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

Fei Ma,[1, a)] Wen Zhang,[2, b] and Thushara D. Abhayapala[1]

[1]*College of Engineering and Computer Science,The Australian National University,*

Canberra, Australian Capital Territory 0200, Australia

[2]*Center of Intelligent Acoustics and Immersive Communications,*

Northwestern Polytechnical University, Xi'an 710072, Shaanxi,

China

1   The sound field separation methods can separate the target field from the interfering

2   noises, facilitating the study of the acoustic characteristics of the target source, which

3   is placed in a noisy environment. However, most of the existing sound field separa-

4   tion methods are derived in the frequency-domain, thus are best suited for separating

5   stationary sound fields. In this paper, a time-domain sound field separation method

6   is developed that can separate the non-stationary sound field generated by the target

7   source over a sphere in real-time. A spherical array sets up a boundary between the

8   target source and the interfering sources, such that the outgoing field on the array

9   is only generated by the target source. The proposed method decomposes the pres-

10 sure and the radial particle velocity measured by the array into spherical harmonic

11 coefficients, and recoveries the target outgoing field based on the time-domain rela-

12 tionship between the decomposition coefficients and the theoretically derived spatial

13 filter responses. Simulations show the proposed method can separate non-stationary

14 sound fields both in free field and room environments, and over a longer duration

15 with small errors. The proposed method could serve as a foundation for developing

16 future time-domain spatial sound field manipulation algorithms.

[a] Author to whom correspondence should be addressed. Electronic mail: fei.ma@ieee.org

[b] Also at : College of Engineering and Computer Science,The Australian National University, Canberra,

Australian Capital Territory 0200, Australia

I.   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 sound 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 sound fields. The disturbing sound fields and the room reflected sound fields will interfere with the target sound 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 characteristics of the sound sources, it is possible to separate the target sound field from the interfering sound fields using a sensor array. Thus we can study the target source through analyzing the separated target sound field.

Over the years, a number of sound field separation (SFS) methods have been reported in the literature. The spherical wave expansion based method uses the spherical harmonic modes as the basic modeling functions of a sound field, and are capable of separating the outgoing and the incoming field on a spherical sensor array.[1–5] The Spatial Fourier Transform based method uses the two-dimensional Fourier Transform to decompose the sound field into plane-wave components, and can separates the incident and the reflected sound field on a planar sensor array.[6] The statistically optimized near field holography method is also capable of performing SFS on a planar sensor array, and can mitigate the spatial window leakage problem of the Spatial Fourier Transform based method.[7] The recently developed

38 boundary element based method[8,9] and the equivalent source based method[10–12] extend the

39 application of SFS to arbitrary shape sensor arrays. Further, these two methods can separate

40 the scattering from the target source surface and recover the free-field radiation of the target

41 source.

42 The above mentioned SFS methods can separate the sound fields incoming from two

43 sides of a sensor array (or two sensor arrays) apart. However, they are all derived in the

44 frequency-domain. That is, they first accumulate and transfer a frame of the time-domain

45 pressure (or particle velocity) measurements into the time-frequency domain using the Short

46 Time Fourier Transform. They then conduct the SFS at each time-frequency bin. The

47 Short Time Fourier Transform process inevitably introduces the spectrum leakage problem

48 and frame-delays into the SFS methods. The frame-delay is acceptable if the target sound

49 field is stationary (or quasi-stationary). Nonetheless, when dealing with fast changing non-

50 stationary sound field, the frame-delay will make the SFS methods unable to recover the

51 target field in real-time. That further makes the separated target field unable to be used in

52 time-critical applications, such as active noise control,[13] real-time beamforming, and machine

53 anomaly diagnosis.[8] A time-domain SFS method, on the other head, can track the changes

54 of the target sound field without introducing the frame latency issue.

55 Manipulations of the sound fields over spatial regions in the time-domain are difficult,

56 and time-domain SFS methods are seldom developed. The only existing time-domain SFS

57 methods are developed by Bi and his coauthors.[14–17] Based on the Spatial Fourier Transform,

58 the method they developed can separate the sound field coming from two sides of a planar

59 sensor array.[14,15,17] Based on the interpolated time-domain equivalent source method,[16] they

[60] further developed a method that can separate the sound generated by a particular source in a

[61] multiple non-stationary sources scene. However, these time-domain methods are best suited

[62] for SFS in free field (or semi-anechoic rooms). Because first, in room environments, due to

[63] wall reflections the interfering sounds and the target sound can arrive at a planar sensor

[64] array in the same directions,[18,19] such that the target sound and the interfering noise become

[65] indistinguishable. Second, the equivalent source based method requires perfect knowledge of

[66] the location and geometry information of all the sources. However, in a room environment,

[67] the image source effect exists, and it is difficult to characterize the image source especially

[68] for an arbitrary shape room with irregular boundaries.[18,19]

[69] In this paper, based on the spherical wave expansion, we develop a time-domain SFS

[70] method that can separate the non-stationary outgoing and incoming sound fields on a

[71] sphere in both free field and room environments. We first decompose the pressure and

[72] the radial particle velocity into corresponding spherical harmonics coefficients. The time-

[73] domain convolution between the spherical harmonic coefficients and the derived impulse

[74] response functions results in the outgoing and the incoming field on the array. The perfor-

[75] mance of the proposed method is confirmed by simulations, and compared with the Spatial

[76] Fourier Transform based method. Possible applications of the proposed method include (1)

[77] reference signal generation for active noise control system,[13] (2) real-time beamforming, (3)

[78] de-reverberation for speech recognition and sound field reproduction system, (4) machine

[79] working condition monitoring in a noisy environment.

[80] This paper is organized as follows. We introduce the problem formulation and the

[81] frequency-domain SFS method based on spherical harmonic expansion in Sec. II. In Sec. III,

82 we provide the theoretical derivation of the time-domain SFS method. Sec. IV introduces

83 the practical implementations of the proposed method, whose effectiveness is validated by

84 simulations in Sec. V, and Sec. VI concludes this paper.

II. SYSTEM MODEL

A. Problem formulation

Consider the systems shown in Fig. 1, where in Fig. 1 (a) the target sources (denoted as $\triangle$) are placed in free field, and in Fig. 1 (b) the target sources (denoted as $\triangle$) are placed inside a room. We use (x, y, z) and (r, θ, ϕ) to denote the Cartesian and spherical coordinates of a point with respect to the point O, respectively. The sound field $p(t, R, \theta, \phi)$ on the sphere $\mathbb{S}_2$ of radius R is the superposition of the outgoing field $p^{\mathrm{o}}(t, R, \theta, \phi)$ and the incoming field $p^{\mathrm{i}}(t, R, \theta, \phi)$

$$p(t, R, \theta, \phi) = p^{\mathrm{o}}(t, R, \theta, \phi) + p^{\mathrm{i}}(t, R, \theta, \phi), \tag{1}$$

where t is the continuous time. The frequency-domain representation of (1) is

$$P(\omega, R, \theta, \phi) = P^{\mathrm{o}}(\omega, R, \theta, \phi) + P^{\mathrm{i}}(\omega, R, \theta, \phi), \tag{2}$$

87 where $\omega = 2\pi f$ is the angular frequency (f is the frequency). In the paper, the temporal

88 reference $e^{i\omega t}$ is implicitly assumed.

89 As shown in Fig. 1, the incoming field $p^{\mathrm{i}}(t, R, \theta, \phi)$ on the sphere $\mathbb{S}_2$ is due to the external

90 sources or room reflections. The outgoing field $p^{\mathrm{o}}(t, R, \theta, \phi)$ on the sphere $\mathbb{S}_2$, on the other

91 hand, is generated by the target sources, thus characterizes the target sources. In this paper,

92 we aim to separate the outgoing field $p^{\mathrm{o}}(t, R, \theta, \phi)$ on the sphere $\mathbb{S}_2$ to facilitate the study of

93 the target sources.

