Active control of outgoing noise fields in rooms

Fei Ma

M.Sc. Engineering, Harbin Institute of Technology
B.Sc. Engineering, Harbin Institute of Technology

May 2020

A thesis submitted for the degree of Doctor of Philosophy
of The Australian National University

Audio & Acoustic Signal Processing Group
Research School of Electrical, Energy & Materials Engineering
The Australian National University

Declaration

The contents of this thesis are the results of original research and have not been submitted for a higher degree to any other university or institution. Much of this work has either been published or submitted for publications as journal papers and conference proceedings. Following is a list of these papers.

Journal Publications

- **Fei Ma**, Wen Zhang, Thushara D. Abhayapala, "Active control of outgoing noise fields in rooms," The Journal of the Acoustic Society of America, vol. 144, no. 3, pp. 1589-1599, Sep. 2018.
- **Fei Ma**, Wen Zhang, Thushara D. Abhayapala, "Real-time separation of non-stationary sound fields on spheres," The Journal of the Acoustic Society of America, vol. 146, no. 1, pp. 11-21, Jul. 2019.
- **Fei Ma**, Wen Zhang, Thushara D. Abhayapala, "Active control of broadband outgoing noise fields in rooms," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 529-539, 2020.
- **Fei Ma**, Thushara D. Abhayapala, Wen Zhang, "Multiple circular arrays of vector sensors for real-time 3D sound field analysis," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 29, pp. 286-299, 2021.

Conference Proceedings

- **Fei Ma**, Wen Zhang, and Thushara D. Abhayapala, "Reference signal generation for broadband ANC systems in reverberant rooms," 2018 IEEE Inter-

national Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 216-220, Apr. 2018.

The research work presented in this thesis has been performed jointly with Prof. Thushara D. Abhayapala and Prof. Wen Zhang. Approximately 75% of this work is my own.

Fei Ma
Research School of Electrical, Energy & Materials Engineering
The Australian National University
Canberra ACT 2600
May 2020

Acknowledgments

Without the support of the many faces in my life, this work would not have been possible. I would like to acknowledge and thank each of the followings:

- I would like to thank my supervisors Porf. Thushara D. Abhayapala and Porf. Wen Zhang. They encouraged me to do works with real impacts. They allowed me to make mistakes. I will never be able to derive the equations in the paper "Real-time separation of non-stationary sound fields on spheres" without their guidance.

- I would like to thank Dr. Prasanga N. Samarasinghe for her suggestions on communication skills.

- I would like to thank Dr. Hanchi Chen, who uses his rich hardware knowledge to help me design the tonal active noise control experiment.

- I would like to thank Dr. Jihui Zhang for her financial support.

- I would like to thank Dr. Qingbo Sun, his wife, and his daughters. It has been a wonderful time to be a close friend with them in the past five years.

- I would like to thank Lachlan Birnie, Huiyuan Sun, Huanyu Zuo, Yonggang Hu, for the interesting discussion during the lunch time.

- I would like to thank Dr. Buye Xu, Dr. Terry Cho, and Dr. Tony Miller, Dr. Chuming Zhao, engineers from Facebook, Inc. It was a wonderful time to work with them.

Abstract

Active noise control is a strategy to suppress a noise by superimposing it with a carefully designed secondary noise. This strategy has been under research over the past half century with active noise control aided devices surging on the market over the last decade. However, up to now, the most successful applications of active noise control are still limited to the single channel systems, where noises propagate in ducts or in the human ear canals.

Many researchers attempted to extend the applications of active noise control to spatial noise fields, such as controlling the tire rolling noise in cars, the ventilation noise at workplaces, or the pump engine noise outdoors, which account for the majority of noises we encounter in our everyday lives. They developed spatial active noise control systems based on room modes, spherical modes, or the Helmholtz integral equation. The attempts have found limited success in the real world because of two problems. The first is that a spatial noise field is the complicated interaction of a number of noise sources with the environment, which can be non-stationary and time-varying. This problem makes it extremely difficult to obtain clean reference signals for spatial active noise control systems. The second is that due to the lack of a time-domain spatial sound field control theory, the existing spatial active noise control systems process the acoustic quantities in the time-frequency domain. The time-frequency domain processing introduces the frame delay and thus probably makes the systems violate the causal control constraint.

This thesis proposes an outgoing noise field control system based on the frequency-domain sound field separation method. The method decouples the outgoing field (due to the noise sources) from the incoming field (due to the environment) on a sphere surrounding the noise sources. By canceling the outgoing field only, the proposed system reduces the noise entirely in a room without estimating the secondary paths in real-time and with negligible influence on the desired sound field in the room.

This thesis further derives a time-domain sound field separation method, based on which a low latency outgoing field control system with random noise field cancellation capacity is developed. Multiple circular arrays of vector sensors for three-dimensional sound field analysis are developed based on the time-domain method. The designed arrays have a compact geometry, and thus can be integrated with small sized wearable devices and provide them with real-time sound field analysis capacity.

Notations and Symbols

i	$\sqrt{-1}$, the unit imaginary number
n	the discrete time index
t	the continuous time index
f	frequency, Hz
f_s	the sampling frequency, Hz
ω	the angular frequency, $\omega = 2\pi * f$
c	the speed of sound, 343 m/s
ρ	the density of air, 1.225 kg/m^3
(x, y, z)	the Cartesian coordinates
(r, θ, ϕ)	the spherical coordinates , $\Theta = (\theta, \phi)$
τ_R, τ	R/c, r/c
$j_u(\cdot), j'_u(\cdot)$	the spherical Bessel function and its derivative with respect to the argument
$h_u(\cdot), h'_u(\cdot)$	the spherical Hankel function and its derivative with respect to the argument
$\mathcal{P}_u(\cdot)$	the Legendre function
$\mathfrak{P}_{u,v}(\cdot)$	the associated Legendre function
$\mathcal{P}_{u,v}(\cdot)$	the normalized associated Legendre function
$Y_{u,v}(\cdot,\cdot)$	the spherical harmonics

$\lceil \cdot \rceil$	the ceiling operator
$\lvert \cdot \rvert$	the absolute value operator
$(\cdot)^T$	the transpose operator
$\lvert\lvert \cdot \rvert\rvert_2$	the 2-norm operator
$P(\cdot)$, $p(\cdot)$	the frequency-domain and the time-domain pressure
$V(\cdot)$, $\nu(\cdot)$	the frequency-domain and the time-domain radial particle velocity
$\frac{dp(\cdot)}{dt}$, $\frac{d\nu(\cdot)}{dt}$	the time derivatives of the pressure and the velocity
$\mathsf{P}_{u,v}(\cdot)$, $\mathsf{p}_{u,v}(\cdot)$	the frequency-domain and the time-domain pressure field coefficients
$\mathsf{V}_{u,v}(\cdot)$, $\mathsf{v}_{u,v}(\cdot)$	the frequency-domain and the time-domain velocity field coefficients
$\mathsf{A}_{u,v}(\cdot)$	the frequency-domain Hankel function coefficients
$\mathsf{B}_{u,v}(\cdot)$	the frequency-domain Bessel function coefficients
$\delta(\cdot)$	the delta function
$U(\cdot)$	the unit-step function
ANC	active noise control
PNC	passive noise control
DSP	digital signal processing
SFS	sound field separation
$\mathcal{F}(\cdot)$, $\mathcal{F}^{-1}(\cdot)$	the Fourier transform and the inverse Fourier transform
STFT	short-time-Fourier-transform
2D, 3D	two-dimensional, three-dimensional

Contents

Chapter 1

Introduction

1.1 Background

In our everyday lives, we are always surrounded by noise. Long time exposure to noise can cause auditory fatigue and downgrade the working efficiency of a person. Short time exposure to loud noise, on the other hand, can damage the person's hearing.

To protect people from harmful noise, researchers have been actively developing noise control systems, and currently there are two kinds of noise control systems: the passive noise control (PNC) system and the active noise control (ANC) system. A PNC system uses sound absorbing materials to reduce noise, and is most effective at the high frequency range, i.e., $f > 1000$ Hz, because the sound absorbing material works best when the thickness of the material is comparable with the wavelength of noise. At the low frequency range, i.e., $f < 1000$ Hz, where the wavelength of noise is large, the PNC system is uneconomical as we have to use bulky sound absorbing materials [1].

An ANC system, on the other hand, uses a secondary noise to destructively interfere with the original noise. The destructive interference requires the secondary noise and the original noise to be perfectly aligned on time at the control point (or within the control region). This requirement makes the ANC system most effective at the low frequency range, i.e., $f < 1000$ Hz, where the original noise and the secondary noise can align over a larger region [1].

Though the PNC system and the ANC system cancel noise from different ways, the two systems can be integrated as a single system, which reduces noise over a broad frequency band. The noise cancellation headphone is one example of such integrations. The Bose QuietComfort 20, the Sony WH-1000XM3, the recently released Apple Airpod Pro, and many other noise cancellation headphones constitute a large market and protect millions of people from the noise pollution.

Tonal noises vs Random noises

Here, we present the definitions of tonal noises and random noises, which will be frequently referred to hereafter. Tonal noises are the stationary periodic noises whose energies are concentrated at harmonic frequencies [1]. Random noises are the non-stationary or quasi-stationary noises, whose energies are distributed over a broad frequency band (e.g., [50, 1000] Hz) [1]. The repeatable sounds generated by engines and the tire-rolling noises in cars are typical examples of tonal noises and random noises, respectively. Note that in the field of noise cancellation, we conventionally classify a noise as a narrowband noise or a broadband noise [1]. However, there are different definitions for narrowband and broadband in other fields. To avoid misunderstandings, in this thesis, we refer narrowband noises and broadband noises as tonal noises and random noises, respectively.

1.2 Problem statement

The tire rolling noise in cars, the ventilation noise at workplaces, and the engine noise outdoors, i.e., noise fields over space, are prevalent in our everyday lives and account for the majority of noises we encounter. However, the existing spatial ANC system have a hard time controlling such noises, which damage our hearing.

This leads to the main research problem of this thesis:

"We wish to control the noise in space, specifically in a room, and create a global quiet environment for people."

Note that, in this thesis, we deal with ANC problems, and thus we focus on noises below 1000 Hz, or the low frequency part of noises.

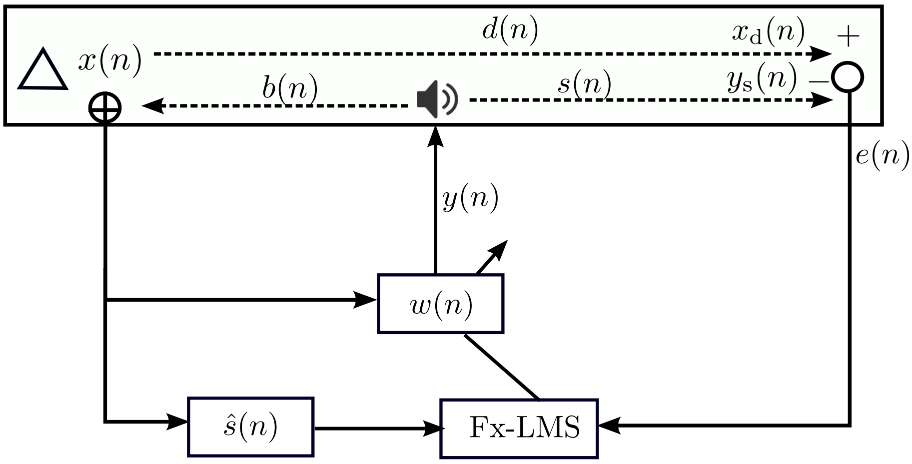


Figure 1.1: The layout of a single channel ANC system in a duct: We mark the primary source by $\triangle$, the secondary source by 🔊, the reference sensor by $\oplus$, and the error sensor by $\bigcirc$. The secondary path and its estimation are denoted as $s(n)$ and $\widehat{s}(n)$, respectively. The primary path and the feedback path are denoted as $d(n)$ and $b(n)$, respectively. The secondary source driving signal $y(n)$ is the convolution between the reference signal $x(n)$ and the filer $w(n)$. The system tries to minimize the error signal $e(n) = x_{\mathrm{d}}(n) - y_{\mathrm{s}}(n)$, where $x_{\mathrm{d}}(n)$ is the primary noise and $y_{\mathrm{s}}(n)$ is the secondary noise.

1.3 An overview of spatial ANC problems

In this section, we first use the single channel ANC system in a duct as an example to introduce the basic terminology of ANC. We then overview spatial ANC problems.

Figure 1.1 shows the layout of a single channel ANC system in a duct. We mark the primary source by $\triangle$, the secondary source by 🔊, the reference sensor by $\oplus$, and the error sensor by $\bigcirc$. The secondary path and its estimation are denoted as $s(n)$ and $\widehat{s}(n)$, respectively. The primary path and the feedback path are denoted as $d(n)$ and $b(n)$, respectively. The secondary source driving signal $y(n)$ is the convolution between the reference signal $x(n)$ and the filter $w(n)$. A filtered-x least mean square (LMS, Fx-LMS) algorithm is typically used to update the filter coefficients. The system tries to minimize the error signal $e(n) = x_{\mathrm{d}}(n) - y_{\mathrm{s}}(n)$, where $x_{\mathrm{d}}(n) = x(n) * d(n)$ is the primary noise, $y_{\mathrm{s}}(n) = y(n) * s(n)$ is the secondary noise, and $*$ is the convolution operation.

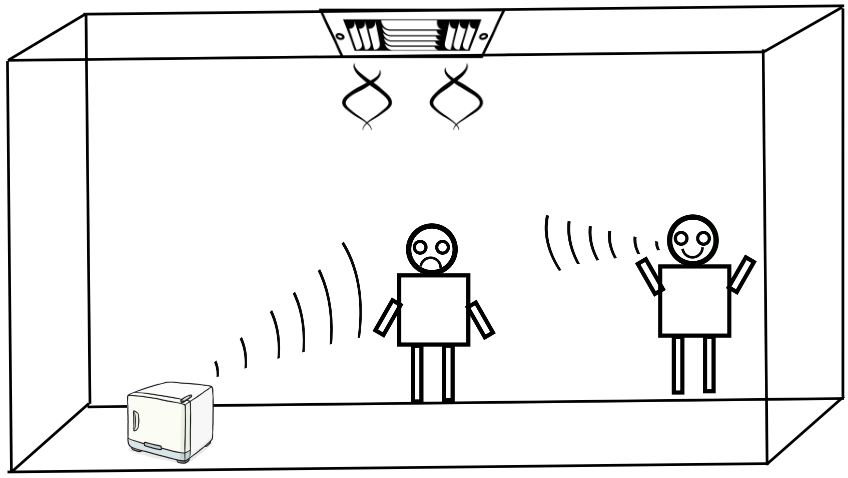


Figure 1.2: The spatial noise field in a room environment is the complicated interaction between a number of noise sources with the room environment.

We present a spatial noise field in a room environment in Fig. 1.2. The spatial noise field is the complicated interaction between a number of noise sources with the room environment, both of which can be non-stationary and time-varying. This fact makes the active control of the spatial noise field challenging. We present the spatial ANC problems in the rest of this section.

1. **Reference signal generation**

   An ANC system uses a secondary noise, which is of the same amplitude but opposite phase with the primary noise, to destructively interfere with the primary noise. To design the secondary source driving signal, we need information, i.e., a reference signal, of the primary noise.

   For a single channel ANC system as shown in Fig. 1.1, we have one reference signal and one primary noise. The coherence between them is given by [1]

$$C_{\mathrm{dx}}(\omega) = \frac{S_{\mathrm{dx}}(\omega)}{\sqrt{S_{\mathrm{dd}}(\omega) S_{\mathrm{xx}}(\omega)}}, \tag{1.1}$$

   where $S_{\mathrm{dx}}(\omega)$ is the complex cross-power between the reference signal $x(n)$ and the primary noise $x_{\mathrm{d}}(n)$, $S_{\mathrm{dd}}(\omega)$ and $S_{\mathrm{xx}}(\omega)$ are the auto-power spectra of the primary noise and the reference signal, respectively. Equation (1.1) determines the maximum achievable noise cancellation level for a single channel ANC system [1]. In the followings, we introduce ways to generate the reference signal for a single channel ANC system, and then explain the challenges of reference signal generation for a spatial ANC system.

The primary noise is tonal if it is generated by a rotating primary source. In this case, we can use non-acoustic sensors to detect the rotating frequency of the primary source. We generate the sinusoidal reference signal based on the measured frequency and the harmonic frequencies [1]. The artificially generated reference signal is fully coherent with the primary noise, and thus allows a significant reduction of the primary noise [1].

In case the primary noise is random, we normally do not use non-acoustic sensors. They may detect the non-acoustic characteristics of the primary source [2,3], and thus provide reference signals which do not fully characterize the primary noise. We use acoustic sensors which are placed close to the primary source to measure the primary source output, and use it as the reference signal. The problem with the acoustic reference sensors is that they will pick up background noises, which will downgrade the coherence between the reference signal and the primary noise.

Based on the feedback control principle, it is possible to generate the reference signal using the error sensor measurements, which consist of both the primary noise and the secondary noise [1,3]. Nonetheless, the controllable bandwidth of a feedback ANC system is limited by the distance between the secondary source and the error sensor [3]. Furthermore, we have to trade the noise cancellation performance of a feedback ANC system for its stability [1,3].

We can see from the above that it is difficult to generate the reference signal for a single channel ANC system unless the primary noise is tonal. However, it is even more challenging to generate the reference signal for a spatial ANC system in a room environment. As shown in Fig. 1.2, the spatial noise field is the interaction of the room environment with a number of noise field components, which can be tonal or random, and the source of each component can be in the nearfield or in the farfield. We may have to use non-acoustic sensors and acoustic sensors which are placed in the nearfield or in the farfield to provide a number of reference signals. We also need to derive a spatial coherence function similar to Eq. (1.1), which can guide us in designing the spatial ANC system.

To sum up, it is difficult to generate the reference signals for spatial ANC systems [4–7].

2. **Acoustic feedback**

   The acoustic feedback problem is part of the reference signal generation problem. It is a barely explored topic in the field of spatial ANC, and thus warrants a separate discussion.

   As shown in Fig. 1.1, the secondary source will radiate sounds not only to the error sensor, but also to the reference sensor, which provides the reference signal. This reference sensor-secondary source feedback loop can make the whole ANC system unstable [1]. This problem becomes even more severer for a spatial ANC system, because the feedback paths can be long and time-varying [8]. A spatial ANC system, which has not taken the acoustic feedback problem into consideration will not achieve its designed ANC performance [9].

   Note that in case the primary noise is tonal, we do not have to use the acoustic reference sensor, and thus the acoustic feedback will not be a problem.

3. **Real-time secondary path estimation**

   The secondary noise is expected to destructively interfere with the primary noise. However, the secondary noise may not work as expected, because in a spatial environment, such as in a room or in a car, the secondary path drifts slowly with the varying temperature and changes abruptly with the opening of doors [10]. The variations of the secondary path will cause the misalignment between the secondary noise and the primary noise. Thus, we need to estimate the secondary path in real-time and modify the secondary source driving signal accordingly [11–14]. However, real-time secondary path estimation is challenging for a spatial ANC system, which can have multiple time-varying long secondary paths [15].

   We need to develop novel real-time secondary path estimation methods, which exploit the spatial noise field characteristics [16, 17].

4. **Time-domain spatial algorithms**

   The alignment of the secondary noise and the primary noise requires the overall latency of the ANC system, including reference signal generation, transformation of the reference signal to the secondary source driving signal, and propagation of the secondary source driving signal through the secondary path, to be the same as the acoustic-path propagation time from the reference sensor to the error sensor [1]. This is the causal control constraint for an ANC system [1]. For most ANC systems the acoustic-path propagation time is small, and thus we need to design the ANC system with low latency.

   Most of the control algorithms for existing spatial ANC systems are frequency-domain algorithms. They use the short-time-Fourier-transform (STFT) to transform the time-domain acoustic quantities into the time-frequency-domain and vice-versa. The STFT process unavoidably introduces the frame delay, which is one of the main contributions of latency to a spatial ANC system. The frame delay probably makes the system violate the causal control constraint. Thus to control a spatial noise field, it is desirable to develop low latency time-domain spatial algorithms.

   Note that in case the primary noise only contains tonal components, which are predictable, the low latency requirement for the ANC system may be alleviated. We can add latency to the control algorithm such that the secondary noise aligns with the primary noise [1].

5. **Selective cancellation**

   A spatial ANC system probably uses a number of secondary sources and sensors. This leads to the possibility of selective noise cancellation, which can be direction-oriented or source-oriented.

   For direction-oriented selective cancellation, we cancel noise from some particular directions without canceling sound from other directions. For source-oriented selective cancellation, we cancel noise from some sources without canceling sound from other sources. This can be explained by Fig. 1.2, where we wish to cancel the noise form the ventilation system or the micro oven and preserve the human voices.

By selective cancellation, we can enhance the human hearing experience in a noisy environment. We need to develop novel real-time source separation methods for separating noise from sound-noise mixture [18]. This is an interesting and promising research direction [19, 20].

6. **High frequency noise cancellation**
   For a conventional noise control system, it is the integration of the ANC system and the PNC system that results in noise cancellation over a broad frequency band. However, the integration is not always possible. For drones, it is not suitable for arranging bulky sound absorbing materials around them. We may have to use a spatial ANC system to control noise at the high frequency range for them.

   Active control of high frequency noise is a new and challenging problem, but not impossible [21, 22].

7. **Novel sound field description theory**
   Most of the existing spatial ANC systems decompose the noise field into some basic functions, such as room modes or spherical modes, and controlling the noise field by manipulating the decomposition coefficients. The decompositions are appropriate for the noise field in an empty rectangular room or on a spherical surface, but may not be appropriate when the noise source or the noise field are of irregular shapes, such as the noise propagating through windows [23–27], the drone noises [28], and the noise around a scattering object [29].

   We may have to develop novel sound field description theory, which is tailored to the geometry of sound sources or noise fields.

8. **Hardware design**
   A spatial ANC system involves a number of sources and sensors, and requires the overall system latency to be small. This requests for an embedded platform, which support multiple channels, i.e., $\geq$ 16 input channels and $\geq$ 8 output channels, with low latency, i.e., $\leq$ 50 μs, and good computational capacity, i.e., $\geq$ 100 million floating-point multiply operations per second.

Embedded platforms that meet such requests are still barely available on the market [30].

Furthermore, to accurately characterize the primary source and evaluate the noise cancellation performance, we may have to design novel microphone arrays to measure the spatial sound field [31–34]. We may need the arrays to measure not only the sound pressure but also the sound velocity or intensity [35].

1.4 Thesis outline

This thesis consists of two parts. For the first part, we propose tonal and random spatial ANC systems, which can cancel the noise fields in a room globally. We develop a time-domain sound field separation (SFS) method, which is used by the random ANC system for outgoing and incoming SFS. For the second part, we design multiple circular arrays, which allows the analysis of a 3D sound field by a frequency-domain method or a time-domain method.

Chapter 2: Background theory and literature review
We introduce the classical single channel ANC systems. We overview the spatial ANC systems based on room modes, spherical modes, the Helmholtz integral, and identify their shortcomings.

Chapter 3: Active control of tonal outgoing noise fields in rooms
This chapter develops a spatial ANC system that can cancel tonal noise fields produced by a primary source in a room. The problem of tonal noise field control is formulated as estimating and canceling the outgoing field on a sphere surrounding the primary source. The proposed system limits the energy of the primary source radiating out of the sphere, thereby creating a global quiet zone inside the room. In addition, it removes the need for real-time secondary path estimation with negligible influence on the desired field in the room. A method for estimating the outgoing field on a sphere is presented, together with a wave-domain algorithm for controlling the outgoing field. Simulations and hardware demonstrations show the

proposed system can reduce tonal noise fields in a room and over a wide frequency range.

Chapter 4: Real-time separation of non-stationary sound fields on spheres
The SFS method presented in Chap. 3 is a frequency-domain method. Because of the frame delay introduced by the STFT process, the SFS method is best suitable for separating stationary sound field. In this chapter, we derive the corresponding time-domain SFS method, which can separate the non-stationary sound field on a sphere in real-time. The method decomposes the pressure and the radial particle velocity measured on the sphere into coefficients, and recoveries the outgoing field through the convolution between the decomposition coefficients and the theoretically derived spatial filters. Simulations show the proposed method can separate non-stationary sound fields both in free field and in room environments, and over a longer duration with small errors.

Chapter 5: Active control of random outgoing noise fields in rooms
In this chapter, based on the idea of outgoing noise field cancellation presented in Chap. 3 and the time-domain SFS method derived in Chap. 4, we develop a spatial ANC system that can cancel random outgoing noise fields in rooms. The proposed system decomposes the noise field on a sphere surrounding the noise sources into spherical modes, exploiting their directionality to generate the reference signals and remove the need for real-time secondary path estimation. Simulation results demonstrate that the proposed system can cancel a random noise field globally in a room without suffering from the secondary source feedback problem.

Chapter 6: Multiple circular arrays of vector sensors for 3D sound field analysis
We propose multiple circular arrays of vector sensors for analyzing the three-dimensional (3D) sound field. By exploiting the fact that the dimensionality of a sound field is limited by its size and the highest frequency component, the designed arrays analyze the sound field around the arrays by either a frequency-domain method or a time-domain method. The time-domain method further reduces the latency of the sound field analysis process. The proposed arrays can provide small

sized wearable devices with real-time 3D sound field analysis capacity. The effectiveness of the proposed arrays and the corresponding sound field analysis methods are confirmed by simulations.

Chapter 7: Conclusions and future directions

This chapter summarizes this thesis and points out future directions to work on.

Chapter 2

Background theory and literature review

In this chapter, we introduce the classical single channel ANC systems. We introduce room modes, spherical modes, the Helmholtz integral, and the spatial ANC systems based on them.

Here, we present the coordinate systems and the free-space Green functions, which will be referred to frequently in the rest of this thesis.

The coordinate systems

We denote the Cartesian coordinates and the spherical coordinates of a point as $\boldsymbol{x} = (x, y, z)$ and $\boldsymbol{r} = (r, \theta, \phi)$, respectively, and we denote a sphere as $\mathbb{S}$. The setup of the coordinate systems is shown in Fig. 2.1. The Cartesian coordinates and the spherical coordinates are related through the following equations:

$$x = r \sin\theta \cos\phi, \tag{2.1}$$

$$y = r \sin\theta \sin\phi, \tag{2.2}$$

$$z = r \cos\theta, \tag{2.3}$$

$$r = \sqrt{x^2 + y^2 + z^2}, \tag{2.4}$$

$$\theta = \tan^{-1} \frac{\sqrt{x^2 + y^2}}{z}, \tag{2.5}$$

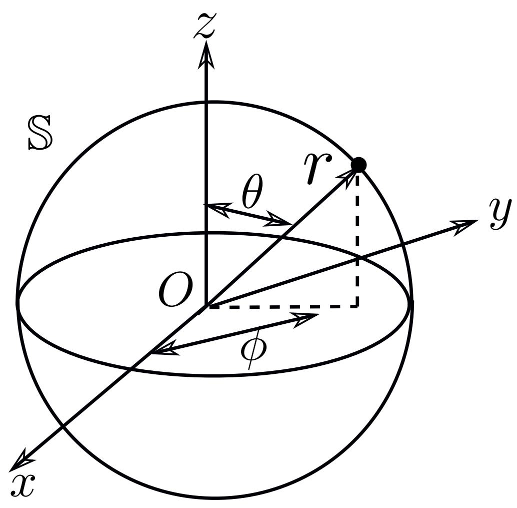


Figure 2.1: The coordinate systems: We denote the Cartesian coordinates and the spherical coordinates of a point as (x, y, z) and (r, θ, ϕ), respectively, and we denote a sphere as $\mathbb{S}$.

$$\phi = \tan^{-1} \frac{y}{x}. \tag{2.6}$$

The Green function

We denote the spherical coordinates of a point source as $\boldsymbol{r}_\mathrm{s} = (r_\mathrm{s}, \theta_\mathrm{s}, \phi_\mathrm{s})$, and the spherical coordinates of another point as $\boldsymbol{r} = (r, \theta, \phi)$. The frequency-domain and the time-domain free-space Green functions between the two points are given by [36]

$$G(\omega, \boldsymbol{r}, \boldsymbol{r}_\mathrm{s}) = \frac{e^{-i\omega \frac{||\boldsymbol{r}-\boldsymbol{r}_\mathrm{s}||_2}{c}}}{4\pi||\boldsymbol{r} - \boldsymbol{r}_\mathrm{s}||_2}, \tag{2.7}$$

and

$$G(t, \boldsymbol{r}, \boldsymbol{r}_\mathrm{s}) = \frac{\delta(t - \frac{||\boldsymbol{r}-\boldsymbol{r}_\mathrm{s}||_2}{c})}{4\pi||\boldsymbol{r} - \boldsymbol{r}_\mathrm{s}||_2}, \tag{2.8}$$

respectively, where $i = \sqrt{-1}$ is the unit imaginary number, and

$$||\boldsymbol{r} - \boldsymbol{r}_\mathrm{s}||_2 = \sqrt{r^2 + r_s^2 - 2rr_s(\cos(\theta - \theta_s)\sin(\phi)\sin(\phi_s) + \cos(\phi)\cos(\phi_s))}, \tag{2.9}$$

is the distance between the two points. The free-space Green functions govern the

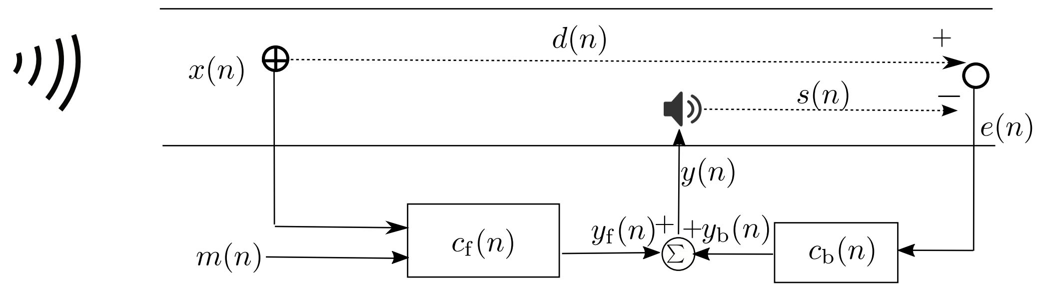


Figure 2.2: The structure of a noise cancellation headphone: $\oplus$ is the reference sensor, 🔊 is the secondary source, and $\bigcirc$ is the error sensor, $d(n)$ and $s(n)$ are the primary path and the secondary path, respectively, $x(n)$, $y(n)$, and $e(n)$ are the reference signal, the secondary source driving signal, and the error signal, respectively, $c_{\mathrm{f}}(n)$ and $c_{\mathrm{b}}(n)$ are the feedforward controller and the feedback controller, respectively, and $m(n)$ is the measurement noise of the reference sensor.

acoustic propagation (the spatial decay and the temporal delay) between the point source and the other point.

2.1 The single channel ANC system

In this section, we introduce the fixed and the adaptive single channel ANC systems and constraints that limit their performance.

2.1.1 The single channel fixed ANC system

We present the structure of a noise cancellation headphone in Fig. 2.2, and use it as an example to introduce the single channel fixed ANC system.

As shown in Fig. 2.2, a noise cancellation headphone consists of the acoustic part and the electrical part. The acoustic part consists of the secondary source, the reference sensor and the error sensor. The electrical part consists of the feedback controller $c_{\mathrm{b}}(n)$ and the feedforward controller $c_{\mathrm{f}}(n)$. We denote the impulse response between the reference sensor and the error sensor as the primary path $d(n)$, and the impulse response between the secondary source driving signal and the error sensor as the secondary path $s(n)$. We denote the reference signal as $x(n)$, the reference sensor measurement noise as $m(n)$, the secondary source driving signal as

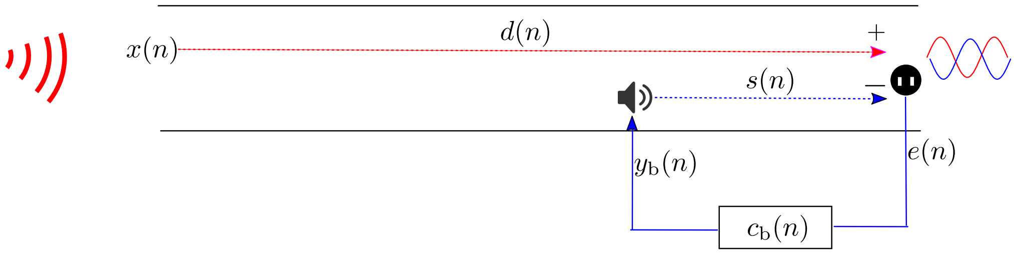


Figure 2.3: The feedback control system of a noise cancellation headphone: the primary noise $x(n) * d(n)$, the error signal $e(n)$, the secondary source driving signal $y_{\mathrm{b}}(n)$, the feedback controller $c_{\mathrm{b}}(n)$, and the secondary path $s(n)$.

$y(n)$, the feedforward part and the feedback part of the secondary source driving signal as $y_{\mathrm{f}}(n)$ and $y_{\mathrm{b}}(n)$, respectively, and the error signal as $e(n)$. We present details of the feedforward control system and the feedback control system in the followings.

The feedback control system

We present the structure of the feedback control system of a noise cancellation headphone in Fig. 2.3. As shown in Fig. 2.3, the feedback control system consists of the secondary source, the error sensor, and the feedback controller.

The feedback control system converts the error signal $e(n)$ into the secondary source driving signal $y_{\mathrm{b}}(n)$ through the feedback controller $c_{\mathrm{b}}(n)$, and uses the filtered secondary source driving signal $y_{\mathrm{b}}(n) * s(n)$ to destructively interfere with the primary noise $x_{\mathrm{d}}(n) = x(n) * d(n)$ at the error sensor.

Fundamentally, the feedback controller serves as a predictor, which predicts the future primary noise based on the current error signal [3,37–40]. The overall latency of the secondary path and the feedback controller limits the predict capability of the feedback controller. Denoting the overall latency as τ_{se}, the upper frequency limit where a significant noise cancellation performance the feedback control system can achieve is [3]

$$f_{\mathrm{m}} = \frac{1}{20}\frac{1}{\tau_{\mathrm{se}}}. \tag{2.10}$$

Based on (2.10), to control a noise up to 1000 Hz using the feedback control system, we need the overall latency to be less than

$$\tau_{\mathrm{se}} = \frac{1}{20 \times 1000} = 50\ \mu\mathrm{s}\ . \tag{2.11}$$

Such a low latency requires the distance between the secondary source and the error sensor to be small, i.e., < 1.0 cm, and specially designed low latency DSP processors to implement the feedback controller.

Based on Fig. 2.3, we have the following relationship

$$e(n) = x(n) * d(n) - e(n) * c_{\mathrm{b}}(n) * s(n), \tag{2.12}$$

or

$$e(n)[1 + c_{\mathrm{b}}(n) * s(n)] = x_{\mathrm{d}}(n). \tag{2.13}$$

Based on (2.13), we define the closed loop transfer function between the primary noise $x_{\mathrm{d}}(n)$ and the error signal $e(n)$ as [37–40]

$$\begin{aligned} L_c(\omega) &= \frac{E(\omega)}{X_{\mathrm{d}}(\omega)} \\ &= \frac{1}{1 + C_{\mathrm{b}}(\omega)S(\omega)}, \end{aligned} \tag{2.14}$$

where $E(\omega)$, $X_{\mathrm{d}}(\omega)$, $C_{\mathrm{b}}(\omega)$ and $S(\omega)$ are the frequency-domain error signal, primary noise, feedback controller, and secondary path, respectively. Equation (2.14) denotes the reduction of the error signal with respect to the primary noise, or the feedback ANC system performance.

The feedforward control system

We present the structure of the feedforward control system of a noise cancellation headphone in Fig. 2.4. The feedforward control system consists of the secondary source, the reference sensor, the error sensor, and the feedforward controller.

The feedforward control system converts the reference signal $x(n)$ into the secondary source driving signal $y_{\mathrm{f}}(n)$ through the feedforward controller $c_{\mathrm{f}}(n)$, and

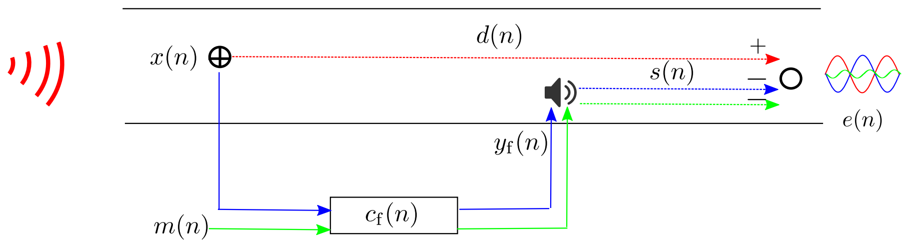


Figure 2.4: The feedforward control system of a noise cancellation headphone: the reference signal $x(n)$, the secondary source driving signal $y_{\mathrm{f}}(n)$, the error signal $e(n)$, the primary path $d(n)$, the secondary path $s(n)$, the feedforward controller $c_{\mathrm{f}}(n)$, and the measurement noise of the reference sensor $m(n)$.

uses the filtered secondary source driving signal $y_{\mathrm{f}}(n) * s(n)$ to destructively interfere with the primary noise $x_{\mathrm{d}}(n) = x(n) * d(n)$ at the error sensor [41, 42].

While the performance of the feedback system is limited by the overall latency of the secondary path and the feedback controller, the performance of the feedforward system is limited by the measurement noise of the reference sensor [43]. Based on Fig. 2.4, we have the following relationship

$$\begin{aligned} e(n) &= x(n) * d(n) - (x(n) + m(n)) * c_{\mathrm{f}}(n) * s(n) \\ &= x(n) * (d(n) - c_{\mathrm{f}}(n) * s(n)) - m(n) * c_{\mathrm{f}}(n) * s(n). \end{aligned} \tag{2.15}$$

Equation (2.15) reveals that if we design the feedforward filter as

$$c_{\mathrm{f}}(n) = \mathcal{F}^{-1}\left[\frac{D(\omega)}{S(\omega)}\right], \tag{2.16}$$

where $D(\omega)$ and $S(\omega)$ are the transfer functions of the primary path and the secondary path, respectively, and $\mathcal{F}^{-1}$ is the inverse Fourier transform, we can cancel the primary noise at the error sensor, but unavoidably we introduce an additional noise to the error sensor, i.e., $e(n) = -m(n) * c_{\mathrm{f}}(n) * s(n)$. The ideal noise cancellation performance of the feedforward control system is given by the reference signal to measurement noise power ratio [1]

$$\frac{||x(n) * d(n)||_2^2}{||m(n) * c_{\mathrm{f}}(n) * s(n)||_2^2} = \frac{||x(n)||_2^2}{||m(n)||_2^2}. \tag{2.17}$$

If the reference signal to measurement noise power ratio is less than 1, we will add more noises to the error sensor.

Note that the secondary source driving signal can propagate upstream and get detected by the reference sensor, causing the acoustic feedback problem [1]. However, for a well designed noise cancellation headphone, the gain of the acoustic feedback is small, and thus can be neglected.

In the above, we introduced the feedback system and the feedforward system for a noise cancellation headphone. Due to its limited computational resources a noise cancellation headphone normally implement a fixed feedback controller $c_{\mathrm{b}}(n)$ and a fixed feedforward controller $c_{\mathrm{f}}(n)$ [43]. However, the impulse responses of the primary path and the secondary path differ from one person to another. In the next section, we introduce the adaptive ANC system, which can account for the variations of the primary path and the secondary path, and thus achieves better ANC performance.

2.1.2 The single channel adaptive ANC system

In this section, we introduce the single channel adaptive ANC system. An adaptive ANC system can be feedback or feedforward. However, the adaptive feedback ANC system has limited applications. Because it can achieve a significant noise cancellation reduction only if the secondary source is very close to the error sensor or if the primary noise is tonal. Nonetheless, in case the primary noise is tonal the adaptive feedback system is in essence an approximation of the adaptive feedforward system [1]. Thus in this section, we will focus on the adaptive feedforward system, and specifically the tonal adaptive system and the random adaptive system.

The tonal adaptive system

We present the structure of a tonal adaptive system (the l-th harmonic component) in Fig. 2.5 [44]. We denote the primary source output as $x(n)$, which consists of a number of harmonic components of frequencies $\{\omega_l\}_{l=1}^{L}$.

The error signal at the error sensor is

$$e(n) = x_{\mathrm{d}}(n) - y_{\mathrm{s}}(n), \tag{2.18}$$

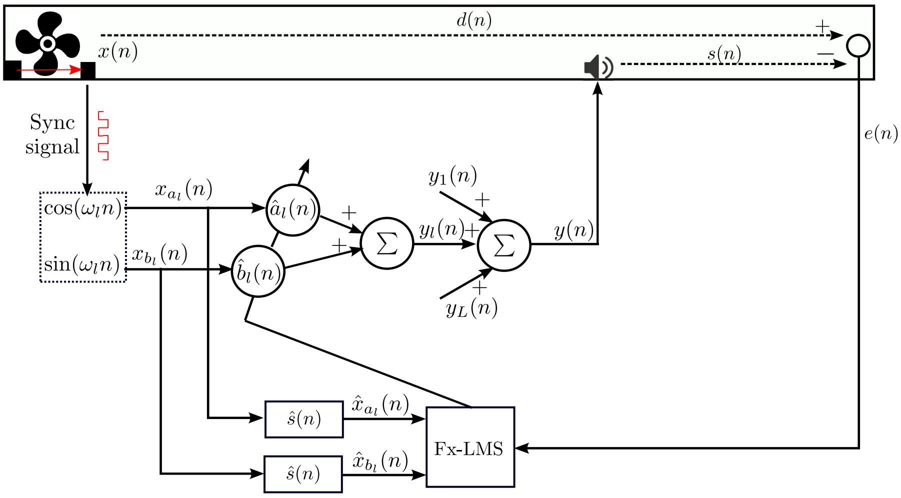


Figure 2.5: The structure of a tonal adaptive system (the l-th harmonic component): We denote the non-acoustic reference sensor as ■, which is the key of the tonal adaptive system.

where $x_\mathrm{d}(n)$ is the primary noise, and $y_\mathrm{s}(n)$ is the secondary noise. The primary noise can be expressed as

$$\begin{aligned} x_\mathrm{d}(n) &= x(n) * d(n) \\ &= \sum_{l=1}^{L} (a_l(n)\cos(\omega_l n) + b_l(n)\sin(\omega_l n)), \end{aligned} \tag{2.19}$$

where $d(n)$ is the primary path, $a_l(n)$ and $b_l(n)$ are strengths of the cos-part and the sin-part of the l-th harmonic component, respectively. The secondary noise can be expressed as

$$\begin{aligned} y_\mathrm{s}(n) &= y(n) * s(n) \\ &= \sum_{l=1}^{L} y_l(n) * s(n) \\ &= \sum_{l=1}^{L} (\widehat{a}_l(n) x_{a_l}(n) + \widehat{b}_l(n) x_{b_l}(n)) * s(n) \\ &= \sum_{l=1}^{L} (\widehat{a}_l(n)\cos(\omega_l n) + \widehat{b}_l(n)\sin(\omega_l n)) * s(n), \end{aligned} \tag{2.20}$$

where $y(n)$ is the secondary source driving signal, $s(n)$ is the secondary path, $y_l(n)$ is the l-th harmonic component of the secondary source driving signal, $x_{a_l}(n) = \cos(\omega_l n)$ and $x_{b_l}(n) = \sin(\omega_l n)$ are the reference signals for the l-th harmonic component, and are generated based on the synchronization signal provided by the non-acoustic sensor, $\widehat{a}_l(n)$ and $\widehat{b}_l(n)$ are strengths for the reference signals.

To minimize the error signal, we first define a cost function

$$J(n) = e^2(n). \tag{2.21}$$

Taking partial derivative of the cost function with respect to the reference signal strengths, we obtain

$$\frac{\partial J(n)}{\partial \widehat{a}_l(n)} = -e(n)x_{a_l}(n) * \widehat{s}(n), \tag{2.22}$$

$$\frac{\partial J(n)}{\partial \widehat{b}_l(n)} = -e(n)x_{b_l}(n) * \widehat{s}(n). \tag{2.23}$$

We update the reference signal strengths according to the partial derivatives

$$\widehat{a}_l(n+1) = \widehat{a}_l(n) + \mu_l e(n)\widehat{x}_{a_l}(n), \tag{2.24}$$

$$\widehat{b}_l(n+1) = \widehat{b}_l(n) + \mu_l e(n)\widehat{x}_{b_l}(n), \tag{2.25}$$

where μ_l is the step size parameter, and

$$\widehat{x}_{a_l}(n) = x_{a_l}(n) * \widehat{s}(n), \tag{2.26}$$

$$\widehat{x}_{b_l}(n) = x_{b_l}(n) * \widehat{s}(n), \tag{2.27}$$

are the reference signals filtered by $\widehat{s}(n)$, an estimation of the secondary path. This is the Fx-LMS algorithm, the most widely used filter updating algorithm for the single channel ANC system [1].

The random adaptive system

We present the structure of a random adaptive system in Fig. 2.6. The primary path is denoted $d(n)$. The secondary path and its estimation are denoted as $s(n)$

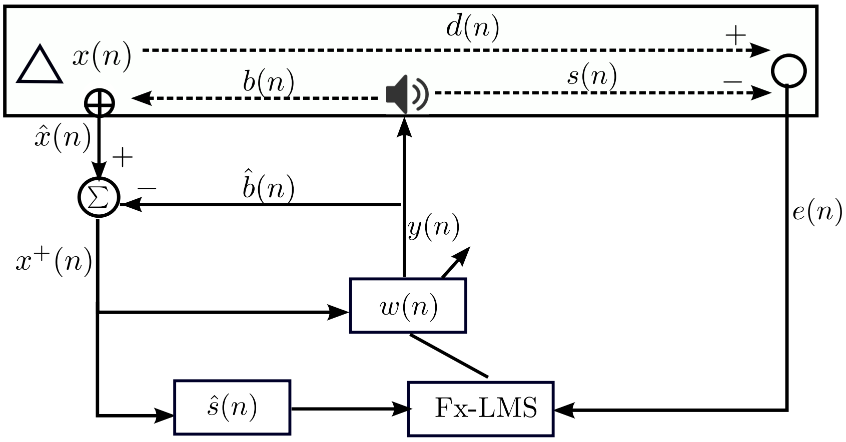


Figure 2.6: The structure of a random adaptive system: The primary path is denoted $d(n)$. The secondary path and its estimation are denoted as $s(n)$ and $\widehat{s}(n)$, respectively. The feedback path and its estimation are denoted as $b(n)$ and $\widehat{b}(n)$, respectively.

and $\widehat{s}(n)$, respectively. The feedback path and its estimation are denoted as $b(n)$ and $\widehat{b}(n)$, respectively. The system tries to minimize the error signal $e(n)$.

