Modelling Room Reverberation Directivity using von Mises-Fisher Mixture Distribution

Amy Bastine, Thushara D. Abhayapala, Jihui Zhang, and Huiyuan Sun
Audio and Acoustic Signal Processing Group, The Australian National University, Canberra, Australia
E-mail:{amy.bastine, thushara.abhayapala, jihui.zhang, huiyuan.sun}@anu.edu.au

***Abstract*—In this paper, we propose a novel approach to room reverberation analysis based on the angular power distribution model of the reverberant field developed using a von Mises-Fisher (vMF) mixture function. The statistical features of the complex reverberant power distribution convey the directional strength of inhomogeneous room reflections. The model is primarily designed for directivity study and also helps in encapsulating the extensive raw data into a convenient set of parameters of the density function. Initially, we conducted impulse response measurements of the test room and determined the reverberant field power values using the spatial correlation model in the spherical harmonics domain. The proposed technique transforms this data into directional power vectors to estimate the parameters of a convex vMF mixture function. A power distribution model is generated from these parameters to represent the directional characteristics of room reverberation. The paper evaluates the performance of the model for a test room, and the results conform to the real room environment. The directivity model identifies directions of distinct reflections and exhibits the potential for further functional enhancements to evolve as a reliable statistical room acoustic model.**

***Index Terms*—Reverberation directivity, room acoustic modelling, vMF mixture distribution, spatial correlation**

I. Introduction

Room acoustic modelling and reverberation analysis play significant roles in the development of virtual reality environments and immersive audio technologies. The classical room acoustic models consider a reverberant room as a linear time-invariant (LTI) acoustical transmission system. The collection of Room Impulse Responses (RIR) or Room Transfer Functions (RTF) measured for different source-receiver positions across the room represent the dynamic behavior of the reverberant field and acts as its acoustic fingerprint [1].

The primitive reverberation analysis methods used the RIR/RTF directly to compute many objective parameters and performance measures [2]–[4]. A more intricate examination of spatial sound field features was enabled by later studies using methods like sound intensity mapping [5], 3D analysis using higher-order spherical harmonics [6]–[9], and plane-wave decomposition [10], [11]. In [12], Hioka et al. proposed a technique to directly process the microphone measurements using a spatial correlation matrix model. However, such models required a higher number of microphones and suffered spatial aliasing errors at higher frequencies. In [13], Samarasinghe et al. refined the spatial-correlation model and overcame the above drawbacks by incorporating wave-based spherical harmonic modelling with higher modal analysis of Eigenbeams. This model was robust to surrounding noises and clearly identified the dominant reflections, but it exhibited a few negative power values in the reflection power profile.

The latest room acoustic models [14], [15] have tended to rely on synthetic RIRs generated from reconstructed 3D room geometry due to the practical limitations in performing large-scale acoustic measurements for various source-receiver positions across the 3D room space. However, the substantial use of image processing and deep learning algorithms in these models make them numerically complex and computationally very demanding. Hence, the current scenario puts forth the requirement for a computationally modest and pragmatic model based on sound field measurements without compromising data authenticity.

In this paper, we propose a novel approach to reverberation analysis by representing the reflected field directivity as an angular power distribution based on von Mises-Fisher (vMF) mixture model. Over the last two decades, researchers have established the competence of vMF mixture models to process directional data in the fields of biomedical imaging [16]–[18], renewable energy [19], computer vision [20], [21], speaker identification [22], and speech modelling [23]. Through this paper, we are introducing vMF mixture modelling in the field of room acoustics for higher-dimensional analysis of directional data by exploiting its multivariate density function.

The proposed method converts the reverberant field powers corresponding to different reflection directions into directional power vectors for fitting a vMF mixture model. We determine the field powers beforehand using the state-of-the-art estimation algorithm proposed in [13] on the grounds of its aforementioned advantages. The vMF mixture density function produces a non-negative reflection power profile, thereby overcoming the shortcoming of [13] and encapsulates the extensive sound field information into suitable statistical parameters. The transformation of standard RIRs into a power distribution function reduces the redundancy of the raw data. Consequently, it scales down the measurement effort since the RIRs between sufficiently sampled points across the room space are adequate for faithful reproduction of the angular power distribution model.

The remainder of this paper is organized as follows: Section II presents the system model and background framework, followed by the problem definition and objectives of this research

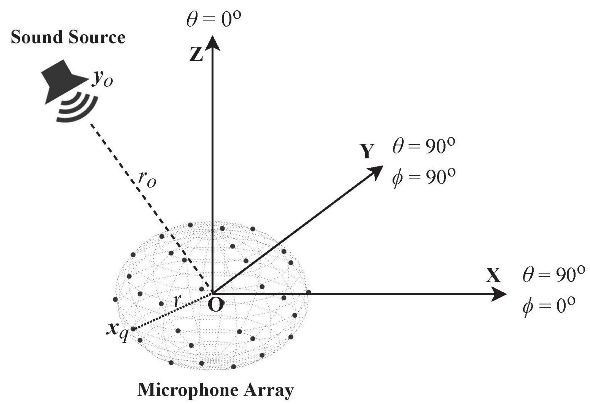


Fig. 1: Geometrical illustration of the source-receiver system

work. In Section III, we describe the procedure for estimating the reverberant field power using the spatial correlation model in the spherical harmonics domain. Section IV explains the new directivity model and highlights the transformation of reverberant field power data into directional power vectors. Section V outlines the experimental setup and evaluates the preliminary results of the proposed model. Finally, in Section VI, we conclude the paper and present future research plans.