94   In this work, we assume the target sources are point sources, which do not generate the

95 scattering sound field. For real world target source of finite size, the source's surface can

96 scatter the incoming field and transform it into part of the outgoing field. To recover the

97 outgoing field generated only by the target source, we need to further separate the scattered

98 sound field using prior knowledge, such as the incoming field direction and the target source

99 geometry. Such investigations are beyond the scope of the paper and will be our future

100 work.

101   **B.   The frequency-domain sound field separation on a sphere**

    In the frequency-domain, the sound field $P(\omega, R, \theta, \phi)$ on a sphere $\mathbb{S}_2$ of radius R consists

of the outgoing field $P^{\mathrm{o}}(\omega, R, \theta, \phi)$ and the incoming field $P^{\mathrm{i}}(\omega, R, \theta, \phi)$[1]

$$
\begin{aligned}
P(\omega, R, \theta, \phi) &= P^{\mathrm{o}}(\omega, R, \theta, \phi) + P^{\mathrm{i}}(\omega, R, \theta, \phi) \\
&= \sum_{\mu=0}^{\infty} \sum_{\nu=-\mu}^{\mu} A_{\mu\nu}(\omega, R) Y_{\mu\nu}(\theta, \phi), \\
&= \sum_{\mu=0}^{\infty} \sum_{\nu=-\mu}^{\mu} [\underbrace{L_{\mu\nu}(\omega) h_{\mu}(\omega R/c)}_{A_{\mu\nu}^{\mathrm{o}}(\omega, R)} + \underbrace{K_{\mu\nu}(\omega) j_{\mu}(\omega R/c)}_{A_{\mu\nu}^{\mathrm{i}}(\omega, R)}] Y_{\mu\nu}(\theta, \phi), \quad (3)
\end{aligned}
$$

where c is the speed of sound, $A_{\mu\nu}(\omega, R)$ are spherical harmonic coefficients correspond

the sound field $P(\omega, R, \theta, \phi)$ measured on the sphere $\mathbb{S}_2$, $A_{\mu\nu}^{\mathrm{o}}(\omega, R)$ and $A_{\mu\nu}^{\mathrm{i}}(\omega, R)$ are the

outgoing and incoming field coefficients, respectively. $h_{\mu}(\cdot)$ is the second kind spherical

Hankel function of order u, $j_{\mu}(\cdot)$ is the first kind spherical Bessel function of order u, $L_{\mu\nu}(\cdot)$

and $K_{\mu\nu}(\cdot)$ are the (radial-independent) spherical Hankel and Bessel function coefficients,

respectively, $Y_{\mu\nu}(\cdot, \cdot)$ is the real spherical harmonic of order n and degree m,[20] i.e.,

$$Y_{\mu\nu}(\theta, \phi) \equiv \sqrt{\frac{(2\mu+1)(\mu-|\nu|)!}{4\pi(\mu+|\nu|)!}} \mathcal{P}_{\mu|\nu|}(\cos(\theta))$$

$$\times \begin{cases} \sqrt{2}\cos(|\nu|\phi), & \nu > 0, \\[2mm] 1, & \nu = 0, \\[2mm] \sqrt{2}\sin(|\nu|\phi), & \nu < 0, \end{cases} \tag{4}$$

$\mathcal{P}_{\mu|\nu|}(\cdot)$ is the associated Legendre function of order μ and degree $|\nu|$. Hereafter, we abbreviate (θ, ϕ) as a single symbol Θ to simplify the notations.

The frequency-domain radial particle velocity on the sphere $\mathbb{S}_2$ can be similarly expressed as[1]

$$\begin{aligned} V(\omega, R, \Theta) &= \frac{i}{\rho_0 \omega} \left. \frac{\partial P(\omega, R, \Theta)}{\partial r} \right|_{r=R} \\ &= \sum_{\mu=0}^{\infty} \sum_{\nu=-\mu}^{\mu} B_{\mu\nu}(\omega, R) Y_{\mu\nu}(\Theta) \\ &= \frac{i}{\rho_0 c} \sum_{\mu=0}^{\infty} \sum_{\nu=-\mu}^{\mu} [L_{\mu\nu}(\omega) h_\mu'(\omega R/c) + K_{\mu\nu}(\omega) j_\mu'(\omega R/c)] Y_{\mu\nu}(\Theta), \end{aligned} \tag{5}$$

where $B_{\mu\nu}(\omega, R)$ are spherical harmonic coefficients corresponding to the radial particle velocity $V(\omega, R, \Theta)$ measured on the sphere $\mathbb{S}_2$, $h_\mu'(\cdot)$ and $j_\mu'(\cdot)$ are derivatives of $h_\mu(\cdot)$ and $j_\mu(\cdot)$ about the argument, respectively, i is the unit imaginary number, and ρ_0 is the density of air.

Based on (3) and (5), we obtain the outgoing and incoming field coefficients as[1]

$$A_{\mu\nu}^{o}(\omega, R) = -i(\omega R/c)^2 j_\mu'(\omega R/c) h_\mu(\omega R/c) A_{\mu\nu}(\omega, R)$$

$$+\rho_0 c(\omega R/c)^2 j_\mu(\omega R/c) h_\mu(\omega R/c) B_{\mu\nu}(\omega, R), \tag{6}$$

$$A_{\mu\nu}^{\mathrm{i}}(\omega, R) = i(\omega R/c)^2 j_\mu(\omega R/c) h'_\mu(\omega R/c) A_{\mu\nu}(\omega, R)$$

$$-\rho_0 c(\omega R/c)^2 j_\mu(\omega R/c) h_\mu(\omega R/c) B_{\mu\nu}(\omega, R), \tag{7}$$

respectively. Substitution of the coefficients $A_{\mu\nu}^{\mathrm{o}}(\omega, R)$ and $A_{\mu\nu}^{\mathrm{i}}(\omega, R)$ back into (3) results in the frequency-domain 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 effectiveness of the frequency-domain SFS method based on the spherical wave expansion has been validated by both simulations and experiments.[3,4,21]

The outgoing field separated by the frequency-domain method reveals the characteristics of the target source. However, the frequency-domain method needs the Short Time Fourier Transform to transfer the time-domain acoustic quantities into the time-frequency domain. The inherent frame-delay of the Short Time Fourier Transform process makes the separated outgoing field not suitable to use in time-critical applications, such as active noise control and real-time beamforming.[13] A time-domain method, which does not introduce the frame-delay, is more appropriate to separate the non-stationary outgoing field for applications which have strict low-latency requirements.

III. THE TIME-DOMAIN SOUND FIELD SEPARATION 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 and the incoming field on a sphere. The detailed derivation of the theorem is in the appendix.

9

Theorem 1. *The time-domain outgoing field $p^{\mathrm{o}}(t, R, \Theta)$ and the incoming field $p^{\mathrm{i}}(t, R, \Theta)$*

on a sphere $\mathbb{S}_2$ of radius R are

$$p^{\mathrm{o}}(t, R, \Theta) = \sum_{\mu=0}^{\infty} \sum_{\nu=-\mu}^{\mu} a_{\mu\nu}^{\mathrm{o}}(t, R) Y_{\mu\nu}(\Theta), \tag{8}$$

$$p^{\mathrm{i}}(t, R, \Theta) = \sum_{\mu=0}^{\infty} \sum_{\nu=-\mu}^{\mu} a_{\mu\nu}^{\mathrm{i}}(t, R) Y_{\mu\nu}(\Theta), \tag{9}$$

where the outgoing field coefficients $a_{\mu\nu}^{\mathrm{o}}(t, R)$ and the incoming field coefficients $a_{\mu\nu}^{\mathrm{i}}(t, R)$ are