As shown in Fig. 2.6, for a random adaptive system, the reference signal generation process is much more complicated than the tonal adaptive system. We denote the primary source output as $x(n)$. The reference sensor measurement is given by

$$\widehat{x}(n) = x(n) + y(n) * b(n), \tag{2.28}$$

which contains the convolution between the secondary source driving signal $y(n)$ and the feedback path $b(n)$. If $\widehat{x}(n)$ is used as the reference signal directly, the system can be unstable because of the reference signal-secondary source feedback loop. Thus we obtain $\widehat{b}(n)$, an estimation of the feedback path, and use

$$\begin{aligned} x^+(n) &= \widehat{x}(n) - \widehat{y}(n) * b(n) \\ &= x(n) + y(n) * [b(n) - \widehat{b}(n)], \end{aligned} \tag{2.29}$$

as the reference signal. In the above equation, we normally have $b(n) \neq \widehat{b}(n)$. The residual feedback $y(n) * [b(n) - \widehat{b}(n)]$ reduces the coherence between the reference signal $x^+(n)$ and the primary noise $x_{\mathrm{d}}(n)$, and limits the maximum achievable random noise cancellation level.

The error signal can be expressed as

$$\begin{aligned} e(n) &= x_{\mathrm{d}}(n) - y_{\mathrm{s}}(n) \\ &= x(n) * d(n) - y(n) * s(n) \\ &= x(n) * d(n) - x^{+}(n) * w(n) * s(n), \end{aligned} \tag{2.30}$$

where $w(n)$ is the adaptive filter. To minimize the error signal $e(n)$, we update the adaptive filter $w(n)$ based on the the Fx-LMS algorithm as follows

$$w(n+1) = w(n) + \mu e(n) x_s^{+}(n), \tag{2.31}$$

where μ is the step size parameter, and

$$x_s^{+}(n) = x^{+}(n) * \widehat{s}(n) \tag{2.32}$$

is the reference signal filtered by the secondary path estimation, $\widehat{s}(n)$.

Here are our comments on the adaptive ANC systems:

1. **Reference signal generation**
   As shown in Fig. 2.5, for a tonal adaptive system, we can use non-acoustic sensors to detect the rotating frequency of the primary source. We artificially generate the sinusoidal reference signal based on the measured frequency and its harmonic frequencies. This reference signal is free from the background noise and the secondary source feedback, and thus is highly coherent with the primary noise. The high coherence allows a significant reduction of the tonal primary noise.

   For a random adaptive system, we normally use an acoustic sensor placed close to the primary source as the reference sensor. Unavoidably, the acoustic reference sensor will detect the background noise and the secondary source feedback, which will downgrade the coherence between the reference sensor measurement and the primary noise, and may make the system unstable [1]. We have to estimate the feedback path, and compensate for its influence on the reference sensor measurement as (2.29). The accuracy of the feedback path estimation limits the noise cancellation performance of the random adaptive system.

2. **Adaptive filter updating**
The Fx-LMS algorithm, specifically the LMS algorithm, is normally used as the filter updating algorithm due to its simplicity and robustness. However, the filter updating speed can be slow. To speed up the filter updating, we can use the variable step size parameter LMS algorithm [45, 46], the normalized LMS (NLMS) algorithm [47], the affine projection adaptive algorithm [48], the recursive least square algorithm [47], or the proportionate NLMS algorithm [49–51].

2.1.3 Summary of the single channel ANC systems

In the previous sections, we introduced the single channel ANC systems, which are used for canceling noises in headphones and in ducts. The introduction is not exhaustive, and interested readers should refer to the reference for a comprehensive overview of this field [1,3,52–55]. Here we provide a summary of the single channel ANC systems:

1. **The feedback**
The application of feedback ANC systems is limited, because they work best when the distance between the secondary source and the error sensor is very small or if the primary noise is tonal [1].

   Hereafter we focus on the feedforward ANC systems only.

2. **The adaptation**
By adjusting the filter coefficients with respect to the primary path and the secondary path, the adaptive ANC system tends to achieve better noise cancellation performance than the fixed ANC system [43]. Nonetheless, an adaptive ANC system can be unstable if the step size parameter is not chosen properly [1, 52].

3. **The reference signal**
The performance of a feedforward ANC system is fundamentally determined

by the reference signal to measurement noise power ratio, or the coherence between the reference signal and the primary noise [1]. To achieve a good noise cancellation performance, a feedforward system must reduce the reference sensor measurement noise and the secondary source feedback as much as possible.

4. **The causal control constraint**
   The destructive interference of the primary noise and the secondary noise requires the perfect alignment of them at the error sensor. The alignment further requires the overall latency of the ANC system to be the same as the acoustic propagation time from the reference sensor to the error sensor. This is the causal control constraint for an ANC system.

   For most practical ANC systems, the acoustic propagation time from the reference sensor to the error sensor is small, and thus the systems normally process the signals in the time-domain to lower the latency. The frequency-domain algorithms can introduce additional frame delay, which may make the generated secondary noise unable to align with the primary noise on time.

   Note that for tonal ANC systems, by exploiting the periodicity of the primary noise, we can add delay to the secondary noise and thus it aligns with the primary noise. That is, a tonal ANC system can process the signal in the frequency-domain.

2.2 The room mode and the corresponding spatial ANC systems

2.2.1 The room mode decomposition

The acoustic wave equation in the Cartesian coordinates is given by [56]

$$\frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} + \frac{\partial^2 P}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 P}{\partial t^2}. \tag{2.33}$$

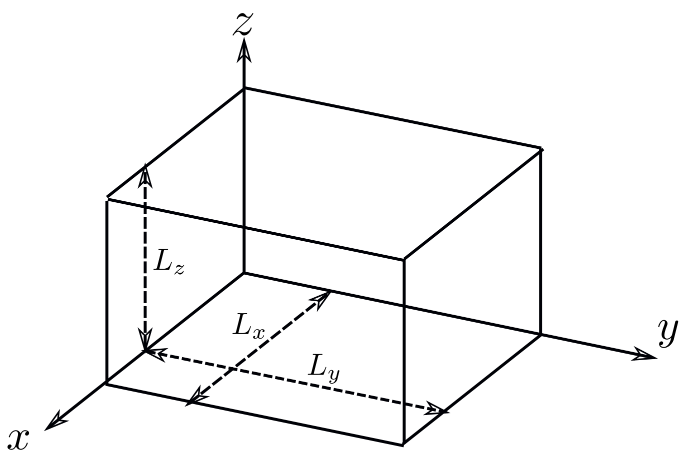


Figure 2.7: A rectangular room of size $L_x \times L_y \times L_z$.

Inside an empty rectangular room of size $L_x \times L_y \times L_z$ with rigid walls as shown in Fig. 2.7, eigen-functions of (2.33) are room modes [56]

$$\psi_{\boldsymbol{m}}(\boldsymbol{x}) = \sqrt{\epsilon_{m1}\epsilon_{m2}\epsilon_{m3}} \cos\left(\frac{m_1 \pi x}{L_x}\right) \cos\left(\frac{m_2 \pi y}{L_y}\right) \cos\left(\frac{m_3 \pi z}{L_z}\right), \tag{2.34}$$

where $\boldsymbol{x} = (x, y, z)$ is the Cartesian coordinates of a point, $\boldsymbol{m} = [m_1, m_2, m_3]$ is the mode index (three non-negative integer numbers), $\{\epsilon_{m_i}\}_{i=1}^{3}$ are mode dependent coefficients, and

$$\epsilon_{m_i} = \begin{cases} 2, & m_i > 0 \\ 1, & m_i = 0. \end{cases} \tag{2.35}$$

The eigen-value associated with a room mode is

$$w_{\boldsymbol{m}} = c\sqrt{\left(\frac{m_1\pi}{L_x}\right)^2 + \left(\frac{m_2\pi}{L_z}\right)^2 + \left(\frac{m_3\pi}{L_z}\right)^2}, \tag{2.36}$$

where c is the speed of sound. If one dimension is a multiplier of another dimension, i.e., $L_x = 2\ L_y$, there are degenerate modes [57], where two modes with different indices share the same eigen-value $w_{\boldsymbol{m}}$.

The room modes are orthogonal with respect to each other. That is,

$$\int_V \psi_{\boldsymbol{m_1}}(\boldsymbol{x})\psi_{\boldsymbol{m_2}}(\boldsymbol{x})dV = V\delta_{\boldsymbol{m_1},\boldsymbol{m_2}}, \tag{2.37}$$

where $V = L_x \times L_y \times L_z$ is the room volume, and $\delta_{\boldsymbol{m_1},\boldsymbol{m_2}}$ is the Kronecker delta function

$$\delta_{\boldsymbol{m_1},\boldsymbol{m_2}} = \begin{cases} 1, & \boldsymbol{m_1} = \boldsymbol{m_2}, \\ 0, & \boldsymbol{m_1} \neq \boldsymbol{m_2}. \end{cases} \tag{2.38}$$

The sound field in the room can be decomposed into room modes and their strengths

$$P(\omega, \boldsymbol{x}) = \sum_{\boldsymbol{m}=[0,0,0]}^{[\infty,\infty,\infty]} a_{\boldsymbol{m}}(\omega)\psi_{\boldsymbol{m}}(\boldsymbol{x}), \tag{2.39}$$

where the mode strengths are given by [56]

$$\begin{aligned} a_{\boldsymbol{m}}(\omega) = \frac{c^2}{V(\omega_{\boldsymbol{m}}^2 - \omega^2)}\Big[\int_S i\omega\rho\psi_{\boldsymbol{m}}(\boldsymbol{y})\boldsymbol{u}_i(\boldsymbol{y})\mathbf{n}dS \\ - \int_S i\frac{\omega}{c}\psi_{\boldsymbol{m}}(\boldsymbol{y})\boldsymbol{\beta}(\boldsymbol{y}) \sum_{\boldsymbol{m}=[0,0,0]}^{[\infty,\infty,\infty]} a_{\boldsymbol{m}}\psi_{\boldsymbol{m}}(\boldsymbol{y})dS\Big], \end{aligned} \tag{2.40}$$

ρ is the density of air, $\boldsymbol{y} = [y_x, y_y, y_z]$ are the Cartesian coordinates of a point on the room surface, $\boldsymbol{u}_i(\boldsymbol{y})$ is the complex acoustic surface velocity normal to the surface, the direction of $\boldsymbol{u}_i(\boldsymbol{y})$ is specified by the unit vector $\mathbf{n}$, $\boldsymbol{\beta}(\boldsymbol{y})$ is the acoustic admittance, and S stands for the room surface.

We present the shapes of several room modes, where $m_1 \in [0, 1, 2]$, $m_2 \in [0, 1, 2]$, and $m_3 = 0$, for a room with $L_x = 5$ m and $L_y = 4$ m in Fig. 2.8. As shown in Fig. 2.8 and by (2.39), the room mode decomposition of the sound field in a room is a global decomposition. That is, the sound field in the entire room is fully determined by the room modes and their strengths [56].

2.2.2 The room mode based spatial ANC systems

Instead of minimizing the error signals directly, a room mode based spatial ANC system minimizes the mode strengths, which allow the system to exploit the structure of the noise field for ANC.

The room mode concept was first introduced by Prof. Elliott, Prof. Nelson, and their co-authors to the field of ANC. In three consecutive papers [58–60],

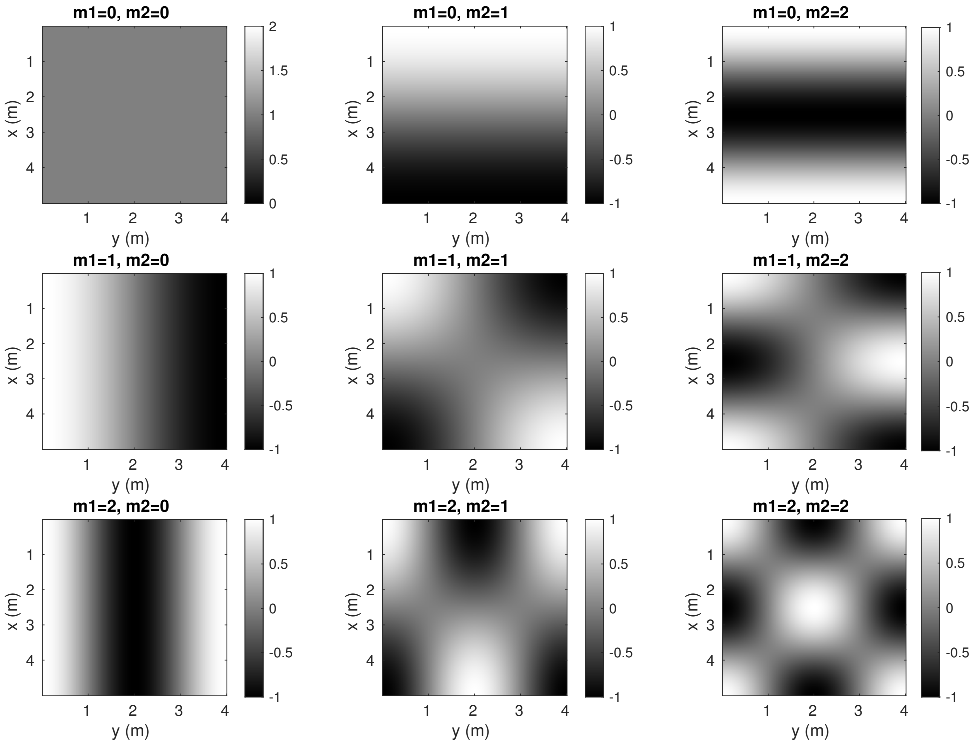


Figure 2.8: Shapes of room modes, where $m_1 \in [0, 1, 2]$, $m_2 \in [0, 1, 2]$, and $m_3 = 0$, for a room with $L_x = 5$ m and $L_y = 4$ m.

they explored the global cancellation of tonal noise fields in a shallow rectangular enclosure. They conducted pioneering ANC experiments in a plane cabin [61], explored the arrangement of sources and sensors, and the maximum achievable noise cancellation level in rooms [5, 11, 62–64] and in cars [65–67].

The room mode concept was further extended by other researchers. Bao worked on the problem of real-time secondary path estimation in a time-varying room environment [12]. Allahyar investigated the degenerate modes, the arrangement of sources and sensors and their influence on the global noise cancellation performance of a multiple-input-multiple-output ANC system inside a telephone kiosk [57, 68]. Parkin proposed to control the noise energy density instead of the noise pressure, aiming to achieve more uniform noise reduction over space [6, 69]. Geng explored the cancellation of interior noise fields in irregular shape enclosures [70].

Though a number of publications mentioned above showed promising simulation and experimental results of global (or local) noise cancellation, there are two problems with the room mode based spatial ANC systems. The first problem is the object-free assumption. The room mode decomposition of a noise field presumes that the room is empty. This assumption is not practical in a real room. The second problem is that the number of room modes grows exponentially with the frequency and the room size [57]. The greater the number of room modes is, the more sources and sensors are needed to control and measure the noise field, respectively. This fact makes the room mode based spatial ANC systems best suitable for small sized rooms and at the low frequency range [57].

2.3 The spherical mode and the corresponding spatial ANC systems

2.3.1 The spherical mode decomposition

The acoustic wave equation in the spherical coordinates is given by [36]

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial P}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial P}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 P}{\partial\phi^2}=\frac{1}{c^2}\frac{\partial^2 P}{\partial t^2}, \tag{2.41}$$

whose eigen-functions are spherical modes $h_u(\omega r/c)Y_{u,v}(\theta,\phi)$ and $j_u(\omega r/c)Y_{u,v}(\theta,\phi)$, where u and v are order and degree of a mode, respectively, $h_u(\cdot)$ and $j_u(\cdot)$ are the spherical Hankel function (the second kind) of order u and the spherical Bessel function of order u, respectively [36], $Y_{u,v}(\theta,\phi)$ are spherical harmonics [71]

$$Y_{u,v}(\theta,\phi)\equiv\mathcal{P}_{u,v}(\theta)E_v(\phi), \tag{2.42}$$

with

$$E_v(\phi)=\begin{cases}\frac{\cos(|v|\phi)}{\sqrt{\pi}}, & v>0,\\ \frac{1}{\sqrt{2\pi}}, & v=0,\\ \frac{\sin(|v|\phi)}{\sqrt{\pi}}, & v<0,\end{cases} \tag{2.43}$$

$$\mathcal{P}_{u,v}(\theta) = \sqrt{\frac{(2u+1)(u-|v|)!}{4\pi(u+|v|)!}}\mathfrak{P}_{u,|v|}(\cos\theta), \tag{2.44}$$

$\mathfrak{P}_{u,|v|}(\cdot)$ is the associated Legendre function of order u and degree $|v|$, $\mathcal{P}_{u,v}(\cdot)$ is the corresponding normalized associated Legendre function, and $(\cdot)!$ is the factorial of a non-negative number. Functions $E_v(\phi)$ are orthogonal with respect each other

$$\int_0^{2\pi} E_v(\phi)E_{\overline{v}}(\phi)d\phi = \begin{cases} 1, & v = \overline{v}, \\ 0, & v \neq \overline{v}. \end{cases} \tag{2.45}$$

The sound field on a sphere can be expressed as [36]

$$\begin{aligned} P(\omega, r, \theta, \phi) = & \underbrace{\sum_{u=0}^{\infty}\sum_{v=-u}^{u} \mathsf{A}_{u,v}(\omega)h_u(\omega\tau_r)Y_{u,v}(\theta,\phi)}_{P^{\mathrm{o}}(\omega,r,\theta,\phi)} \\ & + \underbrace{\sum_{u=0}^{\infty}\sum_{v=-u}^{u} \mathsf{B}_{u,v}(\omega)j_u(\omega\tau_r)Y_{u,v}(\theta,\phi)}_{P^{\mathrm{i}}(\omega,r,\theta,\phi)}, \end{aligned} \tag{2.46}$$

where the time-dependence of $e^{i\omega t}$ is implicitly assumed [36], $\tau_r = r/c$, $P^{\mathrm{o}}(\omega, r, \theta, \phi)$ and $P^{\mathrm{i}}(\omega, r, \theta, \phi)$ denote the outgoing sound field and the incoming sound field, respectively, $\mathsf{A}_{u,v}(\omega)$ and $\mathsf{B}_{u,v}(\omega)$ are the spherical Hankel function coefficients and the spherical Bessel function coefficients, respectively,

In the rest of this section, we preset some properties of the above mentioned functions.

The spherical Bessel function and the Legendre function

The spherical Bessel function relates with the Legendre function through the Fourier transform [72]

$$j_u(\omega\tau_r) = \mathcal{F}^{-1}\Big[\frac{i^u}{2\tau_r}\mathcal{P}_u(\frac{t}{\tau_r})\Big]. \tag{2.47}$$

We present the plots of the Legendre function $\mathcal{P}_u(\frac{t}{\tau_r})$ up to the third order in Fig. 2.9.

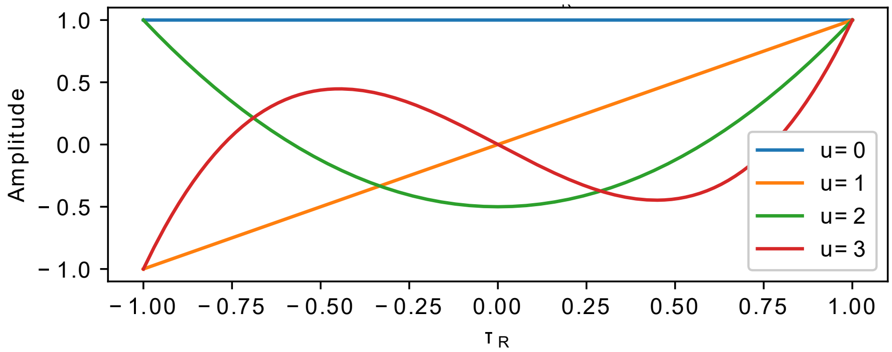


Figure 2.9: Plots of the Legendre function $\mathcal{P}_u(\frac{t}{\tau_r})$ up to the third order.

The spherical Bessel function and the spherical Hankel function

Here we present the explicit expression of the spherical Hankel function, the properties of the spherical Hankel function and the spherical Bessel function

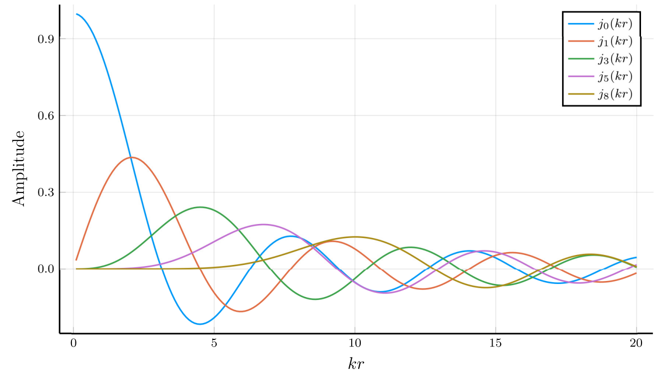
$$h_u(x) = e^{-ix} i^{u+2} \sum_{v=0}^{u} \frac{\varphi_v(u)}{(ix)^{v+1}}, \tag{2.48}$$

$$\varphi_v(u) = \frac{(u+v)!}{2^v v!(u-v)!}, \tag{2.49}$$

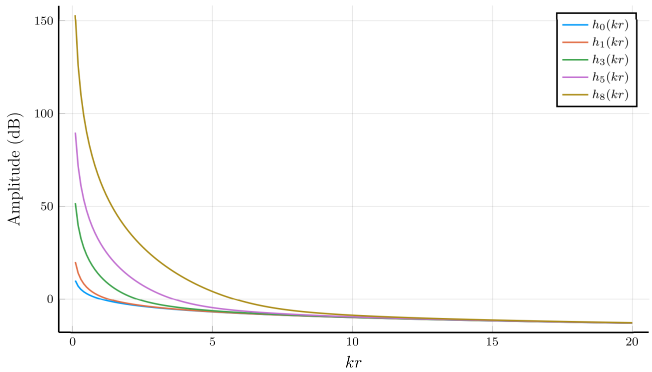
$$j_u(x) = \frac{1}{2}[h_u(x)^* + h_u(x)], \tag{2.50}$$

$$h'_u(x) = h_{u-1}(x) - \frac{u+1}{x} h_u(x), \tag{2.51}$$

$$h_{-u-1}(x) = i(-1)^{u+1} h_u(x), \quad u \geq 0, \tag{2.52}$$



(a)



(b)

Figure 2.10: The amplitudes of the spherical Bessel functions and the spherical Hankel functions up to the eighth order. The spherical Bessel functions exhibit a bandpass filter characteristic, and the spherical Hankel functions exhibit a lowpass filter characteristic.

$$j_u'(x) = \frac{u}{x} j_u(x) - j_{u+1}(x), \tag{2.53}$$

where $(\cdot)^*$ is the complex conjugate operation, and $(\cdot)'(x)$ stands for the derivative with respect the argument x.

We present the amplitudes of the spherical Bessel functions and the spherical Hankel functions in Fig. 2.10. Note that the amplitudes of the spherical Hankel functions are expressed in dB scale. As shown in Fig. 2.10, the spherical Bessel functions exhibit a bandpass filter characteristic, and the spherical Hankel functions

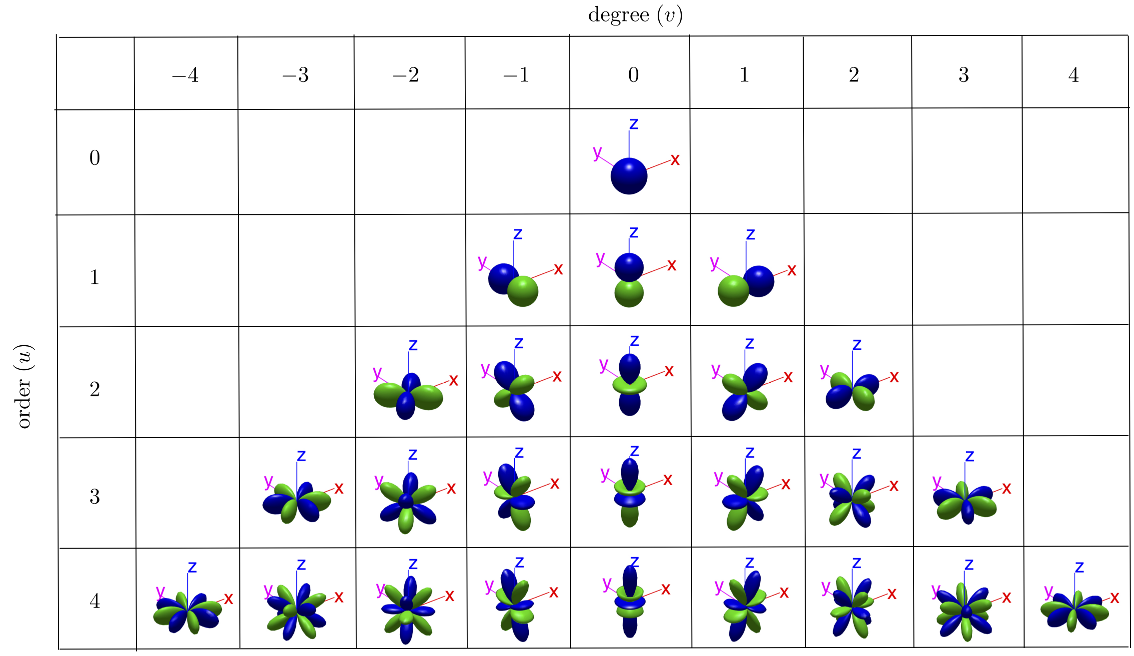


Figure 2.11: The shapes of the spherical harmonics up to the fourth order, where the positive values are colored in blue and the negative values in green [74].

exhibit a lowpass filter characteristic.

Combing the spherical Bessel functions with the spherical harmonics, we obtain the incoming modes on a sphere. We can exploit the bandpass filter characteristic to limit the number of modes needed to represent the incoming sound field [73].

The Wronskian equation

The spherical Bessel function and the spherical Hankel function are related through the Wronskian equation [72]

$$h_u(x)j'_u(x) - h'_u(x)j_u(x) = \frac{-i}{x^2}, \tag{2.54}$$

where $j'_u(\cdot)$ and $h'_u(\cdot)$ are derivatives of the spherical Bessel function and the spherical Hankel function with respect to the argument, respectively.

The spherical harmonics

The spherical harmonics are orthogonal with respect each other

$$\int_0^{2\pi}\int_0^{\pi} Y_{u,v}(\theta,\phi)Y_{\overline{u},\overline{v}}(\theta,\phi)\sin\theta d\theta d\phi = \begin{cases} 1, & (u,v)=(\overline{u},\overline{v}), \\ 0, & (u,v)\neq(\overline{u},\overline{v}). \end{cases} \quad (2.55)$$

A continuous and integrable function $P(\theta,\phi)$ over a sphere can be decomposed into the spherical harmonics and the corresponding strengths

$$P(\theta,\phi) = \sum_{u=0}^{\infty}\sum_{v=-u}^{u} P_{u,v}Y_{u,v}(\theta,\phi), \quad (2.56)$$

where the strengths are recovered based on the orthogonal relationship among the spherical harmonics [71]

$$P_{u,v} = \int_0^{2\pi}\int_0^{\pi} Y_{u,v}(\theta,\phi)P(\theta,\phi)\sin\theta d\theta d\phi. \quad (2.57)$$

For practical implementations, we can replace the continuous integral over the sphere by a finite sampling

$$P_{u,v} \approx \sum_{q=1}^{Q} \gamma_q Y_{u,v}(\theta_q,\phi_q)P(\theta_q,\phi_q), \quad (2.58)$$

where $\{\gamma_q\}_{q=1}^{Q}$ are sampling weights, and $\{\theta_q,\phi_q\}_{q=1}^{Q}$ are spherical coordinates of the sampling points [75]. To accurately recover the strengths up to order u, we require $Q > (u+1)^2$.

We present shapes of spherical harmonics up to the fourth order in Fig. 2.11, which shows that the higher the order is the more complex the shape of a spherical harmonic is. A spherical harmonic is symmetric with respect to xy-plane if $u+v$ is even, and is anti-symmetric with respect to the xy-plane if $u+v$ is odd.

2.3.2 The spherical mode based spatial ANC systems

The spherical modes are suitable for decomposing the sound field on or inside a sphere. This decomposition does not have to take the environment characteristics into consideration, and makes it easier to determine the number of modes needed to represent a sound field [73].

Kempton first expanded the free-space Green function in a spherical mode fashion, and demonstrated the possibility of far-field noise cancellation using a multipole source [76]. Since then a great number of works have been done in this area. Martin expanded the primary noise field using spherical harmonics to facilitate the arrangement of secondary sources and error sensors [77]. Rafaely investigated the use of a spherical secondary source array for local ANC [78]. Waleed developed a family of algorithms for noise cancellation over a continuous zone [79]. Michael analyzed the spatial coherence of a noise field to provide a sparse representation of the noise field [80]. Chen used the spherical harmonic analysis theory to evaluate the spatial ANC system performance [81]. Zhang compared the performance of spatial ANC systems, which uses either the spherical mode coefficient or the acoustic potential energy over a region as the error signal [82]. Xie proposed to speed up the convergence and lower the complexity of a spherical mode based multichannel ANC system by adding the L1-norm constraint [83].

2.4 The Helmholtz integral and the corresponding spatial ANC systems

Apart from room modes and spherical modes based ANC systems, there is another family of spatial ANC systems, which are based on the Helmholtz integral [85].

Let V be a source-free bounded region, S_0 be the boundary, $\boldsymbol{x}_0$ be coordinates of a point on the boundary, $\boldsymbol{x}$ be coordinates of a point inside of the region, and $G(\cdot,\cdot)$ be an appropriate Green function. The Helmholtz integral relates the pressure field $P(\boldsymbol{x},\omega)$ inside of the region with the pressure $P(\boldsymbol{x}_0,\omega)$ and the pressure gradient $\frac{\partial}{\partial \boldsymbol{n}}P(\boldsymbol{x}_0,\omega)$ on the boundary

$$P(\boldsymbol{x},\omega) = -\oint_{\partial V} G(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}}P(\boldsymbol{x}_0,\omega) - P(\boldsymbol{x}_0,\omega)\frac{\partial}{\partial \boldsymbol{n}}G(\boldsymbol{x}|\boldsymbol{x}_0,\omega)\Big)dS_0, \quad (2.59)$$

where $\frac{\partial}{\partial \boldsymbol{n}}$ denotes the inward gradient in the direction of the normal vector $\boldsymbol{n}$. Based on this equation, we can recreate a pressure field $P(\boldsymbol{x},\omega)$ inside of the region V by specifying the pressure $P(\boldsymbol{x}_0,\omega)$ and the pressure gradient $\frac{\partial}{\partial \boldsymbol{n}}P(\boldsymbol{x}_0,\omega)$ on the

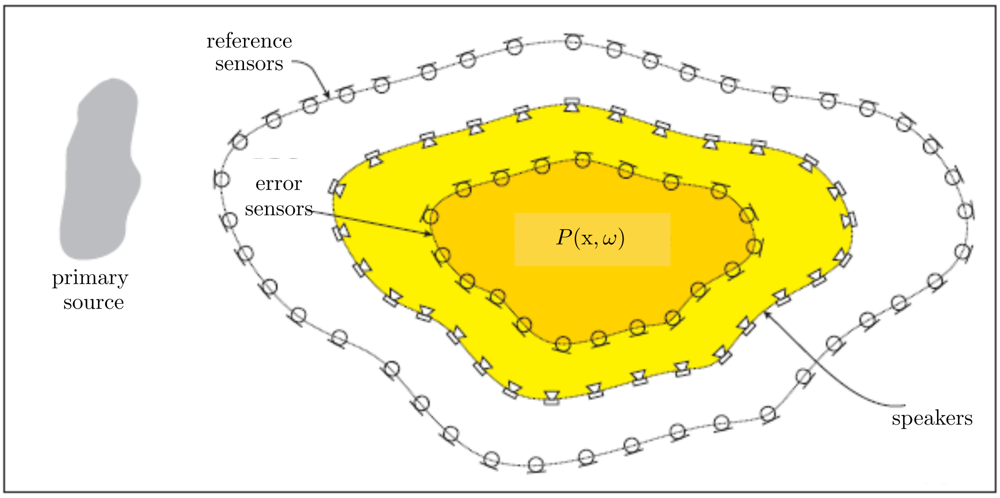


Figure 2.12: An example of the spatial ANC system based on the Helmholtz integral: We recreate an anti-noise field within the error sensor array using the speaker array [84].

boundary S_0. We can exploit this fact to design spatial ANC systems as shown in Fig. 2.12, where we use a reference sensor array to detect the primary noise field, and use the speaker array to generate an anti-noise field within the error sensor array [84, 86–89].

The advantage of the Helmholtz integral based spatial ANC systems is that they can cancel noise fields of arbitrary shapes. However, the Helmholtz integral based spatial ANC systems may require a large number of sources and sensors to implement.

2.5 The problems with existing spatial ANC systems

A spatial noise field, generally speaking, is the complicated interaction of a number of noise sources with the environment. This fact significantly complicates the active control of the spatial noise field. However, to the author's best knowledge, most of the existing spatial ANC systems have neglected this fact. In this section, based on our understanding of the fundamentals of ANC, we identify the problems

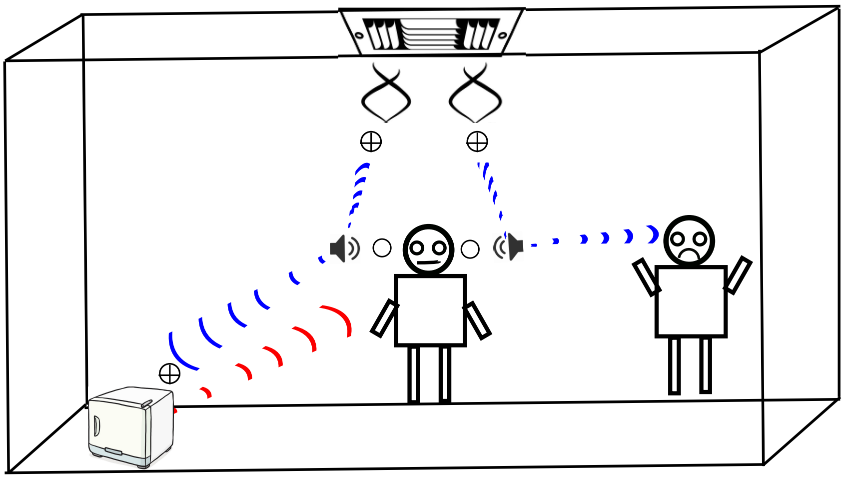


Figure 2.13: The noise field in a room environment and the active control of the noise field around one person's head. The reference sensor (⊕) measurements are contributed by the noise sources, the background noises, and the secondary source feedback.

with the existing spatial ANC systems. Note that the content of this section is a reformated version of the Sec. 1.3.

The reference signal generation

The first and the most serious problem with the existing spatial ANC systems is the lack of a proper way to generate the reference signal. We explain this problem using Fig. 2.13 as an example. One can use the reference sensors which are placed close to the noise sources to detect the noise source outputs, and use them as the reference signals. However, as shown in Fig. 2.13 the reference sensor measurements are the contributions from the noise sources, the background noises, and the secondary source feedback. The background noises and the secondary source feedback reduce the coherence between the reference signals and the primary noises, and can make the ANC system unstable. The feedback paths from secondary sources to reference sensors are long and time-varying in a room environment [10], and thus it is challenging to estimate the feedback paths in real-time and then compensate for their influence on the reference sensor measurements. Without a clean reference signal, it is difficult to cancel the spatial noise field in the room.

The causal control

Most of the existing spatial ANC systems shown in Sec. 2.2.2, 2.3.2, and 2.4 process the signals in the frequency-domain. Due to the introduced frame delay, the frequency-domain processing is suitable for controlling tonal noise fields only and is not suitable for controlling random noise fields, which are more prevalent.

Local sound field control

Most of the existing spatial ANC systems are local sound field control systems [90–92]. As shown in Fig. 2.12 and 2.13, the systems try to minimize the noise field in a small region, specifically around the human head. The arrangement significantly limits the movement of that person, and thus is highly undesirable.

Real-time secondary path estimation

As shown in Fig. 2.12, a spatial ANC system may deploy a large number of sources and sensors for noise field control and measurement, respectively. The complicated interactions between the sources, the sensors, and the environment make real-time secondary path estimation difficult. Nonetheless, the secondary path information is necessary for adjusting the secondary source driving signals.

Selective cancellation

A spatial noise field as shown in Fig. 2.13, may contain some desirable components, such as the human speech. The existing spatial ANC systems, which have not taken this into consideration, may cancel the desired human speech along with the annoying noises.

2.6 Summary

In this section, we first introduced the classical single channel ANC systems and the constraints that limited their performance. We then reviewed the spatial ANC systems based on the room mode, the spherical mode, and the Helmholtz integral, and identified the shortcomings of the systems.

Chapter 3

Active control of tonal outgoing noise fields in rooms

3.1 Introduction

In this chapter, we develop an ANC system that can cancel tonal noise fields in a room. We formulate the problem of tonal noise field control in a room as estimating and canceling the outgoing field produced by the primary source [93]. We reduce the primary outgoing field by superposing it with the secondary outgoing field. As the secondary outgoing field is independent of the room environment, the proposed system removes the need for real-time room characteristic identification, i.e., real-time secondary path estimation. By controlling the outgoing field only, the proposed system achieves global noise reduction with negligible influence on the desired fields in the room. This chapter also presents a method to estimate the outgoing field on a sphere and a wave-domain algorithm for outgoing field cancellation. Simulations and hardware demonstrations confirm the effectiveness of the proposed system for controlling tonal noise fields in rooms over a wide frequency range.

The rest of this chapter is organized as follows. We introduce the problem formulation in Sec. 3.2, the frequency-domain SFS method in Sec. 3.3. Section 3.4 present a spatial ANC system that cancels tonal noise fields globally in a room. The system performance is validated by simulations in Sec. 3.5 and experiments in

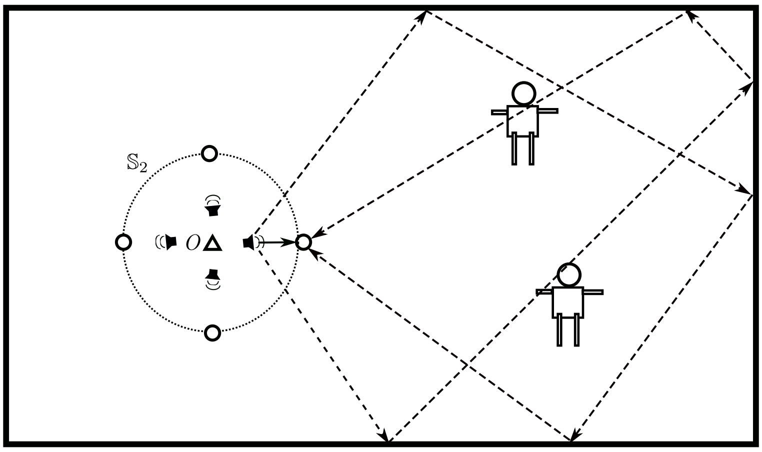


Figure 3.1: An example of the proposed ANC system in a room: the primary source is placed at the point O and marked by $\triangle$, the secondary sources are marked by 🔊, the error sensors are marked by $\circ$, and $\mathbb{S}_2$ denotes a sphere of radius R. A secondary path consists of a secondary source, an error sensor, embedded circuits, and an acoustic path. The acoustic path consists of the outgoing path $\longrightarrow$, and the reverberant path $\dashrightarrow$. There are desired sources (human) in the room.

Sec. 3.6. Section 3.7 concludes this chapter.

3.2 Problem formulation

Consider a room with a primary source placed at the point O and marked as $\triangle$, as shown in Fig. 3.1. A set of secondary sources and error sensors are marked as 🔊 and $\circ$, respectively. The radius of the sphere $\mathbb{S}_2$ is R. There are desired sources (human) in the room. A secondary path consists of a secondary source, an error sensor, embedded circuits, and an acoustic path. The acoustic path consists of the outgoing path $\longrightarrow$ and the reverberant path $\dashrightarrow$. As shown in Fig. 3.1, the reverberant path is susceptible to the movement of sound scattering objects, and thus is time-varying.

Let the noise pressure at a point (r, θ, ϕ) due to the primary source be $P(\omega, r, \theta, \phi)$. The problem considered in this chapter is to cancel the primary noise field $P(\omega, r, \theta, \phi)$ outside the sphere $\mathbb{S}_2$ and inside the room, without real-time estimation of the secondary paths, and with negligible influences on the desired fields in the room.

We express the the primary noise field $P(\omega, R, \theta, \phi)$ on the sphere $\mathbb{S}_2$ surrounding the primary source as a spherical harmonic expansion [36, 94]

$$P(\omega, R, \theta, \phi) = \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{P}_{u,v}(\omega, R) Y_{u,v}(\theta, \phi), \tag{3.1}$$

where $Y_{u,v}(\theta, \phi)$ is the spherical harmonic of order u and degree v, and $\mathsf{P}_{u,v}(\omega, R)$ are pressure field coefficients evaluated at the radius R. Using the spherical harmonic analysis equation, the pressure field coefficients $\mathsf{P}_{u,v}(\omega, R)$ are given by [36]

$$\mathsf{P}_{u,v}(\omega, R) = \int_0^{2\pi} \int_0^{\pi} P(\omega, R, \theta, \phi) Y_{u,v}(\theta, \phi) \sin\theta d\theta d\phi, \tag{3.2}$$

Hereafter, we abbreviate (θ, ϕ) as a single symbol Θ to simplify the notation.

The primary noise field $P(\omega, R, \Theta)$ on the sphere $\mathbb{S}_2$ consists of two parts: the primary outgoing field $P^{\mathrm{o}}(\omega, R, \Theta)$ and the primary incoming field $P^{\mathrm{i}}(\omega, R, \Theta)$ [36, 93]

$$\begin{aligned} P(\omega, R, \Theta) = & \underbrace{\sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{A}^{\mathrm{p}}_{u,v}(\omega) h_u(\omega \tau_R) Y_{u,v}(\Theta)}_{P^{\mathrm{o}}(\omega, R, \Theta)} \\ & + \underbrace{\sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{B}^{\mathrm{p}}_{u,v}(\omega) j_u(\omega \tau_R) Y_{u,v}(\Theta)}_{P^{\mathrm{i}}(\omega, R, \Theta)}, \end{aligned} \tag{3.3}$$

where $\mathsf{A}^{\mathrm{p}}_{u,v}(\omega)$ and $\mathsf{B}^{\mathrm{p}}_{u,v}(\omega)$ denote the spherical Hankel function coefficients and the spherical Bessel function coefficients, respectively, and $\tau_R = R/c$.

Note that the primary outgoing field $P^{\mathrm{o}}(\omega, R, \Theta)$ consists of the sound generated directly by the primary source and the sound scattered from the primary source surface and its surroundings [95]. (At the low frequency range, where the wavelength of noise is large compared with the size of the source, the scattered sound field is negligible [96].) The primary incoming field $P^{\mathrm{i}}(\omega, R, \Theta)$ is due to room reflections and desired sources [96].

We formulate the problem of controlling the primary noise field $P(\omega, r, \Theta)$ in the room as estimating and canceling the primary outgoing field $P^{\mathrm{o}}(\omega, R, \Theta)$ on the sphere $\mathbb{S}_2$. By canceling the primary outgoing field $P^{\mathrm{o}}(\omega, R, \Theta)$, we effectively cancel the undesired noise outside of the sphere $\mathbb{S}_2$ and in the entire room. Further, given that the desired sources are located outside of the sphere $\mathbb{S}_2$, the desired fields do not contribute to primary outgoing field $P^{\mathrm{o}}(\omega, R, \Theta)$ on the sphere $\mathbb{S}_2$. Thus, by canceling the primary outgoing field $P^{\mathrm{o}}(\omega, R, \Theta)$ only, a noise cancellation system can reduce its influence on the desired fields in the room.

3.3 Estimation of the outgoing field on a sphere

In this section, we introduce a method to estimate the outgoing field on a sphere. The effectiveness of this method has been confirmed by simulations and experiments [96]. We present detailed derivations of this method here for completeness.

3.3.1 The pressure and the velocity expansions

Let the primary noise field on the sphere $\mathbb{S}_2$ be sampled by an array of error sensors placed at $\{R, \Theta_q\}_{q=1}^{Q_{\mathrm{e}}}$, as shown in Fig. 1. Using (3.3), the primary noise pressures at the sampling points can be expressed as [36]

$$\begin{aligned} P(\omega, R, \Theta_q) &\approx \sum_{u=0}^{N_R} \sum_{v=-u}^{u} \mathsf{P}_{u,v}(\omega, R) Y_{u,v}(\Theta_q) \\ &= \sum_{u=0}^{N_R} \sum_{v=-u}^{u} [\mathsf{A}^{\mathrm{p}}_{u,v}(\omega) h_u(\omega\tau_R) + \mathsf{B}^{\mathrm{p}}_{u,v}(\omega) j_u(\omega\tau_R)] Y_{u,v}(\Theta_q), \end{aligned} \tag{3.4}$$

where both the outgoing field and the incoming field are truncated to order N_R [73, 97–99]. The truncation order N_R is determined by the radius of the sphere $\mathbb{S}_2$ and the frequency ω, and by the bounds on the spherical Bessel function, which also decides the upper limit of $\mathsf{A}^{\mathrm{p}}_{u,v}(\omega)$ in (3.4) [100].

The radial particle velocities at the sampling points can be expressed as [36]

$$\begin{aligned} V(\omega, R, \Theta_q) &= \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{V}_{u,v}(\omega, R) Y_{u,v}(\Theta_q) \\ &= \frac{i}{\rho\omega} \frac{\partial P(\omega, r, \Theta_q)}{\partial r}\bigg|_{r=R} \end{aligned}$$

$$\approx \frac{i}{\rho c} \sum_{u=0}^{N_R} \sum_{v=-u}^{u} [\mathsf{A}^{\mathrm{p}}_{u,v}(\omega) h'_u(\omega \tau_R) + \mathsf{B}^{\mathrm{p}}_{u,v}(\omega) j'_u(\omega \tau_R)] Y_{u,v}(\Theta_q), \quad (3.5)$$

where $\mathsf{V}_{u,v}(\omega, R)$ are the velocity field coefficients.

Theoretically, the coefficients $\mathsf{P}_{u,v}(\omega, R)$ and $\mathsf{V}_{u,v}(\omega, R)$ can be computed using the analytical equation (3.2) if the pressure and the radial particle velocity are known over the continuous sphere $\mathbb{S}_2$ [75]. However, in practice, only finite samples of the pressure and the radial particle velocity are available. Hence, we obtain the coefficients $\mathsf{P}_{u,v}(\omega, R)$ and $\mathsf{V}_{u,v}(\omega, R)$ through

$$\mathsf{P}_{u,v}(\omega, R) \approx \sum_{q=1}^{Q_{\mathrm{e}}} \tau_q P(\omega, R, \Theta_q) Y_{u,v}(\Theta_q), \quad (3.6)$$

$$\mathsf{V}_{u,v}(\omega, R) \approx \sum_{q=1}^{Q_{\mathrm{e}}} \tau_q V(\omega, R, \Theta_q) Y_{u,v}(\Theta_q), \quad (3.7)$$

where $u \in [0, N_R]$, $v \in [-u, u]$, $\{\tau_q\}_{q=1}^{Q_{\mathrm{e}}}$ are the sampling weights [75], and $\{\Theta_q\}_{q=1}^{Q_{\mathrm{e}}}$ are the sampling point coordinates. To accurately estimate the coefficients up to order N_R, we request $Q_{\mathrm{e}} \geq (N_R + 1)^2$ [75].

