[T-5] Sound Field Estimation: Recent Advances and Applications – Sections 1 & 2 –

Shoichi Koyama[1], Natsuki Ueno[2]

[1] The University of Tokyo

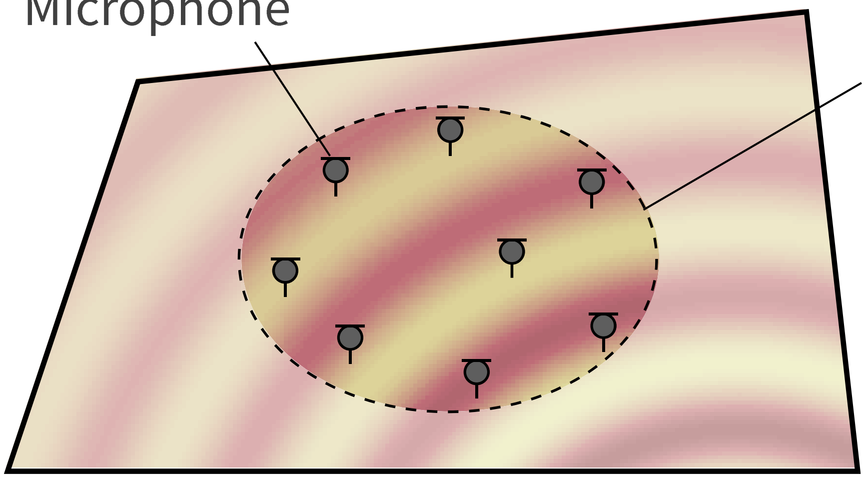
[2] Tokyo Metropolitan University

22-27 MAY 2022

ICASSP 2022

SINGAPORE

What is sound field estimation?



Estimating sound field inside target region using observations of multiple mics

Sound field estimation is one of fundamental problems in acoustic signal processing

Applications of sound field estimation

VR/AR audio

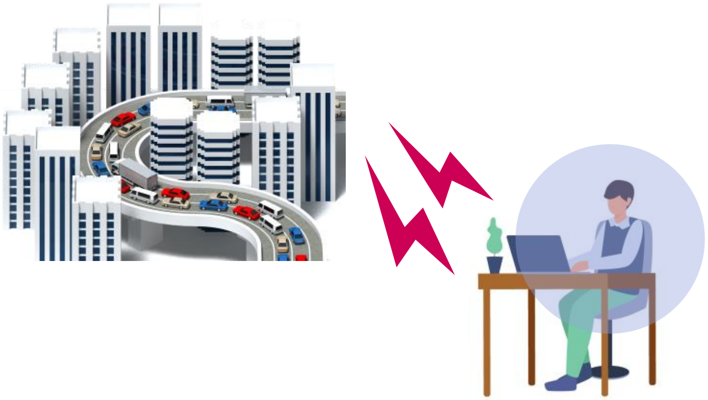
Room acoustic analysis

Signal enhancement

Basic Technologies of Sound Field Estimation

Local-field recording and reproduction

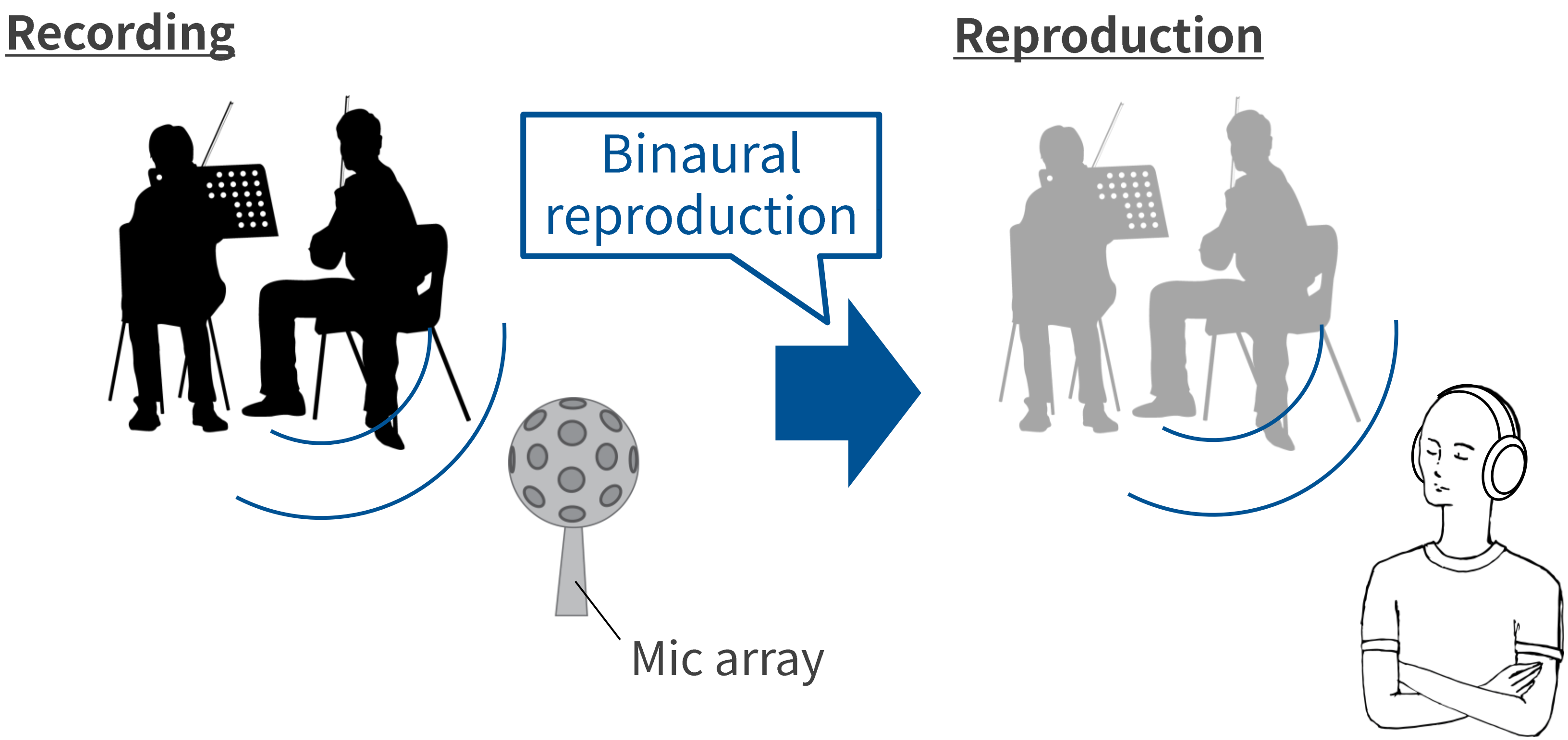
Active noise control



Visualization/auralization

Example: sound field estimation in spatial audio

Sound field estimation technique is necessary for capturing spatial sound



Sound field is estimated by multiple mics and reproduced by headphones

Table of contents

1. Preliminaries on acoustics
   1. Wave equation and Helmholtz equation
   2. Representations of acoustic field
2. Integral-equation-based sound field estimation
   1. Problem formulation
   2. Sound field estimation with spherical array
   3. Methods to avoid forbidden frequency problem
3. Least-squares-based sound field estimation
   1. Basic framework
   2. Infinite-dimensional extension
   3. Other extensions and related works
4. Applications
   1. Spatial audio reproduction by headphones
   2. Spatial audio reproduction by loudspeakers
   3. Spatial active noise control

Notation

- Scalar, vector, and matrix

$$x, \quad \boldsymbol{x}, \quad \boldsymbol{X}$$

- Euler's number, imaginary unit, complex conjugate

$$\mathrm{e}, \quad \mathrm{j}, \quad (\cdot)^*$$

- Transpose, conjugate transpose, and inverse of matrix

$$(\cdot)^{\mathsf{T}}, \quad (\cdot)^{\mathsf{H}}, \quad (\cdot)^{-1}$$

- Inner product and ℓ_p-norm

$$\langle \cdot, \cdot \rangle, \quad \| \cdot \|_p \quad (\| \cdot \| \text{ for Euclidean norm})$$

- Sets of real numbers, complex numbers, and unit sphere in 3D

$$\mathbb{R}, \quad \mathbb{C}, \quad \mathbb{S}_2$$

Table of contents

1. Preliminaries on acoustics
    1. Wave equation and Helmholtz equation
    2. Representations of acoustic field
2. Integral-equation-based sound field estimation
    1. Problem formulation
    2. Sound field estimation with spherical array
    3. Methods to avoid forbidden frequency problem
3. Least-squares-based sound field estimation
    1. Basic framework
    2. Infinite-dimensional extension
    3. Other extensions and related works
4. Applications
    1. Spatial audio reproduction by headphones
    2. Spatial audio reproduction by loudspeakers
    3. Spatial active noise control

Wave equation

- Wave equation is partial differential equation (PDE) governing acoustic wave propagation in medium.

Sound pressure u at position $\boldsymbol{r} \in \mathbb{R}^3$ and time t satisfies the following PDE:

$$\nabla^2 u(\boldsymbol{r}, t) - \frac{1}{c^2} \frac{\partial^2 u(\boldsymbol{r}, t)}{\partial t^2} = 0$$

where c is sound speed, and ∇^2 is Laplacian.

– Same applies to particle velocity vector $\boldsymbol{v}$:

$$\nabla^2 \boldsymbol{v}(\boldsymbol{r}, t) - \frac{1}{c^2} \frac{\partial^2 \boldsymbol{v}(\boldsymbol{r}, t)}{\partial t^2} = 0$$

Helmholtz equation

- In stationary field, sound field is represented in frequency domain.

$$u(\boldsymbol{r}, \omega) = \mathcal{F}_t\left[u(\boldsymbol{r}, t)\right] = \int_{-\infty}^{\infty} u(\boldsymbol{r}, t) \mathrm{e}^{\mathrm{j}\omega t} \mathrm{d}t$$

- Helmholtz equation is derived by Fourier transform of both sides of wave equation

Sound pressure u at position $\boldsymbol{r} \in \mathbb{R}^3$ and angular frequency ω satisfies the following PDE:

$$(\nabla^2 + k^2) u(\boldsymbol{r}, \omega) = 0$$

where $k = \omega / c$ is wave number.

– Hereafter, all the formulations in Sects. 1-3 are in freq domain.

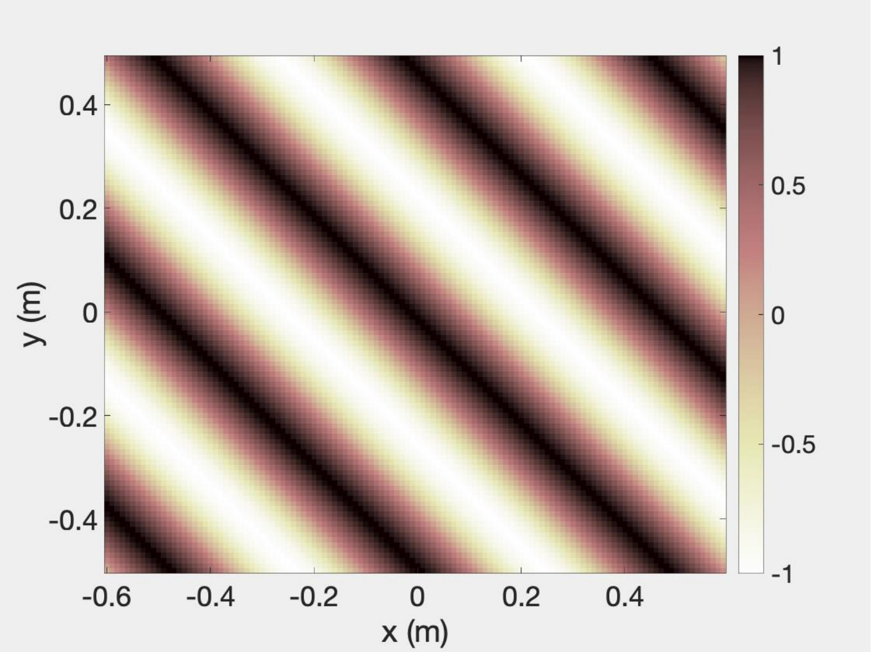
Plane wave

- One of elementary waves is plane wave:

$$u(\boldsymbol{r}) = u_0 \mathrm{e}^{\mathrm{j}(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)}$$

where $\boldsymbol{k} = [k_x, k_y, k_z]^{\mathsf{T}}$ ($\|\boldsymbol{k}\| = k$) is called wave vector.

- Unit vector $\boldsymbol{k}/k$ represents propagation direction of plane wave.

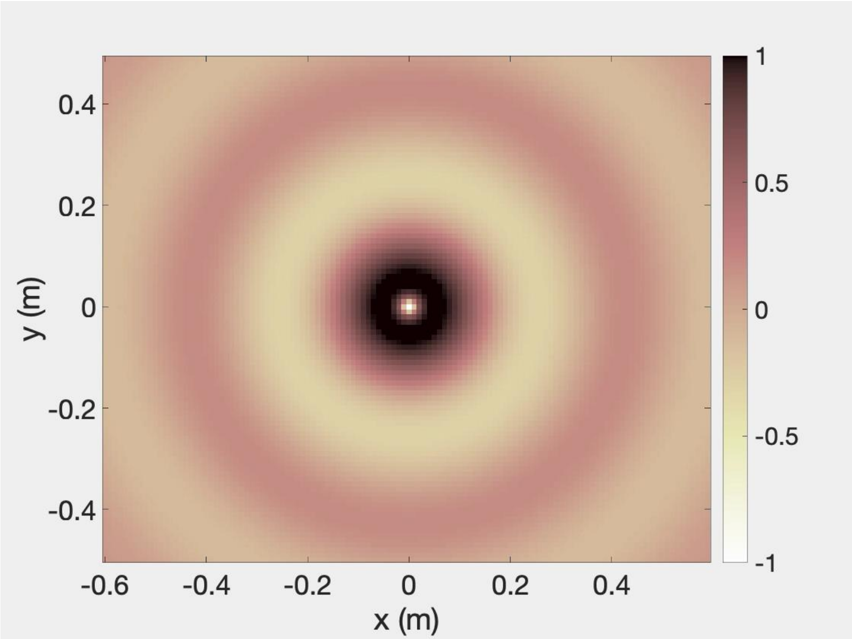


Spherical wave

- Another example of elementary waves is spherical wave.
- Spherical wave propagating from point source / monopole at origin:

$$u(\boldsymbol{r}) = -\mathrm{j}\omega\rho\frac{Q}{4\pi\|\boldsymbol{r}\|}\mathrm{e}^{\mathrm{j}(k\|\boldsymbol{r}\|-\omega t)}$$

where Q is source strength, and ρ is density of medium.



Green's function

- Point source function is equivalent to free-field Green's function $G(\boldsymbol{r}|\boldsymbol{r}')$ that is fundamental solution of the following inhomogeneous Helmholtz equation:

$$\nabla^2 G(\boldsymbol{r}|\boldsymbol{r}') + k^2 G(\boldsymbol{r}|\boldsymbol{r}') = -\delta(\boldsymbol{r} - \boldsymbol{r}')$$

where

$$G(\boldsymbol{r}|\boldsymbol{r}') = \frac{\mathrm{e}^{\mathrm{j}k\|\boldsymbol{r}-\boldsymbol{r}'\|}}{4\pi\|\boldsymbol{r}-\boldsymbol{r}'\|}$$

- Relation between $G(\boldsymbol{r}|\boldsymbol{r}')$ and point source excluding time-dependent term u :

$$u(\boldsymbol{r}) = -\mathrm{j}\rho c k Q G(\boldsymbol{r}|\boldsymbol{r}')$$

Green's function

- Green's function $G(\boldsymbol{r}|\boldsymbol{r}')$ is a function between two positions that describes effects on the field quantity (e.g., pressure) at $\boldsymbol{r}$ from unit excitation at $\boldsymbol{r}'$ (i.e., transfer function).
- Examples of boundary conditions (BCs) imposed in addition to source of Dirac's delta ($\boldsymbol{n}$ is unit normal vector)
  - Neumann BC (sound-hard):

$$\frac{\partial u}{\partial \boldsymbol{n}} = 0$$

  - Dirichlet BC (sound-soft):

$$u = 0$$

  - Robin BC (or impedance condition):

$$\alpha u + \beta \frac{\partial u}{\partial \boldsymbol{n}} = 0$$

    with constants α and β ($Z = -\mathrm{j}k\beta/\alpha$ corresponds to acoustic impedance ratio).

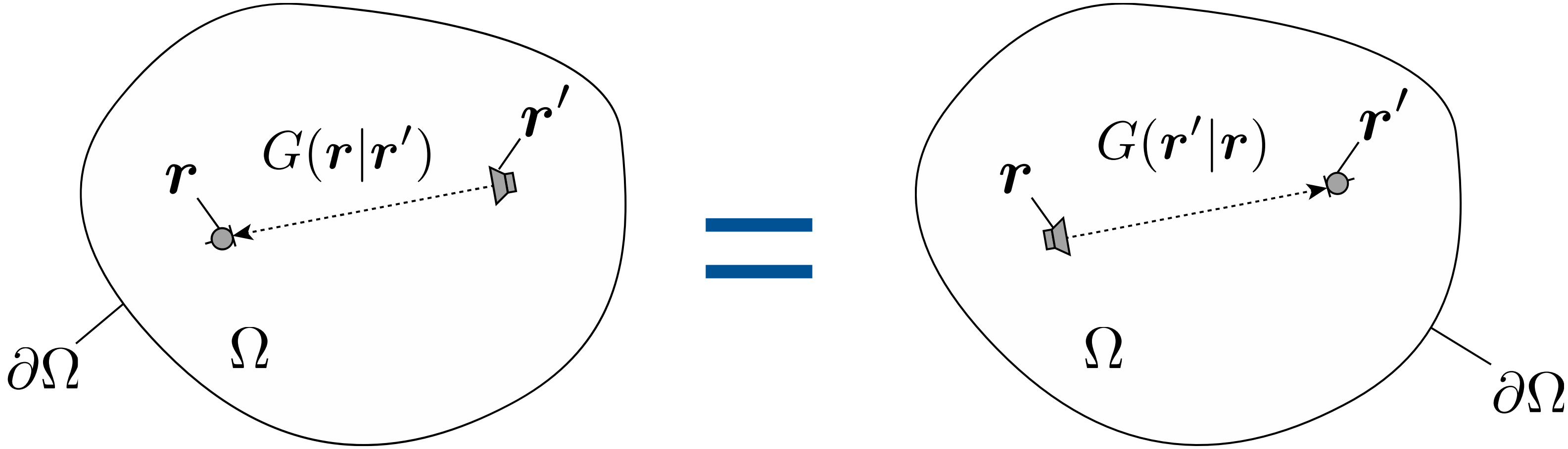
Green's function

- Suppose Green's function satisfying

$$\begin{cases} (\nabla^2 + k^2) G(\boldsymbol{r}|\boldsymbol{r}') = -\delta(\boldsymbol{r} - \boldsymbol{r}') & \text{in } \Omega \\ \alpha G(\boldsymbol{r}|\boldsymbol{r}') + \beta \dfrac{\partial G(\boldsymbol{r}|\boldsymbol{r}')}{\partial \boldsymbol{n}} = 0 & \text{on } \partial\Omega \end{cases}$$

- This Green's function satisfies reciprocity:

$$G(\boldsymbol{r}|\boldsymbol{r}') = G(\boldsymbol{r}'|\boldsymbol{r})$$



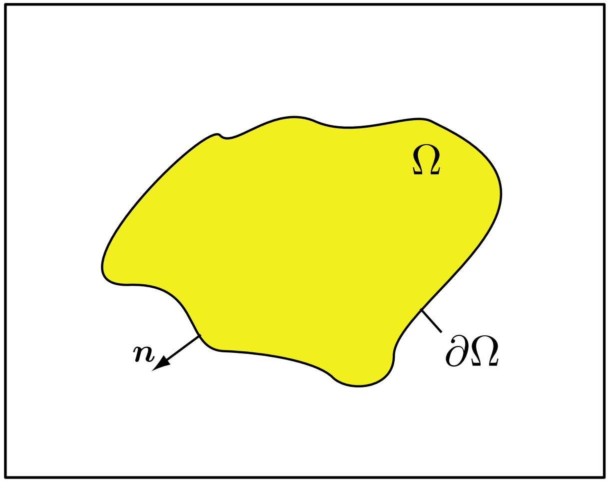
Representations of acoustic field

- Two important acoustic-field representations:
  - Boundary-integral representations
    - Describing sound propagation from boundary surface to its interior/exterior region.
    - Sound field representation without explicit source parameters
  - Wavefunction expansions
    - Sound field is represented by superposition of wavefunctions, i.e., elementary solutions of Helmholtz equation
    - Complete set of wavefunctions fairly approximates any solutions of homogeneous Helmholtz equation

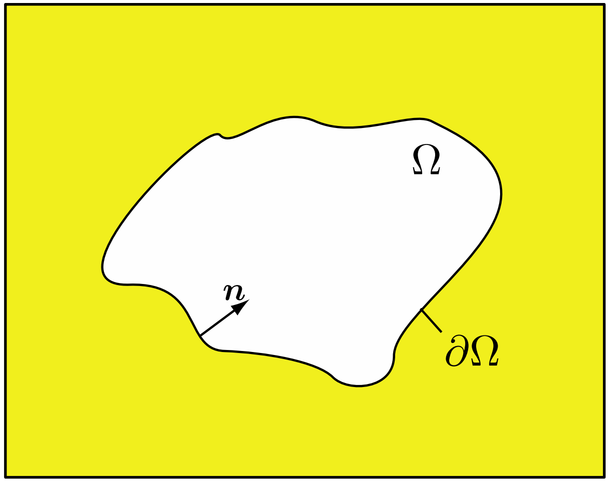
Most of sound field estimation methods are based on these two representations

Representations of acoustic field

- Interior and exterior problems
  - Interior problem
    - Representing interior sound field in Ω from its boundary $\partial\Omega$
    - Sources exist outside Ω, i.e., $\mathbb{R}^3 \backslash \Omega$

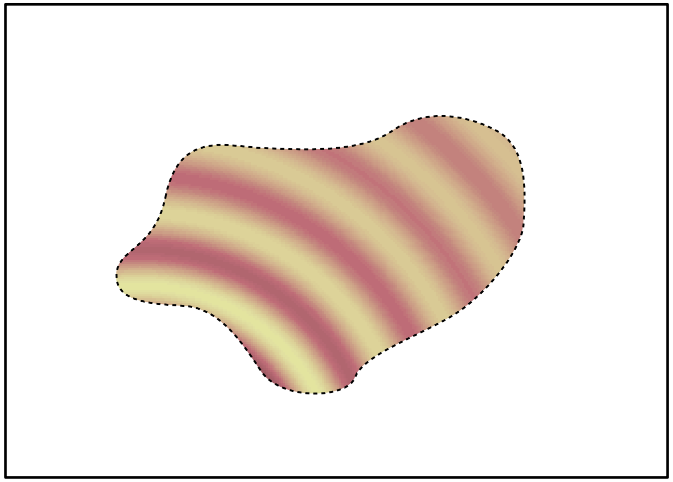


  - Exterior problem
    - Representing exterior sound field of Ω, i.e., $\mathbb{R}^3 \backslash \Omega$, from its boundary $\partial\Omega$
    - Sources exist inside Ω

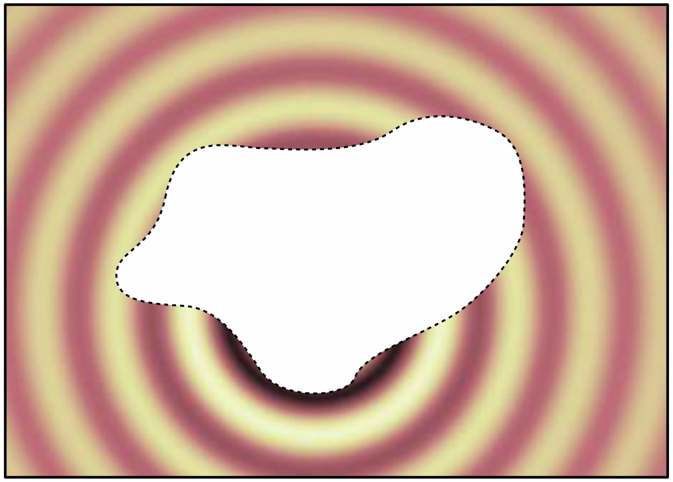


Boundary-integral representation

- Boundary integral equations for Helmholtz equation allow predicting interior/exterior sound field from boundary values
  - Kirchhoff—Helmholtz (KH) integral
  - Single/double layer potential



Interior problem



Exterior problem

Kirchhoff—Helmholtz integral

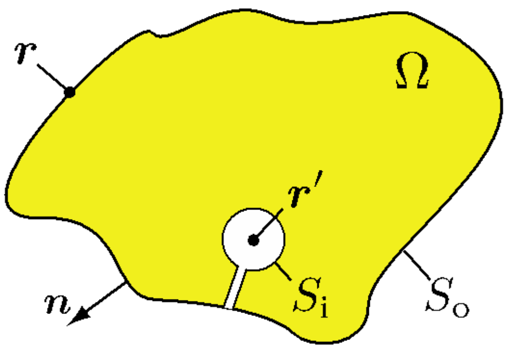
- KH integral is derived by applying Green's theorem to function satisfying Helmholtz equation

Green's (second) identity for bounded and continuous functions in Ω, $\Phi(\boldsymbol{r})$ and $\Psi(\boldsymbol{r})$

$$\int_{\boldsymbol{r}\in\Omega} \left(\Phi\nabla^2\Psi - \Psi\nabla^2\Phi\right) \mathrm{d}\boldsymbol{r} = \int_{\boldsymbol{r}\in\partial\Omega} \left(\Phi\frac{\partial\Psi}{\partial\boldsymbol{n}} - \Psi\frac{\partial\Phi}{\partial\boldsymbol{n}}\right) \mathrm{d}\boldsymbol{r}$$

- For interior problem, singular point is set at $\boldsymbol{r} = \boldsymbol{r}'$; then, Green's theorem is applied.

$$\int_{\boldsymbol{r}'\in S_\mathrm{o}} \left(\Phi\frac{\partial G(\boldsymbol{r}|\boldsymbol{r}')}{\partial\boldsymbol{n}} - G(\boldsymbol{r}|\boldsymbol{r}')\frac{\partial\Phi}{\partial\boldsymbol{n}}\right) \mathrm{d}S_\mathrm{o} + \lim_{\epsilon\to0}\int_{\boldsymbol{r}'\in S_\mathrm{i}} \left(\Phi\frac{\partial G(\boldsymbol{r}|\boldsymbol{r}')}{\partial\boldsymbol{n}} - G(\boldsymbol{r}|\boldsymbol{r}')\frac{\partial\Phi}{\partial\boldsymbol{n}}\right) \mathrm{d}S_\mathrm{i} = 0$$



Kirchhoff—Helmholtz integral

Kirchhoff—Helmholtz integral for interior problem:

Pressure $u(\boldsymbol{r})$ in source-free interior region Ω is represented as

$$u(\boldsymbol{r}) = \int_{\boldsymbol{r}' \in \partial\Omega} \left(G(\boldsymbol{r}|\boldsymbol{r}') \frac{\partial u(\boldsymbol{r}')}{\partial \boldsymbol{n}'} - u(\boldsymbol{r}') \frac{\partial G(\boldsymbol{r}|\boldsymbol{r}')}{\partial \boldsymbol{n}'} \right) \mathrm{d}\boldsymbol{r}' \quad (\boldsymbol{r} \in \Omega)$$

- KH integral for interior problem means that interior sound field can be uniquely determined by pressure and its gradient on the boundary
- Pressure gradient $\partial u(\boldsymbol{r})/\partial \boldsymbol{n}$ is quantity propotional to particle velocity in normal direction

$$\frac{\partial u(\boldsymbol{r})}{\partial \boldsymbol{n}} = \mathrm{j}\omega\rho v_n(\boldsymbol{r})$$

Kirchhoff—Helmholtz integral

- KH integral for exterior problem can be obtained in a similar manner by imposing Sommerfeld radiation condition

$$\lim_{r\to\infty} r\left(\frac{\partial u(\boldsymbol{r})}{\partial r} - \mathrm{j}ku(\boldsymbol{r})\right) = 0$$

Kirchhoff—Helmholtz integral for exterior problem:

Pressure $u(\boldsymbol{r})$ in source-free exterior region $\mathbb{R}^3\backslash\Omega$ is represented as

$$u(\boldsymbol{r}) = \int_{\boldsymbol{r}'\in\partial\Omega}\left(G(\boldsymbol{r}|\boldsymbol{r}')\frac{\partial u(\boldsymbol{r}')}{\partial \boldsymbol{n}'} - u(\boldsymbol{r}')\frac{\partial G(\boldsymbol{r}|\boldsymbol{r}')}{\partial \boldsymbol{n}'}\right)\mathrm{d}\boldsymbol{r}' \qquad (\boldsymbol{r}\in\mathbb{R}^3\backslash\Omega)$$

- KH integral for exterior problem means that exterior sound field can be uniquely determined by pressure and its gradient on the boundary

Single/double layer potential

- Another boundary-integral representation is single/double layer potential that use only the term of free-field Green's function or its normal derivative in the integrand.