given by

$$
\begin{aligned}
a_{\mu\nu}^{\mathrm{o}}(t, R) = \int_{\Theta} \Big\{ & \int_{0}^{2\tau_R} g_0^{\mu}(\tau) p(t - \tau, R, \Theta) d\tau \\
& + \int_{0}^{2\tau_R} g_1^{\mu}(\tau) \frac{dp(t - \tau, R, \Theta)}{dt} d\tau \\
& + \rho_0 c \int_{0}^{2\tau_R} g_2^{\mu}(\tau) \frac{dv(t - \tau, R, \Theta)}{dt} d\tau \Big\} \\
& Y_{\mu\nu}(\Theta) d\Theta,
\end{aligned}
\tag{10}
$$

$$
\begin{aligned}
a_{\mu\nu}^{\mathrm{i}}(t, R) = \int_{\Theta} \Big\{ & \int_{0}^{2\tau_R} g_3^{\mu}(\tau) p(t - \tau, R, \Theta) d\tau \\
& + \int_{0}^{2\tau_R} g_4^{\mu}(\tau) \frac{dp(t - \tau, R, \Theta)}{dt} d\tau \\
& - \rho_0 c \int_{0}^{2\tau_R} g_2^{\mu}(\tau) \frac{dv(t - \tau, R, \Theta)}{dt} d\tau \Big\} \\
& Y_{\mu\nu}(\Theta) d\Theta,
\end{aligned}
\tag{11}
$$

respectively, $\tau_R = R/c$, $p(t, R, \Theta)$ and $v(t, R, \Theta)$ are the pressure and the radial particle

velocity on the sphere at the point (R, Θ), respectively, and $\int_{\Theta} d\Theta \equiv \int_{0}^{2\pi} \int_{0}^{\pi} \sin\theta d\theta d\phi$. The

expressions of $g_0^{\mu}(t)$, $g_1^{\mu}(t)$, $g_2^{\mu}(t)$, $g_3^{\mu}(t)$, and $g_4^{\mu}(t)$ are

$$
\begin{aligned}
g_0^{\mu}(t) = \frac{\mu}{4\tau_R} \sum_{\nu=0}^{\mu} \sum_{\varsigma=0}^{\mu} & \frac{\varphi_{\nu}(\mu)\varphi_{\varsigma}(\mu)}{(\nu + \varsigma)!} \Big[(-1)^{\nu} (\frac{t}{\tau_R})^{\nu+\varsigma} \mathtt{Sign}(t) \\
& + (-1)^{\mu+1} (\frac{t - 2\tau_R}{\tau_R})^{\nu+\varsigma} \mathtt{Sign}(t - 2\tau_R) \Big],
\end{aligned}
\tag{12}
$$

$$g_1^\mu(t) = \frac{1}{4} \sum_{\nu=0}^{\mu+1} \sum_{\varsigma=0}^{\mu} \frac{\varphi_\nu(\mu+1)\varphi_\varsigma(\mu)}{(\nu+\varsigma)!} \left[(-1)^\nu (\frac{t}{\tau_R})^{\nu+\varsigma} \mathtt{Sign}(t) \right.$$
$$\left. + (-1)^\mu (\frac{t-2\tau_R}{\tau_R})^{\nu+\varsigma} \mathtt{Sign}(t-2\tau_R) \right], \qquad (13)$$

$$g_2^\mu(t) = \frac{1}{4} \sum_{\nu=0}^{\mu} \sum_{\varsigma=0}^{\mu} \frac{\varphi_\nu(\mu)\varphi_\varsigma(\mu)}{(\nu+\varsigma)!} \left[(-1)^\nu (\frac{t}{\tau_R})^{\nu+\varsigma} \mathtt{Sign}(t) \right.$$
$$\left. + (-1)^{\mu+1} (\frac{t-2\tau_R}{\tau_R})^{\nu+\varsigma} \mathtt{Sign}(t-2\tau_R) \right], \qquad (14)$$

$$g_3^\mu(t) = \frac{\mu+1}{4\tau_R} \sum_{\nu=0}^{\mu} \sum_{\varsigma=0}^{\mu} \frac{\varphi_\nu(\mu)\varphi_\varsigma(\mu)}{(\nu+\varsigma)!} \left[(-1)^\nu (\frac{t}{\tau_R})^{\nu+\varsigma} \mathtt{Sign}(t) \right.$$
$$\left. + (-1)^{\mu+1} (\frac{t-2\tau_R}{\tau_R})^{\nu+\varsigma} \mathtt{Sign}(t-2\tau_R) \right], \qquad (15)$$

$$g_4^\mu(t) = \begin{cases} \frac{1}{4} \sum_{\nu=0}^{\mu} \sum_{\varsigma=0}^{\mu-1} \frac{\varphi_\nu(\mu)\varphi_\varsigma(\mu-1)}{(\nu+\varsigma)!} \left[(-1)^\nu (\frac{t}{\tau_R})^{\nu+\varsigma} \mathrm{Sign}(t) \right. \\ \qquad \left. + (-1)^{\mu-1} (\frac{t-2\tau_R}{\tau_R})^{\nu+\varsigma} \mathrm{Sign}(t-2\tau_R) \right], \quad \mu > 0 \\ \\ \frac{1}{4}\mathrm{Sign}(t) - \frac{1}{4}\mathrm{Sign}(t-2\tau_R), \qquad\qquad\qquad \mu = 0, \end{cases}$$
$$(16)$$

125 *respectively, where* $\mathtt{Sign}(\cdot)$ *is the sign function,* $\varphi_\nu(\mu) = (\mu+\nu)!/(2^\nu \nu!(\mu-\nu)!)$ *for* $\nu =$

126 $0,1,...,\mu$, $\mu \geq 0$.[22] *Note that* $g_0^\mu(t) = \mu g_2^\mu(t)/\tau_R$ *and* $g_3^\mu(\tau_R) = (\mu+1)g_2^\mu(t)/\tau_R$.

127 We have the following remarks on the time-domain SFS method:

128 1. Equations (8), (9), (10), (11) show that the outgoing (incoming) field at a particular

129   point (R,Θ) can be described by the pressure, the pressure gradient, and the radial

130   partial velocity gradient observed over the sphere within a certain time frame.

131 2. The spatial filter functions $g_0^u(t)$, $g_1^u(t)$, $g_2^u(t)$, $g_3^u(t)$, and $g_4^u(t)$ are specially derived

132   such that they take finite values inside of the time interval $[0, 2\tau_R]$ and are zeros

outside of the interval. The spatial filters derived in this way have finite lengths and are easy to implement in practice, such as on a digital-signal-processor (DSP). The special derivation of the spatial filters makes the expressions of the outgoing and the incoming field coefficients contain the pressure.

In the previous paper,[13] the spatial filters are derived such that the expressions of the outgoing and the incoming field coefficients do not contain the pressure. However, the spatial filter derived in this way have non-zero constant values of outside of the time interval $[0, 2\tau_R]$. To use the spatial filters, we need to truncate their lengths. This truncation process unavoidably introduces additional sound field separation error.

Thus we choose to derive the spatial filters to have finite lengths in the paper.

3. Figure 2 depicts $g_0^{\mu}(t)$, $g_1^{\mu}(t)$, $g_2^{\mu}(t)$ and their estimations $\hat{g}_0^{\mu}(t)$, $\hat{g}_1^{\mu}(t)$, $\hat{g}_2^{\mu}(t)$ for $u = 0, 1, 2$, $R = 0.5$ m, $c = 343$ m/s, and $\tau_R = R/c \approx 1.458$ ms. We plot $g_0^{\mu}(t)$, $g_1^{\mu}(t)$, $g_2^{\mu}(t)$ based on the theoretical expressions (12), (13), and (14), respectively. We plot $\hat{g}_0^{\mu}(t)$, $\hat{g}_1^{\mu}(t)$, and $\hat{g}_2^{\mu}(t)$ based on the numerical calculation of (A6), (A7), and (A8), respectively. As shown in Fig. 2, the theoretically derived $g_0^{\mu}(t)$, $g_1^{\mu}(t)$, $g_2^{\mu}(t)$ agree with the numerically calculated $\hat{g}_0^{\mu}(t)$, $\hat{g}_1^{\mu}(t)$, and $\hat{g}_2^{\mu}(t)$, respectively.

