Spatial Acoustics Library for MATLAB (SALM): A Computational Toolkit for Spatial Audio Processing

César D. Salvador
Research Department and
School of Electronic Engineering, Faculty of Engineering
Universidad Peruana de Ciencias Aplicadas
Lima, Peru
cesar.salvador@upc.edu.pe

Jorge Treviño
Research Institute of Electrical Communication (RIEC) and
Graduate School of Information Sciences (GSIS)
Tohoku University
Sendai, Japan
jorge.trevino@music.yamaha.com

Shuichi Sakamoto
Research Institute of Electrical Communication (RIEC) and
Graduate School of Information Sciences (GSIS)
Tohoku University
Sendai, Japan
saka@ais.riec.tohoku.ac.jp

Abstract—The Spatial Acoustics Library for MATLAB (SALM) is a computational toolkit for efficient spatial audio signal processing. Its principal contribution is providing a unified and extensible framework that integrates Fourier transform techniques for spatial filtering. A defining characteristic of SALM is its ability to seamlessly handle both fundamental tasks, such as visualizing spherical harmonic functions, and advanced applications, including array processing in the circular and spherical Fourier transform domains. Key use cases include spatial interpolation of head-related transfer functions and room impulse responses, essential for high-fidelity spatial sound reproduction. Rigorously tested and validated in previous publications by the authors, SALM has demonstrated its reliability for immersive acoustic research and prototyping. By equipping researchers and developers with a structured and versatile library, SALM advances spatial acoustics with applications in binaural rendering for virtual reality and sound field analysis for architectural acoustics.

Index Terms—Spatial audio, spherical Fourier transform, head-related transfer function, room impulse response.

I. INTRODUCTION

The spatial acoustics library for MATLAB (SALM) is a computational toolkit for rapid prototyping of audio signal processing pipelines in the spatial and transform domains. SALM provides a unified and extensible framework that supports both fundamental analysis and advanced algorithm development. Designed for experimental workflows, SALM emphasizes reproducibility, clarity, and modularity. Developed from the foundational work of the authors at Tohoku University, it integrates signal processing algorithms into a coherent MATLAB-based environment suitable for both rapid prototyping and long-term research.

Table I outlines existing MATLAB/Octave libraries relevant to spatial audio. Each library has unique strengths and has advanced spatial audio functions. Within this landscape, SALM complements previous efforts by offering a unified framework that integrates both analysis and synthesis operations consistently across spatial and transform domains. This integration allows for: 1) direct manipulation of acoustic data, such as head-related transfer functions (HRTF) and room impulse responses (RIR), in the spatial domain, and 2) advanced filtering operations of acoustic data in the transform domain. This enables coherent workflows for advanced tasks such as sound field translation, array signal conversion, angle interpolation, and distance extrapolation. Furthermore, SALM has been rigorously tested and validated in multiple peer-reviewed studies by the authors [1]–[12], demonstrating its reliability for immersive audio research and rapid prototyping.

At its core, SALM integrates circular and spherical Fourier-based techniques for spatial filtering and sound field manipulation. A defining characteristic of the library is its ability to seamlessly handle a wide range of tasks, from basic operations such as visualizing spherical harmonic functions, to advanced applications that require array signal processing in circular and spherical Fourier transform domains. A key strength of SALM is its ease of use and integration. No installation or compilation is required; users simply add the SALM folder to the MATLAB path, save it for future sessions, and begin calling functions directly. The library is implemented entirely in native syntax and does not depend on external toolboxes. It has been successfully tested on MATLAB versions from 2019 onward, ensuring compatibility with recent releases.

New users are encouraged to start with the included example scripts, which illustrate typical use cases and provide a practical introduction to the core functionality of the library. For

Library	Description	Spatial-domain processing	Transforms	Transform-domain processing
SOFiA Toolbox [13]	MATLAB library to analyze a sound field captured with a microphone array.	Microphone-array handling; spherical-grid management; visualization of measured fields.	Spherical Fourier transforms.	Modal beamforming; plane-wave decomposition; radial filtering for rigid/open spheres.
Sound field synthesis toolbox [14]	MATLAB/Octave library to synthesize a sound field in an area surrounded by a loudspeaker array.	Driving functions for loudspeaker-array rendering; convolution engine.	Spherical Fourier transforms.	Modal driving filters; binaural rendering from arrays.
AKtools [15]	MATLAB library for the capture, processing, analysis and rendering of spatial audio signals.	Signal generation; convolution engine; room analysis.	Spherical Fourier transforms.	Radial filtering; binaural rendering.
ITA Toolbox [16]	MATLAB toolbox for acoustic measurements and audio signal processing.	Impulse-response measurement; convolution engines for auralization.	Spherical Fourier transforms.	Radial filtering; modal analysis.
SALM	MATLAB library for sound field analysis, processing and synthesis with circular and spherical arrays.	Diffuse-field filter for equalization; free-field translation operator for acoustic centering.	Circular, semicircular, and spherical Fourier transforms.	Distance-varying filter (DVF) for radial extrapolation; boundary-matching filter (BMF) for array signal conversion.

instance, the following scripts exemplify two core functions through their geometrical visualization:

- `exampleYnm.m` — Visualizes real-valued spherical harmonic functions on the unit sphere, demonstrating their angular structure and orthogonality.
- `exampleSFT.m` — Demonstrates the application of the spherical Fourier transform (SFT) and its inverse to noisy functions defined on the surface of the unit sphere, highlighting the filtering and reconstruction accuracy.

This paper is organized as follows. Section II describes the overall structure of the library. Section III outlines the mathematical foundations underlying the core functions. Section IV presents two representative use cases: diffuse-field equalization of HRTFs and distance extrapolation of HRTFs. Section V briefly announces future work. Finally, Section VI provides the conclusion.