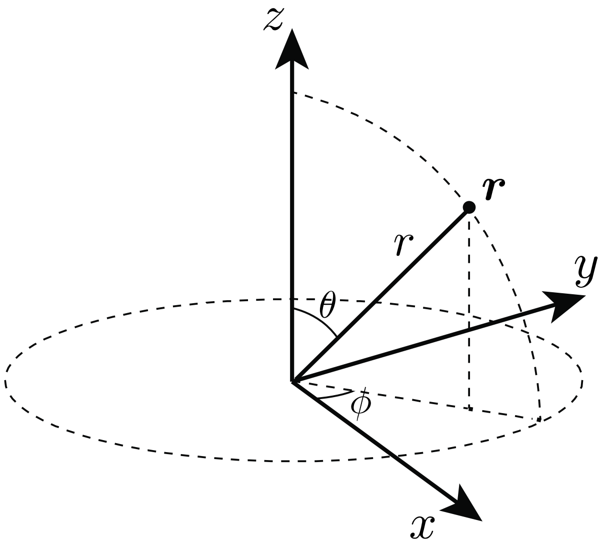
Single layer potential: $$u(\boldsymbol{r}) = \int_{\boldsymbol{r}' \in \partial\Omega} \mu(\boldsymbol{r}') G(\boldsymbol{r}|\boldsymbol{r}') \mathrm{d}\boldsymbol{r}'$$

Double layer potential: $$u(\boldsymbol{r}) = \int_{\boldsymbol{r}' \in \partial\Omega} \mu(\boldsymbol{r}') \frac{\partial G(\boldsymbol{r}|\boldsymbol{r}')}{\partial \boldsymbol{n}} \mathrm{d}\boldsymbol{r}'$$

- Single layer potential is also called simple source formulation or equivalent source method.
- Analytical formulation of $\mu(\boldsymbol{r})$ is not always possible.

Wavefunction expansion

- Representing solutions of (homogeneous) Helmholtz equation by complete set of eigenfunctions
- Two representative wavefunction expansions
  - Plane wave expansion
    - – Equivalent to general solution in Cartesian coordinate
  - Spherical wavefunction expansion
    - – Equivalent to general solution in spherical coordinate



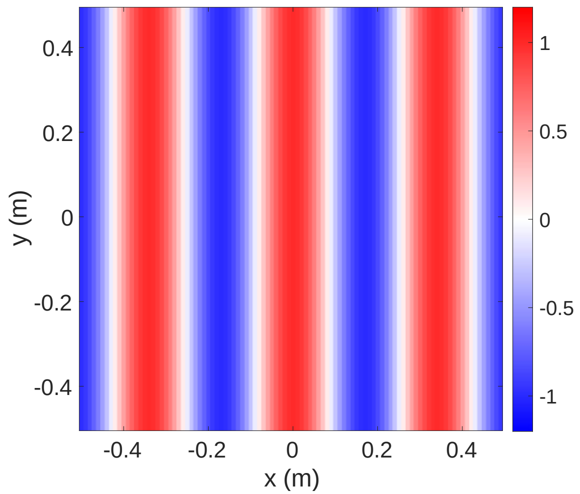
- Plane wave expansion

$$u(\boldsymbol{r}) = \int_{\boldsymbol{x}\in\mathbb{S}_2} \tilde{u}(\boldsymbol{x}) \mathrm{e}^{-\mathrm{j}k\boldsymbol{x}\cdot\boldsymbol{r}} \mathrm{d}\chi$$

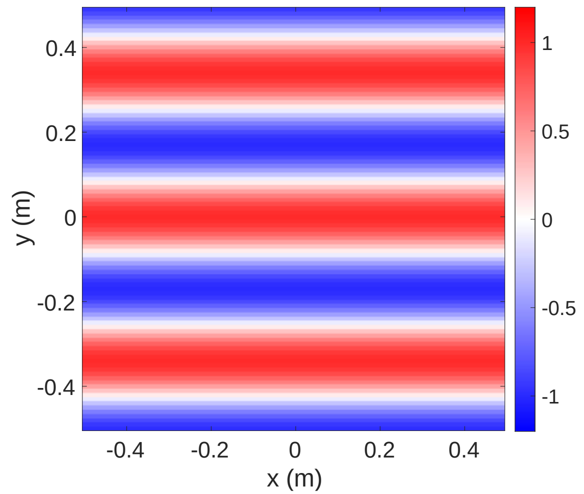
Plane wave function

Expansion coefficient

- $\boldsymbol{x}$: Unit vector of arrival direction $(\boldsymbol{x} := -\boldsymbol{k}/k)$
- $\int_{\boldsymbol{x}\in\mathbb{S}_2} \mathrm{d}\chi$: Integral over unit sphere



$\boldsymbol{x} = [1, 0, 0]^\mathsf{T}$



$\boldsymbol{x} = [0, 1, 0]^\mathsf{T}$

Spherical wavefunction expansion

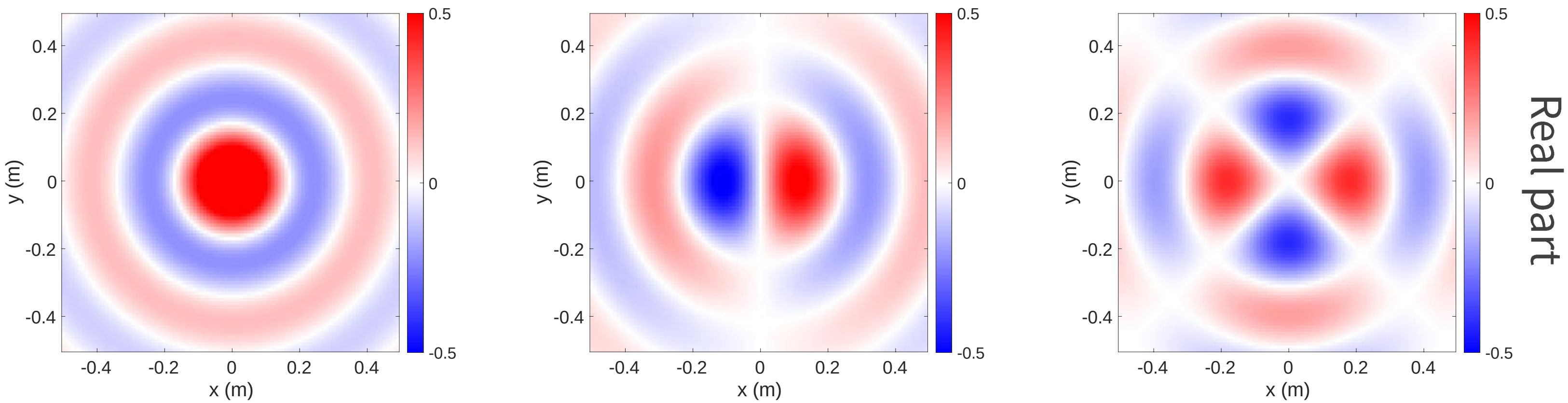
- Spherical wavefunction expansion for interior problem

Spherical wavefunction

$$u(\boldsymbol{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\nu,\mu} \sqrt{4\pi} j_\nu(k\|\boldsymbol{r}\|) Y_{\nu,\mu}(\boldsymbol{r}/\|\boldsymbol{r}\|)$$

Expansion coefficient

  - $j_\nu(\cdot)$: ν th-order spherical Bessel function
  - $Y_{\nu,\mu}(\cdot)$: Spherical harmonic function of order ν and degreee μ



$\nu = 0,\ \mu = 0$ $\nu = 1,\ \mu = 1$ $\nu = 2,\ \mu = 2$

Spherical wavefunction expansion

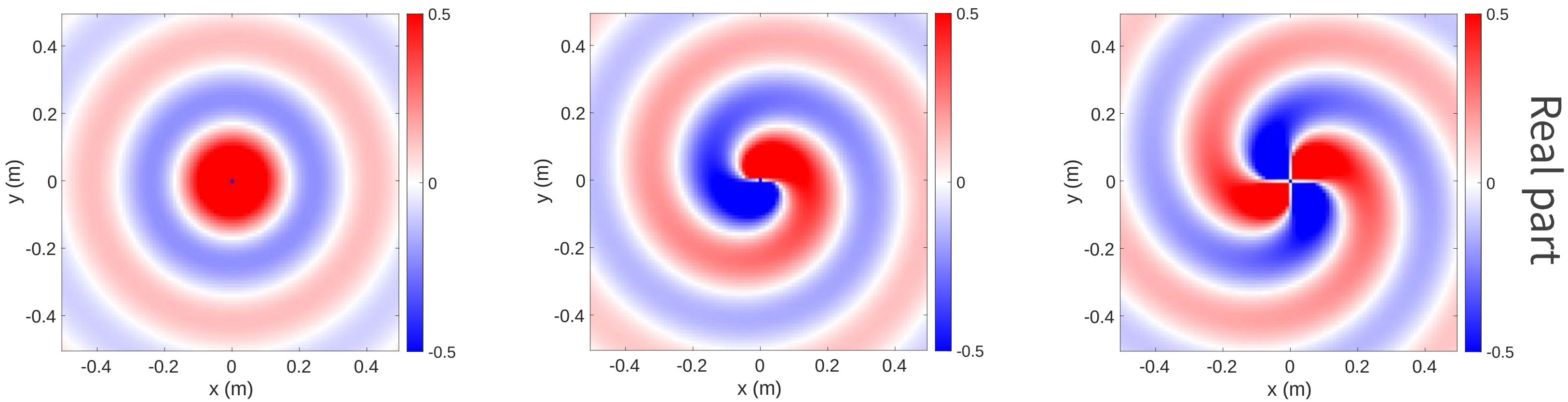
➤ Spherical wavefunction expansion for exterior problem

$$u(\boldsymbol{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\nu,\mu} \sqrt{4\pi} h_\nu(k\|\boldsymbol{r}\|) Y_{\nu,\mu}(\boldsymbol{r}/\|\boldsymbol{r}\|)$$

Spherical wavefunction: $\sqrt{4\pi} h_\nu(k\|\boldsymbol{r}\|) Y_{\nu,\mu}(\boldsymbol{r}/\|\boldsymbol{r}\|)$

Expansion coefficient: $\mathring{u}_{\nu,\mu}$

– $h_\nu(\cdot)$: ν th-order spherical Hankel function of 1st kind
– $Y_{\nu,\mu}(\cdot)$: Spherical harmonic function of order ν and degreee μ



$\nu = 0,\ \mu = 0$ $\nu = 1,\ \mu = 1$ $\nu = 2,\ \mu = 2$

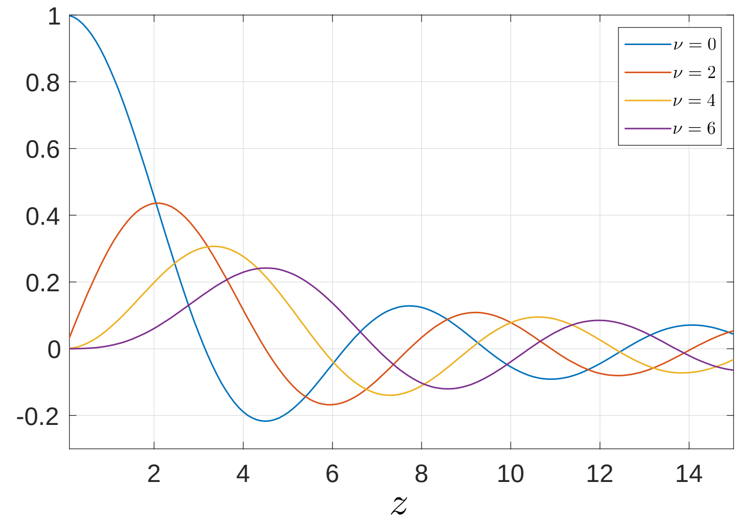
($h_\nu(\cdot)$ has singularity at origin)

Spherical wavefunction expansion

- Spherical Bessel function

$$j_\nu(z) = \sqrt{\frac{\pi}{2z}} J_{\nu+1/2}(z)$$

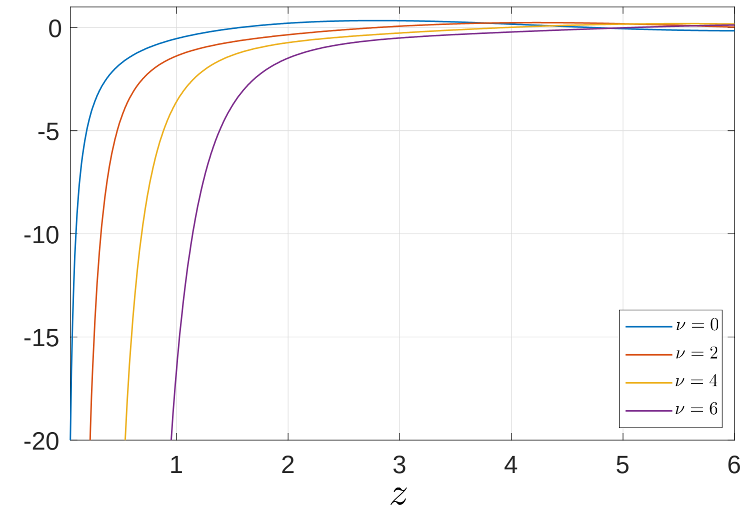
Bessel function



- Spherical Neumann function

$$n_\nu(z) = \sqrt{\frac{\pi}{2z}} N_{\nu+1/2}(z)$$

Neumann function



- Spherical Hankel function of 1st kind

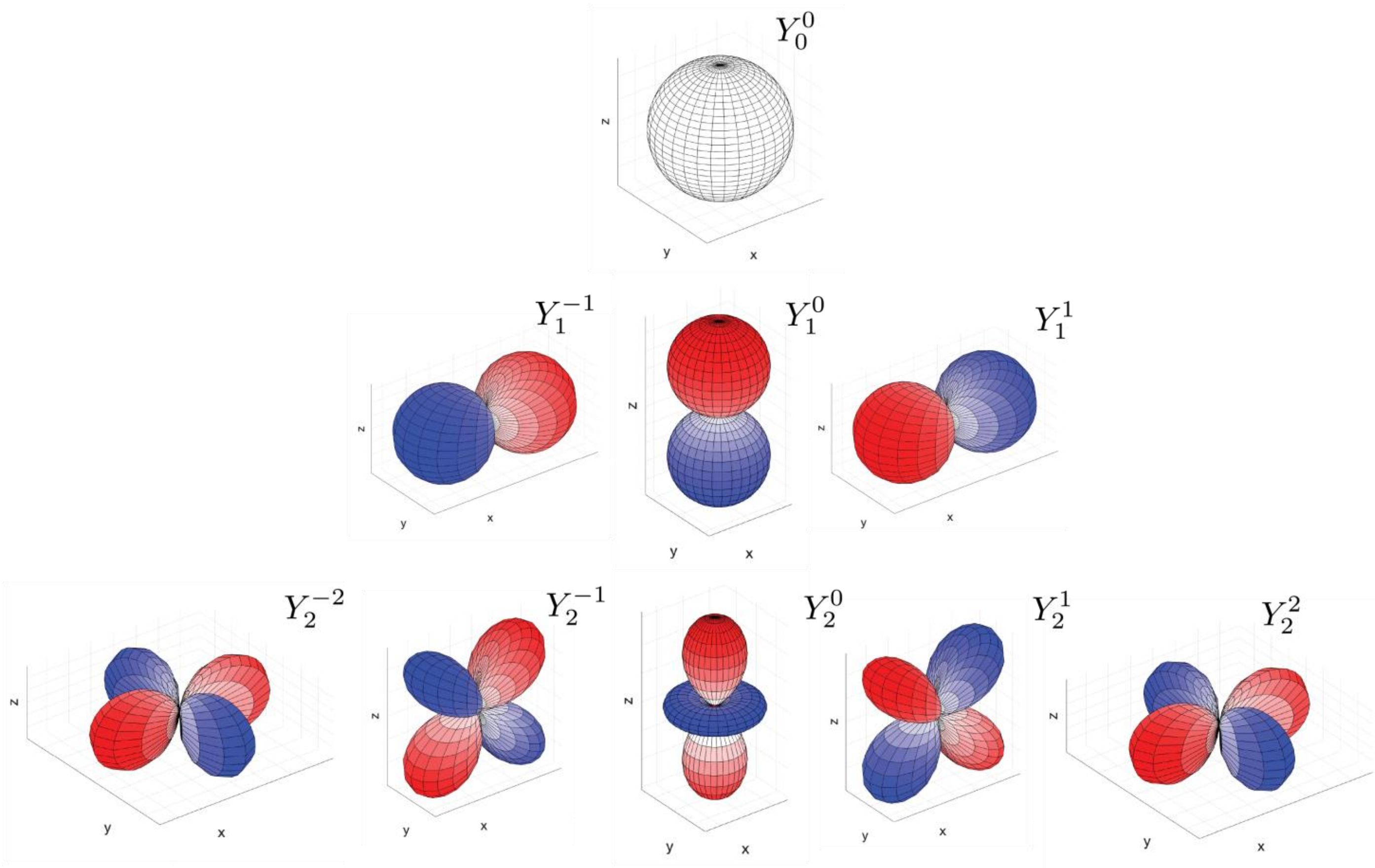
$$h_\nu(z) = j_\nu(z) + \mathrm{j} n_\nu(z)$$

Spherical wavefunction expansion

- Spherical harmonic function

Associated Legendre function

$$Y_{\nu,\mu}(\theta,\phi)=\sqrt{\frac{(2\nu+1)}{4\pi}\frac{(\nu-\mu)!}{(\nu+\mu)!}}P_{\nu}^{\mu}(\cos\theta)\mathrm{e}^{\mathrm{j}\mu\phi}$$



Orthogonality of spherical wavefunctions

Spherical wavefunctions are orthogonal set of functions satisfying homogeneous Helmholtz eq

- Spherical Bessel function

$$\int_{-\infty}^{\infty} j_\nu(kr) j_{\nu'}(kr) \mathrm{d}r = \frac{\pi}{k(2n+1)} \delta_{\nu,\nu'}$$

- Spherical harmonic function

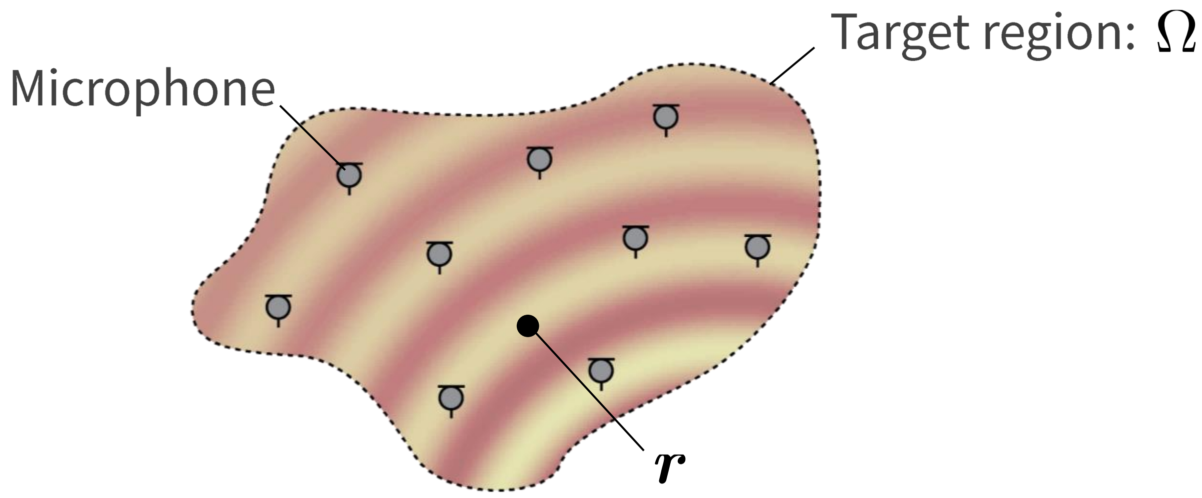
$$\int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} Y_{\nu,\mu}(\theta,\phi) Y_{\nu',\mu'}(\theta,\phi)^* \sin\theta \mathrm{d}\theta = \delta_{\nu,\nu'} \delta_{\mu,\mu'}$$

Table of contents

1. Preliminaries on acoustics
    1. Wave equation and Helmholtz equation
    2. Representations of acoustic field
2. Integral-equation-based sound field estimation
    1. Problem formulation
    2. Sound field estimation with spherical array
    3. Methods to avoid forbidden frequency problem
3. Least-squares-based sound field estimation
    1. Basic framework
    2. Infinite-dimensional extension
    3. Other extensions and related works
4. Applications
    1. Spatial audio reproduction by headphones
    2. Spatial audio reproduction by loudspeakers
    3. Spatial active noise control

Sound field estimation problem

Formulation of sound field estimation problem



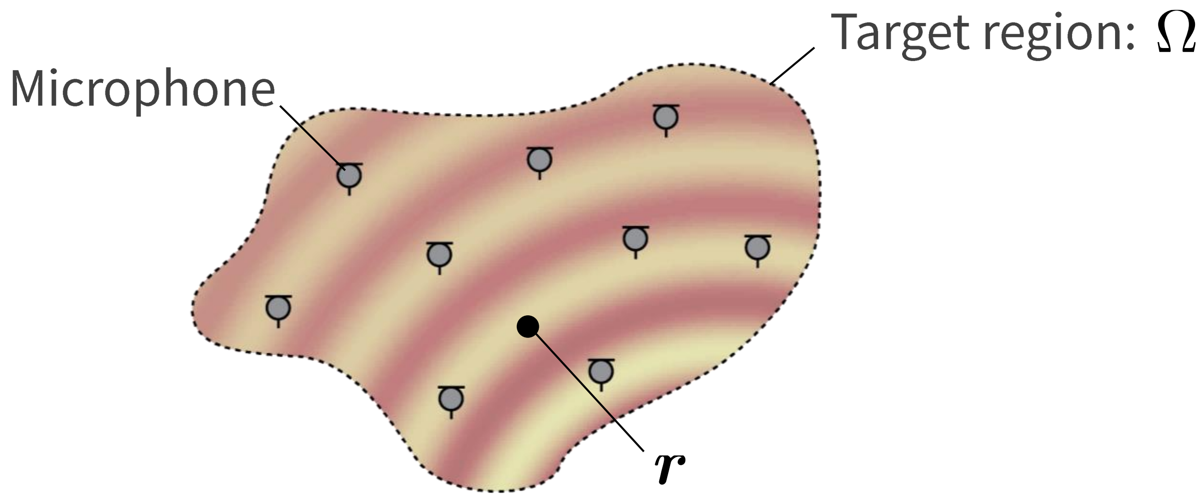
(P1)

Estimate pressure distribution $u(\boldsymbol{r})$ $(\boldsymbol{r} \in \Omega)$ with observations $\{s_m\}_{m=1}^M$ at discrete set of M mics $\{\boldsymbol{r}_m\}_{m=1}^M$

➡ Ω : Source-free and simply-connected interior region

Sound field estimation problem

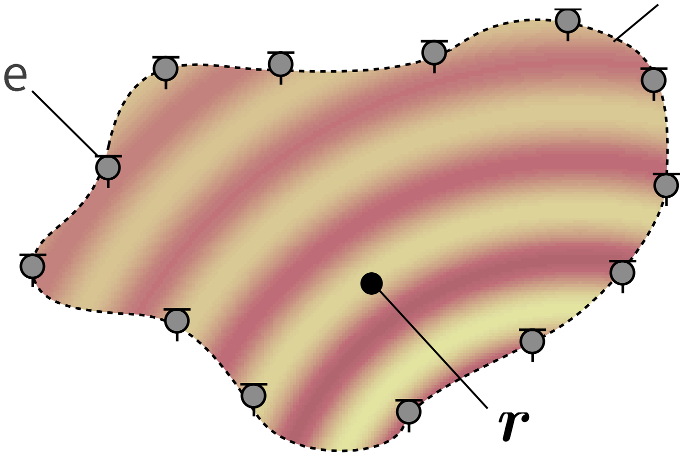
Two major categories of sound field estimation methods



- Integral-equation-based method
  - Based on discretization of boundary integral equation
- Least-squares-based method
  - Based on minimization of squre error

Integral-equation-based method

Sound field estimation based on KH integral…?



$$u(\boldsymbol{r}) = \int_{\boldsymbol{r}' \in \partial\Omega} \left(G(\boldsymbol{r}|\boldsymbol{r}') \frac{\partial u(\boldsymbol{r}')}{\partial \boldsymbol{n}'} - u(\boldsymbol{r}') \frac{\partial G(\boldsymbol{r}|\boldsymbol{r}')}{\partial \boldsymbol{n}'} \right) \mathrm{d}\boldsymbol{r}'$$

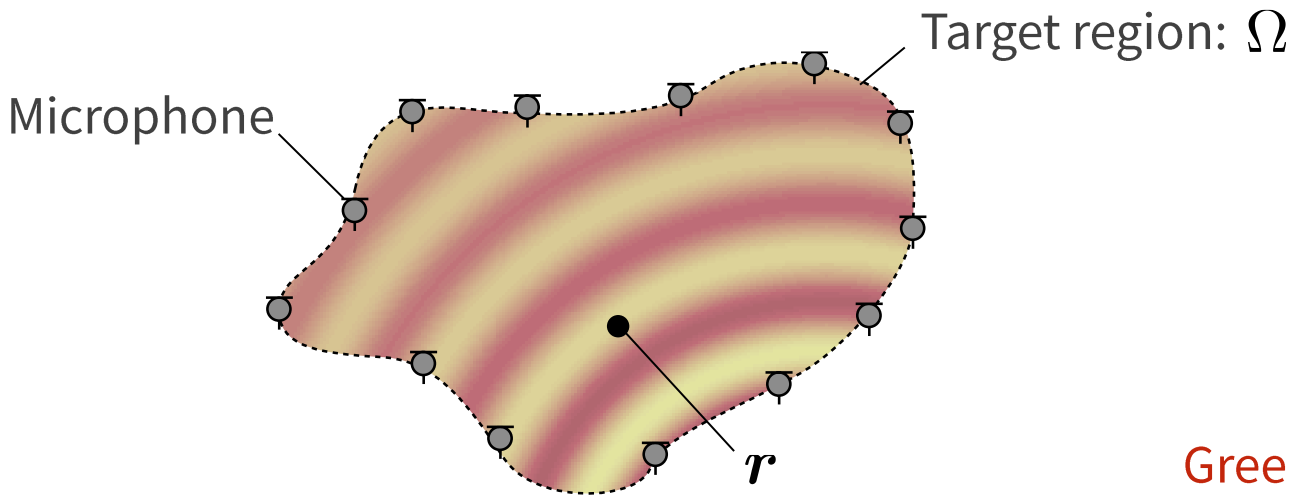
Pressure: $u(\boldsymbol{r}')$

Pressure gradient: $\frac{\partial u(\boldsymbol{r}')}{\partial \boldsymbol{n}'}$

- Pressure and its gradient on the boundary must be measured
- Ordinary omni-directional mics can be used to measure pressure, but measuring pressure gradient is not simple

Integral-equation-based method

Sound field estimation based on KH integral⋯?