3.3.2 Estimation of the spherical Hankel function coefficients

In this section, we estimate the spherical Hankel function coefficients based on the coefficients of the pressure and the radial particle velocity on the sphere $\mathbb{S}_2$.

From the right hand sides of (3.4) and (3.5), the spherical Hankel function coefficients $\mathsf{A}^{\mathrm{p}}_{u,v}(\omega)$ and the spherical Bessel function coefficients $\mathsf{B}^{\mathrm{p}}_{u,v}(\omega)$ relate with the coefficients $\mathsf{P}_{u,v}(\omega, R)$ and $\mathsf{V}_{u,v}(\omega, R)$ through

$$\mathsf{P}_{u,v}(\omega, R) = \mathsf{A}^{\mathrm{p}}_{u,v}(\omega) h_u(\omega \tau_R) + \mathsf{B}^{\mathrm{p}}_{u,v}(\omega) j_u(\omega \tau_R), \quad (3.8)$$

$$-i\rho c \mathsf{V}_{u,v}(\omega, R) = \mathsf{A}^{\mathrm{p}}_{u,v}(\omega) h'_u(\omega \tau_R) + \mathsf{B}^{\mathrm{p}}_{u,v}(\omega) j'_u(\omega \tau_R). \quad (3.9)$$

Using (3.8) and (3.9), we obtain the spherical Hankel function coefficients $\mathsf{A}^{\mathrm{p}}_{u,v}(\omega)$ as

$$\mathsf{A}^{\mathrm{p}}_{u,v}(\omega) = \frac{\mathsf{P}_{u,v}(\omega, R) j'_u(\omega\tau_R) + i\rho c \mathsf{V}_{u,v}(\omega\tau_R) j_u(\omega\tau_R)}{h_u(\omega\tau_R) j'_u(\omega\tau_R) - h'_u(\omega\tau_R) j_u(\omega\tau_R)}. \tag{3.10}$$

Substituting the Wronskian relation (2.54) into (3.10) [36], we obtain

$$\mathsf{A}^{\mathrm{p}}_{u,v}(\omega) = (\omega\tau_R)^2[-i\mathsf{P}_{u,v}(\omega, R) j'_u(\omega\tau_R) + \rho c \mathsf{V}_{u,v}(\omega, R) j_u(\omega\tau_R)]. \tag{3.11}$$

Similarly, using (3.6), (3.8), and (2.54), we obtain the spherical Bessel function coefficients $B^{\mathrm{p}}_{u,v}(\omega)$ as

$$\mathsf{B}^{\mathrm{p}}_{u,v}(\omega) = (\omega\tau_R)^2[i\mathsf{P}_{u,v}(\omega, R) h'_u(\omega\tau_R) - \rho c \mathsf{V}_{u,v}(\omega, R) h_u(\omega\tau_R)]. \tag{3.12}$$

Substituting (3.11) and (3.12) into corresponding parts of (3.3), we obtain estimations of the primary outgoing field and the primary incoming field, respectively.

We have the following comments on (3.11):

1. The outgoing field is uniquely determined by the spherical Hankel function coefficients $\mathsf{A}^{\mathrm{p}}_{u,v}(\omega)$, and can be controlled by manipulating the coefficients.

2. The spherical Hankel function coefficients $\mathsf{A}^{\mathrm{p}}_{u,v}(\omega)$ are radial independent. The outgoing field on one sphere of radius R_1 can be projected to another sphere of radius R_2, where $R_2 > R_1$ [36]. By reducing the outgoing field on a small sphere surrounding the primary source, we essentially reduce the outgoing field on all outer spheres with larger radii. This fact enables the proposed ANC system to cancel the noise field globally in a room.

3.4 Spatial noise field cancellation

In this section, we develop a wave-domain mode matching method to cancel the outgoing field. [1] We preset the implementation of the proposed ANC system in

[1] We take a mode-domain approach for noise cancellation for two reasons. First, the spherical harmonic modes and their coefficients provide a sparse description for the sound field. The sparsity allows the reduction of the computational complexity of the control algorithm. Second,

Sec. 3.4.1.

The idea is to use the secondary outgoing field to destructively interfere with the primary outgoing field. We place Q_s point sources between the primary source and the sphere $\mathbb{S}_2$ as the secondary sources as shown in Fig. 3.1. The secondary outgoing field $S^\mathrm{o}(\omega, R, \Theta)$ on the sphere $\mathbb{S}_2$ is

$$S^\mathrm{o}(\omega, R, \Theta) \approx \sum_{u=0}^{N_R} \sum_{v=-u}^{u} \mathsf{A}^\mathrm{s}_{u,v}(\omega) h_u(\omega \tau_R) Y_{u,v}(\Theta), \tag{3.13}$$

where we also truncate the secondary outgoing field $S^\mathrm{o}(\omega, R, \Theta)$ to order N_R, and the expressions of the secondary spherical Hankel function coefficients $\mathsf{A}^\mathrm{s}_{u,v}(\omega)$ are [36]

$$\mathsf{A}^\mathrm{s}_{u,v}(\omega) = -ik \sum_{q=1}^{Q_\mathrm{s}} d_q(\omega) j_u(kr_q) Y_{u,v}(\Theta_q), \tag{3.14}$$

$u \in [0, N_R]$, $v \in [-u, u]$, (r_q, Θ_q) and $d_q(\omega)$ are the spherical coordinates and the driving signal of the q-th secondary source, respectively.

The matrix form of (3.14) is

$$\boldsymbol{A}^\mathrm{s}(\omega) = \boldsymbol{J}(\omega)\boldsymbol{D}(\omega), \tag{3.15}$$

where $\boldsymbol{A}^\mathrm{s}(\omega) = [\mathsf{A}^\mathrm{s}_{0,0}(\omega), \mathsf{A}^\mathrm{s}_{1,-1}(\omega), \dots, \mathsf{A}^\mathrm{s}_{N_R,N_R}(\omega)]^T$ is a $(N_R+1)^2 \times 1$ vector ($(\cdot)^T$ is the transpose operator), $\boldsymbol{D}(\omega) = [d_1(\omega), d_2(\omega), .., d_{Q_\mathrm{s}}(\omega)]^T$ is a $Q_\mathrm{s} \times 1$ vector, and $\boldsymbol{J}(\omega)$ is a $(N_R + 1)^2 \times Q_\mathrm{s}$ matrix

$$\boldsymbol{J}(\omega) = \begin{bmatrix} J_{0,0}(1) & J_{0,0}(2) & \dots & J_{0,0}(Q_\mathrm{s}) \\ J_{1,-1}(1) & J_{1,-1}(2) & \dots & J_{1,-1}(Q_\mathrm{s}) \\ \vdots & \vdots & \ddots & \vdots \\ J_{N_R,N_R}(1) & J_{N_R,N_R}(2) & \dots & J_{N_R,N_R}(Q_\mathrm{s}) \end{bmatrix}, \tag{3.16}$$

with entries $J_{u,v}(q) = -ikj_n(kr_q)Y_{u,v}(\Theta_q)$.

---

a mode and the corresponding coefficient denote the directivity of the noise field, and thus by minimizing that mode we can control the noise radiated to a particular direction only. This allows us to reduce the number of sources and sensors needed for noise cancellation.

The cancellation of the outgoing field requires

$$S^{\mathrm{o}}(\omega, R, \Theta) + P^{\mathrm{o}}(\omega, R, \Theta) = 0, \tag{3.17}$$

where $S^{\mathrm{o}}(\omega, R, \Theta)$ and $P^{\mathrm{o}}(\omega, R, \Theta)$ are given by (3.13) and (3.3), respectively. Using the spherical harmonic expansion of each term, (3.17) can be expressed as

$$\mathsf{A}^{\mathrm{s}}_{u,v}(\omega) + \mathsf{A}^{\mathrm{p}}_{u,v}(\omega) = 0, \quad u \in [0, N_R], v \in [-u, u]. \tag{3.18}$$

Based on (3.15), (3.18) can be expressed as a matrix equation

$$\boldsymbol{J}(\omega)\boldsymbol{D}(\omega) = \boldsymbol{A}^{\mathrm{p}}(\omega), \tag{3.19}$$

where

$$\boldsymbol{A}^{\mathrm{p}}(\omega) = -[\mathsf{A}^{\mathrm{p}}_{0,0}(\omega), \mathsf{A}^{\mathrm{p}}_{1,-1}(\omega), \ldots, \mathsf{A}^{\mathrm{p}}_{N_R,N_R}(\omega)]^T,$$

is a $(N_R + 1)^2 \times 1$ vector.

We solve (3.19) as a least square problem to obtain the secondary source driving signals

$$\boldsymbol{D}(\omega) = \boldsymbol{J}^{+}(\omega)\boldsymbol{A}^{\mathrm{p}}(\omega), \tag{3.20}$$

where $\boldsymbol{J}^{+}(\omega)$ is the pseudo-inverse of the matrix $\boldsymbol{J}(\omega)$.

The matrix $\boldsymbol{J}(\omega)$ characterizes the combined transfer function of the secondary source, the error sensor, the embedded circuits, and the outgoing path, excluding the time-varying reverberant path. The matrix $\boldsymbol{J}(\omega)$ is relatively stationary and can be estimated offline [95]. Therefore, the proposed ANC system does not need to estimate the secondary path transfer function in real-time.

Note that using the SFS method, both the pressure, the radial particle velocity, and the sound intensity associated with the outgoing field can be estimated. More advanced control strategies, such as the active sound intensity control [101–103], can be applied to further improve the performance of the proposed ANC system.

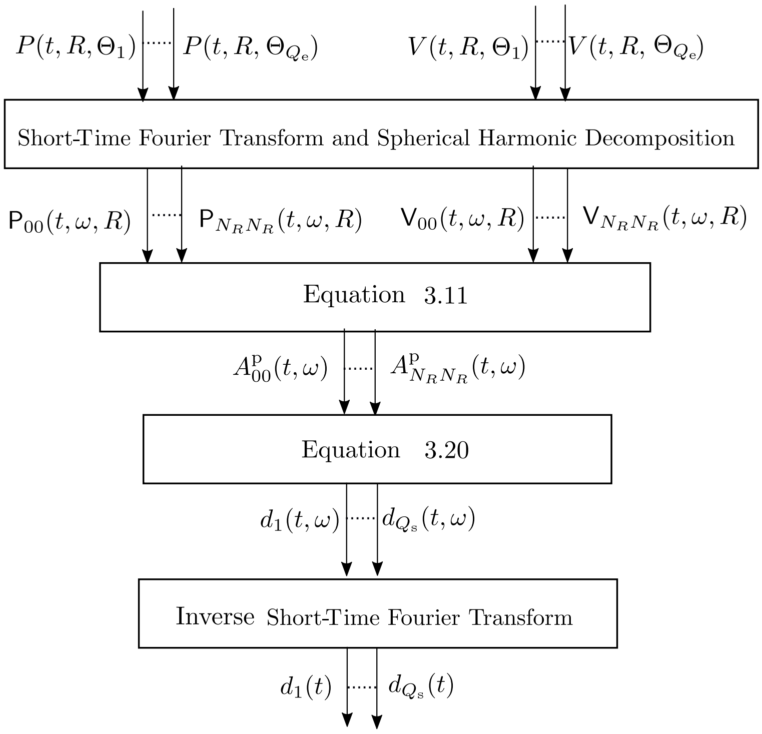


Figure 3.2: Signal flow diagram of the proposed ANC system.

3.4.1 Implementation of the proposed ANC system

In this section, we provide details to facilitate the implementation of the proposed ANC system.

We provide the signal flow diagram of the proposed ANC system in Fig. 3.2. We first transform the pressure $\{P(t,R,\Theta_q)\}_{q=1}^{Q_e}$ and the radial particle velocity $\{V(t,R,\Theta_q)\}_{q=1}^{Q_e}$ measurements of the error sensors into the time-frequency-domain through a STFT process. Then we decompose the results into corresponding coefficients $\{\mathsf{P}_{u,v}(t,\omega,R), \mathsf{V}_{u,v}(t,\omega,R)\}$, where $u \in [0, N_R], v \in [-u, u]$, through (3.6). Next we obtain the spherical Hankel function coefficients $\{\mathsf{A}^{\mathrm{p}}_{u,v}(t,\omega)\}$, where $u \in [0, N_R], v \in [-u, u]$, through (3.11). We calculate the time-frequency-domain secondary source driving signals $\{d_q(t,\omega)\}_{q=1}^{Q_s}$ using (3.20). Last we transform $\{d_q(t,\omega)\}_{q=1}^{Q_s}$ into the time-domain secondary source driving signal $\{d_q(t)\}_{q=1}^{Q_s}$ through an inverse STFT process.

In practical implementations of the proposed ANC system, the primary source is most likely to be attached to a surface. We can arrange the error sensors on an upper hemisphere surrounding the primary source as in examples shown in Sec. 3.5 and 3.6. At the low frequency range, the reflection from of typical surfaces, such as concrete or glass, are high [10]. The measured sound field can be regarded as symmetric with respect to the surface, and can be duplicated to the lower hemisphere to use the full spherical harmonic expansion [96].

We take a mode-domain approach for noise cancellation for two reasons. First, the spherical harmonic modes and their coefficients provide a sparse description for the sound field. That is, generally speaking, the number of modes are less than the number of pressure sampling points that is needed to describe the sound field on a sphere. The sparsity allows the reduction of the computational complexity of the control algorithm. Second, a mode and the corresponding coefficient denote the directivity of the noise field, and thus by minimizing that mode we can control the noise radiated to a particular direction only. This allows the reduction of the number of sources and sensors needed for noise cancellation.

3.5 Simulations

Simulations in this section illustrate the effectiveness of the SFS method and the proposed ANC system.

3.5.1 Simulation settings

The simulation environment is a rectangular room of size 4 m $\times$ 5 m $\times$ 3 m as shown in Fig. 3.3. There is a primary source ($\triangle$) on the floor, three secondary sources (🔊) around the primary source, a number of error sensors ($\circ$) on the upper hemisphere $\mathbb{S}_2$ of radius $R = 0.5$ m, and a desired sound source ($\bullet$). The reflection coefficients of the floor, the ceiling, and walls are all $\gamma = 0.995$. We duplicate the sound field measured on the upper hemisphere to the lower hemisphere [96]. We set up a Cartesian coordinate system and a spherical coordinate system based on the primary source center O. One corner of the room locates at $X = (-1.5, -2, 0)$ m with respect to the point O. The sampling frequency is $f_\mathrm{s} = 48000$ Hz, the speed of sound is $c = 343$ m/s, and the density of air is $\rho = 1.225$ kg/m^3. The room transfer

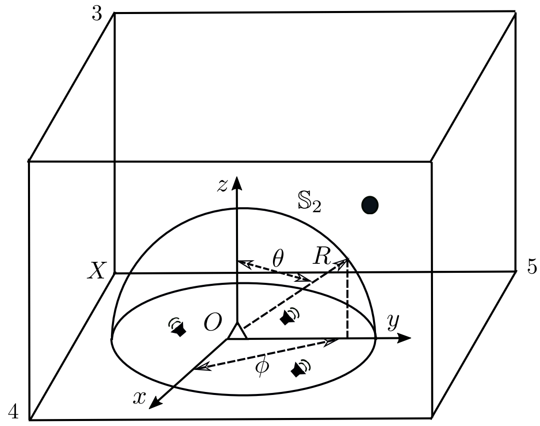


Figure 3.3: The simulation environment: The primary source is placed at a point O and marked by $\triangle$, the secondary sources are marked by 🔊, the error sensors are marked by $\circ$ (which are not shown for ease of illustration), $\mathbb{S}_2$ denotes a hemisphere enclosing the system, and $\bullet$ denotes a desired sound source.

functions (including the radial particle velocity responses) are simulated using the image source method [104], and the first 5832 image sources are considered. We add Gaussian white noises to the error sensor measurements, and the signal to noise power ratio is 40 dB. The simulation results are from the average of 100 independent runs. We use the settings in this paragraph for all simulations unless otherwise stated.

3.5.2 Sound field separation

In the first simulation, we let the primary source produce a unit-amplitude tonal wave of 200 Hz. The amplitudes of the outgoing field $P^{\mathrm{o}}(\omega, R, \Theta)$, the reverberation field $P(\omega, R, \Theta)$, the estimated outgoing field $\widehat{P}^{\mathrm{o}}(\omega, R, \Theta)$, and the field estimation error $P^{\mathrm{err}}(\omega, R, \Theta) = P^{\mathrm{o}}(\omega, R, \Theta) - \widehat{P}^{\mathrm{o}}(\omega, R, \Theta)$ on the upper hemisphere $\mathbb{S}_2$ are presented in Fig. 3.4 (a), (b), (c), and (d), respectively. We obtain the outgoing field $P^{\mathrm{o}}(\omega, R, \Theta)$ and the reverberation field $P(\omega, R, \Theta)$ by multiplying the primary source strength with the half-space Green function and the room transfer function, respectively [96]. We place four error sensors on the upper hemisphere $\mathbb{S}_2$ according to the first order Gauss sampling scheme [75], and obtain the estimated outgoing field $\widehat{P}^{\mathrm{o}}(\omega, R, \Theta)$ by the sound field separation method. The truncation order of

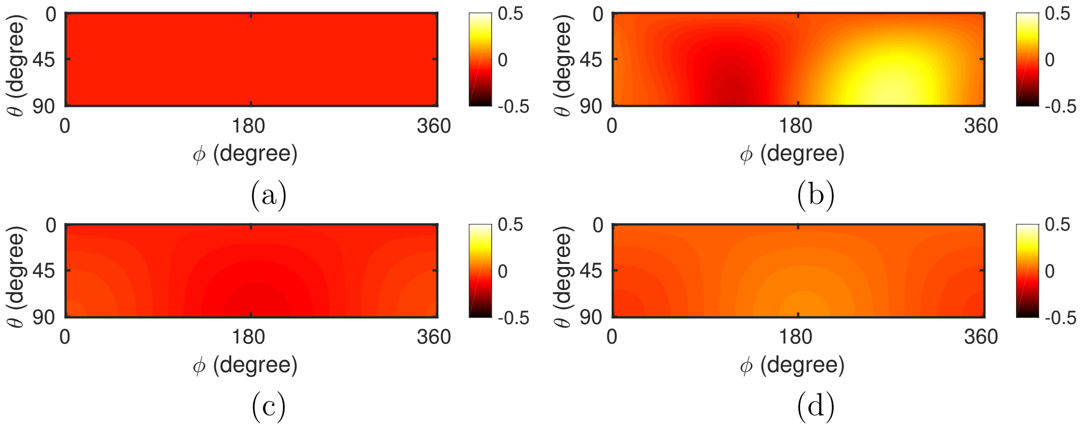


Figure 3.4: Sound field separation on the hemisphere $\mathbb{S}_2$: The amplitudes of (a) the outgoing field $P^{\rm o}(\omega, R, \Theta)$, (b) the reverberation field $P(\omega, R, \Theta)$, (c) the estimated outgoing field $\widehat{P}^{\rm o}(\omega, R, \Theta)$, and (d) the field estimation error $P^{\rm err}(\omega, R, \Theta)$.

sound field on the upper hemisphere is $N_R = 1$. In Fig. 3.4, the outgoing field $P^{\rm o}(\omega, R, \Theta)$ is equal over the hemisphere $\mathbb{S}_2$ as the primary source is at the center. The reverberation field $P(\omega, R, \Theta)$ distributes unequally on the hemisphere $\mathbb{S}_2$ due to wall reflections. The estimated outgoing field $\widehat{P}^{\rm o}(\omega, R, \Theta)$ approximates the outgoing field $P^{\rm o}(\omega, R, \Theta)$, and the field estimation error $P^{\rm err}(\omega, R, \Theta)$ is small over the hemisphere $\mathbb{S}_2$.

In the second simulation, we examine the field estimation error as a function of frequency. Define the normalized estimation error as

$$\sigma(\omega) = 10 \log_{10} \frac{\sum_{q=1}^{Q} \|P^{\rm o}(\omega, R, \Theta_q) - \widehat{P}^{\rm o}(\omega, R, \Theta_q)\|_2^2}{\sum_{q=1}^{Q} \|P^{\rm o}(\omega, R, \Theta_q)\|_2^2}, \tag{3.21}$$

where the sound fields are sampled at $Q = 100 \times 400$ equal-angel points on the upper hemisphere, and (R, Θ_q) are the spherical coordinates of q-th sampling point [75]. Denote $\sigma_1(\omega)$, $\sigma_2(\omega)$ and $\sigma_3(\omega)$ as the normalized estimation errors when the SFS method is realized by the first, second, and third order Gauss sampling scheme, i.e., four, nine, and sixteen error sensors are placed on the upper hemisphere $\mathbb{S}_2$, respectively. In computations of the normalized estimation errors $\sigma_1(\omega)$, $\sigma_2(\omega)$ and $\sigma_3(\omega)$, we choose the truncation orders for the sound field on the upper hemisphere $\mathbb{S}_2$ as $N_R = 1, 2, 3$, respectively. We depict $\sigma_1(\omega)$, $\sigma_2(\omega)$, and $\sigma_3(\omega)$ as functions

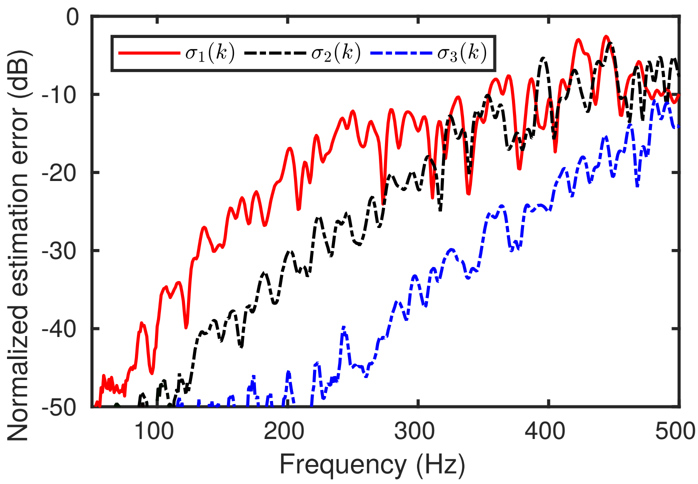


Figure 3.5: Sound field separation error as a functions of frequency: The normalized estimation errors (3.21) using the first, the second, and the third order Gauss sampling schemes for sound field separation are labeled as $\sigma_1(\omega)$, $\sigma_2(\omega)$, and $\sigma_3(\omega)$, respectively.

of frequency f in Fig. 3.5. As shown in Fig. 3.5, using the first, the second, and the third order Gauss sampling scheme, we are able to accurately ($\sigma_q(\omega) < -20$ dB, $q = 1, 2, 3$) estimate the outgoing field on the hemisphere $\mathbb{S}_2$ up to 200, 300, and 400 Hz, respectively. This simulation result shows that, by realizing the SFS method with an appropriate Gauss sampling scheme, the estimated outgoing field agrees well with the outgoing field over a wide frequency range.

Next we examine the normalized estimation error as a functions of the reflection coefficient. We keep the reflection coefficient of the floor constant as $\gamma = 0.995$, and change the reflection coefficients γ of the ceiling and walls from 0.5 to 0.995 to simulate the variation from a weak reverberant room to a strong reverberant room. We place nine error sensors on the hemisphere $\mathbb{S}_2$ according to the second order Gauss sampling scheme, and choose the truncation order of sound field on the upper hemisphere as $N_R = 2$. In Fig. 3.6, the normalized estimation errors at frequencies 200, 300, 400 Hz are labeled as σ_{200}, σ_{300}, and σ_{400}, respectively. As shown in Fig. 3.6, the normalized estimation error increases along with the reflection coefficients.

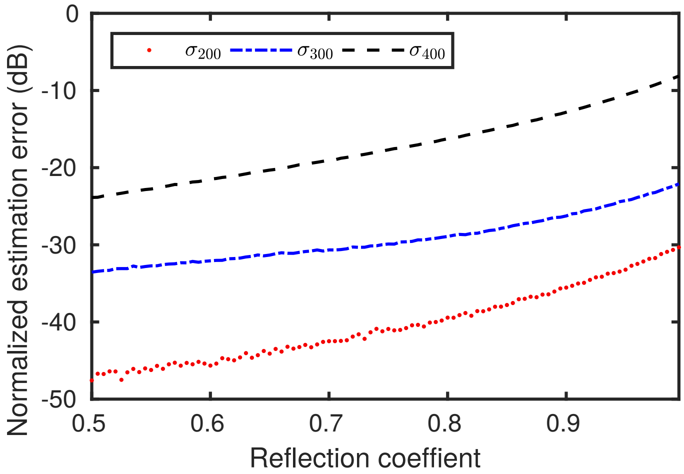


Figure 3.6: Sound field separation error as a function of the reflection coefficient: The normalized estimation errors at frequencies 200, 300, and 400 Hz are labeled as σ_{200}, σ_{300}, and σ_{400}, respectively.

3.5.3 Comparisons with a classical ANC system

In this section, we compare the performance of the proposed ANC system with a classical multichannel ANC system [11].

In the first simulation, we let the primary source produce a unit-amplitude tonal wave of 230 Hz. Three secondary sources are located at $(0.22, 0.0, 0.1)$ m, $(-0.11, 0.19, 0.1)$ m, and $(-0.12, -0.2, 0.1)$ m, respectively. For the proposed system, we design the secondary source driving signals to cancel the outgoing field on the upper hemisphere $\mathbb{S}_2$ based on (3.20). We obtain the spherical Hankel function coefficients by the SFS method, which is realized by four error sensors placed on the upper hemisphere $\mathbb{S}_2$ according to the first order Gauss sampling scheme. The sound field truncation order is $N_R = 1$. For the classical system, we calculate the secondary source driving signals using Eq. (12.2.4) from [11]. Rather than canceling the outgoing field as in the proposed system, the classical system tries to minimize the noise pressures at $(-1.4, -1.9, 0.1)$ m, $(2.4, -1.9, 0.1)$ m, $(2.4, 2.9, 0.1)$ m, $(-1.4, 2.9, 0.1)$ m, $(-1.4, -1.9, 2.9)$ m, $(2.4, -1.9, 2.9)$ m, $(2.4, 2.9, 2.9)$ m, and $(-1.4, 2.9, 2.9)$ m. (For the proposed system, we can use a microphone pair as the error sensor as shown in Section 3.6. The microphone pair can measure the noise pressure at two points. To make the comparison fair, we design the classical multichannel ANC system to cancel the noise pressure at 4×2 points.)

We present the results of the first simulation in Fig. 3.7. Figure 3.7 (a) depicts the primary field energy $\|P(\omega, \boldsymbol{x})\|_2^2$. Figure 3.7 (b) depicts the residual field energy $\|E(\omega, \boldsymbol{x})\|_2^2 = \|P(\omega, \boldsymbol{x}) + S(\omega, \boldsymbol{x})\|_2^2$ after cancellation by the proposed system, where $S(\omega, \boldsymbol{x})$ is the secondary field, $\boldsymbol{x} = (x, y, z)$ are Cartesian coordinates of 120000 sampling points which are arranged uniformly on the xz-plane. The residual field energy after cancellation by the classical system is depicted in Fig. 3.7 (c). In Fig. 3.7, the semi-circle denotes the hemisphere $\mathbb{S}_2$. As shown in Fig. 3.7, after cancellation by both the proposed system and the classical system, the residual field energy $\|E(\omega, \boldsymbol{x})\|_2^2$ is about 20 dB less than the primary field energy $\|P(\omega, \boldsymbol{x})\|_2^2$ on the xz-plane and outside of the hemisphere.

In the second simulation, we add a desired unit-amplitude point source (marked as $\bullet$ in Fig. 3.8) into the room at $(1.5, 0.0, 1.8)$ m. The desired point source also generates a tonal wave of 230 Hz. Other simulating settings are the same as in the first simulation. The desired field energy $\|D(\omega, \boldsymbol{x})\|_2^2$ and the total field energy $\|T(\omega, \boldsymbol{x})\|_2^2 = \|P(\omega, \boldsymbol{x}) + D(\omega, \boldsymbol{x})\|_2^2$ on the xz-plane are depicted in Fig. 3.8 (a) and (b), respectively. The residual field energy $\|E(\omega, \boldsymbol{x})\|_2^2 = \|P(\omega, \boldsymbol{x}) + D(\omega, \boldsymbol{x}) + S(\omega, \boldsymbol{x})\|_2^2$ after cancellation by the proposed system and the classical system are given by Fig. 3.8 (c) and (d), respectively. In Fig. 3.8, the semi-circle also denotes the hemisphere $\mathbb{S}_2$.

Comparisons of Fig. 3.8 (a), (b), and (c) reveal that the proposed system reduces the total field energy $\|T(\omega, \boldsymbol{x})\|_2^2$; whilst the residual field energy approximates the desired field energy, i.e., $\|E(\omega, \boldsymbol{x})\|_2^2 \approx \|D(\omega, \boldsymbol{x})\|_2^2$, on the xz-plane and outside of the hemisphere. Comparisons of Fig. 3.8 (a), (b), and (d) reveal that the classical system reduces the total field energy $\|T(\omega, \boldsymbol{x})\|_2^2$; but the residual field energy differs from the desired field energy, i.e., $\|E(\omega, \boldsymbol{x})\|_2^2 \not\approx \|D(\omega, \boldsymbol{x})\|_2^2$, on the xz-plane.

The simulation results demonstrate the advantage of the proposed system compared with the classical system. By arranging the noise control system inside a small hemisphere and controlling the outgoing field only, the proposed system can reduce its influence on the desired fields in the room. The classical system, on the other hand, may cancel the desired fields together with the primary noise field.

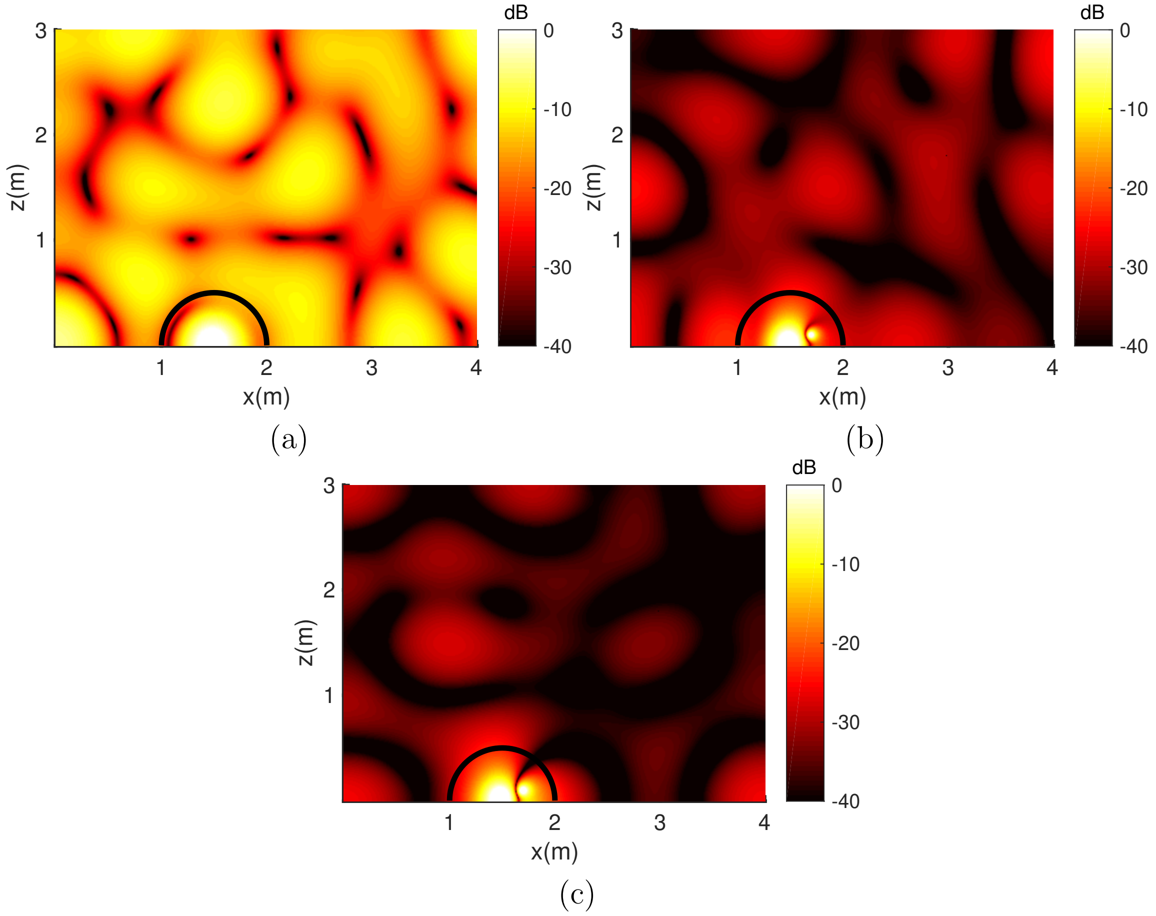


Figure 3.7: Comparisons between the proposed system and the classical system without any desired sound sources: (a) the primary field energy $\|P(\omega, \boldsymbol{x})\|_2^2$, the residual field energy $\|E(\omega, \boldsymbol{x})\|_2^2$ after cancellation by (b) the proposed system and (c) the classical system.

3.5.4 Noise cancellation over a wide frequency range

In this section, we conduct noise cancellation over a wide frequency range using the proposed system. For simplification, we do not consider the desired sound sources. We let the primary source produce a unit-amplitude tonal wave at a single frequency f ($f \in [50, 500]$ Hz), and use three secondary sources to control the primary noise field in two cases:

1. The three secondary sources are located at $(0.22, 0.0, 0.1)$ m, $(-0.11, 0.19, 0.1)$ m, and $(-0.12, -0.2, 0.1)$ m, respectively. The distances from the secondary

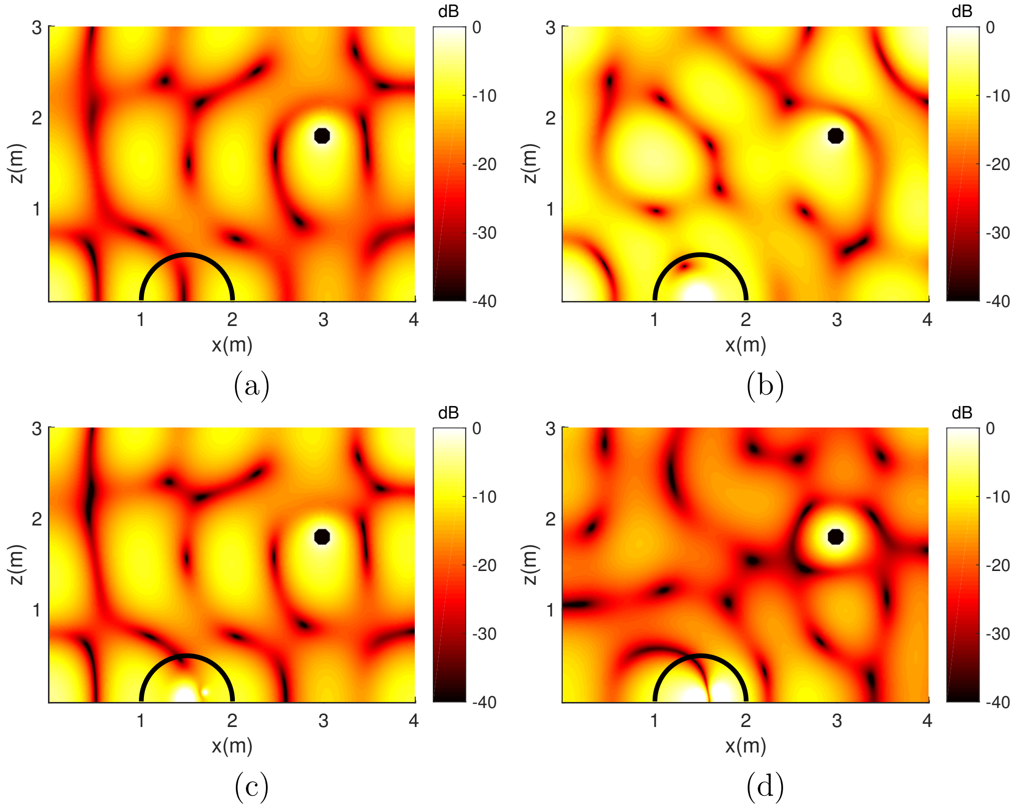


Figure 3.8: Comparisons between the proposed and the classical system with desired sound source: (a) the desired field energy $\|D(\omega,\boldsymbol{x})\|_2^2$, (b) the total field energy $\|T(\omega,\boldsymbol{x})\|_2^2$, the residual field energy $\|E(\omega,\boldsymbol{x})\|_2^2$ after cancellation by (c) the proposed and (d) the classical system.

sources to the origin are 0.24 m, 0.25 m, and 0.26 m, respectively.

2. The three secondary sources are located at $(0.173, 0.0, 0.1)$ m, $(-0.075, 0.13, 0.1)$ m, and $(-0.081, -0.14, 0.1)$ m, respectively. The distances from the secondary sources to the origin are 0.2 m, 0.18 m, and 0.19 m, respectively.

Other simulation settings are same as in the first simulation of Sec. 3.5.3.

The performances of the proposed ANC system are characterized by the noise field energy reduction in the room

$$\xi(\omega) = 10\ \log_{10} \frac{\sum_{q=1}^{Q} \|P(\omega,\boldsymbol{x}_q)\|^2}{\sum_{q=1}^{Q} \|E(\omega,\boldsymbol{x}_q)\|^2}, \tag{3.22}$$

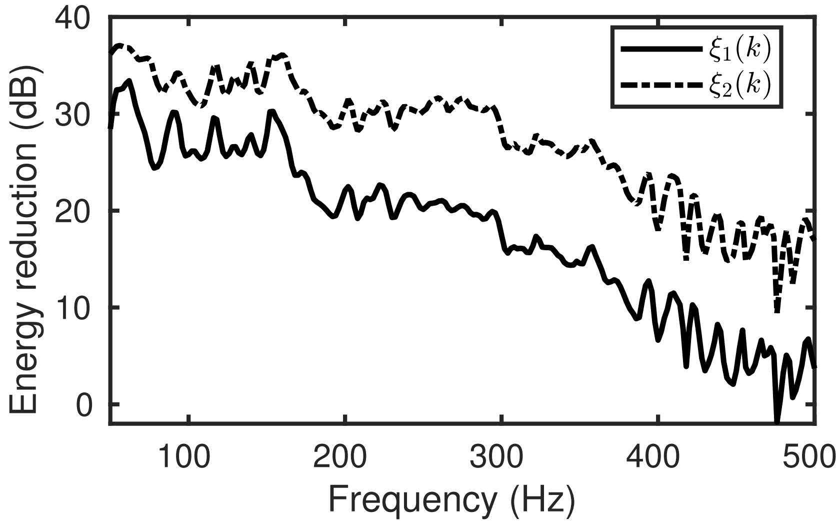


Figure 3.9: Noise cancellation over a wide frequency range: Noise field energy reductions achieved by the proposed ANC system over the frequency range [50, 500] Hz in the first case $\xi_1(\omega)$ and in the second case $\xi_2(\omega)$.

where $P(\omega, \boldsymbol{x}_q)$ is the primary noise pressure, $E(\omega, \boldsymbol{x}_q)$ is the residual noise pressure, $Q = 60000$ is the number of sampling points, $\boldsymbol{x}_q = (x_q, y_q, z_q)$ denotes the q-th sampling point position, $x_q = -1.4, -1.3, ..., 2.5$ m, $y_q = -1.9, -1.8, ..., 3.0$ m, and $z_q = 0.1, 0.2, ..., 3.0$ m. The sampling points inside of the hemisphere $\mathbb{S}_2$ are excluded in summations of (3.22). The simulation results of the first case and the second case are presented in Fig. 3.9 as $\xi_1(\omega)$ and $\xi_2(\omega)$, respectively.

As shown in Fig. 3.9, in the first case, the proposed system reduces the noise field energy in the room by more than 10 dB over the frequency range [50, 400] Hz. In the second case, where the secondary sources are placed closer to the primary source, the proposed system reduces the noise field energy in the room by more than 10 dB over the frequency range [50, 500] Hz. [2]

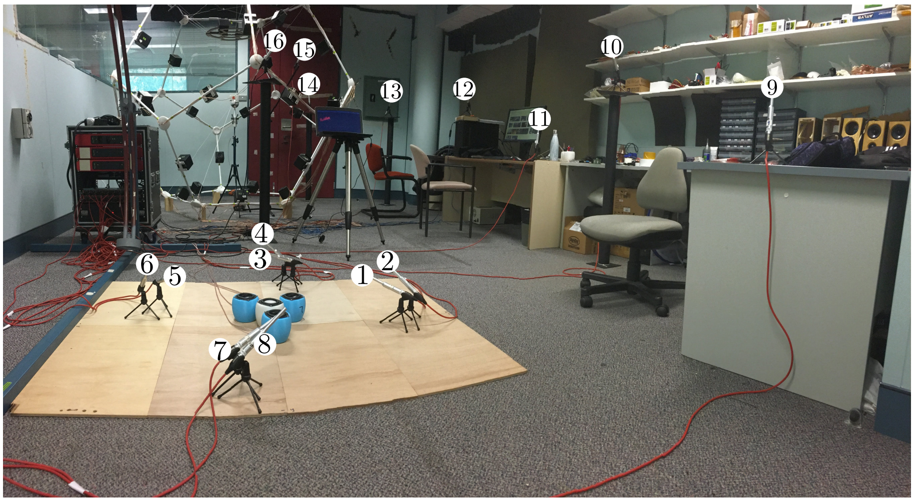
[2]It is possible to further reduce the noise field energy by using more sources and sensors. However, the arrangement of the sources and sensors influences the noise cancellation performance of the ANC system in a complicated way, which we do not have a good understanding yet. The theoretical investigation will be conducted in the future.

3.6 Hardware demonstration

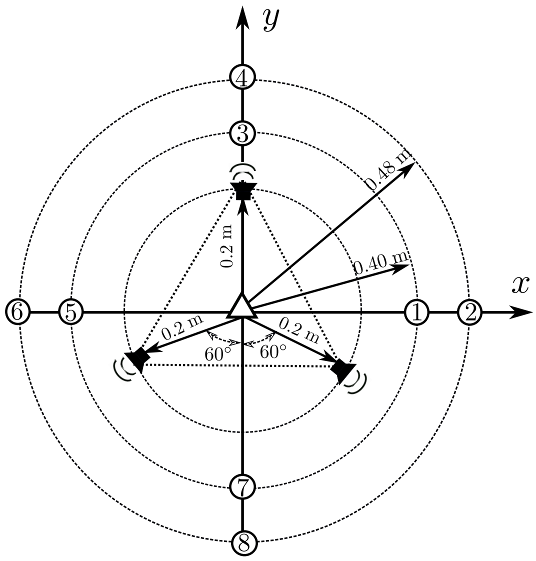
We implement the proposed ANC system to validate its effectiveness for reducing real noises. The demonstration environment is a room of size about 3.6 m $\times$ 6.7 m $\times$ 2.8 m as shown in Fig. 3.10. There is a thin layer of carpet on the concrete floor, and many objects in the room. The room reverberation time is $T_{60} \approx 1.5$ s. We have four loudspeakers on the ply-woods. The loudspeaker at the center is the primary source, and the other three loudspeakers are secondary sources. All loudspeakers are approximately circular cylinders, with each having a height of 0.11 m and a radius of 0.055 m. The loudspeaker drivers are Dayton Audio ND90-8. We set up a spherical coordinate system with the origin at the center of the bottom of the primary source. The distances from the secondary sources to the origin are about 0.2 m, and the distances between secondary sources are the same. We have four pairs of microphones labeled as 1,2; 3,4; 5,6; 7,8 as shown in Fig. 3.10. The inner four microphones (labeled as 1, 3, 5, and 7) are placed on a hemisphere of radius $R_1 = 0.45$ m according to the first order Gauss sampling scheme. The distance between these inner four microphones to the ground is about 0.2 m. The outer four microphones (labeled as 2, 4, 6, and 8) are placed on a hemisphere of radius $R_2 = 0.55$ m according the first order Gauss sampling scheme. The distance between these outer four microphones to the ground is about 0.25 m. These four microphone pairs are the error sensors, providing the pressure and the velocity information on a hemisphere of radius $R = 0.5$ m [96]. We have eight more microphones labeled as 9, 10, ..., 16 in the room as shown in Fig. 3.10. These eight microphones monitor the noise pressure levels in the room but are not part of the ANC system. All sixteen microphones are Dayton Audio EMM-6 precision electric condenser microphones, and have been calibrated up to 1000 Hz. The analogy-to-digital converter is Behringer Ultragain ADA 8200, and the digital-to-analogy converter is Rednet 2. The precision of both ADA8200 and Rednet 2 is 24 bit. We use a desktop computer to process the signals. The sampling frequency is $f_s = 48000$ Hz and the speed of sound is $c \approx 343$ m/s.

The overall electrical latency of the ANC system is 5.3 ms.

We conduct two hardware demonstrations. In the first demonstration, we use the secondary sources to control the noise produced by the primary source at a



(a)



(b)

Figure 3.10: The hardware demonstration: (a) An overview of the sources and microphones in the room. The loudspeaker at the center is the primary source, the other three loudspeakers are secondary sources, the microphones pairs labeled 1,2; 3,4; 5,6; 7,8 are the error sensors, and the microphones labeled 9, 10, 11, ..., 16 monitor the noise pressure levels in the room. (b) A top view of the primary source (△), the secondary sources (🔊), and the errors sensors ∘.

single frequency f, where $f = 100, 110, ..., 500$ Hz. Due to their limited volume, the loudspeakers are unable to radiate very low frequency, $f < 100$ Hz, sound effectively. Thus the demonstrations are conducted for higher frequency, where $100 < f < 500$ Hz. Prior to the operation of the ANC system, we let each secondary source produces a noise of the form $\cos(2\pi f t)$ over a period of $T = 8$ seconds. Recordings from the four error sensors, the microphone pairs labelled 1,2; 3,4; 5,6; 7,8, are transformed into the frequency-domain. Denote the frequency-domain noise pressures at a microphones pair be $P_{2q-1}(\omega, R_1, \Theta_q)$ and $P_{2q}(\omega, R_2, \Theta_q)$, where $q \in [1, 4]$. We obtain the pressure and the radial particle velocity at (R, Θ_q) through

$$P_q(\omega, R, \Theta_q) \approx \frac{1}{2}(P_{2q}(\omega, R_2, \Theta_q) + P_{2q-1}(\omega, R_1, \Theta_q)), \tag{3.23}$$

$$V_q(\omega, R, \Theta_q) \approx \frac{i}{\rho\omega} \frac{P_{2q}(\omega, R_2, \Theta_q) - P_{2q-1}(\omega, R_1, \Theta_q)}{\delta_d}, \tag{3.24}$$

where $\delta_d = R_2 - R_1 = 0.1$ m. To use the full spherical harmonic expansion, we mirror the pressure and the radial particle velocity on the upper hemisphere to the lower hemisphere [96]. We use the sound field separation method to obtain the spherical Hankel function coefficients which constitute the matrix $\boldsymbol{J}(\omega)$ in (3.16). We then let the primary source produce a unit-amplitude tonal noise and cancel the primary noise field in the room following the signal flow diagram as shown in Fig. 3.2.