4. In the paper, we take a pressure-velocity formulation. That is we recover the outgoing and the incoming field based on the pressure and the radial particle velocity measurement on a sphere.[4] This is mainly to use the Wronskian equation to simplify the expression of the frequency-domain outgoing and incoming field coefficients ((6) and (7)),[1,21] such that they do not contain the divide-operations. The expressions of the

154 corresponding time-domain outgoing and incoming field coefficients ((10) and (11))

155 then do not include any de-convolution operations. It is possible to take a pressure-

156 pressure formulation,[4] and recover the outgoing and the incoming field based on the

157 pressure measurement on two adjacent spheres. However, the resulting time-domain

158 outgoing and incoming field coefficients contain the de-convolution operations, which

159 are computational expensive to implement.

5. 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 and the expansion of the Spherical Hankel function,[22] we can derive the spatial
filter $g_0^\mu(t)$ as (the detailed derivations are not shown for brevity)

$$
\begin{aligned}
g_0^\mu(t) &= \mathcal{F}^{-1}\Big[-\mu(i\omega\tau_R)j_\mu(\omega\tau_R)h_\mu(\omega\tau_R)\Big] \\
&= \mu\frac{(-1)^\mu}{2}\Big[\frac{1}{\tau_R}\mathcal{P}_\mu(\frac{t-\tau_R}{\tau_R})\Big] + \mu\frac{(-1)^\mu}{2}\sum_{\nu=1}^{\mu}\frac{\varphi_\nu(\mu)}{(\nu-1)!}\frac{1}{\tau_R^{\nu+1}}\Big[\mathcal{P}_\mu(\frac{t-\tau_R}{\tau_R}) * t^{\nu-1}\mathrm{U}(t)\Big], (17)
\end{aligned}
$$

160 which is the accumulated convolutions between the Legendre function $\mathcal{P}_\mu(\cdot)$ and terms

161 like $t^{\nu-1}\mathrm{U}(t)$ ($\mathrm{U}(t)$ is the step function). We can also express the spatial filters $g_1^\mu(t)$,

162 $g_2^\mu(t)$, $g_3^\mu(t)$, and $g_4^\mu(t)$ in a similar form as (17) using the Legendre functions. Equa-

163 tion (17) reveals the Legendre function shapes of the spatial filters. However, the

164 convolution operation in (17) makes its evaluation not straightforward. Thus in the

165 paper we choose to derive the spatial filters using the `Sign` function, and the resulting

166 expressions of the spatial filters are easy to implement in practice, such as on a DSP.

6. The frequency-domain outgoing and incoming field coefficients $\{A_{\mu\nu}^{\mathrm{o}}(\omega, R), A_{\mu\nu}^{\mathrm{i}}(\omega, R)\}$

168 and their time-domain counterparts $\{a_{\mu\nu}^{\mathrm{o}}(t, R), a_{\mu\nu}^{\mathrm{i}}(t, R)\}$ depend on the radius R of

169 the sound field separation sphere $\mathbb{S}_2$. We can use the radial-independent spherical

170 Bessel and Hankel function coefficients $\{L_{\mu\nu}(\omega), K_{\mu\nu}(\omega)\}$ in (3) to characterize the

171 outgoing and the incoming field, and derive their corresponding time-domain counter-

172 parts $\{l_{\mu\nu}(t), k_{\mu\nu}(t)\}$. By using $\{l_{\mu\nu}(t), k_{\mu\nu}(t)\}$, we may recover the outgoing and the

173 incoming field on a sphere, and project the outgoing and the incoming field from one

174 sphere to another.[1] However, the non-robustness issue of the spherical Bessel function

175 inversion at low frequency range and the problem of spherical Bessel function zeros

176 still exist. This follow up work is out the scope of the paper and will be our future

177 work.

178 **IV. REALIZATION OF THE TIME-DOMAIN SOUND FIELD SEPARATION**

179 **METHOD**

180 The theoretically derived (8), (9), (10), and (11) of the proposed method involve con-

181 tinuous integrals, time derivatives of the sound pressure and the radial particle velocity,

182 which are acoustic quantities that can not be measured directly using existing sensors. In

183 this section, we further develop approaches to facilitate the implementation of the proposed

184 method.

A. Discretization and sampling

We first approximate the continuous integral in (10) by a discrete form as follows

$$
\begin{aligned}
a_{\mu\nu}^{\mathrm{o}}(t, R) \approx \int_{\Theta} \Big\{ &\sum_{n=0}^{T_n} g_0^{\mu}(\tau_n) p(t - \tau_n, R, \Theta) \delta_{\tau} \\
&+ \sum_{n=0}^{T_n} g_1^{\mu}(\tau_n) \frac{p(t - \tau_n, R, \Theta) - p(t - \delta_t - \tau_n, R, \Theta)}{\delta_t} \delta_{\tau} \\
&+ \rho_0 c \sum_{n=0}^{T_n} g_2^{\mu}(\tau_n) \frac{v(t - \tau_n, R, \Theta) - v(t - \delta_t - \tau_n, R, \Theta)}{\delta_t} \delta_{\tau} \Big\} \\
&Y_{\mu\nu}(\Theta) d\Theta \\
\approx \int_{\Theta} \Big\{ &\sum_{n=0}^{T_n} g_0^{\mu}(\tau_n) p(t - \tau_n, R, \Theta) \delta_t \\
&+ \sum_{n=0}^{T_n} g_1^{\mu}(\tau_n) [p(t - \tau_n, R, \Theta) - p(t - \delta_t - \tau_n, R, \Theta)] \\
&+ \rho_0 c \sum_{n=0}^{T_n} g_2^{\mu}(\tau_n) [v(t - \tau_n, R, \Theta) - v(t - \delta_t - \tau_n, R, \Theta)] \Big\} \\
&Y_{\mu\nu}(\Theta) d\Theta,
\end{aligned}
\tag{18}
$$

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, and $\lceil \cdot \rceil$ is the ceiling operator).

We further develop (18) 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 t and τ_n using the discrete time n and n'. By denoting $p(n, R, \Theta_q)$ and $v(n, R, \Theta_q)$ as the pressure and the radial partial velocity at (R, Θ_q) at discrete time instant n, we have the outgoing field coefficient $a_{\mu\nu}^{\mathrm{o}}(n, R)$

15

at discrete time n as

$$
\begin{aligned}
a_{\mu\nu}^{\mathrm{o}}(n, R) \approx \sum_{q=1}^{Q} \gamma_q Y_{\mu\nu}(\Theta_q) \sum_{n'=0}^{T_n} \Big\{ & g_0^{\mu}(n') p(n-n', R, \Theta_q) \delta_t \\
& + g_1^{\mu}(n')[p(n-n', R, \Theta_q) - p(n-1-n', R, \Theta_q)] \\
& + \rho_0 c g_2^{\mu}(n')[v(n-n', R, \Theta_q) - v(n-1-n', R, \Theta_q)] \Big\},
\end{aligned}
\tag{19}
$$

188 where $\{\gamma_q\}_{q=1}^{Q}$ are the sampling weights.[28]

189 Substitution of $a_{\mu\nu}^{\mathrm{o}}(t, R)$ by $a_{\mu\nu}^{\mathrm{o}}(n, R)$ into (8) results in an estimation of the outgoing

190 field $p^{\mathrm{o}}(n, R, \Theta)$ at discrete time instance on the sphere $\mathbb{S}_2$. We can obtain the discrete form

191 of (11) in a similar manner.