II. STRUCTURE OF SALM

SALM is structured to provide users with a modular, mathematically grounded, and application-oriented toolkit for the analysis and processing of spatial audio data on spherical domains. It comprises a hierarchy of functions organized according to their domain of operation and mathematical specialization. Figure 1 illustrates the overall structure.

A. Spatial-Domain Functions

These functions operate directly on impulse responses or transfer functions in the spatial domain:

- `diffuseFieldFilter`: designs equalization filters based on the spatial average of a set of transfer functions, commonly used for diffuse-field flattening.
- `freeFieldtranslationOperator`: implements sound field translation under free-field conditions, using either plane-wave or spherical-wave models depending on the scenario.

B. Special Functions

These core analytical functions are fundamental to modal representations and are used throughout the library:

- `pnm`: associated Legendre polynomials, used in spherical harmonic evaluations.
- `ynm`: real-valued and complex-valued spherical harmonic functions with multiple normalization factors.
- `besseljsph`: spherical Bessel function of the first kind.
- `besselysph`: spherical Bessel function of the second kind, also known as the spherical Neumann function.
- `besselhsph`: spherical Hankel functions of the first and second kind, frequently used in scattering and radiation problems.
- `dbesseljsph`: derivative of `besseljsph`.
- `dbesselhsph`: derivative of `besselhsph`.

C. Transform Functions

These functions facilitate conversions between spatial and modal (transform) domains:

- `cft` / `icft`: compute the forward and inverse circular Fourier transforms.
- `flt` / `iflt`: implement the forward and inverse semicircular Fourier–Legendre transforms.
- `sft` / `isft`: perform the forward and inverse spherical Fourier transforms.
- `pinvreg`: calculates a Tikhonov-regularized pseudoinverse, which is particularly effective in stabilizing ill-posed inverse problems that arise when computing circular and spherical transforms.

D. Transform-Domain Functions

These functions apply spatial filtering directly to modal coefficients:

- `dvf`: generates distance-varying filters (DVF) to extrapolate near-field HRTFs from measurements at a fixed radius.
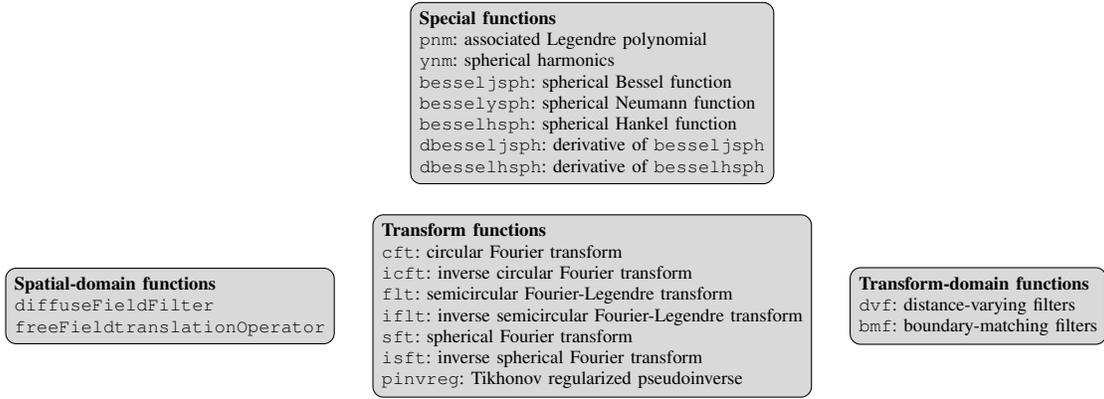
Fig. 1. Structure of SALM.

- `bmf`: implements boundary-matching filters (BMF) for reconstructing sound fields from spatial samples.

This layered architecture allows SALM to maintain both flexibility and mathematical rigor, enabling advanced spatial audio workflows that span from theoretical modeling to practical signal processing implementations.

III. MATHEMATICAL FOUNDATIONS

This section introduces the key mathematical concepts underlying the SALM library, in accordance with the structure described in Section II. The aim is to provide the theoretical foundation that supports each functional layer.

Figure 2 shows the vertical-polar spherical coordinate system used throughout this document. In these coordinates, a point in space $\vec{r} = (r, \theta, \phi)$ is specified by its radial distance r, azimuthal angle $\theta \in [-\pi, \pi]$, and elevation angle $\phi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. For brevity, the angles are merged into the variable $\Omega = (\theta, \phi)$ in such a way that a point in space is also represented by $\vec{r} = (r, \Omega)$. Conversions between Cartesian coordinates and vertical-polar spherical coordinates can be performed using the native MATLAB functions `cart2sph.m` and `sph2cart.m`.

To emphasize a listener position, it is convenient to place the center of their head at the origin of the spherical coordinate system. Figure 3 shows a generic head model. The positive x–axis points to the front of the listener. The y–axis is the interaural axis connecting the eardrums. The center of the head lies at the midpoint between the eardrums and coincides with the center of coordinates. Radial distances are measured from the center of the head, whereas positive azimuthal angles are measured from the positive x–axis toward the left.

SALM also supports the interaural-polar spherical coordinate system. Conversions between Cartesian coordinates and interaural-polar spherical coordinates can be performed using the SALM functions `cart2isph.m` and `isph2cart.m`. For details of interaural-polar spherical coordinates, we refer the reader to [1].

