\alpha}_{\mathrm{local}}$ and $\boldsymbol{\tau}$ consist of Q vectors and matrices, whose q-th block elements are $\boldsymbol{\alpha}_{\mathrm{local}}^{(q)} = [\alpha_{0,0}^{(q)}, \ldots, \alpha_{V,V}^{(q)}]^{\mathsf{T}}$ and

$$\mathbf{T}^{(q)} = \begin{bmatrix} T_{0,0}^{0,0}(\mathbf{r}_q) & \cdots & T_{N,0}^{N,0}(\mathbf{r}_q) \\ \vdots & \ddots & \vdots \\ T_{0,V}^{0,V}(\mathbf{r}_q) & \cdots & T_{N,V}^{N,V}(\mathbf{r}_q) \end{bmatrix} \in \mathbb{C}^{(V+1)^2 \times (N+1)^2}. \tag{6.28}$$

The truncation order of the global coefficient N is determined 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 $(N+1)^2 \geq Q(V+1)^2$. As shown in Figure 6.6a, 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 $\alpha_{u,v}^{(q)}$ can be calculated by measuring the sound pressure on a boundary of a sphere to the local origin O_q:

$$\alpha_{u,v}^{(q)} \approx \frac{1}{b_\nu(k\bar{r}_q)} \sum_{a=1}^{A} \Lambda_a S_r^{(q)}(\mathbf{r}_a) Y_{\nu,\mu}^*(\theta_a, \phi_a), \tag{6.29}$$

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

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

$$\underset{\boldsymbol{\gamma}}{\text{minimize}} \, \|\boldsymbol{\alpha}_{\text{local}} - \boldsymbol{\tau}\boldsymbol{\psi}\boldsymbol{\gamma}\|_2^2 + \lambda\|\boldsymbol{\gamma}\|_p^p. \tag{6.30}$$

6.3.4 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. The ROI is assumed to be free of any sound sources in the interior problem. When this assumption is violated (i.e., sound sources exist in the region), sound field estimation fails. To improve the robustness of the ANC system, we devised method for taking into account both an interior and an exterior sound field. The considered array structure and the grid points are illustrated in Figures 6.6b and 6.6c. It is worth noting that the boundary of the interior and exterior fields can be defined in a variety of ways, and it is determined by the local origin on which the focus is placed.

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

$$S_{\text{P}}(\mathbf{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \breve{\beta}_{\nu,\mu} h_\nu^{(1)}(kr) Y_{\nu,\mu}(\theta, \phi), \tag{6.31}$$

where $\breve{\beta}_{\nu,\mu}$ is the SH coefficient of the exterior field, and $h_\nu^{(1)}(\cdot)$ is the first kind of the ν-th-order spherical Hankel function. From Eqs. (2.4) and (6.31), a composite noise field at a boundary of the interior and exterior sound field can be written as

$$\begin{aligned} S(\mathbf{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} [\, & \alpha_{\nu,\mu} j_\nu(kr) Y_{\nu,\mu}(\theta, \phi) \\ & + \breve{\beta}_{\nu,\mu} h_\nu^{(1)}(kr) Y_{\nu,\mu}(\theta, \phi) \,]. \end{aligned} \tag{6.32}$$

The translation relationship for the interior field in Eq. (6.23), as well as for the exterior field, can be applied to Eq. (6.32), which leads to

$$S(\mathbf{r}) = \sum_{v=0}^{\infty} \sum_{u=-v}^{v} [\, \underbrace{\alpha_{v,u}^{(q)} + \beta_{v,u}^{(q)}}_{\zeta_{v,u}^{(q)}} \,] \, j_v(kr') Y_{v,u}(\theta', \phi'), \tag{6.33}$$

where

$$\beta_{v,u}^{(q)} = \sum_{n=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \breve{\beta}_{\nu,\mu} \breve{T}_{\nu,v}^{\mu,u}(\mathbf{r}_q). \tag{6.34}$$

$\breve{T}_{\nu,v}^{\mu,u}(\mathbf{r}_q)$ is the translation coupling coefficient [38, 96] between the local interior SH coefficient $\beta_{v,u}^{(q)}$ originating from the interior sources and the global exterior SH coefficient $\breve{\beta}_{\nu,\mu}$.

Finally, taking into account all Q HOMs and applying the translation relationships in Eqs. (6.24) and (6.34), we derive the following matrix formulation:

$$\boldsymbol{\zeta}_{\text{local}} = \begin{bmatrix} \boldsymbol{\tau}^{(1)} & & & \\ & \boldsymbol{\tau}^{(2)} & & \\ & & \ddots & \\ & & & \boldsymbol{\tau}^{(Q)} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Psi}^{(1)} \\ \boldsymbol{\Psi}^{(2)} \\ \vdots \\ \boldsymbol{\Psi}^{(Q)} \end{bmatrix} \boldsymbol{\gamma}, \tag{6.35}$$

where $\boldsymbol{\zeta}_{\text{local}} = [\boldsymbol{\zeta}_{\text{local}}^{(1)\mathsf{T}}, \ldots, \boldsymbol{\zeta}_{\text{local}}^{(Q)\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{Q(V+1)^2}$, $\boldsymbol{\tau}^{(q)} = [\mathbf{T}^{(q)}, \breve{\mathbf{T}}^{(q)}]$, and

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

$\breve{\boldsymbol{\psi}}^{(q)}$ corresponds to the q-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 Eq. (6.28), $\breve{\mathbf{T}}^{(q)}$ can be written as

$$\breve{\mathbf{T}}^{(q)} = \begin{bmatrix} \breve{T}_{0,0}^{0,0}(\mathbf{r}_q) & \cdots & \breve{T}_{N,0}^{N,0}(\mathbf{r}_q) \\ \vdots & \ddots & \vdots \\ \breve{T}_{0,V}^{0,V}(\mathbf{r}_q) & \cdots & \breve{T}_{N,V}^{N,V}(\mathbf{r}_q) \end{bmatrix} \in \mathbb{C}^{(V+1)^2 \times (N+1)^2}. \tag{6.36}$$

In contrast to Eq. (6.25), we define different sensing matrices $\boldsymbol{\Psi}^{(q)}$ 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 Figures 6.6b and 6.6c. A sparse set of weights can be estimated by solving the following minimization problem:

$$\underset{\boldsymbol{\gamma}}{\text{minimize}} \, \|\boldsymbol{\zeta}_{\text{local}} - \boldsymbol{\tau}\boldsymbol{\Psi}\boldsymbol{\gamma}\|_2^2 + \lambda \|\boldsymbol{\gamma}\|_p^p. \tag{6.37}$$

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

$$\boldsymbol{\zeta}_{\text{local}} = \begin{bmatrix} \boldsymbol{\tau}^{(1)} \\ \boldsymbol{\tau}^{(2)} \\ \vdots \\ \boldsymbol{\tau}^{(Q)} \end{bmatrix} \boldsymbol{\Psi}^{(0)} \boldsymbol{\gamma}. \tag{6.38}$$

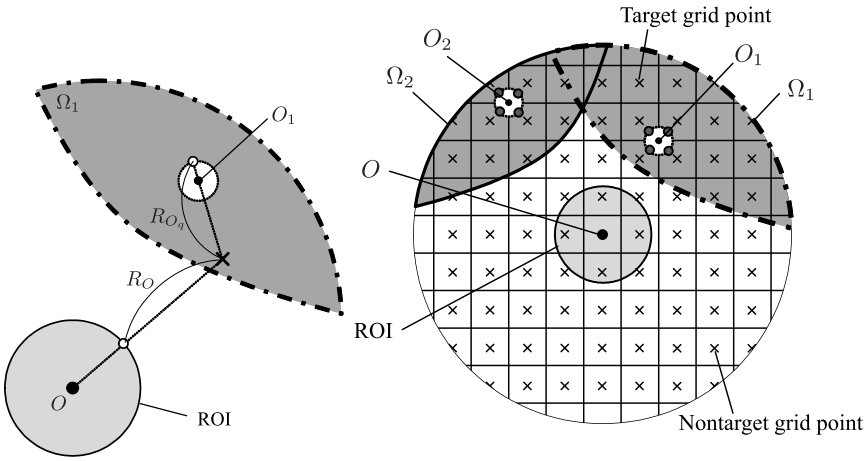


Figure 6.7: Definition example of target and nontarget grid points: (a) Illustration of a boundary of the target region. (b) Overall configuration.

