Sound Field Recording Using Distributed Microphones Based on Harmonic Analysis of Infinite Order

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

***Abstract*—A sound field recording method based on spherical or circular harmonic analysis for arbitrary array geometry and directivity of microphones is proposed. In current methods based on harmonic analysis, a sound field is decomposed into harmonic functions with a center given in advance, which is called a global origin, and their coefficients are obtained up to a certain truncation order using microphone measurements. However, the accuracy of the reconstructed sound field depends on the predefined position of the global origin and the truncation order, which makes it difficult to apply this technique to an asymmetric array since the criterion to determine the position of the global origin and the truncation order is not obvious. We formulate an estimate of the harmonic coefficients on the basis of infinite-order analysis. This formulation enables us to estimate the harmonic coefficients at an arbitrary desired position independently of the position of the global origin without truncation errors. Numerical simulation results indicated that the proposed method makes it possible to avoid performance degradation caused by inappropriate setting of the global origin.**

***Index Terms*—Circular/spherical harmonics, microphone array, sound field recording, sound field reproduction.**

I. Introduction

SOUND field recording aims to estimate a sound field inside a region of interest by using measurements of multiple microphones, which can be used to reproduce the sound field by using multiple loudspeakers or a headphone. To obtain these driving signals for reproducing the sound field, it is inherently necessary to estimate the entire sound field from the microphone measurements. A typical strategy for this sound field capturing problem is to decompose the sound field into spatial Fourier basis functions [1]. The basis functions are generally chosen on the basis of the array geometry of the microphones, i.e., plane-wave functions for planar and linear arrays [2] and spherical, cylindrical, or circular harmonics for spherical, cylindrical, or circular arrays, respectively [3]–[7]. Note that the sound field inside a closed surface cannot be uniquely determined by pressure measurements on its boundary surface [8]. This nonuniqueness problem is typically avoided by using microphones mounted on an acoustically rigid object, a microphone array with multiple layers, or an array of directional microphones [3], [5], [7], [9]–[14].

We here address the problem of estimating circular or spherical harmonic coefficients using distributed microphones. These coefficients can be directly used in many sound field reproduction methods [5], [7], [15], [16]. Laborie *et al.* [17] proposed a comprehensive method for estimating spherical harmonic coefficients at a predefined global origin up to a certain truncation order by solving a linear equation of the coefficients and the microphone measurements. The use of distributed higher order microphones (HOM) was proposed by Samarasinghe *et al.* [18], [19], in which the global harmonic coefficients are estimated from multiple local harmonic coefficients on the basis of the translation theorem. The disadvantage of these methods lies in the necessity to predefine the global origin and the truncation order because the criterion of these settings is not obvious.

We propose a sound field recording method based on harmonic analysis of infinite order. The proposed method enables us to estimate the harmonic coefficients at an arbitrary position independently of the global origin; therefore, it is useful for capturing a sound field using distributed microphones.

II. Preliminaries

We first introduce the basic theory of harmonic representation of a sound field and its translation.

A solution of the homogeneous Helmholtz equation $u(\boldsymbol{r}, k)$ of wavenumber k at position $\boldsymbol{r}$ can be expanded around $\boldsymbol{r}_0$ by using circular or spherical harmonics as

$$u(\boldsymbol{r}, k) = \begin{cases} \sum_{\mu=-\infty}^{\infty} \alpha_\mu(\boldsymbol{r}_0, k)\varphi_\mu(\boldsymbol{r}-\boldsymbol{r}_0, k) & \text{in 2-D} \\ \sum_{\nu=0}^{\infty}\sum_{\mu=-\nu}^{\nu} \alpha_\nu^\mu(\boldsymbol{r}_0, k)\varphi_\nu^\mu(\boldsymbol{r}-\boldsymbol{r}_0, k) & \text{in 3-D} \end{cases}$$
$$= \boldsymbol{\alpha}(\boldsymbol{r}_0, k)^{\mathsf{T}}\boldsymbol{\varphi}(\boldsymbol{r}-\boldsymbol{r}_0, k) \quad (1)$$

where $\boldsymbol{\alpha}(\boldsymbol{r}_0, k) \in \mathbb{C}^\infty$ and $\boldsymbol{\varphi}(\boldsymbol{r}-\boldsymbol{r}_0, k) \in \mathbb{C}^\infty$ are infinite-dimensional vectors of harmonic coefficients and basis functions, respectively. The basis function $\varphi_\mu(\boldsymbol{r}, k)$ in

