Directionally Weighted Wave Field Estimation Exploiting Prior Information on Source Direction

Natsuki Ueno, *Student Member, IEEE*, Shoichi Koyama, *Member, IEEE*, and Hiroshi Saruwatari, *Member, IEEE*

***Abstract*—A wave field estimation method exploiting prior information on source direction is proposed. First, we formulate a wave field estimation problem as regularized least squares, where the norm of the wave field is used for a regularization term. The norm of the wave field is defined on the basis of the weighting function that reflects the prior information on the source direction. We derive the closed-form solution using theories on Hilbert spaces. Results of numerical experiments indicated that high estimation accuracy can be achieved by using the proposed method in comparison with other current methods that do not use any prior information.**

***Index Terms*—Helmholtz equation, Hilbert space, inverse problem, sensor array, wave field estimation.**

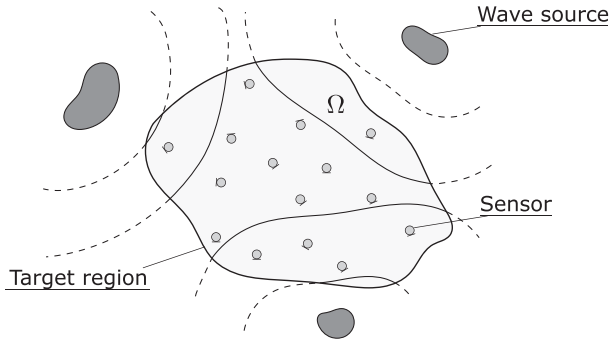


Fig. 1. Problem setting of wave field estimation.

I. Introduction

CAPTURING a three-dimensional wave field is an essential technique for its analysis, visualization, and other related applications. In the field of spatial audio, in particular, sound field estimation techniques using multiple sensors (i.e., microphones) have attracted considerable attention owing to their wide variety of applications, such as the reproduction of a captured sound field using loudspeakers [1]–[4] or headphones [5]–[7] and the spatial active noise control [8]–[10].

Wave field estimation methods are classified on the basis of whether wave sources are allowed to exist inside the target region [11]–[15] or not [2], [16]–[20]. In general, the former case is much more difficult than the latter and often requires additional assumptions, such as spatial sparsity of source distribution and a reverberant-free condition. Here, we focus on the wave field estimation inside a source-free target region as shown in Fig. 1. Note that wave sources may exist outside the target region also in this context.

There are a large number of wave field estimation methods for source-free target regions, which can be further classified on the basis of whether the linearity between the observed signals and estimated wave fields at each frequency, i.e., the *linear time-invariant* (LTI) property, holds. For example, many methods based on the higher-order Ambisonics (HOA) approach in spatial audio have the LTI property [2], [18]–[20]. In these methods, the wave fields are expanded by spherical wavefunctions up to a certain order, and their coefficients are estimated linearly from the observed signals [2], [18], [20]. In our previous work [19], the spherical wavefunction expansion is considered up to infinite dimensions, which is essentially equivalent to the expansion by infinite number of plane-wave functions [21]. This makes the estimation process independent of the settings of the expansion center and truncation order. Owing to the LTI property, these methods allow a fast implementation using LTI digital filters, which is preferable to non-LTI ones from the viewpoint of computational efficiency and almost essential for real-time systems (e.g., spatial active noise control).

On the other hand, sparsity-based methods [15], [22]–[24] are typical examples of non-LTI approaches. In these methods, wave fields are decomposed into sparse basis functions, and their coefficients are determined by sparse optimization methods. In contrast to the HOA approach, the plane-wave functions and monopole functions (free-field Green's functions) are often used as the basis functions. These methods may improve estimation accuracy in various practical situations since they exploit the prior information on the target wave field, i.e., the sparsity of the coefficients. However, these non-LTI methods require iterative calculations for optimization, which makes a real-time implementation difficult. In addition, the function space in which we seek the solution has to be approximated by a finite number of basis functions, e.g., by discretizing the directions of plane-wave functions or the positions of monopole functions, which causes

Manuscript received June 18, 2020; revised December 18, 2020 and February 15, 2021; accepted March 25, 2021. Date of publication April 2, 2021; date of current version April 26, 2021. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. M. R. Petraglia. This work was supported in part by JSPS KAKENHI under Grant JP18J21926, in part by JST PRESTO under Grant JPMJPR18J4, and in part by JSPS-CAS Joint Research Program under Grant JPJSBP120197203. *(Corresponding author: Natsuki Ueno.)*

The authors are with the Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan (e-mail: natsuki.ueno@ieee.org; koyama.shoichi@ieee.org; hiroshi_saruwatari@ipc.i.u-tokyo.ac.jp).

Digital Object Identifier 10.1109/TSP.2021.3070228

a tradeoff relationship between the approximation accuracy and the computational cost. Therefore, an LTI method exploiting some prior information on a target wave field is desired to achieve both high estimation accuracy and low computational cost.

In this paper, we present a wave field estimation method exploiting prior information on source direction. Although it is not always easy to obtain the exact source positions beforehand, their rough estimates are available in many practical situations (e.g., when the possible source position is limited within a certain region owing to the physical constraint or when the source position can be estimated by using some other sensors such as image sensors). We formulate the wave field estimation problem as the regularized least squares in a Hilbert space [25], where the prior information is incorporated as a *directional weighting* into the regularization term, and the closed-form solution can be obtained using the representer theorem [26], [27]. In particular, this method

- has the LTI property,
- incorporates prior information on source direction in a flexible form (e.g., single direction, multiple directions, and diffuse field), and
- does not require the finite-dimensional approximation of the function space composed of wave fields as in our previous work [19], [28], which is made possible by utilizing theories of functional analysis.

Note that the formulation in this paper includes those in our previous works [19], [28] as special cases; the proposed method can be regarded as their generalization. The proposed method is also closely related to our other previous work [29], where the kernel interpolation of two-dimensional sound fields based on directional weighting is presented. Whereas only omnidirectional sensors in two-dimensional wave fields are considered in [29], the sensor directivity is extended to general cases and three-dimensional wave fields are considered in this paper.

The rest of this paper is organized as follows. In Section II, several notations and preliminaries are introduced. In Section III, the proposed method is described. In Section IV, results of several numerical experiments for comparison between the proposed method and other LTI methods are reported. Finally, we present our conclusions in Section V.

II. Notations and Preliminaries

First, we provide several basic notations and preliminaries on the representation of wave fields and modeling of sensor directivities.

A. Notations

The set of natural numbers including zero are denoted by $\mathbb{N}$, and the set of positive integers are denoted by $\mathbb{N}_+$. For any integers m and n satisfying $m \leq n$, $[\![m, n]\!]$ denotes the integer interval between m and n inclusive.

The sets of real and complex numbers are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. The imaginary unit is denoted by i, and the complex conjugate of a complex number is denoted by $(\cdot)^*$. The unit sphere in $\mathbb{R}^3$ is denoted by $\mathbb{S}_2$. The transpose and Hermitian transpose of a matrix are denoted by $(\cdot)^{\mathsf{T}}$ and $(\cdot)^{\mathsf{H}}$, respectively. For any two vectors $\boldsymbol{x} \coloneqq [x_1, x_2, x_3]^{\mathsf{T}}$ and $\boldsymbol{y} \coloneqq [y_1, y_2, y_3]^{\mathsf{T}}$ in $\mathbb{S}_2$, $\mathbb{R}^3$, or $\mathbb{C}^3$, the operator $\circ$ is defined as $\boldsymbol{x} \circ \boldsymbol{y} \coloneqq x_1 y_1 + x_2 y_2 + x_3 y_3$.

An integral of a measurable function $f : \mathbb{S}_2 \to \mathbb{C}$ over the sphere $\mathbb{S}_2$ (with respect to the natural measure) is denoted by $\int_{\boldsymbol{x} \in \mathbb{S}_2} f(\boldsymbol{x}) \, \mathrm{d}\chi$. The set of square-integrable functions from $\mathbb{S}_2$ to $\mathbb{C}$ is denoted by $L_2(\mathbb{S}_2)$. According to conventions, we do not distinguish functions in $L_2(\mathbb{S}_2)$ that are equal almost everywhere on $\mathbb{S}_2$.

B. Representation of Wave Fields

Let $\Omega \subseteq \mathbb{R}^3$ be a simply connected open subset of $\mathbb{R}^3$ and $u : \Omega \to \mathbb{C}$ be a wave field in Ω at a fixed angular frequency $\omega \in \mathbb{R}$, i.e., $u(\boldsymbol{r})$ denotes a pressure at $\boldsymbol{r} \in \Omega$.[1] If Ω does not include any wave sources, u can be well modeled as a solution of the following (homogeneous) Helmholtz equation [30]:

$$(\Delta + k^2) u = 0, \tag{1}$$

where Δ denotes the Laplace operator and $k \coloneqq \omega / c$ is the wavenumber with $c \in (0, \infty)$ being the phase velocity, which is assumed to be constant in Ω. A typical modeling of u satisfying (1) is a superposition of plane-wave functions (also called the Herglotz integral [31]) as follows:

$$u(\boldsymbol{r}) = \int_{\boldsymbol{x} \in \mathbb{S}_2} \tilde{u}(\boldsymbol{x}) \exp(-ik \boldsymbol{x} \circ \boldsymbol{r}) \, \mathrm{d}\chi \quad (\boldsymbol{r} \in \Omega). \tag{2}$$

Here, $\tilde{u} \in L_2(\mathbb{S}_2)$ represents the complex amplitude of plane waves arriving from each direction. Let $\mathcal{A}$ denote a transform of functions from $\tilde{u}$ to u defined as (2), and we define a function space $\mathscr{H}$ as

$$\mathscr{H} \coloneqq \{ \mathcal{A} \tilde{u} \mid \tilde{u} \in L_2(\mathbb{S}_2) \}. \tag{3}$$

In what follows, the tilde symbol over a letter denotes the inverse of $\mathcal{A}$ (e.g., $\tilde{u} \coloneqq \mathcal{A}^{-1} u$).

Although $\mathscr{H}$ does not include all solutions of (1), any solution of (1) can be approximated arbitrarily by functions in $\mathscr{H}$ in the sense of the uniform convergence on compact sets (see Appendix VII for a proof). Therefore, a wave field estimation problem can be regarded practically as a process of determining a function u within $\mathscr{H}$ on the basis of some criterion.

C. Modeling of Sensor Directivity

Suppose a single sensor with a certain directivity located at $\boldsymbol{r}_0 \in \Omega$ in a wave field $u \in \mathscr{H}$. Its directivity can be modeled as the directional function $\gamma \in L_2(\mathbb{S}_2)$, i.e., $\gamma(\boldsymbol{x})$ denotes the sensor's response to the plane wave arriving from the direction $\boldsymbol{x} \in \mathbb{S}_2$. For example, omnidirectional, bidirectional, and first-order sensors are respectively modeled with constant parameters $\boldsymbol{y} \in \mathbb{S}_2$ and $\zeta \in [0, 1]$ as follows.

Omnidirectional:

$$\gamma(\boldsymbol{x}) \coloneqq 1 \quad (\boldsymbol{x} \in \mathbb{S}_2)$$

[1] Harmonic time dependence $\exp(-i\omega t)$ with $t \in \mathbb{R}$ denoting the time will be assumed in this paper according to conventions [30].