6.4 Reconstruction of reference SH coefficient

In this section, we describe the reference signal reconstruction stage of our proposed method shown in Figure 6.3. The estimated spatial noise distribution $\hat{\boldsymbol{\gamma}}$ and a reconstruction matrix can be used to reconstruct the reference SH coefficients. By defining the appropriate reconstruction matrix, the reference signal can be reconstructed in a variety of ways. We discuss several definitions of the reconstruction matrix, which is required to run the SH-domain FXLMS algorithm.

First, let us assume that only incoming noise sources exist in the exterior field, which is the same configuration discussed in Section 6.3.3. Note that the grid points of the point source are defined to be outside of the outermost-positioned RMA, as shown in Figure 6.6a. Since the global interior SH coefficients of the reference signals are required, we define a reconstruction matrix as $\boldsymbol{\Pi} = \tilde{\boldsymbol{\psi}}$, 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

$$\mathbf{x} = \boldsymbol{\Pi}\hat{\boldsymbol{\gamma}}. \tag{6.39}$$

The proposed algorithm is summarized in Algorithm 1. We use M-FOCUSS algorithm [109] to estimate $\hat{\boldsymbol{\gamma}}$. As shown in Algorithm 1, M-FOCUSS is nested in the adaptive ANC processing, hence this structure is computationally expensive in general. We did, however, confirm that by initializing $\bar{\boldsymbol{\gamma}}$ with $\hat{\boldsymbol{\gamma}}$, which was the last value in the previous iteration, only a few iterations were required to satisfy M-FOCUSS's termination condition. This is because when the noise sources have the same distribution, the spatial sparsity does not change. As a result, if the distribution of noise sources does not change rapidly, the computational complexity can be reduced.

Algorithm 1: Proposed algorithm

$i = 1$; Initialize $\mathbf{w}[i] \leftarrow \mathbf{0}$;
while i reaches last frame number **do**
  $\boldsymbol{\alpha}_{\text{local}}[i] = [\alpha_{0,0}^{(1)}[i], \ldots, \alpha_{N,N}^{(Q)}[i]]^{\mathsf{T}}$;
  Update $\boldsymbol{\alpha}_{\text{local}}[i]$ by calculating Eq. (6.29);
  Update $\mathbf{e}[i]$ by calculating Eq. (6.12);
  $j = 1$; 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**
    $W[j] \leftarrow \mathrm{diag}(\bar{\boldsymbol{\gamma}}[j]^{(1-p/2)})$;
    $Q[j] \leftarrow \left[(\boldsymbol{\psi}W[j])^{\mathsf{H}}\boldsymbol{\psi}W[j] + \lambda I\right]^{-1}(\boldsymbol{\psi}W[j])^{\mathsf{H}}\boldsymbol{\alpha}_{\text{local}}[i]$;
    $\bar{\boldsymbol{\gamma}}[j] \leftarrow W[j]Q[j]$;
    if Termination condition is satisfied **then**
      $\hat{\boldsymbol{\gamma}}[i] \leftarrow \bar{\boldsymbol{\gamma}}[j]$;
      stop $= 1$;
    end
    $j \leftarrow j + 1$;
  end
  $\mathbf{x}[i] \leftarrow \boldsymbol{\Pi}\hat{\boldsymbol{\gamma}}[i]$;
  $\mathbf{z}[i] \leftarrow \mathbf{g}\mathbf{x}[i]\mathbf{w}[i]$;
  $\mathbf{w}[i+1] \leftarrow \mathbf{w}[i] - \eta(\mathbf{g}\mathbf{x}[i])^{\mathsf{H}}\mathbf{e}[i]$;
  $i \leftarrow i + 1$;
end

Next, we consider the general setup shown in Figures 6.6b and 6.6c, where both the interior and exterior noise sources exist. The exterior noise sources are represented by target grid points, from which the noise source is expected to be attenuated, whereas nontarget grid points represent the interior noise sources. We define the reconstruction matrix as $\boldsymbol{\Pi} = [\tilde{\boldsymbol{\psi}}, O] \in \mathbb{C}^{(\tilde{N}+1)^2 \times L}$, where $\boldsymbol{\Pi}$ is now a reconstruction matrix consisting only of the interior field basis functions $\tilde{\boldsymbol{\psi}}$. The global SH coefficients of the reference signals can be reconstructed using Eq. (6.39). 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 $\mathbf{r}_{\text{tar}}$ based on the distance between the local origins and the global origin. Because of the causality constraint, noise attenuation cannot be achieved in a feedforward structure when a grid point is closer to the global origin than to any local origins where an RMA exists. This is because a noise wavefront propagates to the ROI before it is captured by the RMA. Let us consider the distance R_{O_q}, which is from a grid point to the farthest reference microphone of q-th RMA, and R_O, which is from a grid point to the nearest point of ROI. The two distances are shown in Figure 6.7a. The noise

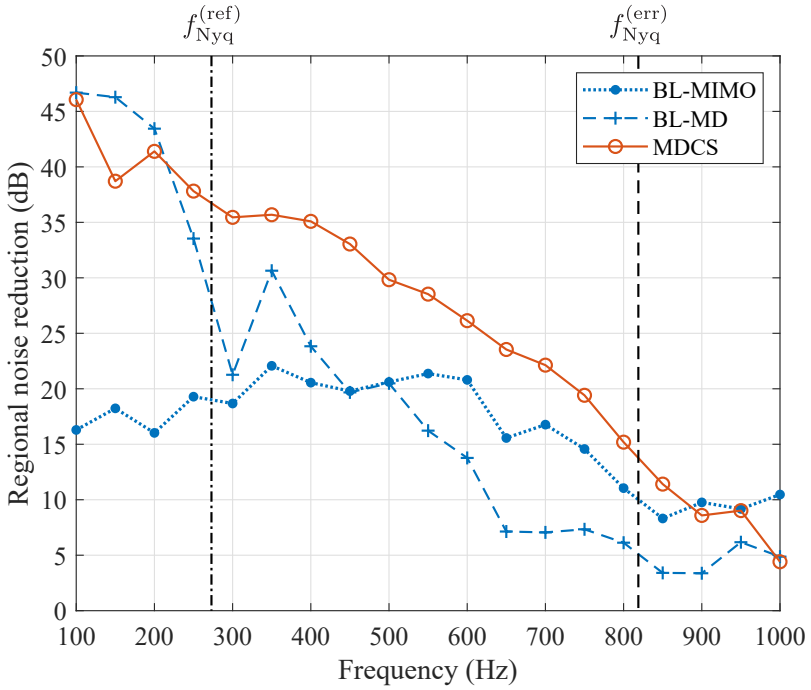


Figure 6.8: Regional noise reduction with the on-grid condition after 100 iterations of FXLMS algorithm for various frequencies. $f_{\mathrm{Nyq}}^{(\mathrm{ref})}$ and $f_{\mathrm{Nyq}}^{(\mathrm{err})}$ are the Nyquist frequencies of RMA and EMA, respectively.

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 $\mathbf{r}_{\mathrm{tar}}^{(q)} \in \Omega_q$ as the target grid point for the q-th local origin, where Ω_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 Ω_q as $\mathbf{r}_{\mathrm{tar}} \in \{\Omega_1 \cup \cdots \cup \Omega_Q\}$. The overall model is described in Figure 6.7b. We can then define the reconstruction matrix by using only the interior basis functions corresponding to the grid point $\mathbf{r}_{\mathrm{tar}}$.

The target and nontarget grid points can be arbitrarily defined in the reconstruction stage. The reconstruction matrix can be defined using prior information from the target and nontarget regions. As a result, nontarget noise sources are passed through, while only target noise sources propagating from the target region are attenuated.

6.5 Validation

6.5.1 Experimental conditions

Methods

We evaluated and compared the noise attenuation level among the following methods:

(i) "BL-MIMO" corresponds to the temporal-frequency-domain feedforward FXLMS algorithm [3];

(ii) "BL-MD" corresponds to the SH-domain feedforward FXLMS algorithm, which is a modification of the feedback version based on [23] 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 [78];

(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, $\boldsymbol{\gamma}$. Each method's filter update equation was modified by its normalized version [111], which divided the update term by the power of the filtered reference signals.