In Fig. 3.11, we depict the noise energy reduction at the sixteen microphones as $\xi(\omega)$, and

$$\xi(\omega) = 10 \ \log_{10} \frac{\sum_{q=1}^{16} \|P_q(\omega)\|^2}{\sum_{q=1}^{16} \|E_q(\omega)\|^2}, \tag{3.25}$$

where $P_q(\omega)$ and $E_q(\omega)$ are the primary noise pressure and the residual noise pressure at the q-th microphone, respectively. As shown in Fig. 3.11, we reduce the noise energy at the sixteen microphones by more than 13 dB over the frequency range [100, 450] Hz. Measurements using a sound level meter reveal that the noise pressure levels are reduced by about 10 dB over the frequency range $[100, 450]$ Hz

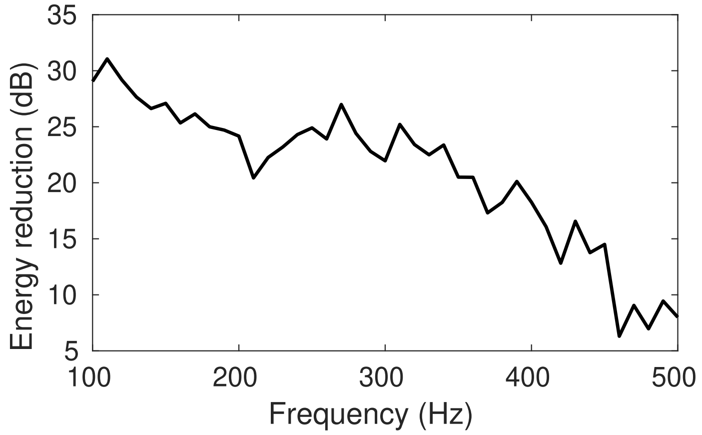


Figure 3.11: Noise cancellation over a wide frequency range: Noise energy reduction $\xi(\omega)$ at the sixteen microphones over the frequency range [100, 500] Hz.

in the room.

The arrangement of the sources and error sensors in the first hardware demonstration is similar to the Case 2 in Sec. 3.5.4. Comparison of Fig. 3.11 with $\xi_2(\omega)$ in Fig. 3.9 reveals that the energy reduction level in the hardware demonstration is about $5 \sim 8$ dB less than in the simulation. This is because, as shown in Fig. 3.10, our lab is much more complicated than the simulation environment. Overall, the simulation and the experimental results are consistent and demonstrate that the proposed ANC system can reduce outgoing noise fields over a wide frequency range.

In the second demonstration, the basic settings are the same as in the first demonstration, except that the primary noise consists of three unit-amplitude tonal waves of 150, 200, and 310 Hz. Denote the time-domain primary noise pressure at the q-th microphone as $p_q(t)$, and the residual noise pressure at the microphone as $e_q(t)$. We record the primary noise pressures $\{p_q(t)\}_{q=1}^{16}$ and the residual noise pressures $\{e_q(t)\}_{q=1}^{16}$ over 20 seconds, respectively. Let the noise energy reduction at the microphones be defined as

$$\xi_q = 10 \log_{10} \frac{\sum_{t=t_1}^{T=t_1+20\ s} p_q(t)^2}{\sum_{t=t_2}^{T=t_2+20\ s} e_q(t)^2}, \quad q \in [1, 16], \tag{3.26}$$

where t_1 and t_2 are the time instants that we start recording the primary noise

Table 3.1: Noise energy reductions $\{\xi_q\}_{q=1}^{16}$ at the sixteen microphones over 20 seconds.

q	1	2	3	4	5	6	7	8
ξ_q (dB)	18.4	18.7	19.4	17.2	21.4	19.2	14.3	12.5
q	9	10	11	12	13	14	15	16
ξ_q (dB)	11.7	8.4	8.8	14.7	10	8.1	9.6	13.6

pressures and the residual noise pressures, respectively. We present the noise energy reduction at the microphones in Table I. As shown in Table I, the proposed system reduces noise energy at the four microphone pairs (labeled as 1,2; 3,4; 5,6; 7,8 in Fig. 10) by more than 12.5 dB, and at the eight monitoring microphones (labeled as 9, 10,...,16 in Fig. 10) by more than 8 dB.

In all simulations and hardware demonstrations, the proposed system reduces the noise fields globally in the entire room and exterior to the hemisphere $\mathbb{S}_2$ (the error microphones) without real-time estimation of the secondary paths.

Note that for ease of conducting the experiment, we assume the primary noise is known, and we use the primary noise as the reference signal directly. Narrowband noises are mainly generated by the rotating part (such as fan or motor) of the primary source, and thus for practical implementation of the proposed system, we non-acoustic sensors, such as tachometer, to measure the rotation speed of the corresponding part. We can artificially generate the sinusoidal reference signal based on the measured rotation speed.

3.7 Conclusion

This chapter proposed an ANC system that can cancel tonal noise fields in rooms without the need for real-time secondary path estimation. The idea is to cancel the outgoing field produced by the primary source on a small sphere surrounding the primary source, instead of controlling the noise pressure at multiple points [12] or the room mode coefficients [57, 58]. We use a frequency-domain SFS method to separate the outgoing field from the interfering incoming field for the ANC system, enabling it to cancel the noise field globally without canceling the desired field in the room.

However, the frequency-domain SFS method, which relies on the STFT process to transfer the time-domain acoustic quantities into the time-frequency-domain and vice-versa, unavoidably introduces the frame delay. This frame delay makes the ANC system only suit for controlling tonal noise fields. In the next chapter, we derive the time-domain SFS method for separating the non-stationary sound fields on spheres. The method will enable a spatial ANC to control the random noise field globally in a room.

Publications related to this chapter

- **Fei Ma**, Wen Zhang, Thushara D. Abhayapala, "Active control of outgoing noise fields in rooms," The Journal of the Acoustic Society of America, vol. 144, no. 3, pp. 1589-1599, Sep. 2018.

Chapter 4

Real-time separation of non-stationary sound fields on spheres

4.1 Introduction

To investigate the working condition of a machine or the acoustic characteristics of a musical instrument, we can use a sensor array to measure the target field generated by them and analyze the sensor array measurement. However, in practical acoustic environments, such as inside rooms, there are disturbing sources, which produce disturbing fields. The disturbing fields and the room reflected fields will interfere with the target field and contaminate the sensor array measurements, making it difficult to study the target source characteristics. Nonetheless, the target source does not necessarily co-locates with the disturbing sources. By exploiting the spatial differences of the sound sources, it is possible to separate the target field from the interfering fields using a sensor array. Thus we can study the target source through analyzing the separated target field.

Over the years, a number of SFS methods have been reported in the literature. The spherical mode decomposition based method uses the spherical modes as the basic modeling functions of a field, and are capable of separating the outgoing field and the incoming field on a spherical sensor array [36, 96, 105–107]. The Spatial

Fourier Transform based method uses the two-dimensional Fourier Transform to decompose the field into plane-wave components, and can separates the incident field and the reflected field on a planar sensor array [108]. The statistically optimized near field holography method can separate the sound field on a planar sensor array, and can mitigate the spatial window leakage problem of the Spatial Fourier Transform [109]. The recently developed boundary element based method [110,111] and the equivalent source based method [112–114] extend the application of SFS to arbitrary shape sensor arrays. Further, these two methods can separate the scattering from the target source surface and recover the free-field radiation of the target source.

The above mentioned SFS methods can separate the fields coming from two sides of a sensor array (or two sensor arrays) apart. However, they are all frequency-domain methods. They accumulate and transfer a frame of the time-domain pressure (or particle velocity) measurements into the time-frequency domain using the STFT. They conduct the SFS at each time-frequency bin. The STFT process inevitably introduces the spectrum leakage problem and the frame delay. The frame delay is acceptable if the target field is stationary. Nonetheless, when dealing with fast changing non-stationary field, the frame delay will make the SFS methods unable to recover the target field in real-time. That further makes the separated target field unable to be used in time-critical applications, such as active noise control [8], real-time beamforming, and machine anomaly diagnosis [110]. A time-domain SFS method, on the other head, can track the changes of the target field without introducing the frame delay.

Manipulations of the sound fields in the time-domain are difficult, and a limited number of time-domain SFS methods have been reported up to date. The only existing time-domain SFS methods are developed by Prof. Bi and his coauthors [115–119]. Based on the Spatial Fourier Transform, they developed a method that can separate the field coming from two sides of a planar sensor array [115–117, 119]. Based on the interpolated time-domain equivalent source method [118], they further developed a method that can separate the sound generated by a particular source in a multiple non-stationary sources scene. However, these time-domain methods are best suitable for SFS in free field (or semi-anechoic rooms). Because first, in room environments, due to wall reflections the interfering sounds and the target sound can arrive at a planar sensor array in the same directions [10,104], and thus the target sound and the interfering sound become indistinguishable. Second, the equivalent source based method requires perfect knowledge of the location and the geometry of all the sources. However, in a room environment, the image source effect exists, and it is difficult to locate the image source especially for an arbitrary shape room with irregular boundaries [10, 104].

In this chapter, based on the spherical mode expansion, we develop a time-domain SFS method that can separate the non-stationary outgoing field and incoming field on a sphere in both free field and room environments. We first decompose the pressure and the radial particle velocity into corresponding coefficients. The convolution between the coefficients and the derived spatial filters results in the outgoing field and the incoming field on the array. The performance of the proposed method is confirmed by simulations, and compared with the Spatial Fourier Transform based method. Possible applications of the proposed method include (1) reference signal generation for active noise control system [8], (2) real-time beamforming, (3) de-reverberation for speech recognition and sound field reproduction system, (4) machine working condition monitoring in a noisy environment.

This chapter organized as follows. We introduce the problem formulation in Sec. 4.2. In Sec. 4.3, we provide the theoretical derivation of the time-domain SFS method. Section 4.4 introduces the practical implementations of the proposed method, whose effectiveness is validated by simulations in Sec. 4.5, and Sec. 4.6 concludes this chapter.

4.2 System model

4.2.1 Problem formulation

Consider the systems shown in Fig. 4.1, where in Fig. 4.1 (a) the target sources (denoted as $\triangle$) are placed in free field, and in Fig. 4.1 (b) the target sources (denoted as $\triangle$) are placed inside a room. We use (x, y, z) and (r, Θ) $(\Theta = (\theta, \phi))$ to denote the Cartesian coordinates and the spherical coordinates of a point with respect to the point O, respectively. The field $p(t, R, \Theta)$ on the sphere $\mathbb{S}_2$ of radius R is the

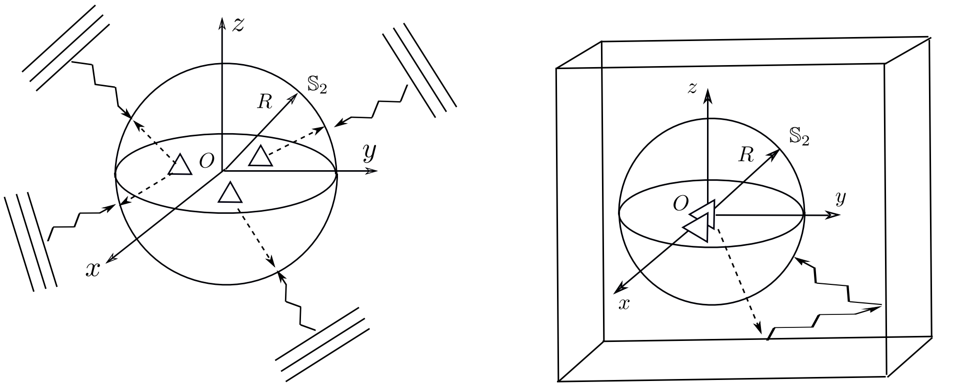


Figure 4.1: System models: (a) the target sources △ are placed in free field, (b) the target sources △ are placed inside a room. The outgoing field and the incoming field on the sphere $\mathbb{S}_2$ of radius R are denoted as $\dashrightarrow$ and $\rightsquigarrow$, respectively.

superposition of the outgoing field $p^{\mathrm{o}}(t, R, \Theta)$ and the incoming field $p^{\mathrm{i}}(t, R, \Theta)$

$$p(t, R, \Theta) = p^{\mathrm{o}}(t, R, \Theta) + p^{\mathrm{i}}(t, R, \Theta). \tag{4.1}$$

As shown in Fig. 4.1, the incoming field $p^{\mathrm{i}}(t, R, \Theta)$ on the sphere $\mathbb{S}_2$ is due to the external sources or room reflections. The outgoing field $p^{\mathrm{o}}(t, R, \Theta)$ on the sphere $\mathbb{S}_2$, on the other hand, is generated only by the target sources. In this paper, we aim to separate the outgoing field $p^{\mathrm{o}}(t, R, \Theta)$ on the sphere $\mathbb{S}_2$ to facilitate the study of the target sources.

Note that, we assume the target sources are point sources, which do not scatter the incoming field. For real world target source of finite size, the source's surface scatters the incoming field and transforms it into part of the outgoing field. To recover the outgoing field generated only by the target source, we need to further separate the scattered field using prior knowledge, such as the incoming field direction and the target source geometry. Such investigations are beyond the scope of the thesis and will be our future work.

4.2.2 The frequency-domain sound field separation on a sphere

In the frequency-domain, the sound field $P(\omega, R, \Theta)$ on a sphere $\mathbb{S}_2$ of radius R consists of the outgoing field $P^{\mathrm{o}}(\omega, R, \Theta)$ and the incoming field $P^{\mathrm{i}}(\omega, R, \Theta)$ [36]

$$\begin{aligned} P(\omega, R, \Theta) &= P^{\mathrm{o}}(\omega, R, \Theta) + P^{\mathrm{i}}(\omega, R, \Theta) \\ &= \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{P}_{u,v}(\omega, R) Y_{u,v}(\Theta) \\ &= \sum_{u=0}^{\infty} \sum_{v=-u}^{u} [\underbrace{\mathsf{A}_{u,v}(\omega) h_u(\omega\tau_R)}_{\mathsf{P}^{\mathrm{o}}_{u,v}(\omega,R)} + \underbrace{\mathsf{B}_{u,v}(\omega) j_u(\omega\tau_R)}_{\mathsf{P}^{\mathrm{i}}_{u,v}(\omega,R)}] Y_{u,v}(\Theta), \end{aligned} \tag{4.2}$$

where $\tau_R = R/c$, $\mathsf{P}_{u,v}(\omega, R)$ are the pressure field coefficients corresponding to the sound field $P(\omega, R, \Theta)$ measured on the sphere $\mathbb{S}_2$, $\mathsf{P}^{\mathrm{o}}_{u,v}(\omega, R)$ and $\mathsf{P}^{\mathrm{i}}_{u,v}(\omega, R)$ are the outgoing field coefficients and the incoming field coefficients, respectively, $\mathsf{A}_{u,v}(\omega)$ and $\mathsf{B}_{u,v}(\omega)$ are the spherical Hankel function coefficients and the spherical Bessel function coefficients, respectively.

The frequency-domain radial particle velocity on the sphere $\mathbb{S}_2$ can be similarly expressed as [36]

$$\begin{aligned} V(\omega, R, \Theta) &= \frac{i}{\rho\omega} \frac{\partial P(\omega, R, \Theta)}{\partial r}\Big|_{r=R} \\ &= \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{V}_{u,v}(\omega, R) Y_{u,v}(\Theta) \\ &= \frac{i}{\rho c} \sum_{u=0}^{\infty} \sum_{v=-u}^{u} [\mathsf{A}_{u,v}(\omega) h'_u(\omega\tau_R) + \mathsf{B}_{u,v}(\omega) j'_u(\omega\tau_R)] Y_{u,v}(\Theta), \end{aligned} \tag{4.3}$$

where $\mathsf{V}_{u,v}(\omega, R)$ are the velocity field coefficients corresponding to the radial particle velocity $V(\omega, R, \Theta)$ measured on the sphere $\mathbb{S}_2$.

Based on (4.2) and (4.3), we obtain the outgoing field coefficients and the incoming field coefficients as [36]

$$\begin{aligned} \mathsf{P}^{\mathrm{o}}_{u,v}(\omega, R) = &-i(\omega\tau_R)^2 j'_u(\omega\tau_R) h_u(\omega\tau_R) \mathsf{P}_{u,v}(\omega, R) \\ &+\rho c(\omega\tau_R)^2 j_u(\omega\tau_R) h_u(\omega\tau_R) \mathsf{V}_{u,v}(\omega, R), \end{aligned} \tag{4.4}$$

and

$$\begin{aligned}\mathsf{P}^{\mathrm{i}}_{u,v}(\omega,R) = &\, i(\omega R/c)^2 j_u(\omega\tau_R)h'_u(\omega\tau_R)\mathsf{P}_{u,v}(\omega,R)\\ &-\rho c(\omega\tau_R)^2 j_u(\omega\tau_R)h_u(\omega\tau_R)\mathsf{V}_{u,v}(\omega,R),\end{aligned} \tag{4.5}$$

respectively. Substitution of the coefficients $\mathsf{P}^{\mathrm{o}}_{u,v}(\omega,R)$ and $\mathsf{P}^{\mathrm{i}}_{u,v}(\omega,R)$ back into (4.2) results in the outgoing field $P^{\mathrm{o}}(\omega,R,\Theta)$ and the incoming field $P^{\mathrm{i}}(\omega,R,\Theta)$ on the sphere $\mathbb{S}_2$, respectively. The detailed derivation of (4.4) and (4.5) can be found in Chap. 3.

Note that the frequency-domain SFS method in this section is similar but is not identical to FSF method in Chap. 3. The main difference is that we use the radial-dependent outgoing field coefficients $\mathsf{P}^{\mathrm{o}}_{u,v}(\omega,R)$ and the radial-dependent incoming field coefficients $\mathsf{P}^{\mathrm{i}}_{u,v}(\omega,R)$ rather than the radial-independent spherical Hankel function coefficients $\mathsf{A}_{u,v}(\omega)$ and the radial-independent spherical Bessel function coefficients $\mathsf{B}_{u,v}(\omega)$ to denote the outgoing field and the incoming field, respectively. The reason for the modification is given in later part of this chapter.

The outgoing field separated by the frequency-domain method reveals the characteristics of the target source. However, the frequency-domain method needs the STFT process to transfer the time-domain acoustic quantities into the time-frequency domain. The inherent frame delay of the STFT process makes the separated outgoing field not suitable for time-critical applications, such as active noise control and real-time beamforming [8]. A time-domain method, which does not introduce the frame delay, is more appropriate to separate the non-stationary sound field for applications which have low-latency requirements.

4.3 The time-domain sound field separation method on a sphere

In this section, we derive the time-domain SFS method on a sphere. We start with a theorem, which provides the expressions of the outgoing field and the incoming field on a sphere. The detailed derivation of the theorem is at the end of this chapter.

Theorem 1 *The time-domain outgoing field $p^{\mathrm{o}}(t,R,\Theta)$ and the time-domain incoming field $p^{\mathrm{i}}(t,R,\Theta)$ on a sphere $\mathbb{S}_2$ of radius R are*

$$p^{\mathrm{o}}(t,R,\Theta) = \sum_{u=0}^{\infty}\sum_{v=-u}^{u} \mathsf{p}^{\mathrm{o}}_{u,v}(t,R) Y_{u,v}(\Theta), \tag{4.6}$$

and

$$p^{\mathrm{i}}(t,R,\Theta) = \sum_{u=0}^{\infty}\sum_{v=-u}^{u} \mathsf{p}^{\mathrm{i}}_{u,v}(t,R) Y_{u,v}(\Theta), \tag{4.7}$$

respectively, where the outgoing field coefficients $\mathsf{p}^{\mathrm{o}}_{u,v}(t,R)$ and the incoming field coefficients $\mathsf{p}^{\mathrm{i}}_{u,v}(t,R)$ are given by

$$\begin{aligned}\mathsf{p}^{\mathrm{o}}_{u,v}(t,R) = \int_{\Theta}\Big\{&\int_0^{2\tau_R} g^u_0(\tau)p(t-\tau,R,\Theta)d\tau\\ &+\int_0^{2\tau_R} g^u_1(\tau)\frac{dp(t-\tau,R,\Theta)}{dt}d\tau\\ &+\rho c\int_0^{2\tau_R} g^u_2(\tau)\frac{d\nu(t-\tau,R,\Theta)}{dt}d\tau\Big\}Y_{u,v}(\Theta)d\Theta,\end{aligned} \tag{4.8}$$

and

$$\begin{aligned}\mathsf{p}^{\mathrm{i}}_{u,v}(t,R) = \int_{\Theta}\Big\{&\int_0^{2\tau_R} g^u_3(\tau)p(t-\tau,R,\Theta)d\tau\\ &+\int_0^{2\tau_R} g^u_4(\tau)\frac{dp(t-\tau,R,\Theta)}{dt}d\tau\\ &-\rho c\int_0^{2\tau_R} g^u_2(\tau)\frac{d\nu(t-\tau,R,\Theta)}{dt}d\tau\Big\}Y_{u,v}(\Theta)d\Theta,\end{aligned} \tag{4.9}$$

respectively, $p(t,R,\Theta)$ and $\nu(t,R,\Theta)$ are the pressure and the radial particle velocity on the sphere at the point (R,Θ), respectively, and $\int_{\Theta}(\cdot)d\Theta \equiv \int_0^{2\pi}\int_0^{\pi}(\cdot)\sin\theta d\theta d\phi$. The expressions of $g^u_0(t)$, $g^u_1(t)$, $g^u_2(t)$, $g^u_3(t)$, and $g^u_4(t)$ are

$$\begin{aligned}g^u_0(t) = &\frac{u}{4\tau_R}\sum_{v=0}^{u}\sum_{\varsigma=0}^{u}\frac{\varphi_v(u)\varphi_\varsigma(u)}{(v+\varsigma)!}\\ &\times\Big[(-1)^v\Big(\frac{t}{\tau_R}\Big)^{v+\varsigma}\mathtt{Sign}(t)+(-1)^{u+1}\Big(\frac{t-2\tau_R}{\tau_R}\Big)^{v+\varsigma}\mathtt{Sign}(t-2\tau_R)\Big],\end{aligned} \tag{4.10}$$

$$g_1^u(t) = \frac{1}{4}\sum_{v=0}^{u+1}\sum_{\varsigma=0}^{u}\frac{\varphi_v(u+1)\varphi_\varsigma(u)}{(v+\varsigma)!} \times\Big[(-1)^v\Big(\frac{t}{\tau_R}\Big)^{v+\varsigma}\mathtt{Sign}(t) + (-1)^u\Big(\frac{t-2\tau_R}{\tau_R}\Big)^{v+\varsigma}\mathtt{Sign}(t-2\tau_R)\Big], \quad (4.11)$$

$$g_2^u(t) = \frac{1}{4}\sum_{v=0}^{u}\sum_{\varsigma=0}^{u}\frac{\varphi_v(u)\varphi_\varsigma(u)}{(v+\varsigma)!} \times\Big[(-1)^v\Big(\frac{t}{\tau_R}\Big)^{v+\varsigma}\mathtt{Sign}(t) + (-1)^{u+1}\Big(\frac{t-2\tau_R}{\tau_R}\Big)^{v+\varsigma}\mathtt{Sign}(t-2\tau_R)\Big], \quad (4.12)$$

$$g_3^u(t) = \frac{u+1}{4\tau_R}\sum_{v=0}^{u}\sum_{\varsigma=0}^{u}\frac{\varphi_v(u)\varphi_\varsigma(u)}{(v+\varsigma)!} \times\Big[(-1)^v\Big(\frac{t}{\tau_R}\Big)^{v+\varsigma}\mathtt{Sign}(t) + (-1)^{u+1}\Big(\frac{t-2\tau_R}{\tau_R}\Big)^{v+\varsigma}\mathtt{Sign}(t-2\tau_R)\Big], \quad (4.13)$$

$$g_4^u(t) = \begin{cases} \frac{1}{4}\sum_{v=0}^{u}\sum_{\varsigma=0}^{u-1}\frac{\varphi_v(u)\varphi_\varsigma(u-1)}{(v+\varsigma)!}\Big[(-1)^v(\frac{t}{\tau_R})^{v+\varsigma}\mathrm{Sign}(t) \\ \qquad +(-1)^{u-1}(\frac{t-2\tau_R}{\tau_R})^{v+\varsigma}\mathrm{Sign}(t-2\tau_R)\Big], & u>0, \\ \frac{1}{4}\mathrm{Sign}(t) - \frac{1}{4}\mathrm{Sign}(t-2\tau_R), & u=0, \end{cases} \quad (4.14)$$

respectively, where $\mathtt{Sign}(\cdot)$ *is the sign function,*

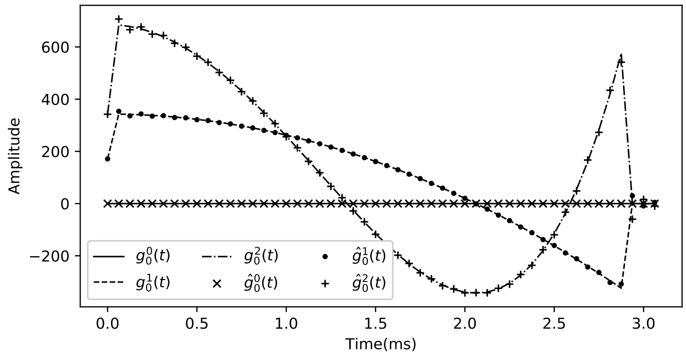
$$\varphi_v(u) = \frac{(u+v)!}{2^v v!(u-v)!},\ v\in[1,u], u\geq 0, \quad (4.15)$$

and $\varphi_\varsigma(\cdot)$ *is defined similar to* $\varphi_v(\cdot)$ *[72].*

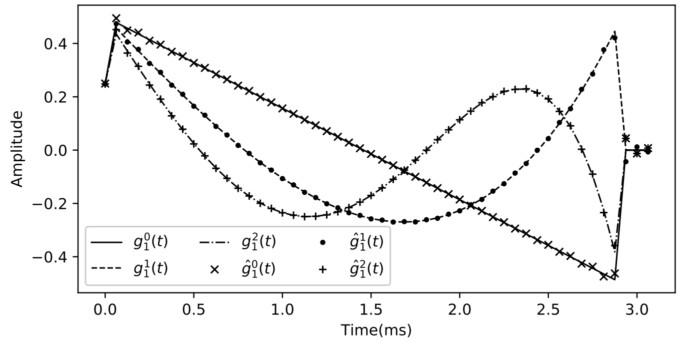
Note that $g_0^u(t) = ug_2^u(t)/\tau_R$ *and* $g_3^u(\tau_R) = (u+1)g_2^u(t)/\tau_R$.

We have the following comments on the time-domain SFS method:

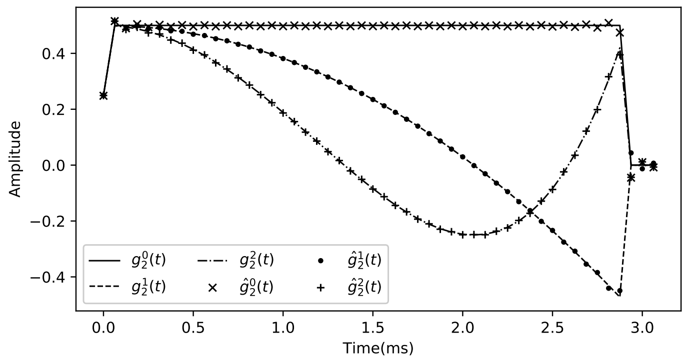
1. Equations (4.6), (4.7), (4.8), (4.9) show that the outgoing (incoming) field at a particular point (R,Θ) can be described by the pressure, the time derivative of the pressure, and the time derivative of the radial partial velocity observed over the sphere within a certain time frame.



(a)



(b)



(c)

Figure 4.2: The plots of the theoretically derived $g_0^u(t)$, $g_1^u(t)$, $g_2^u(t)$ and the numerically calculated $\widehat{g}_0^u(t)$, $\widehat{g}_1^u(t)$, $\widehat{g}_2^u(t)$ for $u = 0, 1, 2$, $R = 0.5$ m, $c = 343$ m/s, and $\tau_R \approx 1.458$ ms.

2. The spatial filter $g_0^u(t)$, $g_1^u(t)$, $g_2^u(t)$, $g_3^u(t)$, and $g_4^u(t)$ are specially derived such that they take finite values inside of the time interval $[0, 2\tau_R]$ and are zero outside of the interval.

3. We take a pressure-velocity formulation. That is, we recover the outgoing field and the incoming field based on the pressure and the radial particle velocity measurement on a sphere [106]. This formulation allows us to use the Wronskian equation to simplify the expressions of the frequency-domain outgoing and incoming field coefficients ((4.4) and (4.5)) [8, 36], such that they do not contain the divide-operations. The corresponding expressions of the time-domain outgoing and incoming field coefficients ((4.8) and (4.9)) then do not include any deconvolution operations. It is possible to take a pressure-pressure formulation [106], and recover the outgoing field and the incoming field based on the pressure measurements on two adjacent spheres. However, the resulting time-domain outgoing and incoming field coefficients contain the deconvolution operations, which are computationally expensive to implement.

4. Note that we can express the spatial filters using the Legendre functions. Based on the fact that the inverse Fourier transform of a spherical Bessel function is a Legendre function (2.47) and the expansion of the spherical Hankel function (2.48) [72], we can derive the spatial filter $g_0^u(t)$ as (the detailed derivations are not shown for brevity)

$$\begin{aligned} g_0^u(t) &= \mathcal{F}^{-1}\Big[- u(i\omega\tau_R) j_u(\omega\tau_R) h_u(\omega\tau_R)\Big] \\ &= u\frac{(-1)^u}{2}\Big[\frac{1}{\tau_R}\mathcal{P}_u(\frac{t-\tau_R}{\tau_R})\Big] \\ &\quad + u\frac{(-1)^u}{2}\sum_{v=1}^{u}\frac{\varphi_v(u)}{(v-1)!}\frac{1}{\tau_R^{v+1}}\Big[\mathcal{P}_u(\frac{t-\tau_R}{\tau_R}) * t^{v-1}\mathtt{U}(t)\Big], \end{aligned} \tag{4.16}$$

which is the accumulated convolutions between the Legendre function $\mathcal{P}_u(\cdot)$ and terms like $t^{v-1}\mathtt{U}(t)$, where $\mathtt{U}(t)$ is the step function. We can also express the spatial filters $g_1^u(t)$, $g_2^u(t)$, $g_3^u(t)$, and $g_4^u(t)$ in a similar form as (4.16) using the Legendre functions. Equation (4.16) reveals the Legendre function shapes of the spatial filters. However, the convolution operation in (4.16) makes its evaluation not straightforward. Thus we choose to derive the spatial filters using the `Sign` function, and the resulting expressions of the spatial filters are easy to implement in practice, such as on a DSP processor.

5. The frequency-domain outgoing and incoming field coefficients $\{\mathsf{P}^{\mathrm{o}}_{u,v}(\omega, R), \mathsf{P}^{\mathrm{i}}_{u,v}(\omega, R)\}$ and their time-domain counterparts $\{\mathsf{p}^{\mathrm{o}}_{u,v}(t, R), \mathsf{p}^{\mathrm{i}}_{u,v}(t, R)\}$ depend on the radius R of the sound field separation sphere $\mathbb{S}_2$. We can use the spherical Bessel function coefficients and the spherical Hankel function coefficients $\{\mathsf{A}_{u,v}(\omega), \mathsf{B}_{u,v}(\omega)\}$ to characterize the outgoing field and the incoming field, and derive their corresponding time-domain counterparts $\{\mathsf{a}_{u,v}(t), \mathsf{b}_{u,v}(t)\}$. Using $\{\mathsf{a}_{u,v}(t), \mathsf{b}_{u,v}(t)\}$, we not only can recover the outgoing field and the incoming field on a sphere, but also can project the outgoing field and the incoming field from one sphere to another [36]. However, the non-robustness issue of the spherical Bessel function inversion at the low frequency range and the problem of spherical Bessel function zeros still exist. This follow up work is out the scope of this thesis and will be our future work.

6. Figure 4.2 depicts $g^u_0(t)$, $g^u_1(t)$, $g^u_2(t)$ and their estimations $\widehat{g}^u_0(t)$, $\widehat{g}^u_1(t)$, $\widehat{g}^u_2(t)$ for $u = 0, 1, 2$, $R = 0.5$ m, $c = 343$ m/s, and $\tau_R = R/c \approx 1.458$ ms. We plot $g^u_0(t)$, $g^u_1(t)$, $g^u_2(t)$ based on the theoretical expressions (4.10), (4.11), and (4.12), respectively. We plot $\widehat{g}^u_0(t)$, $\widehat{g}^u_1(t)$, and $\widehat{g}^u_2(t)$ based on the numerical calculation of (4.31), (4.32), and (4.33), respectively. As shown in Fig. 4.2, the theoretically derived $g^u_0(t)$, $g^u_1(t)$, $g^u_2(t)$ agree with the numerically calculated $\widehat{g}^u_0(t)$, $\widehat{g}^u_1(t)$, and $\widehat{g}^u_2(t)$, respectively.

7. The main limitation of the proposed method is that to implement the method we have to surround the sound source with a spherical (or hemispherical) sensor array. However, that is not always possible.

We provide the signal flow diagram of the proposed method ((4.6) and (4.8)) in Fig. 4.3. The pressure, the time derivative of the pressure, and the time derivative of the radial particle velocity are decomposed into coefficients. By summing the convolutions between the coefficients with the spatial filters $g^u_0(t)$, $g^u_1(t)$, and $g^u_2(t)$, we obtain the outgoing field coefficients $\mathsf{p}^{\mathrm{o}}_{u,v}(t, R)$. By summing the multiplications

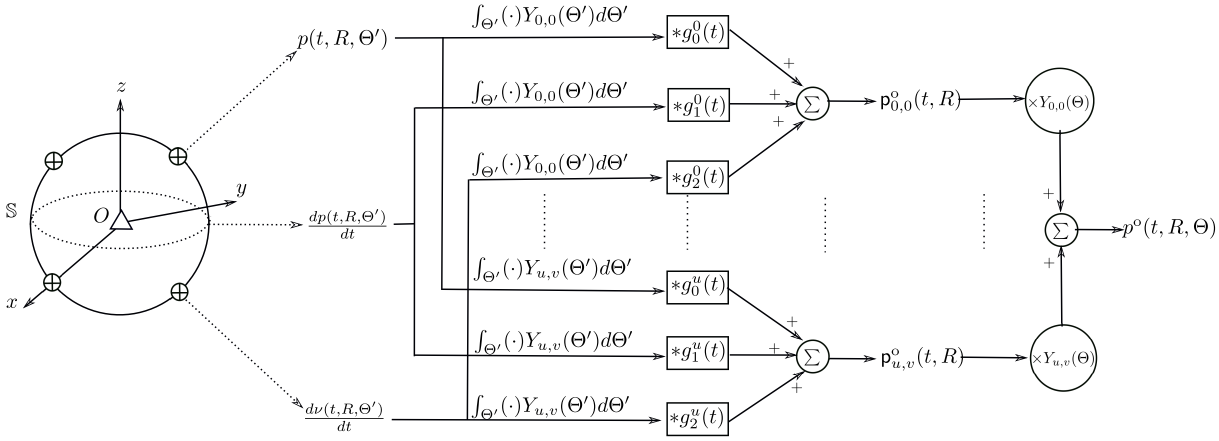


Figure 4.3: The signal flow diagram of the proposed method: We decompose the pressure $p(t, R, \Theta')$, the time derivative of the pressure $dp(t, R, \Theta')/dt$, and the time derivative of the radial particle velocity $d\nu(t, R, \Theta')/dt$ on a sphere into coefficients. By summing the convolutions between the coefficients with the spatial filters $g_0^u(t)$, $g_1^u(t)$, and $g_2^u(t)$, we obtain the outgoing field coefficients $\mathsf{p}_{u,v}^{\mathrm{o}}(t, R)$. By summing the multiplications between the outgoing field coefficients with the spherical harmonics $Y_{u,v}(\Theta)$, we obtain the outgoing field pressure $p^{\mathrm{o}}(t, R, \Theta)$.

between the outgoing field coefficients with the spherical harmonics $Y_{u,v}(\Theta)$, we obtain the outgoing field pressure $p^{\mathrm{o}}(t, R, \Theta)$.

The main limitation of the proposed method is that to implement the method we have to surround the sound source with a sphere (or hemisphere) of microphones or vector sensors. However, that is not always possible. Furthermore, the method works best at the low frequency range, where a small number of microphones or vector sensors are needed for sound field separation.

4.4 Realization of the time-domain sound field separation method

The equations (4.6), (4.7), (4.8), and (4.9) of the proposed method involve continuous integrals, the time derivative of the sound pressure, and the time derivative of the radial particle velocity, which are acoustic quantities that can not be measured directly by existing sensors. In this section, we provide details to facilitate the implementation of the proposed method.

4.4.1 Discretization and sampling

We first approximate the continuous integral in (4.8) by a discrete form as follows

$$
\begin{aligned}
\mathsf{p}^{\mathrm{o}}_{u,v}(t,R) \approx& \int_0^{2\tau_R} g_0^u(\tau) \int p(t-\tau,R,\Theta) Y_{u,v}(\Theta) d\Theta d\tau \\
&+ \int_0^{2\tau_R} g_1^u(\tau) \int \frac{p(t-\tau,R,\Theta)-p(t-\delta_t-\tau,R,\Theta)}{\delta_t} Y_{u,v}(\Theta) d\Theta d\tau \\
&+\rho c \int_0^{2\tau_R} g_2^u(\tau) \int \frac{\nu(t-\tau,R,\Theta)-\nu(t-\delta_t-\tau,R,\Theta)}{\delta_t} Y_{u,v}(\Theta) d\Theta d\tau \\
\approx& \int_\Theta \Big\{ \sum_{n=0}^{T_n} g_0^u(\tau_n) p(t-\tau_n,R,\Theta)\delta_\tau \\
&+\sum_{n=0}^{T_n} g_1^u(\tau_n) \frac{p(t-\tau_n,R,\Theta)-p(t-\delta_t-\tau_n,R,\Theta)}{\delta_t}\delta_\tau \\
&+\rho c \sum_{n=0}^{T_n} g_2^u(\tau_n) \frac{\nu(t-\tau_n,R,\Theta)-\nu(t-\delta_t-\tau_n,R,\Theta)}{\delta_t}\delta_\tau \Big\} Y_{u,v}(\Theta) d\Theta \\
\approx& \int_\Theta \Big\{ \sum_{n=0}^{T_n} g_0^u(\tau_n) p(t-\tau_n,R,\Theta)\delta_t \\
&+ \sum_{n=0}^{T_n} g_1^u(\tau_n)[p(t-\tau_n,R,\Theta)-p(t-\delta_t-\tau_n,R,\Theta)] \\
&+\rho c \sum_{n=0}^{T_n} g_2^u(\tau_n)[\nu(t-\tau_n,R,\Theta)-\nu(t-\delta_t-\tau_n,R,\Theta)] \Big\} Y_{u,v}(\Theta) d\Theta,
\end{aligned} \tag{4.17}
$$

where $\delta_\tau = \delta_t = 1/f_\mathrm{s}$, and $T_n = \lceil 2 f_\mathrm{s} \tau_R \rceil + 1$ is the number of samples corresponding to $2\tau_R$ (f_s is the sampling frequency).

We further develop (4.17) by replacing the continuous integral over the sphere with a finite summation at $\{(R,\Theta_q)\}_{q=1}^{Q}$ and by substituting the continuous time indices t and τ_n using the discrete time indices n and n', respectively. We denote $p(n,R,\Theta_q)$ and $\nu(n,R,\Theta_q)$ as the pressure and the radial partial velocity at (R,Θ_q) at discrete time instant n. We have the outgoing field coefficient $\mathsf{p}^{\mathrm{o}}_{u,v}(n,R)$ at discrete time n as

$$\begin{aligned}\mathsf{p}^{\mathrm{o}}_{u,v}(n,R) \approx \sum_{q=1}^{Q} \gamma_q Y_{u,v}(\Theta_q) \sum_{n'=0}^{T_n} \Big\{ & g_0^u(n') p(n-n',R,\Theta_q)\delta_t \\ & + g_1^u(n')[p(n-n',R,\Theta_q) - p(n-1-n',R,\Theta_q)] \\ & + \rho c g_2^u(n')[\nu(n-n',R,\Theta_q) - \nu(n-1-n',R,\Theta_q)] \Big\},\end{aligned} \tag{4.18}$$

where $\{\gamma_q\}_{q=1}^{Q}$ are the sampling weights [75].

Substitution of $\mathsf{p}^{\mathrm{o}}_{u,v}(t,R)$ by $\mathsf{p}^{\mathrm{o}}_{u,v}(n,R)$ into (4.6) results in an estimation of the outgoing field $p^{\mathrm{o}}(n,R,\Theta)$ at discrete time instance on the sphere $\mathbb{S}_2$.

$$\begin{aligned}\widehat{p}^{\mathrm{o}}(n,R,\Theta) \approx & \sum_{u=0}^{N} \sum_{v=-u}^{u} Y_{u,v}(\Theta) \sum_{q=1}^{Q} \gamma_q Y_{u,v}(\Theta_q) \\ & \times \sum_{n'=0}^{T_n} \Big\{ g_0^u(n') p(n-n',R,\Theta_q)\delta_t \\ & + g_1^u(n')[p(n-n',R,\Theta_q) - p(n-1-n',R,\Theta_q)] \\ & + \rho c g_2^u(n')[\nu(n-n',R,\Theta_q) - \nu(n-1-n',R,\Theta_q)] \Big\},\end{aligned}$$

For practical implementation, we realize the spatial filers $g_0^u(n)$, $g_1^u(n)$, and $g_2^u(n)$ as finite impulse response filters of length $T_n = \lceil 2R/c * f_{\mathrm{s}} \rceil$. The computational complexity, specifically, the number of multiplication operations, involved in above equations is $\mathrm{C_c} = (N+1)^2 \times Q \times 3(T_n+1)$. To avoid spatial aliasing, we choose $Q \geq (N+1)^2$, and thus $\mathrm{C_c} \geq 3(N+1)^4(T_n+1)$. Then for $R = 0.5$ m, $c = 343$ m/s, $f_{\mathrm{s}} = 48000$ Hz, $N = 3$, we have $\mathrm{C_c} \geq 3 \times 256 \times 142 \approx 10^5$. That is, to recover the outgoing field at one point on a sphere of radius 0.5 m, we need more than 10^5 multiplication operations per sampling period. However, as shown in Fig. 4.3, the operation among different mode order u is highly parallel, and thus we can use a field-gate-programming-array to speed up the the implementation of above equation.

We can obtain the discrete form and the computational complexity of (4.7) in a similar manner.

Due to the finite sampling over the sphere, we can only recover the outgoing field coefficients up to N, where $N \leq \sqrt{Q}-1$ is the truncation order of the outgoing field [73].

4.4.2 Pressure and velocity approximation

We may use an array of vector sensors, which can measure both the pressure and the particle velocity [120] to realize the SFS method. Alternatively, we can use two microphone arrays to realize the method.

Place Q microphones on a sphere of radius $R - \delta_R$ at $\{R - \delta_R, \Theta_q\}_{q=1}^{Q}$, and another Q microphones on a sphere of radius $R + \delta_R$ at $\{R + \delta_R, \Theta_q\}_{q=1}^{Q}$, where $\delta_R \ll R$ is a real positive number. We approximate the pressure on the middle sphere of radius R at positions $\{R, \Theta_q\}_{q=1}^{Q}$ by [6]

$$p(n, R, \Theta_q) \approx \frac{1}{2}[p(n, R + \delta_R, \Theta_q) + p(n, R - \delta_R, \Theta_q)]. \tag{4.19}$$

Based on the Euler's equation, we approximate the radial particle velocity on the middle sphere of radius R at positions $\{R, \Theta_q\}_{q=1}^{Q}$ by [6]

$$\nu(n, R, \Theta_q) \approx \nu(n-1, R, \Theta_q) + \frac{1}{\rho}\frac{p(n, R + \delta_R, \Theta_q) - p(n, R - \delta_R, \Theta_q)}{2\delta_R}\delta_t. \tag{4.20}$$

Substituting (4.19) and (4.20) into (4.18), we obtain an estimation of the time-domain outgoing field coefficient $\mathsf{p}_{u,v}^{o}(n, R)$.

Because we approximate the velocity by the pressure difference between two closely spaced microphones, there will be errors when recovering the high frequency components of the outgoing field. However, we can extend the method to different frequency bands by using microphone pairs with different spacing and increasing the number of pairs with frequency. We will explore this in a future study.

4.5 Simulation examples

In this section, we conduct simulations to validate the effectiveness of the time-domain SFS method.

4.5.1 Sound field separation in free field

The simulation environment is shown in Fig. 4.1 (a). We place a point source at $(0, 0, 0.3)$ m as the target source. We let 100 plane waves, whose directions are determined according to the 100-point spherical packing [121], be the incoming fields. We let the target source output a unit-variance Gaussian white noise, which has been filtered by a 64-tap Bandpass Butterworth filter with a frequency range $[100, 600]$ Hz. The incoming plane waves have the same frequency components as the target source output. The strengths of the plane waves are normalized such that their contributions to the sound pressure on the sphere are approximately the same as the contribution from the target source.

We aim to estimate the target (outgoing) field on the sphere $\mathbb{S}_2$ produced by the target source. We choose the radius of the sphere $\mathbb{S}_2$ as $R = 0.65$ m, and the outgoing field truncation order to be $N = 5$. We use a sensor array arranged on the sphere according to the 6-th order Gauss sampling scheme to realize the time-domain SFS method [75]. The reason that we use the 6-th order Gauss sampling scheme to recover the outgoing field (coefficients) up to 5-th order is to reduce the effect of incoming field (coefficients) aliasing. The incoming field, which has high order coefficients, contaminates the estimation of the 6-th order outgoing field (coefficients). Thus the 6-th order outgoing field (coefficients) is not estimated. Through over-sampling on the sphere, the outgoing field coefficients of lower orders ($N \leq 5$) are more accurately estimated.

We simulate the impulse responses, including the radial particle velocity response, between the sources and the sensors based on the free-space Green function and the Euler's equation [36]. The impulse responses are truncated to 1024 taps long under the sampling frequency $f_\mathrm{s} = 48$ kHz. The speed of sound is $c = 343$ m/s, and the air density is $\rho = 1.225$ kg/m^3. The length of the spatial filters $g_0^u(t)$, $g_1^u(t)$, $g_2^u(t)$ is $T_n = \lceil 2Rf_\mathrm{s}/c \rceil + 1 = 183$ taps. We add Gaussian white noise to the sensor measurement such that the signal to noise power ratio is 40 dB. The simulation results are from an average of 100 independent runs. The settings in this paragraph are used in all simulations unless otherwise stated.