II. System Model

Consider the source-receiver arrangement illustrated in Fig.1 with a spherical microphone array of radius r consisting of Q omnidirectional microphones located at the origin O. Let $H(\boldsymbol{x}_q, \boldsymbol{y}_o, k)$ be the RTF between the source located at $\boldsymbol{y}_o = (r_o, \theta_o, \phi_o)$ and the q^{th} microphone element located at $\boldsymbol{x}_q = (r, \theta_q, \phi_q)$, where $q = 1, 2, ..., Q$. Assuming the room as an LTI system, the sound field received by the q^{th} microphone can be expressed in the frequency-domain as

$$\Psi(\boldsymbol{x}_q, k) = S(k) H(\boldsymbol{x}_q, \boldsymbol{y}_o, k) \tag{1}$$

where $S(k)$ is the Short-time Fourier transform (STFT) of the source signal, and k is the wavenumber. Since the incident sound field $\Psi(\boldsymbol{x}_q, k)$ contains the dominant direct path along with the early reflections and the reverberant field, we decompose the RTF $H(\boldsymbol{x}_q, \boldsymbol{y}_o, k)$ into the direct $H_{\text{dir}}(\boldsymbol{x}_q, \boldsymbol{y}_o, k)$ and reflected $H_{\text{rvb}}(\boldsymbol{x}_q, \boldsymbol{y}_o, k)$ components as

$$H(\boldsymbol{x}_q, \boldsymbol{y}_o, k) = H_{\text{dir}}(\boldsymbol{x}_q, \boldsymbol{y}_o, k) + H_{\text{rvb}}(\boldsymbol{x}_q, \boldsymbol{y}_o, k). \tag{2}$$

Assuming that the distance between the source and microphone array is significantly larger than $2L^2/\lambda$, where L is the aperture size of the microphone array and λ is the wavelength, we can consider the sound field as a composition of plane waves [13]. Let the gain of the direct path incoming from direction $\hat{\boldsymbol{y}}_o$ be $G_D(k)$, then the direct component of RTF is a plane wave of the form

$$H_{\text{dir}}(\boldsymbol{x}_q, \boldsymbol{y}_o, k) = G_D(k) e^{ik\hat{\boldsymbol{y}}_o \cdot \boldsymbol{x}_q}. \tag{3}$$

Similarly, let $G_R(k, \hat{\boldsymbol{y}})$ be the gain of the reflected plane wave arriving from direction $\hat{\boldsymbol{y}} = (1, \theta, \phi)$ for $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$, then the reflection component of RTF from all directions is

$$H_{\text{rvb}}(\boldsymbol{x}_q, \boldsymbol{y}_o, k) = \int_0^{2\pi} \int_0^{\pi} G_R(k, \hat{\boldsymbol{y}}) e^{ik\hat{\boldsymbol{y}} \cdot \boldsymbol{x}_q} \sin\theta d\theta d\phi. \tag{4}$$

Substituting (2), (3) and (4) in (1), we expand $\Psi(\boldsymbol{x}_q, k)$ as

$$\begin{aligned}\Psi(\boldsymbol{x}_q, k) &= S(k) G_D(k) e^{ik\hat{\boldsymbol{y}}_o \cdot \boldsymbol{x}_q} \\ &+ S(k) \int_0^{2\pi} \int_0^{\pi} G_R(k, \hat{\boldsymbol{y}}) e^{ik\hat{\boldsymbol{y}} \cdot \boldsymbol{x}_q} \sin\theta d\theta d\phi. \end{aligned} \tag{5}$$

By inspecting the above representation, the total power contributed by the reflected components is:

$$P_R = E\{|S(k)|^2\} \int_0^{2\pi} \int_0^{\pi} E\{|G_R(k, \hat{\boldsymbol{y}})|^2\} \sin\theta d\theta d\phi \tag{6}$$

where $E\{\cdot\}$ is the expectation operator.

A. Harmonic Expansion of Angular Power Distribution

The function $E\{|G_R(k, \hat{\boldsymbol{y}})|^2\}$ across different $\hat{\boldsymbol{y}}$ gives the directional distribution of P_R. For computational convenience and efficient estimation using limited coefficients, we decompose $E\{|G_R(k, \hat{\boldsymbol{y}})|^2\}$ using spherical harmonics as

$$E\{|G_R(k, \hat{\boldsymbol{y}})|^2\} = \sum_{v=0}^{\infty} \sum_{u=-v}^{v} \gamma_{vu}(k) Y_{vu}(\hat{\boldsymbol{y}}) \tag{7}$$

where $Y_{vu}(\cdot)$ is called the spherical harmonic function of v^{th} order and u^{th} mode, and γ_{vu} are the corresponding reflection gain coefficients. In order to model the 3D reflected field power distribution, we write the power of reflected component from look direction $\hat{\boldsymbol{y}}$ as

$$P_R(k, \hat{\boldsymbol{y}}) = \sum_{v=0}^{\infty} \sum_{u=-v}^{v} \Gamma_{vu}(k) Y_{vu}(\hat{\boldsymbol{y}}) \tag{8}$$

where $\Gamma_{vu}(k) = E\{|S(k)|^2\} \gamma_{vu}(k)$ are the reflection power coefficients. Thus, the reflected power from any direction can be determined once we estimate the $\Gamma_{vu}(k)$ coefficients.