Bidirectional:

$$\gamma(\boldsymbol{x}) \coloneqq \boldsymbol{y} \circ \boldsymbol{x} \quad (\boldsymbol{x} \in \mathbb{S}_2)$$

First-order:

$$\gamma(\boldsymbol{x}) \coloneqq \zeta + (1-\zeta)\boldsymbol{y} \circ \boldsymbol{x} \quad (\boldsymbol{x} \in \mathbb{S}_2)$$

In many practical cases including the above examples, the directivity γ can be well represented using finite-order spherical harmonic functions as

$$\gamma(\boldsymbol{x})^* = \sum_{\nu,\mu}^{N} c_{\nu,\mu} Y_{\nu,\mu}(\boldsymbol{x}) \quad (\boldsymbol{x} \in \mathbb{S}_2), \tag{4}$$

where $N \in \mathbb{N}$ is the maximum order, $\sum_{\nu,\mu}^{N}$ is the abbreviated form of $\sum_{\nu=0}^{N}\sum_{\mu=-\nu}^{\nu}$, and $Y_{\nu,\mu}(\cdot) : \mathbb{S}_2 \to \mathbb{C}$ denotes the spherical harmonic function of order $\nu \in \mathbb{N}$ and degree $\mu \in [\![-\nu, \nu]\!]$ (see [21] for the definition and [32] for an efficient computational algorithm). Here, we consider the complex conjugate on the right-hand side of (4) to simplify later discussion.

Since u can be expanded around $\boldsymbol{r}_0$ as

$$u(\boldsymbol{r}) = \int_{\boldsymbol{x}\in\mathbb{S}_2} \tilde{u}(\boldsymbol{x}) \exp(-ik\boldsymbol{x} \circ \boldsymbol{r}_0) \cdot \exp(-ik\boldsymbol{x} \circ (\boldsymbol{r} - \boldsymbol{r}_0))\, \mathrm{d}\chi \quad (\boldsymbol{r} \in \Omega), \tag{5}$$

the observed signal $s \in \mathbb{C}$ of the sensor is given by

$$s = \int_{\boldsymbol{x}\in\mathbb{S}_2} \tilde{u}(\boldsymbol{x}) \exp(-ik\boldsymbol{x} \circ \boldsymbol{r}_0)\gamma(\boldsymbol{x})\, \mathrm{d}\chi + \epsilon, \tag{6}$$

where $\epsilon \in \mathbb{C}$ is the observation noise. Hereafter, we use the simplified notation as

$$s = \mathcal{F}u + \epsilon. \tag{7}$$

Here, $\mathcal{F} : \mathscr{H} \to \mathbb{C}$ is a functional defined as

$$\mathcal{F}u \coloneqq \int_{\boldsymbol{x}\in\mathbb{S}_2} \tilde{u}(\boldsymbol{x}) \exp(-ik\boldsymbol{x} \circ \boldsymbol{r}_0)\gamma(\boldsymbol{x})\, \mathrm{d}\chi \quad (u \in \mathscr{H}). \tag{8}$$

In addition, we refer to $\mathcal{F}$ as an observation operator of the sensor. Note that $\mathcal{F}$ is a linear functional on $\mathscr{H}$, which means the superposition principle in the observation.

III. Proposed Method

A. Formulation

Let $M \in \mathbb{N}_+$ sensors be located arbitrarily in Ω as shown in Fig. 1. For each $m \in [\![1, M]\!]$, the position, directivity, and observed signal of the mth sensor are denoted respectively by $\boldsymbol{r}_m \in \Omega$, $\gamma_m \in L_2(\mathbb{S}_2)$, and $s_m \in \mathbb{C}$, and they are assumed to be given. In this case, the observation operator of the mth sensor, denoted by $\mathcal{F}_m$, is given for each $m \in [\![1, M]\!]$ by

$$\mathcal{F}_m u \coloneqq \int_{\boldsymbol{x}\in\mathbb{S}_2} \tilde{u}(\boldsymbol{x}) \exp(-ik\boldsymbol{x} \circ \boldsymbol{r}_m)\gamma_m(\boldsymbol{x})\, \mathrm{d}\chi \quad (u \in \mathscr{H}). \tag{9}$$

We consider the following formulation of wave field estimation problems:

$$\underset{u\in\mathscr{H}}{\text{minimize}} \quad Q(u) \coloneqq \sum_{m=1}^{M} \frac{1}{\sigma_m^2} |\mathcal{F}_m u - s_m|^2 + \lambda \|u\|_{\mathscr{H}}^2, \tag{10}$$

where $\sigma_1, \ldots, \sigma_M \in (0, \infty)$ are dispersion parameters representing the observational uncertainty, $\lambda \in (0, \infty)$ is a regularization parameter, and $\|\cdot\|_{\mathscr{H}} : \mathscr{H} \to [0, \infty)$ is a norm on $\mathscr{H}$, which is defined later. The first term of (10) is a loss term, which represents a deviation between the predicted values $\mathcal{F}_1 u, \ldots, \mathcal{F}_M u$ and the observed values $s_1, \ldots, s_M$. On the other hand, the second term is a regularization term, which evaluates the reasonability of u independently of the observation. If we can design the norm $\|\cdot\|_{\mathscr{H}}$ so that the wave fields that are likely to occur have small norms, the solution of (10) will be induced to such wave fields. Therefore, it is desirable to design an appropriate norm exploiting prior information of wave fields.

On the basis of the above discussion, we introduce the norm $\|\cdot\|_{\mathscr{H}}$ over $\mathscr{H}$ defined as

$$\|u\|_{\mathscr{H}} \coloneqq \left(\int_{\boldsymbol{x}\in\mathbb{S}_2} \frac{|\tilde{u}(\boldsymbol{x})|^2}{w(\boldsymbol{x})}\, \mathrm{d}\chi \right)^{\frac{1}{2}} \quad (u \in \mathscr{H}). \tag{11}$$

Here, $w : \mathbb{S}_2 \to (0, \infty)$ is a weighting function defined as

$$w(\boldsymbol{x}) \coloneqq \frac{1}{4\pi C(\beta)} \exp(\beta \boldsymbol{\eta} \circ \boldsymbol{x}) \quad (\boldsymbol{x} \in \mathbb{S}_2), \tag{12}$$

where $\beta \in [0, \infty)$ and $\boldsymbol{\eta} \in \mathbb{S}_2$ are constant parameters and $C(\beta)$ is a scaling constant (so that w satisfies $\int_{\boldsymbol{x}\in\mathbb{S}_2} w(\boldsymbol{x})\, \mathrm{d}\chi = 1$) defined as

$$C(\beta) \coloneqq \begin{cases} 1 & (\beta = 0) \\ (\exp(\beta) - \exp(-\beta))/(2\beta) & (\beta \in (0, \infty)) \end{cases}. \tag{13}$$

This function w is also known as the probability density function of the von Mises–Fisher distribution used in directional statistics [33]. Note that $w(\boldsymbol{x})$ takes a large value when $\boldsymbol{x}$ is close to $\boldsymbol{\eta}$, especially in the case of large β. Therefore, by using this norm in (10), we can impose large penalties for wave fields originating from wave sources in the direction away from $\boldsymbol{\eta}$ while imposing small penalties for wave fields from wave sources in the direction close to $\boldsymbol{\eta}$. When $\beta = 0$, on the contrary, the solution of (10) will be induced relatively to a diffuse field. The weighting function w can be further generalized by using a linear combination of (12) for different parameters β and $\boldsymbol{\eta}$, which is useful in cases of multiple wave sources. This extension is discussed in Section III-D after we describe the basic concept of the proposed method.

B. Closed-Form Solution

We provide the closed-form solution of (10) on the basis of theories on Hilbert spaces. First, we define the inner product $\langle\cdot,\cdot\rangle_{\mathscr{H}}$ on $\mathscr{H}$ as

$$\langle u_1, u_2\rangle_{\mathscr{H}} \coloneqq \int_{\boldsymbol{x}\in\mathbb{S}_2} \frac{\tilde{u}_1(\boldsymbol{x})^* \tilde{u}_2(\boldsymbol{x})}{w(\boldsymbol{x})}\, \mathrm{d}\chi \quad (u_1, u_2 \in \mathscr{H}). \tag{14}$$

Then, $(\mathscr{H}, \langle\cdot,\cdot\rangle_{\mathscr{H}})$ is a complex Hilbert space because it is isomorphic to the Hilbert space $L_2(\mathbb{S}_2)$ with the weighted L_2 inner product. Using this inner product $\langle\cdot,\cdot\rangle_{\mathscr{H}}$, we can write the objective function as

$$Q(u) = \sum_{m=1}^{M} \frac{1}{\sigma_m^2} |\langle v_m, u\rangle_{\mathscr{H}} - s_m|^2 + \lambda \langle u, u\rangle_{\mathscr{H}} \quad (u \in \mathscr{H}), \tag{15}$$

where $v_1, \ldots, v_M \in \mathscr{H}$ are given by

$$v_m(\boldsymbol{r}) := \int_{\boldsymbol{x}\in\mathbb{S}_2} w(\boldsymbol{x})\gamma_m(\boldsymbol{x})^* \exp(-ik\boldsymbol{x}\circ(\boldsymbol{r}-\boldsymbol{r}_m))\,\mathrm{d}\chi \quad (\boldsymbol{r}\in\Omega,\ m\in[\![1,M]\!]). \tag{16}$$

Then, we can apply the representer theorem [26], [27], which guarantees that the solution of (10), denoted by $\hat{u}$, has the form of

$$\hat{u} = \sum_{m=1}^{M} \hat{\alpha}_m v_m \tag{17}$$

with some $\hat{\boldsymbol{\alpha}} \coloneqq [\hat{\alpha}_1, \ldots, \hat{\alpha}_M]^{\mathsf{T}} \in \mathbb{C}^M$. Therefore, we need only to seek the optimal coefficients $\hat{\boldsymbol{\alpha}}$ in the finite-dimensional linear space $\mathbb{C}^M$ instead of $\hat{u}$ in the infinite-dimensional linear space $\mathscr{H}$. Here, $\hat{\boldsymbol{\alpha}}$ can be obtained by solving the following optimization problem:

$$\underset{\boldsymbol{\alpha}\in\mathbb{C}^N}{\text{minimize}} \quad Q^{(*)}(\boldsymbol{\alpha}) := (\boldsymbol{K}\boldsymbol{\alpha}-\boldsymbol{s})^{\mathsf{H}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{K}\boldsymbol{\alpha}-\boldsymbol{s}) + \lambda\boldsymbol{\alpha}^{\mathsf{H}}\boldsymbol{K}\boldsymbol{\alpha}, \tag{18}$$

which is given by substituting $u = \sum_{m=1}^{M}\alpha_m v_m$ with $\boldsymbol{\alpha} \coloneqq [\alpha_1, \ldots, \alpha_M]^{\mathsf{T}}$ in (15). Here, $\boldsymbol{s} \in \mathbb{C}^M$, $\boldsymbol{\Sigma} \in \mathbb{C}^{M\times M}$, and $\boldsymbol{K} \in \mathbb{C}^{M\times M}$ are defined as

$$\boldsymbol{s} \coloneqq [s_1, \ldots, s_M]^{\mathsf{T}}, \tag{19}$$

$$\boldsymbol{\Sigma} \coloneqq \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_M^2 \end{bmatrix}, \tag{20}$$

$$\boldsymbol{K} \coloneqq \begin{bmatrix} K_{1,1} & \ldots & K_{1,M} \\ \vdots & \ddots & \vdots \\ K_{M,1} & \ldots & K_{M,M} \end{bmatrix} \tag{21}$$

with

$$K_{m_1,m_2} \coloneqq \langle v_{m_1}, v_{m_2}\rangle_{\mathscr{H}} = \int_{\boldsymbol{x}\in\mathbb{S}_2} w(\boldsymbol{x})\gamma_{m_1}(\boldsymbol{x})\gamma_{m_2}(\boldsymbol{x})^* \cdot \exp(-ik\boldsymbol{x}\circ(\boldsymbol{r}_{m_1}-\boldsymbol{r}_{m_2}))\,\mathrm{d}\chi \quad (m_1, m_2 \in [\![1,M]\!]). \tag{22}$$

Finally, by solving (18), we obtain

$$\hat{\boldsymbol{\alpha}} = (\boldsymbol{K}+\lambda\boldsymbol{\Sigma})^{-1}\boldsymbol{s}, \tag{23}$$

and $\hat{u}$ is given by substituting (23) into (17). Therefore, the remaining problems are the calculations of $v_1, \ldots, v_M$ and $\boldsymbol{K}$.