A. Spatial-Domain Functions

1) Diffuse-Field Equalization: When dealing with a spherical set of transfer functions, diffuse-field equalization [17],
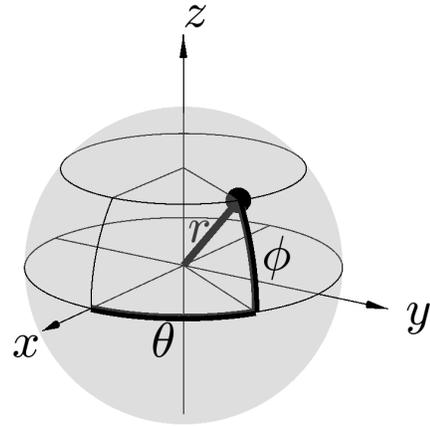

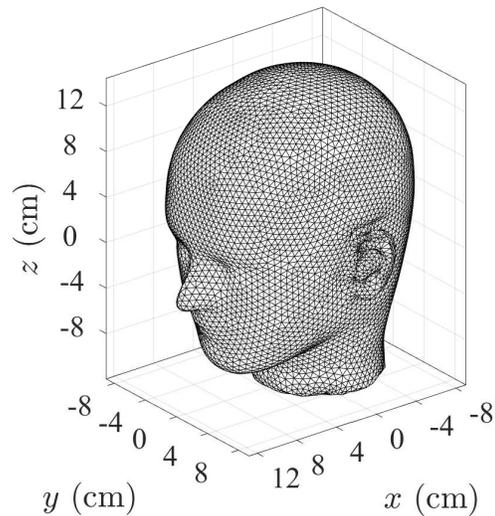Fig. 2. Vertical-polar spherical coordinate system.
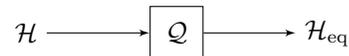

Fig. 3. Head model.

Fig. 4. Diffuse-field equalization.

[18] aims at removing the common component along the direction in such a way that no direction is perfectly characterized, but also no direction is fully suppressed.

Figure 4 shows the diffuse-field equalization process. Given an acoustic transfer function $\mathcal{H}$, its equalized version $\mathcal{H}_{\text{eq}}$ is

$$\mathcal{H}_{\text{eq}} = \mathcal{Q}\mathcal{H}. \tag{1}$$

Calculating the diffuse-field equalizer $\mathcal{Q}$ requires to obtain beforehand a spherical dataset of acoustic transfer functions $\{\mathcal{H}_\ell\}$, $\ell = 1, \ldots, L$, where L is the number of directions. The diffuse-field equalizer is defined by

$$\mathcal{Q} = |\mathcal{Q}| \exp(j\angle\mathcal{Q}). \tag{2}$$

The magnitude of the equalizer is obtained as

$$|\mathcal{Q}| = \left(\frac{1}{\sum_{\ell=1}^{L} |\mathcal{H}_\ell|^2 w_\ell} \right)^{\frac{1}{2}}, \tag{3}$$

where $w_\ell > 0$ are weighting quadratures for integration on the sphere fulfilling the normalization $\sum_\ell w_\ell = 1$. The phase of the equalizer can be of two types:

$$\angle\mathcal{Q} = \begin{cases} 0, & \text{zero-phase}, \\ \mathfrak{H}\left\{ \log \frac{1}{|\mathcal{Q}|} \right\}, & \text{minimum-phase}, \end{cases} \tag{4}$$

where $\mathfrak{H}$ denotes the Hilbert transform.

2) Free-Field Translation Operator: When acoustic data are sparsely measured on a spherical surface and analyzed using circular or spherical transforms, the center of the orthonormal functions does not necessarily coincide with the center of the spherical measurement surface. Acoustic centering [19]–[23] is the process of aligning the measurement and analysis centers by means of an acoustic translation operator $\mathcal{T}$.

Consider the following positions in Cartesian coordinates: $\vec{x}_{\text{s}}$ denotes the position of a sound source outside the origin, $\vec{x}_{\text{i}}$ denotes the initial position at which the radiated sound field is observed, and $\vec{x}_{\text{f}}$ denotes the final position of observation. In practice, $\vec{x}_{\text{i}}$ and $\vec{x}_{\text{f}}$ represent the measurement center and the analysis center, respectively. Two models of $\mathcal{T}$ that translates the sound field from $\vec{x}_{\text{i}}$ to $\vec{x}_{\text{f}}$ in free-field are described below: a plane-wave model and a spherical-wave model.

Let $\hat{x}_{\text{s}}$ be a unit vector in the direction of $\vec{x}_{\text{s}}$. A plane-wave traveling in the direction $-\hat{x}_{\text{s}}$ is observed at $\vec{x}_{\text{i}}$ and $\vec{x}_{\text{f}}$ as $\exp(jk\hat{x}_{\text{s}} \cdot \vec{x}_{\text{i}})$ and $\exp(jk\hat{x}_{\text{s}} \cdot \vec{x}_{\text{f}})$, respectively. The free-field translation operator from $\vec{x}_{\text{i}}$ to $\vec{x}_{\text{f}}$ is defined as a ratio of the plane-wave observations:

$$\mathcal{T}_{\text{PW}}(\vec{x}_{\text{i}}, \vec{x}_{\text{f}}) = \exp\left(jk\hat{x}_{\text{s}} \cdot (\vec{x}_{\text{f}} - \vec{x}_{\text{i}}) \right). \tag{5}$$

The inverse translation operator (from $\vec{x}_{\text{f}}$ to $\vec{x}_{\text{i}}$) is the reciprocal of (5):

$$\mathcal{T}_{\text{PW}}^{-1} = \frac{1}{\mathcal{T}_{\text{PW}}}. \tag{6}$$

A spherical wave front emanating from $\vec{x}_{\text{s}}$ is observed at $\vec{x}_{\text{i}}$ and $\vec{x}_{\text{f}}$, respectively, as $\frac{e^{-jk\|\vec{x}_{\text{s}} - \vec{x}_{\text{i}}\|}}{\|\vec{x}_{\text{s}} - \vec{x}_{\text{i}}\|}$ and $\frac{e^{-jk\|\vec{x}_{\text{s}} - \vec{x}_{\text{f}}\|}}{\|\vec{x}_{\text{s}} - \vec{x}_{\text{f}}\|}$, where $\|\cdot\|$ denotes the Euclidean norm. The free-field translation operator from $\vec{x}_{\text{i}}$ to $\vec{x}_{\text{f}}$ is defined as a ratio of spherical-wave observations:

$$\mathcal{T}_{\text{SW}}(\vec{x}_{\text{i}}, \vec{x}_{\text{f}}) = \frac{\|\vec{x}_{\text{s}} - \vec{x}_{\text{i}}\|}{\|\vec{x}_{\text{s}} - \vec{x}_{\text{f}}\|} e^{jk(\|\vec{x}_{\text{s}} - \vec{x}_{\text{i}}\| - \|\vec{x}_{\text{s}} - \vec{x}_{\text{f}}\|)}. \tag{7}$$