Manuscript received August 24, 2017; accepted November 3, 2017. Date of publication November 21, 2017; date of current version December 11, 2017. This work was supported in part by JSPS KAKENHI under Grant JP15H05312 and Grant JP16H01735, and in part by SECOM Science and Technology Foundation. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Peter Jax. *(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@ipc.i.u-tokyo.ac.jp; shoichi_koyama@ipc.i.u-tokyo.ac.jp; hiroshi_saruwatari@ipc.i.u-tokyo.ac.jp).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2017.2775242

two-dimensional (2-D) space is defined by using polar coordinates $\boldsymbol{r} = (r, \phi)$ as

$$\varphi_\mu(\boldsymbol{r}, k) = J_\mu(kr)\mathrm{e}^{\mathrm{j}\mu\phi}. \tag{2}$$

Similarly, $\varphi_\nu^\mu(\boldsymbol{r}, k)$ in 3-D space is defined by using spherical coordinates $\boldsymbol{r} = (r, \theta, \phi)$ as

$$\varphi_\nu^\mu(\boldsymbol{r}, k) = \sqrt{4\pi} j_\nu(kr) Y_\nu^\mu(\theta, \phi). \tag{3}$$

Here, $J_\mu(\cdot)$ is the μth-order cylindrical Bessel function of the first kind, $j_\nu(\cdot)$ is the νth-order spherical Bessel function of the first kind, and $Y_\nu^\mu(\cdot)$ is the spherical harmonic function of order ν and degree μ [20]. The basis functions are scaled so that $\alpha_0(\boldsymbol{r}, k)$ and $\alpha_0^0(\boldsymbol{r}, k)$ correspond to $u(\boldsymbol{r}, k)$ at any $\boldsymbol{r}$ and k. Hereafter, k is omitted for notational simplicity.

Two vectors $\boldsymbol{\alpha}(\boldsymbol{r}_0)$ and $\boldsymbol{\alpha}(\boldsymbol{r}_0')$ of harmonic coefficients around two different centers $\boldsymbol{r}_0$ and $\boldsymbol{r}_0'$ can be related as

$$\boldsymbol{\alpha}(\boldsymbol{r}_0') = \boldsymbol{T}(\boldsymbol{r}_0' - \boldsymbol{r}_0)\boldsymbol{\alpha}(\boldsymbol{r}_0) \tag{4}$$

where $\boldsymbol{T}(\boldsymbol{r}_0' - \boldsymbol{r}_0) \in \mathbb{C}^{\infty \times \infty}$ is a linear operator, i.e., an infinite-dimensional matrix. This operator in 2-D and 3-D spaces is, respectively, defined as

$$[\boldsymbol{T}(\boldsymbol{r})\boldsymbol{\alpha}]_\mu = \sum_{\mu'=-\infty}^{\infty} (-1)^{\mu-\mu'} J_{\mu-\mu'}(kr)\mathrm{e}^{-\mathrm{j}(\mu-\mu')\phi}\alpha_{\mu'} \tag{5}$$

and

$$[\boldsymbol{T}(\boldsymbol{r})\boldsymbol{\alpha}]_\nu^\mu = \sum_{\nu'=0}^{\infty}\sum_{\mu'=-\nu'}^{\nu'} \Bigg[4\pi(-1)^{\mu'}\mathrm{j}^{\nu-\nu'} \times \sum_{l=0}^{\nu+\nu'} \mathrm{j}^l j_l(kr) Y_l^{\mu-\mu'}(\theta,\phi)^* \mathcal{G}(\nu', \mu'; \nu, -\mu, l)\Bigg]\alpha_{\nu'}^{\mu'} \tag{6}$$

where $[\cdot]_\mu$ is the element of order μ, $[\cdot]_\nu^\mu$ is the element of order ν and degree μ, and $\mathcal{G}(\cdot)$ is the Gaunt coefficient [20]. This operator $\boldsymbol{T}(\boldsymbol{r})$ is referred to as the translation operator and is used in several sound field recording and reproduction methods [15], [18], [19], [21]. As shown in [20], the translation operator satisfies the following equations:

$$\boldsymbol{T}(-\boldsymbol{r}) = \boldsymbol{T}(\boldsymbol{r})^{-1} = \boldsymbol{T}(\boldsymbol{r})^{\mathsf{H}} \tag{7}$$

$$\boldsymbol{T}(\boldsymbol{r} + \boldsymbol{r}') = \boldsymbol{T}(\boldsymbol{r})\boldsymbol{T}(\boldsymbol{r}'). \tag{8}$$

III. Estimation of Infinite-Order Harmonic Coefficients

Our objective is to obtain the harmonic coefficients $\boldsymbol{\alpha}(\boldsymbol{r})$ at an arbitrary position $\boldsymbol{r}$. We derive the posterior probability measure of $\boldsymbol{\alpha}(\boldsymbol{r})$ given the observed signals on the basis of harmonic analysis of infinite order and Bayesian inference.