The step size was set to 1.0 for all methods.

The truncation order for each RMA and EMA can be calculated as $\tilde{N} = \lceil kR_{\mathrm{mic}} \rceil$, where R_{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. Because rotational symmetry can be assumed, we only define grid points on the horizontal plane. It is worth noting that the grid points for the interior basis functions are not defined within the ROI. The truncation order N of the translation matrix described in Eq. (6.25) is heuristically determined, i.e., $N = 11$, to satisfy $(N+1)^2 \geq (V+1)^2 Q$.

Geometrical setup

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 and feeds back the reference noise field, whereas the EMA captures and feeds back the residual noise field to the adaptive filter. Each error and reference microphone output receives a signal-to-noise ratio of 40 dB white Gaussian noise. The SLA generates the antinoise field to dampen the unwanted noise field. For the microphone arrays, we assume an open array structure. A rigid baffle array, on the other hand, is commonly used in the

actual implementation of a microphone array. In that case, reflections among HOMs may be taken into account, which is analytically derived in the literature [112], to improve the accuracy of the sound field estimation.

We considered the following array configurations:

Concentric array structure. A typical array setup of spatial ANC is a concentric array configuration, as shown in Figure 6.1a. RMA, SLA, and EMA are all placed at the global origin in this configuration. Microphones and loudspeakers are placed on a sphere's surface, which is sampled at 20 vertexes of a dodecahedron to sample the sphere uniformly. In theory, a 20-channel microphone array and SLA can control a sound field up to the third order. 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.

Nonconcentric array structure. A generalized array setup has a nonconcentric array configuration, as shown in Figure 6.1b. This configuration allows for greater flexibility in RMA placement. An HOM, for example, can be placed at a local origin. Furthermore, multiple HOMs can be placed as RMAs at multiple local origins that do not necessarily surround the ROI. In this configuration, the SLA and EMA are assumed to be at the global origin, while the RMAs are assumed to be at local origins.

Noise signal

Two monochromatic point sources of white noise were used as primary noise sources. Note that uncorrelated noises were used as the primary source signals.

Evaluation metric

The noise attenuation level was evaluated by defining the RNR inside the region as

$$\epsilon_1(\iota) = 10 \log_{10} \frac{\sum_i |S_{(0)}(\mathbf{r}_i)|^2}{\sum_i |S_{(\iota)}(\mathbf{r}_i)|^2}, \tag{6.40}$$

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

6.5.2 Noise attenuation performance in a concentric array

The RNR for the frequency band from 100 Hz to 1000 Hz is shown in Figures 6.8 and 6.9. An on-grid condition in which the primary sources are located at grid points $(1, 2, 0)$ and $(0, 5, 0)$, as well as an off-grid condition in which the sources are located at non-grid points $(0.8, 2.2, 0)$ and $(0.3, 5, 0)$, are evaluated. The result was obtained by calculating the average of 10 simulation runs. The proposed "MDCS"

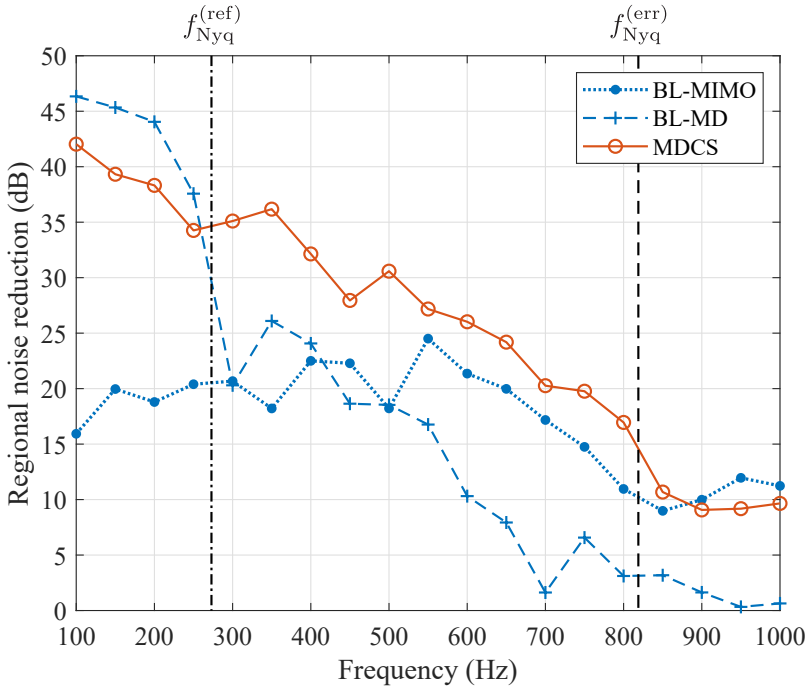


Figure 6.9: Regional noise reduction with off-grid condition after 100 iterations of FXLMS algorithm for various frequencies.

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. Figure 6.10 shows the convergence performance and Figure 6.11 shows the residual noise field on x-y plane after 100 iterations of ANC at 500 Hz. To smooth the fluctuation of the result caused by the random amplitude of the primary noise source and measurement noise, a simple moving average is calculated over 10 neighbor iterations. Figure 6.10 shows that among the methods, the proposed method provides the highest RNR with the fewest iterations.

We compared the computational complexity of the filter update among "BL-MIMO," "BL-MD," and "MDCS," which were $\mathcal{O}(L_{\mathrm{s}}(L_{\mathrm{e}}+L_{\mathrm{r}}))$, $\mathcal{O}((N+1)^2)$, and $\mathcal{O}(LL_{\mathrm{r}}^2)$, 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 Section 6.4. Therefore, the computational order of "MDCS" can be regarded as L times higher than "BL-MIMO."

6.5.3 Evaluation of different RMA structures

The proposed method provides the flexibility of the RMA structure and placement because RMAs can be located at local origins, as shown in Figure 6.1b. We compared the RNR among four setups, i.e., "SDCS," "MDCS," "MDCS-C1," and "MDCS-

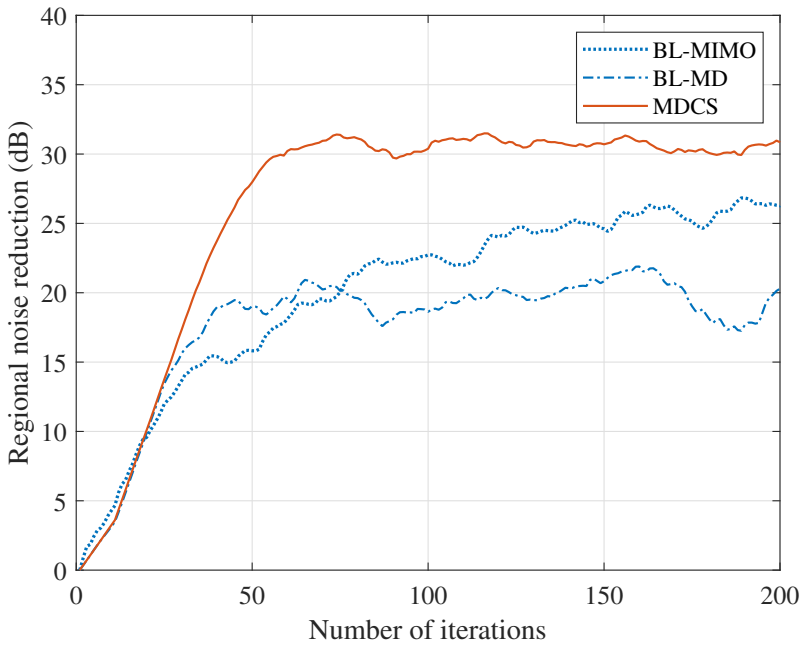


Figure 6.10: Regional noise reduction for each iteration index.