The inverse translation operator (from $\vec{x}_{\text{f}}$ to $\vec{x}_{\text{i}}$) is the reciprocal of (7):

$$\mathcal{T}_{\text{SW}}^{-1} = \frac{1}{\mathcal{T}_{\text{SW}}}. \tag{8}$$

B. Special Functions

1) Associated Legendre Polynomial: The non-normalized associated Legendre polynomial is defined by

$$P_n^m(x) = (-1)^m (1 - x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_n(x), \tag{9}$$

where $P_n(x)$ is the Legendre polynomial,

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n, \tag{10}$$

which in turn obeys the following recursion formula:

$$nP_n(x) = (2n - 1)P_{n-1}(x) - (n - 1)P_{n-2}(x). \tag{11}$$

2) Spherical Harmonics: The angular par of the solution to the acoustic wave equation in spherical coordinates is defined by the spherical harmonic functions. The spherical harmonic function, of order n and degree m, can be defined as complex-valued

$$Y_n^m(\Omega) = N_{nm} P_n^{|m|}(\sin\phi) \exp(jm\theta), \tag{12}$$

or real-valued

$$Y_n^m(\Omega) = N_{nm} P_n^{|m|}(\sin\phi) \times \begin{cases} 1, & m = 0, \\ \sqrt{2}\cos(m\theta), & m > 0, \\ \sqrt{2}\sin(|m|\theta), & m < 0. \end{cases} \tag{13}$$

In (12) and (13), N_{nm} is a normalization factor and P_n^m is the non-normalized associated Legendre polynomial.

The normalization factor is

$$N_{nm} = (-1)^{|m|} \sqrt{\frac{2n + 1}{4\pi} \frac{(n - |m|)!}{(n + |m|)!}}. \tag{14}$$

The normalization factor is chosen in such a way that

$$\int_{\Omega \in \mathbb{S}^2} |Y_n^m(\Omega)|^2 d\Omega = 1, \tag{15}$$

where the integral covers the entire surface of the unit sphere $\mathbb{S}^2$

$$\int_{\Omega \in \mathbb{S}^2} d\Omega = \int_{\theta=-\pi}^{\pi} \int_{\phi=-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(\phi) d\phi d\theta. \tag{16}$$

The following addition theorem relates spherical harmonics and Legendre polynomials:

$$\sum_{m=-n}^{n} Y_n^m(\Omega_1) \overline{Y_n^m}(\Omega_2) = \frac{2n + 1}{4\pi} P_n(\cos\Theta_{12}). \tag{17}$$

Here, Θ_{12} denotes the angle between the directions $\Omega_1 = (\theta_1, \phi_1)$ and $\Omega_2 = (\theta_2, \phi_2)$.

Fig. 5. Interpolation along direction using the direct and inverse spherical Fourier transforms $\mathcal{S}$ and $\mathcal{S}^{-1}$.
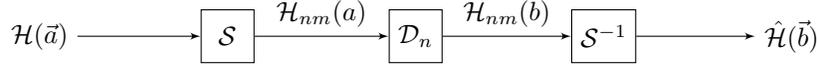

Fig. 6. Extrapolation along distance using the distance-varying filter $\mathcal{D}_n$.

Fig. 7. Regularized extrapolation along distance using the window $\mathcal{W}_n$.

Fig. 8. Ear-centered extrapolation along distance using the direct and inverse translators $\mathcal{T}$ and $\mathcal{T}^{-1}$.

The spherical harmonics compose an orthonormal and complete set on $\mathbb{S}^2$:

$$\text{Orthonormality:} \int_{\Omega \in \mathbb{S}^2} Y_n^m(\Omega)\overline{Y_{n'}^{m'}}(\Omega)d\Omega = \delta_{nm,n'm'}, \quad (18)$$

$$\text{Completeness:} \sum_{n=0}^{\infty}\sum_{m=-n}^{n} Y_n^m(\Omega)\overline{Y_n^m}(\Omega') = \delta(\Omega-\Omega'), \quad (19)$$

In (18), δ is the two-dimensional Kronecker delta, whereas in (19), the two-dimensional Dirac delta.

3) Spherical Hankel Function: The radial part of the solution to the acoustic wave equation in spherical coordinates, for standing waves, is defined by

$$\mathfrak{R}(r) = R_1\mathcal{J}_n(kr) + R_2\mathcal{N}_n(kr), \quad (20)$$

or alternatively, for traveling waves, by

$$\mathfrak{R}(r) = R_3 h_n^{(1)}(kr) + R_4 h_n^{(2)}(kr). \quad (21)$$

In (20), $\mathcal{J}_n$ and $\mathcal{N}_n$ respectively denote the n-th order spherical Bessel functions of the first and second kind. In (21), $h_n^{(1)}$ and $h_n^{(2)}$ respectively denote the n-th order spherical Hankel functions of the first and second kind. In (20) and (21), R_1, R_2, R_3 and R_4 are scalar coefficients.