192 Due to the finite sampling over the sphere, we can only recover the outgoing field coeffi-

193 cients up to N, where $N \le \sqrt{Q} - 1$ is the truncation order of the outgoing field.[23]

194 **B. Pressure and velocity approximation**

195 We may use an array of acoustic vector sensors, which can measure both the pressure

196 and the radial particle velocity,[25] to realize the SFS method. Alternatively, we can use two

197 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[24]

$$
p(n, R, \Theta_q) \approx 0.5[p(n, R + \delta_R, \Theta_q) + p(n, R - \delta_R, \Theta_q)].
\tag{20}
$$

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[24]

$$v(n, R, \Theta_q) \approx v(n-1, R, \Theta_q)$$
$$+\frac{1}{\rho_0} \frac{p(n, R+\delta_R, \Theta_q) - p(n, R-\delta_R, \Theta_q)}{2\delta_R} \delta_t. \tag{21}$$

Substitutions of (20) and (21) into (19) result in an estimation of the time-domain outgoing field coefficient $a_{\mu\nu}^{o}(n, R)$.

V. SIMULATION EXAMPLES

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

A. Sound field separation in free field

The simulation environment is as shown in Fig. 1 (a). We place a point source at $(0, 0, 0.3)$ m as the target source, and 100 plane waves, whose directions are determined according to the 100-point spherical packing,[27] as the incoming fields. We let the target source output a band-limited sound, which is simulated by passing a unit-variance Gauss white noise through a 64-tap Bandpass Butterworth filter of the 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 compared with the contribution from the target source.

212   We aim to estimate the target (outgoing) sound field on the sphere $\mathbb{S}_2$ produced by the

213   target source. We choose the radius of the sphere $\mathbb{S}_2$ as $R = 0.65$ m, and the outgoing field

214   truncation order to be $N = 5$. We use a sensor array arranged on the sphere according

215   to the 6-th order Gauss sampling scheme to realize the time-domain SFS method.[28] The

216   reason that we use the 6-th order Gauss sampling scheme to recover the outgoing field

217   (coefficients) up to 5-th order is to reduce the effect of incoming field (coefficients) aliasing.

218   The incoming field, which has high order coefficients, contaminates the estimation of the

219   6-th order outgoing field (coefficients). Thus the 6-th order outgoing field (coefficients) is

220   not estimated. Through over-sampling on the sphere, the outgoing field coefficients of lower

221   orders ($N \leq 5$) are more accurately estimated.

222   We simulate the impulse responses, including the radial particle velocity response, be-

223   tween the source and the sensors based on the free-space Green function and the Euler's equa-

224   tion.[1] The impulse responses are truncated to 1024 taps long under the sampling frequency

225   $f_{\mathrm{s}} = 48$ kHz. The speed of sound is $c = 343$ m/s, and the air density is $\rho_0 = 1.225$ kg/m^3.

226   The length of impulse response functions $g_0^\mu(t)$, $g_1^\mu(t)$, $g_2^\mu(t)$ is $T_n = \lceil 2R/cf_{\mathrm{s}} \rceil + 1 = 183$ taps.

227   We add Gauss white noise to the sensor measurement such that the signal to noise ratio is

228   40 dB. The simulation results are from an average of 100 independent runs. The settings in

229   this paragraph are used in all simulations unless otherwise stated.

230   We reconstruct the outgoing field on the sphere over a period of 10 ms according to (8).

231   We depict the amplitudes of the outgoing field $p^{\mathrm{o}}(t, R, \Theta_{17})$, the total field $p(t, R, \Theta_{17})$, the

232   separated outgoing field $\hat{p}^{\mathrm{o}}(t, R, \Theta_{17})$, and the outgoing field separation error $p^{\mathrm{e}}(t, R, \Theta_{17}) =$

233 $p^{\mathrm{o}}(t, R, \Theta_{17}) - \hat{p}^{\mathrm{o}}(t, R, \Theta_{17})$, at the 17-th sensor in Fig. 3. The total field is the summation

234 of the outgoing field and the incoming plane wave fields.

As shown in Fig. 3, the total field differs from the outgoing field. The separated outgoing

field approximates the outgoing field, and the field separation error is small over the whole

10 ms period. The normalized separation error at the 17-th sensor over the observation

period is

$$\xi_{17} = 10\mathrm{log}10\frac{\sum_{t=0}^{10 \text{ ms}} ||p^{\mathrm{e}}(t, R, \Theta_{17})||^2}{\sum_{t=0}^{10 \text{ ms}} ||p^{\mathrm{o}}(t, R, \Theta_{17})||^2} = -30.1 \text{ dB}. \tag{22}$$

We depict the normalized sound field separation error ξ_{Θ_q} on the sphere $\mathbb{S}_2$ over the 10

ms in Fig. 4, where ξ_{Θ_q} is defined as

$$\xi_{\Theta_q} = 10\mathrm{log}10\frac{\sum_{t=0}^{10 \text{ ms}} ||p^{\mathrm{o}}(t, R, \Theta_q) - \hat{p}^{\mathrm{o}}(t, R, \Theta_q)||^2}{\sum_{t=0}^{10 \text{ ms}} ||p^{\mathrm{o}}(t, R, \Theta_q)||^2}, \tag{23}$$

235 $p^{\mathrm{o}}(t, R, \Theta_q)$ and $\hat{p}^{\mathrm{o}}(t, R, \Theta_q)$ are the outgoing field and its estimation at the point (R, Θ_q)

236 respectively, $(R, \Theta_q)_{q=1}^{180 \times 360}$ are equal-angle sampling point positions on the sphere $\mathbb{S}_2$.[28] As

237 shown in Fig. 4, the proposed method is able to recover the outgoing field accurately, and

238 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

$$\xi^N(\omega) = 10\mathrm{log}10\frac{\sum_{q=1}^{98} ||p^{\mathrm{o}}(\omega, R, \Theta_q) - \hat{p}^{\mathrm{o},N}(\omega, R, \Theta_q)||^2}{\sum_{q=1}^{98} ||p^{\mathrm{o}}(\omega, R, \Theta_q)||^2},$$

$$\tag{24}$$

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

$$\hat{p}^{\mathrm{o},N}(n, R, \Theta_q) = \sum_{\mu=0}^{N} \sum_{\nu=-\mu}^{\mu} \hat{a}_{\mu\nu}^{\mathrm{o}}(n, R) Y_{\mu\nu}(\Theta_q), \tag{25}$$

239 $\hat{a}_{\mu\nu}^{\mathrm{o}}(n, R)$ are the time-domain outgoing field coefficient obtained through (19), $\{R, \Theta_q\}_{q=1}^{98}$

240 are the sampling point positions from the 6-th order Gauss sampling scheme, and $\omega =$

241 $200\pi, 202\pi, ..., 1200\pi$ rad/s (or $f = 100, 101, ..., 600$ Hz).

242 We plot the normalized separation error $\xi^N(\omega)$ in Fig. 5, which shows that $\xi^N(\omega)$ increases

243 along with the frequency. The higher the outgoing field truncation order N is the smaller the

244 normalized separation error $\xi^N(\omega)$ is. However, in the low frequency range, the truncation

245 order N $= 5$ does not make the normalized separation error $\xi^N(\omega)$ significantly smaller than

246 that in the case of N $= 4$, i.e., $\xi^5(\omega) = \xi^4(\omega)$ for $\omega < 600\ \pi/rad$ or $f < 300$ Hz. To conclude,

247 we should choose an appropriate truncation order N for the separated target (outgoing) field

248 depending on its frequency components such that the separation error is sufficiently small.

249 **B.   Comparisons with the Spatial Fourier Transform based method**

250 In this section, we compare the performance of the proposed method with the Spatial

251 Fourier Transform based method.[15] We make a comparison between the two methods because

252 they both use some theoretically derived spatial filters for sound field separation. The

253 comparison is to show that the derivation of the spatial filters influences the sound field

254 separation accuracy.

255   The setup of the simulation environment is shown in Fig. 6, which has a Cartesian

256   coordinate system with the origin at the point O. On the $z = 0$ m plane H, we equally

257   distribute 15×15 sensors on a square of size 0.7 m × 0.7 m, whose center is at the point O.