B. Problem Formulation

The magnitude of the reflected power $P_R(k, \hat{\boldsymbol{y}})$ conveys the strength of effective reflectance from all surface points located along the $\hat{\boldsymbol{y}}$ direction. Therefore, the $P_R(k, \hat{\boldsymbol{y}})$ values sampled over a set of arbitrary directions $\hat{\boldsymbol{y}}$ can provide the relative strength of reflectance across the room boundary, and its diverse combination from several RIRs can serve as directional data. Our objective is to deduce the reflection gain directivity statistically from this set of $P_R(k, \hat{\boldsymbol{y}})$ values by fitting a probability distribution model capable of characterizing directional data.

The reflections from the inhomogeneous room surfaces contribute to distinct statistical features in the reverberant field power distribution. Also, this distribution should be a multimodal function with multiple peaks corresponding to the highly reflecting surfaces. Hence, a convex mixture model is required to represent the angular power distribution as follows:

$$P_R(k, \hat{\boldsymbol{y}}) \equiv \mathfrak{F}(k, \hat{\boldsymbol{y}}; \Delta) \equiv \sum_{\iota} \sigma_\iota \mathfrak{f}(k, \hat{\boldsymbol{y}}; \delta_\iota) \tag{9}$$

where σ_ι are the mixing coefficients, $\mathfrak{f}(\cdot)$ is the probability density function modelling directional data, and δ_ι are the function parameters of the mixture components. In this paper, we have chosen the vMF density function as $\mathfrak{f}(\cdot)$ since it is the optimal choice for characterizing multi-dimensional directional data [24]. The function $\mathfrak{F}(\cdot)$ represents the compound model and Δ is its parameter comprising of all σ_ι and δ_ι.

The main task in modelling the room reverberation directivity is the estimation of Δ by fitting the compound model $\mathfrak{F}(\cdot)$ on directional power vectors derived from the $P_R(k, \hat{\boldsymbol{y}})$ data. Before that, we determine the $P_R(k, \hat{\boldsymbol{y}})$ values by processing the incident sound field over the spherical microphone array using RIR measurements for different source positions inside the room of interest.

III. Estimation of Reverberant Field Power

In this section, we estimate the $\Gamma_{vu}(k)$ coefficients and subsequently calculate the reflected field powers $P_R(k, \hat{\boldsymbol{y}})$ for any arbitrary direction in the room using the spatial correlation method in the spherical harmonics domain [13] according to the following steps:

Step 1: Estimating spherical harmonic coefficients of the incident sound field

We obtain the incident sound field $\psi(\boldsymbol{x}_q, t)$, where t is the time-dependency, by convolving the RIR $h(\boldsymbol{x}_q, \boldsymbol{y}_o, t)$ captured by the spherical microphone element $\boldsymbol{x}_q$ with any clean signal from the source at $\boldsymbol{y}_o$. We can express the sound field transform function $\Psi(\boldsymbol{x}_q, k)$ as the Helmholtz wave equation solution to the interior sound field problem in the spherical harmonics domain [25]. Since this work uses a rigid microphone array, the radial solution coefficient $j_n(kr)$ is adjusted to obtain the modified expression of $\Psi(\boldsymbol{x}_q, k)$ as

$$\Psi(\boldsymbol{x}_q, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \alpha_{nm}(k) b_n(kr) Y_{nm}(\theta_q, \phi_q) \qquad (10)$$

where α_{nm} are the corresponding spherical harmonic coefficients. The function $b_n(kr) = j_n(kr) - \frac{j_n^{'}(kr)}{h_n^{'}(kr)} h_n(kr)$ where $j_n(\cdot)$ and $h_n(\cdot)$ denote the spherical Bessel and Hankel functions of order n, respectively, and $j_n^{'}(\cdot)$ and $h_n^{'}(\cdot)$ are their corresponding first derivatives.

Using the orthogonal property of spherical harmonics [26], we can derive the incident sound field coefficients $\alpha_{nm}(k)$ as

$$\alpha_{nm}(k) = \frac{\sum_{q=1}^{Q} \Psi(\boldsymbol{x}_q, k) Y_{nm}^*(\theta_q, \phi_q)}{b_n(kr)}. \qquad (11)$$

For practical implementation, we truncate (10) to an order N such that $N = \lceil kr \rceil$ and $Q \geq (N+1)^2$ to avoid the high-pass nature of higher-order Bessel functions and spatial aliasing problems [27].

Step 2: Estimating reflection power coefficients $\Gamma_{vu}(k)$ and power of reflected components $P_R(k, \hat{\boldsymbol{y}})$

From the $\alpha_{nm}(k)$ coefficients obtained using (11), we define a spatial cross-correlation matrix in the modal domain as

$$\boldsymbol{R}(k) \triangleq E\{\boldsymbol{\alpha}(k) \boldsymbol{\alpha}^H(k)\} \qquad (12)$$

where $\boldsymbol{\alpha}(k) = [\alpha_{00}(k)\ \alpha_{1-1}(k)\ \cdots\ \alpha_{NN}(k)]^T_{1\times(N+1)^2}$ and $\boldsymbol{\alpha}^H(k)$ is the Hermitian transpose of $\boldsymbol{\alpha}(k)$.