The spherical Bessel functions in (20) are particularly defined by their series expansions:

$$\mathcal{J}_n(z) = 2^n z^n \sum_{k=0}^{\infty} \frac{(-1)^k(k+n)!}{k!(2k+2n+1)!}z^{2k}, \quad (22)$$

$$\mathcal{N}_n(z) = \frac{(-1)^{n+1}}{2^n z^{n+1}} \sum_{k=0}^{\infty} \frac{(-1)^k(k-n)!}{k!(2k-2n)!}z^{2k}. \quad (23)$$

The spherical Hankel functions in (21) are defined in terms of (22) and (23) by

$$h_n^{(1)}(z) = \mathcal{J}_n(z) + j\mathcal{N}_n(z), \quad (24)$$

$$h_n^{(2)}(z) = \mathcal{J}_n(z) - j\mathcal{N}_n(z). \quad (25)$$

Let $f_n(z)$ denote any of $\mathcal{J}_n(z)$, $\mathcal{N}_n(z)$, $h_n^{(1)}(z)$, or $h_n^{(2)}(z)$. The recurrence relations are as follows:

$$\frac{2n+1}{z}f_n(z) = f_{n+1}(z) + f_{n-1}(z), \quad (26)$$

$$\frac{d}{dz}f_n(z) = f_{n-1}(z) - \frac{n+1}{z}f_n(z). \quad (27)$$

C. Transform Functions

The angular part of a finite-energy acoustic pressure field, distribution, or transfer function $\mathcal{H}$ can be expanded in terms of spherical harmonics,

$$\mathcal{H}(\Omega) = \mathcal{S}^{-1}\{\mathcal{H}_{nm}\} = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \mathcal{H}_{nm}Y_n^m(\Omega), \quad (28)$$

where the expansion coefficients are defined by

$$\mathcal{H}_{nm} = \mathcal{S}\{\mathcal{H}\} = \int_{\Omega \in \mathbb{S}^2} \mathcal{H}(\Omega)\overline{Y_n^m}(\Omega)d\Omega. \quad (29)$$

In the literature [3], [24]–[26], the operator $\mathcal{S}$ in (29) is called the spherical Fourier transform (SFT), whereas $\mathcal{S}^{-1}$ in (28), it is called the inverse spherical Fourier transform (ISFT). Angle interpolation consists on applying $\mathcal{S}$ followed by $\mathcal{S}^{-1}$ as shown in Fig. 5.

In practice, L points distributed on a spherical surface are required to numerically evaluate the integral in (29), which leads to a truncated-sum version of (28) up to a maximum order N. Regular samplings on the sphere satisfy

$$(N+1)^2 \approx L. \quad (30)$$

The discrete SFT in matrix form is formulated as

$$\mathcal{H}_{\mathbf{nm}} = \mathbf{Y}^+\mathcal{H}. \quad (31)$$

The symbol $^+$ denotes Tikhonov regularized pseudoinverse [27]. The matrix $\mathbf{Y} = [Y_q(\Omega_\ell)]$ is of size $L \times Q$, where $q =$

$1, 2, \ldots, Q$ and $\ell = 1, 2, \ldots, L$; its entries are the spherical harmonics $Y_q(\Omega_\ell) = Y_n^m(\Omega_\ell)$ in (12) or (13), where

$$q = n^2 + n + m. \tag{32}$$

The vector of acoustic data $\mathcal{H}$ is of size $L \times 1$, whilst its SFT is the vector $\mathcal{H}_{\mathbf{nm}}$ of size $Q \times 1$. Using the same notation, the discrete ISFT is formulated as

$$\mathcal{H} = \mathbf{Y}\mathcal{H}_{\mathbf{nm}}. \tag{33}$$

For details of additional circular and semicircular transforms, we refer the reader to [1].

D. Transform-Domain Functions

Distance extrapolation synthesizes datapoints at arbitrary radial distances from a sparse dataset at a single radial distance. A wave-based extrapolation method uses the spherical Fourier transforms in (28) and (29) together with the DVFs.

Let $\mathcal{H}(\vec{a})$ denote sound pressure measured on a spherical distribution of points $\vec{a} = (a, \Omega_a)$. Let $\hat{\mathcal{H}}(\vec{b})$ denote sound pressure synthesized at an arbitrary point $\vec{b} = (b, \Omega_b)$. Figure 6 describes the synthesis of $\hat{\mathcal{H}}(\vec{b})$ from $\mathcal{H}(\vec{a})$ by means of $\mathcal{D}_n$, which denotes a DVF in the spherical transform domain.

The DVF from radius a to radius b is defined by

$$\mathcal{D}_n(a, b) = \frac{h_n^{(i)}(kb)}{h_n^{(i)}(ka)}. \tag{34}$$

Here, the kind i of spherical Hankel function (24) or (25) is decided according to

$$i = \begin{cases} 1, & b < a, \\ 2, & b \geq a. \end{cases} \tag{35}$$

The DVF in (34) results from an inverse problem under ideal conditions and yields excessive gains for certain frequencies and distances. To deal with this issue, a regularization window $\mathcal{W}_n$ is applied according to Fig. 7.

The regularization window is formulated as follows:

$$\mathcal{W}_n = \frac{1}{1 + \left(\frac{b}{a}\right)^2 |\mathcal{D}_n|^2}. \tag{36}$$

The conversion of microphone array signals into loudspeaker array signals is another essential process that requires BMFs in the spherical transform domain. BMFs adapt the physical boundary conditions used during recording to those required for reproduction by relying on a theoretical framework provided by the Kirchhoff-Helmholtz integral equation (KHIE). Computationally, array signal conversion is performed in a transform domain where sound fields are represented in terms of spherical harmonic functions. For details of BMFs and their applications, we refer the reader to [3].