258   On the $z = 0.01$ m plane H_1, we place another 15×15 sensors equally on a square of size

259   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

260   sensors on these two planes for SFS using the Spatial Fourier Transform based method. We

261   extend the array size to be 81×81 by zero-padding to reduce the aliasing error of the Spatial

262   Fourier Transform.[15]

263   There is a sphere $\mathbb{S}_2$ of radius $R = 0.35$ m. On the sphere $\mathbb{S}_2$, we place sensors according

264   to the 9-th order Gauss sampling scheme.[28] We use the sensors on this sphere $\mathbb{S}_2$ to estimate

265   the outgoing field up to 6-th order using the proposed method.

266   The sampling schemes used by these two methods are selected such that they can detect

267   approximately the same number of acoustic quantities, and that the sizes of the two arrays

268   are comparable.

269   We have a target point source at $(0.0, 0.0, -0.2)$ m, and another interfering point source

270   at $(0, 0, 0.2)$ m as shown in Fig. 6. Outputs of these two sources are unit-variance Gauss

271   white noises filtered a 64-tap Butterworth Bandpass filter of frequency range $[100, 1000]$

272   Hz. In the Spatial Fourier Transform based method, we truncate the Bessel function, i.e.,

273   (7) from Bi,[15] to 1024 taps long. In the proposed method, the length of impulse response

274   functions $g_0^\mu(t)$, $g_1^\mu(t)$, $g_2^\mu(t)$ is $T_n = \lceil 2R/cf_s \rceil + 1 = 99$ taps.

21

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 error $p^{\mathrm{e}}(t) = p^{\mathrm{d}}(t) - \hat{p}^{\mathrm{d}}(t)$ using the proposed and the Spatial Fourier Transform-based method are shown in Fig. 7. The target (outgoing) sound is produced by convolving the target source output and the free-space Green function. The total sound is the summation of the target (outgoing) sound and the interfering (incoming) sound. The separated target sounds $\hat{p}^{\mathrm{d}}(t)$ in the two methods are not shown for the ease of illustration.

As shown in Fig. 7, the proposed method can recover the target sound accurately over the whole 30 ms period. The Spatial Fourier Transform based method 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.[15] The truncation of the infinite long Bessel function, i.e., (7) from Bi,[15] to a finite length introduces errors. We have conducted simulations using the SFT based method with a longer truncation length (8192 taps) for the Bessel functions. However, we found that the sound field separation errors of the SFT based methods in case the spatial filters are 8192 taps long are similar to the line with the plus sign ($+$) in Fig. 7, thus the simulation results are not shown. In the proposed method we have specially derived the spatial filter functions functions, $g_0^{\mu}(t)$, $g_1^{\mu}(t)$, $g_2^{\mu}(t)$, to be finite long, avoiding SFS error due to the impulse response function truncation.

296 The simulation results demonstrate that the Spatial Fourier Transform-based method,

297 which does not require the sensor array to surround the target source, is best suited for

298 separating transient plane wave sound fields. The proposed method, on the other hand,

299 can accurately separate the target (outgoing) sound field on a spherical array over a longer

300 duration.

301 **C. Sound field separation in a room**

302 In practical implementations of the proposed SFS method, the target sources may be

303 placed on a surface. In case the surface is rigid or highly reflective, we can measure the sound

304 field using a semi-spherical array around the target sources, and duplicate the measurements

305 with respect to the surface to use the full spherical harmonics expansion.[3,4] In this section,

306 we simulate such a case.

307 We have a room of size $(4, 5, 3)$ m as shown in Fig. 8. The reflection coefficients of all the

308 walls, floor, and ceiling are 0.99. One corner of the room is located at $D = (-1.8, -1.5, 0)$

309 m with respect to O, a point on the floor. Based on the point O, we set up a Cartesian

310 and a spherical coordinate system. There is a target point source at the point O inside

311 the semi-sphere $\mathbb{S}_2$ of radius $R = 0.5$. The length of impulse response functions $g_0^\mu(t)$,

312 $g_1^\mu(t)$, $g_2^\mu(t)$ is $T_n = \lceil 2R/cf_\mathrm{s} \rceil + 1 = 141$ taps. There is another interfering point source at

313 $(0.7, 0.8, 0.7)$ m outside of the sphere $\mathbb{S}_2$. These two sources produces band-limited random

314 noises, which are generated by passing unit-variance Gauss white noises through a 64-tap

315 Bandpass Butterworth filter with frequency range [100, 300] Hz.

We measure the sound pressure and the radial particle velocity on the semi-sphere $\mathbb{S}_2$ at four sensors, whose locations are determined according to the first-order Gauss sampling scheme.[28] We simulate the impulse responses, including the radial particle velocity response, between each source and sensor using the image source method.[19] We take the first 2744 image sources into consideration, and the impulse responses are truncated to 8192 taps long under the sampling frequency of 48 kHz. The pressure and the radial partial velocity on the semi-sphere $\mathbb{S}_2$ are duplicated with respect to the floor to use the full spherical harmonic expansions.[3,4] We estimate the outgoing field coefficients up to 0-th order, and reconstruct the outgoing field on the sphere over a period of 10 ms according to (8). 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 $\hat{p}^{\mathrm{o}}(t, R, \Theta_1)$, and the outgoing field separation error $p^{\mathrm{e}}(t, R, \Theta_1) = p^{\mathrm{o}}(t, R, \Theta_1) - \hat{p}^{\mathrm{o}}(t, R, \Theta_1)$, at the first sensor in Fig. 9. The outgoing field is produced by convolving the target source outputs and corresponding the half-space Green function.[3,4] The total field is the summation of the outgoing and incoming fields.

As shown in Fig. 9, in this case, the proposed method is still able to accurately estimate the outgoing field. The normalized separation error at the first sensor is

$$\xi_1 = 10\log 10 \frac{\sum_{t=0}^{10 \text{ ms}} ||p^{\mathrm{e}}(t, R, \Theta_1)||^2}{\sum_{t=0}^{10 \text{ ms}} ||p^{\mathrm{o}}(t, R, \Theta_1)||^2} = -31 \text{ dB}. \tag{26}$$

The proposed method also recovers the outgoing field over the whole semi-sphere. The results are similar to Fig. 4, thus are not shown for brevity.

VI. CONCLUSION

In this paper, we developed a time-domain SFS method that can separate non-stationary sound fields over a sphere. We decompose the sound field and the radial particle velocity measure on the sphere into spherical harmonic coefficients, and recover the outgoing/incoming field based on the time-domain relationship between the coefficients and the derived separation 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. A future extension of the proposed method is to take the scattering from the target source surface into consideration.[8,9,11] Note that, as mentioned in Sec. IV, we can use either the acoustic 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.[17] The sensor calibration and the experimental validation of the proposed method will be our future work.

349 **APPENDIX A: PROOF OF THEOREM 1**

350 *Proof.* In this section, we derive the expression of the time-domain outgoing field coefficients

351 $a_{\mu\nu}^{\mathrm{o}}(t, R)$.