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 Figure 6.12. The proposed methods outperformed "SDCS" below 700 Hz. The high coherence of the sensing matrix in "SDCS," as shown in Figures 6.4 and 6.5, degrades the accuracy of the sparse decomposition of the reference signals at low frequency. Even when a compact RMA was used, "MDCS-C1" performed similarly to "MDCS" above 600 Hz; however, at low frequencies, the performance degraded. This is due to the lack of spatial resolution of the captured reference signal [66, 92]. We can improve low frequency performance by sampling the noise field with higher spatial resolution using distributed microphone arrays. Figure 6.12 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. Figure 6.13 shows the residual noise field on x-y plane after 100 iterations of ANC at 500 Hz.

6.5.4 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

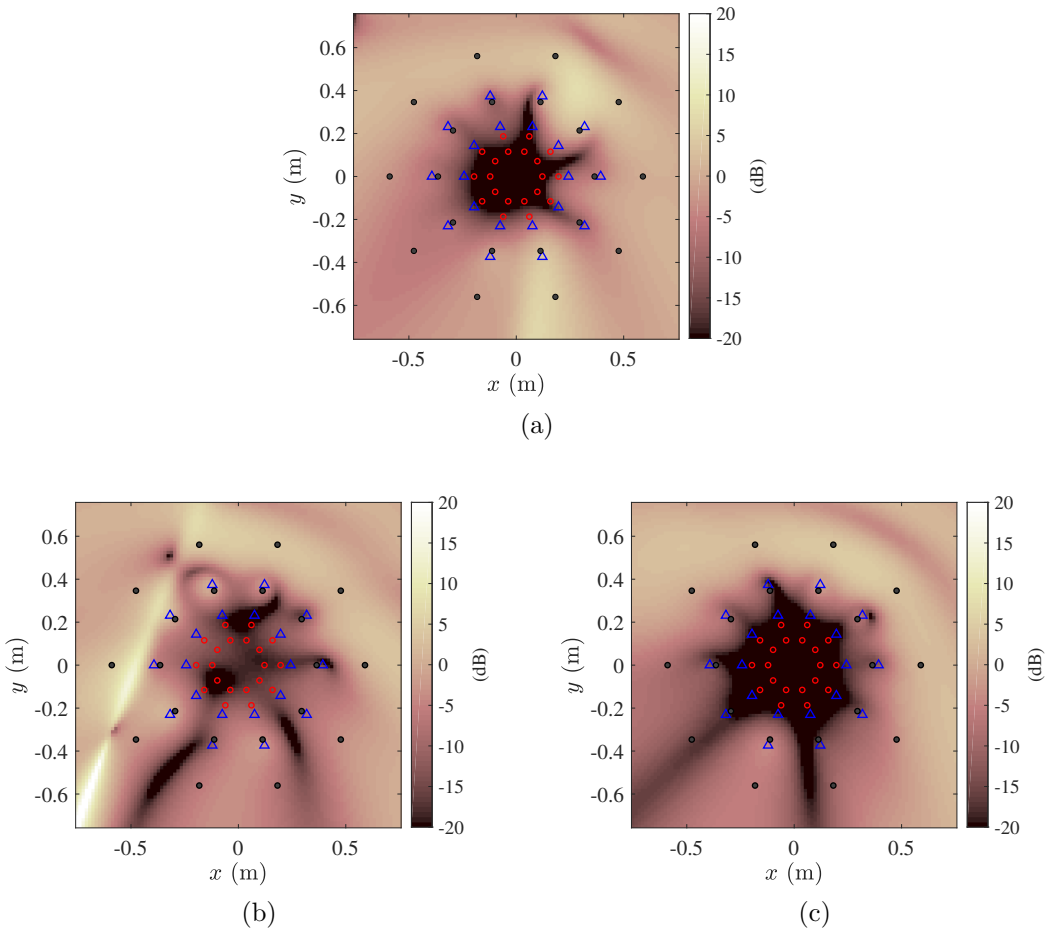


Figure 6.11: Residual noise field after 100 iterations of FXLMS algorithm: (a) "BL-MIMO." (b) "BL-MD." (c) "MDCS." The reference microphones, loudspeakers, and error microphones are denoted by black dots, blue triangles, and red circles, respectively. It should be noted that all microphones and loudspeakers are mapped to the x-y plane.

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 the same as "MDCS-C2" described in Section 6.5.3. 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 30th iteration and simulated a composite noise field consisting of two point sources, as shown in Figure 6.14a. The noise sources had a frequency of 600 Hz. Figure 6.14b depicts the noise field with only the interior noise source. Only the exterior noise source is expected to be attenuated; thus, the interior noise field remains after the adaptive processing convergence. The resulting noise field is shown in Figure 6.14c, from which we can confirm that the noise field of the interior noise source is reproduced in the ROI. The noise attenuation

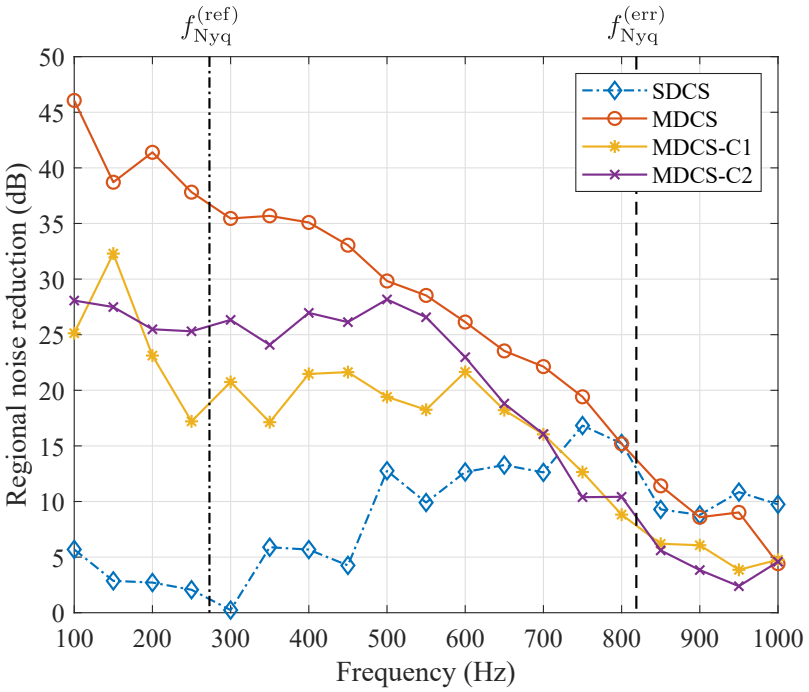


Figure 6.12: Regional noise reduction after 100 iterations of FXLMS algorithm for various frequencies.

performance of the incoming noise field was evaluated by calculating the RNR inside the region using Eq. (6.40) from the first iteration to the 29th iteration. After the 30th iteration, we defined another measure as

$$\epsilon_2(\iota) = 10\log_{10}\frac{\sum_i |S_{(\iota)}(\mathbf{r}_i) - \breve{S}_{(\iota)}(\mathbf{r}_i)|^2}{\sum_i |\breve{S}_{(\iota)}(\mathbf{r}_i)|^2}, \tag{6.41}$$

where $\breve{S}_{(\iota)}(\mathbf{r}_i)$ 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 Figure 6.14d.

Figure 6.15 shows the convergence performance for "BL-MD" and "MDCS-C2." After the 30th iteration, "BL-MD" clearly no longer attenuates the incoming noise field. In contrast, "MDCS-C2" can detect the interior noise source that we do not want to attenuate and remove it from the ANC processing; thus, the presence of the interior noise source does not affect 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 hear a sound, such as music, from your front but not noise, such as 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.