IV. USE CASES

This section presents two representative use cases: diffuse-field equalization and HRTF distance extrapolation.
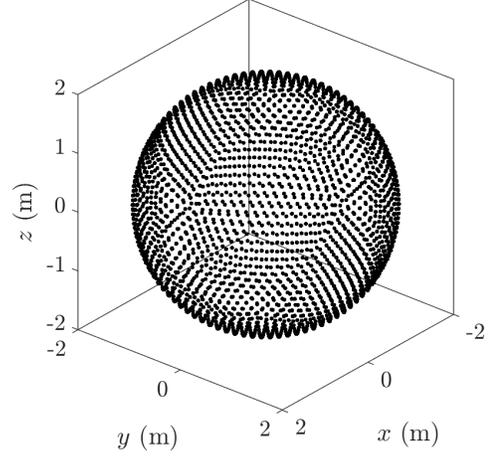

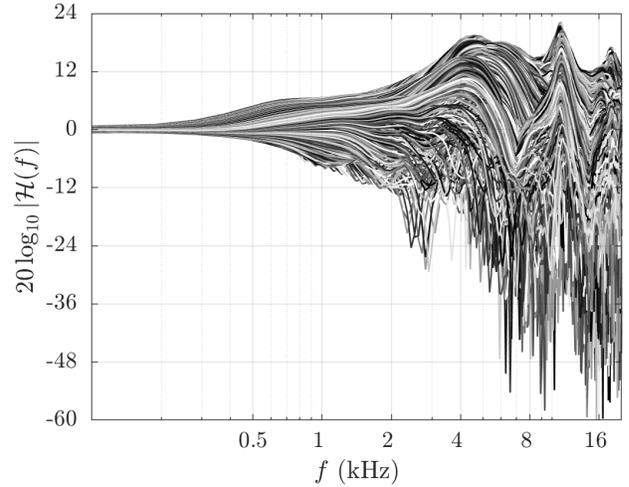Fig. 9. Sound source positions used to obtain the HRTF.
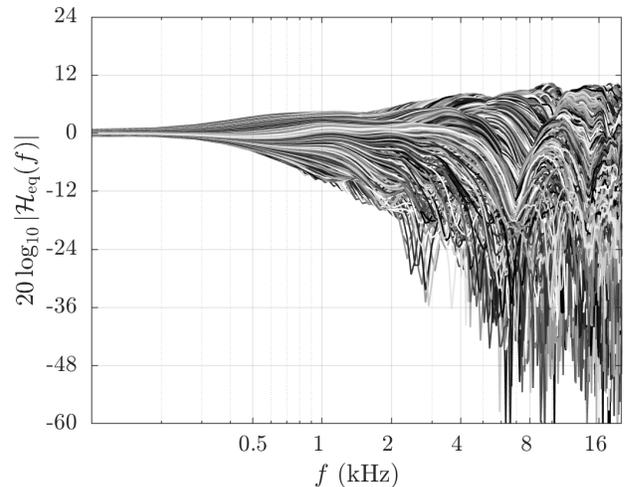

Fig. 10. HRTF before equalization.

Fig. 11. HRTF with equalization.

A. Diffuse-Field Equalization of HRTF

Diffuse-field equalization is a signal processing technique that is used to normalize the frequency response of head-related transfer functions (HRTFs) so that, on average across all directions, the response mimics a flat spectrum. This enhances the perceptual neutrality of binaural rendering, making the resulting binaural audio more natural and spatially consistent.

In SALM, this functionality is demonstrated in the script `exampleDiffuseFieldFilter.m`. The process involves the following stages:

- Import a dense spherical HRTF dataset in the spatially-oriented format for acoustics (SOFA) [28] and preprocess the data.
- Transform HRIRs to the frequency domain and compute quadrature weights for integration over the sphere.
- Design minimum- or zero-phase diffuse-field equalizers using the `diffuseFieldFilter` function and apply them to HRTFs.
- Render and play a binaural audio example with and without equalization, and generate comparative figures for visualization.

Figure 9 shows the spatial distribution of the sound sources used in the measurement of HRTFs. The sources are arranged on a dense icosahedral grid surrounding the listener, which ensures almost uniform angular sampling necessary for accurate spatial averaging and filter design.

Figure 10 displays the magnitude spectra of the HRTFs before applying any equalization. The frequency responses exhibit significant variation across directions, reflecting the natural spectral shaping effects caused by the listener's head, torso, and pinnae.

Figure 11 illustrates the same set of HRTFs after diffuse-field equalization. The equalization process aims to flatten the spatially averaged response, effectively removing coloration introduced by the anatomical features and resulting in a more neutral perceptual baseline for spatial audio reproduction.

B. Distance Extrapolation of HRTF

This use case demonstrates how the SALM library enables HRTF distance extrapolation using spherical DVFs and ear-centering. Figure 8 shows the processing pipeline for distance extrapolation. Figure 12 shows the underlying geometry. By virtue of acoustic reciprocity [29], HRTF datasets can also be regarded as sound pressure field points around the head due to a point source at each of the eardrums.
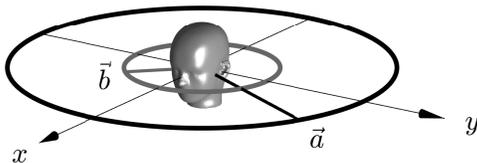

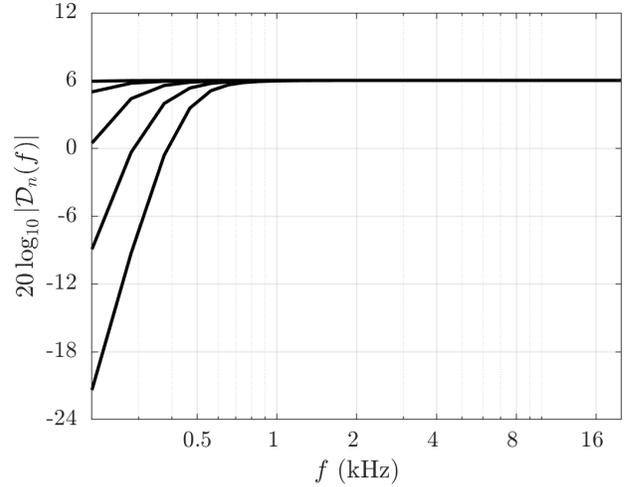Fig. 12. Geometry for distance extrapolation of HRTFs.
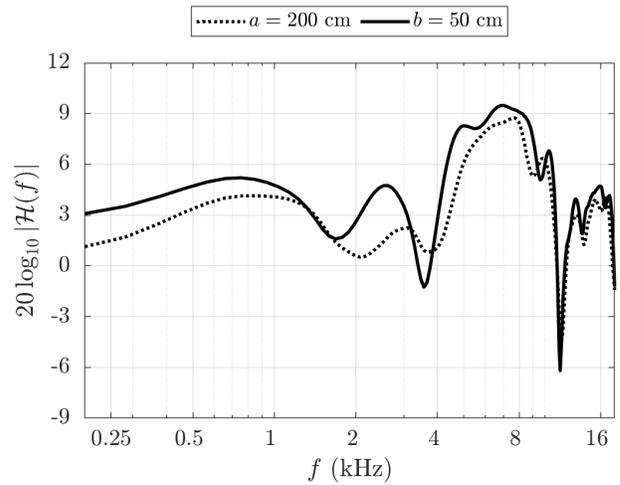