C. Directionally Weighted Spherical Wavefunctions and Directionally Weighted Translation Operators

As noted in Section II-C, the directivities of various sensors can be well modeled by finite-order spherical harmonic functions. Suppose $\gamma_1, \ldots, \gamma_M$ can be represented as

$$\gamma_m(\boldsymbol{x})^* = \sum_{\nu,\mu}^{N_m} c_{m,\nu,\mu} Y_{\nu,\mu}(\boldsymbol{x}) \quad (\boldsymbol{x}\in\mathbb{S}_2,\ m\in[\![1,M]\!]) \tag{24}$$

with $N_1, \ldots, N_M \in \mathbb{N}$. Then, by substituting (12) and (24) into (16) and (22) and solving the integrals (see Appendix IX for detailed derivations), we obtain

$$v_m(\boldsymbol{r}) = \sum_{\nu,\mu}^{N_m} c_{m,\nu,\mu}\varphi_{\nu,\mu}(\boldsymbol{r}-\boldsymbol{r}_m) \quad (\boldsymbol{r}\in\Omega,\ m\in[\![1,M]\!]), \tag{25}$$

$$K_{m_1,m_2} = \sum_{\nu_1,\mu_1}^{N_{m_1}}\sum_{\nu_2,\mu_2}^{N_{m_2}} c^*_{m_1,\nu_1,\mu_1}c_{m_2,\nu_2,\mu_2}T^{\nu_2,\mu_2}_{\nu_1,\mu_1}(\boldsymbol{r}_{m_1}-\boldsymbol{r}_{m_2}) \quad (m_1,m_2\in[\![1,M]\!]\}), \tag{26}$$

where

$$\varphi_{\nu,\mu}(\boldsymbol{r}) := \frac{1}{C(\beta)}\xi_{\nu,\mu}(k\boldsymbol{r}+i\beta\boldsymbol{\eta}) \quad (\boldsymbol{r}\in\mathbb{R}^3,\ \nu\in\mathbb{N},\ \mu\in[\![-\nu,\nu]\!]), \tag{27}$$

$$T^{\nu_2,\mu_2}_{\nu_1,\mu_1}(\boldsymbol{r}) := \frac{1}{C(\beta)}\Theta^{\nu_2,\mu_2}_{\nu_1,\mu_1}(k\boldsymbol{r}+i\beta\boldsymbol{\eta}) \quad \boldsymbol{r}\in\mathbb{R}^3,\ \nu_1,\nu_2\in\mathbb{N}, \quad (\mu_1\in[\![-\nu_1,\nu_1]\!],\ \mu_2\in[\![-\nu_2,\nu_2]\!]) \tag{28}$$

with

$$\xi_{\nu,\mu}(\boldsymbol{z}) := \frac{1}{i^\nu}j_\nu\left(\left(z_1^2+z_2^2+z_3^2\right)^{\frac{1}{2}}\right)y_{\nu,\mu}\left(\frac{\boldsymbol{z}}{(z_1^2+z_2^2+z_3^2)^{\frac{1}{2}}}\right) \quad (\boldsymbol{z}\coloneqq[z_1,z_2,z_3]^{\mathsf{T}}\in\mathbb{C}^3,\ \nu\in\mathbb{N},\ \mu\in[\![-\nu,\nu]\!]), \tag{29}$$

$$\Theta^{\nu_2,\mu_2}_{\nu_1,\mu_1}(\boldsymbol{z}) := \sum_{\nu_3,\mu_3}^{\nu_1+\nu_2}\mathcal{G}(\nu_1,\mu_1;\nu_2,\mu_2;\nu_3,\mu_3)\xi_{\nu_3,\mu_3}(\boldsymbol{z}) \quad (\boldsymbol{z}\in\mathbb{C}^3,\ \nu_1,\nu_2\in\mathbb{N}, \quad \mu_1\in[\![-\nu_1,\nu_1]\!],\ \mu_2\in[\![-\nu_2,\nu_2]\!]). \tag{30}$$

Here, $j_\nu(\cdot):\mathbb{C}\to\mathbb{C}$ is the νth-order spherical Bessel function of the first kind, $y_{\nu,\mu}(\cdot):\mathbb{C}^3\to\mathbb{C}$ is the normalized homogeneous harmonic polynomial [21] of order ν and degree μ (i.e., a homogeneous harmonic polynomial satisfying $y_{\nu,\mu}(\boldsymbol{x}) = Y_{\nu,\mu}(\boldsymbol{x})$ for $\boldsymbol{x}\in\mathbb{S}_2$), and $\mathcal{G}(\cdot)$ denotes the Gaunt coefficient with a slight modification regarding complex conjugation, which is defined as

$$\mathcal{G}(\nu_1,\mu_1;\nu_2,\mu_2;\mu_3;\mu_3) \coloneqq \int_{\boldsymbol{x}\in\mathbb{S}_2} Y_{\nu_1,\mu_1}(\boldsymbol{x})^* Y_{\nu_2,\mu_2}(\boldsymbol{x}) Y_{\nu_3,\mu_3}(\boldsymbol{x})^*\,\mathrm{d}\chi \quad (\nu_1,\nu_2,\nu_3\in\mathbb{N},\ \mu_1\in[\![-\nu_1,\nu_1]\!], \quad \mu_2\in[\![-\nu_2,\nu_2]\!],\ \mu_3\in[\![-\nu_3,\nu_3]\!]). \tag{31}$$

The closed-form expression of (31) can be obtained using the formulae in [21], and an efficient computational algorithm is proposed in [34]. Note that (29) is well defined independently from branches of the $1/2$-power function and can also be defined at $\boldsymbol{z}=[z_1,z_2,z_3]^{\mathsf{T}}\in\mathbb{C}^3$ such that $z_1^2+z_2^2+z_3^2=0$ by using limit values because these points are removable singularities

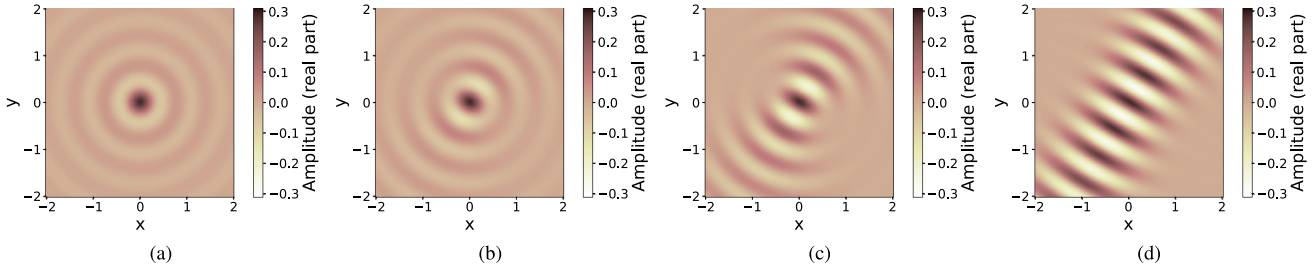


Fig. 2. Directionally weighted spherical wavefunctions of order $\nu = 0$ and degree $\mu = 0$ plotted in xy-plane ($\boldsymbol{r} = [x, y, 0]^\mathsf{T}$). (a) $\beta = 0$; (b) $\beta = 2$; (c) $\beta = 5$; (d) $\beta = 25$.

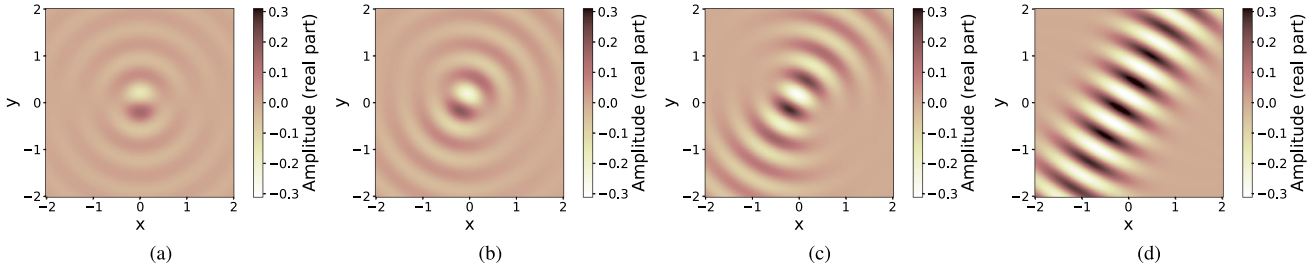


Fig. 3. Directionally weighted spherical wavefunctions of order $\nu = 1$ and degree $\mu = 1$ plotted in xy-plane ($\boldsymbol{r} = [\mathrm{x}, \mathrm{y}, 0]^\mathsf{T}$). (a) $\beta = 0$; (b) $\beta = 2$; (c) $\beta = 5$; (d) $\beta = 25$.

(also note that the term $(z_1^2 + z_2^2 + z_3^2)^{\frac{1}{2}}$ in (29) is not the complex norm given by $(|z_1|^2 + |z_2|^2 + |z_3|^2)^{\frac{1}{2}}$).

When $\beta = 0$, it can be immediately shown that (27) and (28) correspond respectively to the usual spherical wavefunction and translation operator [4], [19], [20] (except for the constant coefficients), and this method corresponds to our previous work [19]. Conversely, the proposed method is obtained formally by replacing the usual spherical wavefunctions and translation operators in [19] with (27) and (28), respectively (therefore, the computational costs are essentially the same in these two methods). Since the differences from the previous work [19] in $\beta \in (0, \infty)$ are derived from the directional weighting function w, we refer to (27) and (28) as a *directionally weighted spherical wavefunction* and a *directionally weighted translation operator*, respectively. We show several examples of directionally weighted spherical wavefunctions in Fig. 2, where $k = 10$ and $\boldsymbol{\eta} = [\cos(\pi/3), \sin(\pi/3), 0]^\mathsf{T}$. One can see that these functions become close to the plane-wave function arriving from $\boldsymbol{\eta}$ as β increases.

As a special case, we consider the case of omnidirectional sensors, i.e., $\gamma_m(\boldsymbol{x}) = 1$ ($\boldsymbol{x} \in \mathbb{S}_2$, $m \in [\![1, M]\!]$). Then, we obtain

$$v_m(\boldsymbol{r}) = \sqrt{4\pi}\varphi_{0,0}(\boldsymbol{r} - \boldsymbol{r}_m) \quad (\boldsymbol{r} \in \Omega,\ m \in [\![1, M]\!]), \tag{32}$$

$$K_{m_1,m_2} = \sqrt{4\pi}\varphi_{0,0}(\boldsymbol{r}_{m_1} - \boldsymbol{r}_{m_2}) \quad (m_1, m_2 \in [\![1, M]\!]). \tag{33}$$

In this case, the proposed method corresponds to the kernel ridge regression [35] with the kernel function $\kappa : \Omega \times \Omega \to \mathbb{C}$ defined as

$$\kappa(\boldsymbol{r}_1, \boldsymbol{r}_2) \coloneqq \sqrt{4\pi}\varphi_{0,0}(\boldsymbol{r}_1 - \boldsymbol{r}_2) \quad (\boldsymbol{r}_1, \boldsymbol{r}_2 \in \Omega), \tag{34}$$

i.e., $(\mathscr{H}, \langle \cdot, \cdot \rangle_{\mathscr{H}})$ is a reproducing kernel Hilbert space [36] generated by κ. Similar discussion for the two-dimensional case is also given in [29].