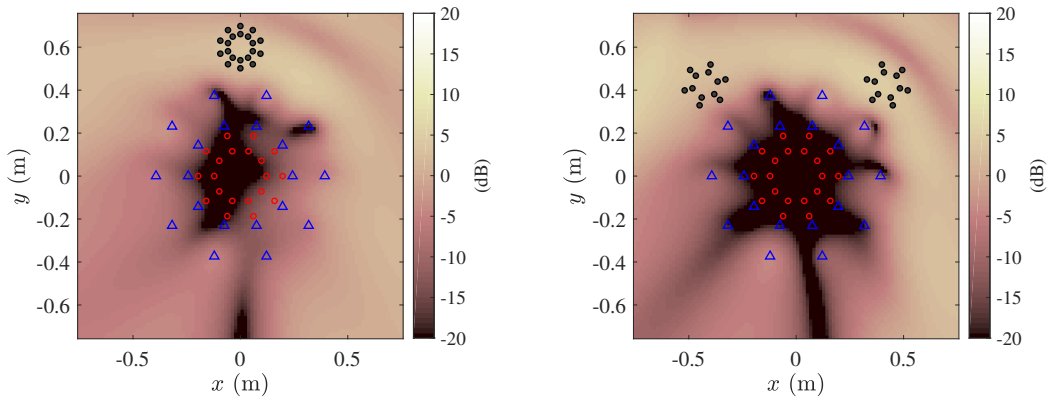


Figure 6.13: Residual noise field after 100 iterations of FXLMS algorithm: (a) "MDCS-C1." (b) "MDCS-C2."

However, narrow directivity is difficult to achieve at low frequencies, affecting ANC performance due to signal capture from nontarget noise sources. This problem can be solved using the proposed method.

6.5.5 Actual room experiment

We conducted measurements in our audio laboratory to evaluate the noise attenuation performance in a real-world environment. The main differences from the ideal setup are that the acoustic transfer function includes reverberations and that primary and secondary sources are not always point sources. The Eigenmike [54] 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. At the center of the SLA, the EMA was placed. The RMA and the primary loudspeaker, which produces the primary noise, were placed 1.7 m and 3.18 m apart from the center point, respectively. The secondary path's SH coefficients can be calculated based on the measured secondary path $\mathbf{G}_{\mathrm{act}} \in \mathbb{C}^{L_{\mathrm{e}} \times L_{\mathrm{s}}}$. We write the matrix form of (6.12) as $\mathbf{e} = \mathbf{T}_{\mathrm{e}}\mathbf{S}_{\mathrm{e}}$, where $\mathbf{T}_{\mathrm{e}} \in \mathbb{C}^{(N+1)^2 \times L_{\mathrm{e}}}$ is the SH transformation matrix. From (2.18) and (2.20), the SH coefficients of the secondary path can be derived in a matrix form as

$$\mathbf{g}_{\mathrm{act}} = \mathbf{T}_{\mathrm{e}}\mathbf{G}_{\mathrm{act}}\mathbf{Y}, \tag{6.42}$$

where $\mathbf{Y} \in \mathbb{C}^{L_{\mathrm{s}} \times (N+1)^2}$ is the matrix consisting of SH functions. We confirmed that the diagonal elements of $\mathbf{g}_{\mathrm{act}}$ were dominant in our experimental condition. Hence, we extracted the diagonal elements and diagonalized the matrix similar to the related work [84].

The average noise attenuation level at the error microphones is shown in Figure 6.16. The noise attenuation was achieved at the frequency range of 100 to

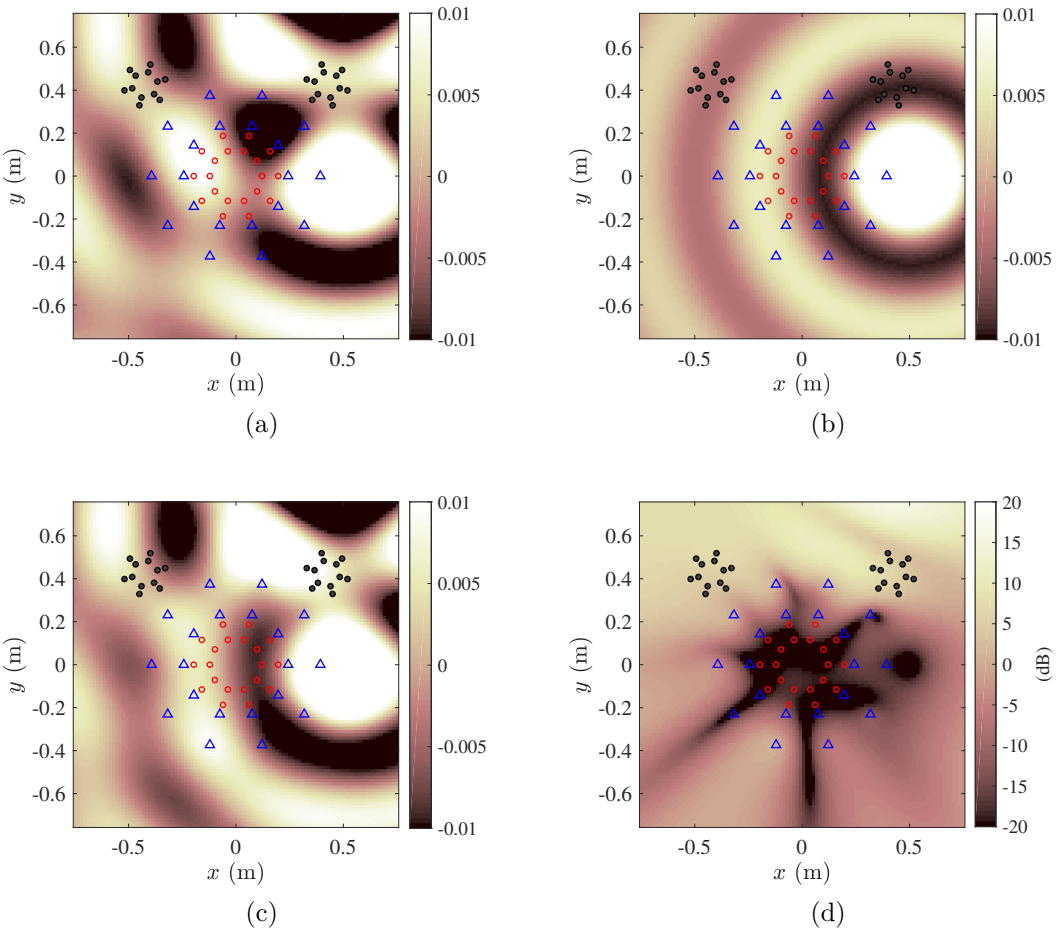


Figure 6.14: Simulated sound fields of interior noise compensation: (a) Composite noise field. (b) Noise field of interior noise source. (c) Residual noise field after noise attenuation. (d) Error distribution based on Eq. (6.41).

1000 Hz. Unlike the results of the simulation in Sections 6.5.2 and 6.5.3, performance degradation at high frequencies cannot be seen. One possible explanation is that the evaluation points in this experiment are only at the error microphone positions. The evaluation in terms of a large region in a real-world environment is something to think about for future research.

6.6 Summary

In this chapter, 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

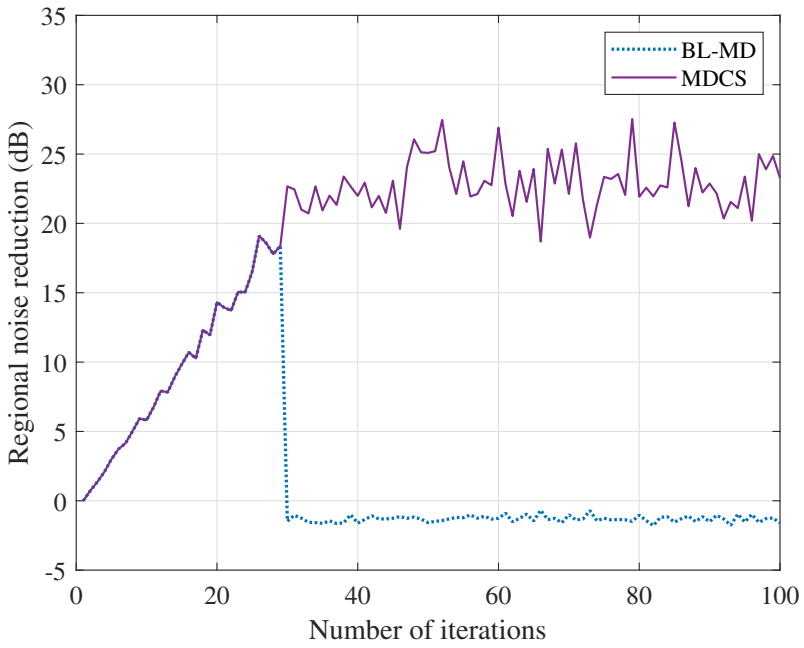


Figure 6.15: Average performance of the incoming noise field attenuation. Interior noise source is added at the 30th iteration.

basis function is one of the efficient spaces to decompose the sound field. We tested the proposed method and found that it outperformed the conventional method for noise attenuation at any frequency above the RMA's Nyquist frequency. 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 aids practical implementation by removing an unexpected noise source from the system. Furthermore, in our method, any arbitrary array structure can be used, resulting in the realization of a viable system.