Fig. 13. DVF from $a = 200$ cm to $b = 50$ cm.

Fig. 14. Left-ear HRTFs for a source at $\theta = 90°$ and $\phi = 0°$.

In SALM, this functionality is demonstrated in the script `exampleDistanceVaryingFilterSpherical.m`. The process involves the stages below:

- Reading a dense spherical dataset of head-related impulse responses (HRIRs) from a SOFA file.
- Applying diffuse-field equalization to the dense HRIR dataset.
- Selecting a sparse spherical subset of HRIRs from the dense set.
- Translating the reference from the center of the head to the position of the ear (direct ear-centering) using the `freeFieldTranslationOperator` function.
- Performing distance extrapolation on the sparse spherical set from an original distance a to a new distance b using the `dvf` function.
- Translating the reference back from the position of the ear to the center of the head (inverse ear-centering) using

the `freeFieldTranslationOperator` function.

The dense HRTF measurements were sampled around the head and torso of a mannequin at multiple radial distances. For simplicity, only one sphere of sources located at $a = 200$ cm from the center of the head is selected. A sparse subset of HRTFs was then selected from the dense dataset, based on a reduced-resolution icosahedral grid. This simulates a typical measurement setup with a limited number of directions.

As illustrated in Figures 13 and 14, DVFs allow HRTF extrapolation using transform-domain processing. Figure 13 shows the magnitude of a DVF, which acts as a radial propagator stabilized by a regularization window. Figure 14 shows an initial far-field HRTF at a radial distance $a = 200$ cm and an extrapolated near-field HRTF at a radial distance $b = 50$ cm. Both HRTFs are for the left ear and for a source at $\theta = 90°$ and $\phi = 0°$. These results demonstrate how DVFs, combined with spherical Fourier analysis, can reliably extend measured datasets to arbitrary listener distances.

V. Future Work

The future development of SALM will focus on three directions. First, performance benchmarking and scalability analysis will be performed to assess computational requirements for large head-related transfer functions and room impulse responses datasets compatible with the spatial-oriented format for acoustics (SOFA). Second, new algorithmic extensions will be introduced, including circular DVFs. Third, portability will be improved through the development of a Python version to increase accessibility and integration with modern workflows.

VI. Conclusion

This paper has introduced SALM, a MATLAB library developed for the analysis and processing of spatial audio using spherical transform techniques. The library offers a modular structure composed of spatial-domain functions, analytical tools, transform operators, and filtering methods specifically designed for the manipulation of acoustic transfer functions. Two principal use cases have been demonstrated: diffuse-field equalization and distance extrapolation of HRTFs using spherical distance-varying filters. These applications highlight the library's ability to bridge physical modeling with transform-domain processing for accurate and perceptually relevant sound field manipulation. By supporting standardized data formats and providing reproducible workflows, SALM facilitates research and development in spatial acoustics, binaural rendering, and virtual auditory environments. Future developments might target the integration of perceptual assessment frameworks.

Acknowledgment

References

[1] C. D. Salvador, S. Sakamoto, J. Treviño, and Y. Suzuki, "Distance-varying filters to synthesize head-related transfer functions in the horizontal plane from circular boundary values," *Acoust. Sci. Technol.*, vol. 38, no. 1, pp. 1–13, Jan. 2017.

[2] S. Sakamoto, C. D. Salvador, J. Treviño, and Y. Suzuki, "Binaural synthesis using a spherical microphone array based on the solution to an inverse problem," *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 259, no. 4, pp. 5983–5988, 2019.

[3] C. D. Salvador, S. Sakamoto, J. Treviño, and Y. Suzuki, "Boundary matching filters for spherical microphone and loudspeaker arrays," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 3, pp. 461–474, Mar. 2018.

[4] S. Sakamoto, F. Monasterolo, C. D. Salvador, Z. Cui, and Y. Suzuki, "Effects of target speech distance on auditory spatial attention in noisy environments," in *Proc. ICA 2019 and EAA Euroregio*, Aachen, Germany, Sep. 2019, pp. 2177–2181.

[5] C. D. Salvador, S. Sakamoto, J. Treviño, and Y. Suzuki, "Design theory for binaural synthesis: Combining microphone array recordings and head-related transfer function datasets," *Acoust. Sci. Technol.*, vol. 38, no. 2, pp. 51–62, Mar. 2017.

[6] F. Monasterolo, S. Sakamoto, C. D. Salvador, Z. Cui, and Y. Suzuki, "The effect of target speech distance on reaction time under multi-talker environment," *IEICE Tech. Rep.*, vol. 118, no. 313, pp. 83–88, Nov. 2018.

[7] C. D. Salvador, S. Sakamoto, J. Treviño, and Y. Suzuki, "Dataset of near-distance head-related transfer functions calculated using the boundary element method," in *Proc. Audio Eng. Soc. Int. Conf. Spatial Reproduction —Aesthetics and Science—*. Tokyo, Japan: Audio Engineering Society, Aug. 2018. [Online]. Available: https://cesardsalvador.github.io/download.html

[8] A. Urviola, S. Sakamoto, and C. D. Salvador, "Ear centering for accurate synthesis of near-field head-related transfer functions," *Appl. Sci.*, vol. 12, no. 16, 2022.