Green's func for Dirichlet BC

Pressure

$$u(\boldsymbol{r}) = \int_{\boldsymbol{r}' \in \partial\Omega} \left(G_{\mathrm{D}}(\boldsymbol{r}|\boldsymbol{r}') \frac{\partial u(\boldsymbol{r}')}{\partial \boldsymbol{n}'} - u(\boldsymbol{r}') \frac{\partial G_{\mathrm{D}}(\boldsymbol{r}|\boldsymbol{r}')}{\partial \boldsymbol{n}'} \right) \mathrm{d}\boldsymbol{r}'$$

Pressure gradient

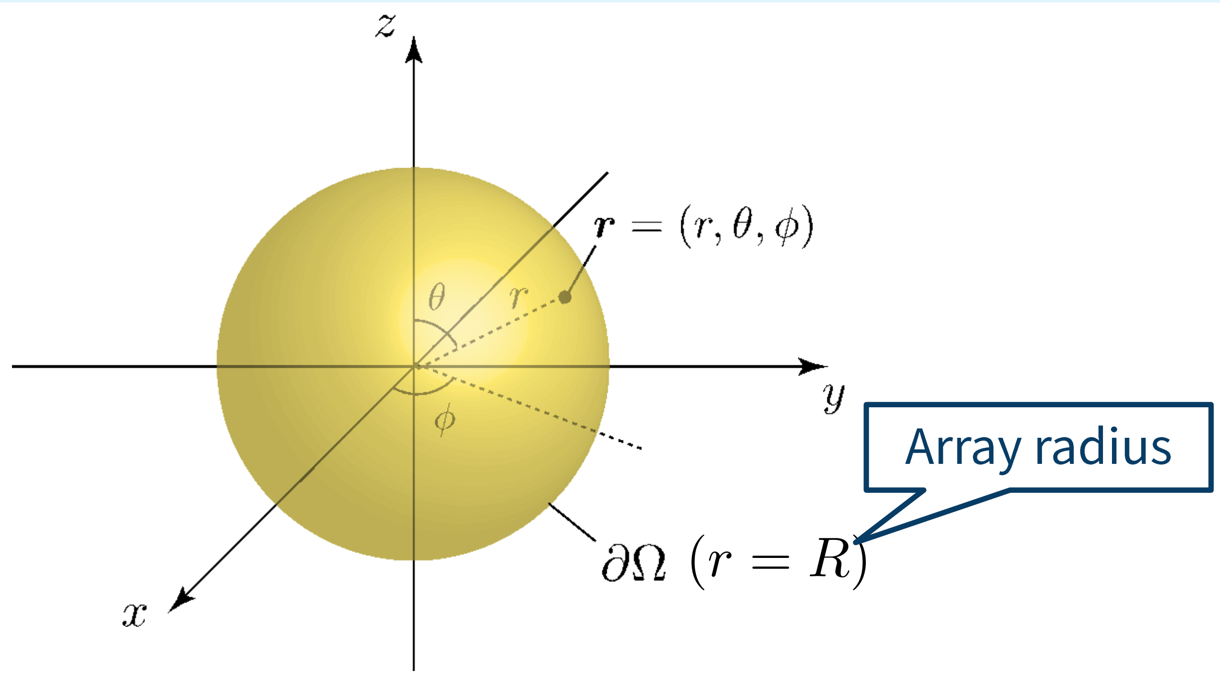
- Using only pressure on the boundary with Dirichlet Green's func
  - Analytical formulation is possible only for simple shape of Ω
  - Forbidden frequency problem for closed-shape of Ω

Sound field estimation with spherical mic array

- Simplify the problem by setting Ω to sphere of radius R
- Spherical array is typically used for spatial audio recording

(P2)

Estimate expansion coefficients at array center $\boldsymbol{r}_0$, i.e., $\mathring{u}_{\nu,\mu}(\boldsymbol{r}_0)$ with observations $\{s_m\}_{m=1}^{M}$ on spherical surface $\partial\Omega$



Sound field estimation with spherical mic array

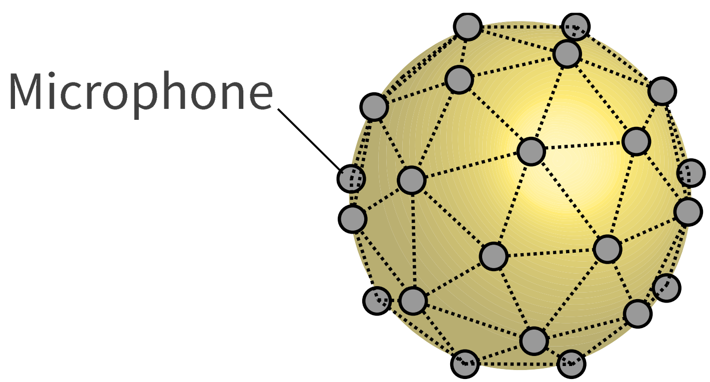
- Given pressure distribution on $\partial\Omega$, spherical harmonic coefficients $U_{\nu,\mu}(R)$ are calculated as

$$U_{\nu,\mu}(R) = \int_0^{2\pi} \int_0^{\pi} u(R,\theta,\phi) Y_{\nu,\mu}(\theta,\phi)^* \sin\theta \mathrm{d}\theta \mathrm{d}\phi$$

- Discretization by M mic positions on $\partial\Omega$

$$U_{\nu,\mu}(R) \approx \sum_{m=1}^{M} w_m u(R,\theta_m,\phi_m) Y_{\nu,\mu}(\theta_m,\phi_m)^*$$

Weight (w_m) Observation s_m ($u(R,\theta_m,\phi_m)$)



Sound field estimation with spherical mic array

- $U_{\nu,\mu}(R)$ and $\mathring{u}_{\nu,\mu}$ are related as

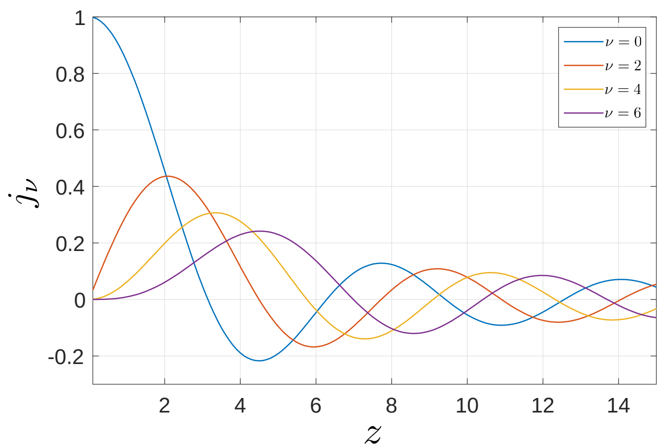
$$\begin{cases} u(R,\theta,\phi) = \sum_{\nu=0}^{\infty}\sum_{\mu=-\nu}^{\nu} U_{\nu,\mu}(R) Y_{\nu,\mu}(\theta,\phi) \\ u(R,\theta,\phi) = \sum_{\nu=0}^{\infty}\sum_{\mu=-\nu}^{\nu} \mathring{u}_{\nu,\mu}\sqrt{4\pi} j_\nu(kR) Y_{\nu,\mu}(\theta,\phi) \end{cases}$$

$$\Rightarrow \quad U_{\nu,\mu}(R) = \mathring{u}_{\nu,\mu}\sqrt{4\pi} j_\nu(kR)$$

- Expansion coefficients $\mathring{u}_{\nu,\mu}$ are estimated as

$$\hat{\mathring{u}}_{\nu,\mu} = \frac{1}{\sqrt{4\pi} j_\nu(kR)} U_{\nu,\mu}(R)$$

$\Rightarrow$ Incomputable when $j_\nu(kR) = 0$!
(forbidden frequency problem)



How to avoid forbidden frequency problem?

- Several established techniques for avoiding forbidden frequency problem
  1. Mics mounted on rigid spherical baffle
  2. Array of directional mics (e.g., unidirectional mics)
  3. Two (or more) layers of spherical mic array

1.

mh acoustics
em32 Eigenmike®

2.

Core Sound
OctoMic™

3.

[Jin+ IEEE/ACM TASLP 2014]

Estimation by rigid spherical mic array

- Sound field scattered by rigid spherical baffle of radius R

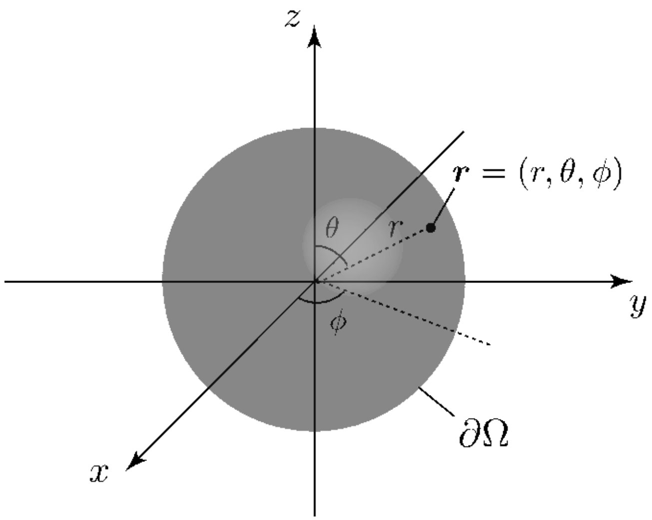
$$\begin{cases} (\nabla^2 + k^2) u(\boldsymbol{r}) = 0 & \text{in } \mathbb{R}^3 \backslash \Omega \\ \left. \dfrac{\partial u(\boldsymbol{r})}{\partial r} \right|_{r=R} = 0 & \text{on } \partial\Omega \end{cases}$$

- Total sound field is represented by sum of incident and scattered fields

Scattered field

$$u(\boldsymbol{r}) = u_{\mathrm{inc}}(\boldsymbol{r}) + u_{\mathrm{sct}}(\boldsymbol{r})$$

Incident field



Estimation by rigid spherical mic array

- Incident and scattered fields can be expanded by interior and exterior spherical wavefunctions, respectively

$$u_{\mathrm{inc}}(\boldsymbol{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\mathrm{inc},\nu,\mu} \sqrt{4\pi} j_\nu(kr) Y_{\nu,\mu}(\theta, \phi)$$

$$u_{\mathrm{sct}}(\boldsymbol{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\mathrm{sct},\nu,\mu} \sqrt{4\pi} h_\nu(kr) Y_{\nu,\mu}(\theta, \phi)$$

- From the boundary condition on $\partial\Omega$

$$\begin{aligned} \left.\frac{\partial u(\boldsymbol{r})}{\partial r}\right|_{r=R} &= \left.\frac{\partial}{\partial r} \left(u_{\mathrm{inc}}(\boldsymbol{r}) + u_{\mathrm{sct}}(\boldsymbol{r})\right)\right|_{r=R} \\ &= \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \sqrt{4\pi} k \left[\mathring{u}_{\mathrm{inc},\nu,\mu} j'_\nu(kR) + \mathring{u}_{\mathrm{sct},\nu,\mu} h'_\nu(kR)\right] Y_{\nu,\mu}(\theta, \phi) \\ &= 0 \end{aligned}$$

$$\Rightarrow \quad \mathring{u}_{\mathrm{sct},\nu,\mu} = -\frac{j'_\nu(kR)}{h'_\nu(kR)} \mathring{u}_{\mathrm{inc},\nu,\mu}$$

Relationship between incident and scattered fields

Estimation by rigid spherical mic array

- Total sound field is represented only by interior expansion coefficients

$$u(\boldsymbol{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\mathrm{inc},\nu,\mu} \sqrt{4\pi} \left[j_\nu(kr) - \frac{j'_\nu(kR)}{h'_\nu(kR)} h_\nu(kr) \right] Y_{\nu,\mu}(\theta,\phi)$$

- Expansion coeffs of incident field $\mathring{u}_{\mathrm{inc},\nu,\mu}$ (i.e., sound field without rigid buffle) is represented by using spherical harmonic coeffs on $\partial\Omega$

$$\begin{cases} u(R,\theta,\phi) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} U_{\nu,\mu}(R) Y_{\nu,\mu}(\theta,\phi) \\ u(R,\theta,\phi) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\mathrm{inc},\nu,\mu} \sqrt{4\pi} \left[j_\nu(kR) - \frac{j'_\nu(kR)}{h'_\nu(kR)} h_\nu(kR) \right] Y_{\nu,\mu}(\theta,\phi) \end{cases}$$

$$\Rightarrow \quad \mathring{u}_{\mathrm{inc},\nu,\mu} = \frac{1}{\sqrt{4\pi} \left[j_\nu(kR) - \frac{j'_\nu(kR)}{h'_\nu(kR)} h_\nu(kR) \right]} U_{\nu,\mu}(R)$$

Estimation by rigid spherical mic array

[Poletti 2005]

- Desired expansion coeffs are estimated as

$$\mathring{\hat{u}}_{\nu,\mu} = \frac{1}{\sqrt{4\pi}\left[j_\nu(kR) - \frac{j'_\nu(kR)}{h'_\nu(kR)}h_\nu(kR)\right]}U_{\nu,\mu}(R)$$

$$= \frac{h'_\nu(kR)}{\sqrt{4\pi}\left[j_\nu(kR)h'_\nu(kR) - j'_\nu(kR)h_\nu(kR)\right]}U_{\nu,\mu}(R)$$

$$= -\frac{\mathrm{j}k^2R^2}{\sqrt{2\pi}}h'_\nu(kR)U_{\nu,\mu}(R)$$

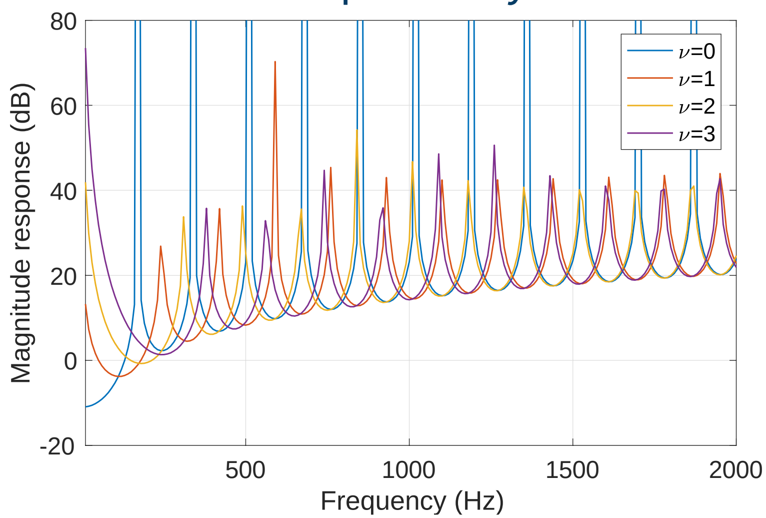
Wronskian relation:

$$j_\nu(z)h'_\nu(z) - j'_\nu(z)h_\nu(z) = \frac{\mathrm{j}}{z^2}$$

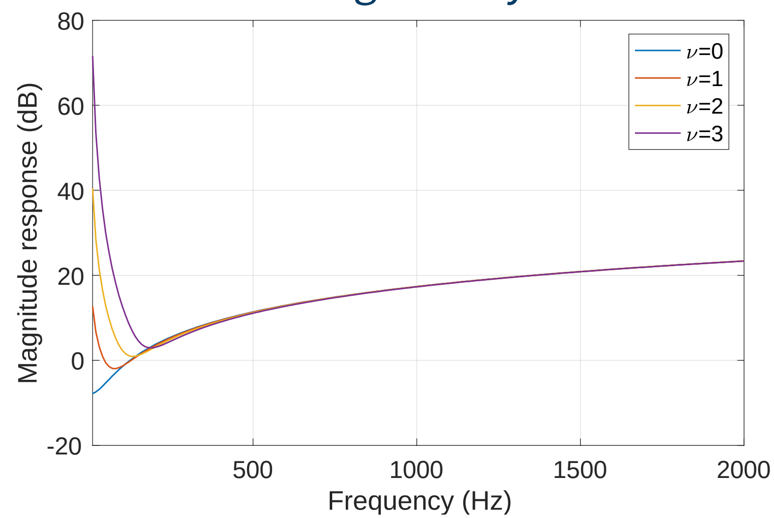
More robust than open spherical mic array

- Comparison of array responses

Open array



Rigid array



Estimation by spherical array of directional mics

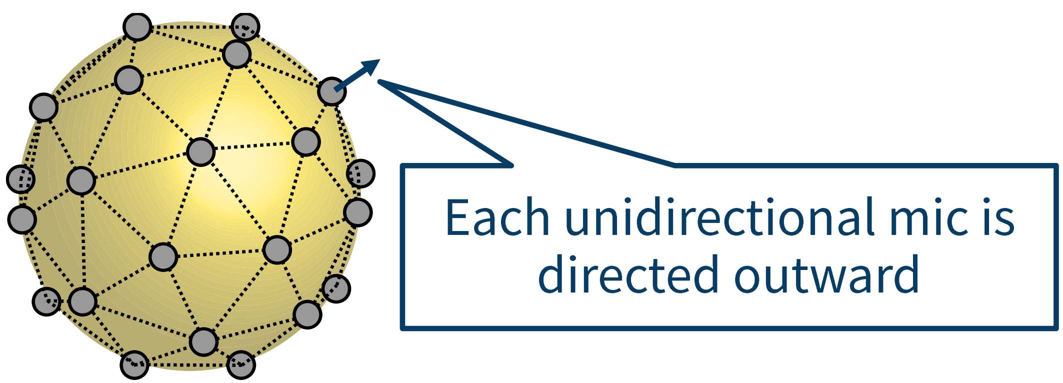
- Suppose that unidirectional mics are directed outward on $\partial\Omega$
- Observation s_m can be modeled as

$$s_m = \alpha u(R, \theta_m, \phi_m) + (1-\alpha)\,\frac{1}{\mathrm{j}k}\,\frac{\partial u(r, \theta_m, \phi_m)}{\partial r}\bigg|_{r=R}$$

where $\alpha \in [0, 1)$ is constant

- By using spherical wavefunction expansion,

$$s_m = \sum_{\nu=0}^{\infty}\sum_{\mu=-\nu}^{\nu} \sqrt{4\pi}\,[\alpha j_\nu(kR) - \mathrm{j}(1-\alpha) j'_\nu(kR)]\,\mathring{u}_{\nu,\mu} Y_{\nu,\mu}(\theta_m, \phi_m)$$



Estimation by spherical array of directional mics

[Poletti 2005]

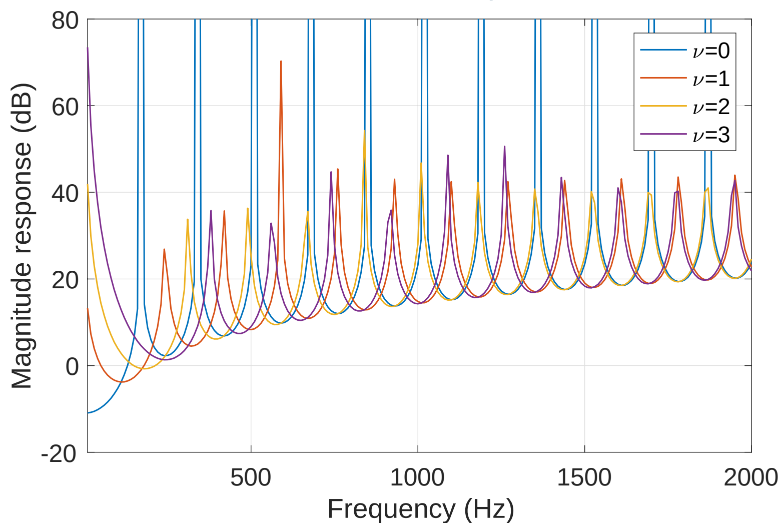
- Desired expansion coeffs are estimated as

$$\hat{\mathring{u}}_{\nu,\mu} = \frac{1}{\sqrt{4\pi}\left[\alpha j_\nu(kR) - \mathrm{j}(1-\alpha) j'_\nu(kR)\right]} U_{\nu,\mu}(R)$$

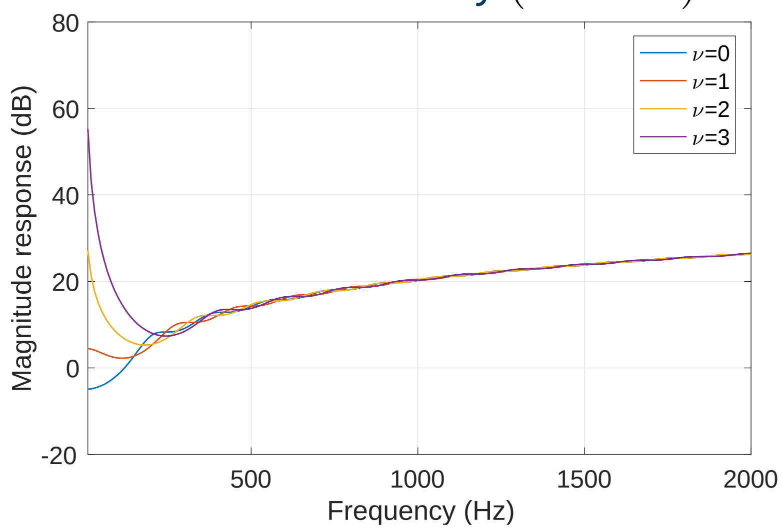
More robust than omnidirectional spherical mic array

- Comparison of array response

Open array



Directional array $(\alpha = 0.5)$



Sampling on sphere

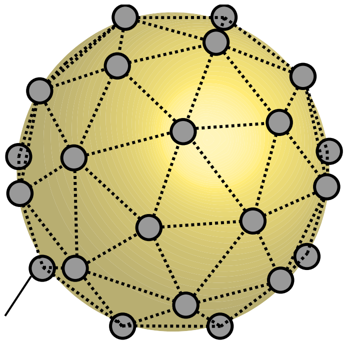
Integral-equation-based methods rely on computation of spherical harmonic coeffs on sphere by discrete set of mics

$$U_{\nu,\mu}(R) \approx \sum_{m=1}^{M} w_m s_m Y_{\nu,\mu}(\theta_m, \phi_m)^*$$

Sampling weight (w_m)

Observation (s_m)

- Several sampling schemes on spherical surface [Rafaely 2015]
  - Equal-angle sampling
  - Gaussian sampling
  - Uniform / Nearly uniform sampling



Conclusion

- Integral-equation-based sound field estimation
  - Stable computation and useful for analyzing properties bacause of analytical formulation of estimator
  - Forbidden frequency problem for closed region of interest, but there exist several establised methods for avoiding it
  - Applicable only to simple array geometry, e.g., sphere, plane, and cylinder, to derive analytical formulation
  - Similar formulation can be derived for linear and circular arrays, but some approximations are necessary for 2D arrays [Koyama+ 2013, 2014, 2016]

References (1)

- E. G. Williams, “Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography,” Academic Press, 1999.
- J. Meyer and G. Elko, “A highly scalable spherical microphone array based on an orthogonal decomposition of soundfield,” in Proc. IEEE ICASSP, 2002.
- T. D. Abhayapala and D. B. Ward, “Theory and design of high order sound field microphones using spherical microphone arrays,” in Proc. IEEE ICASSP, 2002.
- J. Daniel, S. Moureau, and R. Nicol, “Further investigations of higher-order ambisonics and wavefield synthesis for holophonic sound imaging,” in Proc. 114th AES Convention, 2003.
- N. A. Gumerov and R. Duraiswami, “Fast Multipole Methods for the Helmholtz Equation in Three Dimensions,” Elsevier, 2004.
- M. A. Poletti, “Three-dimensional surround sound systems based on spherical harmonics,” J. Audio Eng. Soc., 53(11), pp. 1004-1025, 2005.
- B. Rafaely, “Analysis and design of spherical microphone arrays,” IEEE Trans. Audio, Speech, Lang., Process., 13(1), pp. 135-143, 2005.
- P. A. Martin, ”Multiple Scattering: Interaction of Time-Harmonic Waves with N Obstacles,” Cambridge University Press, 2006.
- H. Kuttruff, “Room Acoustics,” CRC Press, 2009.

References (2)

- D. Colton and R. Kress, “Inverse Acoustics and Electromagnetic Scattering Theory,” Springer, 2013.
- 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., 21(4), pp. 685-696, 2013.
- C. T. Jin, N. Epain, A. Parthy, “Design, optimization and evaluation of a dual-radius spherical microphone array,” IEEE/ACM Trans. Audio, Speech, Lang., Process., 22(1), pp. 193-204, 2014.
- 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., 22(10), pp. 1546-1557, 2014.
- A. Kirsch and F. Hettlich, ”The Mathematical Theory of Time-Harmonic Maxwell’s Equations: Expansion-, Integral-, and Variational Methods,” Springer, 2015.
- B. Rafaely, “Fundamentals of Spherical Array Processing,” Springer, 2015.
- 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. Am., 139(3), pp. 1024-1036, 2016.

T-5

May 22, 2022

Sound Field Estimation: Recent Advances and Applications

— Section 3 —

Shoichi Koyama[1], Natsuki Ueno[2]

[1]The University of Tokyo

[2]Tokyo Metropolitan University

22-27 MAY 2022

ICASSP 2022

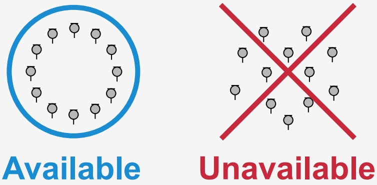
SINGAPORE

Table of Contents

1. **Preliminaries on acoustics**
    1. Wave equation and Helmholtz equation
    2. Representations of acoustic field
2. **Integral-equation-based sound field estimation**
    1. Problem formulation
    2. Sound field estimation with spherical array
    3. Methods to avoid forbidden frequency problem
3. **Least-squares-based sound field estimation**
    1. Basic framework
    2. Infinite-dimensional extension
    3. Other extensions and related works
4. **Applications**
    1. Spatial audio reproduction by headphones
    2. Spatial audio reproduction by loudspeakers
    3. Spatial active noise control

Limitation of integral-equation-based method

- Simple array geometry (e.g., sphere, plane)
- Simple microphone directivity (e.g., omnidirectional, bidirectional)

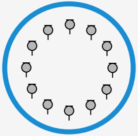


More flexible method is desired.

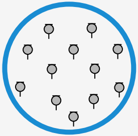
Least-squares-based sound field estimation

[Laborie+, 2003], [Poletti, 2005], etc.

- Applicable to arbitrary array geometry and microphone directivity
- Based on decomposition of sound field into basis functions



Available



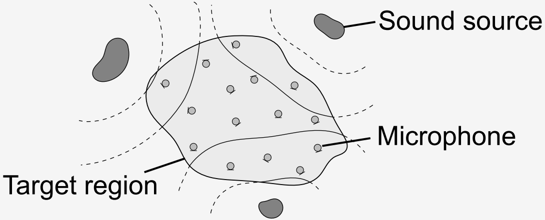
Available

1. Basic framework

2. Infinite-dimensional extension

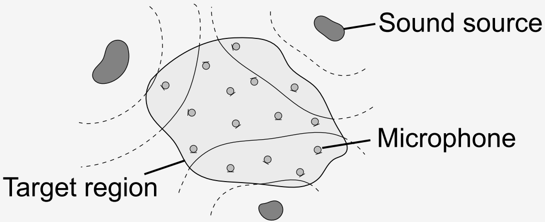
3. Other extensions and related works

Problem settings (1)



Assumption

- Positions and directivities of microphones are given.
- Target region Ω is simply-connected and source-free.



- **Objective** (P1)
  - ▹ To estimate sound field $u(\boldsymbol{r})$ $(\boldsymbol{r} \in \Omega)$ from signals $\{s_m\}_{m=1}^{M}$ observed by M microphones

Four steps in least-squares-based method

- Step 1: decomposition of sound field
  - By spherical wavefunctions or plane wave functions
- Step 2: formulation of observation model
  - Based on microphone position and directivity
- Step 3: formulation of optimization problem
  - As regularized least squares
- Step 4: derivation of optimal solution
  - In closed form

Step 1/4: decomposition of sound field

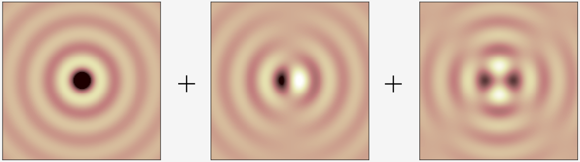
$$u(\boldsymbol{r}) \approx \sum_n \underbrace{a_n}_{\text{expansion coefficient}} \underbrace{\psi_n(\boldsymbol{r})}_{\text{basis function}}$$

- Examples of basis function
  - Spherical wavefunction [Poletti, 2005]
  - Plane wave function [Chardon+, 2012]

Step 1/4: decomposition of sound field

- Expansion by **spherical wavefunctions** [Laborie+, 2003]

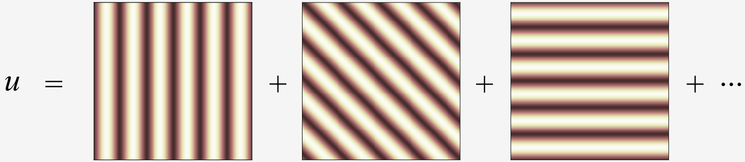
$$u(\boldsymbol{r}) \approx \sum_{\nu=0}^{N} \sum_{\mu=-\nu}^{\nu} \underbrace{\mathring{u}_{\nu,\mu}}_{\text{expansion coefficient}} \underbrace{\sqrt{4\pi} j_{\nu}(k\|\boldsymbol{r}\|) Y_{\nu,\mu}(\boldsymbol{r}/\|\boldsymbol{r}\|)}_{\text{spherical wavefunction}}$$

$u =$  $+ \cdots$

Step 1/4: decomposition of sound field

- Expansion by **plane wave functions** [Chardon+, 2012]

$$u(\boldsymbol{r}) \approx \underbrace{\sum_{n=1}^{N} \tilde{u}_n}_{\text{expansion coefficient}} \underbrace{\exp(-\mathrm{j}k\boldsymbol{x_n} \cdot \boldsymbol{r})}_{\text{plane wave function}}$$



Step 2/4: formulation of observation model

- Decomposition of sound field

$$u(\boldsymbol{r}) \approx \sum_n \underbrace{a_n}_{\text{expansion coefficient}} \underbrace{\psi_n(\boldsymbol{r})}_{\text{basis function}}$$

- Observation by mth microphone (superposition principle)

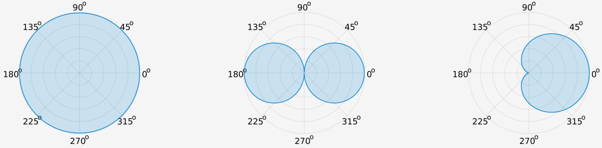
$$\underbrace{s_m}_{\text{observed signal}} = \sum_n a_n \underbrace{c_{m,n}}_{\text{response to } \psi_n} + \underbrace{\epsilon_m}_{\text{sensor noise}}$$

Step 2/4: formulation of observation model

- Observation by mth microphone

$$\underbrace{s_m}_{\text{observed signal}} = \sum_n a_n \underbrace{c_{m,n}}_{\text{response to } \psi_n} + \underbrace{\epsilon_m}_{\text{sensor noise}}$$

- $c_{m,n}$: determined by microphone's position and directivity
  - See [Laborie+, 2003] for detail.



Step 2/4: formulation of observation model

- Observation by mth microphone

$$\underbrace{s_m}_{\text{observed signal}} = \sum_n a_n \underbrace{c_{m,n}}_{\text{response to } \psi_n} + \underbrace{\epsilon_m}_{\text{sensor noise}}$$

⬇

- Matrix-vector representation

$$\boldsymbol{s} = \boldsymbol{C} \underbrace{\boldsymbol{a}}_{\text{to be estimated}} + \boldsymbol{\epsilon}$$

Step 3/4: formulation of optimization problem

$$\underset{\boldsymbol{a}}{\text{minimize}} \;\; \mathcal{J}(\boldsymbol{a}) = \underbrace{\|\boldsymbol{C}\boldsymbol{a} - \boldsymbol{s}\|_2^2}_{\text{loss term}} + \underbrace{\lambda\|\boldsymbol{a}\|_2^2}_{\text{regularization term}}$$

- Minimization of loss term
  ➡ to fit with observation
- Minimization of reguralization term
  ➡ to avoid overfitting

Formulated as regularized least squares

Step 4/4: derivation of optimal solution

- Objective function

$$\mathcal{J}(\boldsymbol{a}) = \|\boldsymbol{C}\boldsymbol{a} - \boldsymbol{s}\|_2^2 + \lambda\|\boldsymbol{a}\|_2^2$$

minimization

- Optimal solution

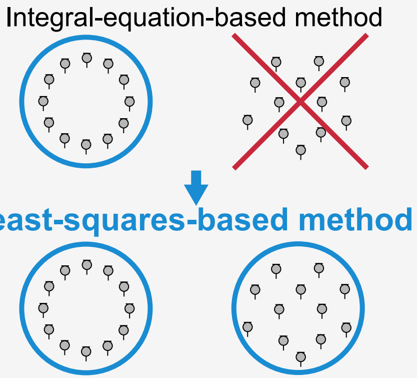
$$\hat{\boldsymbol{a}} = \boldsymbol{C}^{\mathsf{H}}(\boldsymbol{C}\boldsymbol{C}^{\mathsf{H}} + \lambda\boldsymbol{I})^{-1}\boldsymbol{s}$$

- **Estimated sound field**

$$\hat{u}(\boldsymbol{r}) = \sum_n \underbrace{\hat{a}_n}_{\text{estimated coefficient}} \underbrace{\psi_n(\boldsymbol{r})}_{\text{basis function}}$$

Applicability in general cases

- No constraints on microphone positions and directivities
- No necessity to define "boundary" of array



Necessity and availability of regularization

- Numerical unstability in case of insufficient regularization
- Performance improvement by exploiting prior information (will be described later) [Ueno+, 2021]

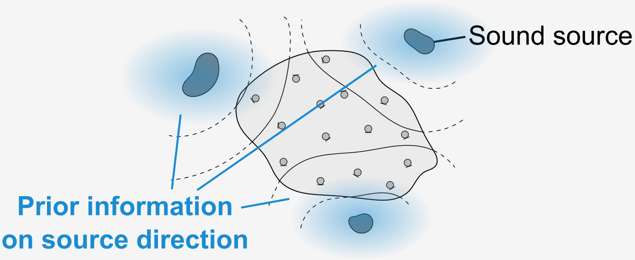


Table of Contents

1. Basic framework

2. Infinite-dimensional extension

3. Other extensions and related works

- **Limitation of finite-dimensional decomposition of sound field**
  - Necessity of parameter setting (in an emprical manner)
    - Number of basis functions
    - Position of expansion center (spherical wavefunction)
    - Direction of xyz-axes (plane wave function)

⬇

Performance degradation for inappropriate parameters

Infinite-dimensional extension [Ueno+, 2018], etc.