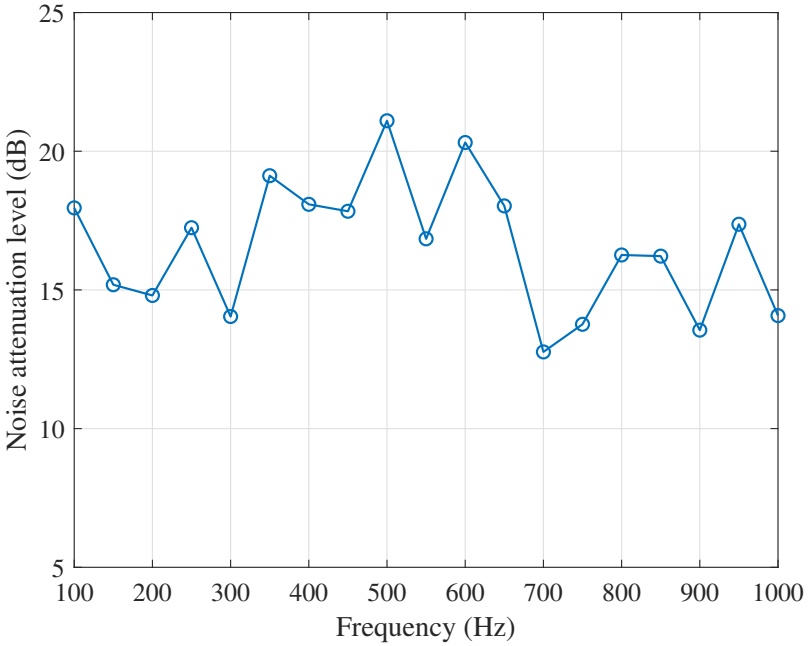


Figure 6.16: Average noise attenuation level after 100 iterations of FXLMS algorithm for various frequencies.

Chapter 7

Conclusion

7.1 Summary

The ANC over a sizable space requires a large number of reference and error microphones to satisfy the spatial Nyquist sampling criterion, which limits the feasibility of practical realization of such systems. This study looked into how to improve system feasibility by introducing several approaches to simplify array structures.

In Chapter 3, the CS approach was introduced in the spatial ANC system. The sparse reference signal representation was adopted in the harmonic-domain signal to precisely calculate the reference harmonic coefficients. The proposed feedforward ANC methods reduced the number of microphones required while attenuating the noise field over a large area. The experimental results showed that the proposed method outperformed conventional methods even above the spatial Nyquist frequency.

The multiple circular arrays were developed in Chapter 4 to attenuate the noise field in 3D space, where the fully surrounding spherical array is usually required for the accurate processing in SH domain. The algorithm for estimating the SH coefficients is based on the properties of the SH-domain representation. Although the array structure must be carefully designed for each condition, the simulation results demonstrated the method's effectiveness.

To leverage the prior information of the array structure, the novel weighted LMS algorithm for the adaptive filtering was proposed in Chapter 5. The DOA weighting method was developed, which can calculate the SH-domain weights from an arbitrary spatial weighting function. Considering the secondary source geometry, the weighting function can be set intuitively.

The CS-based approach was revisited in Chapter 6. In contrast to the spatial sensing matrices proposed in Chapter 3, the new sensing matrices were defined in the SH-domain. The analysis of the sensing matrix's column coherence revealed an improvement in the accuracy of reference signal estimation at low frequencies. Furthermore, the distributed array structure of the RMAs was investigated to improve the flexibility of the array geometry. The simulation considering a real-world application was also conducted. The typical use case of the spatial ANC is to sur-

round the users with an array of microphones to avoid any unwanted noise reaching them. The problem arises when the users make noise, e.g., having a chat, inside the array. The system attempts to dampen the desired sound, which you do not want to dampen. The proposed method demonstrated a solution to this problem by utilizing the estimated distribution of the noise source. The proposed system maintains noise attenuation performance by ignoring the desired outgoing sound in ANC processing.

7.2 Outlook

The uniform sampling of the 3D sound field on the sphere surface is the straightforward scheme to calculate the SH-domain coefficients. However, the drawback is the system's feasibility. The alternative design of the microphone-loudspeaker array geometry was investigated in this thesis. The SH-domain signal characteristics were the focus of one approach described in Chapter 4. The multiple circular arrays were designed to attenuate the horizontal circular region in the specific condition. For other conditions such as the frequency range, truncation order, and size of the ROI, a different array geometry has to be designed. One of the future research directions will be the investigation of the systematic method for determining array geometry. Another method described in Chapter 5 included DOA weighting in adaptive processing. In this thesis, the simple binary window function was used. In contrast, a continuous function may result in an accurate calculation in the SH-domain weights, which may also result in an accurate solution. An investigation of the combination of the array structure and the adequate adaptive algorithms is the possible direction of further study.

As a property of the spatial ANC, the computational complexity should be concerned. The implementation of the CS approach described in Chapters 3 and 6 necessitates iterative processing using IRLS-based algorithms. As a result, the two distinct iterations are nested. The first is an iteration to estimate the sparse weights, and the second is an update to the ANC system's adaptive filter. By focusing on the iterative approach in adaptive filtering, the sparse weights may be estimated more efficiently. More specifically, in the adaptive filtering, the internal variable in the IRLS method, which is updated each iteration, can be estimated efficiently using the information from the previous iteration. The sparse weights represent the noise source distribution. Unless the actual distribution does not change rapidly, starting from the previous internal weights may lead to fast convergence of the CS. One of the important future works for the realization of the spatial ANC is research aimed at reducing computational complexity.

This thesis did not consider any feedback from the secondary sources to the reference microphones as it is assumed to be compensated ideally. To get even closer to real-world situations, however, this feedback path has to be taken into account. Furthermore, the evaluation in a highly reverberant environment needs to be analyzed since the assumption of the sparseness of the source distribution may no longer be valid. Further experiments in the real-world may also aid in evaluating the

system's robustness because it may contain positional errors of the microphones and loudspeakers, sensitivity and gain variations in the microphones and loudspeakers, and other errors that cannot be fully modeled.

Acknowledgment

First and foremost I would like to thank my supervisor Prof. Keisuke Kameyama for all his help and advice with my doctoral course. I would also like to thank Prof. Hiroyuki Kudo, Prof. Ko Sakai, Prof. Takeshi Yamada, and Prof. Naoto Wakatsuki from Systems and Information Engineering. Their insightful comments and concise advice incented me to polish this thesis. My sincere thanks also go to Dr. Yuki Mitsufuji, who made me an opportunity to start this study and gave me continuous support. I thank Mr. Naoki Murata, Mr. Yuhta Takida, and my colleagues at Sony for having invaluable discussions. I would like to thank my esteemed supervisor, Prof. Thushara Abhayapala, who gave me a lot of suggestions in my study with his immense knowledge and plentiful experience. I would also like to express gratitude to Dr. Prasanga Samarasinghe for her insightful discussions which have widened my research. Additionally, I thank Dr. Hanchi Chen, Dr. Jihui Zhang, and lab members at the Australian National University. Their kind help and support have made my study and life in Australia a wonderful time. I would also like to thank Dr. Homare Kon for his understanding of undertaking the doctoral course while working. Last but not the least, I would like to thank my wife and my child for their understanding and encouragement all through my study.

Bibliography

[1] C. M. Harris, *Handbook of Acoustical Measurements and Noise Control*. McGraw-Hill New York, 1991.

[2] I. L. Vér and L. L. Beranek, *Noise and Vibration Control Engineering: Principles and Applications*. John Wiley & Sons, 2005.

[3] S. M. Kuo and D. R. Morgan, *Active Noise Control Systems: Algorithms and DSP Implementations*. New York: Wiley, 1995.