[9] C. D. Salvador, A. Urviola, and S. Sakamoto, "Ear centering in the spatial and transform domains for near-field head-related transfer functions," in *Proceedings of the International Congress on Acoustics (ICA)*, Gyeongju, South Korea, Oct. 2022.

[10] A. Urviola, S. Sakamoto, and C. D. Salvador, "Ear centering for near-distance head-related transfer functions," in *Proc. Int. Conf. Immersive and 3D Audio (I3DA): from Architecture to Automotive*. Bologna, Italy: IEEE, Sep. 2021.

[11] A. Campos, S. Sakamoto, and C. D. Salvador, "Directional early-to-late energy ratios to quantify intelligibility: A case study in a large auditorium," in *Proc. Int. Conf. Immersive and 3D Audio (I3DA): from Architecture to Automotive*. Bologna, Italy: IEEE, Sep. 2021.

[12] J. Alarcón, J. Solis-Lastra, and C. D. Salvador, "Regularized spherical Fourier transform for room impulse response interpolation," in *Proc. IEEE INTERCON*. Lima, Peru: IEEE, Aug. 2021, pp. 1–4.

[13] B. Bernschütz, C. Pörschmann, S. Spors, and S. Weinzierl, "SOFiA Sound Field Analysis Toolbox," in *Proceedings of the ICSA International Conference on Spatial Audio*, Detmold, Germany, Nov. 2011. [Online]. Available: https://code.google.com/archive/p/sofia-toolbox/

[14] H. Wierstorf and S. Spors, "Sound field synthesis toolbox," in *Proceedings of the 132nd Convention of the Audio Engineering Society (AES)*, vol. 4. Budapest, Hungary: Audio Engineering Society (AES), 2012. [Online]. Available: https://sfs.readthedocs.io/

[15] F. Brinkmann and S. Weinzierl, "Aktools—an open software toolbox for signal acquisition, processing, and inspection in acoustics," in *Proceedings of the 142nd Audio Engineering Society (AES) Convention*. Berlin, Germany: Audio Engineering Society, May 2017. [Online]. Available: https://depositonce.tu-berlin.de/items/ce6f29d9-7dc3-45e2-bf77-8af929483fb3

[16] M. Berzborn, R. Bomhardt, J. Klein, J.-G. Richter, and M. Vorländer, "The ITA-Toolbox: An open source MATLAB toolbox for acoustic measurements and signal processing," in *Proceedings of the 43th Annual German Congress on Acoustics (DAGA)*, vol. 6, Kiel, Germany, 2017. [Online]. Available: https://www.ita-toolbox.org/index.php

[17] J. C. Middlebrooks and D. M. Green, "Directional dependence of interaural envelope delays," *J. Acoust. Soc. Am.*, vol. 87, no. 5, pp. 2149–2162, 1990.

[18] J. C. Middlebrooks, "Individual differences in external-ear transfer functions reduced by scaling in frequency," *J. Acoust. Soc. Am.*, vol. 106, no. 3, pp. 1480–1492, 1999.

[19] I. Ben Hagai, M. Pollow, M. Vorländer, and B. Rafaely, "Acoustic centering of sources measured by surrounding spherical microphone arrays," *J. Acoust. Soc. Am.*, vol. 130, no. 4, pp. 2003–2015, 2011.

[20] N. R. Shabtai and M. Vorländer, "Acoustic centering of sources with high-order radiation patterns," *J. Acoust. Soc. Am.*, vol. 137, no. 4, pp. 1947–1961, 2015.

[21] Y. Wang and K. Chen, "Translations of spherical harmonics expansion coefficients for a sound field using plane wave expansions," *J. Acoust. Soc. Am.*, vol. 143, no. 6, pp. 3474–3478, 2018.

[22] Z. Ben-Hur, D. L. Alon, R. Mehra, and B. Rafaely, "Efficient representation and sparse sampling of head-related transfer functions using phase-correction based on ear alignment," *IEEE Trans. Audio, Speech, Language Process.*, pp. 2249–2262, 2019.

[23] M. Kentgens and P. Jax, "Translation of a higher-order ambisonics sound scene by space warping," in *Proc. Audio Eng. Soc. Int. Conf. Audio for Virtual and Augmented Reality*, Aug. 2020.

[24] J. R. Driscoll and D. M. Heally, "Computing Fourier transforms and convolutions on the 2-sphere," *Adv. Appl. Math.*, vol. 15, pp. 202–250, 1994.

[25] D. Healy, Jr., D. Rockmore, P. Kostelec, and S. Moore, "FFTs for the 2-sphere-Improvements and variations," *J. Fourier Anal. Appl.*, vol. 9, no. 4, pp. 341–385, Jul. 2003.

[26] F. J. Simons, F. A. Dahlen, and M. A. Wieczorek, "Spatiospectral concentration on a sphere," *SIAM Rev.*, vol. 48, no. 3, pp. 504–536, 2006.

[27] A. Neumaier, "Solving ill-conditioned and singular linear systems: A tutorial on regularization," *SIAM Review*, vol. 40, no. 3, pp. 636–666, 1998.

[28] P. Majdak, Y. Iwaya, T. Carpentier, R. Nicol, M. Parmentier, A. Roginska, Y. Suzuki, K. Watanabe, H. Wierstorf, H. Ziegelwanger, and M. Noisternig, "Spatially oriented format for acoustics: a data exchange format representing head-related transfer functions," in *AES 134 Convention*, May 2013. [Online]. Available: https://www.sofaconventions.org/

[29] D. N. Zotkin, R. Duraiswami, E. Grassi, and N. A. Gumerov, "Fast head-related transfer function measurement via reciprocity," *J. Acoust. Soc. Am.*, vol. 120, no. 4, pp. 2202–2215, Oct. 2006.