Substitutions of $\tau_R = R/c$ and[22]

$$j_\mu'(x) = \frac{\mu}{x}j_\mu(x) - j_{\mu+1}(x),$$

in (6) result in

$$A_{\mu\nu}^{\mathrm{o}}(\omega, R) = -\mu(i\omega\tau_R)j_\mu(\omega\tau_R)h_\mu(\omega\tau_R) \times A_{\mu\nu}(\omega, R)$$

$$+\tau_R(\omega\tau_R)j_{\mu+1}(\omega\tau_R)h_\mu(\omega\tau_R) \times (i\omega)A_{\mu\nu}(\omega, R)$$

$$-\rho_0 c\tau_R(i\omega\tau_R)j_\mu(\omega\tau_R)h_\mu(\omega\tau_R) \times (i\omega)B_{\mu\nu}(\omega, R) \tag{A1}$$

We define the time-domain outgoing field coefficients $a_{\mu\nu}^{\mathrm{o}}(t, R)$ as

$$a_{\mu\nu}^{\mathrm{o}}(t, R) \equiv \mathcal{F}^{-1}\{A_{\mu\nu}^{\mathrm{o}}(\omega, R)\}$$

$$= g_0^\mu(t) * \lambda_{\mu\nu}(t) + g_1^\mu(t) * \chi_{\mu\nu}(t) + \rho_0 c\, g_2^\mu(t) * \eta_{\mu\nu}(t), \tag{A2}$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform.[29] The terms in the second line of (A2) are

defined based on (A1) as follows

$$\lambda_{\mu\nu}(t) \equiv \mathcal{F}^{-1}\Big\{A_{\mu\nu}(\omega, R)\Big\}, \tag{A3}$$

$$\chi_{\mu\nu}(t) \equiv \mathcal{F}^{-1}\Big\{(i\omega)A_{\mu\nu}(\omega, R)\Big\}, \tag{A4}$$

$$\eta_{\mu\nu}(t) \equiv \mathcal{F}^{-1}\Big\{(i\omega)B_{\mu\nu}(\omega, R)\Big\}, \tag{A5}$$

$$g_0^\mu(t) \equiv \mathcal{F}^{-1}\Big\{-\mu(i\omega\tau_R)j_\mu(\omega\tau_R)h_\mu(\omega\tau_R)\Big\}, \tag{A6}$$

$$g_1^\mu(t) \equiv \mathcal{F}^{-1}\Big\{\tau_R(\omega\tau_R)j_{\mu+1}(\omega\tau_R)h_\mu(\omega\tau_R)\Big\}, \tag{A7}$$

$$g_2^\mu(t) \equiv \mathcal{F}^{-1}\Big\{-\tau_R(i\omega\tau_R)j_\mu(\omega\tau_R)h_\mu(\omega\tau_R)\Big\}. \tag{A8}$$

352 We derive expressions for these six terms in the following.

First, based on properties of the Fourier transform and the spherical harmonic trans-

form,[1,22] we obtain $\lambda_{\mu\nu}(t)$, $\chi_{\mu\nu}(t)$, and $\eta_{\mu\nu}(t)$ by decomposing the the pressure $p(t, R, \Theta)$ and

the radial particle velocity $v(t, R, \Theta)$ as follows

$$
\begin{aligned}
\lambda_{\mu\nu}(t) &= \mathcal{F}^{-1}\{A_{\mu\nu}(\omega, R)\} \\
&= \mathcal{F}^{-1}\left\{ \int_{\Theta} P(\omega, R, \Theta) Y_{\mu\nu}(\Theta) d\Theta \right\} \\
&= \int_{\Theta} p(t, R, \Theta) Y_{\mu\nu}(\Theta) d\Theta,
\end{aligned}
\tag{A9}
$$

$$
\begin{aligned}
\chi_{\mu\nu}(t) &= \mathcal{F}^{-1}\left\{ (i\omega) A_{\mu\nu}(\omega, R) \right\} \\
&= \frac{d}{dt}\left\{ \mathcal{F}^{-1}[A_{\mu\nu}(\omega, R)] \right\} \\
&= \frac{d}{dt}\left\{ \mathcal{F}^{-1}\left[\int_{\Theta} P(\omega, R, \Theta) Y_{\mu\nu}(\Theta) d\Theta \right] \right\} \\
&= \frac{d}{dt}\left\{ \int_{\Theta} p(t, R, \Theta) Y_{\mu\nu}(\Theta) d\Theta \right\} \\
&= \int_{\Theta} \frac{dp(t, R, \Theta)}{dt} Y_{\mu\nu}(\Theta) d\Theta,
\end{aligned}
\tag{A10}
$$

$$
\begin{aligned}
\eta_{\mu\nu}(t) &= \mathcal{F}^{-1}\left\{ (i\omega) B_{\mu\nu}(\omega, R) \right\} \\
&= \frac{d}{dt}\left\{ \mathcal{F}^{-1}[B_{\mu\nu}(\omega, R)] \right\} \\
&= \frac{d}{dt}\left\{ \mathcal{F}^{-1}\left[\int_{\Theta} V(\omega, R, \Theta) Y_{\mu\nu}(\Theta) d\Theta \right] \right\} \\
&= \frac{d}{dt}\left\{ \int_{\Theta} v(t, R, \Theta) Y_{\mu\nu}(\Theta) d\Theta \right\} \\
&= \int_{\Theta} \frac{dv(t, R, \Theta)}{dt} Y_{\mu\nu}(\Theta) d\Theta.
\end{aligned}
\tag{A11}
$$

Then based on the following properties of the spherical Bessel and Hankel function[1,22]

$$h_\mu(x) = e^{-ix} i^{\mu+2} \sum_{\nu=0}^{\mu} \frac{\varphi_\nu(\mu)}{(ix)^{\nu+1}},$$

$$\varphi_\nu(\mu) = \frac{(\mu+\nu)!}{2^\nu \nu! (\mu-\nu)!},$$

$$j_\mu(x) = \frac{1}{2}[h_\mu(x)^* + h_\mu(x)],$$

$$h'_\mu(x) = h_{\mu-1}(x) - \frac{\mu+1}{x} h_\mu(x),$$

$$h_{-\mu-1}(x) = i(-1)^{\mu+1} h_\mu(x), \quad \mu \geq 0,$$

we expand (A6), (A7), and (A8) as

$$g_0^\mu(t) = \mathcal{F}^{-1}\left[-\mu(i\tau_R\omega) j_\mu(\omega\tau_R) h_\mu(\omega\tau_R) \right]$$
$$= \mathcal{F}^{-1}\left[\frac{\mu}{2} \sum_{\nu=0}^{\mu} \sum_{\varsigma=0}^{\mu} \varphi_\nu(\mu)\varphi_\varsigma(\mu) \frac{(-1)^\nu + (-1)^{\mu+1} e^{-2i\tau_R\omega}}{(i\tau\omega)^{\nu+\varsigma+1}} \right], \quad \text{(A12)}$$

$$g_1^\mu(t) = \mathcal{F}^{-1}\left[\tau_R(\omega\tau_R) j_{\mu+1}(\omega\tau_R) h_\mu(\omega\tau_R) \right]$$
$$= \mathcal{F}^{-1}\left[\frac{\tau_R}{2} \sum_{\nu=0}^{\mu+1} \sum_{\varsigma=0}^{\mu} \varphi_\nu(\mu+1)\varphi_\varsigma(\mu) \frac{(-1)^\nu + (-1)^\mu e^{-2i\tau_R\omega}}{(i\tau_R\omega)^{\nu+\varsigma+1}} \right], \quad \text{(A13)}$$

$$g_2^\mu(t) = \mathcal{F}^{-1}\left[-\tau_R(i\omega\tau_R) j_\mu(\omega\tau_R) h_\mu(\omega\tau_R) \right]$$
$$= \mathcal{F}^{-1}\left[\frac{\tau_R}{2} \sum_{\nu=0}^{\mu} \sum_{\varsigma=0}^{\mu} \varphi_\nu(\mu)\varphi_\varsigma(\mu) \frac{(-1)^\nu + (-1)^{\mu+1} e^{-i2\omega\tau_R}}{(i\tau_R\omega)^{\nu+\varsigma+1}} \right], \quad \text{(A14)}$$

respectively. Based on the Fourier transform properties,[29] (A12), (A13), and (A14) can be further developed as (12), (13), and (14), respectively.

The expressions of the incoming field coefficients $a_{\mu\nu}^i(t, R)$, $g_3^\mu(t)$, and $g_4^\mu(t)$ can be derived similarly, and are not shown for brevity.

□

28

358 **References**

359 [1]E. G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*

360 (Academic press, New York, 1999), Chap. 6-7, pp. 183–249.

361 [2]G. Weinreich and E. B. Arnold, "Method for measuring acoustic radiation fields," *J.*

362 *Acoust. Soc. Am.*, **68**(2), pp. 404–411 (1980).