D. Extension to Multiple Wave Sources

The proposed method can be easily extended to cases of multiple wave sources. The weighting function w can be generalized as

$$w(\boldsymbol{x}) \coloneqq \sum_{l=1}^{L} a_l C(\beta_l) \exp(\beta_l \boldsymbol{\eta}_l \circ \boldsymbol{x}) \quad (\boldsymbol{x} \in \mathbb{S}_2), \tag{35}$$

where $L \in \mathbb{N}_+$, $a_1, \ldots, a_L \in [0, 1]$, $\beta_1, \ldots, \beta_L \in [0, \infty)$, and $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_L \in \mathbb{S}_2$ are the constant parameters satisfying $\sum_{l=1}^{L} a_l = 1$. Then, the directionally weighted spherical wavefunctions and the directionally weighted translation operators are given by

$$\varphi_{\nu,\mu}(\boldsymbol{r}) \coloneqq \sum_{l=1}^{L} \frac{a_l}{C(\beta_l)} \xi_{\nu,\mu}(k\boldsymbol{r} + i\beta_l \boldsymbol{\eta}_l) \quad (\boldsymbol{r} \in \mathbb{R}^3,\ \nu \in \mathbb{N},\ \mu \in [\![-\nu, \nu]\!]), \tag{36}$$

$$T_{\nu_1,\mu_1}^{\nu_2,\mu_2}(\boldsymbol{r}) \coloneqq \sum_{l=1}^{L} \frac{a_l}{C(\beta_l)} \Theta_{\nu_1,\mu_1}^{\nu_2,\mu_2}(k\boldsymbol{r} + i\beta_l \boldsymbol{\eta}_l) \quad (\boldsymbol{r} \in \mathbb{R}^3,\ \nu_1, \nu_2 \in \mathbb{N},\ \mu_1 \in [\![-\nu_1, \nu_1]\!],\ \mu_2 \in [\![-\nu_2, \nu_2]\!]), \tag{37}$$

and $v_1, \ldots, v_M$ and $\boldsymbol{K}$ can be obtained by substituting (36) and (37) into (25) and (26), respectively.

E. Comparison With Non-LTI Methods

From (17) and (23), the estimated wave field $\hat{u}$ is immediately shown to be linear with respect to the observed signal $\boldsymbol{s}$ in the proposed method. Here, for a theoretical evaluation of the computational cost, we consider some application represented as a linear system with input $\hat{u}$ and output $\boldsymbol{\phi} \in \mathbb{C}^E$ in the form of

$$\boldsymbol{\phi} = [\Phi_1(\hat{u}), \ldots, \Phi_E(\hat{u})]^{\mathsf{T}}, \tag{38}$$

where $E \in \mathbb{N}_+$ is the dimension of the output variable and $\Phi_1, \ldots, \Phi_E$ are arbitrary linear functionals. For example, in the visualization of wave fields, $\boldsymbol{\phi}$ is the pointwise values of $\hat{u}$, and in sound field reproduction using loudspeakers/headphones, $\boldsymbol{\phi}$ is the loudspeaker/headphone signals that reproduce a sound field $\hat{u}$ (the details of the reproduction methods are beyond the scope of this paper; however, note that the loudspeaker/headphone signals are obtained linearly with respect to the desired field $\hat{u}$ in most methods [1]–[7]). From (17), we can rewrite (38) as

$$\boldsymbol{\phi} = \boldsymbol{V}\hat{\boldsymbol{\alpha}} \tag{39}$$

with $\boldsymbol{V} \in \mathbb{C}^{E\times M}$ defined as

$$\boldsymbol{V} \coloneqq \begin{bmatrix} \Phi_1(v_1) & \ldots & \Phi_1(v_M) \\ \vdots & \ddots & \vdots \\ \Phi_E(v_1) & \ldots & \Phi_E(v_M) \end{bmatrix}. \tag{40}$$

Therefore, from (23), the relationship between $\boldsymbol{\phi}$ and $\boldsymbol{s}$ is represented as

$$\boldsymbol{\phi} = \boldsymbol{M}\boldsymbol{s}, \tag{41}$$

where $\boldsymbol{M} \in \mathbb{C}^{E\times M}$ is defined as $\boldsymbol{M} \coloneqq \boldsymbol{V}(\boldsymbol{K} + \lambda\boldsymbol{\Sigma})^{-1}$. The calculation of $\boldsymbol{M}$ requires $\mathcal{O}(M^3 + EM^2)$ times of scalar multiplication and addition; however, it can be calculated before we obtain $\boldsymbol{s}$. Therefore, once we have obtained $\boldsymbol{M}$, only $\mathcal{O}(EM)$ times of scalar multiplication and addition are required in the calculation of $\boldsymbol{\phi}$ from a given $\boldsymbol{s}$. Also in the time domain, (41) can be implemented using $E \times M$ LTI filters determined independently of $\boldsymbol{s}$.

For comparison, we provide a brief description of the sparsity-based method based on l_p norm regularization [15], [22]–[24]. In this method, the wave field u is first decomposed into a finite number of basis functions as

$$u = \sum_{n=1}^{N_{\mathrm{basis}}} b_n \psi_n, \tag{42}$$

where $N_{\mathrm{basis}} \in \mathbb{N}_+$ is the number of basis functions (typically set to be greater than M), $\psi_1, \ldots, \psi_{N_{\mathrm{basis}}} : \Omega \to \mathbb{C}$ are basis functions and $b_1, \ldots, b_{N_{\mathrm{basis}}} \in \mathbb{C}$ are unknown coefficients that are expected to be sparse (i.e., only a small number of them have nonzero values). The unknown coefficients $b_1, \ldots, b_{N_{\mathrm{basis}}}$ are determined by solving the following minimization problem:

$$\underset{\boldsymbol{b}\in\mathbb{C}^{N_{\mathrm{basis}}}}{\text{minimize}} \quad \|\boldsymbol{b}\|_p \quad \text{subject to} \quad \boldsymbol{D}\boldsymbol{b} = \boldsymbol{s} \tag{43}$$

with $p \in [0, 2)$, where $p = 0$ or 1 is typically used. Here, $\|\cdot\|_p$ denotes the l_p norm [37],[2] and $\boldsymbol{D} \in \mathbb{C}^{M\times N_{\mathrm{basis}}}$ is a dictionary matrix whose (m, n)th element models the observation of ψ_n by the mth sensor for $m \in [\![1, M]\!]$ and $n \in [\![1, N_{\mathrm{basis}}]\!]$. For $p > 0$, the following formulation is also often used instead of (43):

$$\underset{\boldsymbol{b}\in\mathbb{C}^{N_{\mathrm{basis}}}}{\text{minimize}} \quad \|\boldsymbol{D}\boldsymbol{b} - \boldsymbol{s}\|_2^2 + \lambda\|\boldsymbol{b}\|_p^p, \tag{44}$$

where $\lambda \in (0, \infty)$ is a regularization parameter. By defining $\boldsymbol{\mathcal{B}} : \mathbb{C}^M \to \mathbb{C}^{N_{\mathrm{basis}}}$ as the nonlinear map where $\boldsymbol{\mathcal{B}}(\boldsymbol{s}) \in \mathbb{C}^{N_{\mathrm{basis}}}$ denotes the solution of (43) or (44) for a given $\boldsymbol{s} \in \mathbb{C}^M$ and $\boldsymbol{\Psi} \in \mathbb{C}^{E\times N_{\mathrm{basis}}}$ as

$$\boldsymbol{\Psi} \coloneqq \begin{bmatrix} \Phi_1(\psi_1) & \ldots & \Phi_1(\psi_{N_{\mathrm{basis}}}) \\ \vdots & \ddots & \vdots \\ \Phi_E(\psi_1) & \ldots & \Phi_E(\psi_{N_{\mathrm{basis}}}) \end{bmatrix}, \tag{45}$$

we can represent the relationship between $\boldsymbol{\phi}$ and $\boldsymbol{s}$ as

$$\boldsymbol{\phi} = \boldsymbol{\Psi}\boldsymbol{\mathcal{B}}(\boldsymbol{s}). \tag{46}$$

In this case, we cannot utilize the same precomputation as in (41) since $\boldsymbol{\mathcal{B}}$ is a nonlinear mapping. For the calculation of $\boldsymbol{\mathcal{B}}(\boldsymbol{s})$, i.e., the optimization of (43) or (44), the orthogonal matching pursuit [37], [38], (accelerated) proximal gradient [39], and iteratively reweighted least squares [37], [40] are representative algorithms. In total, the following times of scalar multiplication and addition are required in the calculation of (46).

Orthogonal matching pursuit $(p = 0)$:

$$\mathcal{O}(EK + KMN_{\mathrm{basis}})$$

(Accelerated) proximal gradient $(p = 1)$:

$$\mathcal{O}(EK + K_{\mathrm{iter}}MN_{\mathrm{basis}})$$

Iteratively reweighted least squares $(0 < p < 2)$:

$$\mathcal{O}(EK + K_{\mathrm{iter}}M^2N_{\mathrm{basis}})$$

Here, $K \in \mathbb{N}_+$ is the number of nonzero values in $\boldsymbol{b}$, and $K_{\mathrm{iter}} \in \mathbb{N}_+$ is the number of iterations in the algorithm. Particularly when E is small, these computational costs are much larger than that of the proposed method. Also in the time domain, (46) does not allow a fast implementation using LTI filters in the time domain owing to the nonlinearity of $\boldsymbol{\mathcal{B}}$. Therefore, the proposed method is preferable to the above non-LTI methods from the viewpoint of the computational cost, which is an important factor in applications such as real-time systems, where fast implementation is required.

IV. Numerical Experiments

Numerical experiments of sound field estimation using a microphone array were conducted to demonstrate the performance of the proposed method. The sound field in the air was considered, and the phase velocity (i.e., the speed of sound) was set as $c = 340$ m/s. We compared the proposed method with the current LTI method presented in [18] (also introduced as the

[2]We use the term "norm" according to conventions although it is not a norm in the strict sense for $p \in [0, 1)$.

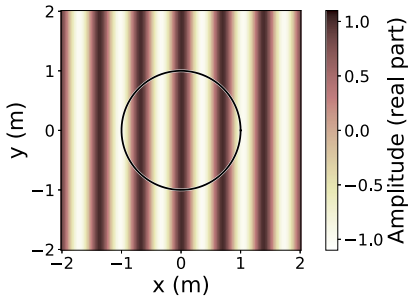


Fig. 4. True sound field under free-field condition at 500 Hz plotted in xy-plane ($\boldsymbol{r} = [\mathrm{x}, \mathrm{y}, 0]^\mathsf{T}$ m).

general sampling array approach in [2]), which is referred to here as the truncation method. In addition, the proposed method for $\beta = 0$, which corresponds to our previous work [19], was also investigated for comparison. Hereafter, the notation of temporal frequency is used instead of angular frequency.