A directivity pattern of a microphone is represented as $c(\phi)$ and $c(\theta, \phi)$ in 2-D and 3-D spaces, respectively. The response of the microphone at $\boldsymbol{r}_0$ in the sound field $u(\boldsymbol{r})$, which is expanded around $\boldsymbol{r}_0$ as in (1), is given by

$$s = \begin{cases} \sum_{\mu=-\infty}^{\infty} c_\mu^* \alpha_\mu(\boldsymbol{r}_0) & \text{in 2-D} \\ \sum_{\nu=0}^{\infty}\sum_{\mu=-\nu}^{\nu} c_\nu^{\mu *}\alpha_\nu^\mu(\boldsymbol{r}_0) & \text{in 3-D} \end{cases} = \boldsymbol{c}^{\mathsf{H}}\boldsymbol{\alpha}(\boldsymbol{r}_0) \tag{9}$$

where $c_\nu \in \mathbb{C}$ and $c_\nu^\mu \in \mathbb{C}$ are, respectively, obtained from the circular and spherical harmonic expansion of $c(\phi)$ and $c(\theta, \phi)$, and $\boldsymbol{c} \in \mathbb{C}^\infty$ is the infinite-dimensional vector of these coefficients. The detailed derivation of (9) is shown in the supplementary material.

Suppose that M microphones with arbitrary directivity $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_M$ are located at $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_M$ with an arbitrary array geometry. The signal observed by the mth microphone is given by

$$\begin{aligned} s_m &= \boldsymbol{c}_m^{\mathsf{H}}\boldsymbol{\alpha}(\boldsymbol{r}_m) + \epsilon_m \\ &= \boldsymbol{c}_m^{\mathsf{H}}\boldsymbol{T}(\boldsymbol{r}_m - \boldsymbol{r}_0)\boldsymbol{\alpha}(\boldsymbol{r}_0) + \epsilon_m \end{aligned} \tag{10}$$

where $\epsilon_m \in \mathbb{C}$ is the sensor noise. Equation (10) can be written in matrix form as

$$\boldsymbol{s} = \boldsymbol{\Xi}(\boldsymbol{r}_0)^{\mathsf{H}}\boldsymbol{\alpha}(\boldsymbol{r}_0) + \boldsymbol{\epsilon} \tag{11}$$

where $\boldsymbol{s} \in \mathbb{C}^M$, $\boldsymbol{\epsilon} \in \mathbb{C}^M$, and $\boldsymbol{\Xi}(\boldsymbol{r}_0) \in \mathbb{C}^{\infty \times M}$ are, respectively, defined as $\boldsymbol{s} = [s_1, \ldots, s_M]^{\mathsf{T}}$, $\boldsymbol{\epsilon} = [\epsilon_1, \ldots, \epsilon_M]^{\mathsf{T}}$, and

$$\begin{aligned} \boldsymbol{\Xi}(\boldsymbol{r}_0) &= \big[\, (\boldsymbol{c}_1^{\mathsf{H}}\boldsymbol{T}(\boldsymbol{r}_1 - \boldsymbol{r}_0))^{\mathsf{H}} \quad \cdots \quad (\boldsymbol{c}_M^{\mathsf{H}}\boldsymbol{T}(\boldsymbol{r}_M - \boldsymbol{r}_0))^{\mathsf{H}} \,\big] \\ &= \big[\, \boldsymbol{T}(\boldsymbol{r}_0 - \boldsymbol{r}_1)\boldsymbol{c}_1 \quad \cdots \quad \boldsymbol{T}(\boldsymbol{r}_0 - \boldsymbol{r}_M)\boldsymbol{c}_M \,\big]. \end{aligned} \tag{12}$$

The second line of (12) follows from (7). Assuming that $\boldsymbol{\epsilon}$ follows a circularly symmetric complex Gaussian measure with covariance $\boldsymbol{\Sigma} \in \mathbb{C}^{M \times M}$ as $\boldsymbol{\epsilon} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{\Sigma})$, the conditional probability measure of the observed signals $\boldsymbol{s}$ given $\boldsymbol{\alpha}(\boldsymbol{r}_0)$ is

$$P_{\boldsymbol{s}|\boldsymbol{\alpha}(\boldsymbol{r}_0)} = \mathcal{N}_{\mathbb{C}}(\boldsymbol{\Xi}(\boldsymbol{r}_0)^{\mathsf{H}}\boldsymbol{\alpha}(\boldsymbol{r}_0), \boldsymbol{\Sigma}) \tag{13}$$

where $P_\cdot$ denotes a probability measure. Additionally, we assume the prior measure of $\boldsymbol{\alpha}(\boldsymbol{r}_0)$ as

$$P_{\boldsymbol{\alpha}(\boldsymbol{r}_0)} = \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \sigma_o^2\boldsymbol{I}) \tag{14}$$