According to the spatial correlation model proposed in [13], the estimation problem of reflection power coefficients $\Gamma_{vu}(k)$ can be formulated as a matrix equation [28]

$$\tilde{\boldsymbol{r}}(k) = \boldsymbol{B}(k) \boldsymbol{p}(k) \qquad (13)$$

where

$$\tilde{\boldsymbol{r}}(k) = \begin{bmatrix} R_{0000} \\ R_{001-1} \\ \vdots \\ R_{00NN} \\ R_{1-100} \\ \vdots \\ R_{NNNN} \end{bmatrix}, \qquad (14)$$

$$\boldsymbol{B}(k) = \begin{bmatrix} b_{0000} & d_{000000} & \cdots & d_{0000VV} \\ b_{001-1} & d_{001-100} & \cdots & d_{001-1VV} \\ \vdots & \vdots & \vdots & \vdots \\ b_{00NN} & d_{00NN00} & \cdots & d_{00NNVV} \\ b_{1-100} & d_{1-10000} & \cdots & d_{1-100VV} \\ \vdots & \vdots & \vdots & \vdots \\ b_{NNNN} & d_{NNNN00} & \cdots & d_{NNNNVV} \end{bmatrix}, \qquad (15)$$

$$\boldsymbol{p}(k) = \begin{bmatrix} P_D \\ \Gamma_{00} \\ \Gamma_{1-1} \\ \vdots \\ \Gamma_{V-V} \\ \vdots \\ \Gamma_{VV} \end{bmatrix}, \qquad (16)$$

$R_{nmn'm'}$ is the $(n^2+n+m+1)^{th}$ row and $(n'^2+n'+m'+1)^{th}$ column component of $\boldsymbol{R}(k)$,

$b_{nmn'm'} = (i)^{n-n'} Y_{nm}^*(\theta_o, \phi_o) Y_{n'm'}(\theta_o, \phi_o)$ can be calculated based on source location,

$d_{nmn'm'vu} = (i)^{n-n'}(-1)^{-m}\sqrt{\frac{(2v+1)(2n+1)(2n'+1)}{4\pi}} W_1 W_2$, where W_1 and W_2 are Wigner 3j symbols [29] given by $W_1 = \begin{pmatrix} v & n & n' \\ 0 & 0 & 0 \end{pmatrix}$ and $W_2 = \begin{pmatrix} v & n & n' \\ u & -m & m' \end{pmatrix}$, and

$P_D = E\{|S(k)|^2 |G_D(k, \hat{\boldsymbol{y}})|^2\}$ is the direct path power.

Since $\boldsymbol{B}(k)$ and $\tilde{\boldsymbol{r}}(k)$ have known elements, we can estimate $\boldsymbol{p}(k)$ which contains the desired reflection power coefficients Γ_{vu} by solving (13) using the least-squares method given by

$$\hat{\boldsymbol{p}}(k) = \boldsymbol{B}^\dagger(k) \tilde{\boldsymbol{r}}(k) \qquad (17)$$

where $[\cdot]^\dagger$ and $\hat{[\cdot]}$ represent pseudo-inverse and estimated values, respectively.

From the estimated $\hat{\boldsymbol{p}}(k)$ array, we filter the Γ_{vu} elements and substitute in (8) to find the reflected powers $P_R(k, \hat{\boldsymbol{y}})$ for different incoming $\hat{\boldsymbol{y}}$ directions across the 3D room space. To avoid an undetermined system while solving (17), we truncate (8) to an order V such that $V = \left\lfloor \sqrt{(N+1)^4 - 1} \right\rfloor$ [13].

IV. Modelling of Reverberant Field Directivity using vMF Mixture Distribution

The $P_R(k, \hat{\boldsymbol{y}})$ values for different $\hat{\boldsymbol{y}}$ sampled across the room define the directional characteristics of the room reverberant field. The proposed method uses the multivariate vMF mixture distribution to extract this information for a faithful representation of the field directivity.

A. Multivariate vMF Mixture Distribution

The vMF distribution is the most straightforward parametric distribution for modelling multi-dimensional directional data distributed over a unit hypersphere [24]. The probability density function of a D-variate vMF distribution is given by

$$f_{\text{vMF}}(\mathbf{z}; \boldsymbol{\mu}, \kappa) = \frac{\kappa^{\left(\frac{D}{2}-1\right)}}{(2\pi)^{\frac{D}{2}} I_{\left(\frac{D}{2}-1\right)}(\kappa)} \exp(\kappa \boldsymbol{\mu}^T \mathbf{z}) \tag{18}$$

where $\mathbf{z}$ is a D-dimensional random unit vector, $\boldsymbol{\mu}$ is the mean direction vector, κ is the concentration parameter, and $I_n(\cdot)$ is the modified Bessel function of the first kind at order n. The function should also satisfy the constraints: $||\boldsymbol{\mu}|| = 1$, $\kappa \geq 0$ and $D \geq 2$. It has minimal parameters compared to other directional models and can be used to design closed-form expressions [16]. Also, the vMF distribution allows decomposition of any function defined on a hypersphere using the property that the product of two vMF distributions is another unnormalized vMF [16], [24].

In accordance with (9), we require a convex mixture of unimodal vMFs to model the angular power distribution of the room reverberant field. Therefore, $P_R(k, \hat{\boldsymbol{y}})$ follows a vMF mixture distribution given by

$$\mathsf{f}(\mathbf{z}; \boldsymbol{M}, \boldsymbol{\kappa}, \boldsymbol{w}) = \sum_{a=1}^{A} w_a f_{\text{vMF}}(\mathbf{z}; \boldsymbol{\mu}_a, \kappa_a) \tag{19}$$

where A is the total number of components in the mixture model $\mathsf{f}(\cdot)$, $\boldsymbol{M} = \{\boldsymbol{\mu}_a\}_{a=1}^{A}$ are the mean vectors of $\mathsf{f}(\cdot)$, $\boldsymbol{\kappa} = \{\kappa_a\}_{a=1}^{A}$ are the concentration parameters of $\mathsf{f}(\cdot)$, and $\boldsymbol{w} = \{w_a\}_{a=1}^{A}$ are the weighting factors that satisfy the constraints $w_a \geq 0$ and $\sum_{a=1}^{A} w_a = 1$.