We reconstruct the outgoing field on the sphere over a period of 10 ms according to (4.6). We depict the amplitudes of the outgoing field $p^\mathrm{o}(t, R, \Theta_{17})$, the total

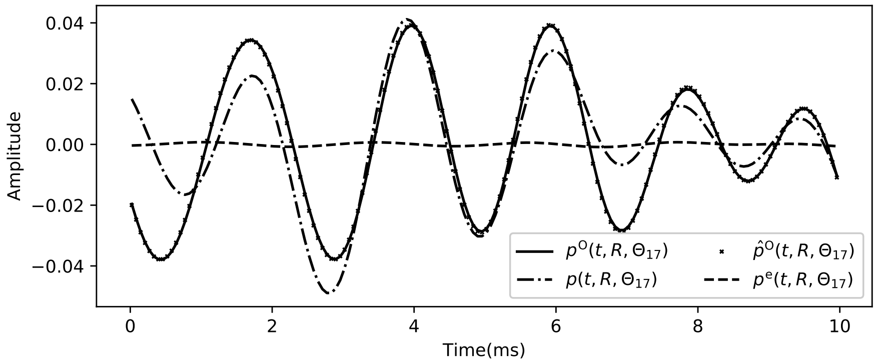


Figure 4.4: Sound field separation at a point: The amplitudes of the outgoing field $p^{\mathrm{o}}(t, R, \Theta_{17})$, the total field $p(t, R, \Theta_{17})$, the separated outgoing field $\widehat{p}^{\mathrm{o}}(t, R, \Theta_{17})$, and the outgoing field separation error $p^{\mathrm{e}}(t, R, \Theta_{17})$, at the 17-th sensor over 10 ms.

field $p(t, R, \Theta_{17})$, the separated outgoing field $\widehat{p}^{\mathrm{o}}(t, R, \Theta_{17})$, and the outgoing field separation error $p^{\mathrm{e}}(t, R, \Theta_{17}) = p^{\mathrm{o}}(t, R, \Theta_{17}) - \widehat{p}^{\mathrm{o}}(t, R, \Theta_{17})$, at the 17-th sensor in Fig. 4.4. The total field is the summation of the outgoing field and the incoming plane wave fields.

As shown in Fig. 4.4, the total field differs from the outgoing field. The separated outgoing field approximates the outgoing field, and the sound field separation error is small over the whole 10 ms period. The normalized sound separation error at the 17-th sensor over the observation period is

$$\begin{aligned} \xi_{17} &= 10 \log 10 \frac{\sum_{t=0}^{10 \text{ ms}} ||p^{\mathrm{e}}(t, R, \Theta_{17})||_2^2}{\sum_{t=0}^{10 \text{ ms}} ||p^{\mathrm{o}}(t, R, \Theta_{17})||_2^2} \\ &= -30.1 \text{ dB}. \end{aligned} \tag{4.21}$$

We depict the normalized sound field separation error ξ_{Θ_q} on the sphere $\mathbb{S}_2$ over the 10 ms in Fig. 4.5, where ξ_{Θ_q} is defined as

$$\xi_{\Theta_q} = 10 \log 10 \frac{\sum_{t=0}^{10 \text{ ms}} ||p^{\mathrm{o}}(t, R, \Theta_q) - \widehat{p}^{\mathrm{o}}(t, R, \Theta_q)||_2^2}{\sum_{t=0}^{10 \text{ ms}} ||p^{\mathrm{o}}(t, R, \Theta_q)||_2^2}, \tag{4.22}$$

$p^{\mathrm{o}}(t, R, \Theta_q)$ and $\widehat{p}^{\mathrm{o}}(t, R, \Theta_q)$ are the outgoing field and its estimation at the point (R, Θ_q) respectively, $(R, \Theta_q)_{q=1}^{180\times 360}$ are equal-angle sampling point positions on the sphere $\mathbb{S}_2$ [75]. As shown in Fig. 4.5, the proposed method is able to recover the

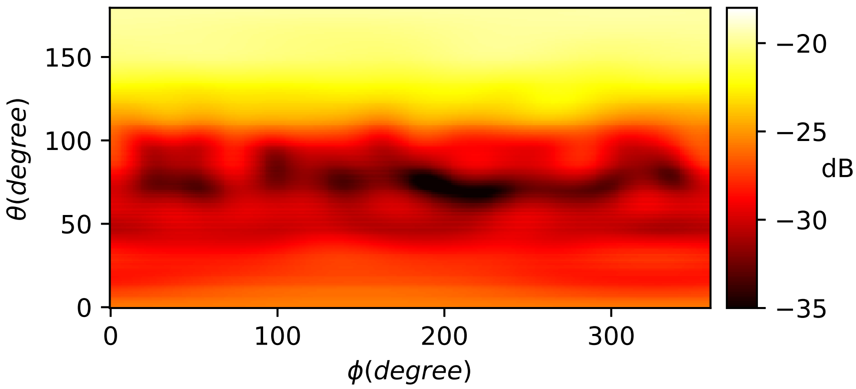


Figure 4.5: Sound field separation over a sphere: Normalized sound field separation error ξ_{Θ_q} on the sphere $\mathbb{S}_2$ over 10 ms, $(R, \Theta_q)_{q=1}^{180\times 360}$ are equal angle sampling point positions on the sphere $\mathbb{S}_2$.

outgoing field accurately, and the normalized sound field separation error ξ_{Θ_q} is less than -18 dB over the whole sphere.

We further investigate the performance of the proposed method in terms of the frequency and the truncation order N. Define the normalized separation error over all the sampling points as

$$\xi^N(\omega) = 10 \log10 \frac{\sum_{q=1}^{98} ||p^{\mathrm{o}}(\omega, R, \Theta_q) - \widehat{p}^{\mathrm{o},N}(\omega, R, \Theta_q)||_2^2}{\sum_{q=1}^{98} ||p^{\mathrm{o}}(\omega, R, \Theta_q)||_2^2}, \tag{4.23}$$

where $p^{\mathrm{o}}(\omega, R, \Theta_q)$ is the frequency-domain outgoing field, $\widehat{p}^{\mathrm{o},N}(\omega, R, \Theta_q)$ is the frequency-domain counterpart of the following equation

$$\widehat{p}^{\mathrm{o},N}(t, R, \Theta_q) = \sum_{u=0}^{N} \sum_{v=-u}^{u} \widehat{\mathsf{P}}^{\mathrm{o}}_{u,v}(t, R) Y_{u,v}(\Theta_q), \tag{4.24}$$

$\widehat{\mathsf{P}}^{\mathrm{o}}_{u,v}(t, R)$ are the time-domain outgoing field coefficient obtained through (4.18), $\{R, \Theta_q\}_{q=1}^{98}$ are the sampling point positions from the 6-th order Gauss sampling scheme, and $\omega = 200\pi, 202\pi, ..., 1200\pi$ rad/s (or $f = 100, 101, ..., 600$ Hz).

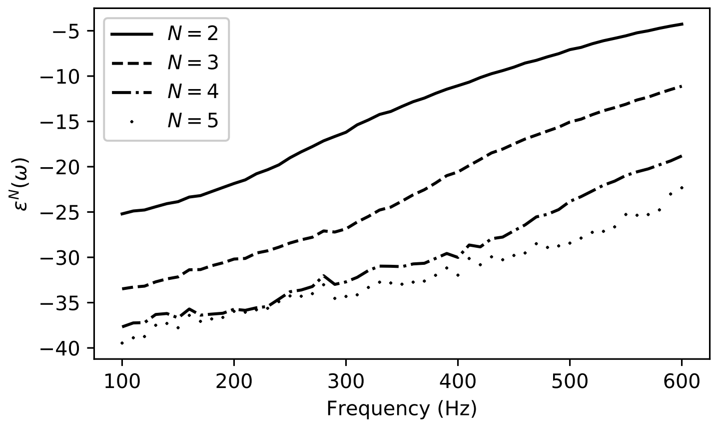


Figure 4.6: Sound field separation at the sampling points: The normalized separation error $\xi^N(\omega)$ as a function of the frequency and the truncation order N.

We plot the normalized separation error $\xi^N(\omega)$ in Fig. 4.6, which shows that $\xi^N(\omega)$ increases along with the frequency. The higher the outgoing field truncation order N is the smaller the normalized separation error $\xi^N(\omega)$ is. However, at the low frequency range, the truncation order $N = 5$ does not make the normalized separation error $\xi^N(\omega)$ significantly smaller than that in the case of $N = 4$, i.e., $\xi^5(\omega) \approx \xi^4(\omega)$ for $\omega < 600\ \pi\ rad/s$ or $f < 300$ Hz. To conclude, we should choose an appropriate truncation order N for the separated target (outgoing) field according to its frequency components such that the separation error is sufficiently small.

4.5.2 Comparisons with the Spatial Fourier Transform based method

In this section, we compare the performance of the proposed method with the Spatial Fourier Transform based method [117]. We make a comparison between the two methods because they both use some theoretically derived spatial filters for SFS. The comparison is to show that the derivation of the spatial filters influences the accuracy of the separated field.

The setup of the simulation environment is shown in Fig. 4.7, which has a Cartesian coordinate system with the origin at the point O. On the $z = 0$ m plane

H, we equally arrange 15×15 sensors on a square of size 0.7 m × 0.7 m, whose center is at the point O. On the $z = 0.01$ m plane H_1, we arrange another 15×15 sensors equally on a square of size 0.7 m × 0.7 m, whose center is the at the point $(x, y, z) = (0.0, 0.0, 0.01)$ m. We use the sensors on these two planes for SFS using the Spatial Fourier Transform based method. We extend the array size to be 81×81 by zero-padding to reduce the aliasing error of the Spatial Fourier Transform [117].

There is a sphere $\mathbb{S}_2$ of radius $R = 0.35$ m. On the sphere $\mathbb{S}_2$, we place sensors according to the 9-th order Gauss sampling scheme [75]. We use the sensors on this sphere $\mathbb{S}_2$ to estimate the outgoing field up to 6-th order using the proposed method.

We use the two sampling schemes because they can detect approximately the same number of acoustic quantities, and the sizes of the two arrays are comparable.

We have a target point source at $(0.0, 0.0, -0.2)$ m, and another interfering point source at $(0, 0, 0.2)$ m as shown in Fig. 4.7. Outputs of these two sources are unit-variance Gaussian white noises filtered a 64-tap Bandpass Butterworth filter of frequency range $[100, 1000]$ Hz. For the Spatial Fourier Transform based method, we truncate the Bessel function, i.e., (7) from [117] to 1024 taps long. In the proposed method, the length of spatial filters $g_0^u(t)$, $g_1^u(t)$, $g_2^u(t)$ is $T_n = \lceil 2Rf_\mathrm{s}/c \rceil + 1 = 99$ taps.

To make the comparison fair, only the outgoing sound at the point O (the sound produced by the target source in the positive z direction), where the plane H is tangent to the sphere $\mathbb{S}_2$, is separated by both methods. The overall duration of the SFS process is about 30 ms.

The target (outgoing) sound $p^\mathrm{d}(t)$, the total sound $p^\mathrm{t}(t)$, and the sound separation errors using the proposed $p^\mathrm{e}_\mathrm{pro}(t) = p^\mathrm{d}(t) - \widehat{p}^\mathrm{d}_\mathrm{pro}(t)$ and the Spatial Fourier Transform based method $p^\mathrm{e}_\mathrm{SFT}(t) = p^\mathrm{d}(t) - \widehat{p}^\mathrm{d}_\mathrm{SFT}(t)$ are shown in Fig. 4.8. $\widehat{p}^\mathrm{d}_\mathrm{pro}(t)$ and $\widehat{p}^\mathrm{d}_\mathrm{SFT}(t)$ are the separated target fields obtained by the proposed and the Spatial Fourier Transform based method, respectively. The target (outgoing) sound is obtained through the convolution between the source output and the free-space Green function. The total sound is the summation of the target (outgoing) sound and the interfering (incoming) sound.

As shown in Fig. 4.8, the proposed method can recover the target sound accurately over the whole 30 ms period. The Spatial Fourier Transform based method

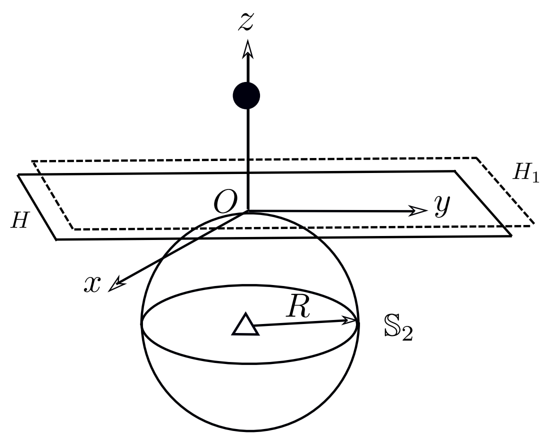


Figure 4.7: Setup of the sound field separation at a point: The proposed method use the sensors on the sphere $\mathbb{S}_2$ to separate the outgoing field at the point O, while the Spatial Fourier Transform based method use sensors on the plane H and the plane H_1 to separate the sound waves at the point O produced by the target source $\triangle$ and the disturbing source $\bullet$.

can separate the target field with small errors before $t = 5$ ms, but after that the SFS errors start to increase. That is because, in the Spatial Fourier Transform based method, the target sound at one time instant depends on all previous sounds [117]. The truncation of the infinitely long Bessel function, i.e., (7) from [117] to a finite length introduces errors. We have conducted simulations with a longer truncation length (8192 taps) for the Bessel functions. However, we found that the field separation errors are similar to the line with the plus sign (+) in Fig. 4.8, and thus the simulation results are not shown. In the proposed method, we have specially derived the spatial filter functions, $g_0^u(t)$, $g_1^u(t)$, $g_2^u(t)$, to be finite long, avoiding SFS error due to the spatial filter truncation.

The simulation results demonstrate that the Spatial Fourier Transform based method is best suitable for separating transient plane wave fields. The proposed method, on the other hand, can accurately separate the target (outgoing) sound field on a spherical array over a longer duration.

4.5.3 Sound field separation in a room

In practical implementations of the proposed SFS method, the target sources may be placed on a surface. In case the surface is rigid or highly reflective, we can mea-

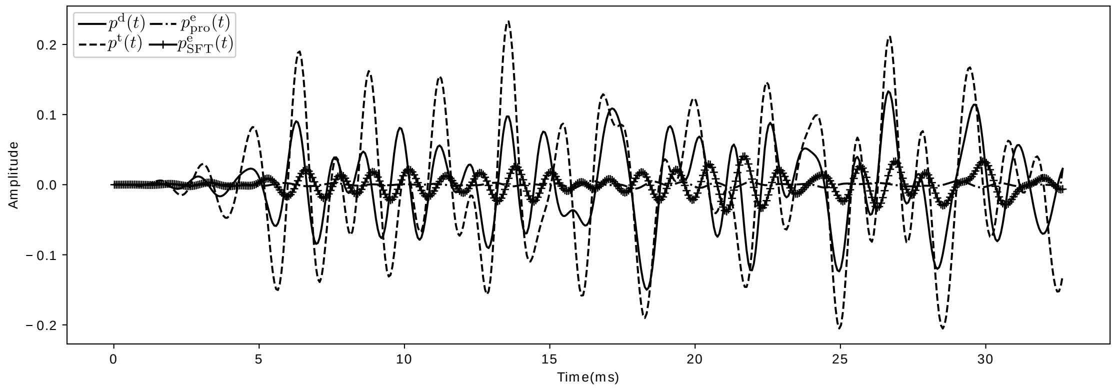


Figure 4.8: Sound field separation at the point O: The amplitudes of the target sound $p^{\mathrm{d}}(t)$, the total sound $p^{\mathrm{t}}(t)$, the sound separation errors using the proposed method $p^{\mathrm{e}}_{\mathrm{pro}}(t)$ and the Spatial Fourier Transform based method $p^{\mathrm{e}}_{\mathrm{SFT}}(t)$ at the point O over 30 ms.

sure the field using a hemispherical array around the target sources, and duplicate the measurements with respect to the surface to use the full spherical harmonic expansion [96, 106]. In this section, we simulate such a case.

We have a room of size 4 m $\times$ 5 m $\times$ 3m as shown in Fig. 4.9. The reflection coefficients of all the walls, floor, and ceiling are 0.99. Based on the point O, a point on the floor, we set up a Cartesian coordinate system and a spherical coordinate system. One corner of the room is located at $D = (-1.8, -1.5, 0)$ m with respect to O. There is a target point source at the point O inside the hemisphere $\mathbb{S}_2$ of radius $R = 0.5$. The length of the spatial filters $g^u_0(t)$, $g^u_1(t)$, $g^u_2(t)$ is $T_n = \lceil 2Rf_{\mathrm{s}}/c \rceil + 1 = 141$ taps. There is another interfering point source at $(0.7, 0.8, 0.7)$ m outside of the sphere $\mathbb{S}_2$. These two sources produce unit-variance Gaussian white noises, which have been filtered by a 64-tap Bandpass Butterworth filter with frequency range [100, 300] Hz.

We measure the sound pressure and the radial particle velocity on the hemisphere $\mathbb{S}_2$ at four sensors, whose locations are determined according to the first-order Gauss sampling scheme [75]. We simulate the impulse responses, including the radial particle velocity response, between each source and sensor using the image source method [104]. We take the first 2744 image sources into consideration, and the impulse responses are truncated to 8192 taps long under the sampling fre-

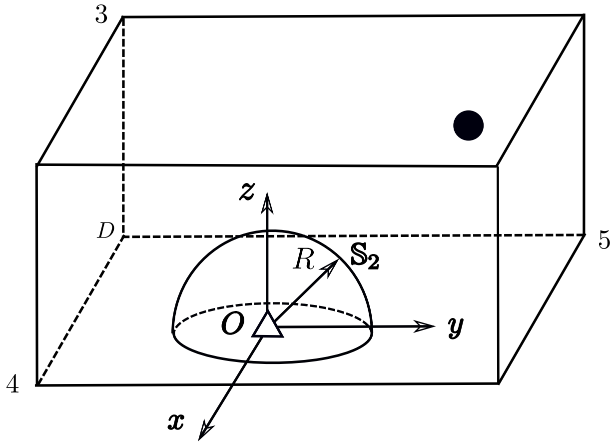


Figure 4.9: Sound field separation in a room: A target source △ is placed at the point O inside a hemisphere $\mathbb{S}_2$ of radius $R = 0.5$ m, and another point source • is placed at $(0.7, 0.8, 0.7)$ m with respective to the point O.

quency of 48 kHz. The pressure and the radial partial velocity on the hemisphere $\mathbb{S}_2$ are duplicated with respect to the floor to use the full spherical harmonic expansion [96, 106]. We estimate the outgoing field coefficients up to zeroth order, and reconstruct the outgoing field on the sphere over a period of 10 ms according to (4.6). We depict the amplitudes of the outgoing field $p^{\mathrm{o}}(t, R, \Theta_1)$, the total field $p(t, R, \Theta_1)$, the separated outgoing field $\widehat{p}^{\mathrm{o}}(t, R, \Theta_1)$, and the outgoing field separation error $p^{\mathrm{e}}(t, R, \Theta_1) = p^{\mathrm{o}}(t, R, \Theta_1) - \widehat{p}^{\mathrm{o}}(t, R, \Theta_1)$, at the first sensor in Fig. 4.10. The outgoing field is obtained through the convolution between the target source outputs and the half-space Green function [96, 106]. The total field is the summation of the outgoing field and the incoming field.

As shown in Fig. 4.10, in this case, the proposed method is still able to accurately estimate the outgoing field. The normalized separation error at the first sensor is

$$\begin{aligned}\xi_1 &= 10\ \log 10 \frac{\sum_{t=0}^{10\ \mathrm{ms}} ||p^{\mathrm{e}}(t, R, \Theta_1)||_2^2}{\sum_{t=0}^{10\ \mathrm{ms}} ||p^{\mathrm{o}}(t, R, \Theta_1)||_2^2} \\ &= -31\ \mathrm{dB}.\end{aligned} \tag{4.25}$$

The proposed method also recovers the outgoing field over the whole hemisphere. The results are similar to Fig. 4.5, and thus are not shown for brevity.

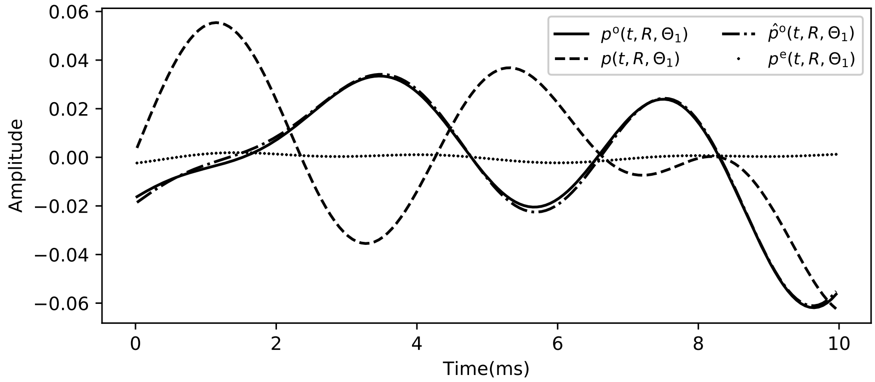


Figure 4.10: Sound field separation at a point: The amplitudes of the outgoing field $p^o(t, R, \Theta_1)$, the total field $p(t, R, \Theta_1)$, the separated outgoing field $\widehat{p}^o(t, R, \Theta_1)$, and the outgoing field separation error $p^e(t, R, \Theta_1)$, at the first sensor over 10 ms.

4.6 Conclusion

In this chapter, we developed a time-domain SFS method that can separate non-stationary sound fields over a sphere. We decompose the sound pressure and the radial particle velocity on the sphere into coefficients, and recover the outgoing field and the incoming field based on the time-domain relationship between the coefficients and the derived spatial filters. The simulations demonstrated that the proposed method can separate non-stationary sound fields in both free field and room environments. Further, the method is able to accurately recover the outgoing field over a long period of time for signals with frequency components below 1000 Hz. A future extension of the proposed method is to take the scattering from the target source surface into consideration [110, 111, 113]. As mentioned in Sec. 4.4, we can use either the vector sensors or microphone pairs to realized the proposed method. In either case, the sensors need to be properly calibrated to compensate for their sensitivity differences [119]. The sensor calibration and the experimental validation of the proposed method will be our future work.

Publications related to this chapter

- **Fei Ma**, Wen Zhang, Thushara D. Abhayapala, "Real time separation of

non-stationary sound fields on spheres," The Journal of the Acoustic Society of America, vol. 146, no. 1, pp. 11-21, Jul. 2019.

- **Fei Ma**, Wen Zhang, and Thushara D. Abhayapala, "Reference signal generation for broadband ANC systems in reverberant rooms," 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 216-220, Apr. 2018.

4.7 Proof of the time-domain sound field separation method

In this section, we derive the expression of the time-domain outgoing field coefficients $\mathsf{p}^{\text{o}}_{u,v}(t,R)$.

Substitutions of [72]

$$j'_u(x) = \frac{u}{x} j_u(x) - j_{u+1}(x),$$

in (4.4) result in

$$\begin{aligned}\mathsf{P}^{\text{o}}_{u,v}(\omega,R) = &-u(i\omega\tau_R)j_u(\omega\tau_R)h_u(\omega\tau_R)\times\mathsf{P}_{u,v}(\omega,R)\\ &+\tau_R(\omega\tau_R)j_{u+1}(\omega\tau_R)h_u(\omega\tau_R)\times(i\omega)\mathsf{P}_{u,v}(\omega,R)\\ &-\rho c\tau_R(i\omega\tau_R)j_u(\omega\tau_R)h_u(\omega\tau_R)\times(i\omega)\mathsf{V}_{u,v}(\omega,R).\end{aligned} \tag{4.26}$$

We define the time-domain outgoing field coefficients $\mathsf{p}^{\text{o}}_{u,v}(t,R)$ as

$$\begin{aligned}\mathsf{p}^{\text{o}}_{u,v}(t,R) &\equiv \mathcal{F}^{-1}\{\mathsf{P}^{\text{o}}_{u,v}(\omega,R)\}\\ &= g^u_0(t)*\lambda_{u,v}(t) + g^u_1(t)*\chi_{u,v}(t) + \rho c\, g^u_2(t)*\eta_{u,v}(t).\end{aligned} \tag{4.27}$$

The terms in the second line of (4.27) are defined based on (4.26) as follows

$$\lambda_{u,v}(t) \equiv \mathcal{F}^{-1}\Big\{\mathrm{P}_{u,v}(\omega,R)\Big\}, \tag{4.28}$$

$$\chi_{u,v}(t) \equiv \mathcal{F}^{-1}\Big\{(i\omega)\mathsf{P}_{u,v}(\omega,R)\Big\}, \tag{4.29}$$

$$\eta_{u,v}(t) \equiv \mathcal{F}^{-1}\Big\{(i\omega)\mathsf{V}_{u,v}(\omega, R)\Big\}, \tag{4.30}$$

$$g_0^u(t) \equiv \mathcal{F}^{-1}\big\{-u(i\omega\tau_R)j_u(\omega\tau_R)h_u(\omega\tau_R)\big\}, \tag{4.31}$$

$$g_1^u(t) \equiv \mathcal{F}^{-1}\big\{\tau_R(\omega\tau_R)j_{u+1}(\omega\tau_R)h_u(\omega\tau_R)\big\}, \tag{4.32}$$

$$g_2^u(t) \equiv \mathcal{F}^{-1}\big\{-\tau_R(i\omega\tau_R)j_u(\omega\tau_R)h_u(\omega\tau_R)\big\}. \tag{4.33}$$

We derive expressions for these six terms in the following.

First, based on properties of the Fourier transform and the spherical harmonic transform [36,72], we obtain $\lambda_{u,v}(t)$, $\chi_{u,v}(t)$, and $\eta_{u,v}(t)$ by decomposing the pressure $p(t, R, \Theta)$ and the radial particle velocity $\nu(t, R, \Theta)$ as follows

$$\begin{aligned}\lambda_{u,v}(t) &= \mathcal{F}^{-1}\{\mathsf{P}_{u,v}(\omega, R)\} \\ &= \mathcal{F}^{-1}\Big\{\int_\Theta P(\omega, R, \Theta)Y_{u,v}(\Theta)d\Theta\Big\} \\ &= \int_\Theta p(t, R, \Theta)Y_{u,v}(\Theta)d\Theta,\end{aligned} \tag{4.34}$$

$$\begin{aligned}\chi_{u,v}(t) &= \mathcal{F}^{-1}\Big\{(i\omega)\mathsf{P}_{u,v}(\omega, R)\Big\} \\ &= \frac{d}{dt}\Big\{\mathcal{F}^{-1}[\mathsf{P}_{u,v}(\omega, R)]\Big\} \\ &= \frac{d}{dt}\Big\{\mathcal{F}^{-1}\Big[\int_\Theta P(\omega, R, \Theta)Y_{u,v}(\Theta)d\Theta\Big]\Big\} \\ &= \frac{d}{dt}\Big\{\int_\Theta p(t, R, \Theta)Y_{u,v}(\Theta)d\Theta\Big\} \\ &= \int_\Theta \frac{dp(t, R, \Theta)}{dt}Y_{u,v}(\Theta)d\Theta,\end{aligned} \tag{4.35}$$

$$\begin{aligned}\eta_{u,v}(t) &= \mathcal{F}^{-1}\Big\{(i\omega)\mathsf{V}_{u,v}(\omega, R)\Big\} \\ &= \frac{d}{dt}\Big\{\mathcal{F}^{-1}[B_{u,v}(\omega, R)]\Big\} \\ &= \frac{d}{dt}\Big\{\mathcal{F}^{-1}\Big[\int_\Theta V(\omega, R, \Theta)Y_{u,v}(\Theta)d\Theta\Big]\Big\} \\ &= \frac{d}{dt}\Big\{\int_\Theta \nu(t, R, \Theta)Y_{u,v}(\Theta)d\Theta\Big\}\end{aligned}$$

$$= \int_{\Theta} \frac{d\nu(t, R, \Theta)}{dt} Y_{u,v}(\Theta) d\Theta. \tag{4.36}$$

Then based on the following properties of the spherical Bessel and Hankel function [36, 72]

$$\begin{aligned}
h_u(x) &= e^{-ix} i^{u+2} \sum_{v=0}^{u} \frac{\varphi_v(u)}{(ix)^{v+1}}, \\
\varphi_v(u) &= \frac{(u+v)!}{2^v v!(u-v)!}, \\
j_u(x) &= \frac{1}{2}[h_u(x)^* + h_u(x)], \\
h'_u(x) &= h_{u-1}(x) - \frac{u+1}{x} h_u(x), \\
h_{-u-1}(x) &= i(-1)^{u+1} h_u(x), \quad u \geq 0,
\end{aligned}$$

we expand (4.31), (4.32), and (4.33) as

$$\begin{aligned}
g_0^u(t) &= \mathcal{F}^{-1}\Big[- u(i\tau_R\omega) j_u(\omega\tau_R) h_u(\omega\tau_R)\Big] \\
&= \mathcal{F}^{-1}\Big[\frac{u}{2} \sum_{v=0}^{u} \sum_{\varsigma=0}^{u} \varphi_v(u)\varphi_\varsigma(u) \frac{(-1)^v + (-1)^{u+1} e^{-2i\tau_R\omega}}{(i\tau_R\omega)^{v+\varsigma+1}}\Big],
\end{aligned} \tag{4.37}$$

$$\begin{aligned}
g_1^u(t) &= \mathcal{F}^{-1}\Big[\tau_R(\omega\tau_R) j_{u+1}(\omega\tau_R) h_u(\omega\tau_R)\Big] \\
&= \mathcal{F}^{-1}\Big[\frac{\tau_R}{2} \sum_{v=0}^{u+1} \sum_{\varsigma=0}^{u} \varphi_v(u+1)\varphi_\varsigma(u) \frac{(-1)^v + (-1)^{u} e^{-2i\tau_R\omega}}{(i\tau_R\omega)^{v+\varsigma+1}}\Big],
\end{aligned} \tag{4.38}$$

$$\begin{aligned}
g_2^u(t) &= \mathcal{F}^{-1}\Big[- \tau_R(i\omega\tau_R) j_u(\omega\tau_R) h_u(\omega\tau_R)\Big] \\
&= \mathcal{F}^{-1}\Big[\frac{\tau_R}{2} \sum_{v=0}^{u} \sum_{\varsigma=0}^{u} \varphi_v(u)\varphi_\varsigma(u) \frac{(-1)^v + (-1)^{u+1} e^{-i2\omega\tau_R}}{(i\tau_R\omega)^{v+\varsigma+1}}\Big],
\end{aligned} \tag{4.39}$$

respectively. Based on the Fourier transform properties, (4.37), (4.38), and (4.39) can be further developed as (4.10), (4.11), and (4.12), respectively.

The expressions of the incoming field coefficients $\mathsf{p}^{\mathsf{i}}_{u,v}(t, R)$, $g_3^u(t)$, and $g_4^u(t)$ can be derived similarly, and are not shown for brevity.

Chapter 5

Active control of random outgoing noise fields in a room

5.1 Introduction

In this chapter, we develop a spatial ANC system that can cancel random noise fields in a room by exploiting the characteristics of the noise fields in rooms. First, to generate the reference signal, we surround the primary source with an array of reference sensors. We separate the outgoing field produced by the primary source from the incoming field due to the secondary source feedback, room reflections, and the background noise. We use the separated outgoing field (coefficients), which is highly coherent with the primary noise (observed at the error sensor) as the reference signal for a random ANC system. Second, we remove the need for real-time secondary path estimation. We surround the primary source and the secondary sources with an array of error sensors. We use the secondary outgoing field (coefficients) to destructively interfere with the primary outgoing field (coefficients) on the error sensor array. As the secondary outgoing field is mainly determined by the secondary sources and is less affected by the room, we remove the need for identifying the room environment characteristics, or real-time secondary path estimation. Third, to meet the causal control constraint, we present a time-wave domain adaptive algorithm. This time-domain algorithm together with the time-domain SFS method derived in Chap. 4 minimize the latency of the proposed

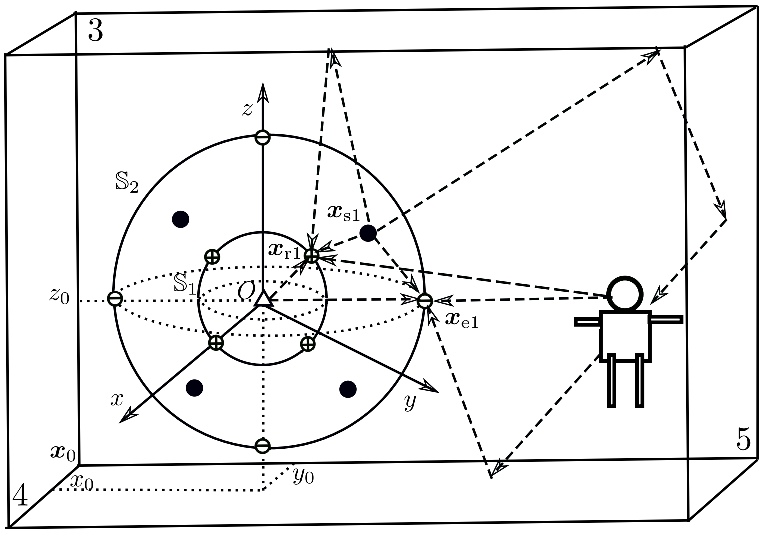


Figure 5.1: The layout of a random ANC system in a room: We mark the primary source by $\triangle$, the secondary sources by $\bullet$, the reference sensors by $\oplus$, and the error sensors by $\ominus$. A reference sensor, a secondary source, and an error sensor are located at $\boldsymbol{x}_{r1}$, $\boldsymbol{x}_{s1}$, and $\boldsymbol{x}_{e1}$, respectively. The dashed lines denote acoustic paths between sources and sensors. The radii of the sphere $\mathbb{S}_1$ and the sphere $\mathbb{S}_2$ are R_1 and R_2, respectively. We aim to cancel the outgoing random noise field produced by the primary source, creating a global quiet zone outside of the sphere $\mathbb{S}_2$ and inside of the room.

system. The effectiveness of the proposed system for canceling random outgoing noise fields globally in a room is confirmed by simulations and compared with the classical point-based ANC system.

The rest of this chapter is organized as follows. The problem of interest is introduced in Sec. 5.2. We present a random ANC system in Sec. 5.3. The effectiveness of the proposed system is shown by simulations in Sec. 5.4. Section 5.5 concludes this paper.

5.2 Problem formulation

Consider the random ANC system in a room environment as shown in Fig. 5.1. We mark the primary source by $\triangle$, the secondary sources by $\bullet$, the reference sensors by $\oplus$, and the error sensors by $\ominus$. We use $\boldsymbol{x} = (x, y, z)$ and $\boldsymbol{x} = (r, \Theta)$ $(\Theta = (\theta, \phi))$

to denote the Cartesian coordinates and the spherical coordinates of a point with respect to the origin O, respectively. A reference sensor, a secondary source, and an error sensor are located at $\boldsymbol{x}_{\mathrm{r}1}$, $\boldsymbol{x}_{\mathrm{s}1}$, and $\boldsymbol{x}_{\mathrm{e}1}$, respectively. The radii of the sphere $\mathbb{S}_1$ and the sphere $\mathbb{S}_2$ are R_1 and R_2, respectively. The dashed lines denote acoustic paths between sources and sensors. Figure 5.1 shows that the measurements of the sensors are the accumulated contributions from the primary source, the secondary sources, the disturbing source (human), and the room reflections. This fact makes it difficult to obtain a clean reference signal of the primary noise, and makes the random ANC system susceptible to the secondary source feedback problem and the room environmental changes [1, 26].

In this chapter, we aim to cancel the outgoing random primary noise field using the secondary sources, creating a global quiet zone outside the sphere $\mathbb{S}_2$ and inside the room.

The sound field $p(t, R, \Theta)$ on a sphere of radius R can be decomposed as [36]

$$\begin{aligned} p(t, R, \Theta) &= \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{p}_{u,v}(t, R) Y_{u,v}(\Theta) \\ &= \underbrace{\sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{p}^{\mathrm{o}}_{u,v}(t, R) Y_{u,v}(\Theta)}_{p^{\mathrm{o}}(t,R,\Theta)} + \underbrace{\sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{p}^{\mathrm{i}}_{u,v}(t, R) Y_{u,v}(\Theta)}_{p^{\mathrm{i}}(t,R,\Theta)}, \end{aligned} \tag{5.1}$$

where the pressure field coefficients $\mathsf{p}_{u,v}(t)$ are obtained by [75]

$$\mathsf{p}_{u,v}(t, R) = \int_0^{2\pi} \int_0^{\pi} p(t, R, \Theta) Y_{u,v}(\Theta) \sin\theta d\theta d\phi, \tag{5.2}$$

$p^{\mathrm{o}}(t, R, \Theta)$ and $p^{\mathrm{i}}(t, R, \Theta)$ are the outgoing field and the incoming field on the sphere, respectively, $\mathsf{p}^{\mathrm{o}}_{u,v}(t, R)$ and $\mathsf{p}^{\mathrm{i}}_{u,v}(t, R)$ are the outgoing field coefficients and the incoming field coefficients, respectively.

Though the measurements of the sensors are contributions from various sources in the room environment, the outgoing fields on the sphere $\mathbb{S}_1$ and the sphere $\mathbb{S}_2$ are only due to the primary source and the secondary sources. In this paper, we propose to cancel the outgoing random primary noise field by exploiting this characteristic. We separate the outgoing fields on the sphere $\mathbb{S}_1$ and the sphere $\mathbb{S}_2$

out of the reference sensor array measurements and the error sensor array measurements, respectively. We use the outgoing fields (coefficients) on the sphere $\mathbb{S}_1$ and the sphere $\mathbb{S}_2$ as the reference signals and the error signals for a random ANC system, respectively. The outgoing field (coefficients) on the sphere $\mathbb{S}_1$ is only generated by the primary source, providing a reference signal of the primary noise free from the secondary source feedback, the disturbing noise, and the room reflected sound. The outgoing field (coefficients) on the sphere $\mathbb{S}_2$ is un-affected by the room environment, and thus there is no need for identifying its characteristics in real-time. We present a time-wave domain adaptive algorithm in the next section. This time-domain algorithm and the time-domain SFS method derived in Chap. 4 do not introduce any frame delay. These two algorithms together with a low latency hardware platform will enable the proposed system to meet the causal control constraint.

5.3 The proposed random ANC system

In this section, we propose an ANC system that can cancel outgoing random noise fields in rooms. The signal flow diagram of the proposed system is depicted in Fig. 5.2. We conduct the SFS process, and use the outgoing field coefficients on the sphere $\mathbb{S}_1$ and the sphere $\mathbb{S}_2$ as the reference signals and the error signals for a time-wave domain adaptive algorithm, respectively. The derivations of the adaptive algorithm are in the rest of this section. In what follows, $\mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n, R_2)$ and $\mathsf{p}^{\mathrm{o}}_{u_1,v_1}(n, R_1)$ correspond to the outgoing field coefficients $\mathsf{p}^{\mathrm{o}}_{u,v}(t, R)$ of (5.2) estimated on two spheres of radii R_2 and R_1, respectively.

First, we define a cost function

$$J(n) = \sum_{u_2=0}^{N_2} \sum_{v_2=-u_2}^{u_2} ||\mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n, R_2)||_2^2, \tag{5.3}$$

where $\mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n, R_2)$ is order u_2 and degree v_2 outgoing field coefficient on the sphere $\mathbb{S}_2$. Denote the pressure and the radial particle velocity at $\{R_2, \Theta_q\}_{q=1}^{Q_2}$ on the sphere $\mathbb{S}_2$ as $\{p(n, R_2, \Theta_q), \nu(n, R_2, \Theta_q)\}_{q=1}^{Q_2}$. We use the SFS method to obtain the outgoing field coefficients $\{\mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n, R_2)\}$, where $u \in [0, N_2], v_2 \in [-u_2, u_2]$, on the

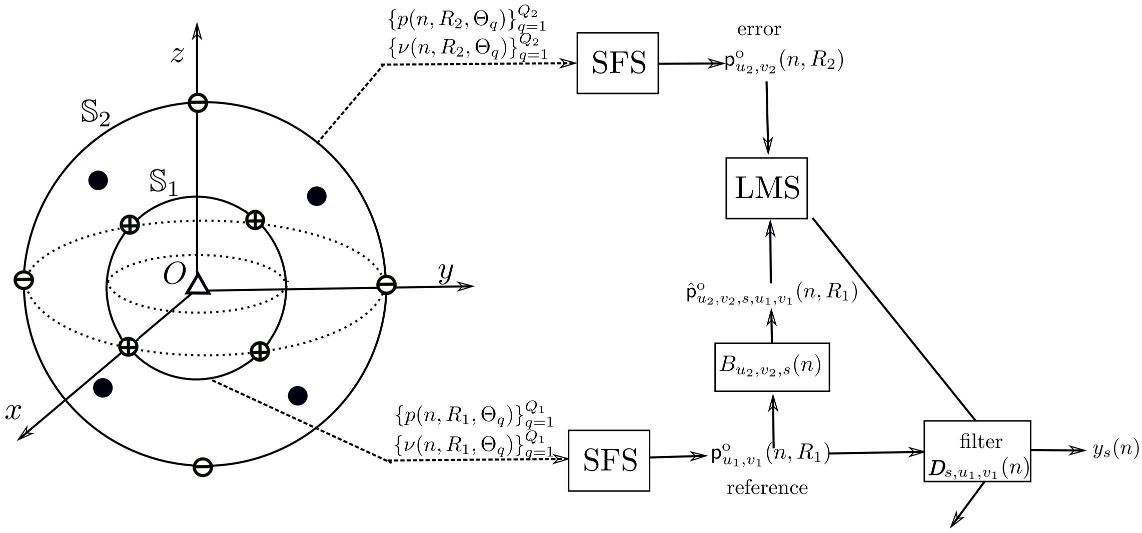


Figure 5.2: The signal flow diagram of the proposed random ANC system: $\{p(n,R_1,\Theta_q),\nu(n,R_1,\Theta_q)\}_{q=1}^{Q_1}$ and $\{p(n,R_2,\Theta_q),\nu(n,R_2,\Theta_q)\}_{q=1}^{Q_2}$ are the pressure and the radial particle velocity on the sphere $\mathbb{S}_1$ and $\mathbb{S}_2$, respectively. The pressure and the radial particle velocity are transformed into the outgoing field coefficients, $\mathsf{p}^{\mathrm{o}}_{u_1,v_1}(n,R_1)$ and $\mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n,R_2)$, which are the reference signals and the error signals for the time-wave domain adaptive algorithm, respectively. The adaptive algorithm updates the adaptive filter based on the error signals $\mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n,R_2)$, and the pre-filtered reference signals, $\widehat{\mathsf{p}}^{\mathrm{o}}_{u_2,v_2,s,u_1,v_1}(n,R_1)$. (The u_1,v_1-s-u_2,v_2 channel.)

sphere $\mathbb{S}_2$. Similarly, we obtain the outgoing field coefficients $\{\mathsf{p}^{\mathrm{o}}_{u_1,v_1}(n,R_1)\}$, where $u\in[0,N_1], v_1\in[-u_1,u_1]$, on the sphere $\mathbb{S}_1$. The truncation orders of the outgoing fields on the sphere $\mathbb{S}_2$ and the sphere $\mathbb{S}_1$ are N_2 and N_1, respectively [73].

Next, we develop the relationship between the outgoing field coefficients $\mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n,R_2)$ on the sphere $\mathbb{S}_2$ and the secondary source driving signals $y_s(n)$. The outgoing field $p^{\mathrm{o}}(n,R_2,\Theta)$ on the sphere $\mathbb{S}_2$ is the summation of the primary outgoing filed and the secondary outgoing field

$$p^{\mathrm{o}}(n,R_2,\Theta)=p^{\mathrm{p,o}}(n,R_2,\Theta)+p^{\mathrm{s,o}}(n,R_2,\Theta). \tag{5.4}$$

Based on (5.1), we express (5.4) as

$$\mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n,R_2)=\mathsf{p}^{\mathrm{p,o}}_{u_2,v_2}(n,R_2)+\mathsf{p}^{\mathrm{s,o}}_{u_2,v_2}(n,R_2)$$

$$= \mathsf{p}^{\mathrm{p,o}}_{u_2,v_2}(n,R_2) + \sum_{s=1}^{Q_\mathrm{s}} \underbrace{\sum_{u_1=0}^{N_1} \sum_{v_1=-u_1}^{u_1} \mathsf{p}^{\mathrm{o}}_{u_1,v_1}(n,R_1) * \boldsymbol{D}_{\mathrm{s},u_1,v_1}(n)}_{y_\mathrm{s}(n)} * \boldsymbol{B}_{u_2,v_2,\mathrm{s}}(n), \tag{5.5}$$

where $\mathsf{p}^{\mathrm{p,o}}_{u_2,v_2}(n)$ and $\mathsf{p}^{\mathrm{s,o}}_{u_2,v_2}(n)$ are the primary outgoing filed coefficients and the secondary outgoing field coefficients on the sphere $\mathbb{S}_2$, respectively, $u_2 \in [0, N_2]$, $v_2 \in [-u_2, u_2]$, $\{y_s(n)\}_{s=1}^{Q_\mathrm{s}}$ are the secondary source driving signals,

$$\boldsymbol{D}_{s,u_1,v_1}(n) = [d_{s,u_1,v_1}(n,0), d_{s,u_1,v_1}(n,1), ..., d_{s,u_1,v_1}(n,l_\mathrm{a}-1)], \\ s \in [1,Q_\mathrm{s}], u_1 \in [0,N_1], v_1 \in [-u_1,u_1], \tag{5.6}$$

are adaptive filters,

$$\boldsymbol{B}_{u_2,v_2,s}(n) = [b_{u_2,v_2,s}(n,0), b_{u_2,v_2,s}(n,1), ..., b_{u_2,v_2,s}(n,l_\mathrm{se}-1)], \\ s \in [1,Q_\mathrm{s}], u_2 \in [0,N_2], v_2 \in [-u_2,u_2], \tag{5.7}$$

are the outgoing path impulse responses between the secondary source outputs $y_\mathrm{s}(n)$ and the secondary outgoing field coefficients $\mathsf{p}^{\mathrm{s,o}}_{u_2,v_2}(n)$ on the sphere $\mathbb{S}_2$ [122]. The lengths of the adaptive filters and the outgoing paths are l_a and l_se, respectively.

Based on (5.6) and (5.7), the outgoing field coefficients $\mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n,R_2)$ in (5.5) can be expanded and expressed in a different form as

$$\begin{aligned} \mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n,R_2) =& \mathsf{p}^{\mathrm{p,o}}_{u_2,v_2}(n,R_2) \\ &+ \sum_{s=1}^{Q_\mathrm{s}} \sum_{u_1=0}^{N_1} \sum_{v_1=-u_1}^{u_1} \sum_{l=0}^{l_\mathrm{se}-1} \sum_{j=0}^{l_\mathrm{a}-1} \mathsf{p}^{\mathrm{o}}_{u_1,v_1}(n-l-j,R_1) d_{\mathrm{s},u_1,v_1}(n,j) b_{u_2,v_2,\mathrm{s}}(n,l) \\ =& \mathsf{p}^{\mathrm{p,o}}_{u_2,v_2}(n,R_2) + \sum_{s=1}^{Q_\mathrm{s}} \sum_{u_1=0}^{N_1} \sum_{v_1=-u_1}^{u_1} \sum_{j=0}^{l_\mathrm{a}-1} d_{\mathrm{s},u_1,v_1}(n,j) \widehat{\mathsf{p}}^{\mathrm{o}}_{u_2,v_2,s,u_1,v_1}(n-j,R_1), \end{aligned} \tag{5.8}$$

where

$$\widehat{\mathsf{p}}^{\mathrm{o}}_{u_2,v_2,s,u_1,v_1}(n-j,R_1) = \sum_{l=0}^{l_{\mathrm{se}}-1} b_{u_2,v_2,s}(n,i)\mathsf{p}^{\mathrm{o}}_{u_1,v_1}(n-l-j,R_1), \tag{5.9}$$

are the reference signals filtered by the outgoing path impulse responses [1].