A. Estimation of Plane-Wave Field Under Free-Field Condition

In a three-dimensional free field, 64 microphones were located on a sphere with a radius of 1.0 m centered at the origin. Their positions were determined according to the spherical t-design [41]. Each microphone was modeled as a cardioid microphone oriented outward, i.e., the observed signals $s_1, \ldots, s_M$ in a sound field u were given by

$$s_m = \frac{1}{2}u(\boldsymbol{r}_m) - \frac{1}{2ik}\frac{\partial}{\partial \boldsymbol{y}_m}u(\boldsymbol{r}_m) + \epsilon_m \quad (m \in [\![1, M]\!]). \tag{47}$$

Here, for each $m \in [\![1, M]\!]$, $\boldsymbol{y}_m$ denotes the outward unit normal vector on the sphere at $\boldsymbol{r}_m$, $\partial/\partial \boldsymbol{y}_m$ denotes the directional derivative along the direction $\boldsymbol{y}_m$, and ϵ_m denotes the observation noise of the mth microphone. For $u \in \mathscr{H}$, these microphones can be equivalently modeled by the following directivities:

$$\gamma_m(\boldsymbol{x}) = \frac{1}{2} + \frac{1}{2}\boldsymbol{y}_m \circ \boldsymbol{x} \quad (\boldsymbol{x} \in \mathbb{S}_2,\ m \in [\![1, M]\!]). \tag{48}$$

Furthermore, these directivities $\gamma_1, \ldots, \gamma_M$ can be represented as (24) with

$$c_{m,\nu,\mu} = \begin{cases} \sqrt{\pi} & (\nu = 0) \\ \frac{2\pi}{3} Y_{1,\mu}(\boldsymbol{y}_m)^* & (\nu = 1) \\ 0 & (\nu \geq 2) \end{cases} \quad (m \in [\![1, M]\!]). \tag{49}$$

The true sound field u_{true} was set as a single plane wave, which was defined as

$$u_{\mathrm{true}}(\boldsymbol{r}) \coloneqq \exp(-ik\boldsymbol{x}_{\mathrm{inc}} \circ \boldsymbol{r}) \quad (\boldsymbol{r} \in \mathbb{R}^3) \tag{50}$$

with $\boldsymbol{x}_{\mathrm{inc}} = [1, 0, 0]^\mathsf{T}$ (note that $\boldsymbol{x}_{\mathrm{inc}}$ denotes the incident direction, not traveling direction). An example of the true sound field is shown in Fig. 4, where the black line denotes the boundary of the microphone array. The observed signals were calculated using (47), and the observation noises were sampled independently from the circularly symmetric Gaussian distribution with zero mean and variance of $10^{-2} \times S$, where S denotes the average power of the noise-free signals (i.e., the signal-to-noise ratio was 20 dB).

In the proposed method, $\sigma_1, \ldots, \sigma_N$ were set as 1, and λ was set as 10^{-2}. For prior information, we used the weighting function defined in (12). Here, we defined $\boldsymbol{\eta} \coloneqq [\cos\theta, \sin\theta, 0]^\mathsf{T}$

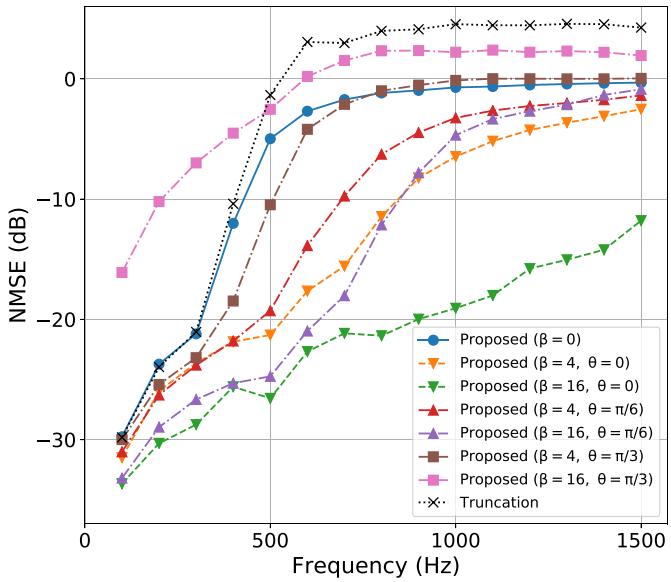


Fig. 5. Normalized mean squared error (NMSE) plotted against frequency under free-field condition.

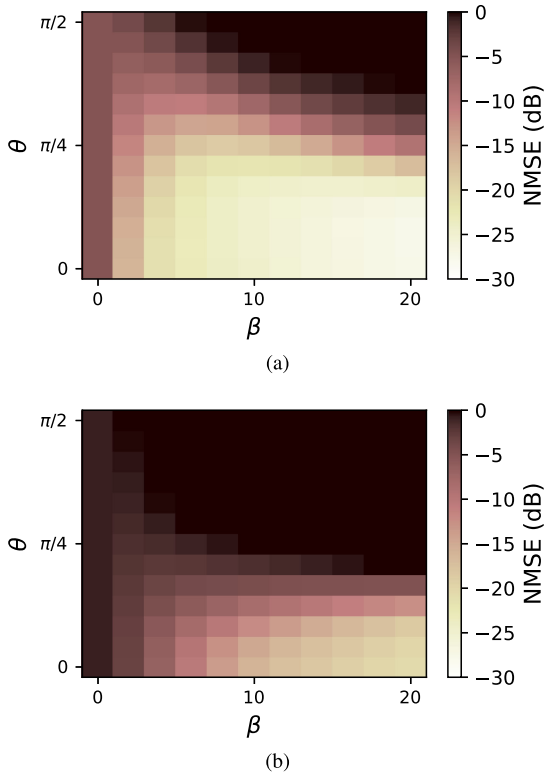


Fig. 6. Relationship between parameters and normalized mean squared error (NMSE) under free-field condition. (a) 500 Hz; (b) 1000 Hz.

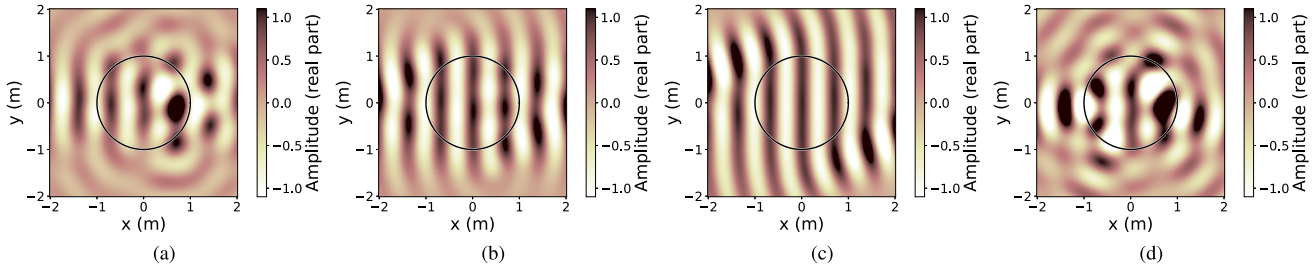


Fig. 7. Estimated sound fields under free-field condition at 500 Hz plotted in xy-plane ($\boldsymbol{r} = [\mathrm{x}, \mathrm{y}, 0]^{\mathsf{T}}$ m). (a) proposed method ($\beta = 0$); (b) proposed method ($\beta = 4$, $\theta = \pi/6$); (c) proposed method ($\beta = 16$, $\theta = \pi/6$); (d) truncation method.

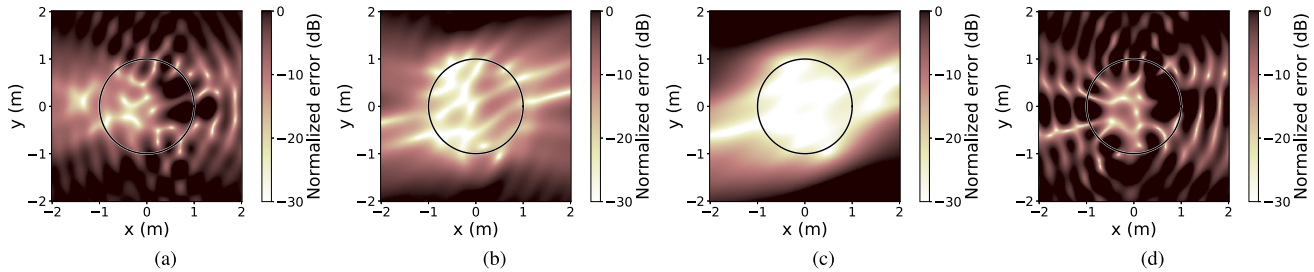


Fig. 8. Normalized error distributions under free-field condition at 500 Hz plotted in xy-plane ($\boldsymbol{r} = [\mathrm{x}, \mathrm{y}, 0]^{\mathsf{T}}$ m). (a) proposed method ($\beta = 0$); (b) proposed method ($\beta = 4$, $\theta = \pi/6$); (c) proposed method ($\beta = 16$, $\theta = \pi/6$); (d) truncation method.

for $\theta \in [0, \pi/2]$ ($\theta = 0$ means the accurate prior information and a large θ means an inaccurate one), and different values of β and θ were investigated.

In the truncation method [18], the truncation order was determined as 7 so that the number of unknown coefficients corresponds to the number of microphones, i.e., 64. The global origin (center of the spherical wave function expansion) was set at the center of the spherical microphone array. The regularization parameter in the matrix inversion was set at 10^{-2} as in the proposed method.

As an evaluation criterion, the normalized mean squared error (NMSE) was used, which was defined as

$$\mathrm{NMSE} \coloneqq 10 \log_{10} \frac{\sum_{\mathrm{i} \in \mathrm{I}_{\mathrm{eval}}} |u_{\mathrm{true}}(\boldsymbol{r}_{\mathrm{eval}}^{(\mathrm{i})}) - u_{\mathrm{est}}(\boldsymbol{r}_{\mathrm{eval}}^{(\mathrm{i})})|^2}{\sum_{\mathrm{i} \in \mathrm{I}_{\mathrm{eval}}} |u_{\mathrm{true}}(\boldsymbol{r}_{\mathrm{eval}}^{(\mathrm{i})})|^2} \ (\mathrm{dB}). \tag{51}$$

Here, u_{est} denotes the estimated sound field, and the evaluation points $\{\boldsymbol{r}_{\mathrm{eval}}^{(\mathrm{i})}\}_{\mathrm{i} \in \mathrm{I}_{\mathrm{eval}}}$ were set as all grid points with an interval of 0.1 m on and inside the surface of the spherical microphone array.

First, the relationship between frequencies and NMSEs was plotted in Fig. 5. We can see that the NMSEs for the proposed method ($\beta = 0$) were almost the same as those for the truncation method at low frequencies and lower than those for the truncation method at high frequencies, which was also reported in [19] for the two-dimensional case. Among the proposed methods, the NMSEs for $\theta = 0$ were lower than those for the other conditions, and even the NMSEs for $\theta = \pi/6$ (i.e., inaccurate prior information) were lower than those for $\beta = 0$. For further investigation, the NMSEs at different θ and β were plotted at frequencies of 500 and 1000 Hz in Fig. 6. It can be seen that for $\theta = 0$, the NMSE decreased as β increased, and for $\theta > 0$, the NMSE took a minimum value at a certain value of β. This is considered to be because the estimation was strongly affected by inaccurate prior information in cases of large β for $\theta > 0$. The best value of β varied depending on θ and the frequency, and their quantitative relationship seems complex; however, at least, many conditions achieved lower NMSEs than the condition of $\beta = 0$. These results indicate that even rough prior information on the source direction may improve the estimation accuracy in the proposed method. We also show an example of the estimated sound fields and the (pointwise) normalized error distributions at 500 Hz in Figs. 7 and 8, respectively. In this example, the NMSEs were (a) -4.87, (b) -18.20, (c) -24.74, and (d) -1.18 dB. The tendencies described above can also be seen in these results.

B. Estimation of Monopole Field Under Free-Field and Reverberant Conditions

In a $6\,\mathrm{m} \times 4\,\mathrm{m} \times 3\,\mathrm{m}$ rectangular room with its center defined as the origin, the same spherical microphone array as used in the previous experiments was located with its center positioned at $[-1, 0, 0]^{\mathsf{T}}$ m. The reverberation in the room was simulated by the image-source method [42], where image sources were considered up to the 20th reflection order. Here, the reflection coefficients were set as $\Gamma \in \{0, 0.4, 0.8\}$ for all six wall surfaces, where each of the above three values was investigated ($\Gamma = 0$ corresponds to the free-field condition). In these settings, the reverberation time (from Sabine's formula) and the reverberation radius (critical distance) [43] were respectively $0.13\,\mathrm{s}$ and $1.35\,\mathrm{m}$ for $\Gamma = 0.4$ and 0.30 s and 0.88 m for $\Gamma = 0.8$.