The function $\mathsf{f}(\mathbf{z}; \boldsymbol{M}, \boldsymbol{\kappa}, \boldsymbol{w})$ forms the compound model $\mathfrak{F}(k, \hat{\boldsymbol{y}}; \varDelta)$. Consequently, the main task of estimating $\varDelta$ involves the fitting of $\mathsf{f}(\cdot)$ on $P_R(k, \hat{\boldsymbol{y}})$ data to find the parameters $\boldsymbol{M}$, $\boldsymbol{\kappa}$ and $\boldsymbol{w}$.

B. Fitting the vMF Mixture Model

Initially, we follow the method described in Section III to form a comprehensive training data set $\{P_R(k, \hat{\boldsymbol{y}})^{(j)}\}_{j=1}^{J}$ from $P_R(k, \hat{\boldsymbol{y}})$ values corresponding to different $\hat{\boldsymbol{y}}$ directions and source positions. Before using this data set to train the vMF mixture model $\mathsf{f}(\cdot)$, we have to convert each $P_R(k, \hat{\boldsymbol{y}})^{(j)}$ into a unit vector conveying directional reflection power magnitude using the rules of coordinate geometry.

The first step in this data transformation is the normalization of $\{P_R(k, \hat{\boldsymbol{y}})^{(j)}\}_{j=1}^{J}$ to a range of $[0, 1]$. If the elevation and azimuth angles of the look direction $\hat{\boldsymbol{y}}$ are θ and ϕ, respectively, then the unit-vector components of $P_R(k, \hat{\boldsymbol{y}})^{(j)}$ are defined as:

$$z_1^{(j)} = |P_R(k, \hat{\boldsymbol{y}})^{(j)}| \sin\theta \cos\phi, \tag{20a}$$

$$z_2^{(j)} = |P_R(k, \hat{\boldsymbol{y}})^{(j)}| \sin\theta \sin\phi, \tag{20b}$$

$$z_3^{(j)} = |P_R(k, \hat{\boldsymbol{y}})^{(j)}| \cos\theta, \tag{20c}$$

$$z_4^{(j)} = \sqrt{1 - \left[\left(z_1^{(j)}\right)^2 + \left(z_2^{(j)}\right)^2 + \left(z_3^{(j)}\right)^2\right]}. \tag{20d}$$

Thus the reverberant field power $P_R(k, \hat{\boldsymbol{y}})^{(j)}$ is converted into a 4D unit vector $\boldsymbol{z}^{(j)} = \left[z_1^{(j)}, z_2^{(j)}, z_3^{(j)}, z_4^{(j)}\right]^T$ and $\{P_R(k, \hat{\boldsymbol{y}})^{(j)}\}_{j=1}^{J}$ is transformed to a set of independent and identically distributed vectors, denoted by $\mathbb{Z} = \{\boldsymbol{z}^{(j)}\}_{j=1}^{J}$.

We use $\mathbb{Z}$ as the training data to estimate the parameters of a 4-variate vMF mixture model using the Expectation-Maximization (EM) algorithm given in [30]. It is also integrated with the clustering algorithm based on the Bayesian Minimum Message Length (MML) criterion proposed in [31] to decide the number of mixture components that best fit $\mathbb{Z}$. Once the parameters are estimated, we use the vMF mixture model $\mathsf{f}(\mathbf{z}; \boldsymbol{M}, \boldsymbol{\kappa}, \boldsymbol{w})$ to visualize the angular power distribution of the room reverberant field.

V. Experimental Results

This section presents the experimental setup to acquire the RIRs and evaluates the performance of the proposed directivity model for a rectangular test room of size $3.7 \times 4.2 \times 2.7$ m. The upper-half portions of the right and front walls of this room had glass surfaces. We expect the angular power distribution model to identify the reflection variations from various objects across the room boundary, particularly the significant reflections from the glass surfaces.

The experiment used an em32 Eigenmike [32], the 32-element rigid spherical microphone array with radius $r = 0.042$ m, to measure the RIRs. We placed the Eigenmike at the geometric center of the room and considered the center of the receiver array as the origin O of the coordinate system adopted for our model shown in Fig. 1. The RIRs were measured with three repeats of a 2-second linear sweep signal emitted from a source placed at a distance $r_o = 1$ m from O and the same elevation as the Eigenmike, i.e., in the XY plane with $\theta_o = \frac{\pi}{2}$. We repeated the measurements for different source azimuth positions $\phi_o \in [0, 2\pi)$ with 20^o increments to obtain a range of RIRs.

In the next stage of the experiment, we convoluted a 3-second clean speech signal sampled at 48 kHz with a 32-channel 19200-tap RIR sequence to obtain the incident sound field $\psi(\cdot, t)$. We transformed it into the frequency-domain sound field $\varPsi(\cdot, k)$ through STFT to find the $\alpha_{nm}(k)$ coefficients using (11). The frequency components ranging from 2000 Hz to 5000 Hz were selected for the subsequent processing according to Section III to deduce $P_R(k, \hat{\boldsymbol{y}})$ values for $\phi \in [0, 2\pi)$ with 2^o increments and $\theta = \frac{\pi}{2}$.