[4] ——, "Active noise control: a tutorial review," *Proc. IEEE*, vol. 87, no. 6, pp. 943–973, Jun. 1999.

[5] Y. Kajikawa, W.-S. Gan, and S. M. Kuo, "Recent advances on active noise control: open issues and innovative applications," *APSIPA Trans. Signal Inform. Process.*, vol. 1, pp. 1–21, Dec. 2012.

[6] 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.

[7] C. Hansen, S. Snyder, X. Qiu, L. Brooks, and D. Moreau, *Active Control of Noise and Vibration*. CRC Press, 2012.

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

[9] J. Daniel, "Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format," in *Proc. Audio Eng. Soc. 23rd Int. Conf.: Signal Processing in Audio Recording and Reproduction*, 2003.

[10] *AES recommended practice for digital audio engineering – Serial multichannel audio digital interface (MADI)*. Doc. AES10-2008 (R2019), Audio Engineering Society, New York, 2008.

[11] 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.

[12] 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.

[13] M. Bouchard and B. Paillard, "A transform domain optimization to increase the convergence speed of the multichannel filtered-x least-mean-square algorithm," *J. Acoust. Soc. Amer.*, vol. 100, no. 5, pp. 3203–3214, 1996.

[14] T. Kosaka, S. J. Elliott, and C. C. Boucher, "A novel frequency domain filtered-x lms algorithm for active noise reduction," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 1. IEEE, 1997, pp. 403–406.

[15] D. Das, G. Panda, and S. Kuo, "New block filtered-X LMS algorithms for active noise control systems," *IET Signal Processing*, vol. 1, no. 2, pp. 73–81, 2007.

[16] S. M. Kuo, R. K. Yenduri, and A. Gupta, "Frequency-domain delayless active sound quality control algorithm," *J. Sound Vib.*, vol. 318, no. 4-5, pp. 715–724, 2008.

[17] X. Tang and C.-M. Lee, "Time-frequency-domain filtered-x LMS algorithm for active noise control," *J. Sound Vib.*, vol. 331, no. 23, pp. 5002–5011, 2012.

[18] F. Yang, Y. Cao, M. Wu, F. Albu, and J. Yang, "Frequency-domain filtered-x LMS algorithms for active noise control: A review and new insights," *Applied Sciences*, vol. 8, no. 11, p. 2313, 2018.

[19] D. R. Morgan, "An adaptive modal-based active control system," *J. Acoust. Soc. Amer.*, vol. 89, no. 1, pp. 248–256, Jan. 1991.

[20] S. Spors and H. Buchner, "An approach to massive multichannel broadband feedforward active noise control using wave-domain adaptive filtering," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, 2007, pp. 171–174.

[21] ——, "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.

[22] 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.

[23] 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, Language Process.*, vol. 26, no. 4, pp. 774–786, Apr. 2018.

[24] 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, Language Process.*, vol. 26, no. 12, pp. 2355–2370, Dec. 2018.

[25] T. D. Abhayapala and A. Gupta, "Spherical harmonic analysis of wavefields using multiple circular sensor arrays," *IEEE Trans. Audio, Speech, Language Process.*, vol. 18, no. 6, pp. 1655–1666, Aug. 2010.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] Y. Takida, S. Koyama, and H. Saruwatari, "Exterior and interior sound field separation using convex optimization: Comparison of signal models," in *Proc. 26th European Signal Process. Conf.*, Rome, Italy, Sep. 2018, pp. 2549–2553.

[31] 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.

[32] 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.

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

[34] B. Rafaely, B. Weiss, and E. Bachmat, "Spatial aliasing in spherical microphone arrays," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 1003–1010, 2007.

[35] 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.

[36] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. New York: Academic press, 1999.

[37] D. L. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer, 1998, vol. 93.

[38] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Elsevier, Oxford, UK, 2005.

[39] S. J. Elliott and P. A. Nelson, "The application of adaptive filtering to the active control of sound and vibration," *NASA STI/Recon Technical Report N*, vol. 86, p. 32628, 1985.

[40] D. Morgan, "An analysis of multiple correlation cancellation loops with a filter in the auxiliary path," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 4, pp. 454–467, 1980.

[41] J. C. Burgess, "Active adaptive sound control in a duct: A computer simulation," *J. Acoust. Soc. Amer.*, vol. 70, no. 3, pp. 715–726, 1981.

[42] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1985.

[43] L. J. Eriksson and M. C. Allie, "Use of random noise for on-line transducer modeling in an adaptive active attenuation system," *J. Acoust. Soc. Amer.*, vol. 85, no. 2, pp. 797–802, 1989.

[44] M. Zhang, H. Lan, and W. Ser, "Cross-updated active noise control system with online secondary path modeling," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 5, pp. 598–602, 2001.

[45] M. de Diego, A. Gonzalez, M. Ferrer, and G. Piñero, "Multichannel active noise control system for local spectral reshaping of multifrequency noise," *J. Sound Vib.*, vol. 274, no. 1-2, pp. 249–271, 2004.

[46] M. T. Akhtar and W. Mitsuhashi, "Improving performance of FxLMS algorithm for active noise control of impulsive noise," *J. Sound Vib.*, vol. 327, no. 3-5, pp. 647–656, 2009.

[47] M. Bouchard and S. Quednau, "Multichannel recursive-least-square algorithms and fast-transversal-filter algorithms for active noise control and sound reproduction systems," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 606–618, 2000.

[48] G. Chen, H. Wang, K. Chen, and K. Muto, "A preprocessing method for multichannel feedforward active noise control," *Acoustical Science and Technology*, vol. 26, no. 3, pp. 292–295, 2005.

[49] J. Yuan, "Orthogonal adaptation for multichannel feedforward control," *J. Acoust. Soc. Amer.*, vol. 120, no. 6, pp. 3723–3729, 2006.

[50] J. K. Thomas, S. P. Lovstedt, J. D. Blotter, and S. D. Sommerfeldt, "Eigenvalue equalization filtered-x algorithm for the multichannel active noise control of stationary and nonstationary signals," *J. Acoust. Soc. Amer.*, vol. 123, no. 6, pp. 4238–4249, 2008.

[51] 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.

[52] 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.

[53] Z. Li, R. Duraiswami, E. Grassi, and L. S. Davis, "Flexible layout and optimal cancellation of the orthonormality error for spherical microphone arrays," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 4, 2004, pp. 41–44.

[54] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 2, Orlando, FL, USA, May 2002, pp. 1781–1784.

[55] L. Gandemer, G. Parseihian, C. Bourdin, and R. Kronland-Martinet, "Perception of surrounding sound source trajectories in the horizontal plane: A comparison of VBAP and basic-decoded HOA," *Acta Acustica United With Acustica*, vol. 104, no. 2, pp. 338–350, 2018.

[56] G. D. Romigh, D. S. Brungart, R. M. Stern, and B. D. Simpson, "Efficient real spherical harmonic representation of head-related transfer functions," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 5, pp. 921–930, 2015.

[57] J. M. Arend, P. Stade, and C. Pörschmann, "Binaural reproduction of self-generated sound in virtual acoustic environments," in *Proc. Mtgs. Acoust.*, vol. 30, no. 1, 2017, p. 015007.

[58] M. A. Poletti, "A unified theory of horizontal holographic sound systems," *J. Audio Eng. Soc.*, vol. 48, no. 12, pp. 1155–1182, 2000.

[59] T. D. Abhayapala and D. B. Ward, "Theory and design of high order sound field microphones using spherical microphone array," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 2, 2002, pp. 1949–1952.

[60] J. Ahrens, *Analytic Methods of Sound Field Synthesis*. New York: Springer, 2012.

[61] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[62] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[63] J. Zhang, T. D. Abhayapala, P. N. Samarasinghe, W. Zhang, and S. Jiang, "Multichannel active noise control for spatially sparse noise fields," *J. Acoust. Soc. Amer.*, vol. 140, no. 6, pp. EL510–EL516, 2016.

[64] A. Moiola, R. Hiptmair, and I. Perugia, "Plane wave approximation of homogeneous helmholtz solutions," *Zeitschrift für angewandte Mathematik und Physik*, vol. 62, no. 5, p. 809, 2011.