The true sound field u_{true} was set as a superposition of two monopole functions, whose direct wave component was defined

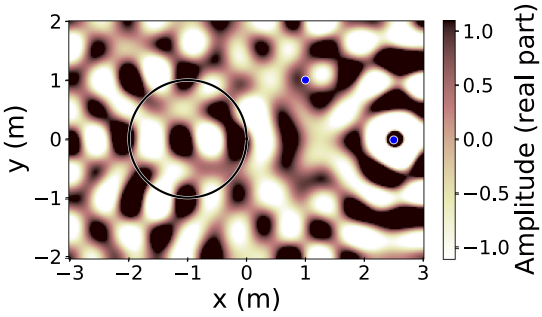
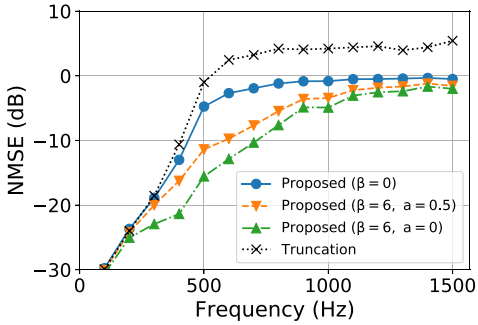
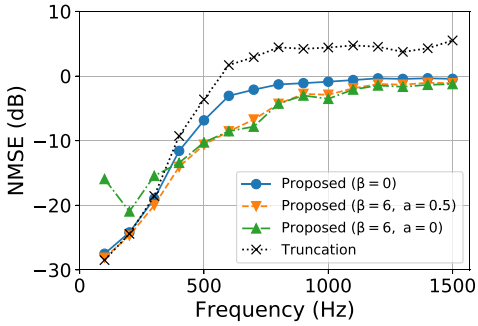


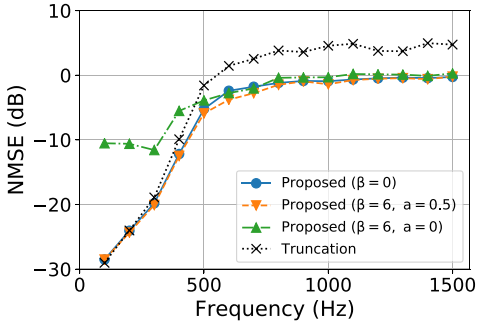
Fig. 9. True sound field under reverberant condition ($\Gamma = 0.8$) at 500 Hz plotted in xy-plane ($\boldsymbol{r} = [\mathrm{x}, \mathrm{y}, 0]^{\mathsf{T}}$ m).



(a)



(b)



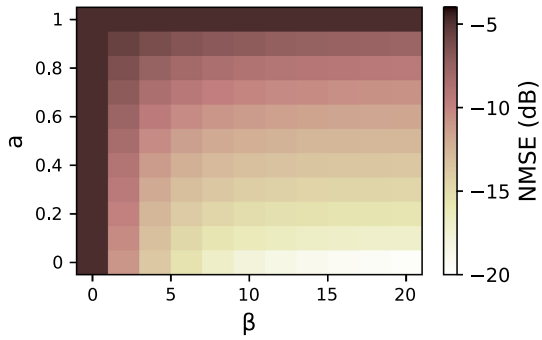
(c)

Fig. 10. Normalized mean squared error (NMSE) plotted against frequency. (a) $\Gamma = 0$; (b) $\Gamma = 0.4$; (c) $\Gamma = 0.8$.

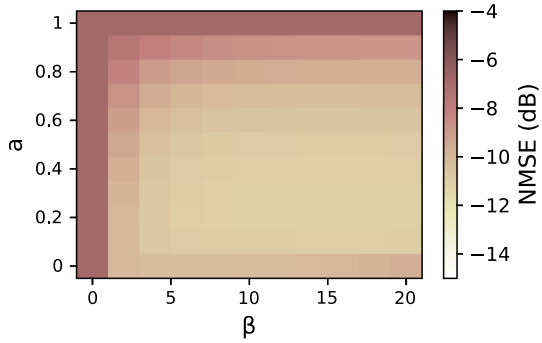
as

$$u_{\mathrm{true}}(\boldsymbol{r}) = \sum_{q=1}^{2} A^{(q)} \frac{\exp(ik\|\boldsymbol{r} - \boldsymbol{r}_{\mathrm{src}}^{(q)}\|_2)}{4\pi\|\boldsymbol{r} - \boldsymbol{r}_{\mathrm{src}}^{(q)}\|_2} \quad (\boldsymbol{r} \in \mathbb{R}^3) \tag{52}$$

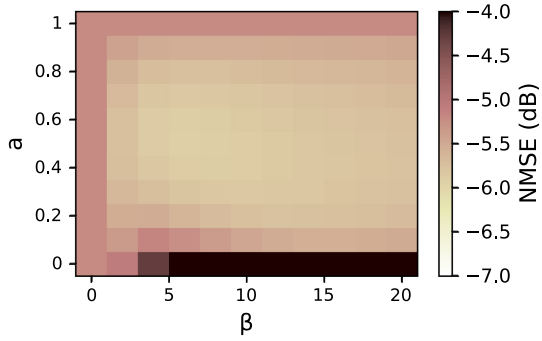
with $A^{(1)} = 15$ m, $\boldsymbol{r}_{\mathrm{scr}}^{(1)} = [2.5, 0, 0]^{\mathsf{T}}$ m, $A^{(2)} = 10i$ m, and $\boldsymbol{r}_{\mathrm{scr}}^{(1)} = [1, 1, 1]^{\mathsf{T}}$ m. An example of the true sound field is shown



(a)



(b)



(c)

Fig. 11. Relationship between parameters and normalized mean squared error (NMSE). (a) $\Gamma = 0$; (b) $\Gamma = 0.4$; (c) $\Gamma = 0.8$.

in Fig. 9, where the black line denotes the boundary of the microphone array and the blue dots denote the position of the sound sources projected into the xy-plane. The observed signals were calculated using (47), and the observation noises were added in the same way as in the previous experiments.

In the proposed method, $\sigma_1, \ldots, \sigma_N$ and λ were set to be the same as in the previous experiments. For prior information, we used the weighting function defined in (35) with $L = 3$, where we defined

$$\beta_1 = \beta_2 = \beta, \ \beta_3 = 0, \tag{53}$$

$$\boldsymbol{\eta}_1 = [1, 0, 0]^{\mathsf{T}}, \ \boldsymbol{\eta}_2 = \left[\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right]^{\mathsf{T}}, \tag{54}$$

$$a_1 = a_2 = \frac{1-a}{2}, \ a_3 = a. \tag{55}$$

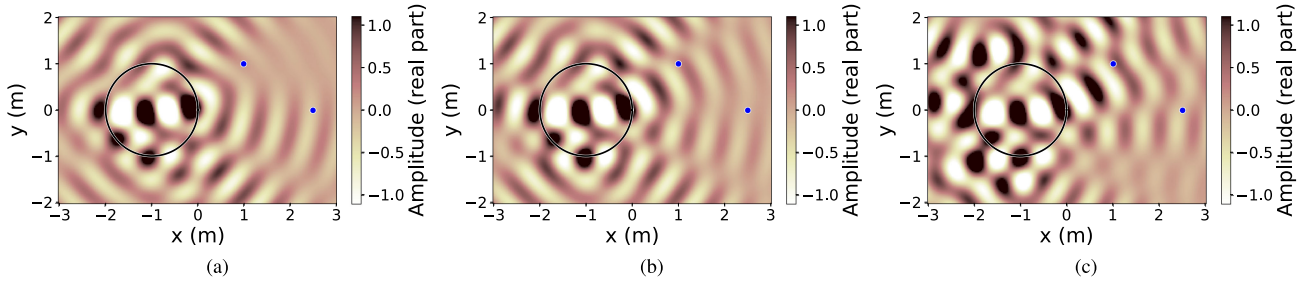


Fig. 12. Estimated sound fields under reverberant condition ($\Gamma = 0.8$) at 500 Hz plotted in xy-plane ($\boldsymbol{r} = [\mathrm{x}, \mathrm{y}, 0]^{\mathsf{T}}$ m). (a) proposed method ($\beta = 0$); (b) proposed method ($\beta = 6$, $a = 0$); (c) proposed method ($\beta = 6$, $a = 0.5$).

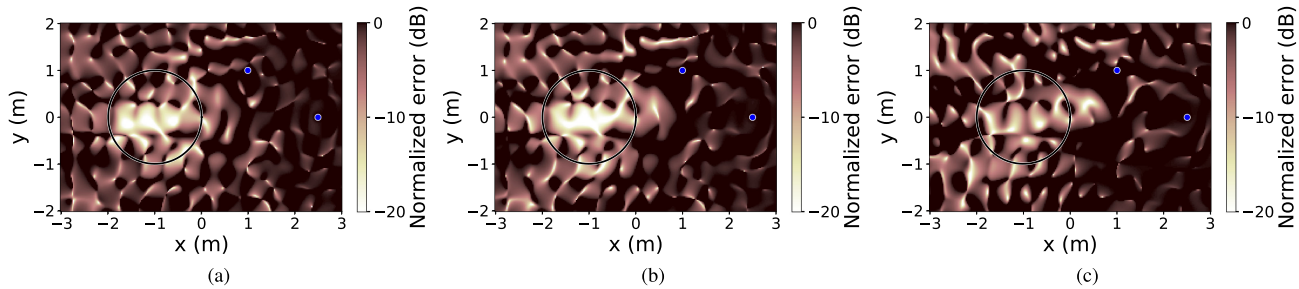


Fig. 13. Normalized error distributions under reverberant condition ($\Gamma = 0.8$) at 500 Hz plotted in xy-plane ($\boldsymbol{r} = [\mathrm{x}, \mathrm{y}, 0]^{\mathsf{T}}$ m). (a) proposed method ($\beta = 0$); (b) proposed method ($\beta = 4$, $a = 0.5$); (c) proposed method ($\beta = 4$, $a = 0$).

using the parameters $\beta \in [0, \infty)$ and $a \in [0, 1]$. We defined $\beta_3 = 0$ for the third weighting to represent the reverberant component, and different β and a were investigated. In the truncation method, the same parameters as in the previous experiments were used. The NMSE was used as an evaluation criterion.

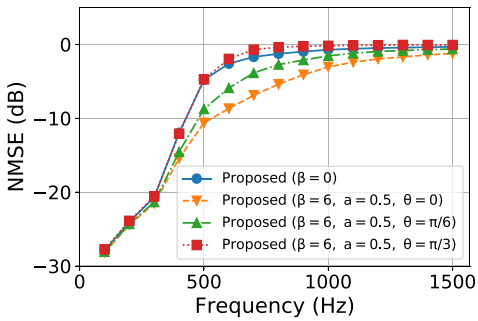
First, the relationship between frequencies and NMSEs was plotted in Fig. 10. In the free-field condition, i.e., $\Gamma = 0$, one can see similar tendencies to the experimental results in Section IV-A. In the reverberant conditions, the NMSEs for the proposed method ($\beta = 6$, $a = 0$) became large as Γ increased, which was because the regularization term for $a = 0$ takes a large value for reverberant components. On the other hand, the proposed method ($\beta = 6$, $a = 0.5$) achieved the lower NMSEs than the proposed method ($\beta = 0$) at most frequencies although the performance improvement decreased as Γ increased. The reduced effectiveness is considered to be related to the extent of reverberation; in highly reverberant environments, even sound fields originated from monopole sources become close to diffuse fields, which makes it difficult to improve estimation performance by using a directional weighting.

For further investigation, the NMSEs for different a and β at 500 Hz were plotted in Fig. 11. Also in this case, most conditions achieved lower NMSEs than $\beta = 0$, which indicates again that even rough prior information may contribute to the improvement of estimation performance. We also show an example of the estimated sound fields and normalized error distributions for $\Gamma = 0.8$ at 500 Hz in Figs. 12 and 13, respectively. In this example, the NMSEs were (a) -5.22, (b) -5.90, and (c) -3.90 dB. The tendencies described above can also be seen in these results.