To minimize the outgoing field (coefficients) on the sphere $\mathbb{S}_2$, we update the adaptive filter coefficients according to the gradient of (5.3) [11]

$$\begin{aligned} d_{s,u_1,v_1}(n+1,j) &= d_{s,u_1,v_1}(n,j) \\ &\quad -\mu \sum_{u_2=0}^{N_2} \sum_{v_2=-u_2}^{u_2} \mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n,R_2)\widehat{\mathsf{p}}^{\mathrm{o}}_{u_2,v_2,s,u_1,v_1}(n-j,R_1), \end{aligned} \tag{5.10}$$

where μ is the step size parameter [11].

By summing the convolutions between the reference signals $\{\mathsf{p}^{\mathrm{o}}_{u_1,v_1}(n,R_1)\}$ with the adaptive filters $\{d_{s,u_1,v_1}(n,0), d_{s,u_1,v_1}(n,1), ..., d_{s,u_1,v_1}(n,l_a-1)\}$, we obtain the secondary source driving signals $y_s(n)$.

As shown by Fig 5.2, the secondary source driving signal computation for different channels is highly parallel. We can implement the algorithm using a field gate programming array system, which supports parallel computation [30].

The proposed algorithm is an advanced version of the multiple-error least mean square (Me-LMS) algorithm [11]. The differences between the proposed algorithm and the Me-LMS algorithm are as follows:

- The Me-LMS algorithm uses the room impulse responses between secondary sources and error sensors as secondary paths. The proposed algorithm uses the direct (outgoing) parts of the room impulse responses only. The room impulse responses are susceptible to the movement of objects, and thus the Me-LMS algorithm needs an extra real-time secondary path identification algorithm [11–14]. The proposed algorithm, on the other hand, does not need that extra algorithm, because the direct parts of the room impulse responses are relatively stationary and can be estimated offline.

  Note that, we assume the secondary sources are point sources, but real world secondary sources can be complex. However, given the directivity pattern of the secondary sources, we can model them as an equivalent point source set, and then estimate the direct paths. The model of the secondary sources by point sources, and the estimation of the corresponding direct paths is beyond the scope of this work.

- The room impulse responses are at least several thousand taps long in a room environment [15]. For the Me-LMS algorithm, the computational burden of filtering the reference signals by the room impulse responses are demanding [11]. On the other hand, the direct paths of the room impulse responses are relatively short. This makes the proposed algorithm more computationally friendly, especially on embedded processors.

5.4 Simulation examples

The simulation environment is a room of size of 4 m $\times$ 5 m $\times$ 3 m, as shown in Fig. 5.1. Based on the point O in the room, we set up a Cartesian coordinate system and a spherical coordinate system. One corner of the room is located at $\boldsymbol{x}_0 = (-1.7, -1.8, -1.5)$ m with respect to the point O. The reflection coefficients of the walls are $[0.94, 0.95, 0.96, 0.94, 0.95, 0.96]$. The sampling frequency is $f_\mathrm{s} = 48$ kHz, the speed of sound is $c = 343$ m/s, and the density of air is $\rho = 1.225$ kg/m^3. We use the sampling frequency of 48 kHz because a high sampling frequency helps to reduce the Analog-to-Digital and Digital-to-Analog converter latency. The primary sources are four point sources at $(-0.01, 0.01, -0.01)$ m, $(0.01, 0.01, 0.01)$ m, $(-0.01, -0.01, 0.01)$ m, and $(0.01, -0.01, -0.01)$ m, respectively. The secondary sources are four point sources at $(-0.11, 0.11, -0.11)$ m, $(0.12, 0.12, 0.12)$ m, $(-0.11, -0.11, 0.11)$ m, and $(0.12, -0.12, -0.12)$ m, respectively. Eight reference sensors and eight error sensors are placed on the sphere $\mathbb{S}_1$ of radius $R_1 = 0.1$ m and the sphere $\mathbb{S}_2$ of radius $R_2 = 0.5$ m according to the first order Gauss sampling scheme, respectively [75]. The truncation orders of the outgoing sound fields on the sphere $\mathbb{S}_1$ and the sphere $\mathbb{S}_2$ are $N_1 = 1$ and $N_2 = 1$, respectively [73]. Since the primary sources and the secondary sources are placed close to the origin, the outgoing fields can be characterized using the low order spherical modes. For real sources, which can not be placed close to the origin, we have to enlarge the two spheres. We need high order modes to represent the sound fields on the enlarged spheres, and more sensors and sources to implement the proposed system. Specifically, we need at least $(N_1+1)^2$ reference sensors, $(N_2+1)^2$ secondary sources, and $(N_2+1)^2$ error sensors. To estimate the outgoing fields (coefficients) on the sphere $\mathbb{S}_1$ and the sphere $\mathbb{S}_2$, (4.10), (4.11), (4.12) are implemented as finite impulse response filters of length $\lceil 2 * R_1/c * f_\mathrm{s} \rceil + 1 = 29$ taps and $\lceil 2 * R_2/c * f_\mathrm{s} \rceil + 1 = 141$ taps, respectively. The room impulse responses (including the radial particle velocity responses) are 3072 taps long and simulated using the image source method [104]. We take the first 1000 image sources into consideration. We add background Gaussian white noises to the reference sensor measurements and the error sensor measurements, and the signal-to-noise power ratio is 40 dB. The simulation results are from an average of 100 independent runs. The settings in this paragraph are used in all simulations unless otherwise stated.

5.4.1 Cancellation of outgoing random noise fields over a sphere

In this section, we demonstrate the cancellation of outgoing random noise fields over a sphere and inside the entire room. The outputs of the primary sources are four mutual-independent unit-variance Gaussian white noises filtered by a 64 taps Bandpass Butterworth filter of frequency range [100, 400] Hz. We compare the performance of the proposed system with the multiple-point system which uses the Me-LMS algorithm for noise cancellation [11]. In simulation results shown here and in the next section, the noise energy does not reduce much after T=100 s. Thus we choose the noise energy reduction level at T=100 s as the final noise energy reduction level, and we further choose $T > 90$ s as the steady state, where the noise reduction level is within ±10% of the final noise energy reduction level.

For the proposed system, we conduct the SFS process using both the reference sensor array and the error sensor array. We use the outgoing field coefficients $\{\mathsf{p}^{\mathrm{o}}_{u_1,v_1}(n, R_1)\}$ where $u_1 \in [0, N_1], v_1 \in [-u_1, u_1]$ on the sphere $\mathbb{S}_1$ as the reference signals and the outgoing field coefficients $\{\mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n, R_2)\}$ where $u_2 \in [0, N_2]$, $v_2 \in [-u_2, u_2]$ on the sphere $\mathbb{S}_2$ as the error signals. We use the time-wave domain adaptive algorithm to cancel the total outgoing field (coefficients) on the sphere

$\mathbb{S}_2$. The length of $B_{u_2,v_2,s}(n)$ in (5.7) is 256 taps, and $B_{u_2,v_2,s}(n)$ is estimated offline [1]. The length of the adaptive filters in (5.5) is 500 taps, and the step size parameter is $\mu = 0.00001$.

We design the multiple-point system to cancel the noise pressures at the error sensors on the sphere $\mathbb{S}_2$. We arrange the error sensors on the sphere for two reasons. First, this arrangement allows both the proposed system and the multiple-point system to achieve noise energy reduction in the whole room. Second, because there is no available procedure for determining the optimal arrangement of sensors and sources for the multiple-point system, specifically for the Me-LMS algorithm, this arrangement allows a fair comparison of the two systems. The secondary paths are the room impulse responses between secondary sources the error sensors, and are estimated offline [1]. The length of the adaptive filters are 500 taps, and the step size parameter is $\mu = 0.00001$. We obtain the reference signals in two cases. In the first case, we use the sound pressure measured by the first reference sensor as the reference signal. In the second case, we conduct the SFS process using the reference sensor array, and use the outgoing field coefficients $\{\mathsf{p}^{\mathrm{o}}_{u_1,v_1}(n, R_1)\}$ where $u_1 \in [0, N_1], v_1 \in [-u_1, u_1]$ on the sphere $\mathbb{S}_1$ as the reference signals.

We characterize the performance of the proposed system by the noise energy reduction at the eight error sensors

$$\xi_{\mathrm{sh}}(n) = 10\,\log_{10}\frac{\sum_{q=1}^{8}\mathcal{E}[e_q^2(n)]}{\sum_{q=1}^{8}\mathcal{E}[p_q^2(n)]}, \tag{5.11}$$

where $p_q(n)$ and $e_q(n)$ are the primary noise pressure and the residual noise pressure at q-th error sensor respectively, and $\mathcal{E}[\cdot]$ is the expectation operator. Similarly, we define and denote the noise energy reductions achieved by the multiple-point system for the two cases as $\xi^{\mathrm{r1}}_{\mathrm{me}}(n)$ (without SFS) and $\xi^{\mathrm{ref}}_{\mathrm{me}}(n)$ (with SFS), respectively.

We conduct noise cancellation using these systems, and present the noise energy reductions at the error sensors in Fig. 5.3. Figure 5.3 shows that, using the sound pressure measured by the first reference sensor as the reference signal, the multiple-point system is unable to converge to the best noise energy reduction level. The final noise energy reduction achieved in this case oscillates around -14 dB. This phenomenon exemplifies the secondary source feedback problem of random ANC

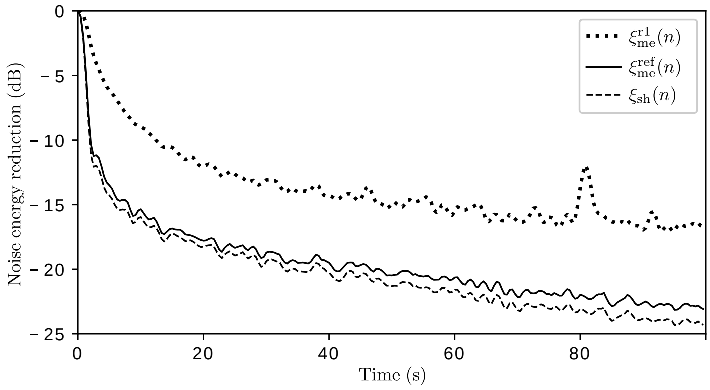


Figure 5.3: Cancellation of outgoing random noise fields over a sphere: $\xi_{\mathrm{me}}^{\mathrm{r1}}(n)$ and $\xi_{\mathrm{me}}^{\mathrm{ref}}(n)$ are the noise energy reductions at the error sensors achieved by the multiple-point system with and without the SFS process, respectively, $\xi_{\mathrm{sh}}(n)$ is the noise energy reduction at the error sensors achieved by the proposed system,

systems [1]. Figure 5.3 demonstrates that using the outgoing field coefficients as the reference signals can make both the multiple-point system and the proposed system converge to the best noise energy reduction level. In the steady state ($t > 90$ s), the multiple-point (with SFS) system and the proposed system achieve noise energy reductions at the error sensors by about -23 dB and -24 dB, respectively.

We further characterize the performance of the proposed system by the noise energy reduction in the room

$$\xi_{\mathrm{sh,room}}(n) = 10 \log_{10} \frac{\sum_{q=1}^{60000} \mathcal{E}[e_q^2(n, \boldsymbol{x}_q)]}{\sum_{q=1}^{60000} \mathcal{E}[p_q^2(n, \boldsymbol{x}_q)]},$$

where $\boldsymbol{x}_q$ are the positions of 60000 uniform sampling points in the room. Note that, the points inside of the sphere $\mathbb{S}_2$ are excluded in the above equation. Similarly, we define and denote the noise energy reduction achieved by the multiple-point system (with SFS) in the room as $\xi_{\mathrm{me,room}}^{\mathrm{ref}}(n)$. As shown in Fig. 5.4, the multiple-point system and the proposed system reduce the noise energy in the room by about -21 dB and -22 dB in the steady state ($t > 90$ s), respectively.

We conduct another simulation and use either the outgoing field coefficients or the outgoing pressures as the reference signals or the error signals for the proposed system. Four cases are considered:

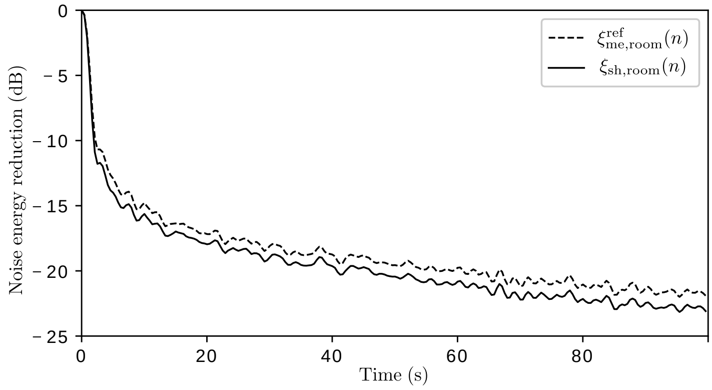


Figure 5.4: Cancellation of outgoing random noise fields in the room: $\xi_{\mathrm{sh,room}}(n)$ and $\xi^{\mathrm{ref}}_{\mathrm{me,room}}(n)$ are the noise energy reductions in the room achieved by the proposed and the multiple-point system (with SFS), respectively.

1. the outgoing field coefficients on the sphere $\mathbb{S}_1$ $\{\mathsf{p}^{\mathrm{o}}_{u_1,v_1}(n, R_1)\}$, where $u_1 \in [0, N_1]$, $v_1 \in [-u_1, u_1]$, as the reference signals and the outgoing field coefficients on the sphere $\mathbb{S}_2$ $\{\mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n, R_2)\}$, where $u_2 \in [0, N_2], v_2 \in [-u_2, u_2]$, as the error signals;

2. the outgoing field coefficients on the sphere $\mathbb{S}_1$ $\{\mathsf{p}^{\mathrm{o}}_{u_1,v_1}(n, R_1)\}$, where $u_1 \in [0, N_1]$, $v_1 \in [-u_1, u_1]$, as the reference signals and the outgoing pressures at the eight error sensors on the sphere $\mathbb{S}_2$ as the error signals;

3. the outgoing pressures at the 1-th, 3-th, 5-th, 7-th reference sensors on the sphere $\mathbb{S}_1$ as the reference signals and the outgoing field coefficients on the sphere $\mathbb{S}_2$ $\{\mathsf{p}^{\mathrm{o}}_{u_2,v_2}(n, R_2)\}$, where $u_2 \in [0, N_2], v_2 \in [-u_2, u_2]$, as the error signals;

4. the outgoing pressures at the 1-th, 3-th, 5-th, 7-th reference sensors on the sphere $\mathbb{S}_1$ as the reference signals and the outgoing pressures at the eight error sensors on the sphere $\mathbb{S}_2$ as the error signals.

The noise energy reductions at the eight error sensors for Case 1, Case 2, Case 3, and Case 4 are denoted as $\xi^{\mathrm{cc}}_{\mathrm{sh}}(n)$, $\xi^{\mathrm{cp}}_{\mathrm{sh}}(n)$, $\xi^{\mathrm{pc}}_{\mathrm{sh}}(n)$, and $\xi^{\mathrm{pp}}_{\mathrm{sh}}(n)$, respectively, which are defined similar to (5.11). For Case 2 and Case 4, we choose the outgoing pressures at the 1-th, 3-th, 5-th, 7-th reference sensors as the reference signals for two reasons. The first is that for the four cases the number of reference signals are the same. The second is that the distances between the reference sensors for Case 2 and Case 4 are maximized, and thus the outgoing pressures at the four sensors are diversified and can characterize the primary source output better.

For the four cases, the basic simulation settings are the same as the proposed system in the former part of this section. For Case 2, Case 3, and Case 4, modifications to the proposed system are made. Specifically, we express the reference signals, the error signals, the adaptive filters, and the outgoing path impulse responses either in the spherical mode domain or in the time-domain based on whether the reference signals and the error signals are expressed using the outgoing field coefficients or the outgoing pressures.

We present the simulation results in Fig. 5.5. As shown in Fig. 5.5, in Case 1 where the outgoing field coefficients are used as both the reference signals and the error signals, the noise energy reduction of the proposed system is the best. In Case 4 where the outgoing pressures are used as both the reference signals and the error signals, the noise energy reduction of the proposed system is the worst. Specifically, the steady-state ($t > 90$ s) noise energy reduction achieved in Case 1 is 8 dB more than in Case 4. The simulation results demonstrate that using the outgoing field coefficients as the reference signals and the error signals enables the proposed system to achieve better noise cancellation performance.

We have one more simulation about the proposed system. We consider two cases. The first case is the same as proposed system in the first simulation of this section. We use the outgoing path impulse response $B_{u_2,v_2,s}(n)$ to multiply with the reference signals and the error signals, and use the results to update the adaptive filter coefficients. The outgoing path impulse response $B_{u_2,v_2,s}(n)$ carries the information of the spatial decay and the temporal delay from a secondary source to the secondary outgoing field. The second case differs from the first case by replacing the outgoing path impulse response $B_{u_2,v_2,s}(n)$ using its temporal delay when we update the adaptive filer coefficients in (5.10). We have the noise energy reductions, which are defined similar to (5.11), of these two cases in Fig. 5.6. Figure 5.6 shows that the using the temporal delay in place of the outgoing path impulse response $B_{u_2,v_2,s}(n)$ results in about 3 dB less noise energy reduction level.

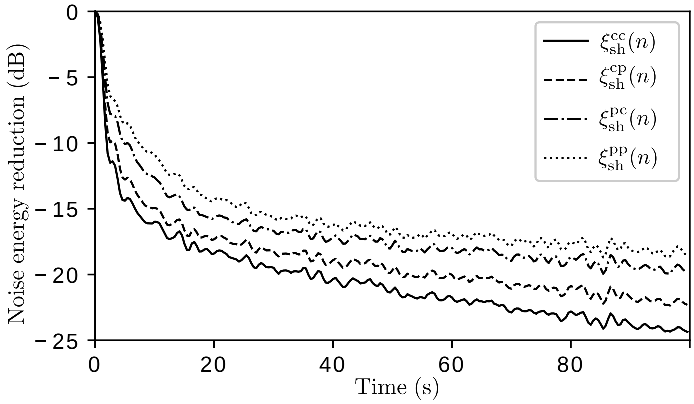


Figure 5.5: Cancellation of outgoing random noise fields over a sphere: $\xi_{\rm sh}^{\rm cc}(n)$ is the noise energy reduction when the outgoing field coefficients are used as both the reference signals and the error signals, $\xi_{\rm sh}^{\rm pp}(n)$ is the noise energy reduction when the outgoing pressures are used as both the reference signals and the error signals, $\xi_{\rm sh}^{\rm cp}(n)$ and $\xi_{\rm sh}^{\rm pc}(n)$ are the noise energy reductions when either the outgoing pressures or the outgoing field coefficients are used as the reference signals or the error signals.

5.4.2 Cancellation of outgoing random noise fields over a hemisphere

In real world implementation of the proposed system, the primary source can be attached to a surface. In case the surface is rigid or highly reflective, we can measure and cancel the outgoing noise field on a hemisphere covering the source [96, 106]. In this section, we simulate the cancellation of outgoing random noise fields in such a scenario.

The simulation environment is shown in Fig. 5.7. The room size is 4 m $\times$ 5 m $\times$ 3 m. On the floor, there is a point O, based on which we set up a Cartesian coordinate system and a spherical coordinate system. One corner of the room is located at $\boldsymbol{x}_0 = (-1.7, -1.8, 0)$ m with respect to the point O. A point source at the point O is the primary source. A point source at $(0, 0, 0.2)$ m is the secondary source. Four reference sensors and four error sensors are placed on the hemisphere $\mathbb{S}_1$ of radius $R_1 = 0.1$ m and the hemisphere $\mathbb{S}_2$ of radius $R_2 = 0.5$ m according to the first order Gauss sampling scheme, respectively [75]. The truncation orders of

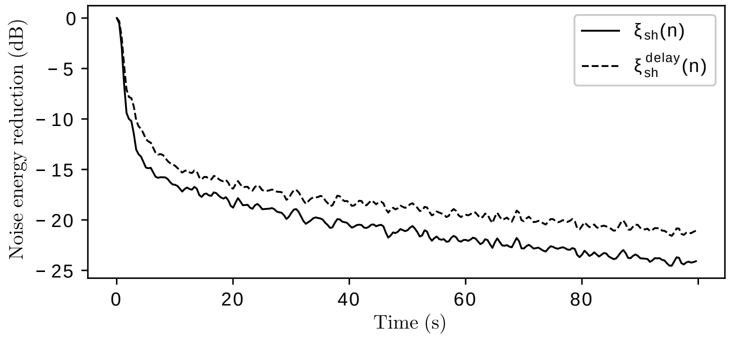


Figure 5.6: Cancellation of outgoing random noise fields on the sphere $\mathbb{S}_2$: $\xi_{\text{sh}}(n)$ and $\xi_{\text{sh}}^{\text{delay}}(n)$ are the noise energy reductions at the error sensors achieved by the proposed system, $\xi_{\text{sh}}^{\text{delay}}(n)$ differs from $\xi_{\text{sh}}(n)$ by that it is achieved when we replace the outgoing path impulse response $B_{u_2,v_2,s}(n)$ by its temporal delay to update the adaptive filter coefficients.

the outgoing fields on the hemisphere $\mathbb{S}_1$ and the hemisphere $\mathbb{S}_2$ are $N_1 = 0$ and $N_2 = 1$, respectively. Same as the above section, to estimate the outgoing fields (coefficients) on the hemisphere $\mathbb{S}_1$ and the hemisphere $\mathbb{S}_2$, (4.10), (4.11), (4.12) are implemented as finite impulse response filters of length 29 taps and 141 taps, respectively. The primary source output is a unit-variance Gaussian white noise filtered by a 64 tap Butterworth Bandpass filter with frequency range [100, 300] Hz. The sound fields measured by the reference sensors and the error sensors on the upper hemispheres are mirrored to the lower hemispheres [96, 106].

We design the proposed system to reduce the outgoing sound field (coefficients) on the hemisphere $\mathbb{S}_2$. We use the outgoing field coefficient $\mathsf{p}^{\text{o}}_{0,0}(n, R_1)$ on the hemisphere $\mathbb{S}_1$ as the reference signal, and the outgoing field coefficients $\{\mathsf{p}^{\text{o}}_{u_2,v_2}(n, R_2)\}$ where $u_2 \in [0, N_2], v_2 \in [-u_2, u_2]$ on the hemisphere $\mathbb{S}_2$ as the error signals.

We compare the performance of the proposed system with the multiple-point system. We generate the reference signal for the multiple-point system in two cases. The first is to use the sound pressure measured by the first reference sensor as the reference signal. The second is to use the SFS process to obtain the outgoing field coefficient $\mathsf{p}^{\text{o}}_{0,0}(n, R_1)$ and use it as the reference signal.

We characterize the performance of the proposed system by the noise energy reduction $\xi_{\text{sh}}(n)$ at the four error sensors, and $\xi_{\text{sh}}(n)$ is defined similar to (5.11). Sim-

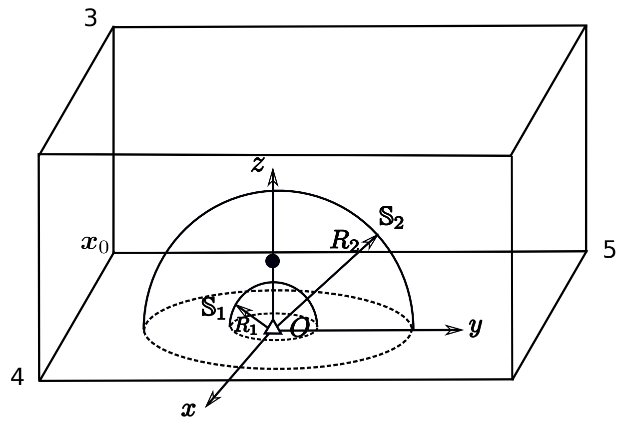


Figure 5.7: Cancellation of outgoing random noise fields over a hemisphere: The simulation environment is a room of size 4 m $\times$ 5 m $\times$ 3 m. There is a point O on the floor, and one corner of the room is located at $\boldsymbol{x}_0 = (-1.7, -1.8, 0)$ m with respect to the point O. There is a primary source $\triangle$, a secondary source $\bullet$, four reference sensors and four error sensors (which are not shown for ease of illustrations) on the hemisphere $\mathbb{S}_1$ and the hemisphere $\mathbb{S}_2$, respectively. The radii of two hemispheres are $R_1 = 0.1$ m and $R_2 = 0.5$ m, respectively.

ilarly, we define and denote the noise energy reductions achieved by the multiple-point system at the four error sensors for the two cases as $\xi_{\text{me}}^{\text{r1}}(n)$ (without SFS) and $\xi_{\text{me}}^{\text{ref}}(n)$ (with SFS), respectively.

The secondary path, the adaptive filter length, and the step size parameter for the proposed system and the multiple-point systems in this section are the same as in Sec. 5.4.1.

We present the noise energy reductions achieved by the systems in Fig. 5.8. As shown in Fig. 5.8, without the SFS process to provide the reference signal, the multiple-point system is unstable to converge to the best noise energy reduction level. Aided with the SFS process, both the proposed system and the multiple-point system can reduce the noise energy at the error sensors by about -25 dB in the steady state ($t > 90$ s). The convergence speed of the proposed system is faster than the multiple-point system (with the SFS).

We further characterize the performance of the systems by the noise energy reductions in the room. The proposed system and the multiple-point system (with SFS) reduce the noise energy in the room by about -24 dB in the steady state. The detailed simulation results are not shown for brevity.

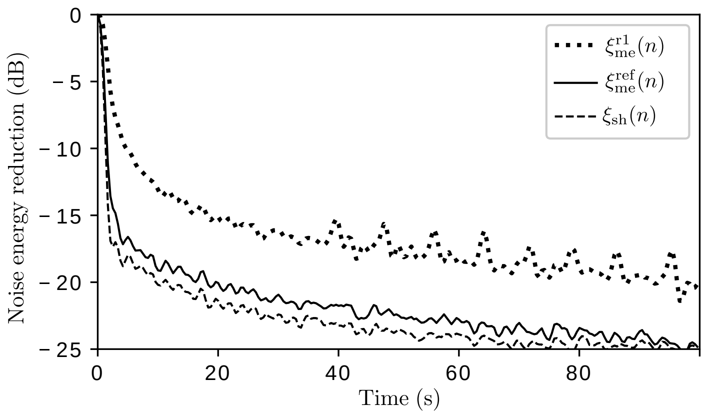


Figure 5.8: Cancellation of outgoing random noise fields over a hemisphere: $\xi_{\text{me}}^{\text{r1}}(n)$ and $\xi_{\text{me}}^{\text{ref}}(n)$ are the noise energy reductions at the error sensors achieved by the multiple-point system with and without the SFS process, respectively, $\xi_{\text{sh}}(n)$ is the noise energy reduction at the error sensors achieved by the proposed system.

We conduct another simulation to show the influence of the system latency on the noise cancellation performance of a random ANC system. We artificially add latency to the reference sensor measurements, the error sensor measurements, and the secondary source outputs to account for the time delay introduced by the Analog-to-Digital and the Digital-to-Analog converters, which contribute the majority of the system latency of an embedded ANC system.

We depict in Fig. 5.9, the noise energy reductions achieved by the proposed system at the four error sensors on the hemisphere $\mathbb{S}_2$. The other simulation settings for the proposed system are the same as former parts of this section. Three cases are considered. For the three cases, the outputs of the primary source are the same unit-variance Gaussian white noise filtered by three 64 taps Butterworth filters. The frequency ranges of the Butterworth filters are [100, 300] Hz, [100, 400] Hz, and [100, 500] Hz, respectively. In Fig. 5.9, $\xi_{\text{sh}}^{300}(n)$, $\xi_{\text{sh}}^{400}(n)$, and $\xi_{\text{sh}}^{500}(n)$ denote the steady-state noise energy reductions for the three cases.

As shown in Fig. 5.9, the noise energy reduction performance of the system downgrades along with the increment of the system latency. The higher the upper frequency limit of the primary source outputs is, the less the noise energy reduction the system can achieve. To reduce a random noise with higher frequency compo-

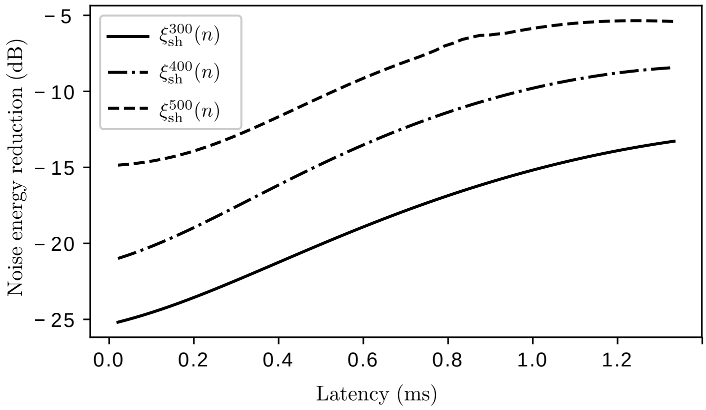


Figure 5.9: Noise cancellation performance of the proposed ANC system as a function of the system latency: $\xi_{\mathrm{sh}}^{300}(n)$, $\xi_{\mathrm{sh}}^{400}(n)$, and $\xi_{\mathrm{sh}}^{500}(n)$ denote the steady-state noise energy reductions where the upper frequency limits of the primary source outputs are 300, 400, and 500 Hz, respectively.

nents (500 Hz) up to -15 dB, we need the ANC system to have an overall system latency less than 0.2 ms, or 9 samples under the sampling frequency of 48 kHz. Such a low system latency is difficult to achieve if the system uses a frequency-domain noise control algorithm, which introduces frame delay. We also conduct simulations to investigate the influence of the system latency on the noise cancellation performance of the multiple-point system. The simulation results are similar to Fig. 5.9, and thus are not shown for brevity. The simulation results demonstrate the necessity of the time-domain spatial algorithms.

5.5 Conclusion

In this chapter, by exploiting the characteristics of spatial noise fields, an ANC system was developed to control outgoing random noise fields in rooms. The proposed system resolves the problems of the reference signal generation and the real-time secondary path estimation by a time-domain SFS algorithm. This algorithm together with a time-wave domain adaptive algorithm allow the proposed system to meet the causal control constraint. Simulations demonstrated the superior noise cancellation performance of the proposed system compared with the multiple-point

ANC system. The influences of the reference signal and the system latency on the noise cancellation performance of random ANC systems were also shown.

Publications related to this chapter

- **Fei Ma**, Wen Zhang, Thushara D. Abhayapala, "Active control of broadband outgoing noise fields in rooms," IEEE Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 529-539, 2020.

Chapter 6

Multiple circular arrays of vector sensors for 3D sound field analysis

6.1 Introduction

The spherical modes, the combinations of the spherical harmonics and the spherical Bessel functions (or the spherical Hankel functions), are a complete and orthogonal eigen-function set of the acoustic wave equation, which fundamentally governs the sound field in a 3D region [36]. The modal decomposition allows the manipulation of a sound field by exploiting its underlying structure [94]. Thus, the theory of modal decomposition and analysis has become a popular tool among researchers, proliferating a number of advanced spatial algorithms in sound field reproduction [73], beamforming [75,94], active noise control [8,78,79], spatial filtering [123], and sound field separation [124].

As spherical harmonics are functions defined over the surface of a sphere, spherical arrays are the natural selections for implementing the above mentioned algorithms [75]. Nonetheless, the spherical arrays are not preferred in many applications, such as smart speakers, wearable devices, and augmented reality devices, which are of limited size and space. Compact non-spherical arrays are preferred over the standard spherical arrays for mounting on these devices and providing them with sound field analysis capacity.

In the past decade, a number of alternative array structures have been proposed.

Abhayapala *et al.* first presented a multiple circular array structure, offering increased sensor position flexibility compared with the standard spherical array [125]. Chen *et al.* designed the first multiple circular array on two-dimensional (2D) plane with 3D sound field analysis capacity [34]. Recently, Berge introduced the acoustically hard 2D array for 3D high-order harmonic signal acquisition [126], and Okamoto proposed to use multiple co-centered omni-directional arrays for 3D sound field recording [127].

Fundamentally, the above mentioned alternative array structures are approximations to the standard spherical arrays, and they have similar problems as the standard spherical arrays. For example, the decomposition coefficients of the sound fields are obtained through the matrix inversion operations, but due to the problem of spherical Bessel function zeros the matrix can be ill-conditioned [36, 73]. The matrix inversion operations are also not desirable to real-time embedded systems. Furthermore, the existing spatial algorithms all process the signal in the frequency-domain, which unavoidably introduces the latency problem and the spectrum leakage problem [128].

In this chapter, we design multiple circular arrays of vector sensors and present two methods of 3D sound field analysis based on the array measurement. A vector sensor is a compact first order sensor [129], which can measure both the sound pressure and the particle velocities, and are finding increased applications in spatial audio [35, 130–132]. We use a single vector sensor to obtain the zeroth and the first order decomposition coefficients directly. We then obtain the high order decomposition coefficients through an iterative process based on the measurements of multiple circular arrays on two planes. We present the method in the frequency-domain first, then derive the corresponding time-domain method. The time-domain method has much lower latency, making it suitable for real-time implementations on embedded systems. The time-domain method further allows the development of advanced time-domain spatial algorithms, which do not suffer from the spectrum leakage problem. In comparison with the existing 3D sound field analysis methods based on the spherical and the non-spherical arrays, the proposed methods are computationally simple, as no matrix inversion operations are involved. The effectiveness of the proposed array for sound field analysis is confirmed by simulation examples.

The rest of the chapter is organized as follows. We present the system description and the problem formulation in Sec. II. We present the design of multiple circular arrays and the frequency-domain sound field analysis method in Sec. III. We present the time-domain sound field analysis method in Sec IV. Simulations are conducted in Sec. V and Sec. VI concludes this chapter.

6.2 System Description and problem formulation

Consider the acoustic system shown in Fig. 6.1. We have two circles $\mathrm{c}_{0,2}$ and $\mathrm{c}_{1,2}$ on two planes h_0 and h_1, respectively. These two circles are also on the sphere $\mathbb{S}_2$ of radius r_2. We denote the Cartesian coordinates and the spherical coordinates of a point with respect to the origin O as (x, y, z) and (r, θ, ϕ), respectively [75].

The frequency-domain and the time-domain sound pressure at the point (r, θ, ϕ) can be expressed as

$$P(\omega, r, \theta, \phi) = \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{P}_{u,v}(\omega, r) Y_{u,v}(\theta, \phi), \tag{6.1}$$

and

$$p(t, r, \theta, \phi) = \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{p}_{u,v}(t, r) Y_{u,v}(\theta, \phi), \tag{6.2}$$

respectively, where $\mathsf{P}_{u,v}(\omega, r)$ and $\mathsf{p}_{u,v}(t, r)$ are the frequency-domain and the time-domain pressure field coefficients, respectively, and $Y_{u,v}(\theta, \phi)$ are spherical harmonics.

A vector sensor can measure the sound pressure and the particle velocities in the x, y, z directions at a point $\boldsymbol{x}_0 = (x_0, y_0, z_0)$ as

$$p(t, \boldsymbol{x}_0),\ \nu_{ox}(t, \boldsymbol{x}_0),\ \nu_{oy}(t, \boldsymbol{x}_0),\ \nu_{oz}(t, \boldsymbol{x}_0). \tag{6.3}$$

In this chapter, we wish to analyze the 3D sound field both in the frequency-domain and in the time-domain based on the measurements of multiple circular arrays located on two planes.

We use the array shown in Fig. 6.1 as an example. We first present the design

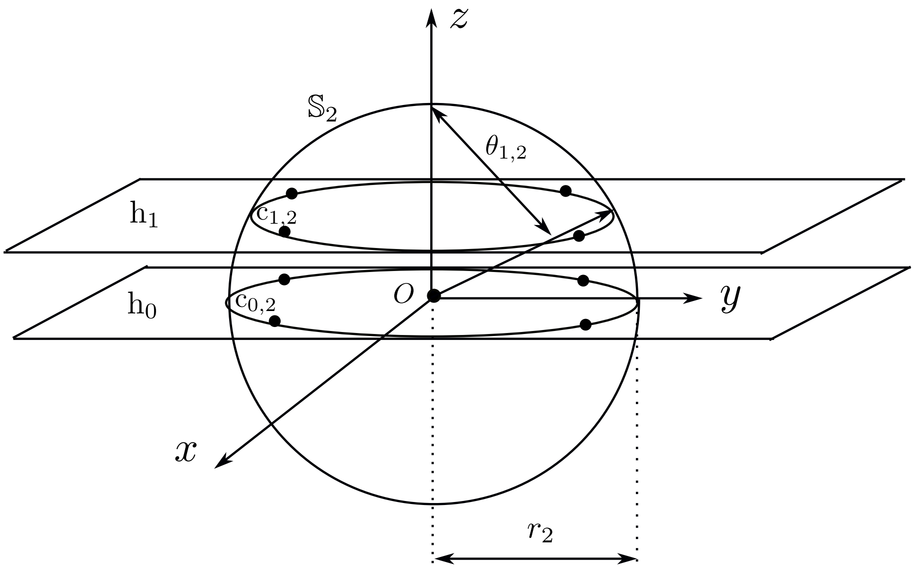


Figure 6.1: System description: We have two circles $\mathrm{c}_{0,2}$ and $\mathrm{c}_{1,2}$ on two planes h_0 and h_1, respectively. These two circles are also on the sphere $\mathbb{S}_2$ of radius r_2. We wish to analyze the 3D sound field within the sphere $\mathbb{S}_2$ based on the measurements of the vector sensors ($\bullet$).

of the array and the sound field analysis method in the frequency-domain, where the relationship among acoustic quantities can be expressed in concise forms. We then present the corresponding time-domain method.

6.3 Frequency-domain derivation

In this section, we first introduce the basic theory of spherical modal analysis, then provide the design of multiple circular arrays and the frequency-domain derivation of a sound field analysis method.

The sound pressure and the radial particle velocity on a sphere of radius r at position (r, θ, ϕ) can be expressed as [36]

$$P(\omega, r, \theta, \phi) = \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{P}_{u,v}(\omega, r) Y_{u,v}(\theta, \phi)$$

$$= \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{B}_{u,v}(\omega) j_u(\omega r/c) Y_{u,v}(\theta,\phi), \tag{6.4}$$

$$\begin{aligned} V(\omega, r, \theta, \phi) &= \frac{i}{\rho\omega} \frac{\partial P(\omega, r, \theta, \phi)}{\partial r} \\ &= \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{V}_{u,v}(\omega, r) Y_{u,v}(\theta, \phi) \\ &= \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \frac{i}{\rho c} \mathsf{B}_{u,v}(\omega) j'_u(\omega r/c) Y_{u,v}(\theta,\phi), \end{aligned} \tag{6.5}$$

where $i = \sqrt{-1}$ is the unit imaginary number, $\omega = 2\pi f$ is the angular frequency (f is the frequency), ρ is the ambient density, c is the speed of sound, $\mathsf{B}_{u,v}(\omega)$ are the spherical Bessel function coefficients, $\mathsf{P}_{u,v}(\omega, r)$ and $\mathsf{V}_{u,v}(\omega, r)$ are the pressure field coefficients and the velocity field coefficients, respectively, $j_u(\cdot)$ and $j'_u(\cdot)$ are the spherical Bessel function of order u and its derivative with respective to the argument, respectively, $Y_{u,v}(\cdot,\cdot)$ are the real value spherical harmonics [71]

$$Y_{u,v}(\theta,\phi) \equiv \mathcal{P}_{u,v}(\theta) E_v(\phi), \tag{6.6}$$

where

$$E_v(\phi) = \begin{cases} \frac{\cos(|v|\phi)}{\sqrt{\pi}}, & v > 0, \\ \frac{1}{\sqrt{2\pi}}, & v = 0, \\ \frac{\sin(|v|\phi)}{\sqrt{\pi}}, & v < 0, \end{cases} \tag{6.7}$$

and

$$\mathcal{P}_{u,v}(\theta) = \sqrt{\frac{(2u+1)(u-|v|)!}{4\pi(u+|v|)!}} \mathfrak{P}_{u,|v|}(\cos(\theta)), \tag{6.8}$$

$\mathfrak{P}_{u,|v|}(\cdot)$ is the associated Legendre function of order u and degree $|v|$, $\mathcal{P}_{u,v}(\cdot)$ is the corresponding normalized associated Legendre function.

As shown by (6.4), the spherical Bessel function coefficients $\mathsf{B}_{u,v}(\omega)$ are radial and angular independent. Given sufficient number of $\mathsf{B}_{u,v}(\omega)$ [73], we can reconstruct and analyze the sound pressure at an arbitrary position. In the rest of this section, we present the design of multiple circular arrays and a method to estimate the spherical Bessel function coefficients $\mathsf{B}_{u,v}(\omega)$ in the frequency-domain based on the array measurement.

6.3.1 The spherical Bessel function coefficients $\mathsf{B}_{0,0}(\omega)$, $\mathsf{B}_{1,-1}(\omega)$, $\mathsf{B}_{1,0}(\omega)$, $\mathsf{B}_{1,1}(\omega)$

Placing a vector sensor at the origin, as shown in Fig. 6.1, we can measure the pressure and the particle velocities at the origin. Based on (6.4), and the fact that [72]

$$j_u(0) = \begin{cases} 1, & u = 0, \\ 0, & \text{else}, \end{cases} \tag{6.9}$$

we can express the sound pressure at the origin as

$$\begin{aligned} P_o(\omega) &= P(\omega, 0, \frac{\pi}{2}, 0) \\ &= \mathsf{B}_{0,0}(\omega)\frac{1}{\sqrt{8\pi^2}}. \end{aligned} \tag{6.10}$$

Based on (6.5), and the fact that [72]

$$j'_u(0) = \begin{cases} \frac{1}{3}, & u = 1, \\ 0, & \text{else}, \end{cases} \tag{6.11}$$

we can express the particle velocities in the x, y, and z directions at the origin as

$$\begin{aligned} V_{ox}(\omega) &= V(\omega, 0, \frac{\pi}{2}, 0) \\ &= \frac{i}{\rho c}\mathsf{B}_{1,1}(\omega)\frac{-1}{\sqrt{24\pi^2}}, \end{aligned} \tag{6.12}$$

$$V_{oy}(\omega) = V(\omega, 0, \frac{\pi}{2}, \frac{\pi}{2})$$

$$= \frac{i}{\rho c} \mathsf{B}_{1,-1}(\omega) \frac{-1}{\sqrt{24\pi^2}}, \tag{6.13}$$

$$\begin{aligned} V_{oz}(\omega) &= V(\omega, 0, 0, 0) \\ &= \frac{i}{\rho c} \mathsf{B}_{1,0}(\omega) \frac{1}{\sqrt{24\pi^2}}. \end{aligned} \tag{6.14}$$

Based on (6.10), (6.12), (6.13), and (6.14), we can obtain the spherical Bessel function coefficients $\mathsf{B}_{0,0}(\omega)$, $\mathsf{B}_{1,-1}(\omega)$, $\mathsf{B}_{1,0}(\omega)$, and $\mathsf{B}_{1,1}(\omega)$ as

$$\mathsf{B}_{0,0}(\omega) = \sqrt{8\pi^2} P_o(\omega), \tag{6.15}$$

$$\mathsf{B}_{1,-1}(\omega) = i\rho c \sqrt{24\pi^2} V_{oy}(\omega), \tag{6.16}$$

$$\mathsf{B}_{1,0}(\omega) = -i\rho c \sqrt{24\pi^2} V_{oz}(\omega), \tag{6.17}$$

$$\mathsf{B}_{1,1}(\omega) = i\rho c \sqrt{24\pi^2} V_{ox}(\omega). \tag{6.18}$$

6.3.2 The spherical Bessel function coefficients $\mathsf{B}_{2,-2}(\omega)$, $\mathsf{B}_{2,0}(\omega)$, $\mathsf{B}_{2,2}(\omega)$

We choose a radius r_2 such that $\lceil 2\pi f_\mathrm{m} r_2/c \rceil = 2$, where $\lceil \cdot \rceil$ is the ceiling operation, f_m is the highest frequency component of the target sound field [73]. That is, the sound field on a sphere of radius r_2 can be accurately described using the spherical harmonics up to the second order. Based on (6.4) and (6.5), the sound pressure and the radial particle velocity on the circle $\mathrm{c}_{0,2}$ on the h_0 plane (as shown in Fig. 6.1) can be expressed as [36, 73]

$$\begin{aligned} P(\omega, r_2, \frac{\pi}{2}, \phi) \approx \sum_{u=0}^{2} \sum_{v=-u}^{u} \mathsf{P}_{u,v}(\omega, r_2) Y_{u,v}(\frac{\pi}{2}, \phi), \\ u, v \neq (1,0), (2,-1), (2,1), \end{aligned} \tag{6.19}$$

$$V(\omega, r_2, \frac{\pi}{2}, \phi) \approx \frac{i}{\rho c} \sum_{u=0}^{2} \sum_{v=-u}^{u} \mathsf{V}_{u,v}(\omega, r_2) Y_{u,v}(\frac{\pi}{2}, \phi),$$
$$u, v \neq (1,0), (2,-1), (2,1), \tag{6.20}$$

where $Y_{1,0}(\frac{\pi}{2}, \phi) = 0$, $Y_{2,-1}(\frac{\pi}{2}, \phi) = 0$, $Y_{2,1}(\frac{\pi}{2}, \phi) = 0$ such that the corresponding sound field components do not present in (6.19) and (6.20).

Define the following intermediate quantities

$$\widetilde{P}(\omega, r_2, \frac{\pi}{2}, \phi) = P(\omega, r_2, \frac{\pi}{2}, \phi) - \sum_{u=0}^{1} \sum_{v=-u}^{u} \mathsf{P}_{u,v}(\omega, r_2) Y_{u,v}(\frac{\pi}{2}, \phi), \tag{6.21}$$

$$\widetilde{V}(\omega, r_2, \frac{\pi}{2}, \phi) = V(\omega, r_2, \frac{\pi}{2}, \phi) - \sum_{u=0}^{1} \sum_{v=-u}^{u} \mathsf{V}_{u,v}(\omega, r_2) Y_{u,v}(\frac{\pi}{2}, \phi), \tag{6.22}$$

where in above two equations $u, v \neq (1,0)$. We can obtain $\widetilde{P}(\omega, r_2, \frac{\pi}{2}, \phi)$ and $\widetilde{V}(\omega, r_2, \frac{\pi}{2}, \phi)$ by subtracting the zeroth and the first order sound field components obtained in Sec. 6.3.1 from the the sound pressure and the radial particle velocity measurements on the circle $\mathrm{c}_{0,2}$ ((6.19) and (6.20)).