[65] C. Fan, S. M. A. Salehin, and T. D. Abhayapala, "Practical implementation and analysis of spatial soundfield capture by higher order microphones," in *Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2014, pp. 1–8.

[66] E. J. Candes, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. and Comput. Harmonic Anal.*, vol. 31, no. 1, pp. 59–73, Jul. 2011.

[67] R. E. Carrillo and K. E. Barner, "Iteratively re-weighted least squares for sparse signal reconstruction from noisy measurements," in *Proc. 43rd IEEE Annu. Conf. Inf. Sci. Syst. (CISS)*, 2009, pp. 448–453.

[68] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Process. Lett.*, vol. 14, no. 10, pp. 707–710, Oct. 2007.

[69] 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.

[70] C. Chen, J. Huang, L. He, and H. Li, "Fast iteratively reweighted least squares algorithms for analysis-based sparsity reconstruction," *arXiv preprint arXiv:1411.5057*, 2014.

[71] P. Samarasinghe, T. Abhayapala, and M. Poletti, "Wavefield analysis over large areas using distributed higher order microphones," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 22, no. 3, pp. 647–658, Mar. 2014.

[72] D. T. M. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms," *IEEE Trans. Signal Process.*, vol. 41, no. 9, pp. 2811–2825, Sep. 1993.

[73] S. Elliot, P. Nelson, I. Stothers, and C. Boucher, "In-flight experiments on the active control of propeller-induced cabin noise," *J. Sound Vib.*, vol. 140, no. 2, pp. 219–238, 1990.

[74] D. Shi, J. He, C. Shi, T. Murao, and W.-S. Gan, "Multiple parallel branch with folding architecture for multichannel filtered-x least mean square algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2017, pp. 1188–1192.

[75] P. A. Nelson, F. Orduna-Bustamante, and H. Hamada, "Inverse filter design and equalization zones in multichannel sound reproduction," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 3, pp. 185–192, 1995.

[76] "Sound field synthesis toolbox," https://sfs-matlab.readthedocs.io/.

[77] H. Wierstorf and S. Spors, "Sound field synthesis toolbox," in *Proc. Audio Eng. Soc. 132nd Conv.*, 2012.

[78] 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.

[79] E. W. Start, "Direct sound enhancement by wave field synthesis," *Ph.D. thesis, Delft University of Technology*, 1997.

[80] S. Spors and J. Ahrens, "Analysis and improvement of pre-equalization in 2.5-dimensional wave field synthesis," in *Proc. Audio Eng. Soc. 128th Conv.*, 2010.

[81] J. Ahrens and S. Spors, "Applying the ambisonics approach on planar and linear arrays of loudspeakers," in *2nd Int. Symp. on Ambisonics and Spherical Acoustics*, 2010.

[82] A. Gupta and T. D. Abhayapala, "Three-dimensional sound field reproduction using multiple circular loudspeaker arrays," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 5, pp. 1149–1159, 2011.

[83] S. Moreau, J. Daniel, and S. Bertet, "3D sound field recording with higher order ambisonics–objective measurements and validation of a 4th order spherical microphone," in *Proc. Audio Eng. Soc. 120th Conv.*, 2006, pp. 20–23.

[84] 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, UK, May 2019, pp. 526–530.

[85] H. Ito, S. Koyama, N. Ueno, and H. Saruwatari, "Feedforward spatial active noise control based on kernel interpolation of sound field," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Brighton, UK, May 2019, pp. 511–515.

[86] ——, "Three-dimensional spatial active noise control based on kernel-induced sound field interpolation," in *Proc. Int. Congress Acoust.*, Aachen, Germany, Sep. 2019, pp. 1101–1108.

[87] Y. Maeno, Y. Mitsufuji, P. N. Samarasinghe, N. Murata, and T. D. Abhayapala, "Spherical-harmonic-domain feedforward active noise control using sparse decomposition of reference signals from distributed sensor arrays," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 28, pp. 656–670, Dec. 2019.

[88] Y. Maeno, Y. Mitsufuji, P. N. Samarasinghe, and T. D. Abhayapala, "Mode-domain spatial active noise control using multiple circular arrays," in *Proc. Int. Workshop Acoust. Signal Enhancement*, Tokyo, Japan, Sep. 2018, pp. 441–445.

[89] B. Rafaely, *Fundamentals of Spherical Array Processing.* New York: Springer, 2015.

[90] L. F. Shampine, "Vectorized adaptive quadrature in MATLAB," *J. Comput. Appl. Math.*, vol. 211, no. 2, pp. 131–140, Feb. 2008.

[91] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[92] A. Xenaki, P. Gerstoft, and K. Mosegaard, "Compressive beamforming," *J. Acoust. Soc. Amer.*, vol. 136, no. 1, pp. 260–271, Jul. 2014.

[93] 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.

[94] 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.

[95] ——, "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.

[96] P. N. Samarasinghe, T. D. Abhayapala, and M. A. Poletti, "3D spatial soundfield recording over large regions," in *Proc. Int. Workshop Acoust. Signal Enhancement*, Aachen, Germany, Sep. 2012, pp. 1–4.

[97] 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.

[98] 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.

[99] Y. Takida, S. Koyama, N. Ueno, and H. Saruwatari, "Gridless sound field decomposition based on reciprocity gap functional in spherical harmonic domain," in *IEEE Int. Sensor Array Multichannel Signal Process. Workshop (SAM)*, Sheffield, UK, Jul. 2018, pp. 627–631.

[100] S. Koyama and L. Daudet, "Sparse representation of a spatial sound field in a reverberant environment," *IEEE J. Sel. Topics in Signal Process.*, vol. 13, no. 1, pp. 172–184, Mar. 2019.

[101] 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.

[102] P. Samarasinghe, T. Abhayapala, M. Poletti, and T. Betlehem, "An efficient parameterization of the room transfer function," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 23, no. 12, pp. 2217–2227, Dec. 2015.

[103] Y. Iwamatsu, K. Fujii, and M. Muneyasu, "Frequency domain method to estimate the coefficients of feedback control filter for active noise control systems," *Acoustical Science and Technology*, vol. 27, no. 5, pp. 264–269, 2006.

[104] M. T. Akhtar, M. Abe, and M. Kawamata, "On active noise control systems with online acoustic feedback path modeling," *IEEE Trans. Audio, Speech, Language Process.*, vol. 15, no. 2, pp. 593–600, Feb. 2007.

[105] I. Balmages and B. Rafaely, "Open-sphere designs for spherical microphone arrays," *IEEE Trans. Audio, Speech, Language Process.*, vol. 15, no. 2, pp. 727–732, Feb. 2007.

[106] 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.

[107] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society, Series B*, vol. 58, no. 1, pp. 267–288, 1996.

[108] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 600–616, Mar 1997.

[109] 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.

[110] V. I. Lebedev and D. Laikov, "A quadrature formula for the sphere of the 131st algebraic order of accuracy," *Doklady Mathematics*, vol. 59, pp. 477–481, 1999.

[111] S. Elliott, *Signal Processing for Active Control*. Elsevier, 2000.

[112] P. A. Martin, *Multiple Scattering: Interaction of Time-Harmonic Waves with N Obstacles*. Cambridge University Press, 2006.

Own Publications

1. Y. Maeno, Y. Takida, N. Murata, Y. Mitsufuji, "Array-geometry-aware spatial active noise control based on direction-of-arrival weighting," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Barcelona, pp. 8414-8418, May, 2020.

2. Y. Maeno, Y. Mitsufuji, P. N. Samarasinghe, N. Murata, T. D. Abhayapala, "Spherical-harmonic-domain feedforward active noise control using sparse decomposition of reference signals from distributed sensor arrays," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 28, pp. 656-670, 2019.

3. Y. Maeno, Y. Mitsufuji, P. N. Samarasinghe, T. D. Abhayapala, "Mode-domain spatial active noise control using multiple circular arrays," in *Proc. Int. Workshop Acoust. Signal Enhancement*, Tokyo, pp. 441-445, Sep, 2018.

4. Y. Maeno , Y. Mitsufuji, T. D. Abhayapala, "Mode domain spatial active noise control using sparse signal representation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Calgary, pp.211-215, Apr, 2018.