where $\boldsymbol{I} \in \mathbb{C}^{\infty \times \infty}$ is the identity operator. Note that this prior measure is invariant with respect to $\boldsymbol{r}_0$. Given (14), the probability measure of $\boldsymbol{\alpha}(\boldsymbol{r}) = \boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}_0)\boldsymbol{\alpha}(\boldsymbol{r}_0)$ at arbitrary $\boldsymbol{r}$ is derived as [22]

$$\begin{aligned} P_{\boldsymbol{\alpha}(\boldsymbol{r})} &= \mathcal{N}_{\mathbb{C}}(\boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}_0)\boldsymbol{0}, \sigma_o^2\boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}_0)\boldsymbol{I}\boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}_0)^{\mathsf{H}}) \\ &= \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \sigma_o^2\boldsymbol{I}). \end{aligned} \tag{15}$$

The second line of (15) follows from (7). Applying Bayes' theorem for infinite-dimensional Gaussian measures [23], the posterior measure of $\boldsymbol{\alpha}(\boldsymbol{r}_0)$ is derived as

$$P_{\boldsymbol{\alpha}(\boldsymbol{r}_0)|\boldsymbol{s}} = \mathcal{N}_{\mathbb{C}}(\boldsymbol{\eta}_\alpha(\boldsymbol{r}_0), \boldsymbol{\Sigma}_\alpha(\boldsymbol{r}_0)) \tag{16}$$

where $\boldsymbol{\eta}_\alpha(\boldsymbol{r}_0) \in \mathbb{C}^\infty$ and $\boldsymbol{\Sigma}_\alpha(\boldsymbol{r}_0) \in \mathbb{C}^{\infty \times \infty}$ are given by

$$\boldsymbol{\eta}_\alpha(\boldsymbol{r}_0) = \boldsymbol{\Xi}(\boldsymbol{r}_0)(\boldsymbol{\Psi} + \sigma_o^{-2}\boldsymbol{\Sigma})^{-1}\boldsymbol{s} \tag{17}$$

$$\boldsymbol{\Sigma}_\alpha(\boldsymbol{r}_0) = \sigma_o^2[\boldsymbol{I} - \boldsymbol{\Xi}(\boldsymbol{r}_0)(\boldsymbol{\Psi} + \sigma_o^{-2}\boldsymbol{\Sigma})^{-1}\boldsymbol{\Xi}(\boldsymbol{r}_0)^{\mathsf{H}}] \tag{18}$$

with $\boldsymbol{\Psi} = \boldsymbol{\Xi}(\boldsymbol{r}_0)^{\mathsf{H}}\boldsymbol{\Xi}(\boldsymbol{r}_0) \in \mathbb{C}^{M\times M}$. Using (8), $\boldsymbol{\Psi}$ is represented as

$$(\boldsymbol{\Psi})_{m,m'} = \boldsymbol{c}_m^{\mathsf{H}}\boldsymbol{T}(\boldsymbol{r}_m - \boldsymbol{r}_0)\boldsymbol{T}(\boldsymbol{r}_0 - \boldsymbol{r}_{m'})\boldsymbol{c}_{m'} \\ = \boldsymbol{c}_m^{\mathsf{H}}\boldsymbol{T}(\boldsymbol{r}_m - \boldsymbol{r}_{m'})\boldsymbol{c}_{m'} \tag{19}$$

where $(\cdot)_{m,m'}$ denotes the (m,m')th element of the matrix. Equation (19) indicates that $\boldsymbol{\Psi}$ does not depend on the position $\boldsymbol{r}_0$, and only depends on the directivity and positions of the microphones. Furthermore, when each $\boldsymbol{c}_m$ has a finite number of nonzero elements, $\boldsymbol{\Psi}$ can be calculated accurately without truncation errors. Note that $\boldsymbol{c}_m$ is typically modeled by low-order harmonic coefficients.

The posterior measure of $\boldsymbol{\alpha}(\boldsymbol{r})$ at an arbitrary position $\boldsymbol{r}$ given $\boldsymbol{s}$ is derived as [22]

$$P_{\boldsymbol{\alpha}(\boldsymbol{r})|\boldsymbol{s}} = \mathcal{N}_{\mathbb{C}}(\boldsymbol{\eta}_\alpha(\boldsymbol{r}), \boldsymbol{\Sigma}_\alpha(\boldsymbol{r})) \tag{20}$$

where $\boldsymbol{\eta}_\alpha(\boldsymbol{r}) \in \mathbb{C}^\infty$ and $\boldsymbol{\Sigma}_\alpha(\boldsymbol{r}) \in \mathbb{C}^{\infty\times\infty}$ are given by