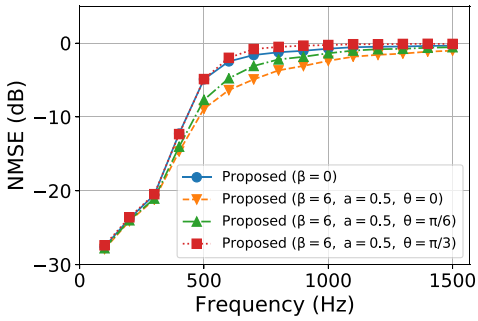
Finally, to investigate how the above results generalize, we conducted the same evaluations for several different source positions, i.e., $\boldsymbol{r}_{\mathrm{src}}^{(1)}, \boldsymbol{r}_{\mathrm{src}}^{(2)}$, and several accuracies of prior information, i.e., $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2$. The source positions were sampled randomly according to the uniform distribution on the entire room excluding the ball with a radius of 2.0 m centered at $[-1, 0, 0]^{\mathsf{T}}$ m. The directions $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2$ were sampled randomly so that the angle between $\boldsymbol{\eta}_l$ and the true lth source's direction from the center of the spherical microphone array was $\theta \in \{0, \pi/6, \pi/3\}$ for each $l \in \{1, 2\}$, where each of the above three values of θ was investigated. Figure 14 shows the relationship between NMSEs averaged over 20 trials and frequencies for different values of a, β, and Γ. For $\theta = 0$ (i.e., accurate prior information), one can see that the performance improvement using the directional weighting can be generalized in various settings of source positions. Moreover, the proposed method ($\beta = 6$, $a = 0.5$, $\theta = \pi/6$) exhibited lower NMSEs than the proposed method ($\beta = 0$). For $\Gamma = 0.8$, the NMSEs for these two conditions were very close; however, the proposed method ($\beta = 6, a = 0.5, \theta = \pi/6$) showed slightly lower NMSEs at most frequencies. These results mean a certain degree of inaccurate prior information can be used in the proposed method.

V. Conclusion

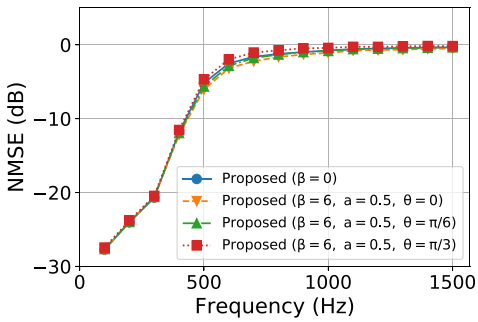
We proposed a wave field estimation method exploiting prior information of source directions by introducing a directional weighting function. The closed-form solution was obtained using the directionally weighted spherical wavefunctions and



(a)



(b)



(c)

Fig. 14. Relationship between parameters and normalized mean squared error (NMSE) averaged over 20 trials. (a) $\Gamma = 0$; (b) $\Gamma = 0.4$; (c) $\Gamma = 0.8$.

directionally weighted translation operators. In the numerical experiments, we confirmed that a higher estimation accuracy can be achieved by incorporating prior information in the proposed method than by using other existing methods that do not use any prior information.

Appendix A
Approximation by Plane-Wave Functions

We show that for any solution u of (1), bounded closed set $\Omega_K \subset \Omega$, and positive number ϵ, there is a function $u_{\mathrm{aprox}} \in \mathscr{H}$ satisfying

$$|u(\boldsymbol{r}) - u_{\mathrm{approx}}(\boldsymbol{r})| < \epsilon \quad \forall \boldsymbol{r} \in \Omega_K. \tag{56}$$

First, from the boundedness of Ω_K, there is an open ball $B \supset \Omega_K$ centered at the origin. For such B, from the Lax–Malgrange theorem [44], [45] (note that the Helmholtz equation is an elliptic partial differential equation), there is a function $f : B \to \mathbb{C}$ satisfying $(\Delta + k^2)f = 0$ and

$$|u(\boldsymbol{r}) - f(\boldsymbol{r})| < \epsilon/2 \quad \forall \boldsymbol{r} \in \Omega_K. \tag{57}$$

Moreover, from [46], f admits the spherical wavefunction expansion as

$$f(\boldsymbol{r}) = \sum_{\nu,\mu} \mathring{f}_{\nu,\mu} \xi_{\nu,\mu}(k\boldsymbol{r}) \quad (\boldsymbol{r} \in B), \tag{58}$$

where $\sum_{\nu,\mu}$ is the abbreviated form of $\sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu}$. Here, for $\nu \in \mathbb{N}$ and $\mu \in [\![-\nu, \nu]\!]$, $\mathring{f}_{\nu,\mu} \in \mathbb{C}$ denotes the expansion coefficients, and $\xi_{\nu,\mu}(\cdot) : \mathbb{C}^3 \to \mathbb{C}$ is defined in (29). This series converges uniformly on Ω_K; therefore, there is some $N \in \mathbb{N}$ satisfying

$$|f(\boldsymbol{r}) - f_N(\boldsymbol{r})| < \epsilon/2 \quad \forall \boldsymbol{r} \in \Omega_K, \tag{59}$$

where $f_N : \Omega \to \mathbb{C}$ is defined as

$$f_N(\boldsymbol{r}) \coloneqq \sum_{\nu,\mu}^{N} \mathring{f}_{\nu,\mu} \xi_{\nu,\mu}(k\boldsymbol{r}) \quad (\boldsymbol{r} \in \Omega). \tag{60}$$

From (57) and (59), we obtain

$$|u(\boldsymbol{r}) - f_N(\boldsymbol{r})| < \epsilon \quad \forall \boldsymbol{r} \in \Omega_K. \tag{61}$$

On the other hand, from the equality (see Appendix IX)

$$\xi_{\nu,\mu}(k\boldsymbol{r}) = \frac{1}{4\pi} \int_{\boldsymbol{x} \in \mathbb{S}_2} \exp(-ik\boldsymbol{r} \circ \boldsymbol{x}) Y_{\nu,\mu}(\boldsymbol{x}) \, \mathrm{d}\chi, \tag{62}$$

f_N is shown to be in $\mathscr{H}$. Therefore, by taking $u_{\mathrm{approx}} \coloneqq f_N$, we obtain (56).

Appendix B
Derivation of Directionally Weighted Spherical Wavefunctions and Directionally Weighted Translation Operators

First, we prove the following equality:

$$\xi_{\nu,\mu}(\boldsymbol{z}) = \frac{1}{4\pi} \int_{\boldsymbol{x} \in \mathbb{S}_2} \exp(-i\boldsymbol{z} \circ \boldsymbol{x}) Y_{\nu,\mu}(\boldsymbol{x}) \, \mathrm{d}\chi \quad (\boldsymbol{z} \in \mathbb{C}^3,\ \nu \in \mathbb{N},\ \mu \in [\![-\nu, \nu]\!]). \tag{63}$$

Let $\Xi(\cdot) : \mathbb{C}^3 \to \mathbb{C}$ be defined as

$$\Xi(\boldsymbol{z}) \coloneqq \frac{1}{4\pi} \int_{\boldsymbol{x} \in \mathbb{S}_2} \exp(-i\boldsymbol{z} \circ \boldsymbol{x}) \, \mathrm{d}\chi \quad (\boldsymbol{z} \in \mathbb{C}^3). \tag{64}$$

Since $\exp(-i\boldsymbol{z} \circ \boldsymbol{x})$ can be represented as a convergent power series with respect to $\boldsymbol{x}$, we can apply the integral formula in [47] and obtain the following equality:

$$\Xi(\boldsymbol{z}) = \sum_{n=0}^{\infty} \frac{(-1)^n n!}{(2n+1)!} (z_1^2 + z_2^2 + z_3^2)^n = j_0\left((z_1^2 + z_2^2 + z_3^2)^{\frac{1}{2}}\right) \quad (\boldsymbol{z} \coloneqq [z_1, z_2, z_3]^{\mathsf{T}} \in \mathbb{C}^3). \tag{65}$$

Moreover, let $\mathcal{Y}_{\nu,\mu}$ be differential operators obtained by replacing variables of the polynomials $y_{\nu,\mu}$ formally with corresponding partial differentials. Then, from (64), we obtain

$$(\mathcal{Y}_{\nu,\mu}\Xi)(\boldsymbol{z}) = \frac{(-i)^\nu}{4\pi}\int_{\boldsymbol{x}\in\mathbb{S}_2}\exp(-i\boldsymbol{z}\circ\boldsymbol{x})Y_{\nu,\mu}(\boldsymbol{x})\,\mathrm{d}\chi \quad (\boldsymbol{z}\in\mathbb{C}^3,\ \nu\in\mathbb{N},\ \mu\in[\![-\nu,\nu]\!]). \tag{66}$$

On the other hand, from Hobson's theorem [21] and (65), the following equality holds:

$$\begin{aligned}(\mathcal{Y}_{\nu,\mu}\Xi)(\boldsymbol{z}) &= \left[\left(\frac{1}{z}\frac{\mathrm{d}}{\mathrm{d}z}\right)^\nu j_0(z)\bigg|_{z=\left(z_1^2+z_2^2+z_3^2\right)^{\frac{1}{2}}}\right] y_{\nu,\mu}(\boldsymbol{z})\\ &= \frac{(-1)^\nu}{\left(z_1^2+z_2^2+z_3^2\right)^{\frac{\nu}{2}}} j_\nu\left(\left(z_1^2+z_2^2+z_3^2\right)^{\frac{1}{2}}\right) y_{\nu,\mu}(\boldsymbol{z})\\ &= (-1)^\nu j_\nu\left(\left(z_1^2+z_2^2+z_3^2\right)^{\frac{1}{2}}\right)\\ &\quad\cdot y_{\nu,\mu}\left(\frac{\boldsymbol{z}}{\left(z_1^2+z_2^2+z_3^2\right)^{\frac{1}{2}}}\right)\\ &\quad(\boldsymbol{z}\coloneqq[z_1,z_2,z_3]^{\mathsf{T}}\in\mathbb{C}^3,\ \nu\in\mathbb{N},\ \mu\in[\![-\nu,\nu]\!]).\end{aligned} \tag{67}$$

From (66) and (67), we obtain (63).