- No necessity of parameter setting
  - Number of basis functions
    ➡ infinite dimensions
  - Position of expansion center (spherical wavefunction)
    ➡ translation invariant
  - Direction of xyz-axes (plane wave function)
    ➡ rotation invariant

Four steps in infinite-dimensional extension

- Step 1: infinite-dimensional representation of sound field
  - As vector in Hilbert space
- Step 2: formulation of observation model
  - Based on microphone position and directivity
- Step 3: formulation of optimization problem
  - As regularized least squares
- Step 4: derivation of optimal solution
  - In closed form

Step 1/4: infinite-dimensional representation of sound field

- Expansion by spherical wavefunction

$$u(\boldsymbol{r}) \approx \sum_{\nu=0}^{N} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\nu,\mu} \sqrt{4\pi} j_{\nu}(k\|\boldsymbol{r}\|) Y_{\nu,\mu}(\boldsymbol{r}/\|\boldsymbol{r}\|)$$

⬇ **Infinite-dimensional extension**

$$u(\boldsymbol{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\nu,\mu} \sqrt{4\pi} j_{\nu}(k\|\boldsymbol{r}\|) Y_{\nu,\mu}(\boldsymbol{r}/\|\boldsymbol{r}\|)$$

Step 1/4: infinite-dimensional representation of sound field

- Expansion by spherical wavefunction

$$u(\boldsymbol{r}) \approx \sum_{\nu=1}^{N} \tilde{u}_n \exp(-\mathrm{j}k\boldsymbol{x}_n \cdot \boldsymbol{r})$$

Infinite-dimensional extension

$$u(\boldsymbol{r}) = \int_{\boldsymbol{x} \in \mathbb{S}_2} \tilde{u}(\boldsymbol{x}) \exp(-\mathrm{j}k\boldsymbol{x} \cdot \boldsymbol{r})\, \mathrm{d}\chi$$

- $\int_{\boldsymbol{x} \in \mathbb{S}_2} \mathrm{d}\chi$: spherical integral

Step 1/4: infinite-dimensional representation of sound field

- Equivalence of two representations

$$\mathring{u}_{\nu,\mu} = \frac{\sqrt{4\pi}}{\mathrm{j}^{\nu}} \int_{\boldsymbol{x}\in\mathbb{S}_2} \tilde{u}(\boldsymbol{x}) Y_{\nu,\mu}(\boldsymbol{x})^* \,\mathrm{d}\chi$$

⇕ **spherical harmonic expansion**

$$\tilde{u}(\boldsymbol{x}) = \frac{\mathrm{j}^{\nu}}{\sqrt{4\pi}} \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\nu,\mu} Y_{\nu,\mu}(\boldsymbol{x})$$

Step 1/4: infinite-dimensional representation of sound field

- Norm on sound field

$$\|u\|_{\mathscr{H}} = \left(\sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} |\mathring{u}_{\nu,\mu}|^2 \right)^{\frac{1}{2}} = \left(4\pi \int_{\boldsymbol{x} \in \mathbb{S}_2} |\tilde{u}(\boldsymbol{x})|^2 \, \mathrm{d}\chi \right)^{\frac{1}{2}}$$

  - rotation/translation invariant

Step 1/4: infinite-dimensional representation of sound field

- Hilbert space (set of sound fields)

$$\mathscr{H} = \left\{ u(\boldsymbol{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\nu,\mu} \sqrt{4\pi} j_\nu(k\|\boldsymbol{r}\|) Y_{\nu,\mu}(\boldsymbol{r}/\|\boldsymbol{r}\|) \;\middle|\; \|u\|_{\mathscr{H}} = \left(\sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} |\mathring{u}_{\nu,\mu}|^2 \right)^{\frac{1}{2}} < \infty \right\}$$

$$= \left\{ u(\boldsymbol{r}) = \int_{\boldsymbol{x}\in\mathbb{S}_2} \tilde{u}(\boldsymbol{x}) \exp(-\mathrm{j}k\boldsymbol{x}\cdot\boldsymbol{r}) \, \mathrm{d}\chi \;\middle|\; \|u\|_{\mathscr{H}} = \left(4\pi \int_{\boldsymbol{x}\in\mathbb{S}_2} |\tilde{u}(\boldsymbol{x})|^2 \, \mathrm{d}\chi \right)^{\frac{1}{2}} < \infty \right\}$$

Step 1/4: infinite-dimensional representation of sound field

- Representation capability of $\mathscr{H}$
  - **Any solution of Helmholtz equation in Ω can be approximated arbitrarily by functions in $\mathscr{H}$ in sense of uniform convergence on compact sets.** [Ueno+, 2021]

⬇

Sufficient representation capability without any parameter

Step 2/4: formulation of observation model

- Observation by mth microphone

$$\underbrace{s_m}_{\text{observed signal}} = \underbrace{\mathcal{F}_m}_{\text{linear functional}} u + \underbrace{\epsilon_m}_{\text{sensor noise}}$$

- $\mathcal{F}_m$: determined by microphone's position and directivity

$$\mathcal{F}_m u = \int_{\boldsymbol{x} \in \mathbb{S}_2} \tilde{u}(\boldsymbol{x}) \exp(-\mathrm{j}k\boldsymbol{x} \cdot \underbrace{\boldsymbol{r}_m}_{\text{position}}) \underbrace{\gamma_m(\boldsymbol{x})}_{\text{directivity}} \mathrm{d}\chi$$

  - See [Ueno+, 2021] for detail.

Step 3/4: formulation of optimization problem

$$\underset{u \in \mathscr{H}}{\text{minimize}} \;\; \mathcal{J}(u) = \underbrace{\sum_{m=1}^{M} \frac{1}{\sigma_m^2} |\mathcal{F}_m u - s_m|^2}_{\text{loss term}} + \underbrace{\lambda \|u\|_{\mathscr{H}}^2}_{\text{regularization term}}$$

Formulated as regularized least squares in infinite-dimensional Hilbert space $\mathscr{H}$

How to solve?

Step 4/4: derivation of optimal solution

- Objective function

$$\mathcal{J}(u) = \sum_{m=1}^{M} \frac{1}{\sigma_m^2} \big| \underbrace{\mathcal{F}_m}_{\text{linear functional}} u - s_m \big|^2 + \lambda \|u\|_{\mathscr{H}}^2$$

reformulation

$$\mathcal{J}(u) = \sum_{m=1}^{M} \frac{1}{\sigma_m^2} \big| \underbrace{\langle v_m, u \rangle_{\mathscr{H}}}_{\text{inner product}} - s_m \big|^2 + \lambda \|u\|_{\mathscr{H}}^2$$

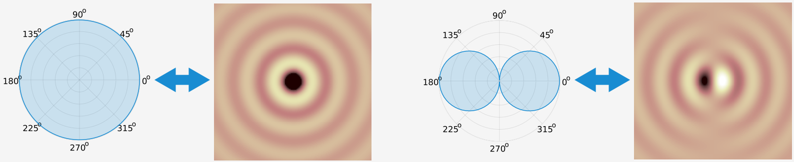
Infinite-dimensional extension (11)

Step 4/4: derivation of optimal solution

- v_m: determined by microphone's position and directivity

$$v_m(\boldsymbol{r}) = \frac{1}{4\pi} \int_{\boldsymbol{x} \in \mathbb{S}_2} \underbrace{\gamma_m(\boldsymbol{x})}_{\text{directivity}} \exp(-\mathrm{j}k\boldsymbol{x} \cdot (\boldsymbol{r} - \underbrace{\boldsymbol{r}_m}_{\text{position}})) \, \mathrm{d}\chi$$

  - Observation $\mathcal{F}_m$ ⟺ sound field v_m



Step 4/4: derivation of optimal solution

- v_m: calculated in closed form for finite-order directivity
  - e.g., omnidirectional, bidirectional, cardioid

$$\gamma_m(\boldsymbol{x})^* = \sum_{\nu=0}^{N_m} \sum_{\mu=-\nu}^{\nu} \underbrace{c_{m,\nu,\mu}}_{\text{harmonic coefficient of directivity}} Y_{\nu,\mu}(\boldsymbol{x})$$

⬇

$$v_m(\boldsymbol{r}) = \sum_{\nu,\mu}^{N_m} \frac{1}{\mathrm{j}^\nu} c_{m,\nu,\mu} j_\nu(k\|\boldsymbol{r}-\boldsymbol{r}_m\|) Y_{\nu,\mu}\left(\frac{\boldsymbol{r}-\boldsymbol{r}_m}{\|\boldsymbol{r}-\boldsymbol{r}_m\|}\right)$$

Step 4/4: derivation of optimal solution

- Objective function

$$\mathcal{J}(u) = \sum_{m=1}^{M} \frac{1}{\sigma_m^2} | \langle v_m, u \rangle_{\mathscr{H}} - s_m |^2 + \lambda \|u\|_{\mathscr{H}}^2$$

⬇ **representer theorem**

- Optimal solution

$$\hat{u}(\boldsymbol{r}) = \sum_{m=1}^{M} \underbrace{\hat{\alpha}_m}_{\text{to be determined}} v_m(\boldsymbol{r})$$

Step 4/4: derivation of optimal solution

- **Estimated sound field**

$$\hat{u}(\boldsymbol{r}) = \sum_{m=1}^{M} \hat{\alpha}_m v_m(\boldsymbol{r})$$

- **Optimal coefficient**

$$\hat{\boldsymbol{\alpha}} = \big(\underbrace{\boldsymbol{K}} + \lambda \underbrace{\boldsymbol{\Sigma}} \big)^{-1} \boldsymbol{s}$$

$$\boldsymbol{K} = \begin{bmatrix} \langle v_1, v_1 \rangle_{\mathscr{H}} & \cdots & \langle v_1, v_M \rangle_{\mathscr{H}} \\ \vdots & \ddots & \vdots \\ \langle v_M, v_1 \rangle_{\mathscr{H}} & \cdots & \langle v_M, v_M \rangle_{\mathscr{H}} \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1^2, \dots, \sigma_M^2)$$

Step 4/4: derivation of optimal solution

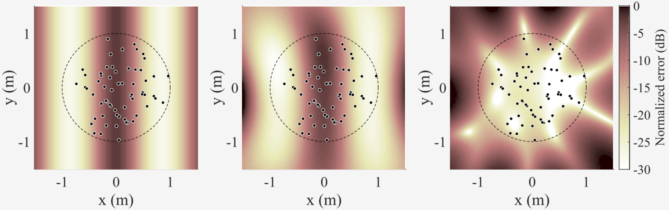
- **Estimated sound field**

$$\hat{u}(\boldsymbol{r}) = \sum_{m=1}^{M} \hat{\alpha}_m v_m(\boldsymbol{r})$$

- **Main points**
  - No necessity to set parameters for finite-dimensional decomposition (e.g., truncation order and expansion center)
  - Closed-form solution ➡ **easy to emplement**

Sound field estimation with irregularly ditributed microphones

- Experimental results using simulated data [Ueno+, 2019]
  - Black dots denote microphone positions.

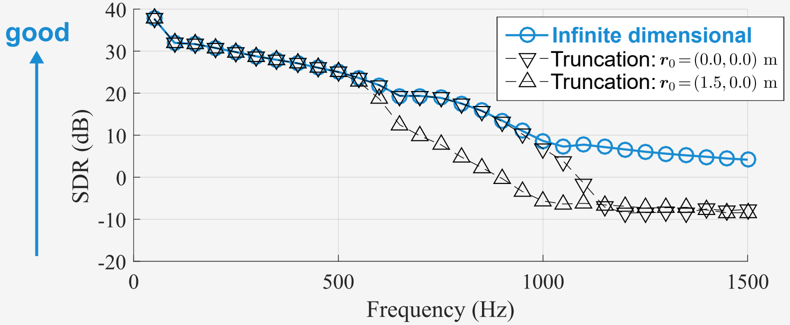


True   Estimated   Estimation error

Applicable to arbitrary array geometry

Finite- vs. infinite-dimensional modeling

- ▷ Experimental results using simulated data [Ueno+, 2018]
  - Estimation of plane wave field using 63 microphones



- **In case of pressure microphones (= interpolation problem of sound field)**
  - Estimated sound field

$$\hat{u}(\boldsymbol{r}) = \sum_{m=1}^{M} \hat{\alpha}_m \underbrace{\kappa(\boldsymbol{r}, \boldsymbol{r}_m)}_{= j_0(k\|\boldsymbol{r} - \boldsymbol{r}_m\|)}$$

$$\hat{\boldsymbol{\alpha}} = \Big(\underbrace{\boldsymbol{K}}_{= \begin{bmatrix} \kappa(\boldsymbol{r}_1, \boldsymbol{r}_1) & \cdots & \kappa(\boldsymbol{r}_1, \boldsymbol{r}_M) \\ \vdots & \ddots & \vdots \\ \kappa(\boldsymbol{r}_M, \boldsymbol{r}_1) & \cdots & \kappa(\boldsymbol{r}_M, \boldsymbol{r}_M) \end{bmatrix}} + \lambda \boldsymbol{\Sigma} \Big)^{-1} \boldsymbol{s}$$

In case of pressure microphones (= interpolation problem of sound field)

- Equivalent to kernel ridge regression with kernel function $\kappa(\boldsymbol{r}, \boldsymbol{r}') = j_0(k\|\boldsymbol{r} - \boldsymbol{r}'\|)$ [Ueno+, 2018]

⬇

Interpreted as kernel ridge regression with constraint of Helmholtz equation

- Demonstration with real data will be shown in Section 4.

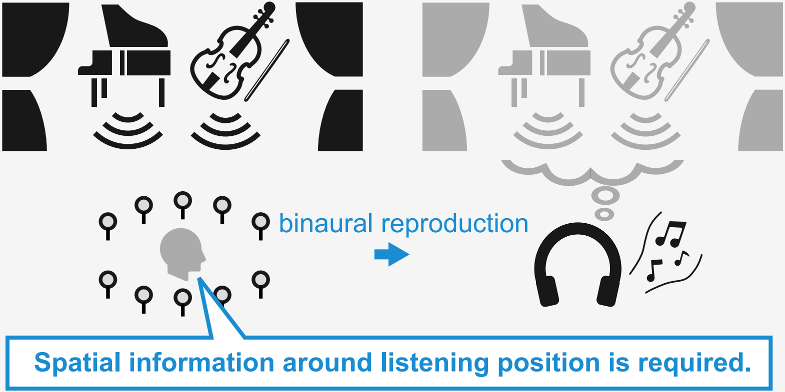
1. Basic framework

2. Infinite-dimensional extension

3. Other extensions and related works

Harmonic analysis of sound field

- e.g., sound field reproduction by loudspeakers [Ueno+, 2019] or headphones [IIjima+, 2021]



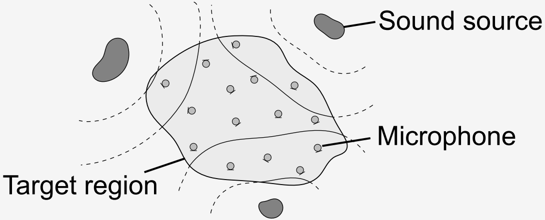
Harmonic analysis of sound field

- Spherical wavefunction expansion at listening position $\boldsymbol{r}_0$

$$u(\boldsymbol{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \underbrace{\mathring{u}_{\nu,\mu}(\boldsymbol{r}_0)}_{\text{spatial information around } \boldsymbol{r}_0} \sqrt{4\pi} j_\nu(k\|\boldsymbol{r}-\boldsymbol{r}_0\|) Y_{\nu,\mu}\left(\frac{\boldsymbol{r}-\boldsymbol{r}_0}{\|\boldsymbol{r}-\boldsymbol{r}_0\|}\right)$$

- For arbitrary listening position

  ➡ Estimation of $\mathring{u}_{\nu,\mu}(\boldsymbol{r}_0)$ at arbitrary $\boldsymbol{r}_0$ is required.



Objective (P3)

- To estimate expansion coefficient $\mathring{u}_{\nu,\mu}(\boldsymbol{r}_0)$ at arbitrary position $\boldsymbol{r}_0 \in \Omega$ from signals $\{s_m\}_{m=1}^{M}$ observed by M microphones

Translation of expansion coefficient

[Martin, 2006], [Samarasinghe+, 2014]

- Relationship between expansion coefficients at two different expansion center $\boldsymbol{r}_0$ and $\boldsymbol{r}_0'$

$$\mathring{u}_{\nu,\mu}(\boldsymbol{r}_0') = \sum_{\nu'=0}^{\infty} \sum_{\mu'=-\nu'}^{\nu'} \underbrace{T_{\nu,\mu}^{\nu',\mu'}(\boldsymbol{r}_0' - \boldsymbol{r}_0)}_{\text{translation operator}} \mathring{u}_{\nu',\mu'}(\boldsymbol{r}_0)$$

⬇

Desired to avoid infinite sums in estimation

Estimation of expansion coefficient at arbitrary expansion center [Ueno+, 2019]

$$\underbrace{\hat{\mathring{\boldsymbol{u}}}(\boldsymbol{r}_0)}_{\substack{\text{estimated coefficients} \\ (\infty \times 1 \text{ vector})}} = \underbrace{\boldsymbol{\Xi}(\boldsymbol{r}_0)}_{\substack{\text{coefficients of } v_1, \ldots, v_M \\ (\infty \times M \text{ matrix})}} (\boldsymbol{K} + \lambda \boldsymbol{\Sigma})^{-1} \boldsymbol{s}$$

- No truncation in expansion of sound field or in translation of expansion coefficient

Source code for Matlab is available at
https://codeocean.com/capsule/7630421

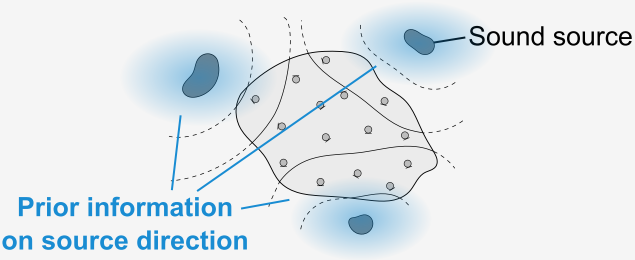
Regularization in sound field estimation

$$\underset{u \in \mathscr{H}}{\text{minimize}} \;\; \mathcal{J}(u) = \sum_{m=1}^{M} \frac{1}{\sigma_m^2} |\mathcal{F}_m u - s_m|^2 + \underbrace{\lambda \|u\|_{\mathscr{H}}^2}_{\text{regularization term}}$$

- Design of regularization exploiting prior information
  - Likely to occur ➡ **small norm**
  - Unlikely to occur ➡ **large norm**

Sound field estimation exploiting prior information on source direction [Ueno+, 2021]

- Based on directional weighting for norm of sound field



Sound field estimation exploiting prior information on source direction [Ueno+, 2021]

- Norm of sound field

$$\|u\|_{\mathscr{H}}^2 = \int_{\boldsymbol{x}\in\mathbb{S}_2} \underbrace{\frac{|\tilde{u}(\boldsymbol{x})|^2}{w(\boldsymbol{x})}}_{\text{directional weighting}} \, \mathrm{d}\chi$$

  - Large $w(\boldsymbol{x})$
    ➡ small $\|u\|_{\mathscr{H}}$ for sound field originating from direction $\boldsymbol{x}$
  - Small $w(\boldsymbol{x})$
    ➡ large $\|u\|_{\mathscr{H}}$ for sound field originating from direction $\boldsymbol{x}$

Sound field estimation exploiting prior information on source direction [Ueno+, 2021]

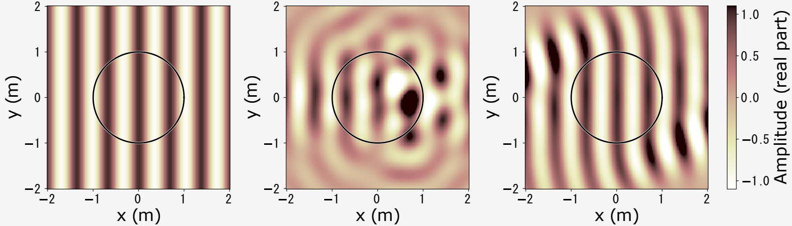
- Directional weighting based on prior information

$$w(\boldsymbol{x}) = \underbrace{\frac{1}{4\pi C(\beta)}}_{\text{normalization constant}} \exp(\beta \underbrace{\boldsymbol{\eta}}_{\text{prior source direction}} \cdot \boldsymbol{x})$$

  - $\boldsymbol{x}$ close to $\boldsymbol{\eta}$ ➡ large $w(\boldsymbol{x})$
  - $\boldsymbol{x}$ away from $\boldsymbol{\eta}$ ➡ small $w(\boldsymbol{x})$

■ Sound field estimation exploiting prior information on source direction [Ueno+, 2021]

▹ Numerical simulation



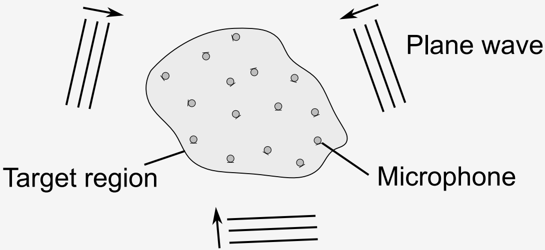
True **w/o prior information** **w/ prior information** ($\boldsymbol{\eta}$ = [cos30°, sin30°, 0]^T)

Performance improvement using prior information

Sparsity-based sound field estimation

[Chardon+, 2012], [Koyama+, 2019]

- Sparsity assumption for sound field
  - Sound field is originated from a small number of basis functions (e.g., plane waves and monopole fields).



Sparsity-based sound field estimation

[Chardon+, 2012], [Koyama+, 2019]

- Decomposition of sound field

$$u(\boldsymbol{r}) \approx \sum_n \underbrace{a_n}_{\text{expansion coefficient}} \underbrace{\psi_n(\boldsymbol{r})}_{\text{basis function}}$$

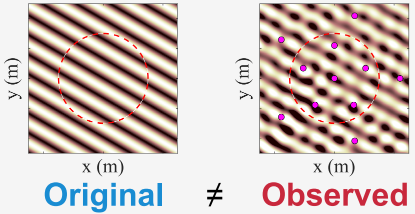
- Formulation

$$\underset{\boldsymbol{a}}{\text{minimize}} \ \mathcal{J}(\boldsymbol{a}) = \|\boldsymbol{C}\boldsymbol{a} - \boldsymbol{s}\|_2^2 + \underbrace{\lambda\|\boldsymbol{a}\|_p^p}_{\text{sparsity-inducing regularization } (p<2)}$$

  - Solved by various iterative algorithms

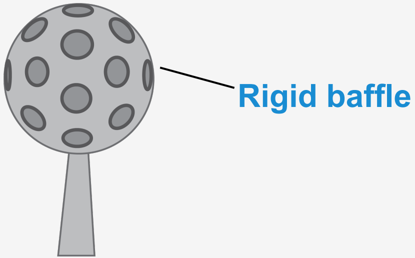
Scattering effect in sound field estimation

- Reflection, absorption, diffraction, etc.
  - Caused by physical existence of microphone array



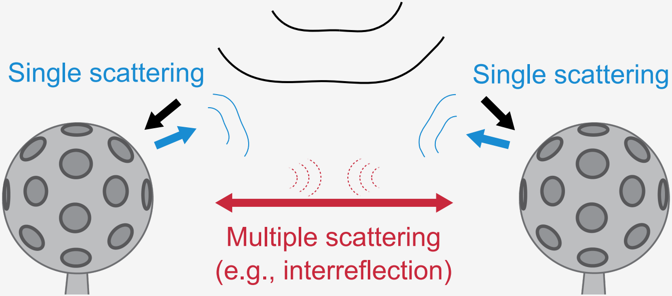
Necessity to compensate scattering effect

- **Microphone array mounted on spherical baffle**
  [Meyer+, 2002], [Abhayapala+, 2002]
  - Based on single scattering problem
  - Applicable for avoiding forbidden frequency problem (also described in Section 2)



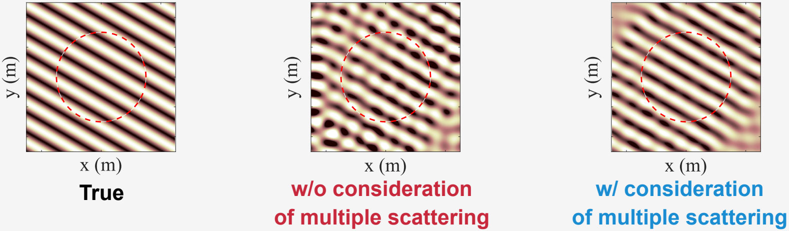
Scattering effect caused by multiple baffles

- Multiple scattering effect
  $\neq$ superposition of each single scattering effect



Sound field estimation considering multiple scattering effect [Nakanishi+, 2019], [Kaneko+, 2021]

- Solving relationship between original and scattered sound fields using translation operator for expansion coefficients [Martin, 2006]



Least-squares-based sound field estimation

- Based on decomposition of sound field
  - Spherical wavefunctions or plane wave functions
- Infinite-dimensional extension
  - No necessity of parameter setting
    (expansion center, truncation order, etc.)
- Other extensions and related works
  - Harmonic analysis for arbitrary position
  - Use of prior information
  - Consideration of multiple scattering effect

- J. Meyer and G. Elko, “A highly scalable spherical microphone array based on an orthogonal decomposition of soundfield,” in Proc. IEEE ICASSP, 2002.
- T. D. Abhayapala and D. B. Ward, “Theory and design of high order sound field microphones using spherical microphone arrays,” in Proc. IEEE ICASSP, 2002.
- A. Laborie, R. Bruno, and S. Montoya, “A new comprehensive approach of surround sound recording,” in Proc. AES Convention, 2003.
- M. A. Poletti, “Three-dimensional surround sound systems based on spherical harmonics,” J. Audio Eng. Soc., 53(11), pp. 1004–1025, 2005.
- P. A. Martin, “Multiple scattering: Interaction of time-harmonic waves with and N obstacles,” Cambridge Univ. Press, 2006.
- G. Chardon, L. Daudet, A. Peillot, et al., “Near-field acoustic holography using sparse regularization and compressive sampling principles,” J. Acoust. Soc. Am., 132, pp. 1521–1534, 2012.
- 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., 22(3), pp. 647–658, 2014.

References (2)

- N. Ueno, S. Koyama, and H. Saruwatari, "Sound field recording using distributed microphones based on harmonic analysis of infinite order," IEEE Signal Process. Lett., 25(1), pp. 135–139, 2018.
- N. Ueno, S. Koyama, and H. Saruwatari, "Kernel ridge regression with constraint of Helmholtz equation for sound field interpolation," in Proc. IWAENC, 2018.
- S. Koyama and L. Daudet, "Sparse representation of a spatial sound field in a reverberant environment," IEEE J. Select. Topics Signal Process., 13(1), pp. 172–184, 2019.
- N. Ueno, S. Koyama, and H. Saruwatari, "Three-dimensional sound field reproduction based on weighted mode-matching method," IEEE/ACM Trans. Audio, Speech, Lang. Process., 27(12), pp. 1852–1867, 2019.
- M. Nakanishi, N. Ueno, S. Koyama, and H. Saruwatari, "Two-dimensional sound field recording with multiple circular microphone arrays considering multiple scattering," in Proc. IEEE WASPAA, 2019.
- N. Ueno, S. Koyama, and H. Saruwatari, "Directionally weighted sound field estimation exploiting prior information on source direction," IEEE Trans. Signal Process., 69, pp. 2383–2395, 2021.

References (3)

- N. Iijima, S. Koyama, and H. Saruwatari, “Binaural rendering from microphone array signals of arbitrary geometry,” J. Acoust. Soc. Am., 150(4), pp. 2479–2491, 2021.
- S. Kaneko and R. Duraiswami, “Multiple scattering ambisonics: Three-dimensional sound field estimation using interacting spheres,” JASA Express Lett., 1(8), 2021.
- T. Nishida, N. Ueno, S. Koyama, and H. Saruwatari, “Region-restricted sensor placement based on Gaussian process for sound field estimation,” IEEE Trans. Signal Process., 70, pp. 1718–1733, 2022.