Based on (6.21), (6.19), (6.4), (6.6), (6.7), and (6.8), we can obtain the pressure field coefficient $\mathsf{P}_{2,-2}(\omega, r_2)$ as follows

$$\begin{aligned}
\mathsf{P}_{2,-2}(\omega, r_2) =& \frac{1}{\mathcal{P}_{2,-2}(\frac{\pi}{2})} \int_0^{2\pi} \widetilde{P}(\omega, r_2, \frac{\pi}{2}, \phi) E_{-2}(\phi) d\phi \\
=& \frac{\mathsf{P}_{2,-2}(\omega, r_2)\mathcal{P}_{2,-2}(\frac{\pi}{2})}{\mathcal{P}_{2,-2}(\frac{\pi}{2})} \underbrace{\int_0^{2\pi} E_{-2}(\phi) E_{-2}(\phi) d\phi}_{=1} \\
&+ \frac{\mathsf{P}_{2,0}(\omega, r_2)\mathcal{P}_{2,0}(\frac{\pi}{2})}{\mathcal{P}_{2,-2}(\frac{\pi}{2})} \underbrace{\int_0^{2\pi} E_{0}(\phi) E_{-2}(\phi) d\phi}_{=0} \\
&+ \frac{\mathsf{P}_{2,2}(\omega, r_2)\mathcal{P}_{2,2}(\frac{\pi}{2})}{\mathcal{P}_{2,-2}(\frac{\pi}{2})} \underbrace{\int_0^{2\pi} E_{2}(\phi) E_{-2}(\phi) d\phi}_{=0}.
\end{aligned} \tag{6.23}$$

Note that, in practical implementation of (6.23), we need to replace the continuous integral over $[0, 2\pi)$ by a summation over a finite number of points [34].

Similar to (6.23), we can obtain the pressure field coefficients

$$\mathsf{P}_{2,v}(\omega, r_2)= \frac{1}{\mathcal{P}_{2,v}(\frac{\pi}{2})} \int_0^{2\pi} \widetilde{P}(\omega, r_2, \frac{\pi}{2}, \phi) E_v(\phi) d\phi, \tag{6.24}$$

where $v = 0, 2$, and the velocity field coefficients

$$\mathsf{V}_{2,v}(\omega, r_2)= \frac{1}{\mathcal{P}_{2,v}(\frac{\pi}{2})} \int_0^{2\pi} \widetilde{V}(\omega, r_2, \frac{\pi}{2}, \phi) E_v(\phi) d\phi, \tag{6.25}$$

where $v = -2, 0, 2$.

Based on (6.4) and (6.5), we have the relationship between the pressure filed coefficients and the velocity field coefficients

$$\mathsf{P}_{2,-2}(\omega, r_2) = \mathsf{B}_{2,-2}(\omega) j_u(\omega\tau_2), \tag{6.26}$$

$$\mathsf{V}_{2,-2}(\omega, r_2) = \frac{i}{\rho c}\mathsf{B}_{2,-2}(\omega) j_u'(\omega\tau_2), \tag{6.27}$$

where $\tau_2 = r_2/c$. Multiplying both sides of (6.26) by $h_2'(\omega\tau_2)$ and both sides of (6.27) by $h_2(\omega\tau_2)$, we can obtain the following equations

$$\mathsf{P}_{2,-2}(\omega, r_2) h_2'(\omega\tau_2) = \mathsf{B}_{2,-2}(\omega) j_2(\omega\tau_2) h_2'(\omega\tau_2), \tag{6.28}$$

$$\mathsf{V}_{2,-2}(\omega, r_2) h_2(\omega\tau_2)=\frac{i}{\rho c}\mathsf{B}_{2,-2}(\omega) j_2'(\omega\tau_2) h_2(\omega\tau_2). \tag{6.29}$$

Based on (6.28) and (6.29), we obtain the spherical Bessel function coefficients $\mathsf{B}_{2,-2}(\omega)$ as

$$\begin{aligned}\mathsf{B}_{2,-2}(\omega) &= \frac{\mathsf{P}_{2,-2}(\omega, r_2) h_2'(\omega\tau_2) + i\rho c \mathsf{V}_{2,-2}(\omega, r_2) h_2(\omega\tau_2)}{[j_2(\omega\tau_2) h_2'(\omega\tau_2) - j_2'(\omega\tau_2) h_2(\omega\tau_2)]} \\ &= i(\omega\tau_2)^2 [\mathsf{P}_{2,-2}(\omega, r_2) h_2'(\omega\tau_2) + i\rho c \mathsf{V}_{2,-2}(\omega, r_2) h_2(\omega\tau_2)], \end{aligned} \tag{6.30}$$

where the last two lines of the above equation are obtained based on the Wronskian equation [36].

Similarly, we can obtain the following spherical Bessel function coefficients

$$\mathsf{B}_{2,v}(\omega) = i(\omega\tau_2)^2[\mathsf{P}_{2,v}(\omega, r_2)h_2'(\omega\tau_2) + i\rho c\mathsf{V}_{2,v}(\omega, r_2)h_2(\omega\tau_2)], \tag{6.31}$$

where $v = 0, 2$.

6.3.3 The spherical Bessel function coefficients $\mathsf{B}_{2,-1}(\omega)$, $\mathsf{B}_{2,1}(\omega)$

The sound pressure and the radial particle velocity on the circle $c_{1,2}$ with polar angle $\theta_{1,2}$ on the h_1 plane (as shown in Fig. 6.1) can be expressed as [73]

$$P(\omega, r_2, \theta_{1,2}, \phi) \approx \sum_{u=0}^{2} \sum_{v=-u}^{u} \mathsf{P}_{u,v}(\omega, r_2) Y_{u,v}(\theta_{1,2}, \phi), \tag{6.32}$$

$$V(\omega, r_2, \theta_{1,2}, \phi) \approx \sum_{u=0}^{2} \sum_{v=-u}^{u} \mathsf{V}_{u,v}(\omega, r_2) Y_{u,v}(\theta_{1,2}, \phi). \tag{6.33}$$

Define the following intermediate quantities

$$\begin{aligned}\widetilde{P}(\omega, r_2, \theta_{1,2}, \phi) = {} & P(\omega, r_2, \theta_{1,2}, \phi) \\ & - \underbrace{\sum_{u=0}^{2} \sum_{v=-u}^{u}}_{uv \neq 21, 2-1} \mathsf{P}_{u,v}(\omega, r_2) Y_{u,v}(\theta_{1,2}, \phi),\end{aligned} \tag{6.34}$$

$$\begin{aligned}\widetilde{V}(\omega, r_2, \theta_{1,2}, \phi) = {} & V(\omega, r_2, \theta_{1,2}, \phi) \\ & - \underbrace{\sum_{u=0}^{2} \sum_{v=-u}^{u}}_{uv \neq 21, 2-1} \mathsf{V}_{u,v}(\omega, r_2) Y_{u,v}(\theta_{1,2}, \phi).\end{aligned} \tag{6.35}$$

We can obtain $\widetilde{P}(\omega, r_2, \theta_{1,2}, \phi)$ and $\widetilde{V}(\omega, r_2, \theta_{1,2}, \phi)$ by subtracting the sound field components obtained in Sec. 6.3.1 and 6.3.2 from the the sound pressure and the radial particle velocity measurements on circle $c_{1,2}$ ((6.32) and (6.33)).

Similar to Sec. 6.3.2, we can obtain the following pressure field coefficients and velocity field coefficients

$$\mathsf{P}_{2,v}(\omega, r_2)= \frac{1}{\mathcal{P}_{2,v}(\theta_{1,2})} \int_0^{2\pi} \widetilde{P}(\omega, r_2, \theta_{1,2}, \phi) E_v(\phi) d\phi, \tag{6.36}$$

$$\mathsf{V}_{2,v}(\omega, r_2)= \frac{1}{\mathcal{P}_{2,v}(\theta_{1,2})} \int_0^{2\pi} \widetilde{V}(\omega, r_2, \theta_{1,2}, \phi) E_v(\phi) d\phi, \tag{6.37}$$

where $v = -1, 1$.

Similar to Sec. 6.3.2 based on (6.36) and (6.37), we can obtain the spherical Bessel function coefficients

$$\mathsf{B}_{2,v}(\omega) = i(\omega\tau_2)^2 [\mathsf{P}_{2,v}(\omega, r_2) h_2'(\omega\tau_2) + i\rho c \mathsf{V}_{2,v}(\omega, r_2) h_2(\omega\tau_2)], \tag{6.38}$$

where $v = -1, 1$.

6.3.4 The general expression for a high order spherical Bessel function coefficient

Based on (6.30), (6.31), and (6.38), we have the general expressions for a higher order ($u > 1$) spherical Bessel function coefficient as

$$\mathsf{B}_{u,v}(\omega) = i(\omega\tau_u)^2 [\mathsf{P}_{u,v}(\omega, r_u) h_u'(\omega\tau_u) + i\rho c \mathsf{V}_{u,v}(\omega, r_u) h_u(\omega\tau_u)], \tag{6.39}$$

where $\tau_u = r_u/c$. Then based on (6.4), we have the general expression for a high order ($u > 1$) pressure field coefficient as

$$\begin{aligned} \mathsf{P}_{u,v}(\omega, r) = i(\omega\tau_u)^2 \big(&\mathsf{P}_{u,v}(\omega, r_u) h_u'(\omega\tau_u) \\ &+ i\rho c \mathsf{V}_{u,v}(\omega, r_u) h_u(\omega\tau_u)\big) j_u(\omega r/c). \end{aligned} \tag{6.40}$$

Substituting (6.40) into (6.4), then we can reconstruct and analyze the frequency-domain sound pressure within the sphere $\mathbb{S}_2$. Similar to (6.40), we can obtain the velocity field coefficients as follows

$$\begin{aligned}\mathsf{V}_{u,v}(\omega,r) = \frac{-1}{\rho c}(\omega\tau_u)^2\big(&\mathsf{P}_{u,v}(\omega,r_u)h'_u(\omega\tau_u)\\ &+i\rho c\mathsf{V}_{u,v}(\omega,r_u)h_u(\omega\tau_u)\big)j'_u(\omega r/c).\end{aligned} \tag{6.41}$$

Substituting (6.41) into (6.5), we can obtain the frequency-domain radial particle velocity within the sphere $\mathbb{S}_2$.

The sound field analysis method is an iterative process. The matrix inversion process of the previous 3D sound field analysis methods is avoided [34], making the proposed method computationally friendly on embedded processors.

6.3.5 Design of arbitrary order arrays

In Sec. 6.2, 6.3.1, 6.3.2, 6.3.3, we use the second order arrays as an example to show how to obtain the spherical Bessel function coefficients iteratively. Here, we summarize the general procedure for designing arbitrary order arrays and obtaining the spherical Bessel function coefficients.

1. We place a single vector sensor at the origin. We have circles $\{\mathrm{c}_{0,u}, \mathrm{c}_{1,u}\}_{u=2}^{N}$ on two planes h_0 and h_1 as shown in Fig. 6.2, where N is the highest sound field order of interest. The circles are also on spheres $\{\mathbb{S}_u\}_{u=2}^{N}$ of radii $\{r_u\}_{u=2}^{N}$, where $r_u \approx uc/2\pi f_m$ [73]. That is, the sound field on the circles $\mathrm{c}_{0,u}$ and $\mathrm{c}_{1,u}$ can be accurately described by spherical harmonics up to order u. We place at least $u+1$ vector sensors uniformly on the circle $\mathrm{c}_{0,u}$, and at least u vector sensors uniformly on the circle $\mathrm{c}_{1,u}$ [34].

2. We obtain the zeroth and the first order spherical Bessel function coefficients based on the measurement of the vector sensor at the origin as shown in Sec. 6.3.1.

3. We subtract the low order ($\overline{u} < u$) sound field components from the measurements of the array on the circle $\mathrm{c}_{0,u}$. Integrating and dividing the results over $[0, 2\pi]$ and by $\mathcal{P}_{u,v}(\frac{\pi}{2})$, respectively, then we can obtain the pressure field coefficient $\mathsf{P}_{u,v}(\omega, r_u)$ and the velocity field coefficient $\mathsf{V}_{u,v}(\omega, r_u)$. Next

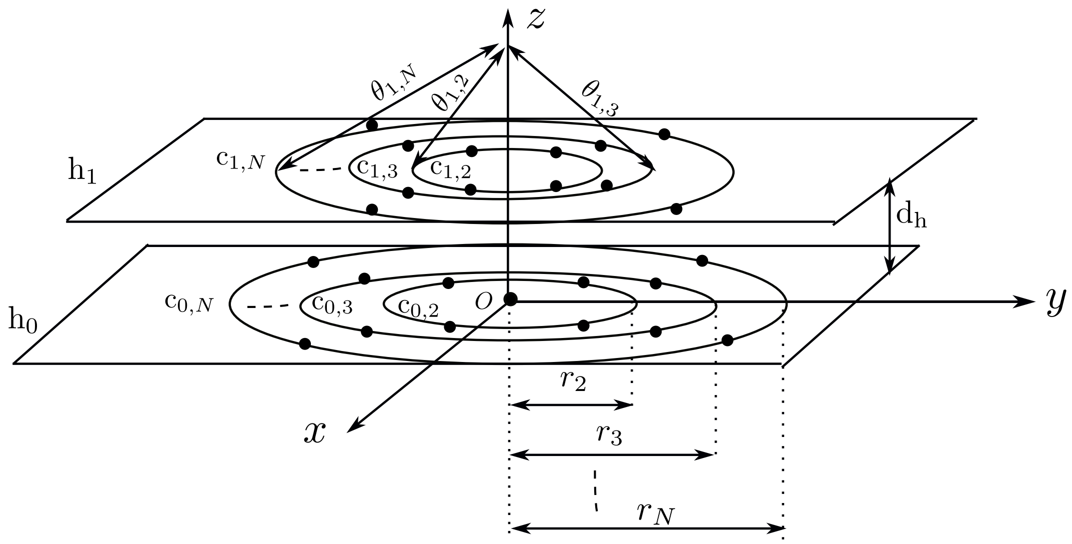


Figure 6.2: System description: We have a single vector sensor at the origin, and circles $\{c_{0,u}, c_{1,u}\}_{u=2}^{N}$ on two planes h_0 and h_1. The distance of the planes h_1 with respect to the plane plane h_0 is d_h. The circles are also on spheres $\{\mathbb{S}_u\}_{u=2}^{N}$ (which are not shown for ease of illustration) of radii $\{r_u\}_{u=2}^{N}$. We place at least $u+1$ vector sensors (•) uniformly on the circle $c_{0,u}$, and at least u vector sensors uniformly on the circle $c_{1,u}$. We can estimate the spherical Bessel function coefficients up to order N based on the array measurement.

based on the Wronskian equation, we can obtain the spherical Bessel function coefficients $\mathsf{B}_{u,v}(\omega)$, where $u+v$ is even [34]. (Sec. 6.3.2)

We summarize this process as a single equation

$$\begin{aligned}\mathsf{B}_{u,v}(\omega) = & \frac{i(\omega\tau_u)^2 h'_u(\omega\tau_u)}{\mathcal{P}_{u,v}(\frac{\pi}{2})} \int_0^{2\pi} E_v(\phi)\Big[P(\omega, r_u, \frac{\pi}{2}, \phi) \\ & - \sum_{\overline{u}=0}^{u-1} \sum_{\overline{v}=-\overline{u}}^{\overline{u}} \mathsf{B}_{\overline{uv}}(\omega) j_{\overline{u}}(\omega\tau_{\overline{u}}) Y_{\overline{uv}}(\frac{\pi}{2}, \phi)\Big] d\phi \\ & - \frac{\rho c(\omega\tau_u)^2 h_u(\omega\tau_u)}{\mathcal{P}_{u,v}(\frac{\pi}{2})} \int_0^{2\pi} E_v(\phi)\Big[V(\omega, r_u, \frac{\pi}{2}, \phi) \\ & - \sum_{\overline{u}=0}^{u-1} \sum_{\overline{v}=-\overline{u}}^{\overline{u}} \frac{i}{\rho c} \mathsf{B}_{\overline{u},\overline{v}}(\omega) j'_u(\omega\tau_{\overline{u}}) Y_{\overline{uv}}(\frac{\pi}{2}, \phi)\Big] d\phi. \end{aligned} \tag{6.42}$$

4. We subtract the low order ($\overline{u} < u$) and the u-th order even sound field

components (obtained in Step 3) from the measurements of the array on the circle $c_{1,u}$. Processing the results similar to Step 3, then we can obtain the spherical Bessel function coefficients $\mathsf{B}_{u,v}(\omega)$, where $u+v$ is odd [34]. (Sec. 6.3.3)

5. Repeating Step 3 and Step 4 for $u = 2, 3, 4, ..., N$, then we can obtain the spherical Bessel function coefficients $\{\mathsf{B}_{u,v}(\omega)\}_{u=1}^{N}$.

Note that the equations (6.36)-(6.37) involve the divide operations, which may cause numerical problems when $\mathcal{P}_{u,v}(\theta) \approx 0$ [36]. However, this problem can be avoided by choosing appropriate values for the polar angles $\{\theta_{1,u}\}_{u=2}^{N}$ for the circles on the plane h_1, or the distance $\mathrm{d_h}$ of plane h_1 with respect to plane h_0. For example, for the fourth order arrays, we can choose the distance to be $\mathrm{d_h} = r_4/6$.

6.4 Time-domain derivation

The frequency-domain sound field analysis methods suffer from the latency problem [128]. For applications with low-latency requirement, a time-domain method is desirable. In this section, we present the time-domain sound field analysis method corresponding to the frequency-domain method in Sec. 6.3. We start with the following two theorems. The proofs of the two theorems are at the end of this chapter.

Theorem 2 *The time-domain sound pressure at* (r_s, θ, ϕ) *is*

$$p(t, r_s, \theta, \phi) = \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{p}_{u,v}(t, r_s) Y_{u,v}(\theta, \phi), \tag{6.43}$$

where the low order ($u \leq 1$) *time-domain pressure field coefficients are given by*

$$\mathsf{p}_{0,0}(t, r_s) = \sqrt{8\pi^2} p_o(t) * \frac{1}{2\tau_s} \mathcal{P}_0(\frac{t}{\tau_s}), \tag{6.44}$$

$$\mathsf{p}_{1,-1}(t, r_s) = i\rho c \sqrt{24\pi^2} \nu_{oy}(t) * \frac{i}{2\tau_s} \mathcal{P}_1(\frac{t}{\tau_s}), \tag{6.45}$$

$$\mathsf{p}_{1,0}(t, r_s) = -i\rho c \sqrt{24\pi^2} \nu_{oz}(t) * \frac{i}{2\tau_s} \mathcal{P}_1(\frac{t}{\tau_s}), \tag{6.46}$$

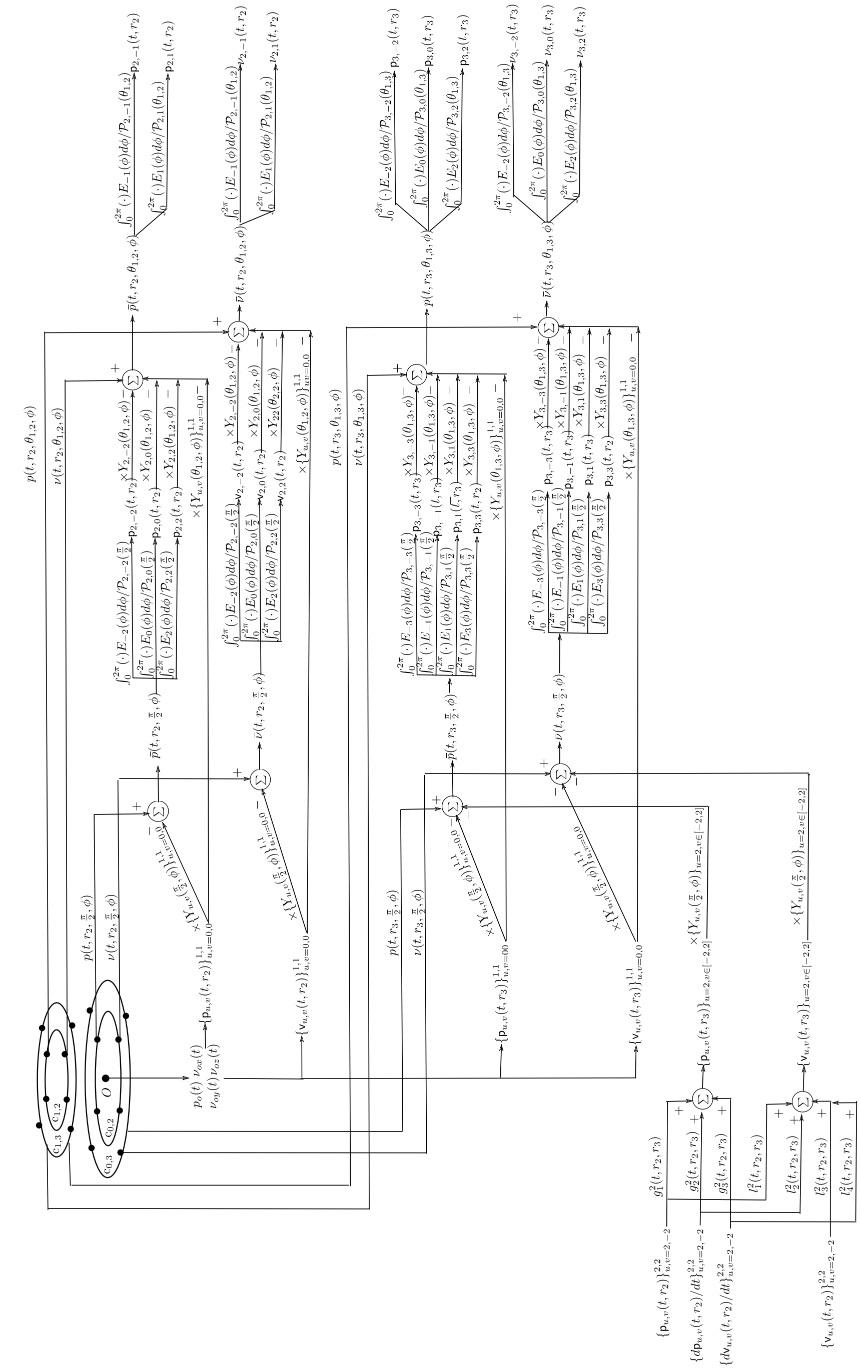


Figure 6.3: Signal flow of the proposed time-domain sound field analysis method for a third order array.

$$\mathsf{p}_{1,1}(t, r_s) = i\rho c\sqrt{24\pi^2}\nu_{ox}(t) * \frac{i}{2\tau_s}\mathcal{P}_1(\frac{t}{\tau_s}), \tag{6.47}$$

$\tau_s = r_s/c$, $*$ is the convolution operation, $p_o(t)$, $\nu_{ox}(t)$, $\nu_{oy}(t)$, $\nu_{oz}(t)$ are the time-domain pressure, the particle velocities in the x, y, and z directions, respectively, which are measured by the vector sensor at the origin, and $\mathcal{P}_u(\cdot)$ is the Legendre function of order u.

The higher order ($u > 1$) time-domain pressure field coefficients are given by

$$\begin{aligned} \mathsf{p}_{u,v}(t, r_s) = &\ \mathsf{p}_{u,v}(t, r_u) * g_1^u(t, r_u, r_s) \\ &+ \frac{d\mathsf{p}_{u,v}(t, r_u)}{dt} * g_2^u(t, r_u, r_s) \\ &+ \frac{d\mathsf{v}_{u,v}(t, r_u)}{dt} * g_3^u(t, r_u, r_s), \end{aligned} \tag{6.48}$$

where the spatial filters $g_1^u(t, r_u, r_s)$, $g_2^u(t, r_u, r_s)$, and $g_3^u(t, r_u, r_s)$ are theoretically derived

$$\begin{aligned} g_1^u(t, r_u, r_s) = &\ (u+1)(i)^{u+2}\frac{i^u}{2\tau_s}\mathcal{P}_u(\frac{t}{\tau_s}) \\ &* \Big[\delta(t-\tau_u) + \sum_{v=1}^{u}\frac{\varphi_v(u)}{(\tau_u)^v}\frac{(t-\tau_u)^{v-1}\mathrm{U}(t-\tau_u)}{(v-1)!}\Big], \end{aligned} \tag{6.49}$$

$$\begin{aligned} g_2^u(t, r_u, r_s) = &\ (\tau_u)(i)^{u}\frac{i^u}{2\tau_s}\mathcal{P}_u(\frac{t}{\tau_s}) \\ &* \Big[\delta(t-\tau_u) + \sum_{v=1}^{u-1}\frac{\varphi_v(u-1)}{(\tau_u)^v}\frac{(t-\tau_u)^{v-1}\mathrm{U}(t-\tau_u)}{(v-1)!}, \end{aligned} \tag{6.50}$$

$$\begin{aligned} g_3^u(t, r_u, r_s) = &\ (\rho c\tau_u)(i)^{u+2}\frac{i^u}{2\tau_s}\mathcal{P}_u(\frac{t}{\tau_s}) \\ &* \Big[\delta(t-\tau_u) + \sum_{v=1}^{u}\frac{\varphi_v(u)}{(\tau_u)^v}\frac{(t-\tau_u)^{v-1}\mathrm{U}(t-\tau_u)}{(v-1)!}\Big], \end{aligned} \tag{6.51}$$

$\delta(\cdot)$ is the delta function, $U(\cdot)$ is the step function, $\varphi_v(u) = [(u+v)!]/[2^v v!(u-v)!]$, $\mathsf{p}_{u,v}(t, r_u)$ and $\mathsf{v}_{u,v}(t, r_u)$ are the time-domain pressure field coefficients and velocity field coefficients, respectively.

Theorem 3 *The time-domain radial particle velocity at (r_s, θ, ϕ) is*

$$\nu(t, r_s, \theta, \phi) = \sum_{u=0}^{\infty} \sum_{v=-u}^{u} \mathsf{v}_{u,v}(t, r_s) Y_{u,v}(\theta, \phi), \tag{6.52}$$

where the low order ($u \leq 1$) time-domain velocity field coefficients are given by

$$\mathsf{v}_{0,0}(t, r_s) = \frac{1}{\rho c}\sqrt{8\pi^2} p_o(t) * \frac{t}{\tau_s}\frac{1}{2\tau_s}\mathcal{P}_0(\frac{t}{\tau_s}), \tag{6.53}$$

$$\mathsf{v}_{1,-1}(t, r_s) = i\sqrt{24\pi^2}\nu_{oy}(t) * \frac{t}{\tau_s}\frac{i}{2\tau_s}\mathcal{P}_1(\frac{t}{\tau_s}), \tag{6.54}$$

$$\mathsf{v}_{1,0}(t, r_s) = -i\sqrt{24\pi^2}\nu_{oz}(t) * \frac{t}{\tau_s}\frac{i}{2\tau_s}\mathcal{P}_1(\frac{t}{\tau_s}), \tag{6.55}$$

$$\mathsf{v}_{1,1}(t, r_s) = i\sqrt{24\pi^2}\nu_{ox}(t) * \frac{t}{\tau_s}\frac{i}{2\tau_s}\mathcal{P}_1(\frac{t}{\tau_s}). \tag{6.56}$$

The high order ($u > 1$) time-domain velocity field coefficients are given by

$$\begin{aligned}\mathsf{v}_{u,v}(t, r_s) = &\ \mathsf{p}_{u,v}(t, r_u) * l_1^u(t, r_u, r_s)\\ &+\frac{d\mathsf{p}_{u,v}(t, r_u)}{dt} * l_2^u(t, r_u, r_s)\\ &+\mathsf{v}_{u,v}(t, r_u) * l_3^u(t, r_u, r_s)\\ &+\frac{d\mathsf{v}_{u,v}(t, r_u)}{dt} * l_4^u(t, r_u, r_s),\end{aligned} \tag{6.57}$$

where the spatial filters $l_1^u(t, r_u, r_s)$, $l_2^u(t, r_u, r_s)$, $l_3^u(t, r_u, r_s)$, and $l_4^u(t, r_u, r_s)$, are theoretically derived

$$\begin{aligned}l_1^u(t, r_u, r_s) = &\frac{-1}{\rho c}\Big\{(-\frac{\tau_u}{\tau_s}u(u+1)) \times (i)^{u+2}\frac{i^u}{2\tau_s}\mathcal{P}_u(\frac{t}{\tau_s})\\ &* \sum_{v=0}^{u}\frac{\varphi_v(u)}{(\tau_u)^{v+1}}\frac{(t-\tau_u)^v \mathrm{U}(t-\tau_u)}{v!}\\ &+(u+1) \times (i)^{u+1}\frac{i^{u+1}}{2\tau_s}\mathcal{P}_{u+1}(\frac{t}{\tau_s})\end{aligned}$$

$$\begin{aligned}
&* \Big[\delta(t-\tau_u) + \sum_{v=1}^{u} \frac{\varphi_v(u)}{(\tau_u)^v} \frac{(t-\tau_u)^{v-1}\mathrm{U}(t-\tau_u)}{(v-1)!}\Big] \\
&+(-u\frac{\tau_u}{\tau_s}) \times (i)^u \frac{i^u}{2\tau_s} \mathcal{P}_u(\frac{t}{\tau_s}) \\
&* \Big[\delta(t-\tau_u) + \sum_{v=1}^{u-1} \frac{\varphi_v(u-1)}{(\tau_u)^\nu} \frac{(t-\tau_u)^{v-1}\mathrm{U}(t-\tau_u)}{(v-1)!}\Big]\Big\}, \qquad (6.58)
\end{aligned}$$

$$\begin{aligned}
l_2^u(t, r_u, r_s) = &\frac{-\tau_u}{\rho c} \times (i)^u \frac{i^{u+1}}{2\tau_s} \mathcal{P}_{u+1}(\frac{t}{\tau_s}) \\
&* \Big[\delta(t-\tau_u) + \sum_{v=1}^{u-1} \frac{\varphi_v(u-1)}{(\tau_u)^v} \frac{(t-\tau_u)^{v-1}\mathrm{U}(t-\tau_u)}{(v-1)!}, \qquad (6.59)
\end{aligned}$$

$$\begin{aligned}
l_3^u(t, r_u, r_s) = u&\frac{-\tau_u}{\tau_s} \times (i)^{u+2} \frac{i^u}{2\tau_s} \mathcal{P}_u(\frac{t}{\tau_s}) \\
&* \Big[\delta(t-\tau_u) + \sum_{v=1}^{u} \frac{\varphi_v(u)}{(\tau_u)^v} \frac{(t-\tau_u)^{v-1}\mathrm{U}(t-\tau_u)}{(v-1)!}\Big], \qquad (6.60)
\end{aligned}$$

$$\begin{aligned}
l_4^u(t, r_u, r_s) = (\tau_u) &\times (i)^{u+1} \frac{i^{u+1}}{2\tau_s} \mathcal{P}_{u+1}(\frac{t}{\tau_s}) \\
&* \Big[\delta(t-\tau_u) + \sum_{v=1}^{u} \frac{\varphi_v(u)}{(\tau_u)^v} \frac{(t-\tau_u)^{v-1}\mathrm{U}(t-\tau_u)}{(v-1)!}\Big]. \qquad (6.61)
\end{aligned}$$

The above two theorems relate the pressure field coefficients and the velocity field coefficients which are depend on different radii, and are the key of the proposed time-domain method. Here, we use the third order arrays as an example to illustrate the implementation of the method. We assume the arrays have been designed according to Sec. 6.3.5 and Fig. 6.2. We present the signal flow of the proposed time-domain method for the third order arrays in Fig. 6.3:

1. We measure the sound pressure and the particle velocities in the x, y, z directions at the origin. Processing the measurements according to (6.44) -(6.47), and (6.53)-(6.56), then we can obtain the pressure field coefficients $\{\mathsf{p}_{u,v}(t, r_2)\}$ and the velocity field coefficients $\{\mathsf{v}_{u,v}(t, r_2)\}$ where $u \in [0, 1], v \in$

$[-u, u]$.

2. Subtracting the products between the pressure field coefficients and the spherical harmonics $\sum_{u=0}^{1}\sum_{v=-u}^{u} \mathsf{p}_{u,v}(t, r_2) Y_{u,v}(\frac{\pi}{2}, \phi)$, where $u, v \neq (1, 0)$, from the pressure measurements $p(t, r_2, \frac{\pi}{2}, \phi)$ on the circle $c_{0,2}$, then we can obtain the intermediate pressure quantities $\widetilde{p}(t, r_2, \frac{\pi}{2}, \phi)$. Integrating the intermediate pressure quantities with the trigonometric functions $\{E_v(\phi)\}$, where $v = -2, 0, 2$, over the circle, and dividing the results by corresponding normalized associated Legendre functions $\{\mathcal{P}_{2,v}(\frac{\pi}{2})\}$, where $v = -2, 0, 2$, we can obtain the pressure field coefficients $\{\mathsf{p}_{u,v}(t, r_2)\}$, where $u = 2, v = -2, 0, 2$.

3. Subtracting the products between the pressure field coefficients and the spherical harmonics $\sum_{u=0}^{2}\sum_{v=-u}^{u} \mathsf{p}_{u,v}(t, r_2) Y_{u,v}(\theta_{1,2}, \phi)$, where $u, v \neq (2, -1), (2, 1)$, from the pressure measurements $p(t, r_2, \theta_{1,2}, \phi)$ on the circle $c_{1,2}$, then we can obtain the intermediate pressure quantities $\widetilde{p}(t, r_2, \theta_{1,2}, \phi)$. Processing the intermediate pressure quantities similarly as in Step 3, we can obtain the pressure field coefficients $\{\mathsf{p}_{u,v}(t, r_2)\}$, where $u = 2, v = -1, 1$.

4. Now we have all the second order pressure field coefficients $\{\mathsf{p}_{u,v}(t, r_2)\}_{u,v=0,0}^{2,2}$ available. In a similar way, we can obtain the second order velocity field coefficients $\{\mathsf{v}_{u,v}(t, r_2)\}_{u,v=0,0}^{2,2}$. Substituting $\{\mathsf{p}_{u,v}(t, r_2)\}_{u,v=0,0}^{2,2}$ and $\{\mathsf{v}_{u,v}(t, r_2)\}_{u,v=0,0}^{2,2}$ into (6.43) and (6.52), respectively, we have the pressure and the radial particle velocity on a sphere of radius r_2 around the origin.

   Note that the pressure field coefficients and the velocity field coefficients obtain from Step 1 to Step 4 are dependent on the radius r_2.

5. Similar to Step 1, we can obtain the pressure field coefficients $\{\mathsf{p}_{u,v}(t, r_3)\}_{u,v=0,0}^{1,1}$ and the velocity field coefficients $\{\mathsf{v}_{u,v}(t, r_3)\}_{u,v=0,0}^{1,1}$.

6. We obtain the time derivatives of the pressure field coefficient $d\mathsf{p}_{u,v}(t, r_2)/dt$ and the time derivatives of the velocity field coefficients $d\mathsf{v}_{u,v}(t, r_2)/dt$ by a finite difference process [124]. Then project the coefficients $\{\mathsf{p}_{u,v}(t, r_2), d\mathsf{p}_{u,v}(t, r_2)/dt, \mathsf{v}_{u,v}(t, r_2), d\mathsf{v}_{u,v}(t, r_2)/dt\}$) to be $\{\mathsf{p}_{u,v}(t, r_3), \mathsf{v}_{u,v}(t, r_3)\}$, where $u = 2, v \in [-2, 2]$ based on (6.48) and (6.57).

7. Subtracting the contributions of the zeroth, the first, and the second order pressure field coefficients and velocity field coefficients, $\{\mathsf{p}_{u,v}(t, r_3), \mathsf{v}_{u,v}(t, r_3)\}_{u,v=0,0}^{2,2}$ from the pressure and velocity measurements on circles $c_{0,3}$ and $c_{1,3}$. Then based on the orthogonality among the trigonometric functions, we can obtain the third order pressure field coefficients $\{\mathsf{p}_{u,v}(t, r_3)\}$ and velocity field coefficients $\{\mathsf{v}_{u,v}(t, r_3)\}$, where $u = 3, v \in [-3, 3]$.

8. Now we have the zeroth, the first, the second, and the third order pressure field coefficients $\{\mathsf{p}_{u,v}(t, r_3)\}_{u,v=0,0}^{33}$ and velocity field coefficients $\{\mathsf{v}_{u,v}(t, r_3)\}_{u,v=0,0}^{33}$ available. Bringing the coefficients into (6.43) and (6.52), we can analyze the pressure and the radial particle velocity on a sphere of radius r_3 around the origin.

   Note that the pressure field coefficients and the velocity field coefficients obtain from Step 5 to Step 8 are dependent on the radius r_3.

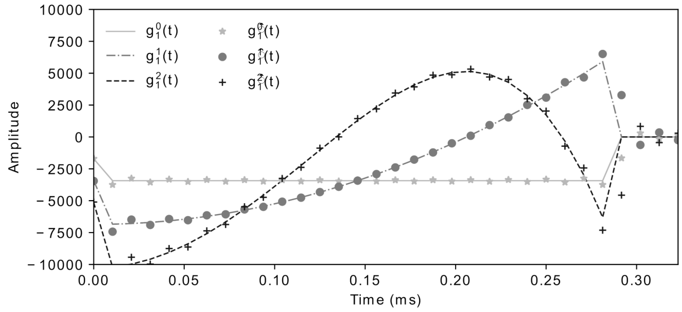
We have the following comments on the time-domain method:

- We have specially derived the spatial filters (6.49)-(6.51), (6.58)-(6.61) to have stable impulse responses [124]. This makes us do not have direct access to the radial-independent time-domain spherical Bessel function coefficients $\mathsf{b}_{u,v}(t)$. This also makes the time-domain pressure field coefficients and velocity field coefficients to be radial-dependent.

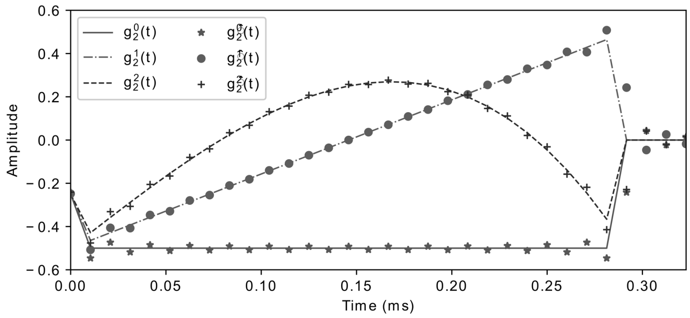
- We present the theoretically derived spatial filters (6.49)-(6.51) and their estimations (obtained through direct inverse discrete Fourier transform) in Fig. 6.4. As shown in Fig. 6.4, the theoretically derived spatial filters and their estimations agree well. The theoretically derived spatial filters (6.58)-(6.61) also agree well with their estimations, and the corresponding figures are not shown for brevity.

- To simplify the expressions, in this chapter we choose to express the spatial filters using the Legendre function rather than the Sign function as in Chap. 3.

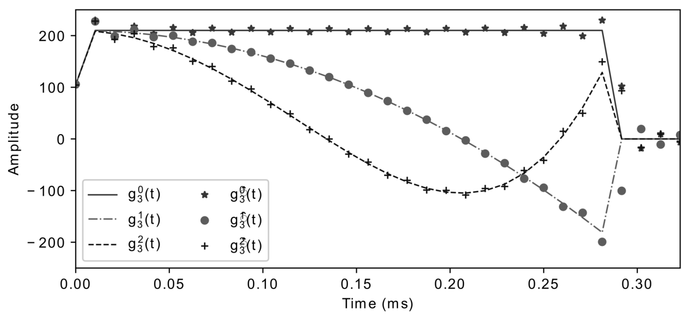
- The spatial filters (6.49)-(6.51), (6.58)-(6.61) are non-causal, and introduce a time advancement of $\tau_N = r_N/c$. When implementing (6.48) and (6.57), we



(a)



(b)



(c)

Figure 6.4: The plots of the theoretically derived spatial filters $g_1^\mu(t)$, $g_2^\mu(t)$, $g_3^\mu(t)$ and the numerically calculated $\widehat{g}_1^\mu(t)$, $\widehat{g}_2^\mu(t)$, $\widehat{g}_3^\mu(t)$ for $\mu = 0, 1, 2$, $r_u = r_s = 0.05$ m, $c = 343$ m/s, ρ= 1.225 kg/m^3.

need to delay the pressure field coefficients and the velocity field coefficients by τ_N. For a compact array, where r_N is small, τ_N is negligible compared with the frame delay introduced by the frequency-domain method in Sec. 6.3.

6.5 Simulation examples

In this section, we design second order arrays, and validate the effectiveness of the proposed methods by simulation examples.

We have a number of target sources, whose spherical coordinates are (r_q, θ_q, ϕ_q), where $r_q = 0.5$ m, θ_q can be either entry in the array $[\pi/5,\ \pi/3,\ \pi/2,\ 4\pi/5,\ 7\pi/8]$, ϕ_q can be either entry in the array $[0.0,\ \pi/3,\ 5\pi/6,\ 7\pi/6,\ 7\pi/4]$. The outputs of the target sources are 25 mutual independent unit-variance Gaussian white noises, which have been filtered by a 64 tap Bandpass Butterworth filter with frequency rage [100, 2000] Hz.

We have a single vector sensor at the origin [120]. We let the radius $r_2 = 0.05$ m, such that $\lceil 2\pi f_{\mathrm{m}} r_2/c \rceil = 2$. The polar angle of the plane h_1 is $\theta_{1,2} = \pi/3$. We equally place 5 vector sensors on the circles $\mathrm{c}_{0,2}$ and $\mathrm{c}_{1,2}$, respectively. The array setup is shown by Fig. 6.1.

We conduct two simulations, and the first simulation is to reconstruct (analyze) the target sound field over the sphere $\mathbb{S}_2$ based on the array measurement in the frequency-domain. We simulate the transfer function between the target sources and the sensors (and the sampling points on the sphere $\mathbb{S}_2$) using the free-space Green function and the Euler's equation [36]. The ambient density is $\rho = 1.225$ kg/m^3, the speed of sound is $c = 343$ m/s, and the sampling frequency is $f_{\mathrm{s}} = 48$ kHz. We add Gaussian white noises to the sensor measurements such that the signal to noise power ratio is 40 dB. The simulation results are from an average of 100 independent runs.

We present the normalized sound field reconstruction error $\xi(\theta_q, \phi_q)$ in Fig. 6.5, where

$$\xi(\theta_q, \phi_q) = 10\ \log 10 \frac{\sum_{\omega=100\times 2\pi}^{2000\times 2\pi} ||e(\omega, r_2, \theta_q, \phi_q)||_2^2}{\sum_{\omega=100\times 2\pi}^{2000\times 2\pi} ||p(\omega, r_2, \theta_q, \phi_q)||_2^2},$$

$||\cdot||_2$ is the 2-norm, $p(\omega, r_2, \theta_q, \phi_q)$ is the target sound field at frequency ω and at position (r_2, θ_q, ϕ_q), $e(\omega, r_2, \theta_q, \phi_q)$ is the corresponding sound field reconstruction

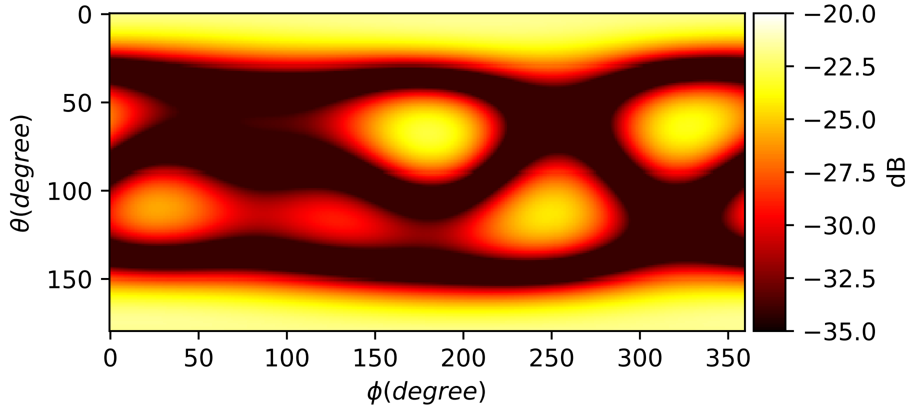


Figure 6.5: The normalized sound field reconstruction error $\xi(\theta_q, \phi_q)$ over the sphere $\mathbb{S}_2$.

error. As shown in Fig. 6.5, the normalized sound field reconstruction error is less than -20 dB over the whole sphere.

We further present the normalized sound field reconstruction error as a function of frequency in Fig. 6.6,

$$\xi(\omega) = 10 \log10 \frac{\sum_{q=1}^{180\times 360} ||e(\omega, r_2, \theta_q, \phi_q)||_2^2}{\sum_{q=1}^{180\times 360} ||p(\omega, r_2, \theta_q, \phi_q)||_2^2},$$

where (r_2, θ_q, ϕ_q) are the spherical coordinates of 180×360 uniform sampling points on the sphere $\mathbb{S}_2$ [75]. As shown in Fig. 6.6, the normalized sound field reconstruction error $\xi(\omega)$ is about -33 dB below 1000 Hz, and increases linearly above 1250 Hz.

We conduct the second simulation to validate the proposed time-domain method. The basic simulation settings are the same as in the first simulation, except that the spatial filters are realized as finite impulse response filters of length $\lceil 2r_2/c * f_\mathrm{s} \rceil + 1 = 15$ taps [8]. We conduct the simulation over a period of $T = 1.0$ s, and present the sound pressure $p(t, r_2, \theta_q, \phi_q)$, its reconstruction $\widehat{p}(t, r_2, \theta_q, \phi_q)$, and the reconstruction error $e(t, r_2, \theta_q, \phi_q) = p(t, r_2, \theta_q, \phi_q) - \widehat{p}(t, r_2, \theta_q, \phi_q)$ over a period of 8 ms in

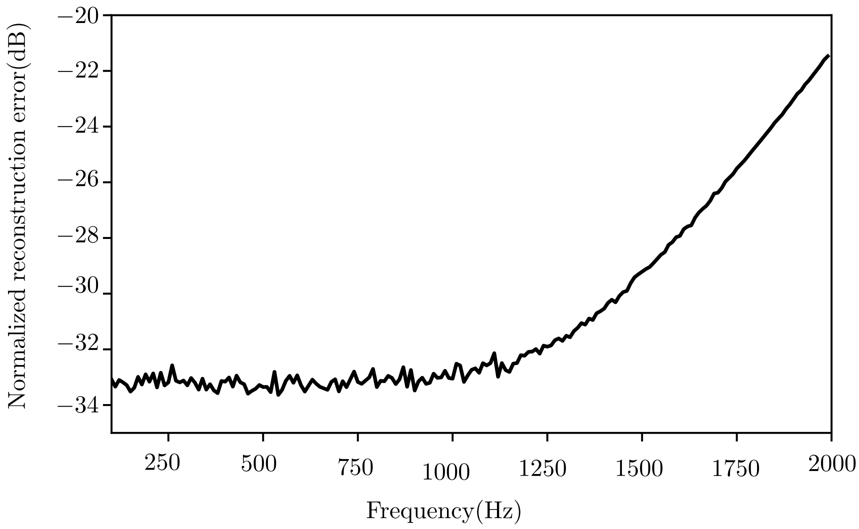


Figure 6.6: The normalized sound field reconstruction error $\xi(\omega)$ over the sphere $\mathbb{S}_2$ as a function of frequency

Fig. 6.7, where $(\theta_q, \phi_q) = (\pi/4, \pi/3)$. The normalized reconstruction error is

$$\xi(\theta_q, \phi_q) = 10 \log 10 \frac{\sum_{t=0}^{1.0\text{ s}} ||e(t, r_2, \theta_q, \phi_q)||_2^2}{\sum_{t=0}^{1.0\text{ s}} ||p(t, r_2, \theta_q, \phi_q)||_2^2} = -28.3 \text{ dB}.$$

The time-domain sound field reconstruction error at other positions on the sphere are similar to Fig. 6.7, and thus are not shown for brevity.