$$\boldsymbol{\eta}_\alpha(\boldsymbol{r}) = \boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}_0)\boldsymbol{\eta}_\alpha(\boldsymbol{r}_0) \\ = \boldsymbol{\Xi}(\boldsymbol{r})(\boldsymbol{\Psi} + \sigma_o^{-2}\boldsymbol{\Sigma})^{-1}\boldsymbol{s} \tag{21}$$

$$\boldsymbol{\Sigma}_\alpha(\boldsymbol{r}) = \boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}_0)\boldsymbol{\Sigma}_\alpha(\boldsymbol{r}_0)\boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}_0)^{\mathsf{H}} \\ = \sigma_o^2[\boldsymbol{I} - \boldsymbol{\Xi}(\boldsymbol{r})(\boldsymbol{\Psi} + \sigma_o^{-2}\boldsymbol{\Sigma})^{-1}\boldsymbol{\Xi}(\boldsymbol{r})^{\mathsf{H}}]. \tag{22}$$

The second lines of (21) and (22) are derived by applying $\boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}_0)\boldsymbol{\Xi}(\boldsymbol{r}_0) = \boldsymbol{\Xi}(\boldsymbol{r})$ to the first lines. Equations (17) and (18), respectively, correspond to (21) and (22) by substituting $\boldsymbol{r}$ for $\boldsymbol{r}_0$.

In conclusion, the harmonic coefficients $\hat{\boldsymbol{\alpha}}(\boldsymbol{r})$ at any position $\boldsymbol{r}$ can be estimated as the posterior mean of $\boldsymbol{\alpha}(\boldsymbol{r})$ as

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{r}) = \boldsymbol{\Xi}(\boldsymbol{r})(\boldsymbol{\Psi} + \sigma_o^{-2}\boldsymbol{\Sigma})^{-1}\boldsymbol{s}. \tag{23}$$

Furthermore, the reliability of the estimated harmonic coefficients can be calculated on the basis of the posterior covariance $\boldsymbol{\Sigma}_\alpha(\boldsymbol{r})$, which will be helpful for evaluating the recordable area and robustness of a microphone array.

As an example, we show the specific formulation for the case where all the microphones have omnidirectional directivity in 2-D space. In this case, $\boldsymbol{c}_m$ and $\boldsymbol{\Psi}$ are given by

$$[\boldsymbol{c}_m]_\mu = \delta_{\mu,0} \tag{24}$$

$$(\boldsymbol{\Psi})_{m,m'} = J_0(k\|\boldsymbol{r}_m - \boldsymbol{r}_{m'}\|) \tag{25}$$

where $\delta_{\mu,\mu'}$ is the Kronecker delta and $\|\cdot\|$ denotes the Euclidean distance. The matrix $\boldsymbol{\Psi}$ for this situation corresponds to the normalized spatial covariance matrix in a diffuse sound field [24].

IV. Comparison With Prior Works

We here discuss the difference between the proposed method and the method proposed in [5] and [17], which is referred to as a truncation method. In the truncation method, the harmonic coefficients at the global origin are obtained up to a certain truncation order N by solving the following equation:

$$\boldsymbol{s} = \bar{\boldsymbol{\Xi}}(\boldsymbol{r}_0)^{\mathsf{H}}\bar{\boldsymbol{\alpha}}(\boldsymbol{r}_0) \tag{26}$$

where $\bar{\boldsymbol{\alpha}}(\boldsymbol{r}_0) \in \mathbb{C}^{\bar{N}}$ and $\bar{\boldsymbol{\Xi}}(\boldsymbol{r}_0) \in \mathbb{C}^{\bar{N}\times M}$ are truncated forms of $\boldsymbol{\alpha}(\boldsymbol{r}_0)$ and $\boldsymbol{\Xi}(\boldsymbol{r}_0)$, respectively, and $\bar{N}$ is defined as $\bar{N} = 2N + 1$ in 2-D space and $\bar{N} = (N+1)^2$ in 3-D space. Its solution including a regularization term is

$$\bar{\boldsymbol{\alpha}}(\boldsymbol{r}_0) = \bar{\boldsymbol{\Xi}}(\boldsymbol{r}_0)(\bar{\boldsymbol{\Xi}}(\boldsymbol{r}_0)^{\mathsf{H}}\bar{\boldsymbol{\Xi}}(\boldsymbol{r}_0) + \lambda\boldsymbol{I}_M)^{-1}\boldsymbol{s} \tag{27}$$

where $\lambda \in [0,\infty)$ is a regularization parameter and $\boldsymbol{I}_M \in \mathbb{C}^{M\times M}$ is the identity matrix. The harmonic coefficients at an arbitrary position $\boldsymbol{r}$ are obtained by