[T-5] Sound Field Estimation: Recent Advances and Applications – Section 4 –

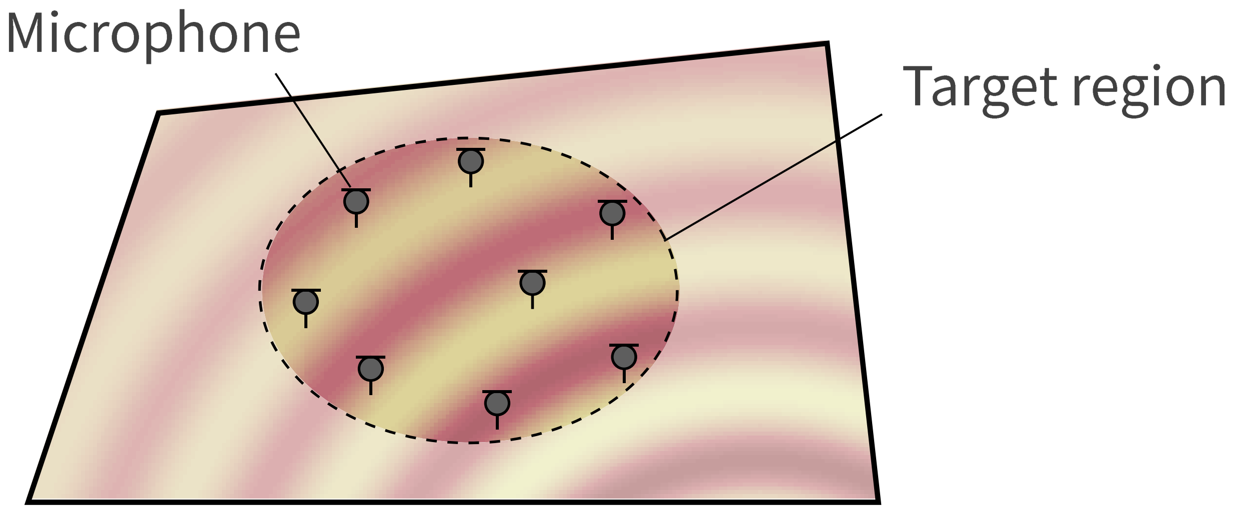
Shoichi Koyama[1], Natsuki Ueno[2]
[1] The University of Tokyo
[2]Tokyo Metropolitan University

22-27 MAY 2022
ICASSP 2022
SINGAPORE

Table of contents

1. Preliminaries on acoustics
    1. Wave equation and Helmholtz equation
    2. Representations of acoustic field
2. Integral-equation-based sound field estimation
    1. Problem formulation
    2. Sound field estimation with spherical array
    3. Methods to avoid forbidden frequency problem
3. Least-squares-based sound field estimation
    1. Basic framework
    2. Infinite-dimensional extension
    3. Other extensions and related works
4. Applications
    1. Spatial audio reproduction by headphones
    2. Spatial audio reproduction by loudspeakers
    3. Spatial active noise control

Sound field estimation

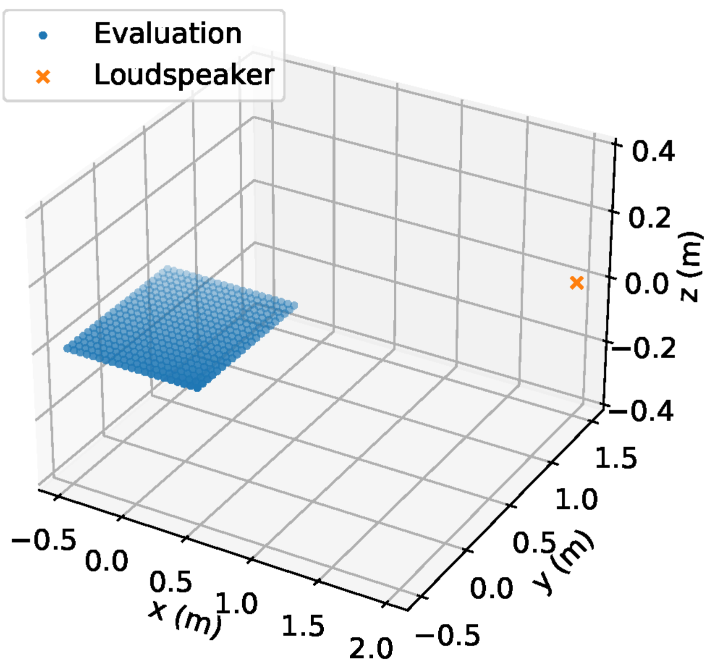


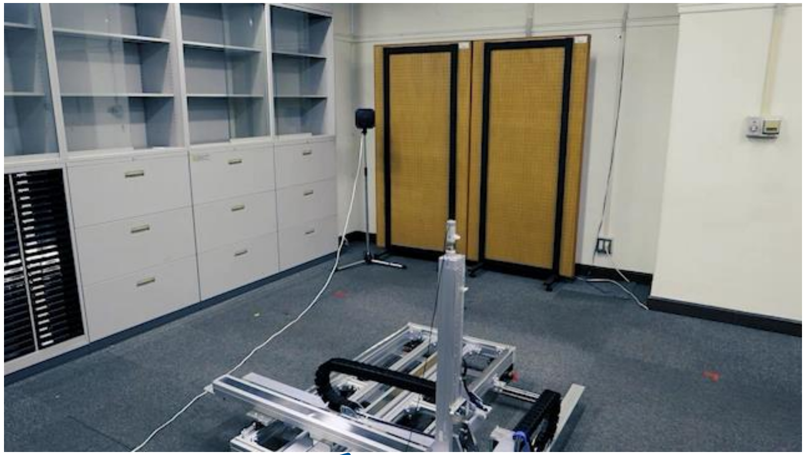
Estimating sound field inside target region using observations of multiple mics

Kernel interpolation of sound field can be applied to estimate pressure distribution (P1) by distributed mics

Demo: Kernel interpolation of sound field

- Evaluation in real environment
  - Impulse responses from single loudspeaker to 441 evaluation points on plane are taken from MeshRIR dataset [Koyama+ 2021]
  - 18 mics are selected from evaluation points by sensor placement metod proposed in [Nishida+ 2022]



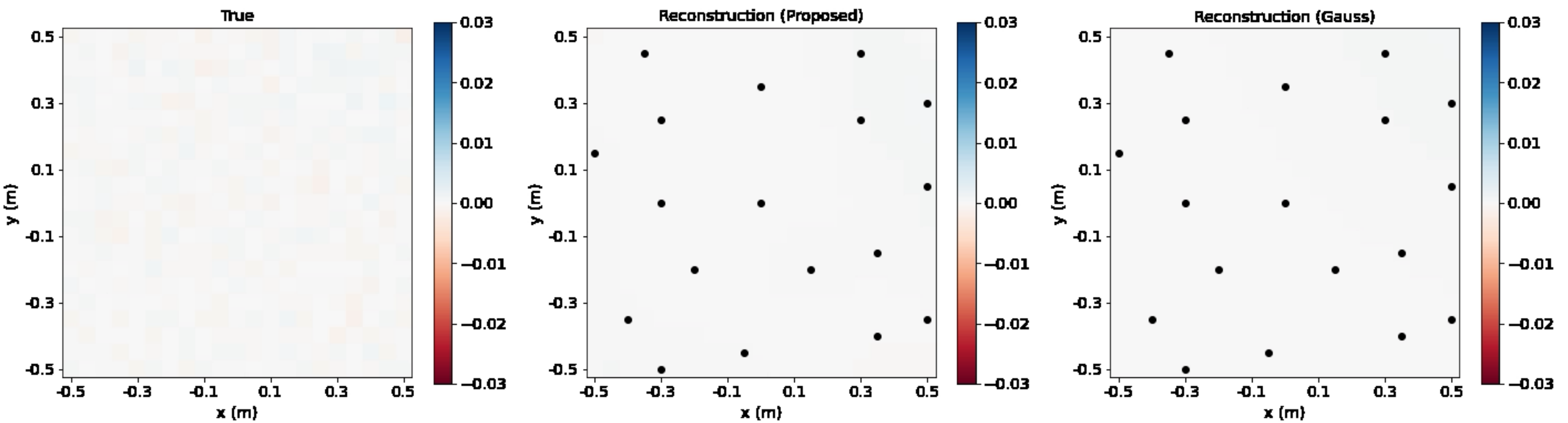


Impulse response measurement system

Demo: Kernel interpolation of sound field

- Estimated pressure distribution
  - Source signal is lowpass-filtered pulse signal < 500Hz
  - Compared with the method using Gaussian kernel

True | Proposed | Gaussian kernel



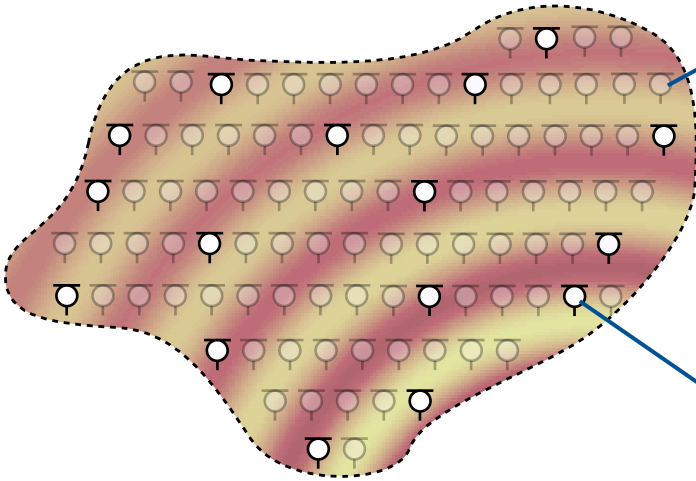
(Black dots indicate mic positions)

High estimation accuracy is achieved by constraint of Helmholtz eq

Sensor placement in sound field estimation

How to optimize sensor placement in sound field estimation?

- Sensor placement problem
  - Discretize target region and set candidate locations
  - Select optimal placement from candidates



$$\underset{\mathcal{S} \subset \mathcal{C}}{\text{minimize}} \quad J(\mathcal{S}) \qquad \text{or} \qquad \underset{\mathcal{S} \subset \mathcal{C}}{\text{maximize}} \quad J(\mathcal{S})$$

Sensor placement in sound field estimation

How to optimize sensor placement in sound field estimation?

- Sensor placement problem
  - Discretize target region and set candidate locations
  - Select optimal placement from candidates

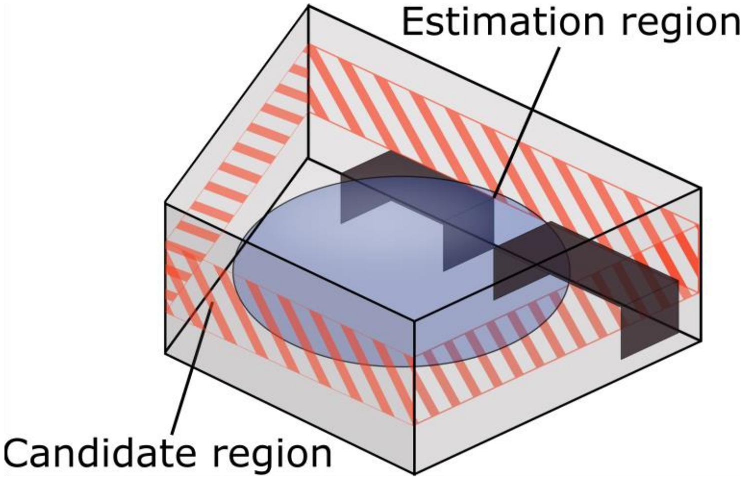
➡ Many sensor placement methods have been investigated in the context of machine learning and sensor network

- Optimization criteria
  - Measures on sensing matrix used in experimental design, e.g., sum of eigenvalues and log determinant
  - Information-theoretic measure, e.g., entropy and mutual information
- Algorithm
  - Greedy method, convex relaxation, heuristics

[Koyama+ 2020]

Sensor placement in sound field estimation

- Our approach in [Nishida+ 2022]
  - Modeling based on Gaussian process to employ kernel function for sound field interpolation
  - Optimization criteria based on expected squared error of estimation
  - Candidate and estimation regions can individually be set
  - Greedy algorithm and mirror descent algorithm for solving it



Useful to obtain optimal sensor placement for sound field estimation

Application of sound field estimation

VR/AR audio

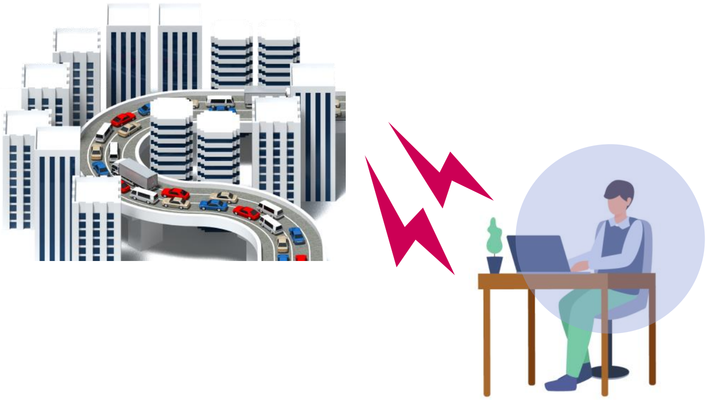
Room acoustic analysis

Signal enhancement

Basic Technologies of
Sound Field
Estimation

Local-field recording
and reproduction

Active noise control



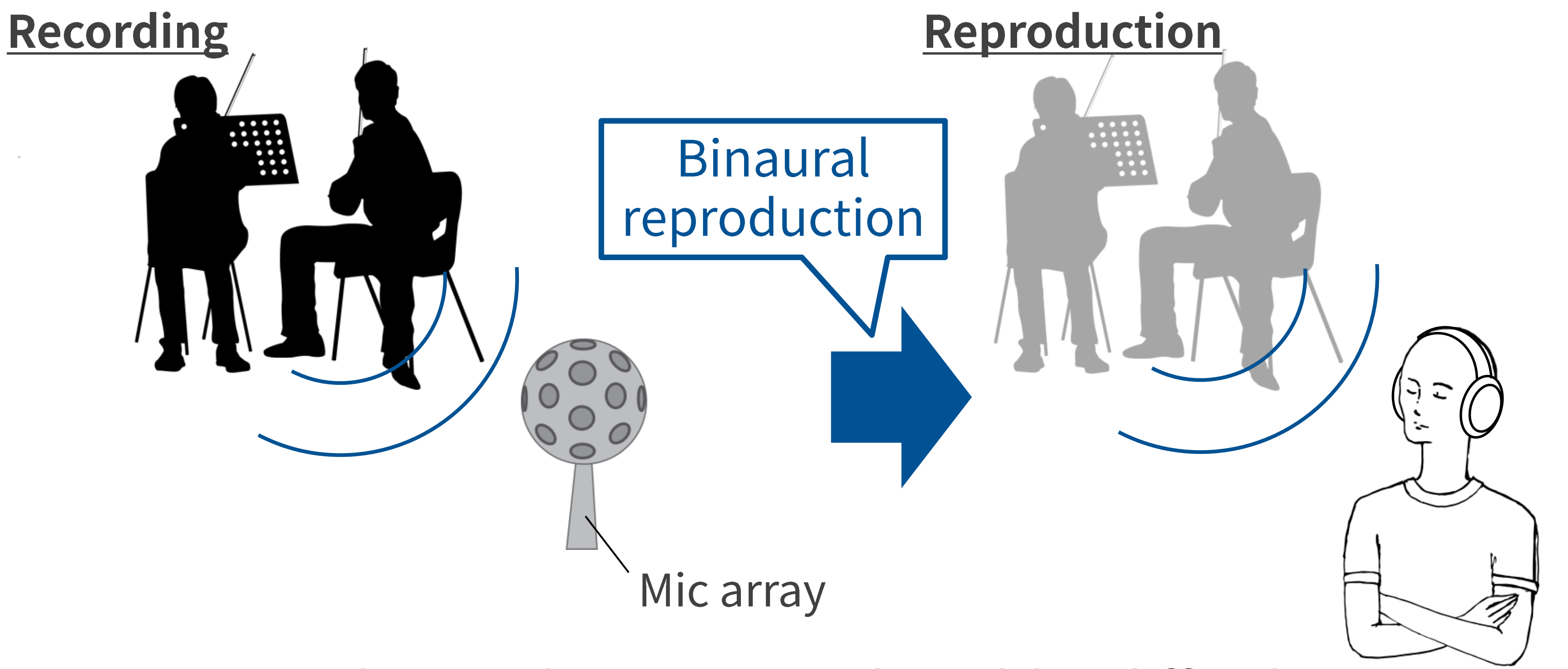
Visualization/auralization

Table of contents

1. Preliminaries on acoustics
    1. Wave equation and Helmholtz equation
    2. Representations of acoustic field
2. Integral-equation-based sound field estimation
    1. Problem formulation
    2. Sound field estimation with spherical array
    3. Methods to avoid forbidden frequency problem
3. Least-squares-based sound field estimation
    1. Basic framework
    2. Infinite-dimensional extension
    3. Other extensions and related works
4. Applications
    1. Spatial audio reproduction by headphones
    2. Spatial audio reproduction by loudspeakers
    3. Spatial active noise control

Binaural reproduction

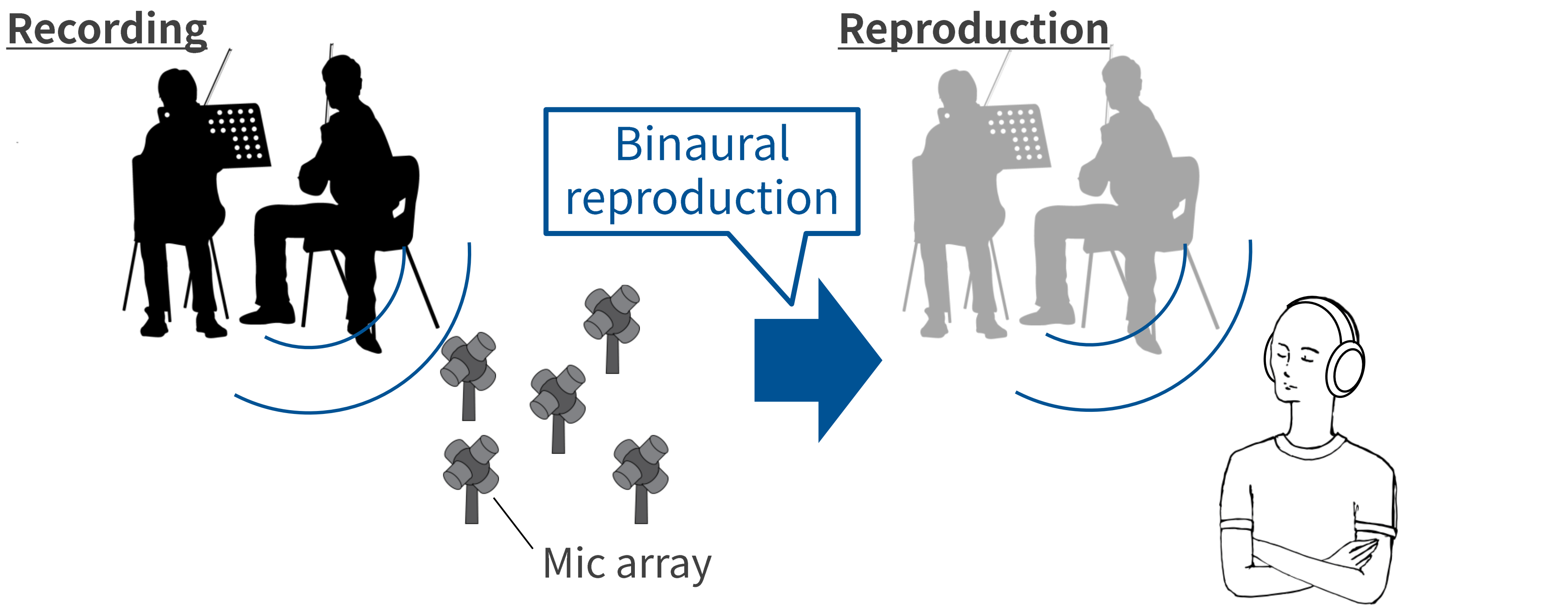
Binaural reproduction from mic array recordings for VR audio



- Binaural reproduction in real world is difficult, compared to binaural synthesis in VR space
- Conventional spherical array processing requires large-scale system to achieve broad listening area

Binaural reproduction

Binaural reproduction from mic array recordings for VR audio

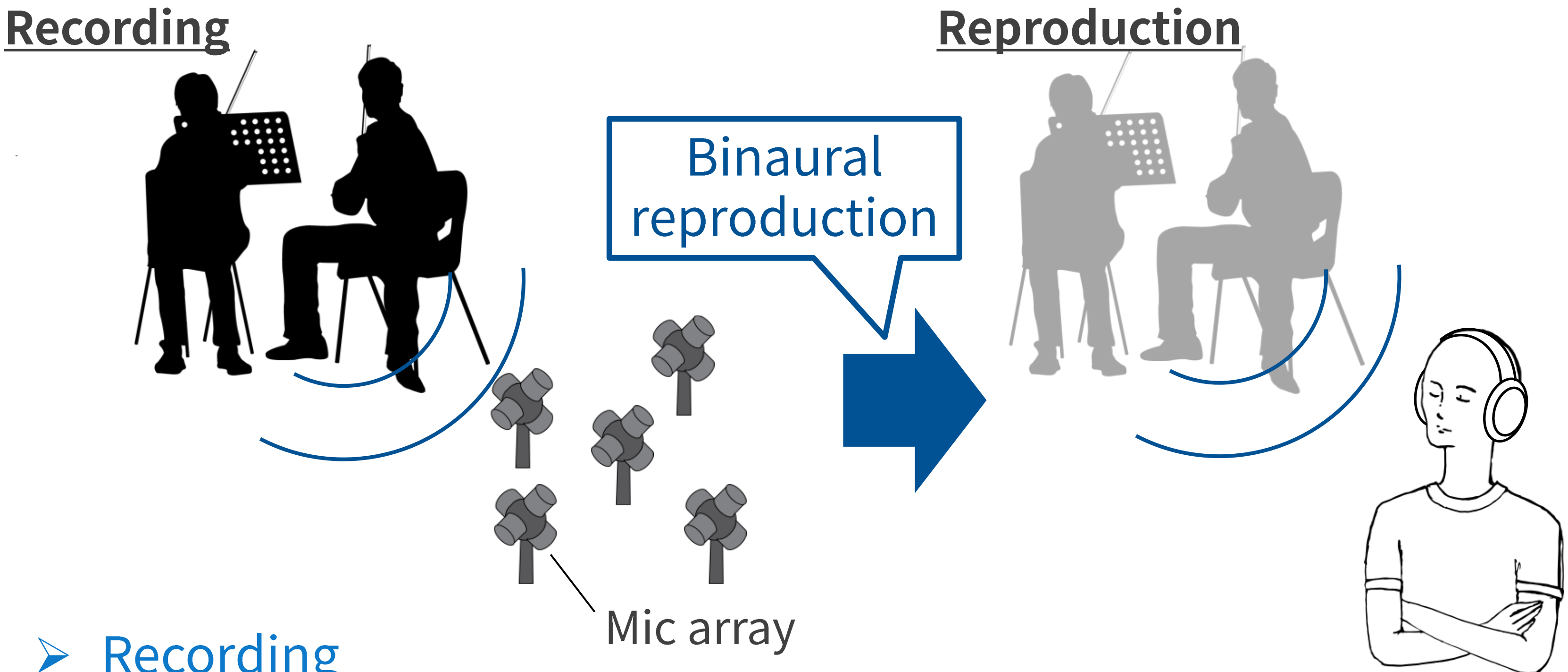


- Binaural reproduction from recordings of multiple small arrays instead of single spherical array
- Broad listening area can be achived by using flexible and scalable recording system

[Iijima+ JASA 2021]

Binaural reproduction

Binaural reproduction from mic array recordings for VR audio



- Recording
  - Sound field estimation based on infinite-dimensional harmonic analysis to estimate expansion coeffs (P3)
- Reproduction
  - Binaural rendering from estimated expansion coeffs with compensation for loudspeaker distance in HRTF measurements

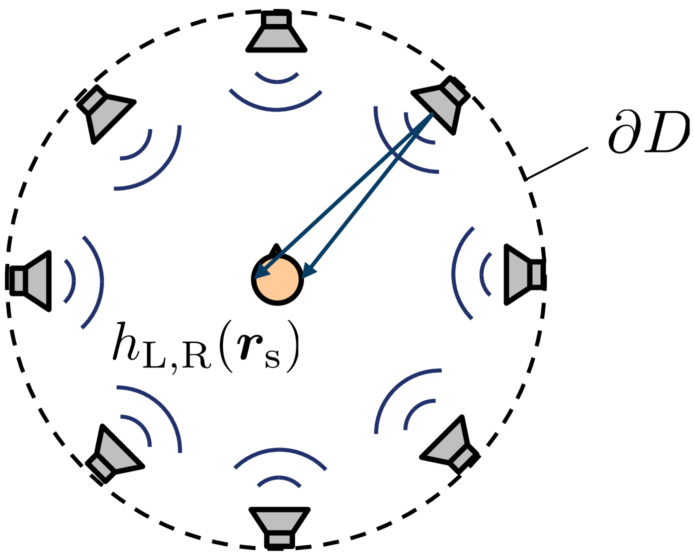
Binaural rendering from estimated expansion coeffs

- HRTFs $h_{\mathrm{L,R}}(\boldsymbol{r}_{\mathrm{s}})$ are assumed to be transfer functions from point sources on spherical surface ∂D to two ears
- Sound field is represented by weigted integral of point sources on ∂D (single layer potential)

Free-field Green's func

$$u(\boldsymbol{r}) = \int_{\partial D} w(\boldsymbol{r}_{\mathrm{s}}) G(\boldsymbol{r} - \boldsymbol{r}_{\mathrm{s}}) \mathrm{d}\boldsymbol{r}_{\mathrm{s}} \qquad (\boldsymbol{r}_{\mathrm{s}} = (R_{\mathrm{s}}, \theta_{\mathrm{s}}, \phi_{\mathrm{s}}) \in \partial D)$$

Weight



Binaural rendering from estimated expansion coeffs

- Weight $w(\boldsymbol{r}_\mathrm{s})$ can be related to expansion coeffs $\mathring{u}_{\nu,\mu}(\boldsymbol{r})$ by using

$$G(\boldsymbol{r}-\boldsymbol{r}_\mathrm{s}) = \frac{\mathrm{e}^{\mathrm{j}k\|\boldsymbol{r}-\boldsymbol{r}_\mathrm{s}\|}}{4\pi\|\boldsymbol{r}-\boldsymbol{r}_\mathrm{s}\|}$$

$$= \mathrm{j}k \sum_{\nu=0}^{\infty} j_\nu(kr) h_\nu(kR_\mathrm{s}) \sum_{\mu=-\nu}^{\nu} Y_{\nu,\mu}(\theta,\phi) Y_{\nu,\mu}(\theta_\mathrm{s},\phi_\mathrm{s})^*$$

and orthogonality of spherical harmonic functions as

$$w(\boldsymbol{r}_\mathrm{s}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \frac{\sqrt{4\pi}\mathrm{j}}{kR_\mathrm{s}^2 h_\nu(kR_\mathrm{s})} \mathring{u}_{\nu,\mu}(\boldsymbol{r}) Y_{\nu,\mu}(\theta_\mathrm{s},\phi_\mathrm{s})$$

Binaural reproduction

Binaural rendering from estimated expansion coeffs

- Spherical harmonic coeffs of HRTFs $h_{\mathrm{L,R}}(\boldsymbol{r}_{\mathrm{s}})$:

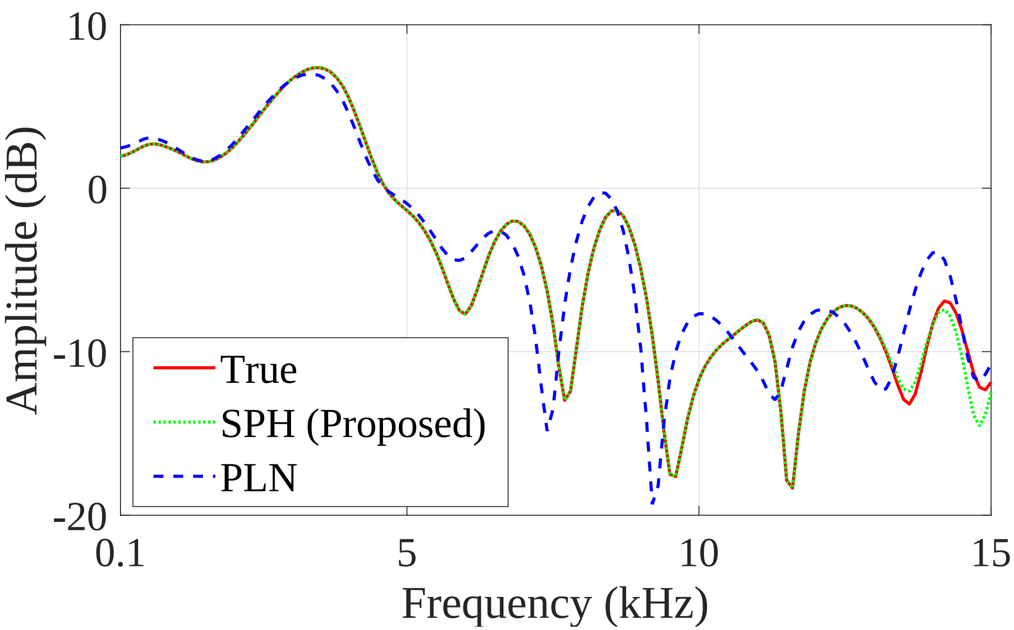
$$h_{\mathrm{L,R}}(\boldsymbol{r}_{\mathrm{s}}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} H_{\mathrm{L,R},\nu,\mu} Y_{\nu,\mu}(\theta_{\mathrm{s}}, \phi_{\mathrm{s}})^*$$