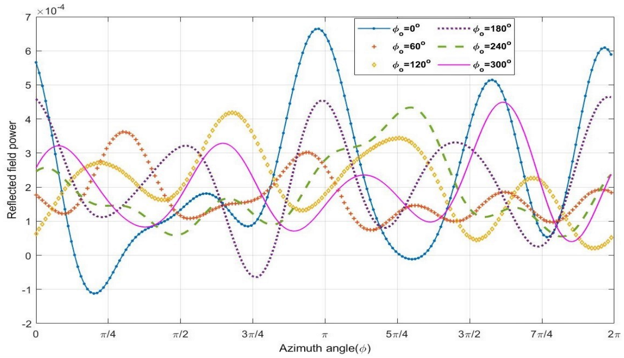


Fig. 2: Reverberant field power in the XY plane for different source positions

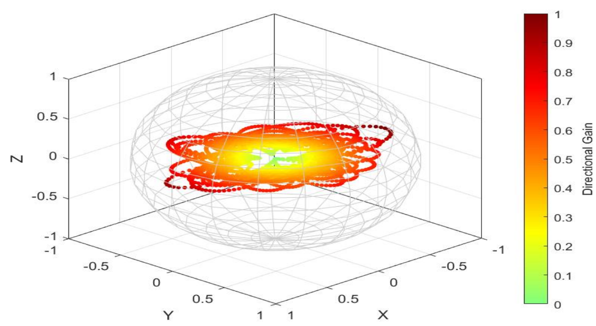


Fig. 3: $\{\boldsymbol{z}^{(j)}\}_{j=1}^{J}$ visualized as spherical data distribution

For the elementary analysis presented in this paper, we have averaged the reflected power values $P_R(k, \hat{\boldsymbol{y}})$ across the frequencies for each $\hat{\boldsymbol{y}}$ direction in the XY plane. Henceforth, the frequency dependency (k) is omitted, and we treat the reverberant field power $P_R(\hat{\boldsymbol{y}})$ as a function of direction. We repeated this process for all the measured RIRs to create a comprehensive training data set $\{P_R(\hat{\boldsymbol{y}})^{(j)}\}_{j=1}^{J}$.

A. Training Data

The simulation results in Fig. 2 show the $P_R(\hat{\boldsymbol{y}})$ plots across the XY plane of the room for a few different source positions. The magnitude variations in the reflection power profile convey the difference in reflective properties across the walls, and its distinct peaks indicate the relative positions of highly-reflecting surfaces inside the room with respect to the source location. If we compare the results of different source locations, the profiles exhibit similar shape but with relative shifts and different peak magnitudes due to the change in displacements between the source and reflecting surfaces. This behavior suggests a fixed reference position of the reflecting surfaces but random reflection paths. This randomness is exploited by the vMF model to extract the directivity pattern of the room reverberant field. The similarity in reflection profiles also reveals the redundancy in the RIR measurements. Hence a proper subset of the RIRs can still produce the same angular power distribution.

The $P_R(k, \hat{\boldsymbol{y}})$ values estimated based on the method given in [13] is only a fraction of the actual reflected power. It shows a few negative values, as visible in Fig. 2, due to the truncation of the infinite summation in (8). However, the truncation error is minimum at high values of V, and since we are interested only in the relative power distribution across the room, this problem is solved by the data transformation mentioned in Section IV. Thus, in the penultimate step of the experiment, we converted $\{P_R(\hat{\boldsymbol{y}})^{(j)}\}_{j=1}^{J}$ into the directional power vectors $\{\boldsymbol{z}^{(j)}\}_{j=1}^{J}$ before introducing to the vMF-mixture parameter estimation algorithm. The training vectors are visualized as spherical data in Fig. 3. The position of each data point corresponds to the direction $\hat{\boldsymbol{y}}$ and the normalized directional power $|P_R(\hat{\boldsymbol{y}})^{(j)}|$ is color-coded according to the scale shown in the figure.

B. Reverberant Field Directivity Model

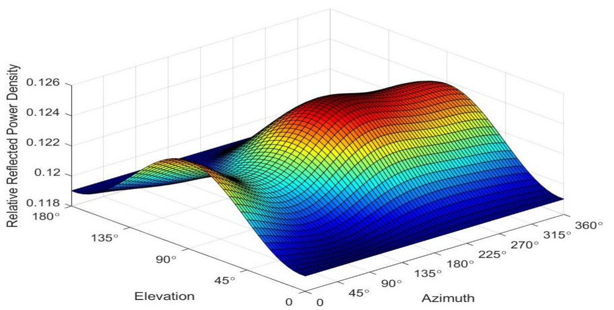
In the final step, we use the vMF mixture model to interpret the directivity information from the distributed training vectors. The algorithm fitted a 4-component vMF mixture model, and the estimated parameters are recorded under Table.I. Thus the proposed approach condensed the extensive sound field information from eighteen 32-channel 19200-tap RIR sequences into the mean $\boldsymbol{M}_{[4\times4]}$, concentration $\boldsymbol{\kappa}_{[4\times1]}$ and weight $\boldsymbol{w}_{[4\times1]}$ parameters of the vMF mixture function. Finally, we substituted the estimated parameters in (19) to generate the power density for testing vectors $\mathbf{z}$ corresponding to arbitrary directions sampled across the room space.