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{r}) = \bar{\boldsymbol{T}}(\boldsymbol{r} - \boldsymbol{r}_0)\bar{\boldsymbol{\Xi}}(\boldsymbol{r}_0)(\bar{\boldsymbol{\Xi}}(\boldsymbol{r}_0)^{\mathsf{H}}\bar{\boldsymbol{\Xi}}(\boldsymbol{r}_0) + \lambda\boldsymbol{I}_M)^{-1}\boldsymbol{s} \tag{28}$$

where $\bar{\boldsymbol{T}}(\boldsymbol{r}) \in \mathbb{C}^{\infty\times\bar{N}}$ is a truncated form of $\boldsymbol{T}(\boldsymbol{r})$. In the proposed method, on the other hand, if the sensor noises are uncorrelated and have the same variance, $\sigma_o^{-2}\boldsymbol{\Sigma}$ is a scalar matrix and (23) becomes

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{r}) = \boldsymbol{\Xi}(\boldsymbol{r})(\boldsymbol{\Psi} + \lambda\boldsymbol{I}_M)^{-1}\boldsymbol{s}. \tag{29}$$

The main difference between the two methods is whether or not $\boldsymbol{\Xi}(\boldsymbol{r}_0)$ and $\boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}_0)$ are truncated. In the truncation method, (28) depends on the position of the global origin $\boldsymbol{r}_0$ and the truncation order N. If $\bar{\boldsymbol{\Xi}}(\boldsymbol{r}_0)$ and $\bar{\boldsymbol{T}}(\boldsymbol{r} - \boldsymbol{r}_0)$ are not truncated, (28) corresponds to (29) in the proposed method, which is derived from $\boldsymbol{T}(\boldsymbol{r} - \boldsymbol{r}_0)\boldsymbol{\Xi}(\boldsymbol{r}_0) = \boldsymbol{\Xi}(\boldsymbol{r})$ and $\boldsymbol{\Xi}(\boldsymbol{r}_0)^{\mathsf{H}}\boldsymbol{\Xi}(\boldsymbol{r}_0) = \boldsymbol{\Psi}$. In the proposed method, these two relations based on the infinite-order analysis enable us to estimate harmonic coefficients independently of $\boldsymbol{r}_0$.

V. Numerical Simulations

We performed numerical simulations in 2-D space to evaluate the proposed method for the reconstruction of a pressure field. We compared the proposed method (*Proposed*) with the truncation method (*Truncation*) [17] and the method using distributed HOM [19]. Nine circular arrays of microphones were distributed under the condition that the arrays do not overlap. Their centers were randomly distributed within a circular area with center $(0.0, 0.0)$ m in Cartesian coordinates and radius 1.5 m. Each circular array consisted of seven omnidirectional microphones equiangularly aligned on a circle of radius 0.2 m; therefore, the total number of microphones was $M = 63$. Open arrays were used in Proposed and Truncation and rigid arrays were used in HOM. Gaussian noise was added to each microphone signal so that the signal-to-noise ratio became 30 dB. The sound speed was set to 340.29 m/s. In Proposed, the regularization parameter was set as $\sigma_o^{-2}\boldsymbol{\Sigma} = \lambda\boldsymbol{I}_M$ with $\lambda = 10^{-3}$. In Truncation, two different global origins, $\boldsymbol{r}_0 = (0.0, 0.0)$ m and $\boldsymbol{r}_0 = (1.5, 0.0)$ m, were investigated. The truncation order was 31 and the regularization parameter in (28) was $\lambda = 10^{-3}$. In HOM, the global origin was $\boldsymbol{r}_0 = (0.0, 0.0)$ m and the truncation order and regularization parameter were determined as described in [19]. In this simulation, mutual scattering effects between the rigid circles were ignored.

Fig. 1 shows the original and reconstructed sound pressure distributions for a single-frequency plane wave traveling along the y-axis in the positive direction at a frequency of 650 Hz. The estimated normalized error distribution based on (22) and the actual normalized error distributions are shown in Fig. 2. The