- Binaural signals are obtained as

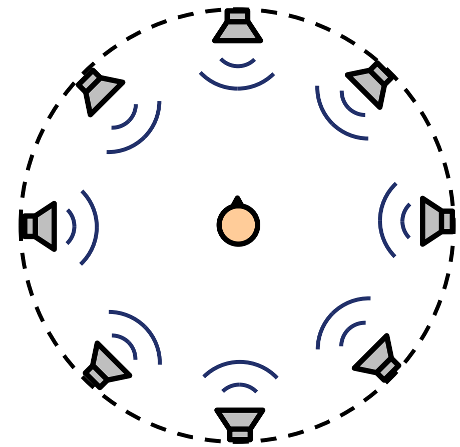
$$\begin{aligned} y_{\mathrm{L,R}}(\boldsymbol{r}) &= \int_{\partial D} w(\boldsymbol{r}_{\mathrm{s}}) h_{\mathrm{L,R}}(\boldsymbol{r}_{\mathrm{s}}) \mathrm{d}\boldsymbol{r}_{\mathrm{s}} \\ &= \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \frac{\sqrt{4\pi}\mathrm{j}}{k h_{\nu}(k R_{\mathrm{s}})} \mathring{u}_{\nu,\mu}(\boldsymbol{r}) H_{\mathrm{L,R},\nu,\mu} \end{aligned}$$

Binaural reproduction

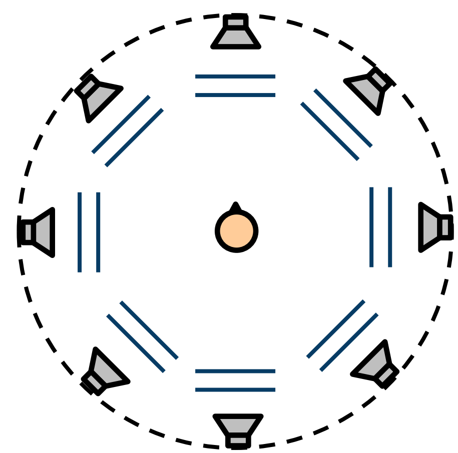
- Comparison of rendering methods



Spherical wave expansion (SPH)



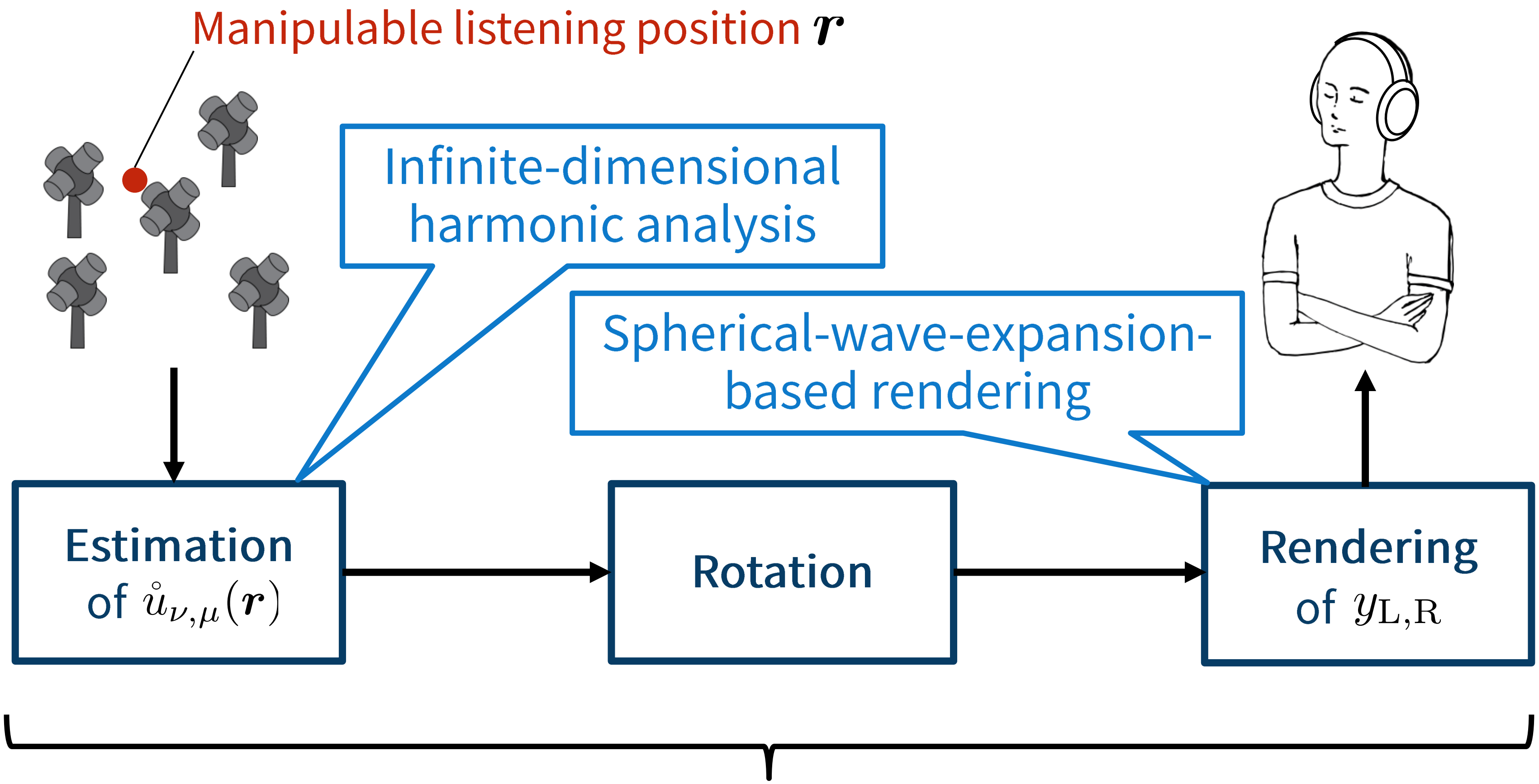
Plane wave expansion (PLN)



Peaks and dips are accurately reproduced by spherical wave expansion

Binaural reproduction

➢ Summary of proposed binaural reproduction method



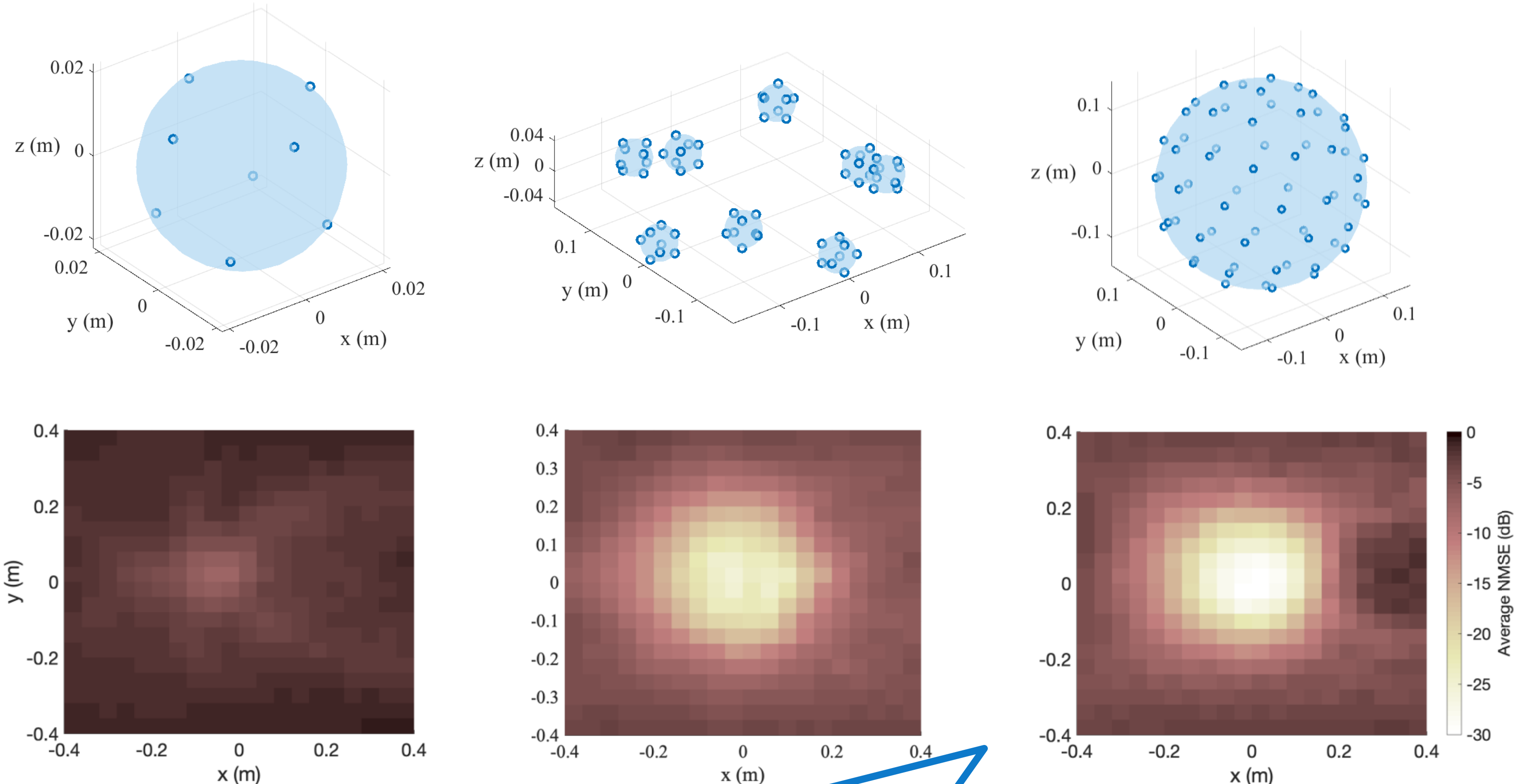
Conversion is achieved by single MIMO FIR filter because all the operations are linear

- Comparison of array geometry

Single array | **Composite array** | **Spherical array**

Array geometry

Reproduction error on xy-plane



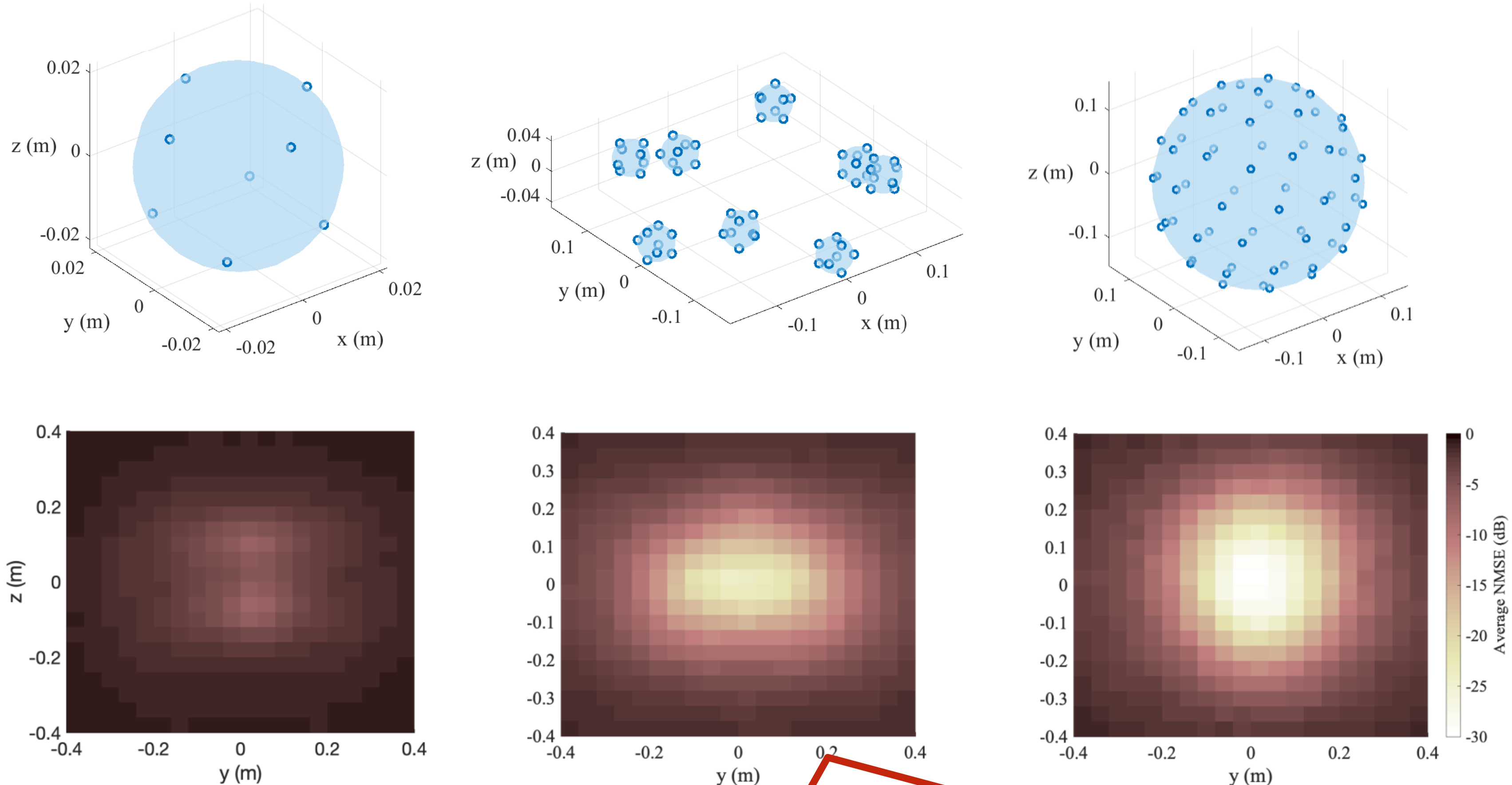
Almost the same binaural reproduction accuracy

- Comparison of array geometry

Single array | Composite array | Spherical array

Array geometry

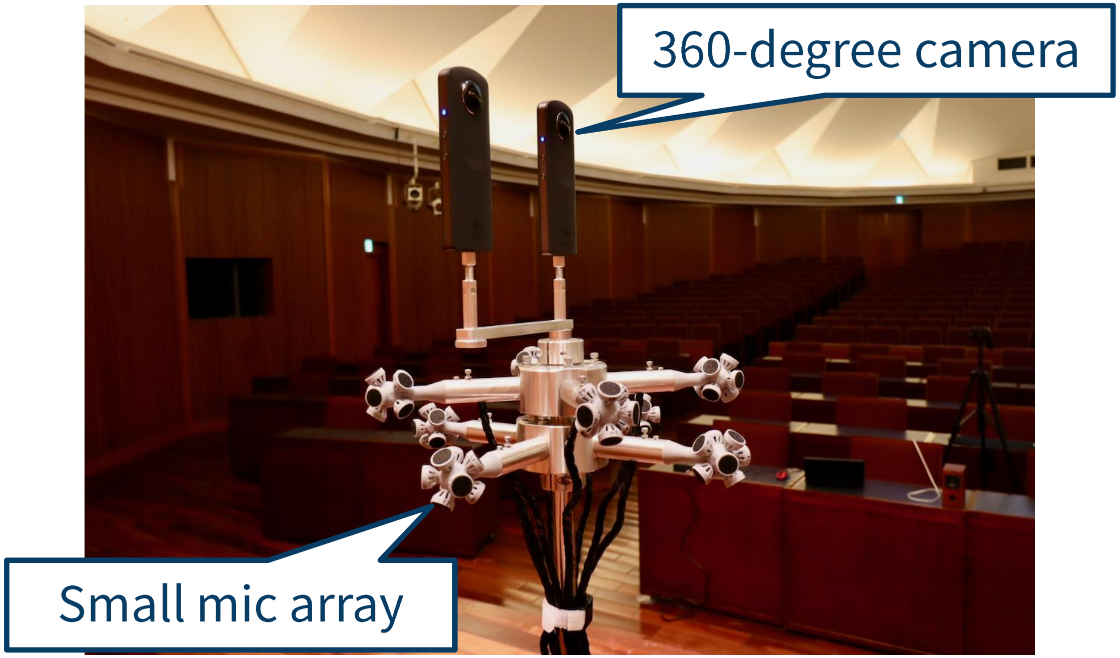
Reproduction error on yz-plane



High reproduction accuracy region is slightly shrunk

Binaural reproduction

- Developed system using composite mic array
  - Composed of 8 small mic arrays
  - 8 unidirectional mics in each array (same placement as 2nd-order Ambisonic mic)
  - 360-degree camera for capturing images

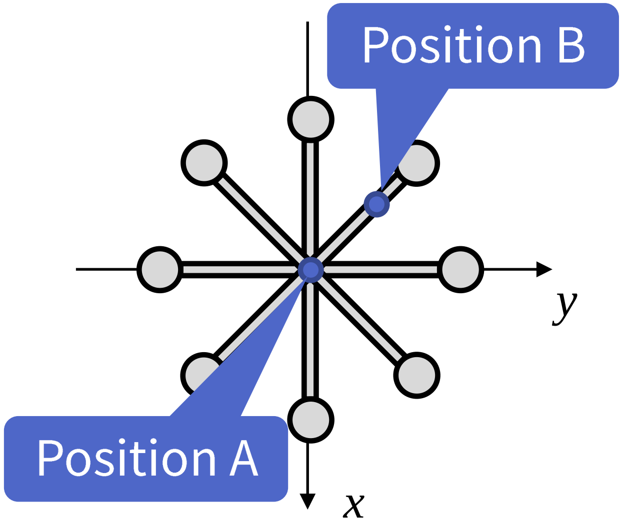


Binaural reproduction

➢ Listening test using MUSHRA

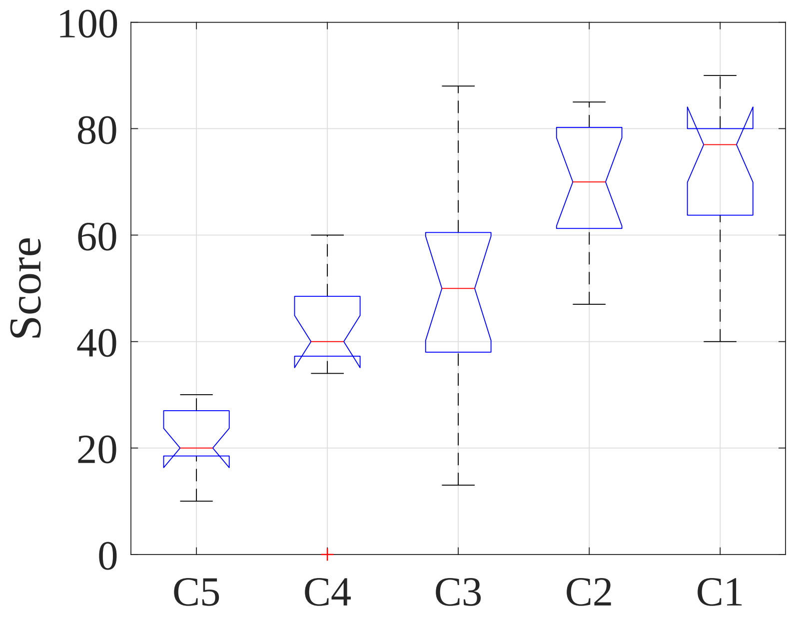
Reference	Dummy head recording
C1	Reproduced by composite array at position A
C2	Reproduced by composite array at position B
C3	Reproduced by single array at position A
C4	Reproduced by single array at position B
C5/anchor	Lowpass-filtered reference (< 1.6 kHz)

Composite array

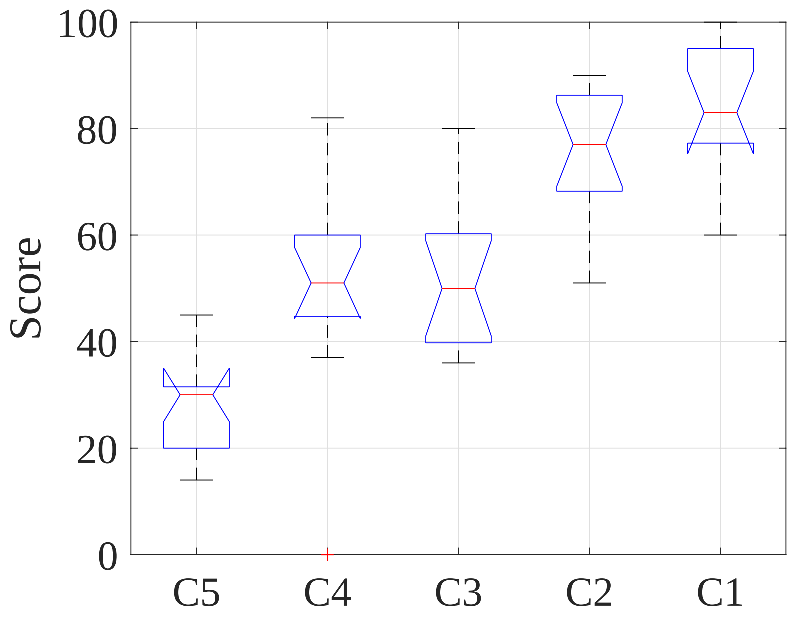


Position A: (0.0, 0.0, 0.0) m
Position B: (-0.071, 0.071, 0.0) m

- Listening test scores



Music



Male speech

- Significant difference between composite array (C1 and C2) and single array (C3 and C4)
- No significant difference between position A and B for composite array (C1 and C2)

Binaural reproduction

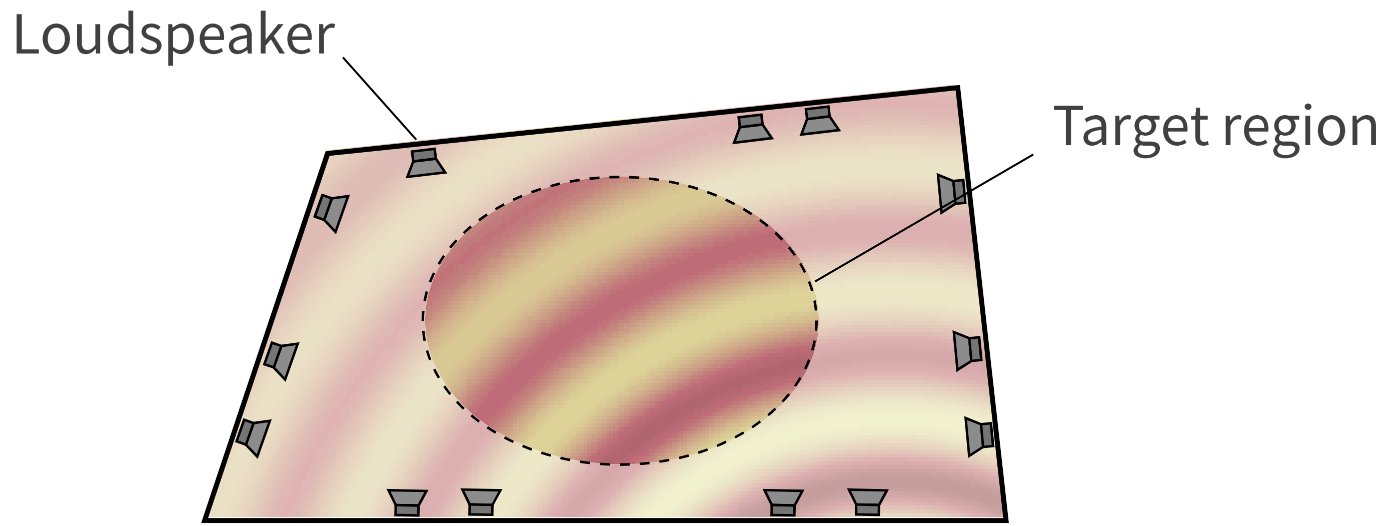
- Demo
  - Recorded classical music (String Quartet)
  - Controllable on Web browser / HMD
  - 360-degree control of viewing and listening directions
  - 2 viewing and listening positions are selectable

https://youtu.be/qhdf18CvFD8

Table of contents

1. Preliminaries on acoustics
   1. Wave equation and Helmholtz equation
   2. Representations of acoustic field
2. Integral-equation-based sound field estimation
   1. Problem formulation
   2. Sound field estimation with spherical array
   3. Methods to avoid forbidden frequency problem
3. Least-squares-based sound field estimation
   1. Basic framework
   2. Infinite-dimensional extension
   3. Other extensions and related works
4. Applications
   1. Spatial audio reproduction by headphones
   2. Spatial audio reproduction by loudspeakers
   3. Spatial active noise control

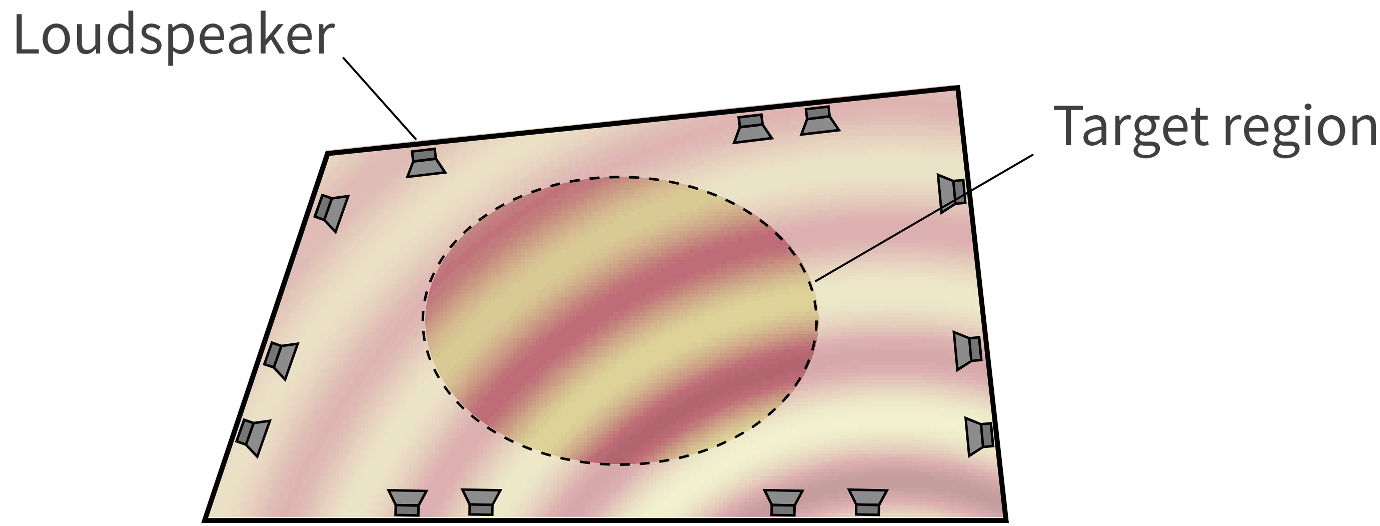
Sound field reproduction



Synthesizing desired sound field inside target region using multiple loudspeakers

Sound field reproduction methods rely on estimation methods when incorporating knowledge on captured sound field

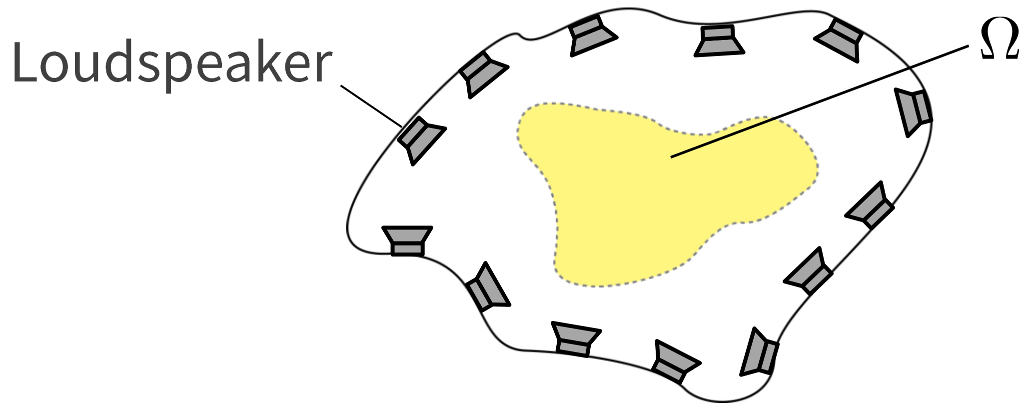
Sound field reproduction



- Two major categories of sound field reproduction:
  - Analytical approach based on integral eq:
    - Fast and stable computation, but array geometry must be simple
  - Numerical approach based on minimization of square error:
    - Flexible array geometry, but computational cost is relatively high

Problem formulation

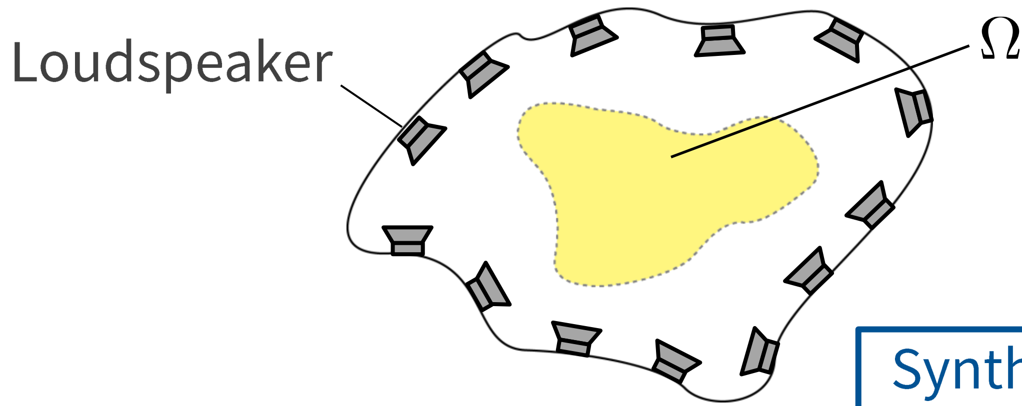
- Optimization problem for sound field synthesis
  - Synthesizing sound field inside Ω with L loudspeakers