6.6 Conclusion

In this chapter, we proposed multiple circular arrays of vector sensors with 3D sound field analysis capacity, and derived the corresponding frequency-domain and time-domain sound field analysis methods. These methods analyze the components of the sound field successively without conducting the matrix inversion process as the previous sound field analysis methods, and thus is numerically robust. Additionally, the time-domain method is suitable for real-time applications because of its low latency. We do acknowledge that the vector sensors are too expensive for constructing an array, but we can use closely spaced microphone pairs to approxi-

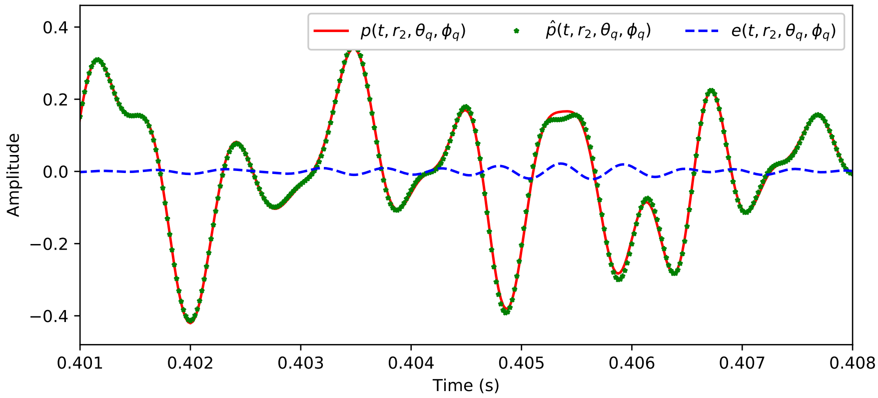


Figure 6.7: The sound pressure $p(t, r_2, \theta_q, \phi_q)$, its reconstruction $\widehat{p}(t, r_2, \theta_q, \phi_q)$, and the reconstruction error $e(t, r_2, \theta_q, \phi_q)$ over a period of 8 ms, where $(\theta_q, \phi_q) = (\pi/4, \pi/3)$.

mate the vector sensors, and this will be one of our future work.

Publications related to this chapter

- **Fei Ma**, Thushara D. Abhayapala, Wen Zhang, "Multiple circular arrays of vector sensors for real-time 3D sound field analysis," IEEE Transactions on Audio, Speech, and Language Processing, (under review).

6.7 Proof of the time-domain sound field analysis method

We first derive the low order time-domain pressure filed coefficients. Based on (6.4) and (6.43), we express the low order ($u \leq 1$) time-domain pressure field coefficients as

$$\mathsf{p}_{u,v}(t, r_s) = \mathcal{F}^{-1}[\mathsf{B}_{u,v}(\omega)] * \mathcal{F}^{-1}[j_u(\omega\tau_s)], \tag{6.62}$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform [128]. Substituting (6.15), (6.16), (6.17), (6.18), and the following [72]

$$\mathcal{F}^{-1}[j_u(\omega\tau_s)] = \frac{i^u}{2\tau_s}\mathcal{P}_u(\frac{t}{\tau_s}). \tag{6.63}$$

into (6.62), we can derive (6.44), (6.45), (6.46), (6.47).

Next we derive the high order ($u > 1$) time-domain pressure filed coefficients. Based on the expansions of the spherical Bessel function and the spherical Hankel function, we express the high order frequency-domain pressure field coefficients (6.40) as [72]

$$\begin{aligned}\mathsf{P}_{u,v}(\omega, r_s) = [&\mathsf{P}_{u,v}(\omega) \times (i\omega\tau_u)h_u(\omega\tau_u)j_u(\omega\tau_s)(u+1)\\ &+i\omega\mathsf{P}_{u,v}(\omega) \times (\omega\tau_u)h_{u-1}(\omega\tau_u)j_u(\omega\tau_s)(\tau_u)\\ &+i\omega\mathsf{V}_{u,v}(\omega) \times (i\omega\tau_u)h_u(\omega\tau_u)j_u(\omega\tau_s)(\rho c\tau_u)].\end{aligned} \tag{6.64}$$

Define the following quantities

$$\mathsf{p}_{u,v}(t, r_u) = \mathcal{F}^{-1}[\mathsf{P}_{u,v}(\omega, r_u)], \tag{6.65}$$

$$\mathsf{v}_{u,v}(t, r_u) = \mathcal{F}^{-1}[\mathsf{V}_{u,v}(\omega, r_u)], \tag{6.66}$$

$$\frac{d\mathsf{p}_{u,v}(t, r_u)}{dt} = \mathcal{F}^{-1}[i\omega\mathsf{P}_{u,v}(\omega, r_u)], \tag{6.67}$$

$$\frac{d\mathsf{v}_{u,v}(t.r_u)}{dt} = \mathcal{F}^{-1}[i\omega\mathsf{V}_{u,v}(\omega, r_u)], \tag{6.68}$$

$$g_1^u(t, r_u, r_s) = \mathcal{F}^{-1}[\underbrace{i(\omega\tau_u)h_u(\omega\tau_u)j_u(\omega\tau_s)(u+1)}_{g_1^u(\omega, r_u, r_s)}], \tag{6.69}$$

$$g_2^u(t, r_u, r_s) = \mathcal{F}^{-1}[\underbrace{(\omega\tau_u)h_{u-1}(\omega\tau_u)j_u(\omega\tau_s)(\tau_u)}_{g_2^u(\omega, r_u, r_s)}], \tag{6.70}$$

$$g_3^u(t, r_u, r_s) = \mathcal{F}^{-1}[\underbrace{i(\omega\tau_u)h_u(\omega\tau_u)j_u(\omega\tau_s)(\rho c\tau_u)}_{g_3^u(\omega, r_u, r_s)}]. \tag{6.71}$$

Based on the expansions of the spherical Bessel function and the spherical Hankel function, we can express $g_1^u(\omega, r_u, r_s)$, $g_2^u(\omega, r_u, r_s)$, and $g_3^u(\omega, r_u, r_s)$ as [72]

$$\begin{aligned} g_1^u(\omega, r_u, r_s) &= (u+1)i(\omega\tau_u)j_u(\omega\tau_s)h_u(\omega\tau_u) \\ &= (u+1)(i)^{u+2}\sum_{v=0}^{u}\frac{\varphi_v(u)}{(\tau_u)^v}\frac{j_u(\omega\tau_s)e^{-i\omega\tau_u}}{(i\omega)^v}, \end{aligned} \tag{6.72}$$

$$\begin{aligned} g_2^u(\omega, r_u, r_s) &= (\tau_u)(\omega\tau_u)j_u(\omega\tau_s)h_{u-1}(\omega\tau_u) \\ &= (\tau_u)(i)^{u}\sum_{v=0}^{u-1}\frac{\varphi_v(u-1)}{(\tau_u)^v}\frac{j_u(\omega\tau_s)e^{-i\omega\tau_u}}{(i\omega)^v}, \end{aligned} \tag{6.73}$$

$$\begin{aligned} g_3^u(\omega, r_u, r_s) &= (\rho c\tau_u)i(\omega\tau_u)j_u(\omega\tau_s)h_u(\omega\tau_u) \\ &= (\rho c\tau_u)(i)^{u+2}\sum_{v=0}^{u}\frac{\varphi_\nu(u)}{(\tau_u)^v}\frac{j_u(\omega\tau_s)e^{-i\omega\tau_u}}{(i\omega)^v}. \end{aligned} \tag{6.74}$$

Based on (6.63), inverse Fourier transform of above three equations, then we can derive (6.49), (6.50), and (6.51) [128]. Based on (6.64)-(6.71), we can derive (6.48).

Hereafter, we derive the low order ($u \leq 1$) time-domain velocity filed coefficients. Based on (6.5) and (6.52), we express the low order ($u \leq 1$) time-domain pressure field coefficients as

$$\begin{aligned} \mathsf{v}_{u,v}(t, r_s) &= \frac{1}{\rho c}\mathcal{F}^{-1}[\mathsf{B}_{u,v}(\omega)ij_u'(\omega\tau_s)] \\ &= \frac{1}{\rho c}\mathcal{F}^{-1}[\mathsf{B}_{u,v}(\omega)] * \mathcal{F}^{-1}[ij_u'(\omega\tau_s)] \end{aligned} \tag{6.75}$$

Substituting (6.15), (6.16), (6.17), (6.18), and the following [72]

$$\mathcal{F}^{-1}[ij_u'(\omega\tau_s)] = \frac{t}{\tau_s}\frac{i^u}{2\tau_s}\mathcal{P}_u(\frac{t}{\tau_s}), \tag{6.76}$$

into (6.75), we can derive (6.53), (6.54), (6.55), (6.56).

Next we derive the high order ($u > 1$) time-domain velocity filed coefficients.

Based on the expansions of the spherical Bessel function and the spherical Hankel function, we express the high order frequency-domain velocity field coefficients (6.41) as [72]

$$
\begin{aligned}
\mathsf{V}_{u,v}(\omega, r_s) &= \mathsf{P}_{u,v}(\omega) \times (\omega\tau_u)^2[j_u'(\omega\tau_s)h_u'(\omega\tau_u)] \times \frac{-1}{\rho c} \\
&\quad -\mathsf{V}_{u,v}(\omega) \times i(\omega\tau_u)^2[j_u'(\omega\tau_s)h_u(\omega\tau_u)] \\
&= \mathsf{P}_{u,v}(\omega) \times \Big[- j_u(\omega\tau_s)h_u(\omega\tau_u) \times \frac{\tau_u}{\tau_s}u(u+1) \\
&\qquad +(\omega\tau_u)j_{u+1}(\omega\tau_s)h_u(\omega\tau_u) \times (u+1) \\
&\qquad -(\omega\tau_u)j_u(\omega\tau_s)h_{u-1}(\omega\tau_u) \times u\frac{\tau_u}{\tau_s}\Big]\frac{-1}{\rho c} \\
&\quad +i\omega\mathsf{P}_{u,v}(\omega) \times (i\omega\tau_u)h_{u-1}(\omega\tau_u)j_{u+1}(\omega\tau_s) \times \frac{-\tau_u}{\rho c} \\
&\quad +\mathsf{V}_{u,v}(\omega) \times i(\omega\tau_u)j_u(\omega\tau_s)h_u(\omega\tau_u) \times u\frac{-\tau_u}{\tau_s} \\
&\quad +(i\omega)\mathsf{V}_{u,v}(\omega) \times (\omega\tau_u)j_{u+1}(\omega\tau_s)h_u(\omega\tau_u) \times (\tau_u). \\
&= \mathsf{P}_{u,v}(\omega) \times l_1^u(\omega, r_u, r_s) + i\omega\mathsf{P}_{u,v}(\omega) \times l_2^u(\omega, r_u, r_s) \\
&\quad +\mathsf{V}_{u,v}(\omega) \times l_3^u(\omega, r_u, r_s) + i\omega\mathsf{V}_{u,v}(\omega) \times l_4^u(\omega, r_u, r_s).
\end{aligned} \tag{6.77}
$$

Based on the above equation, we have the following definitions and expressions [72]

$$
\begin{aligned}
l_1^u(\omega, r_u, r_s) &= \frac{-1}{\rho c}\Big[j_u(\omega\tau_s)h_u(\omega\tau_u) \times (-\frac{\tau_u}{\tau_s}u(u+1)) \\
&\quad +(\omega\tau_u)j_{u+1}(\omega\tau_s)h_u(\omega\tau_u) \times (u+1) \\
&\quad +(\omega\tau_u)j_u(\omega\tau_s)h_{u-1}(\omega\tau_u) \times (-u\frac{\tau_u}{\tau_s})\Big] \\
&= \frac{-1}{\rho c}\Big[(i)^{u+2}\sum_{v=0}^{u}\frac{\varphi_v(u)}{(\tau_u)^{v+1}}\frac{j_u(\omega\tau_s)e^{-i\omega\tau_u}}{(i\omega)^{v+1}} \times (-\frac{\tau_u}{\tau_s}u(u+1)) \\
&\quad +(i)^{u+1}\sum_{v=0}^{u}\frac{\varphi_v(u)}{(\tau_u)^{v}}\frac{j_{u+1}(\omega\tau_s)e^{-i\omega\tau_u}}{(i\omega)^{v}} \times (u+1) \\
&\quad +(i)^{u}\sum_{v=0}^{u-1}\frac{\varphi_v(u-1)}{(\tau_u)^{v}}\frac{j_u(\omega\tau_s)e^{-i\omega\tau_u}}{(i\omega)^{v}} \times (-u\frac{\tau_u}{\tau_s})\Big],
\end{aligned} \tag{6.78}
$$

$$
l_2^u(\omega, r_u, r_s) = (i\omega\tau_u)h_{u-1}(\omega\tau_u)j_{u+1}(\omega\tau_s) \times \frac{-\tau_u}{\rho c}
$$

$$= (i)^u \sum_{v=0}^{u-1} \frac{\varphi_v(u-1)}{(\tau_u)^v} \frac{e^{-i\omega\tau_u} j_{u+1}(\omega\tau_s)}{(i\omega)^v} \times \frac{-\tau_u}{\rho c}, \tag{6.79}$$

$$\begin{aligned} l_3^u(\omega, r_u, r_s) &= i(\omega\tau_u) j_u(\omega\tau_s) h_u(\omega\tau_u) \times u\frac{-\tau_u}{\tau_s} \\ &= (i)^{u+2} \sum_{v=0}^{u} \frac{\varphi_v(u)}{(\tau_u)^v} \frac{j_u(\omega\tau_s) e^{-i\omega\tau_u}}{(i\omega)^v} \times u\frac{-\tau_u}{\tau_s}, \end{aligned} \tag{6.80}$$

$$\begin{aligned} l_4^u(\omega, r_u, r_s) &= (\omega\tau_u) j_{u+1}(\omega\tau_s) h_u(\omega\tau_u) \times (\tau_u) \\ &= (i)^{u+1} \sum_{v=0}^{u} \frac{\varphi_v(u)}{(\tau_u)^v} \frac{j_{u+1}(\omega\tau_s) e^{-i\omega\tau_u}}{(i\omega)^v} \times (\tau_u). \end{aligned} \tag{6.81}$$

Based on (6.63), inverse Fourier transform of above four equations, then we can obtain (6.58), (6.59), (6.60), and (6.61). Based on (6.65)-(6.68), (6.77)-(6.81), we can arrive at (6.57).

Chapter 7

Conclusions and Future Research Directions

7.1 Conclusions

The key of this thesis is introducing the idea of SFS to the field of spatial ANC. Though over the past three decades, researchers have developed a number of spatial ANC systems, most of the systems have neglected the fact that a spatial noise field is the complicated interaction between a number of noise sources and the environment. Thus, the spatial ANC systems may not achieve their designed ANC performance. This thesis propose to pre-process the acoustic measurements by the SFS method, which decouples the interactions between the noise sources and the environment, and significantly facilitates the following ANC process.

The contributions of this thesis are three folds:

1. We proposed a tonal and a random spatial ANC system in Chap. 3 and Chap. 5, respectively.

   Though the idea of outgoing noise field cancellation has been explored by other researchers before [56], their explorations were in a free field scenario without taking the complexity of real noise fields into consideration. We presented a frequency-domain SFS method and derived the corresponding time-domain method. These two methods separate the outgoing field and the incoming field on a sphere, allowing the proposed ANC systems to cancel the outgoing field without real-time estimation of the secondary path and with negligible disturbing to the desired field in the room. These two methods make the global cancellation of noise fields in a room practical.

2. The time-domain SFS method derived in Chap. 4 can serve as a foundation for developing future time-domain spatial algorithms.

   Though there are some other time-domain spatial sound field processing algorithms [113, 118, 119, 133, 134], they involve some impractical spatial filters or the multichannel time-domain deconvolution operations. The proposed time-domain SFS method, on the other hand, has been specially derived to have stable spatial filters and avoid the deconvolution operations. The special derivation makes the method suitable for real-time implementations on embedded systems. We can develop novel time-domain spatial algorithms based on the time-domain SFS method.

3. The multiple circular arrays can provide AR/VR devices with real-time 3D sound field analysis capacity.

   The spatial sound field analysis methods accompanying the existing spherical and non-spherical arrays are all frequency-domain methods. The frame delay of the STFT process makes these methods unable to exploit the spatial feature of a sound field in real-time. The proposed arrays and the time-domain sound field analysis method, on the other hand, do not suffer form the latency problem. The proposed arrays are compact in geometry and thus can be attached to the surface of a wearable device, and provide it with real-time sound field analysis capacity [19, 29, 135]. The proposed arrays can be used for selective ANC for AR/VR devices [19].

7.2 Future research

Here are some problems arising from the research of this thesis.

Theoretical analysis for the outgoing field control systems

In Chap. 3 and Chap. 5, we proposed a tonal and a random spatial ANC system, respectively. In both the simulations and the experiments for these systems, the arrangement of the sources and sensors are determined by a trial-and-error process. In the future, we need to theoretically analyze how does the arrangement of sources and sensors influence the performance of the systems. We need to derive a spatial coherence function, which can predict the performance of a random spatial ANC system.

Time-domain spatial field processing methods
In Chap. 4 and Chap. 7, we derived a time-domain spatial SFS method and a time-domain sound field analysis method, respectively. These time-domain methods avoid the STFT process and thus have lower latency. It maybe possible to further develop other time-domain spatial field processing methods, such as beamforming methods [136–138], or direction of arrival estimation methods [132, 139, 140].

Theoretical analysis of the surface scattering for the multiple circular arrays
In Chap. 7, we designed multiple circular arrays for analyzing the 3D sound field. For practical applications, such as providing an AR glass with real-time 3D sound field analysis capacity, we need to mount the array on a scattering surface, i.e., the glass. We need to analyze how the surface scatter the sound field, and how to modify the array structure correspondingly [141–143].

Novel real-time source separation methods
The SFS methods presented in this thesis have been proved to be useful for spatial ANC systems. However, the methods work best if we can surround the noise sources with a spherical array, which may not be practical for some scenarios. We need to develop novel real-time SFS methods which can be implemented using a small sized sensor array [34]. To further improve the accuracy of the sources separation process, we should integrate the array based methods with the machine learning based methods [144–147].

Bibliography

[1] S. M. Kuo and D. R. Morgan, *Active noise control systems: algorithms and DSP implementations*, John Wiley & Sons, Inc., New York, NY, USA, 1995.

[2] T. J. Sutton, S. J. Elliott, A. M. McDonald, and T. J. Saunders, "Active control of road noise inside vehicles," *Noise Control Eng. J.*, vol. 42, no. 4, pp. 137–147, 1994.

[3] S. J. Elliott, "A review of active noise and vibration control in road vehicles," *ISVR Tech. Memorandum*, vol. 981, 2008.

[4] P. A. Nelson, J. K. Hammond, P. Joseph, and S. J. Elliott, "Active control of stationary random sound fields," *J. Acoust. Soc. Am.*, vol. 87, no. 3, pp. 963–975, Mar. 1990.

[5] S. Laugesen and S. J. Elliott, "Multichannel active control of random noise in a small reverberant room," *IEEE Trans. Audio Speech Lang. Process.*, vol. 1, no. 2, pp. 241–249, Apr. 1993.

[6] J. W. Parkins, S. D. Sommerfeldt, and J. Tichy, "Narrowband and broadband active control in an enclosure using the acoustic energy density," *J. Acoust. Soc. Am.*, vol. 108, no. 1, pp. 192–203, July 2000.

[7] S. J. Elliott and J. Cheer, "Modeling local active sound control with remote sensors in spatially random pressure fields," *J. Acoust. Soc. Am.*, vol. 137, no. 4, pp. 1936–1946, Apr. 2015.

[8] F. Ma, W. Zhang, and T. D. Abhayapala, "Reference signal generation for broadband ANC systems in reverberant rooms," in *2018 IEEE International*

Conference on Acoustics, Speech and Signal Processing (ICASSP), Apr. 2018, pp. 216–220.

[9] K. Iwai, S. Kinoshita, and Y. Kajikawa, "Multichannel feedforward active noise control system combined with noise source separation by microphone arrays," *J. Sound Vib.*, vol. 453, pp. 151–173, Aug. 2019.

[10] H. Kuttruff, *Room Acoustics*, CRC Press, Sept. 2016.

[11] S. J. 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.*, vol. 35, no. 10, pp. 1423–1434, Oct. 1987.

[12] C. Bao, P. Sas, and B. H. Van, "Adaptive active control of noise in 3D reverberant enclosures," *J. Sound Vib.*, vol. 161, no. 3, pp. 501–514, Mar. 1993.

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

[14] M. T. Akhtar, M. Abe, M. Kawamata, and A. Nishihara, "Online secondary path modeling in multichannel active noise control systems using variable step size," *Signal Processing*, vol. 88, no. 8, pp. 2019–2029, Aug. 2008.

[15] Y. Huang, J. Benesty, and J. Chen, "Identification of acoustic MIMO systems: challenges and opportunities," *Signal Processing*, vol. 86, no. 6, pp. 1278–1295, June 2006.

[16] 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. Am.*, vol. 140, no. 6, pp. 510–516, Dec. 2016.

[17] W. Zhang, C. Hofmann, M. Buerger, T. D. Abhayapala, and W. Kellermann, "Spatial noise-field control with online secondary path modeling: a wave-domain approach," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 26, no. 12, pp. 2355–2370, Dec. 2018.

[18] F. R. Stöter, A. Liutkus, and N. Ito, "The 2018 signal separation evaluation campaign," *International Conference on Latent Variable Analysis and Signal Separation*, pp. 293–305, Apr. 2018.

[19] V. Patel, J. Cheer, and S. Fontana, "Design and implementation of an active noise control headphone with directional hear-through capability," *IEEE Trans. Consum. Electron.*, vol. 66, no. 1, pp. 32–40, Feb. 2020.

[20] B. Xu and T. Miller, "Selective active noise cancellation based on directional reference signals," *J. Acoust. Soc. Am.*, vol. 146, no. 4, pp. 2794–2794, Oct. 2019.

[21] J. Khemwong and N. Tangsangiumvisai, "A two-step adaptive noise cancellation system for dental-drill noise reduction," *Engineering Journal*, vol. 22, no. 4, pp. 51–66, July 2018.

[22] T. Xiao, X. Qiu, and B. Halkon, "Ultra-broadband active noise cancellation at the ears via optical microphones," *Sci. Rep.*, 2020.

[23] S. Hu, R. Rajamani, and X. Yu, "Directional cancellation of acoustic noise for home window applications," *Appl. Acoust.*, vol. 74, no. 3, pp. 467–477, Mar. 2013.

[24] B. Lam, S. Elliott, J. Cheer, and W. S. Gan, "The physical limits of active noise control of open windows," *Appl. Acoust.*, vol. 137, pp. 9–17, 2018.

[25] B. Lam and W. S. Gan, "Active acoustic windows: towards a quieter home," *IEEE Potentials*, vol. 35, no. 1, pp. 11–18, Jan. 2016.

[26] T. Murao, C. Shi, W. S. Gan, and M. Nishimura, "Mixed-error approach for multi-channel active noise control of open windows," *Appl. Acoust.*, vol. 127, pp. 305–315, Dec. 2017.

[27] J. He, B. Lam, D. Shi, and W. S. Gan, "Exploiting the underdetermined system in multichannel active noise control for open windows," *Appl. Acoust.*, vol. 9, no. 3, pp. 390, Jan. 2019.

[28] P. Candeloro, T. Pagliaroli, D. Ragni, and F. S. Di, "Small-scale rotor aeroacoustics for drone propulsion: a review of noise sources and control strategies," Oct. 2019.

[29] R. Rishabh, W. S. Gan, and Y. K. Chong, "Applying active noise control technique for augmented reality headphones," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, 2014, vol. 249, pp. 517–525.

[30] R. Scheibler, J. Azcarreta, R. Beuchat, and C. Ferry, "Pyramic: full stack open microphone array architecture and dataset," in *2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC)*, Sept. 2018, pp. 226–230.

[31] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2002, vol. 2, pp. 1781–1784.

[32] F. Borra, S. Krenn, I. D. Gebru, and D. Marković, "1st-order microphone array system for large area sound field recording and reconstruction: discussion and preliminary results," in *2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Oct. 2019, pp. 378–382.

[33] W. Mickiewicz, M. Raczyński, and A. Parus, "Performance analysis of cost-effective miniature microphone sound intensity 2D probe," *Sensors*, vol. 20, no. 1, Jan. 2020.

[34] 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. Am.*, vol. 138, no. 5, pp. 3081–3092, Nov. 2015.

[35] H. Zuo, P. N. Samarasinghe, T. D. Abhayapala, and G. Dickns, "Spatial sound intensity vectors in spherical harmonic domain," *J. Acoust. Soc. Am.*, vol. 145, no. 2, pp. 149–155, Feb. 2019.

[36] E. G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*, Academic Press, June 1999.

[37] J. Laroche, "Optimal constraint-based loop-shaping in the cepstral domain," *IEEE Signal Process. Lett.*, vol. 14, no. 4, pp. 225–227, Apr. 2007.

[38] L. Zhang, L. Wu, and X. Qiu, "An intuitive approach for feedback active noise controller design," *Appl. Acoust.*, vol. 74, no. 1, pp. 160–168, Jan. 2013.

[39] M. Brandner, M. Flock, and M. Schörkmaier, "Feedback loop shaping for active noise control with constraints of the worst-case secondary paths," in *6th Congress of Alps-adria Acoustics Association*, Oct. 2014.

[40] S. Liebich, C. Anemüller, P. Vary, P. Jax, D. Rüschen, and S. Leonhardt, "Active noise cancellation in headphones by digital robust feedback control," in *2016 24th European Signal Processing Conference (EUSIPCO)*, Aug. 2016, pp. 1843–1847.

[41] M. R. Bai, W. Pan, and H. Chen, "Active feedforward noise control and signal tracking of headsets: electroacoustic analysis and system implementation," *J. Acoust. Soc. Am.*, vol. 143, no. 3, pp. 1613–1622, Mar. 2018.

[42] J. Fabry, F. Konig, S. Liebich, and P. Jax, "Acoustic equalization for headphones using a fixed feed-forward filter," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 980–984.

[43] S. Liebich, J. Fabry, P. Jax, and P. Vary, "Signal processing challenges for active noise cancellation headphones," in *Speech Communication; 13th ITG-Symposium*, Oct. 2018, pp. 1–5.

[44] Y. Xiao, A. Ikuta, L. Ma, and K. Khorasani, "Stochastic analysis of the FxLMS-based narrowband active noise control system," *IEEE Trans. Audio Speech Lang. Process.*, vol. 16, no. 5, pp. 1000–1014, July 2008.

[45] M. T. Akhtar, M. Abe, and M. Kawamata, "A new variable step size LMS algorithm-based method for improved online secondary path modeling in ac-

tive noise control systems," *IEEE Trans. Audio Speech Lang. Process.*, vol. 14, no. 2, pp. 720–726, Mar. 2006.

[46] A. Carini and S. Malatini, "Optimal variable step size NLMS algorithms with auxiliary noise power scheduling for feedforward active noise control," *IEEE Trans. Audio Speech Lang. Process.*, vol. 16, no. 8, pp. 1383–1395, Nov. 2008.

[47] A. H. Sayed, *Fundamentals of adaptive filtering*, Wiley-IEEE Press, May 2003.

[48] Hwang K. and Song W., "An affine projection adaptive filtering algorithm with selective regressors," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 54, no. 1, pp. 43–46, Jan. 2007.

[49] Duttweiler D. L., "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. on Speech and Audio Process.*, vol. 8, no. 5, pp. 508–518, Sept. 2000.

[50] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2002, vol. 2, pp. 1881–1884.

[51] Li G., Masahiro F., and Sachio S., "An improved mu-law proportionate NLMS algorithm," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2008, pp. 3797–3800.

[52] S. Elliott, *Signal processing for active control*, Elsevier, Sept. 2000.

[53] 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, Aug. 2012.

[54] C. Hansen, S. Snyder, X. Qiu, L. Brooks, and D. Moreau, *Active control of noise and vibration*, CRC Press, Nov. 2012.

[55] 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.

[56] P. A. Nelson and S. J. Elliott, *Active control of sound*, Academic Press, Aug. 1991.

[57] A. Montazeri, J. Poshtan, and M. H. Kahaei, "Analysis of the global reduction of broadband noise in a telephone kiosk using a MIMO modal ANC system," *Int. J. Eng. Sci.*, vol. 45, no. 2, pp. 679–697, Feb. 2007.

[58] P. A. Nelson, A. R. D. Curtis, S. J. Elliott, and A. J. Bullmore, "The active minimization of harmonic enclosed sound fields, part I: theory," *J. Sound Vib.*, vol. 117, no. 1, pp. 1–13, Aug. 1987.

[59] A. J. Bullmore, P. A. Nelson, A. R. D. Curtis, and S. J. Elliott, "The active minimization of harmonic enclosed sound fields, part II: a computer simulation," *J. Sound Vib.*, vol. 117, no. 1, pp. 15–33, Aug. 1987.

[60] S. J. Elliott, A. R. D. Curtis, A. J. Bullmore, and P. A. Nelson, "The active minimization of harmonic enclosed sound fields, part III: experimental verification," *J. Sound Vib.*, vol. 117, no. 1, pp. 35–58, Aug. 1987.

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

[62] S. J. Elliott and P. A. Nelson, "Multiple-point equalization in a room using adaptive digital filters," *J. Audio Eng. Soc.*, vol. 37, no. 11, pp. 899–907, 1989.

[63] P. Joseph, S. J. Elliott, and P. A. Nelson, "Statistical aspects of active control in harmonic enclosed sound fields," *J. Sound Vib.*, vol. 172, no. 5, pp. 629–655, May 1994.

[64] W. K. Tseng, B Rafaely, and S. J. Elliott, "Combined feedback–feedforward active control of sound in a room," *J. Acoust. Soc. Am.*, vol. 104, no. 6, pp. 3417–3425, Dec. 1998.

[65] S. J. Elliott, I. M. Stothers, P. A. Nelson, A. M. McDonald, D. C. Quinn, and T. Saunders, "The active control of engine noise inside cars," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, 1988, vol. 1988, pp. 987–990.

[66] S. J. Elliott, P. A. Nelson, T. Sutton, A. M. McDonald, D. C. Quinn, I. M. Stothers, and I. Moore, "The active control of low frequency engine and road noise inside automotive interiors," *American Society of Mechanical Engineers, Noise Control and Acoustics Division (Publication) NCA*, vol. 8, pp. 125–129, 1990.

[67] J. Cheer and S. J. Elliott, "The design and performance of feedback controllers for the attenuation of road noise in vehicles," *Int. J. Acoust. Vibr.*, vol. 19, no. 3, pp. 155–164, Sept. 2014.

[68] A. Montazeri and C. J. Taylor, "Modeling and analysis of secondary sources coupling for active sound field reduction in confined spaces," *Mech. Syst. Signal Process.*, vol. 95, pp. 286–309, Oct. 2017.

[69] S. D. Sommerfeldt, J. W. Parkins, and J. Tichy, "Multi-channel active control of acoustic energy density," *J. Acoust. Soc. Am.*, vol. 101, no. 5, pp. 3122–3122, May 1997.

[70] H. C. Geng, Z. S. Rao, and Z. S. Han, "New modeling method and mechanism analyses for active control of interior noise in an irregular enclosure using piezoelectric actuators," *J. Acoust. Soc. Am.*, vol. 113, no. 3, pp. 1439–1447, Mar. 2003.

[71] M. Poletti, "Unified description of ambisonics using real and complex spherical harmonics," in *Proc. Ambisonics Symp.*, 2009.

[72] F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, "*NIST Digital Library of Mathematical Functions*," http://dlmf.nist.gov/, Release 1.0.25 of 2019-12-15.

[73] D. B. Ward and T. D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Trans. Audio Speech Lang. Process.*, vol. 9, no. 6, pp. 697–707, Sept. 2001.

[74] P. Archontis, "Real/complex spherical harmonic transform, gaunt coefficients and rotations," *GitHub.*, 2020.

[75] B. Rafaely, *Fundamentals of spherical array processing*, vol. 8, Springer, 2015.

[76] A. J. Kempton, "The ambiguity of acoustic sources–a possibility for active control?," *J. Sound Vib.*, vol. 48, no. 4, pp. 475–483, Oct. 1976.

[77] T. Martin and A. Roure, "Optimization of an active noise control system using spherical harmonics expansion of the primary field," *J. Sound Vib.*, vol. 201, no. 5, pp. 577–593, Apr. 1997.

[78] B. Rafaely, "Spherical loudspeaker array for local active control of sound," *J. Acoust. Soc. Am.*, vol. 125, no. 5, pp. 3006–3017, Mar. 2009.

[79] I. T. Ardekani and W. H. Abdulla, "Active noise control in three dimensions," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 6, pp. 2150–2159, Nov. 2014.

[80] M. Buerger, T. D. Abhayapala, C. Hofmann, H. Chen, and W. Kellermann, "The spatial coherence of noise fields evoked by continuous source distributions," *J. Acoust. Soc. Am.*, vol. 142, no. 5, pp. 3025–3034, Nov. 2017.

[81] H. Chen, J. Zhang, P. N. Samarasinghe, and T. D. Abhayapala, "Evaluation of spatial active noise cancellation performance using spherical harmonic analysis," in *2016 IEEE International Workshop on Acoustic Signal Enhancement (IWAENC)*, Sept. 2016, pp. 1–5.

[82] 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 and Lang. Proc.*, vol. 26, no. 4, pp. 774–786, Apr. 2018.

[83] J. Xie, D. Jin, W. Zhang, X. Zhang, J. Chen, and D. Wang, "Robust sparse multichannel active noise control," in *2019 IEEE International Conference*

on Acoustics, Speech and Signal Processing (ICASSP), May 2019, pp. 521–525.

[84] S. Spors and H. Buchner, "Efficient massive multichannel active noise control using wave-domain adaptive filtering," in *2008 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP)*, Mar. 2008, pp. 1480–1485.

[85] S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," in *In 124th Convention of the AES*, 2008.

[86] A. Kuntz and R. Rabenstein, "An approach to global noise control by wave field synthesis," in *2004 12th European Signal Processing Conference*, Sept. 2004, pp. 1999–2002.

[87] S. Spors and H. Buchner, "An approach to massive multichannel broadband feedforward active noise control using wave-domain adaptive filtering," in *2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Oct. 2007, pp. 171–174.

[88] A. Lapini, M. Biagini, F. Borchi, M. Carfagni, and F. Argenti, "Active noise control for pulse signals by wave field synthesis," in *2016 24th European Signal Processing Conference (EUSIPCO)*, Aug. 2016, pp. 883–887.

[89] A. Lapini, F. Borchi, M. Carfagni, and F. Argenti, "Application of wave field synthesis to active control of highly non-stationary noise," *Appl. Acoust.*, vol. 131, pp. 220–229, Feb. 2018.

[90] Y. Maeno, Y. Mitsufuji, and T. D. Abhayapala, "Mode domain spatial active noise control using sparse signal representation," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2018, pp. 211–215.

[91] Y. Maeno, Y. Mitsufuji, P. N. Samarasinghe, and T. D. Abhayapala, "Mode-domain spatial active noise control using multiple circular arrays," in *2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC)*, Sept. 2018, pp. 441–445.

[92] H. Ito, S. Koyama, N. Ueno, and H. Saruwatari, "Feedforward spatial active noise control based on kernel interpolation of sound field," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 511–515.

[93] A. Fahim, P. N. Samarasinghe, and T. D. Abhayapala, "Sound field separation in a mixed acoustic environment using a sparse array of higher order spherical microphones," in *2017 Hands-free Speech Communications and Microphone Arrays (HSCMA)*, Mar. 2017, pp. 151–155.

[94] T. D. Abhayapala and Others, "Modal analysis and synthesis of broadband nearfield beamforming arrays," *The Australian National University*, 1999.

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

[96] M. Melon, C. Langrenne, P. Herzog, and A. Garcia, "Evaluation of a method for the measurement of subwoofers in usual rooms," *J. Acoust. Soc. Am.*, vol. 127, no. 1, pp. 256–263, Jan. 2010.

[97] T. D. Abhayapala, T. S. Pollock, and R. A. Kennedy, "Characterization of 3D spatial wireless channels," in *2003 IEEE 58th Vehicular Technology Conference (VTC)*, Oct. 2003, vol. 1, pp. 123–127.

[98] 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, June 2007.

[99] N. A. Gumerov and R. Duraiswami, *Fast multipole methods for the helmholtz equation in three dimensions*, Elsevier, Jan. 2005.

[100] W. Zhang, T. D. Abhayapala, R. A. Kennedy, and R. Duraiswami, "Insights into head-related transfer function: spatial dimensionality and continuous representation," *J. Acoust. Soc. Am.*, vol. 127, no. 4, pp. 2347–2357, Apr. 2010.

[101] X. Qiu, C. H. Hansen, and X. Li, "A comparison of near-field acoustic error sensing strategies for the active control of harmonic free field sound radiation," *J. Sound Vib.*, vol. 215, no. 1, pp. 81–103, Aug. 1998.

[102] N. Han and X. Qiu, "A study of sound intensity control for active noise barriers," *Appl. Acoust.*, vol. 68, no. 10, pp. 1297–1306, Oct. 2007.

[103] J. Zhang, T. D. Abhayapala, W. Zhang, and P. N. Samarasinghe, "Active noise control over space: a subspace method for performance analysis," *NATO Adv. Sci. Inst. Ser. E Appl. Sci.*, vol. 9, no. 6, pp. 1250, Mar. 2019.

[104] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943–950, Apr. 1979.

[105] G. Weinreich and E. B. Arnold, "Method for measuring acoustic radiation fields," *J. Acoust. Soc. Am.*, vol. 68, no. 2, pp. 404–411, Aug. 1980.

[106] Y. Braikia, M. Melon, C. Langrenne, E. Bavu, and A. Garcia, "Evaluation of a separation method for source identification in small spaces," *J. Acoust. Soc. Am.*, vol. 134, no. 1, pp. 323–331, July 2013.

[107] S. Lobréau, E. Bavu, and M. Melon, "Hemispherical double-layer time reversal imaging in reverberant and noisy environments at audible frequencies," *J. Acoust. Soc. Am.*, vol. 137, no. 2, pp. 785–796, Feb. 2015.

[108] M. Tamura, "Spatial fourier transform method of measuring reflection coefficients at oblique incidence: theory and numerical examples," *J. Acoust. Soc. Am.*, vol. 88, no. 5, pp. 2259–2264, Nov. 1990.

[109] F. Jacobsen and V. Jaud, "Statistically optimized near field acoustic holography using an array of pressure-velocity probes," *J. Acoust. Soc. Am.*, vol. 121, no. 3, pp. 1550–1558, Mar. 2007.

[110] C. Langrenne, M. Melon, and A. Garcia, "Boundary element method for the acoustic characterization of a machine in bounded noisy environment," *J. Acoust. Soc. Am.*, vol. 121, no. 5, pp. 2750–2757, May 2007.

[111] C. Langrenne, M. Melon, and A. Garcia, "Measurement of confined acoustic sources using near-field acoustic holography," *J. Acoust. Soc. Am.*, vol. 126, no. 3, pp. 1250–1256, Sept. 2009.

[112] C. Bi, X. Chen, and J. Chen, "Sound field separation technique based on equivalent source method and its application in nearfield acoustic holography," *J. Acoust. Soc. Am.*, vol. 123, no. 3, pp. 1472–1478, Mar. 2008.

[113] C. Bi and B. J. Stuart, "An equivalent source technique for recovering the free sound field in a noisy environment," *J. Acoust. Soc. Am.*, vol. 131, no. 2, pp. 1260–1270, Feb. 2012.

[114] E. Fernandez-Grande, F. Jacobsen, and Q. Leclère, "Sound field separation with sound pressure and particle velocity measurements," *J. Acoust. Soc. Am.*, vol. 132, no. 6, pp. 3818–3825, Dec. 2012.

[115] X. Zhang, J. H. Thomas, C. Bi, and J. C. Pascal, "Separation of nonstationary sound fields in the time-wavenumber domain," *J. Acoust. Soc. Am.*, vol. 131, no. 3, pp. 2180–2189, Mar. 2012.

[116] X. Zhang, J. H. Thomas, C. Bi, and J. C. Pascal, "Reconstruction of nonstationary sound fields based on the time domain plane wave superposition method," *J. Acoust. Soc. Am.*, vol. 132, no. 4, pp. 2427–2436, Oct. 2012.

[117] C. Bi, L. Geng, and X. Zhang, "Real-time separation of non-stationary sound fields with pressure and particle acceleration measurements," *J. Acoust. Soc. Am.*, vol. 135, no. 6, pp. 3474–3482, June 2014.

[118] C. Bi, L. Geng, and X. Zhang, "Separation of non-stationary multisource sound field based on the interpolated time-domain equivalent source method," *Mech. Syst. Signal Process.*, vol. 72, pp. 745–761, May 2016.

[119] C. Bi, L. Geng, and X. Zhang, "Separation of non-stationary sound fields with single layer pressure-velocity measurements," *J. Acoust. Soc. Am.*, vol. 139, no. 2, pp. 781–789, Feb. 2016.

[120] B. De Hans-Elias, "The Microflown: an acoustic particle velocity sensor," *Acoustics Australia*, vol. 31, no. 3, pp. 91–94, Dec. 2003.

[121] J. H. Conway, N. Sloane, and James A., *Sphere packings, lattices and groups*, vol. 290, Springer Science & Business Media, Jan. 2013.

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

[123] E. Fisher and B. Rafaely, "Near-field spherical microphone array processing with radial filtering," *IEEE Trans. Audio Speech Lang. Process.*, vol. 19, no. 2, pp. 256–265, Feb. 2011.

[124] F. Ma, W. Zhang, and T. D. Abhayapala, "Real-time separation of non-stationary sound fields on spheres," *J. Acoust. Soc. Am.*, vol. 146, no. 1, pp. 11–21, July 2019.

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

[126] S. Berge, "Acoustically hard 2D arrays for 3D HOA," in *2019 AES International Conference on Immersive and Interactive Audio*, Mar. 2019.

[127] T. Okamoto, "3D localized sound zone generation with a planar omni-directional loudspeaker array," in *2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Oct. 2019, pp. 110–114.

[128] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals and Systems*, New Jersey: Prentice Hall, 1997.

[129] J. Cao, J. Liu, J. Wang, and X. Lai, "Acoustic vector sensor: reviews and future perspectives," *IET Signal Proc.*, vol. 11, no. 1, pp. 1–9, 2017.

[130] T. Sondergaard and M. Wille, "DOA with first order spherical acoustic vector sensor," *J. Acoust. Soc. Am.*, vol. 138, no. 5, pp. 2820–2828, Nov. 2015.

[131] X. Zhong and A. B. Premkumar, "Particle filtering approaches for multiple acoustic source detection and 2D direction of arrival estimation using a single

acoustic vector sensor," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4719–4733, Sept. 2012.

[132] S. Hafezi, A. H. Moore, and P. A. Naylor, "Augmented intensity vectors for direction of arrival estimation in the spherical harmonic domain," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 25, no. 10, pp. 1956–1968, Oct. 2017.

[133] S. F. Wu, "Transient sound radiation from impulsively accelerated bodies," *J. Acoust. Soc. Am.*, vol. 94, no. 1, pp. 542–553, July 1993.

[134] T. B. Hansen, "Spherical expansions of time-domain acoustic fields: application to near-field scanning," *J. Acoust. Soc. Am.*, vol. 98, no. 2, pp. 1204–1215, Aug. 1995.

[135] V. R. Algazi and R. O. Duda, "Headphone-based spatial sound," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 33–42, Jan. 2011.

[136] Z. Li and R. Duraiswami, "Fast time-domain spherical microphone array beamforming," in *2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Oct. 2007, pp. 155–158.

[137] S. Yan, H. Sun, X. Ma, U. P. Svensson, and C. Hou, "Time-domain implementation of broadband beamformer in spherical harmonics domain," *IEEE Trans. Audio Speech Lang. Process.*, vol. 19, no. 5, pp. 1221–1230, July 2011.

[138] J. Benesty, I. Cohen, and J. Chen, "Beamforming in the time domain," in *Fundamentals of Signal Enhancement and Array Signal Processing*, pp. 397–443. Wiley-IEEE Press, 2018.

[139] A. H. Moore, C. Evers, and P. A. Naylor, "Direction of arrival estimation in the spherical harmonic domain using subspace pseudointensity vectors," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 25, no. 1, pp. 178–192, Jan. 2017.

[140] K. Wu, V. G. Reju, and A. W. H. Khong, "Multisource DOA estimation in a reverberant environment using a single acoustic vector sensor," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 26, no. 10, pp. 1848–1859, Oct. 2018.

[141] O. M. Buyukdura and S. S. Koc, "Two alternative expressions for the spherical wave expansion of the time domain scalar free-space green's function and an application: scattering by a soft sphere," *J. Acoust. Soc. Am.*, vol. 101, no. 1, pp. 87–91, Jan. 1997.

[142] J. Li, D. Dault, and B. Shanker, "A quasianalytical time domain solution for scattering from a homogeneous sphere," *J. Acoust. Soc. Am.*, vol. 135, no. 4, pp. 1676–1685, Apr. 2014.

[143] P. A. Martin, "Acoustic scattering by a sphere in the time domain," *Wave Motion*, vol. 67, pp. 68–80, Dec. 2016.

[144] S. Gannot, E Vincent, S. Markovich-Golan, and A. Ozerov, "A consolidated perspective on multimicrophone speech enhancement and source separation," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 25, no. 4, pp. 692–730, Apr. 2017.

[145] S. Uhlich, M. Porcu, F. Giron, M. Enenkl, T. Kemp, N. Takahashi, and Y. Mitsufuji, "Improving music source separation based on deep neural networks through data augmentation and network blending," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2017, pp. 261–265.

[146] D Wang and J. Chen, "Supervised speech separation based on deep learning: an overview," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 26, no. 10, pp. 1702–1726, Oct. 2018.

[147] Y. Luo and N. Mesgarani, "Conv-tasnet: surpassing ideal time-frequency magnitude masking for speech separation," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 27, no. 8, pp. 1256–1266, May 2019.