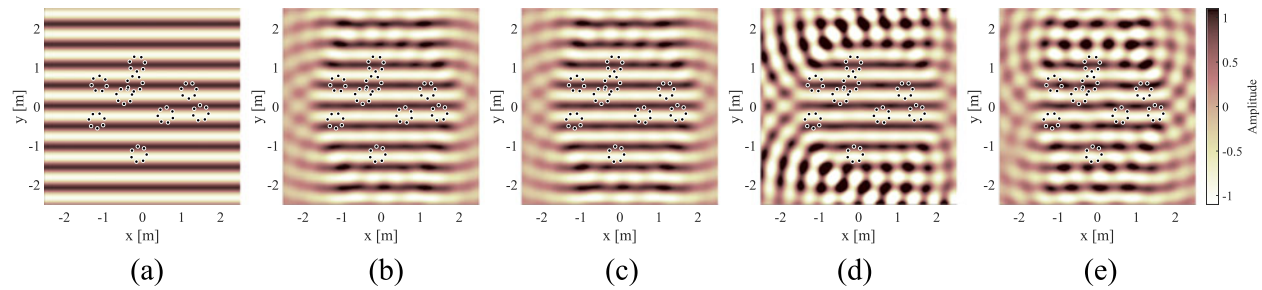

Fig. 1. Sound pressure distributions for (a) original, (b) proposed, (c) truncation: $\boldsymbol{r}_0 = (0.0, 0.0)$ m, (d) truncation: $\boldsymbol{r}_0 = (1.5, 0.0)$ m, (e) HOM.

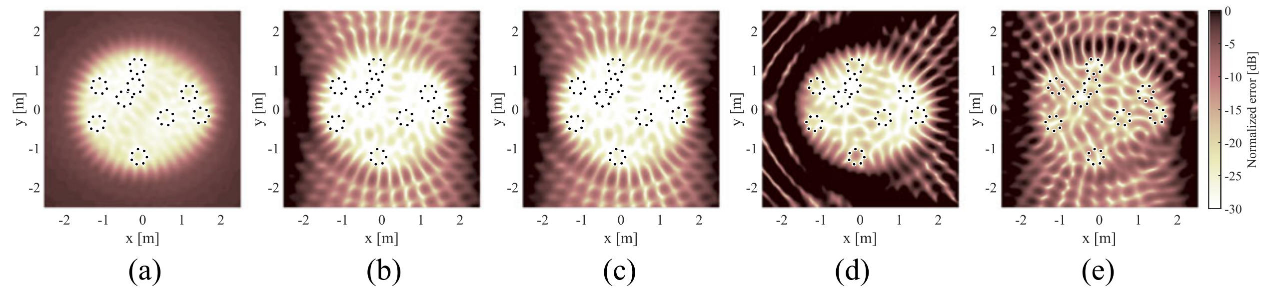

Fig. 2. Normalized error distributions for (a) estimate, (b) proposed, (c) truncation: $\boldsymbol{r}_0 = (0.0, 0.0)$ m, (d) truncation: $\boldsymbol{r}_0 = (1.5, 0.0)$ m, (e) HOM.

black dots represent the positions of the microphones. In spite of a forbidden frequency of the single open circular array (i.e., $J_0(kR) \approx 0$ at 650 Hz and $R = 0.2$ m), high reconstruction accuracy was achieved by Proposed. This is because all 63 microphones are considered as a single large microphone array in the proposed method. Truncation and HOM also achieved high accuracy when $\boldsymbol{r}_0$ was $(0.0, 0.0)$ m; however, unwanted distortion arose in the peripheral area when $\boldsymbol{r}_0$ was $(1.5, 0.0)$ m in Truncation.

To investigate the relationship between the frequency and performance, we define the signal-to-distortion ratio (SDR) $\mathrm{SDR}(k)$ as

$$\mathrm{SDR}(k) = 10 \log_{10} \frac{\int_{\boldsymbol{r} \in V} |u(\boldsymbol{r}, k)|^2 \mathrm{d}\boldsymbol{r}}{\int_{\boldsymbol{r} \in V} |u(\boldsymbol{r}, k) - \hat{u}(\boldsymbol{r}, k)|^2 \mathrm{d}\boldsymbol{r}} \tag{30}$$

for a circular evaluation area V with center $(0.0, 0.0)$ m and radius 1.5 m. Here, $u(\boldsymbol{r}, k)$ and $\hat{u}(\boldsymbol{r}, k)$ are the true and reconstructed sound pressures, respectively. The integrals in (30) were discretized with an interval of 0.02 m. The SDRs were calculated for 16 array geometries of the microphones and 64 plane waves arriving from equiangular directions. Fig. 3 shows the average of all 1024 SDRs at each frequency. We also plotted the estimated SDR given by

$$\mathrm{SDR}_{\mathrm{estimate}}(k) = 10 \log_{10} \frac{\int_{\boldsymbol{r} \in V} \sigma_o^2 \mathrm{d}\boldsymbol{r}}{\int_{\boldsymbol{r} \in V} [\boldsymbol{\Sigma}_\alpha(\boldsymbol{r}, k)]_{0,0} \mathrm{d}\boldsymbol{r}} \tag{31}$$

where $[\boldsymbol{\Sigma}_\alpha(\boldsymbol{r}, k)]_{0,0}$ denotes the posterior variance of the zeroth harmonic coefficient $\alpha_0(\boldsymbol{r}, k)$. At low frequencies, Proposed and Truncation achieved almost the same performance regardless of the position of the global origin. At high frequencies, however, Proposed achieved higher performance than Truncation. This is because the truncation error in Truncation becomes large at high frequencies. In HOM, the reconstruction errors are larger