  - Notation:
    - d_l : Driving signal of l th loudspeaker
      ➡ Vector form $\boldsymbol{d} = [d_1, \ldots, d_L]^{\mathsf{T}} \in \mathbb{C}^L$
    - $g_l(\boldsymbol{r})$: Transfer function of lth loudspeaker
      ➡ Vector form $\boldsymbol{g}(\boldsymbol{r}) = [g_1(\boldsymbol{r}), \ldots, g_L(\boldsymbol{r})]^{\mathsf{T}} \in \mathbb{C}^L$
    - $u_{\mathrm{des}}(\boldsymbol{r})$: Desired sound field inside Ω

Problem formulation

- Optimization problem for sound field synthesis
  - Synthesizing sound field inside Ω with L loudspeakers



Synthesized sound field

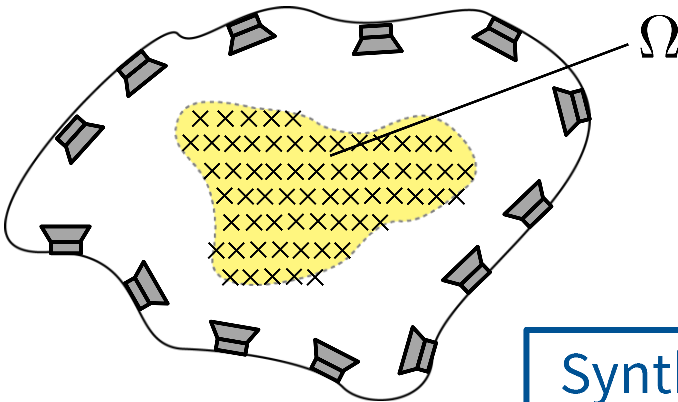
$$\underset{\boldsymbol{d}\in\mathbb{C}^L}{\text{minimize}}\, J := \int_{\boldsymbol{r}\in\Omega} \left| \sum_{l=1}^{L} g_l(\boldsymbol{r}) d_l - u_{\mathrm{des}}(\boldsymbol{r}) \right|^2 \mathrm{d}\boldsymbol{r}$$

$$= \int_{\boldsymbol{r}\in\Omega} \left| \boldsymbol{g}(\boldsymbol{r})^{\mathsf{T}} \boldsymbol{d} - u_{\mathrm{des}}(\boldsymbol{r}) \right|^2 \mathrm{d}\boldsymbol{r}$$

➡ Determine driving signals that minimizes square error between synthesized and desired fields inside target region

Pressure matching

- Discretize target region Ω, and obtain driving signal to control sound pressures at M control points



Synthesized pressure at control points

$$\underset{\boldsymbol{d}\in\mathbb{C}^L}{\text{minimize}} \left\| \boldsymbol{G}\boldsymbol{d} - \boldsymbol{u}^{\text{des}} \right\|^2 + \lambda \|\boldsymbol{d}\|^2$$

- Transfer function matrix: $\boldsymbol{G} = [\boldsymbol{g}(\boldsymbol{r}_1), \ldots, \boldsymbol{g}(\boldsymbol{r}_M)]^{\mathsf{T}} \in \mathbb{C}^{M\times L}$
- Desired pressure vector: $\boldsymbol{u}^{\text{des}} = [u_{\text{des}}(\boldsymbol{r}_1), \ldots, u_{\text{des}}(\boldsymbol{r}_M)]^{\mathsf{T}} \in \mathbb{C}^{M}$

➡ **Closed-form solution:** $\boldsymbol{d} = (\boldsymbol{G}^{\mathsf{H}}\boldsymbol{G} + \lambda\boldsymbol{I})^{-1}\boldsymbol{G}^{\mathsf{H}}\boldsymbol{u}^{\text{des}}$

Weighted mode matching

- In weighted mode matching [Ueno+ 2019], objective function J is approximated by spherical wavefunction expansion
- Weighted mode matching is a generalization of (standard) mode matching
- Spherical wavefunction expansion of $u_{\mathrm{des}}(\boldsymbol{r})$ and $\boldsymbol{g}(\boldsymbol{r})^{\mathsf{T}}$ around $\boldsymbol{r}_{\mathrm{o}}$ up to order N_{tr}

$$u_{\mathrm{des}}(\boldsymbol{r}) \approx \sum_{\nu=0}^{N_{\mathrm{tr}}} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\mathrm{des},\nu,\mu}(\boldsymbol{r}_{\mathrm{o}}) \varphi_{\nu,\mu}(\boldsymbol{r} - \boldsymbol{r}_{\mathrm{o}})$$

$$g_l(\boldsymbol{r}) \approx \sum_{\nu=0}^{N_{\mathrm{tr}}} \sum_{\mu=-\nu}^{\nu} \mathring{g}_{l,\nu,\mu}(\boldsymbol{r}_{\mathrm{o}}) \varphi_{\nu,\mu}(\boldsymbol{r} - \boldsymbol{r}_{\mathrm{o}})$$

where

$$\varphi_{\nu,\mu}(\boldsymbol{r}) = \sqrt{4\pi} j_\nu(k\|\boldsymbol{r}\|) Y_{\nu,\mu}(\boldsymbol{r}/\|\boldsymbol{r}\|)$$

Weighted mode matching

- Matrix form of expansion:

$$u_{\mathrm{des}}(\boldsymbol{r}) \approx \bar{\boldsymbol{\varphi}}(\boldsymbol{r} - \boldsymbol{r}_{\mathrm{o}})^{\mathsf{T}} \boldsymbol{b}(\boldsymbol{r}_{\mathrm{o}})$$

$$\boldsymbol{g}(\boldsymbol{r})^{\mathsf{T}} \approx \bar{\boldsymbol{\varphi}}(\boldsymbol{r} - \boldsymbol{r}_{\mathrm{o}})^{\mathsf{T}} \boldsymbol{C}(\boldsymbol{r}_{\mathrm{o}})$$

  - $\boldsymbol{b}(\boldsymbol{r}_{\mathrm{o}}) \in \mathbb{C}^{(N_{\mathrm{tr}}+1)^2}$: Vector of $\mathring{u}_{\mathrm{des},\nu,\mu}(\boldsymbol{r}_{\mathrm{o}})$
  - $\boldsymbol{C}(\boldsymbol{r}_{\mathrm{o}}) \in \mathbb{C}^{(N_{\mathrm{tr}}+1)^2 \times L}$: Matrix of $\mathring{g}_{l,\nu,\mu}(\boldsymbol{r}_{\mathrm{o}})$
  - $\bar{\boldsymbol{\varphi}}(\boldsymbol{r} - \boldsymbol{r}_{\mathrm{o}}) \in \mathbb{C}^{(N_{\mathrm{tr}}+1)^2}$: Vector of $\varphi_{\nu,\mu}(\boldsymbol{r} - \boldsymbol{r}_{\mathrm{o}})$

Weighted mode matching

- Objective function J is approximated as

$$J = \int_{\boldsymbol{r}\in\Omega} \left| \boldsymbol{g}^{\mathsf{T}}(\boldsymbol{r})\boldsymbol{d} - u_{\mathrm{des}}(\boldsymbol{r}) \right|^2 \mathrm{d}\boldsymbol{r}$$

$$\approx \int_{\boldsymbol{r}\in\Omega} \left| \bar{\varphi}(\boldsymbol{r}-\boldsymbol{r}_{\mathrm{o}})^{\mathsf{T}} \left(\boldsymbol{C}(\boldsymbol{r}_{\mathrm{o}})\boldsymbol{d} - \boldsymbol{b}(\boldsymbol{r}_{\mathrm{o}})\right) \right|^2 \mathrm{d}\boldsymbol{r}$$

$$= (\boldsymbol{C}\boldsymbol{d} - \boldsymbol{b})^{\mathsf{H}} \boldsymbol{W} (\boldsymbol{C}\boldsymbol{d} - \boldsymbol{b})$$

where each element of $\boldsymbol{W} \in \mathbb{C}^{(N_{\mathrm{tr}}+1)^2 \times (N_{\mathrm{tr}}+1)^2}$ is obtained as

$$(\boldsymbol{W})_{i,j} = \int_{\boldsymbol{r}\in\Omega} \varphi_{\nu_i,\mu_i}(\boldsymbol{r})^* \varphi_{\nu_j,\mu_j}(\boldsymbol{r}) \mathrm{d}\boldsymbol{r}$$

➡ Weighting factor determined by setting target region Ω

Weighted mode matching

- Optimization problem of weighted mode matching

$$\underset{\boldsymbol{d}\in\mathbb{C}^L}{\text{minimize}}\,(\boldsymbol{C}\boldsymbol{d}-\boldsymbol{b})^{\mathsf{H}}\boldsymbol{W}(\boldsymbol{C}\boldsymbol{d}-\boldsymbol{b})+\lambda\|\boldsymbol{d}\|^2$$

$$\Rightarrow\quad \hat{\boldsymbol{d}}=\left(\boldsymbol{C}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{C}+\lambda\boldsymbol{I}\right)^{-1}\boldsymbol{C}^{\mathsf{H}}\boldsymbol{W}\boldsymbol{b}$$

Weighting factor for each expansion coef is determined by $\boldsymbol{W}$

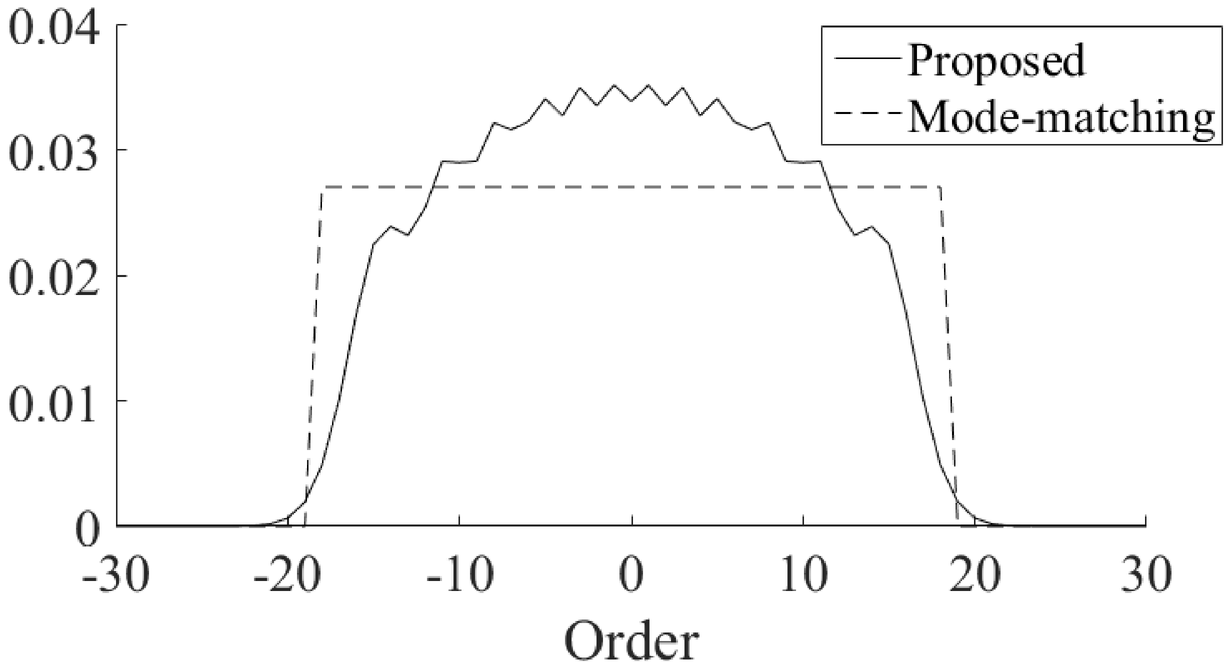
- When $\boldsymbol{W}=\boldsymbol{I}$, weighted mode matching corresponds to (standard) mode matching

$$\boldsymbol{d}=\left(\boldsymbol{C}^{\mathsf{H}}\boldsymbol{C}+\lambda\boldsymbol{I}\right)^{-1}\boldsymbol{C}^{\mathsf{H}}\boldsymbol{b}$$

(Standard) mode matching is sensitive to the setting of truncation order

Weighted mode matching

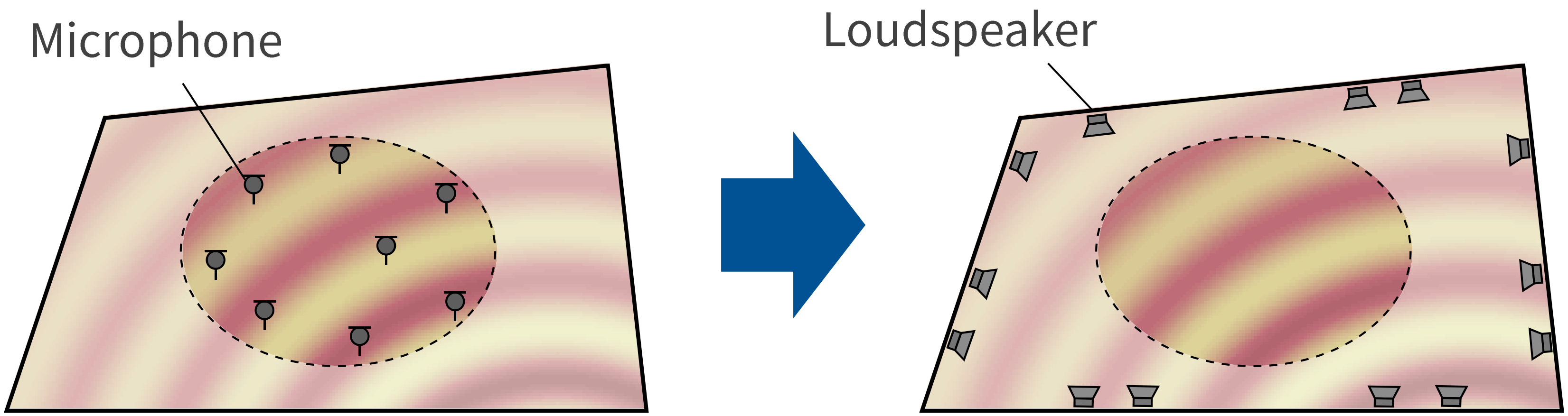
➢ Weighting factor on expansion coeffs for 2D circular target region (0.4 m of radius and k=36.9 rad/m)



Optimal weighting factor is determined by geometry of target region

Sound field estimation for reproduction

- Sound field estimation methods for (P3) can be applied to
  - use transfer functions of loudspeakers measured by mics, i.e., to estimate $\boldsymbol{C}$
  - reproduce sound field captured by mics, i.e., to estimate $\boldsymbol{b}$

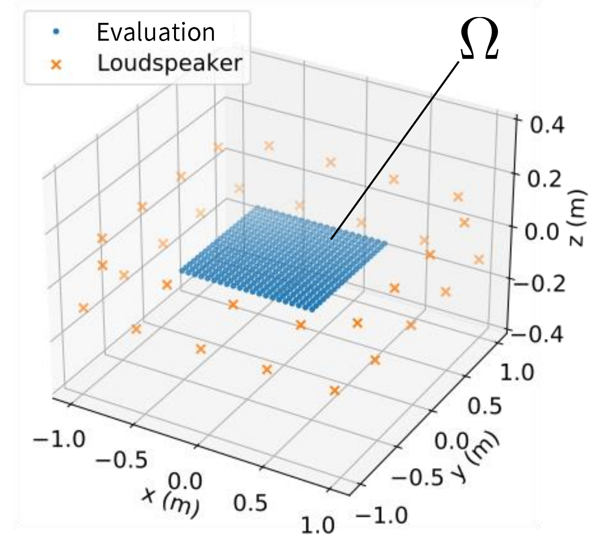


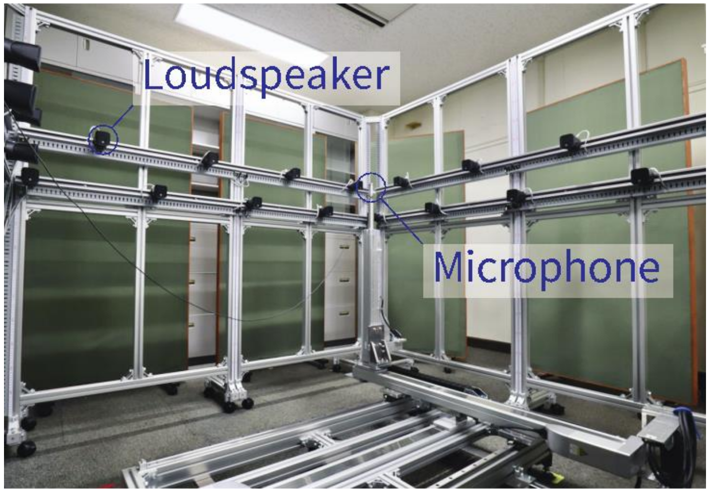
Experiments

[Koyama+ I3DA 2021]

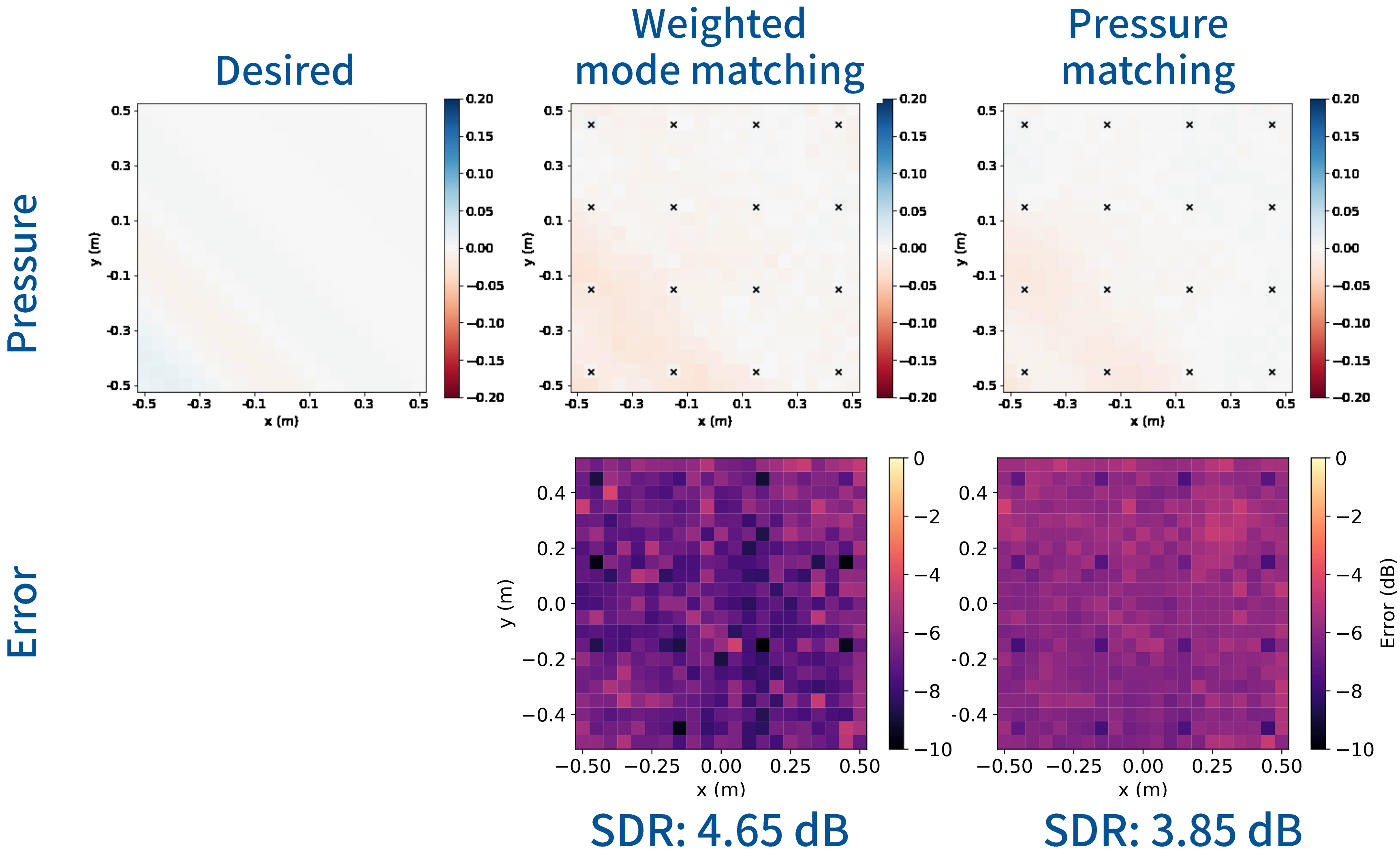
- Synthesizing single planewave field
  - Ω : Square region of 1.0 m x 1.0 m
  - # of loudspeakers: 32
  - Uniform distribution of mics on Ω
  - # of evaluation points: 21 x 21 = 441
  - Compared pressure matching and weighted mode matching
  - Infinite-dimensional harmonic analysis is applied to estimate $\boldsymbol{C}$
  - Evaluation measure:

$$\mathrm{SDR} = \frac{\iint |u_{\mathrm{des}}(\boldsymbol{r},t)|^2 \mathrm{d}\boldsymbol{r}\mathrm{d}t}{\iint |u_{\mathrm{syn}}(\boldsymbol{r},t) - u_{\mathrm{des}}(\boldsymbol{r},t)|^2 \mathrm{d}\boldsymbol{r}\mathrm{d}t}$$

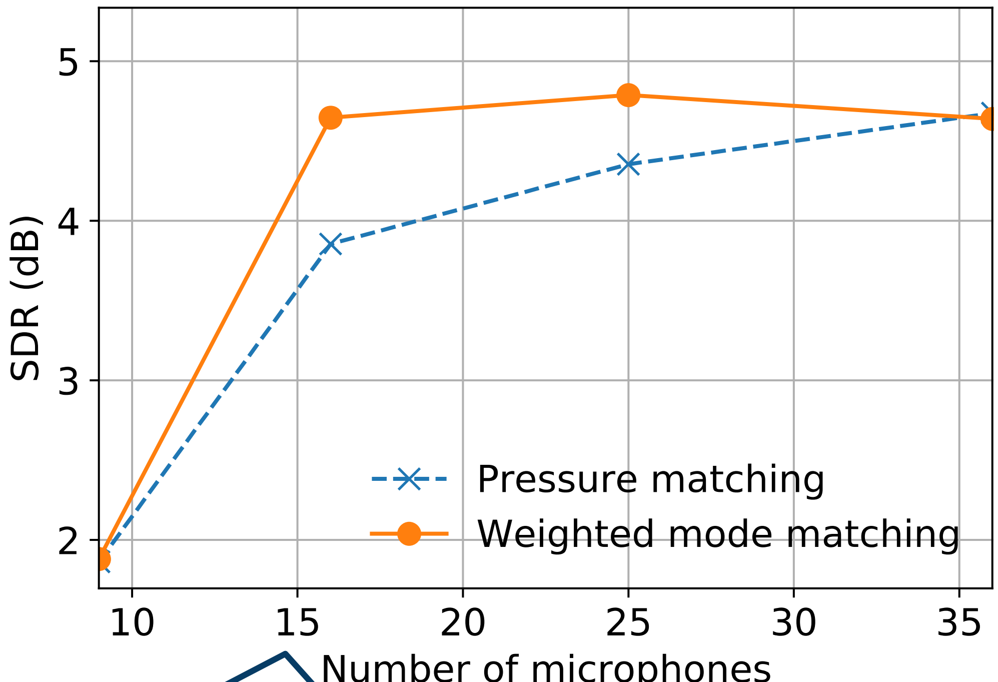




- Synthesizing plane wave by 32 loudspeakers and 16 mics



- # of mics vs. SDR



Weighted mode matching outperforms pressure matching in the case of small number of mics

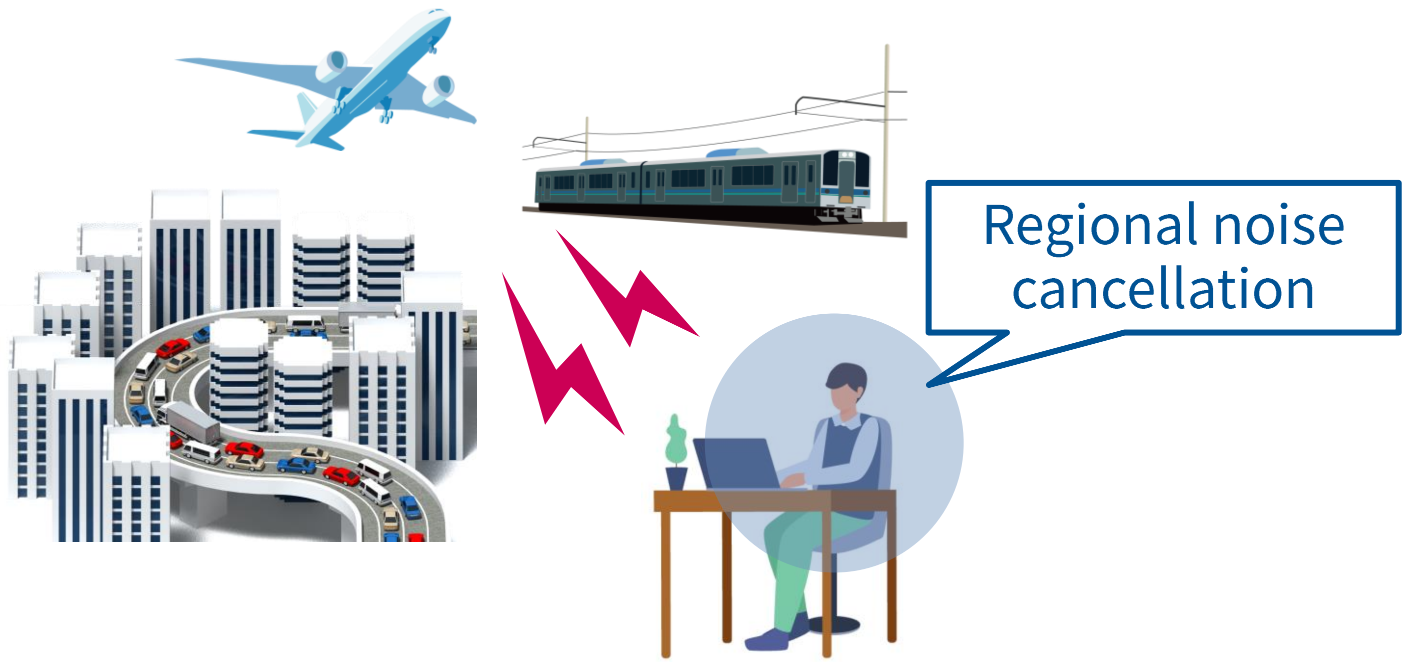
Table of contents

1. Preliminaries on acoustics
   1. Wave equation and Helmholtz equation
   2. Representations of acoustic field
2. Integral-equation-based sound field estimation
   1. Problem formulation
   2. Sound field estimation with spherical array
   3. Methods to avoid forbidden frequency problem
3. Least-squares-based sound field estimation
   1. Basic framework
   2. Infinite-dimensional extension
   3. Other extensions and related works
4. Applications
   1. Spatial audio reproduction by headphones
   2. Spatial audio reproduction by loudspeakers
   3. Spatial active noise control

Spatial active noise control

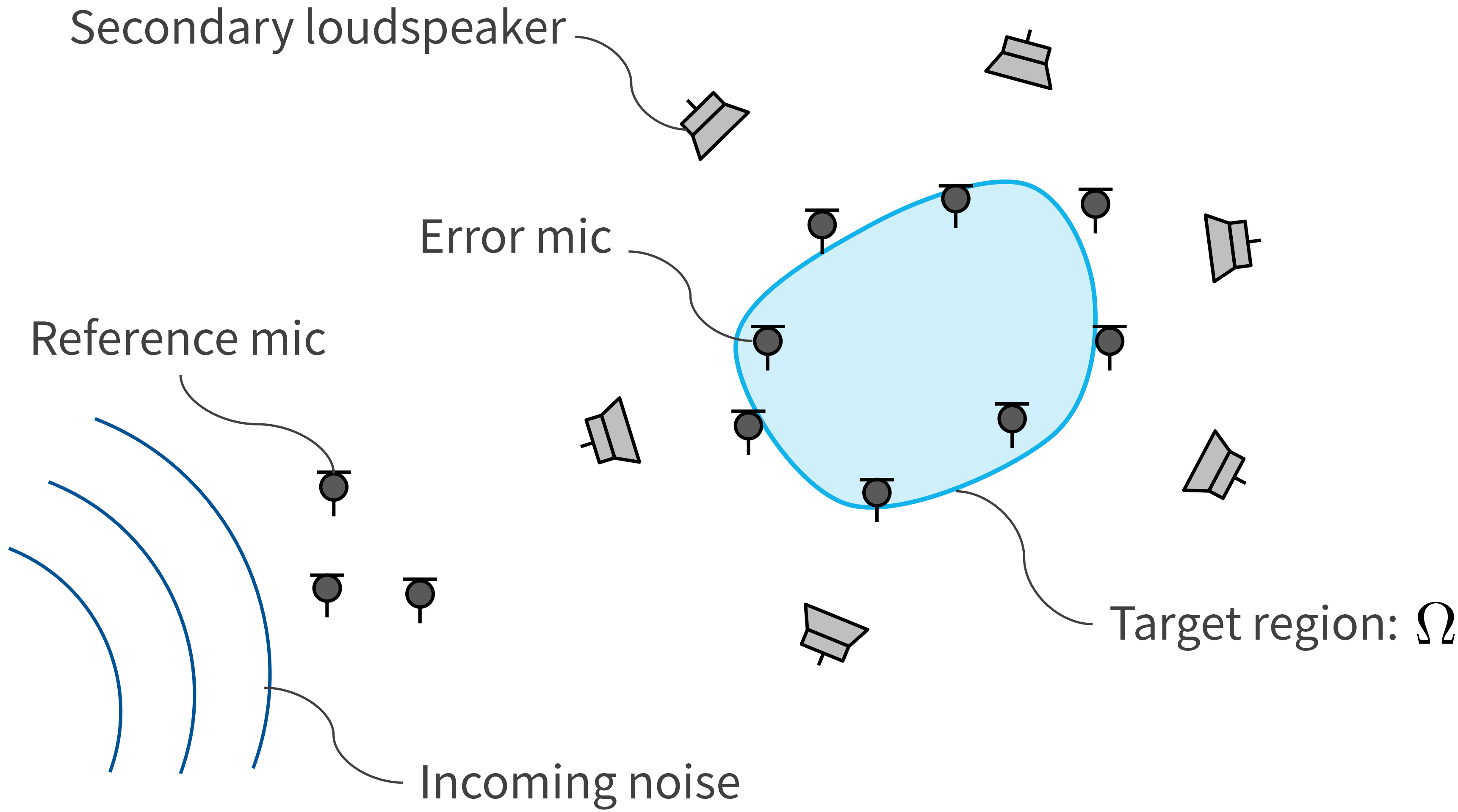
Noise cancelling by loudspeaker signals

- Noise pollution is still major problem for human health
- Active noise control (ANC) is aimed to cancel noise by loudspeaker signals, but its effect is limited to local region
- Goal of spatial ANC is regional noise cancellation in 3D space