363 [3]M. Melon, C. Langrenne, P. Herzog, and A. Garcia, "Evaluation of a method for the

364 measurement of subwoofers in usual rooms," *J. Acoust. Soc. Am.*, **127**(1), pp. 256–263

365 (2010).

366 [4]Y. Braikia, M. Melon, C. Langrenne, and A. Garcia, "Evaluation of a separation method

367 for source identification in small spaces," *J. Acoust. Soc. Am.*, **134**(1), pp. 323–331 (2013).

368 [5]S. Lobréau, É. Bavu, and M. Melon, "Hemispherical double-layer time reversal imaging in

369 reverberant and noisy environments at audible frequencies," *J. Acoust. Soc. Am.*, **137**(2),

370 pp. 785–796 (2015).

371 [6]M. Tamura, "Spatial Fourier Transform method of measuring reflection coefficients at

372 oblique incidence. I: Theory and numerical examples," *J. Acoust. Soc. Am.*, **88**(5), pp.

373 2259–2264 (1990).

374 [7]F. Jacobsen and V. Jaud, "Statistically optimized near field acoustic holography using an

375 array of pressure-velocity probes," *J. Acoust. Soc. Am.*, **121**(3), pp. 1550–1558 (2007).

376 [8]C. Langrenne, M. Melon, and A. Garcia, "Boundary element method for the acoustic

377 characterization of a machine in bounded noisy environment," *J. Acoust. Soc. Am.*, **121**(5),

378 pp. 2750–2757 (2007).

[379] [9]C. Langrenne, M. Melon, and A. Garcia, "Measurement of confined acoustic sources using

[380] near-field acoustic holography," *J. Acoust. Soc. Am.*, **126**(3), pp. 1250–1256 (2009).

[381] [10]C. Bi, X. Chen, and J. Chen, "Sound field separation technique based on equivalent source

[382] method and its application in nearfield acoustic holography," *J. Acoust. Soc. Am.*, **123**(3),

[383] pp. 1472–1478 (2008).

[384] [11]C. Bi and B. J. Stuart, "An equivalent source technique for recovering the free sound field

[385] in a noisy environment," *J. Acoust. Soc. Am.*, **131**(2), pp. 1260–1270 (2012).

[386] [12]E. Fernandez-Grande, F. Jacobsen, and Q. Leclere, "Sound field separation with sound

[387] pressure and particle velocity measurements," *J. Acoust. Soc. Am.*, **132**(6), pp. 3818–3825

[388] (2012).

[389] [13]F. Ma, W. Zhang, and T. D. Abhayapala, "Reference Signal Generation for Broad-

[390] band ANC Systems in Reverberant Rooms," *IEEE International Conference on Acoustics,*

[391] *Speech and Signal Processing (ICASSP)*, pp. 216–220 (2018).

[392] [14]X. Zhang, J. H.Thomas, and J. C. Pascal, "Separation of non-stationary sound fields in

[393] the time-wavenumber domain," *J. Acoust. Soc. Am.*, **131**(3), pp. 2180–2189 (2012).

[394] [15]C. Bi, L. Geng, and X. Zhang, "Real-time separation of non-stationary sound fields with

[395] pressure and particle acceleration measurements," *J. Acoust. Soc. Am.*, **135**(6), pp. 3474–

[396] 3482 (2014).

[397] [16]C. Bi, L. Geng, and X. Zhang, "Separation of non-stationary multi-source sound field based

[398] on the interpolated time-domain equivalent source method," *Mech. Syst. Signal Process.*,

[399] **72**, pp. 745–761 (2016).

[400] [17]C. Bi, L. Geng, and X. Zhang, "Separation of non-stationary sound fields with a single

[401] layer pressure-velocity measurements," *Mech. Syst. Signal Process.*, **139**(2), pp. 781–789

[402] (2016).

[403] [18]H. Kuttruff, *Room Acoustics* (Crc Press, New York, 2014), pp. 1–392.

[404] [19]J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acous-

[405] tics," *J. Acoust. Soc. Am.*, **65**(4), pp. 943–950 (1979).

[406] [20]M. A. Poletti, "Unified description of ambisonics using real and complex spherical har-

[407] monics," *Proc. Ambisonics Symp.*, pp. 1–10 (2009).

[408] [21]F. Ma, and W. Zhang, and T. D. Abhayapala, "Active control of outgoing noise fields in

[409] rooms," *J. Acoust. Soc. Am.*, **144**(3), pp. 1589–1599 (2018).

[410] [22]M. Abramowitz and I. A. Stegun, "Hand book of Mathematical Functions: With Formulas,

[411] Graphs, and Mathematical Tables," (National Bureau of Standards, Washington, DC,

[412] 1964), pp. 355–435.

[413] [23]D. B. Ward and T. D. Abhayapala, "Reproduction of a plane-wave sound field using an

[414] array of loudspeakers," *IEEE Trans. Speech Audio Process.*, **9**(6), pp. 697–707 (2001).

[415] [24]J. W. Parkins, S. D. Sommerfeldt, and J. Tichy, "Narrowband and broadband active

[416] control in an enclosure using the acoustic energy density," *J. Acoust. Soc. Am.*, **108**(1),

[417] pp. 192–203 (2003).

[418] [25]H. E. de Bree, "The microflown: An acoustic particle velocity sensor," *Acoust. Aust.*,

[419] **31**(3), pp. 91–94 (2003).

420 [26]N. Epain and C. T. Jin, "Spherical Harmonic Signal Covariance and Sound Field Diffuse-

421 ness," *IEEE/ACM Trans. Speech Audio Process.*, **24**(10), pp. 1796–1807 (2016).

422 [27]N. Sloane, Spherical codes: Nice arrangements of points on a sphere in various dimensions,

423 (1996), http://neilsloane.com/packings/.

424 [28]B. Rafaely, *Fundamentals of Spherical Array Processing* (Springer, New York, 2015),

425 Chap. 3, pp. 57–78.

426 [29]I. S. Gradshteyn, and I. M. Ryshik, *Table of integrals, series, and products 5th ed.* (Aca-

427 demic Press, New York, 1994), pp. 1185.

Table of Figure captain

Fig. 1 System models: (a) the target sources $\triangle$ are placed in free field, (b) the target sources $\triangle$ are placed inside a room. The outgoing and incoming field on the sphere $\mathbb{S}_2$ of radius R are denoted as $\dashrightarrow$ and $\rightsquigarrow$, respectively.

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

Fig. 3 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 $\hat{p}^{\mathrm{o}}(t, R, \Theta_{17})$, and the outgoing field separation error $p^{\mathrm{e}}(t, R, \Theta_{17})$, at the 17-th sensor over 10 ms.

Fig. 4 (Color online)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 $\mathbb{S}_2$.[28]

Fig. 5 Separation error: The normalized separation error as a function of the frequency and the truncation order N.

Fig. 6 Setup of sound field separation 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 planes H and H_1 to separate the sound waves at the point O produced by the target source $\triangle$ and the disturbing source $\bullet$.

Fig. 7 Sound field separation at the point O: The amplitudes of the target sound, the total sound, the sound separation errors using the proposed and the Spatial Fourier Transform based method over 30 ms.

449   Fig. 8 Sound field separation in a room: A target source $\triangle$ is placed on at the point

450   O inside a semi-sphere $\mathbb{S}_2$ of radius $R = 0.5$ m, and another point source $\bullet$ is placed at

451    $(0.7, 0.8, 0.7)$ m with respective to the point O.

452   Fig. 9 Sound field separation at a point: The amplitudes of the outgoing field $p^{\mathrm{o}}(t, R, \Theta_1)$,

453   the total field $p(t, R, \Theta_1)$, the separated outgoing field $\hat{p}^{\mathrm{o}}(t, R, \Theta_1)$, and the outgoing field

454   separation error $p^{\mathrm{e}}(t, R, \Theta_1)$, at the first sensor over 10 ms.