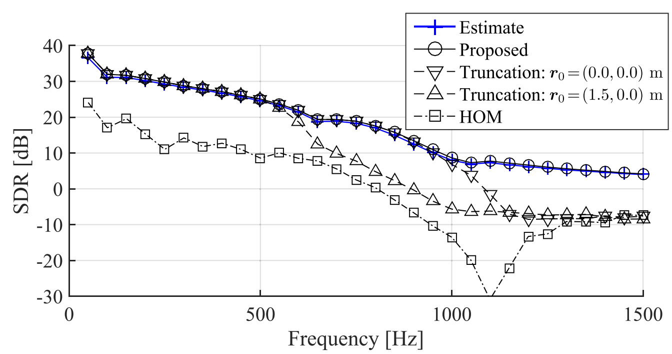

Fig. 3. SDR with respect to frequency.

than those of the other methods. It can be considered that the errors of local coefficients propagate to the global coefficients. Proposed reduces this error propagation by dealing with all the microphones as a single array and maintains relatively high performance up to high frequencies. In addition, the estimated SDRs almost correspond to the actual SDRs in Proposed. This indicates that the proposed method will also be applicable to the evaluation of microphone arrays, although detailed studies remain as future work.

VI. Conclusion

We proposed a new sound field recording method for an array of distributed microphones. The proposed method is based on a formulation using harmonic analysis of infinite order. The harmonic coefficients at an arbitrary position can be obtained independently of the global origin; therefore, it is possible to avoid performance degradation caused by inappropriate setting of the global origin. We also derived the posterior covariance for estimating the performance of a microphone array.

REFERENCES

[1] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustic Holography.* London, U.K.: Academic, 1999.

[2] S. Koyama, K. Furuya, Y. Hiwasaki, and Y. Haneda, "Analytical approach to wave field reconstruction filtering in spatio-temporal frequency domain," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 4, pp. 685–696, Apr. 2013.

[3] J. Daniel, S. Moreau, and R. Nicol, "Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging," in *Proc. 114th AES Conv.*, Amsterdam, Oct. 2003, paper no. 5788.

[4] Z. Li, R. Duraiswami, and N. A. Gumerov, "Capture and recreation of higher-order 3D sound fields via reciprocity," in *Proc. Int. Conf. Audio Display*, 2004, pp. 6–9.

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

[6] S. Koyama, K. Furuya, Y. Hiwasaki, Y. Haneda, and Y. Suzuki, "Wave field reconstruction filtering in cylindrical harmonic domain for with-height recording and reproduction," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 10, pp. 1546–1557, Oct. 2014.

[7] S. Koyama, K. Furuya, K. Wakayama, S. Shimauchi, and H. Saruwatari, "Analytical approach to transforming filter design for sound field recording and reproduction using circular arrays with a spherical baffle," *J. Acoust. Soc. Amer.*, vol. 139, no. 3, pp. 1024–1036, 2016.

[8] F. M. Fazi and P. A. Nelson, "Nonuniqueness of the solution of the sound field reproduction problem with boundary pressure control," *Acta Acustica United Acustica*, vol. 98, no. 1, pp. 1–14, 2012.

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

[10] 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.*, Orlando, May 2002, pp. II-1781–II-1784.

[11] 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.*, 2002, pp. II-1949–II-1952.

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

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

[14] B. Rafaely, "The spherical-shell microphone array," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 4, pp. 740–747, May 2008.

[15] Y. J. Wu and T. D. Abhayapala, "Spatial multizone soundfield reproduction: Theory and design," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 6, pp. 1711–1720, Aug. 2011.

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

[17] A. Laborie, R. Bruno, and S. Montoya, "A new comprehensive approach of surround sound recording," in *Proc. 114th AES Conv.*, Amsterdam, Oct. 2003, paper no. 5717.

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

[19] 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.

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

[21] F. Winter, J. Ahrens, and S. Spors, "On analytic methods for 2.5-D local sound field synthesis using circular distributions of secondary sources," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 5, pp. 914–926, May 2016.

[22] G. D. Prato, *An Introduction to Infinite-Dimensional Analysis*. Berlin, Germany: Springer-Verlag, 2006.

[23] A. M. Stuart, "Inverse problems: A Bayesian perspective," *Acta Numerica*, vol. 19, pp. 451–559, 2010.

[24] R. K. Cook, R. V. Waterhouse, R. D. Berendt, S. Edelman, and M. C. Thompson, Jr., "Measurement of correlation coefficients in reverberant sound fields," *J. Acoust. Soc. Amer.*, vol. 27, no. 6, pp. 1072–1077, 1955.