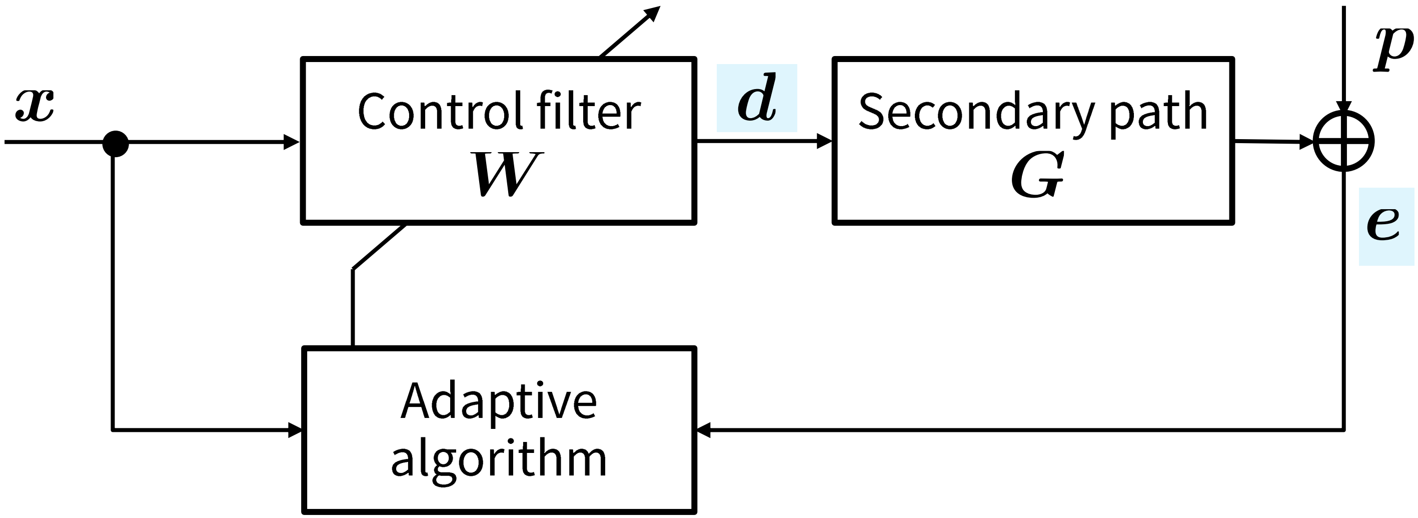
Spatial active noise control

➢ Incoming noise is suppressed in 3D target region by secondary loudspeakers ➡ Spatial Active Noise Control



Spatial active noise control

- Block diagram of multichannel feedforward ANC system

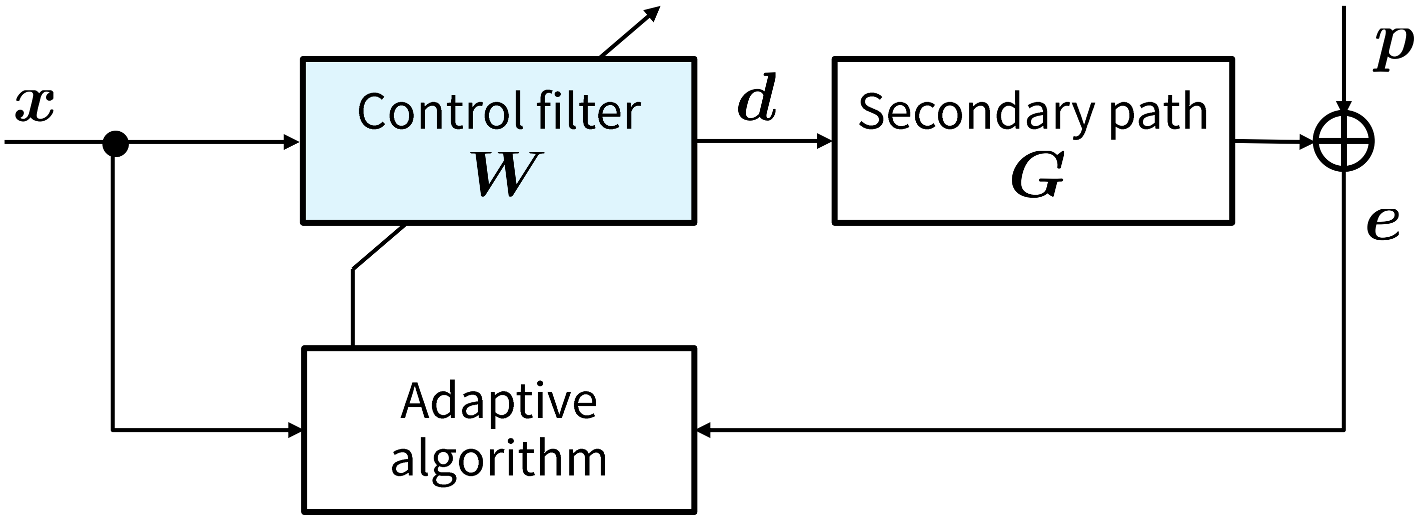


Error mic signal: $$\boldsymbol{e}(n) = \boldsymbol{p}(n) + \sum_{i=0}^{J-1} \boldsymbol{G}(i)\boldsymbol{d}(n-i)$$

Loudspeaker signal: $$\boldsymbol{d}(n) = \sum_{i=0}^{I-1} \boldsymbol{W}(i)\boldsymbol{x}(n-i)$$

Spatial active noise control

- Block diagram of multichannel feedforward ANC system



- **Adaptive computation of control filter**

Cost function

$$\boldsymbol{W}_{n+1}(i) = \boldsymbol{W}_n(i) - \eta \frac{\partial \mathcal{L}}{\partial \boldsymbol{W}_n(i)}$$

Spatial active noise control

Cost function design for spatial ANC

- Conventional multipoint pressure control (MPC):

$$\mathcal{L} := \mathbb{E}\left[\|\boldsymbol{e}(n)\|^2\right]$$

Aimed to reduce noise at error mics only

- Proposed cost function based on regional noise power:

$$\mathcal{L} := \mathbb{E}\left[\int_{\Omega} u(\boldsymbol{r}, n)^2 \mathrm{d}\boldsymbol{r}\right]$$

Aimed to reduce noise over target region Ω

Sound field inside target region must be predicted from error mic signals

Formulation of cost function based on kernel interpolation

[Koyama+ IEEE/ACM TASLP 2021]

- Time-domain kernel interpolation from error mic signals

$$u(\boldsymbol{r}, n) = \sum_{i=-\infty}^{\infty} \boldsymbol{z}^{\mathsf{T}}(\boldsymbol{r}, i)\boldsymbol{e}(n-i)$$

$$\boldsymbol{z}(\boldsymbol{r}, i) = \mathcal{F}^{-1}\left[\left(\left(\boldsymbol{K}(\omega) + \lambda \boldsymbol{I}\right)^{-1}\right)^{\mathsf{T}} \boldsymbol{\kappa}(\boldsymbol{r}, \omega)\right]$$

Kernel interpolation filter

- Cost function based on kernel interpolation

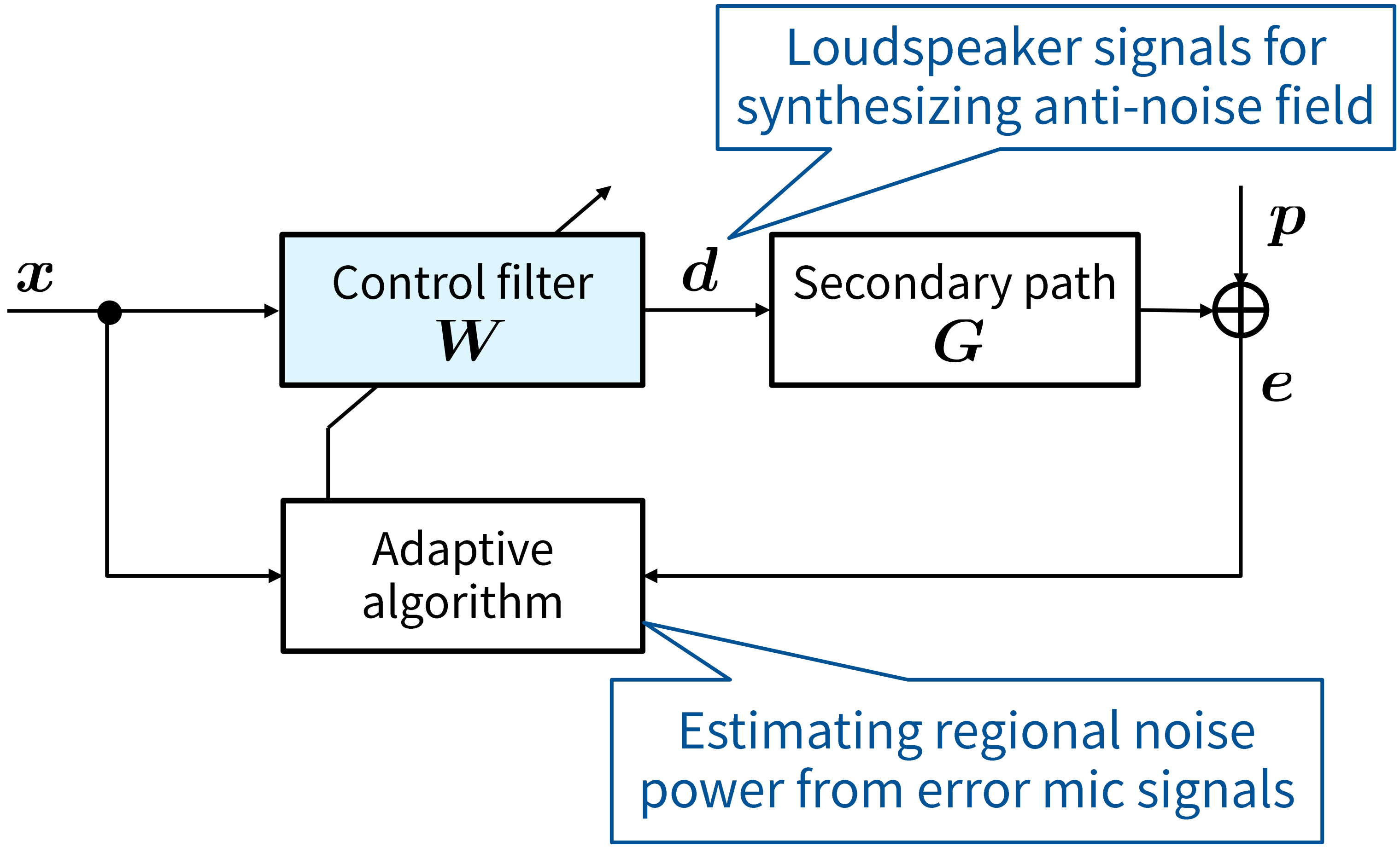
$$\mathcal{L} = \mathbb{E}\left[\sum_{i,j=-\infty}^{\infty} \boldsymbol{e}^{\mathsf{T}}(n-i)\boldsymbol{A}(i,j)\boldsymbol{e}(n-j)\right]$$

$$\boldsymbol{A}(i,j) := \int_{\Omega} \boldsymbol{z}(\boldsymbol{r}, i)\boldsymbol{z}^{\mathsf{T}}(\boldsymbol{r}, j)\mathrm{d}\boldsymbol{r}$$

Weighting matrix for interpolation

Spatial active noise control

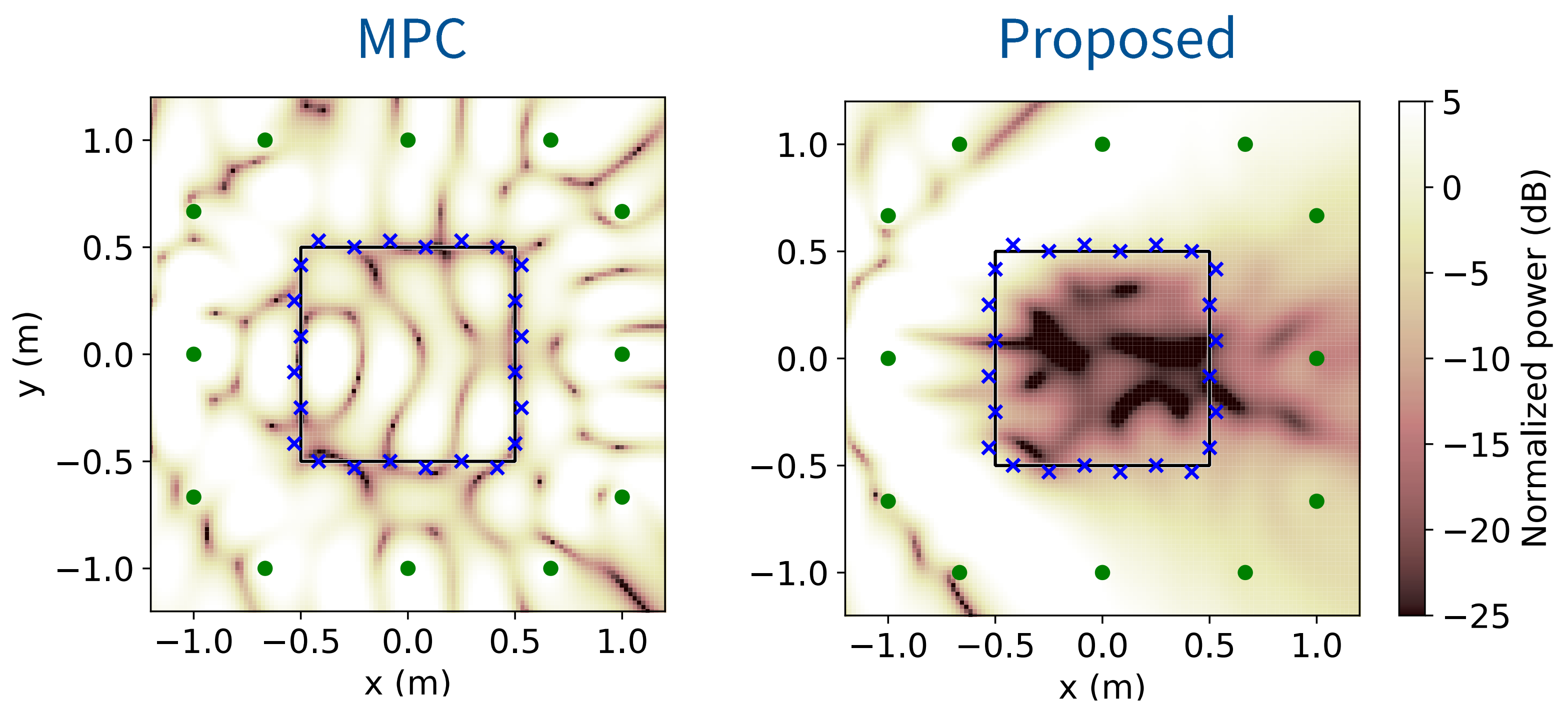
- Feedforward spatial ANC system



Kernel-interpolation-based Filtered-X least mean square (KI-FxLMS) algorithm for spatial ANC

Experiments in frequency domain

- Normalized power distribution at 700 Hz
  - 2D free-field simulation using 24 error mics and 12 loudspeakers



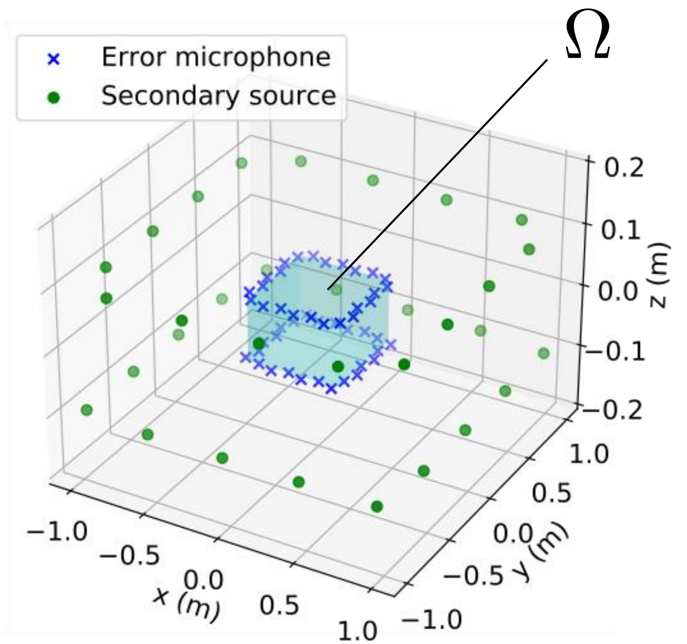
Regional noise reduction is achieved by the proposed method

Experiments in time domain using real data

- Experimental setting:
  - # of loudspeakers: 32
  - # of error mics: 48
  - Reference mic:Directly obtained from primary noises
  - Ω : 0.6 m x 0.6 m x 0.1 m
  - Reverberation time: 380 ms (T_{60})
  - Methods: MPC (FxLMS) and Proposed (KI-FxLMS) are compared
  - Performance measure:

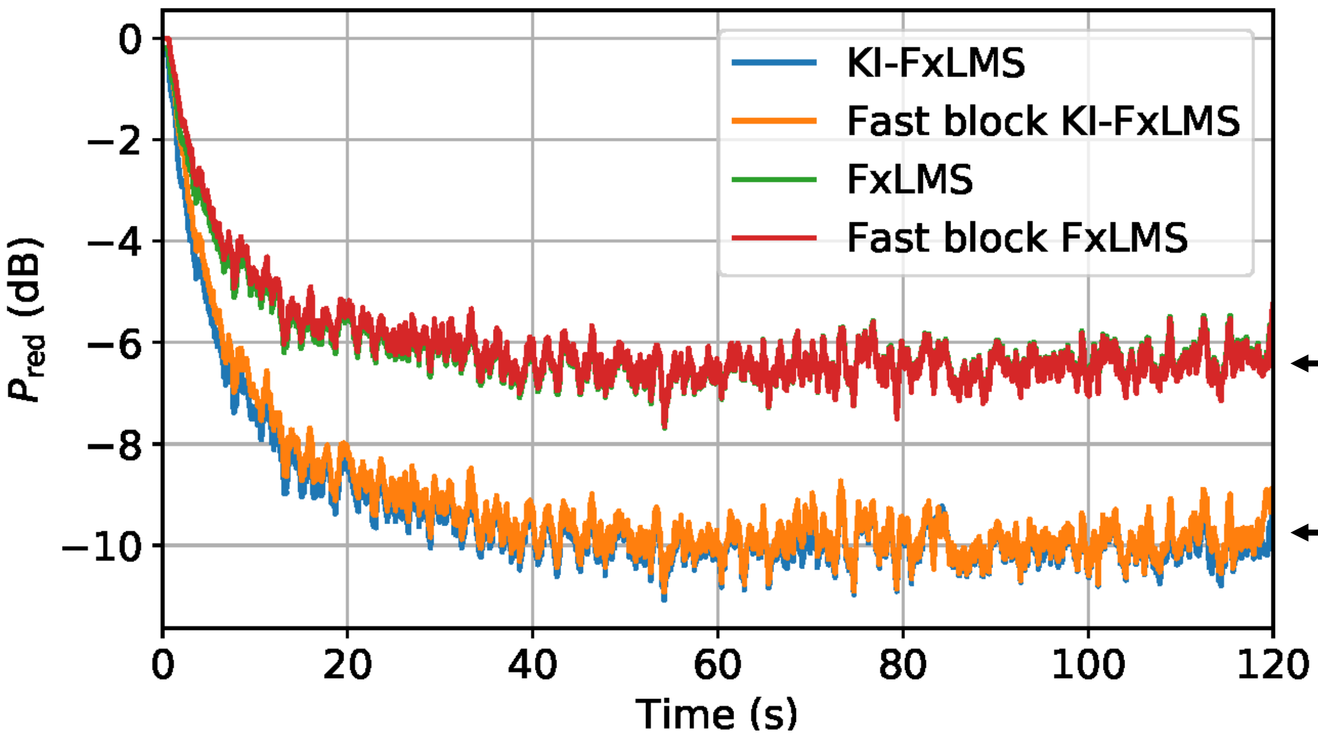
$$P_{\mathrm{red}}(n) = 10\log_{10}\frac{\sum_j \sum_\nu u(\boldsymbol{r}_j, n-\nu)^2}{\sum_j \sum_\nu u_{\mathrm{p}}(\boldsymbol{r}_j, n-\nu)^2}$$

Primary noise field



Experiments in time domain using real data

➢ Regional noise reduction w.r.t. time



FxLMS: –6.41 dB

KI-FxLMS: –10.00 dB

P_{red} of KI-FxLMS is much lower than that of FxLMS

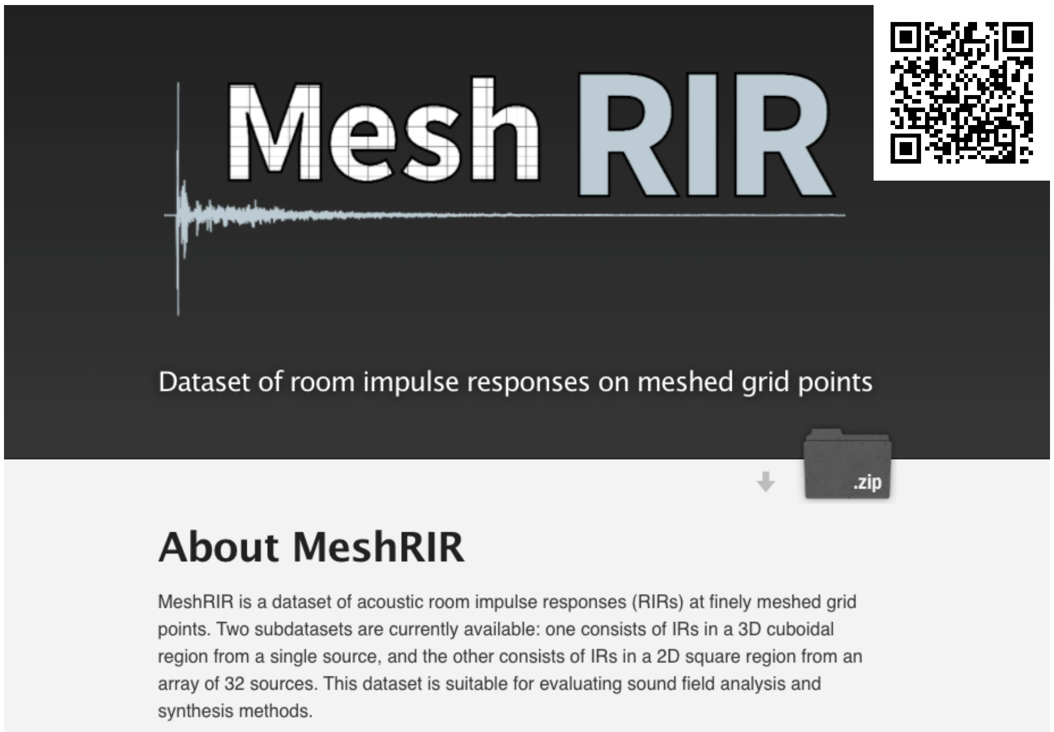
Conclusion

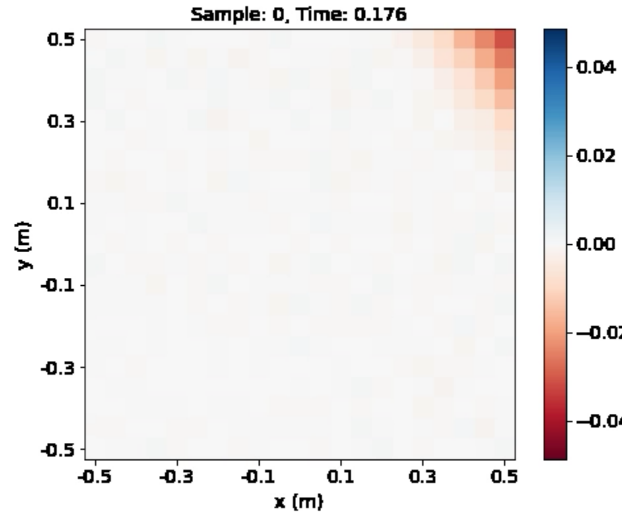
- Application examples of sound field estimation
  - Spatial audio reproduction by headphones [Iijima+ JASA 2021]
    - Binaural reproduction from mic array recordings
    - Binaural signals are redered from estimated expansion coeffs by distributed mics based on spherical wave expansion
  - Spatial audio reproduction by loudspeakers [Ueno+ IEEE/ACM TASLP 2019]
    - Reproducing desired sound field by multiple loudspeakers based on weighted mode matching
    - Estimation of expansion coeffs for reproducing captured sound field and/or using measured transfer functions
  - Spatial active noise control [Ito+ ICASSP 2019, Koyama+ IEEE/ACM TASLP 2021]
    - Noise cancellation over target region based on kernel interpolation of sound field from error mic signals
    - Kernel-interpolation-based FxLMS algorithm outperforms multipoint-pressure-control-based FxLMS

Dataset of room impulse responses (RIRs)

[Koyama+ IEEE WASPAA 2021]

- Released RIR dataset on meshed grid points with example codes
  - https://sh01k.github.io/MeshRIR/





RIR measurement system

Conclusion

- Sound field estimation: recent advances and applications
  - Integral-equation-based sound field estimation
    - Stable computation and useful for analyzing properties because of analytical formulation of estimator
    - Applicable only to simple array geometry to derive analytical formulation
  - Least-squares-based sound field estimation
    - Based on minimization of square error
    - Extentions to infinite-dimensional analysis and its relation to kernel interpolation
    - Appliable to arbitrary array geormetry
  - Appliactions
    - Spatial audio reproduction by headphones
    - Spatial audio reproduction by loudspeakers
    - Spatial active noise control

References (1)

- S. Koyama, G. Chardon, L. Daudet, “Joint source and sensor placement for sound field control based on empirical interpolation method,” in Proc. ICASSP, 2018.
- N. Ueno, S. Koyama, and H. Saruwatari, “Three-dimensional sound field reproduction based on weighted mode-matching method,” IEEE/ACM Trans. Audio, Speech, Lang., Process., 27(12), pp. 1852-1867, 2019.
- H. Ito, S. Koyama, N. Ueno, and H. Saruwatari, “Feedforward spatial active noise control based on kernel interpolation of sound field,” in Proc. IEEE ICASSP, 2019.
- S. Koyama, G. Chardon, and L. Daudet, “Optimizing source and sensor placement for sound field control: an overview,” IEEE/ACM Trans. Audio, Speech, Lang., Process., 28, pp. 686-714, 2020.
- J. G. C. Ribeiro, N. Ueno, S. Koyama, and H. Saruwatari, “Kernel interpolation of acoustic transfer function between regions considering reciprocity,” in Proc. IEEE SAM, 2020.
- N. Iijima, S. Koyama, and H. Saruwatari, “Binaural Rendering from Microphone Array Signals of Arbitrary Geometry,” J. Acoust. Soc. Am., 150(4), pp. 2479-2491, 2021.
- S. Koyama, J. Brunnström, H. Ito, N. Ueno, and H. Saruwatari, “Spatial Active Noise Control Based on Kernel Interpolation of Sound Field,” IEEE/ACM Trans. Audio, Speech, Lang., Process., 29, pp. 3052-3063, 2021.

References (2)

- S. Koyama, T. Nishida, K. Kimura, T. Abe, N. Ueno, and J. Brunnström, "MeshRIR: A Dataset of Room Impulse Responses on Meshed Grid Points for Evaluating Sound Field Analysis and Synthesis Methods," in Proc. IEEE WASPAA, 2021.
- S. Koyama, K. Kimura, and N. Ueno, "Sound Field Reproduction With Weighted Mode Matching and Infinite-Dimensional Harmonic Analysis: An Experimental Evaluation," in Proc. I3DA, 2021.
- R. Horiuchi, S. Koyama, J. G. C. Ribeiro, N. Ueno, and H. Saruawatari, "Kernel learning for sound field estimation with L1 and L2 regularizations," in Proc. IEEE WASPAA, 2021.
- K. Kimura, S. Koyama, N. Ueno, and H. Saruwatari, "Mean-square-error-based secondary source placement in sound field synthesis with prior information on desired field," in Proc. IEEE WASPAA, 2021.
- T. Nishida, N. Ueno, S. Koyama, and H. Saruwatari, "Region-restricted sensor placement based on Gaussian process for sound field estimation," IEEE Trans. Signal Process., 70, pp. 1718-1733, 2022.
- K. Arikawa, S. Koyama, and H. Saruwatari, "Spatial active noise control based on individual kernel interpolation of primary and secondary sound fields," in Proc. IEEE ICASSP, 2022.
- J. G. C. Ribeiro, S. Koyama, and H. Saruwatari, "Region-to-region kernel interpolation of acoustic transfer function with directional weighting," in Proc. IEEE ICASSP, 2022.