Next, we prove the following equality:

$$\Theta_{\nu_1,\mu_1}^{\nu_2,\mu_2}(\boldsymbol{z}) = \frac{1}{4\pi}\int_{\boldsymbol{x}\in\mathbb{S}_2}\exp(-i\boldsymbol{z}\circ\boldsymbol{x})Y_{\nu_1,\mu_1}(\boldsymbol{x})^*Y_{\nu_2,\mu_2}(\boldsymbol{x})\,\mathrm{d}\chi \quad (\boldsymbol{z}\in\mathbb{C}^3,\ \nu_1,\nu_2\in\mathbb{N},\ \mu_1\in[\![-\nu_1,\nu_1]\!],\ \mu_2\in[\![-\nu_2,\nu_2]\!]). \tag{68}$$

From the linearization formula of the spherical harmonic functions [21], we obtain

$$Y_{\nu_1,\mu_1}(\boldsymbol{x})^*Y_{\nu_2,\mu_2}(\boldsymbol{x}) = \sum_{\nu_3,\mu_3}^{\nu_1+\nu_2}\mathcal{G}(\nu_1,\mu_1;\nu_2,\mu_2;\nu_3,\mu_3)Y_{\nu_3,\mu_3}(\boldsymbol{x}) \quad (\boldsymbol{x}\in\mathbb{S}_2,\ \nu_1,\nu_2\in\mathbb{N},\ \mu_1\in[\![-\nu_1,\nu_1]\!],\ \mu_2\in[\![-\nu_2,\nu_2]\!]). \tag{69}$$

Using this formula and (63), we obtain (68) as

$$\begin{aligned}&\frac{1}{4\pi}\int_{\boldsymbol{x}\in\mathbb{S}_2}\exp(-i\boldsymbol{z}\circ\boldsymbol{x})Y_{\nu_1,\mu_1}(\boldsymbol{x})^*Y_{\nu_2,\mu_2}(\boldsymbol{x})\,\mathrm{d}\chi\\ &= \sum_{\nu_3,\mu_3}^{\nu_1+\nu_2}\mathcal{G}(\nu_1,\mu_1;\nu_2,\mu_2;\nu_3,\mu_3)\\ &\quad\cdot\frac{1}{4\pi}\int_{\boldsymbol{x}\in\mathbb{S}_2}\exp(-i\boldsymbol{z}\circ\boldsymbol{x})Y_{\nu_3,\mu_3}(\boldsymbol{x})\,\mathrm{d}\chi\\ &= \sum_{\nu_3,\mu_3}^{\nu_1+\nu_2}\mathcal{G}(\nu_1,\mu_1;\nu_2,\mu_2;\nu_3,\mu_3)\xi_{\nu_3,\mu_3}(\boldsymbol{z})\\ &\quad(\boldsymbol{z}\in\mathbb{C}^3,\ \nu_1,\nu_2\in\mathbb{N},\ \mu_1\in[\![-\nu_1,\nu_1]\!],\ \mu_2\in[\![-\nu_2,\nu_2]\!]).\end{aligned} \tag{70}$$

REFERENCES

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

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

[3] J. Ahrens and S. Spors, "An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions," *Acta Acust. United Acust.*, vol. 94, pp. 988–999, 2008.

[4] N. Ueno, S. Koyama, and H. Saruwatari, "Three-dimensional sound field reproduction based on weighted mode-matching method," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 27, no. 12, pp. 1852–1867, Dec. 2019.

[5] J. Daniel, "Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonics format," in *Proc. Audio Eng. Soc. Int. Conf.*, May 2003.

[6] R. Duraiwami, D. N. Zotkin, Z. Li, E. Grassi, N. A. Gumerov, and L. S. Davis, "High order spatial audio capture and its binaural head-tracked playback over headphones with HRTF cues," in *Proc. Audio Eng. Soc. Conv.*, New York, Oct. 2005.

[7] D. Menzies and M. Al-Akaidi, "Nearfield binaural synthesis and ambisonics," *J. Acoust. Soc. Amer.*, vol. 121, no. 3, pp. 1559–1563, 2007.

[8] J. Zhang, T. D. Abhayapala, W. Zhang, P. N. Samarasinghe, and S. Jiang, "Active noise control over space: A wave domain approach," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 4, pp. 774–786, Apr. 2018.

[9] Y. Maeno, Y. Mitsufuji, P. N. Samarasinghe, N. Murata, and T. D. Abhayapala, "Spherical-harmonic-domain feedforward active noise control using sparse decomposition of reference signals from distributed sensor arrays," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 656–670,, Jan. 2020.

[10] H. Ito, S. Koyama, N. Ueno, and H. Saruwatari, "Feedforward spatial active noise control based on kernel interpolation of sound field," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2019, pp. 466–470.

[11] Y. Takida, S. Koyama, N. Ueno, and H. Saruwatari, "Reciprocity gap functional in spherical harmonic domain for gridless sound field decomposition," *Signal Process.*, vol. 169, 2020, Art. no. 107383.

[12] A. E. Badia and T. Nara, "An inverse source problem for Helmholtz's equation from the Cauchy data with a single wave number," *Inverse Prob.*, vol. 27, no. 10, 2011, Art. no. 105001.

[13] Z. Dogan, V. Tsiminaki, I. Jovanovic, T. Blu, and D. Van DeVille, "Localization of point sources for systems governed by the wave equation," in *Proc. Wavelets Sparsity XIV*, Aug. 2011, Art. no. 81380P.

[14] Z. Dogan, I. Jovanovic, T. Blu, and D. Van DeVille, "3D reconstruction of wave-propagated point sources from boundary measurements using joint sparsity and finite rate of innovation," in *Proc. 9th IEEE Int. Symp. Biomed. Imag.*, May 2012, pp. 1575–1578.

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

[16] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthogonal decomposition of the soundfield," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2002, pp. II-1781–1784.

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

[18] A. Laborie, R. Bruno, and S. Montoya, "A new comprehensive approach of surround sound recording," in *Proc. Audio Eng. Soc. Conv.*, Oct. 2003.

[19] N. Ueno, S. Koyama, and H. Saruwatari, "Sound field recording using distributed microphones based on harmonic analysis of infinite order," *IEEE Signal Process. Lett.*, vol. 25, no. 1, pp. 135–139, Jan. 2018.

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

[21] P. A. Martin, *Multiple Scattering: Interaction of Time-Harmonic Waves With and N Obstacles*. New York, NY, USA: Cambridge Univ. Press, 2006.

[22] A. Wabnitz, N. Epain, A. McEwan, and C. Jin, "Upscaling ambisonic sound scenes using compressed sensing techniques," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, Oct. 2011, pp. 1–4.

[23] D. Y. Hu, H. B. Li, Y. Hu, and Y. Fang, "Sound field reconstruction with sparse sampling and the equivalent source method," *Mech. Syst., Signal Process.*, vol. 108, pp. 317–325, 2018.

[24] N. Murata, S. Koyama, N. Takamune, and H. Saruwatari, "Sparse representation using multidimensional mixed-norm penalty with application to sound field decomposition," *IEEE Trans. Signal Process.*, vol. 66, no. 12, pp. 3327–3338, Jun. 2018.
[25] R. A. Kennedy and P. Sadeghi, *Hilbert Space Methods in Signal Processing*. New York, NY, USA: Cambridge Univ. Press, 2013.
[26] B. Schölkopf, R. Herbrich, and A. J. Smola, "A generalized representer theorem," in *Proc. Comput. Learn. Theory*, Jul. 2001, pp. 416–426.
[27] F. Dinuzzo and B. Schölkopf, "The representer theorem for hilbert spaces: A necessary and sufficient condition," in *Proc. Adv. Neural Inf. Process. Syst.*, Dec. 2012, pp. 189–196.
[28] N. Ueno, S. Koyama, and H. Saruwatari, "Kernel ridge regression with constraint of helmholtz equation for sound field interpolation," in *Proc. Int. Workshop Acoust. Signal Enhancement*, Sep. 2018, pp. 436–440.
[29] H. Ito, S. Koyama, N. Ueno, and H. Saruwatari, "Spatial active noise control based on kernel interpolation with directional weighting," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2020, pp. 8399–8403.
[30] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. London, U.K.: Academic Press, 1999.
[31] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*. New York, NY, USA: Springer, 2013.
[32] G. Masters and K. Richards-Dinger, "On the efficient calculation of ordinary and generalized spherical harmonics," *Geophys. J. Int.*, vol. 135, no. 1, pp. 307–309, 1998.
[33] K. V. Mardia and P. E. Jupp, *Directional Statistics*. Chichester, U.K.: Wiley, 2009.
[34] Y. L. Xu, "Fast evaluation of gaunt coefficients: Recursive approach," *J. Comput. Appl. Math.*, vol. 85, no. 1, pp. 53–65, 1997.
[35] T. Hofmann, B. Schölkopf, and A. J. Smola, "Kernel methods in machine learning," *Ann. Stat.*, vol. 36, no. 3, pp. 1171–1220, 2008.
[36] N. Aronszajn, "Theory of reproducing kernels," *Trans. Am Math. Soc.*, vol. 68, no. 3, pp. 337–404, 1950.
[37] M. Elad, *Sparse and Redundant Representations*. New York, NY, USA: Springer, 2010.
[38] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[39] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
[40] M. A. T. Figueiredo, J. M. Bioucas-Dias, and R. D. Nowak, "Majorization-minimization algorithms for wavelet-based image restoration," *IEEE Trans. Image Process*, vol. 16, no. 12, pp. 2980–2991, Dec. 2007.
[41] X. Chen and R. Womerlsy. Spherical t-design with d=(t 1)^2 points. Accessed: Oct. 26, 2018. [Online]. Available: http://www.polyu.edu.hk/ama/staff/xjchen/sphdesigns.html
[42] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, no. 4, pp. 943–950, 1979.
[43] H. Kuttruff, *Room Acoustics*, 5th ed. London, U.K.: Spon Press, 2009.
[44] R. Narasimhan, *Analysis on Real and Complex Manifolds*, 2nd ed. Amsterdam, The Netherlands: North-Holland, 1985.
[45] A. Dufresnoy, P. N. Gauthier, and W. H. Ow, "Uniform approximation on closed sets by solutions of elliptic partial differential equations," *Complex Var. Elliptic Equ.*, vol. 6, no. 2-4, pp. 235–247, 2007.
[46] A. Kirsch and F. Hettlich, *The Mathematical Theory of Time-Harmonic Maxwell's Equations*. Cham, Switzerland: Springer International Publishing, 2015.
[47] A. Dufresnoy, P. N. Gauthier, and W. H. Ow, "Harmonic polynomials, hyperspherical harmonics, and atomic spectra," *J. Comput. Appl. Math.*, vol. 233, no. 6, pp. 1366–1379, 2007.

Natsuki Ueno (Student Member, IEEE) received the B.E. degree in engineering from Kyoto University, Kyoto, Japan, in 2016, and the M.S. and Ph.D. degrees in information science and technology from the University of Tokyo, Tokyo, Japan, in 2018 and 2021, respectively. His research interests include spatial audio and acoustic signal processing.

Shoichi Koyama (Member, IEEE) received the B.E., M.S., and Ph.D. degrees from the University of Tokyo, Tokyo, Japan, in 2007, 2009, and 2014, respectively. In 2009, he was a Researcher of acoustic signal processing with Nippon Telegraph and Telephone Corporation. In 2014, he moved to the University of Tokyo, and since 2018, he has been a Lecturer. From 2016 to 2018, he was also a Visiting Researcher and JSPS overseas Research Fellow with Institut Langevin, Paris Diderot University (Paris7), Paris, France. His research interests include acoustic inverse problems, sound field analysis and synthesis, and spatial audio. He is a Member of the Acoustical Society of America, the Audio Engineering Society, the Institute of Electronics, Information and Communication Engineers, and the Acoustical Society of Japan (ASJ). He was the recipient of the Itakura Prize Innovative Young Researcher Award by ASJ in 2015 and the Research Award by Funai Foundation for Information Technology in 2018.

Hiroshi Saruwatari (Member, IEEE) received the B.E., M.E., and Ph.D. degrees from Nagoya University, Nagoya, Japan, in 1991, 1993, and 2000, respectively. In 1993, he joined SECOM IS Laboratory, Tokyo, Japan, and in 2000, Nara Institute of Science and Technology, Ikoma, Japan. Since 2014, he has been a Professor with The University of Tokyo, Tokyo, Japan. His research interests include statistical audio signal processing, blind source separation, and speech enhancement. He has put his research into the world's first commercially available independent-component-analysis-based BSS microphone in 2007. He was the recipient of several paper awards from IEICE in 2001 and 2006, from TAF in 2004, 2009, 2012, and 2018, from IEEE-IROS2005 in 2006, and from APSIPA in 2013 and 2018. He was also the recipient of the DOCOMO Mobile Science Award in 2011, the Ichimura Award in 2013, The Commendation for Science and Technology by the Minister of Education in 2015, the Achievement Award from IEICE in 2017, and the Hoko-Award in 2018. He has been professionally involved in various volunteer works for IEEE, EURASIP, IEICE, and ASJ. Since 2018, he has been an APSIPA Distinguished Lecturer.