The resulting angular power density, shown in Fig. 4, illustrates the directional distribution of room reverberation by identifying the regions of varying reflectances. The highest reflection density or dominant reverberation direction coincides with the function maximum around $\phi = 225^o$, which corresponds to the glass surface points on the front wall. We can also observe relatively high magnitudes across $\phi \in [0^o, 45^o] \cap [315^o, 360^o)$ indicating considerable reflections emanating from the glass surface portions on the right wall. The remaining lower magnitude areas are probably the result of normal wall reflections and higher-order reflections from the glass surfaces. The miscellaneous objects present in the test room also influence this distribution. The directions around $\phi = 90^o$ seem to be composed of comparatively absorbent surfaces. Hence, the directivity model in Fig. 4 conforms with the real test room environment and recognizes the directions of distinct reflective surfaces.

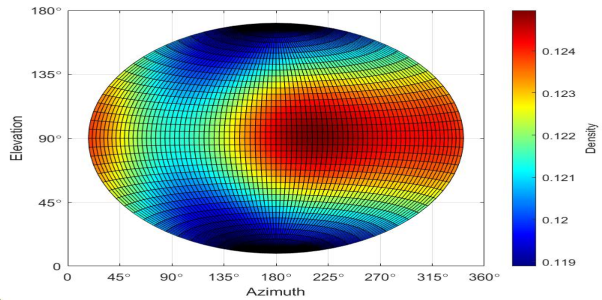
We can create a more intricate directivity model using the training data corresponding to $\hat{\boldsymbol{y}}$ and source positions across the XYZ plane. It also allows the visualization of the frequency-dependent behavior of the reverberant field using

TABLE I: Estimated parameters of vMF mixture distribution

a	$\boldsymbol{\mu}_a$	κ_a	w_a
1	[0.46, 0.02, 0, 0.89]	0.9684	0.2138
2	[-0.03, 0.31, 0, 0.95]	0.9752	0.2991
3	[0.02, -0.37, 0, 0.93]	0.9698	0.2743
4	[-0.47, -0.05, 0, 0.88]	0.9616	0.2128



(a) Angular power density generated from the vMF mixture function



(b) 2D projection of reverberation directivity

Fig. 4: Room reverberant field directivity represented using vMF mixture distribution

the $P_R(k, \hat{\boldsymbol{y}})$ data of each wavenumber k as the training input into the vMF-mixture parameter estimation algorithm.

The processing of reflection power coefficients, training, and production of the directivity model took less than 10 s for the mentioned dataset on an ordinary laptop with an Intel Core i5 1.6GHz CPU and 8GB RAM, whereas the latest techniques of reverberant scene reproduction using image processing and neural networks [14], [15] take a few minutes even on a higher grade processor. Moreover, our model is based on the real acoustic information obtained from the RIRs, which avoids the need for large feature-matching databases for an accurate reverb environment evaluation. The application of the vMF mixture function allows statistical expression of spatial correlation between the higher-order Eigenbeams. It potentially contains in-depth information about the spatial and frequency-related properties of numerous reflection channels inside the room, which were not conceivable earlier using the conventional methods [4]–[13]. We will utilize this information in future works to analyze the frequency-dependent room acoustic behavior under different room environments, source-receiver conditions, and other external factors.

VI. Conclusion

In this paper, we presented an angular power distribution model based on vMF mixture function to represent the room reverberation directivity. The experimental results prove the viability of the model in recognizing significant reflections across the room while overcoming the shortcomings of existing models. The proposed approach reduces data size and redundancy by encapsulating the information from a large number of long RIR sequences into a few parameters of the vMF density function. Therefore, we get a simple model with manageable data requirements for analyzing the directional characteristics of room reverberation.

The future works involve the analysis of high-frequency behavior of room reverberation field and the development of further functionalities to evolve as a reliable statistical room acoustic model. The ability of the directivity model to identify reflecting surfaces can be further extended to estimate wall impedance variations and for 3D room geometry projection in VR algorithms.

References

[1] H. Kuttruff, *Room Acoustics*, 6th ed. Boca Raton, FL: CRC Press, 2017.

[2] *Acoustics — Measurement of room acoustic parameters — Part 1: Performance spaces*, ISO 3382-1:2009, Int. Standards Org., Geneva, Switzerland, 2009.

[3] J. S. Bradley, "Review of objective room acoustics measures and future needs," *Applied Acoust.*, vol. 72, no. 10, pp. 713–720, Oct. 2011.

[4] S. Tervo, T. Korhonen, and T. Lokki, "Estimation of reflections from impulse responses," *Building Acoust.*, vol. 18, no. 1-2, pp. 159–173, 2011.

[5] S. Jaeger and C. Allen, "Two-dimensional sound intensity analysis of jet noise," in *15th Aeroacoustics Conf.*, 1993, p. 4342.

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

[7] B. N. Gover, J. G. Ryan, and M. R. Stinson, "Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array," *J. Acoust. Soc. Amer.*, vol. 116, no. 4, pp. 2138–2148, 2004.

[8] H. Sun, E. Mabande, K. Kowalczyk, and W. Kellermann, "Localization of distinct reflections in rooms using spherical microphone array eigenbeam processing," *J. Acoust. Soc. Amer.*, vol. 131, no. 4, pp. 2828–2840, 2012.

[9] H. Sun, E. Mabande, K. Kowalczyk, and W. Kellermann, "Joint DOA and TDOA estimation for 3D localization of reflective surfaces using eigenbeam MVDR and spherical microphone arrays," in *2011 IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2011, pp. 113–116.

[10] B. Rafaely, "Plane-wave decomposition of the sound field on a sphere by spherical convolution," *J. Acoust. Soc. Amer.*, vol. 116, no. 4, pp. 2149–2157, Oct. 2004.

[11] M. Park and B. Rafaely, "Sound-field analysis by plane-wave decomposition using spherical microphone array," *J. Acoust. Soc. Amer.*, vol. 118, no. 5, pp. 3094–3103, 2005.

[12] Y. Hioka, K. Niwa, S. Sakauchi, K. Furuya, and Y. Haneda, "Estimating Direct-to-Reverberant Energy Ratio Using D/R Spatial Correlation Matrix Model," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 8, pp. 2374–2384, Apr. 2011.

[13] P. N. Samarasinghe and T. D. Abhayapala, "Blind estimation of directional properties of room reverberation using a spherical microphone array," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2017, pp. 351–355.

[14] H. Kim, L. Hernaggi, P. J. Jackson, and A. Hilton, "Immersive Spatial Audio Reproduction for VR/AR Using Room Acoustic Modelling from 360° Images," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces*, 2019, pp. 120–126.

[15] L. Remaggi, A. Neidhardt, A. Hilton, and J. B. J. Philip, "Perceived quality and spatial impression of room reverberation in VR reproduction from measured images and acoustics," in *Proc. 23rd Int. Congr. Acoust.*, Aachen, Germany, 2019.

[16] T. McGraw, B. C. Vemuri, B. Yezierski, and T. Mareci, "von Mises-Fisher mixture model of the diffusion ODF," in *3rd IEEE Int. Symp. Biomedical Imaging: Nano to Macro*, 2006, pp. 65–68.

[17] K. V. Mardia, C. C. Taylor, and G. K. Subramaniam, "Protein Bioinformatics and Mixtures of Bivariate von Mises Distributions for Angular Data," *Biometrics*, vol. 63, no. 2, pp. 505–512, June 2007.

[18] R. E. Røge, K. H. Madsen, M. N. Schmidt, and M. Mørup, "Infinite von Mises–Fisher Mixture Modeling of Whole Brain fMRI Data," *Neural Computation*, vol. 29, no. 10, pp. 2712–2741, Oct. 2017.

[19] N. Masseran, A. M. Razali, K. Ibrahim, and M. T. Latif, "Fitting a mixture of von Mises distributions in order to model data on wind direction in Peninsular Malaysia," *Energy Conversion Manage.*, vol. 72, pp. 94–102, Aug. 2013.

[20] K. Hara, K. Nishino, and K. Ikeuchi, "Mixture of Spherical Distributions for Single-View Relighting," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30, no. 1, pp. 25–35, Nov. 2007.

[21] M. Hasnat, J. Bohné, J. Milgram, S. Gentric, and L. Chen, "von Mises-Fisher Mixture Model-based Deep learning: Application to Face Verification," Dec. 2017. [Online]. Available: https://arxiv.org/abs/1706.04264v2

[22] J. Taghia, Z. Ma, and A. Leijon, "On Von-Mises Fisher Mixture Model in Text-Independent Speaker Identification," in *Proc. 14th Annu. Conf. Int. Speech Commun. Assoc., INTERSPEECH*, 2013, pp. 2499–2503.

[23] Z. Ma and A. Leijon, "Statistical Speech Model Description with VMF Mixture Model," Aug. 2018. [Online]. Available: https://arxiv.org/abs/1808.00960

[24] I. S. Dhillon and S. Sra, "Modeling Data using Directional Distributions," The University of Texas at Austin, Tech. Rep. TR-03-06, Jan. 2003.

[25] P. Samarasinghe, "Modal based Solutions for the Acquisition and Rendering of Large Spatial Soundfields," Ph.D. dissertation, College of Eng. and Comp. Sci., The Australian National Univ., Canberra, Australia, 2014.

[26] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. Academic Press, 1999.

[27] A. Fahim, P. N. Samarasinghe, and T. D. Abhayapala, "PSD Estimation and Source Separation in a Noisy Reverberant Environment Using a Spherical Microphone Array," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 9, pp. 1594–1607, Sept. 2018.

[28] P. N. Samarasinghe, T. D. Abhayapala, and H. Chen, "Estimating the Direct-to-Reverberant Energy Ratio Using a Spherical Harmonics-Based Spatial Correlation Model," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 25, no. 2, pp. 310–319, Feb. 2016.

[29] F. W. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions*. New York, NY, USA: Cambridge Univ. Press, 2010.

[30] A. Banerjee, I. S. Dhillon, J. Ghosh, and S. Sra, "Clustering on the Unit Hypersphere using von Mises-Fisher Distributions," *J. Machine Learning Research*, vol. 6, pp. 1345–1382, Sept. 2005.

[31] P. Kasarapu and L. Allison, "Minimum message length estimation of mixtures of multivariate Gaussian and von Mises-Fisher distributions," *Machine Learning*, vol. 100, no. 2-3, pp. 333–378, Mar. 2015.

[32] *em32 Eigenmike® microphone array release notes (v17.0)*, mh acoustics, Summit, NJ, USA, Oct. 2013. [Online]. Available: https://mhacoustics.com/sites/default/files/ReleaseNotes.pdf