Sparsity-based Sound Field Reconstruction

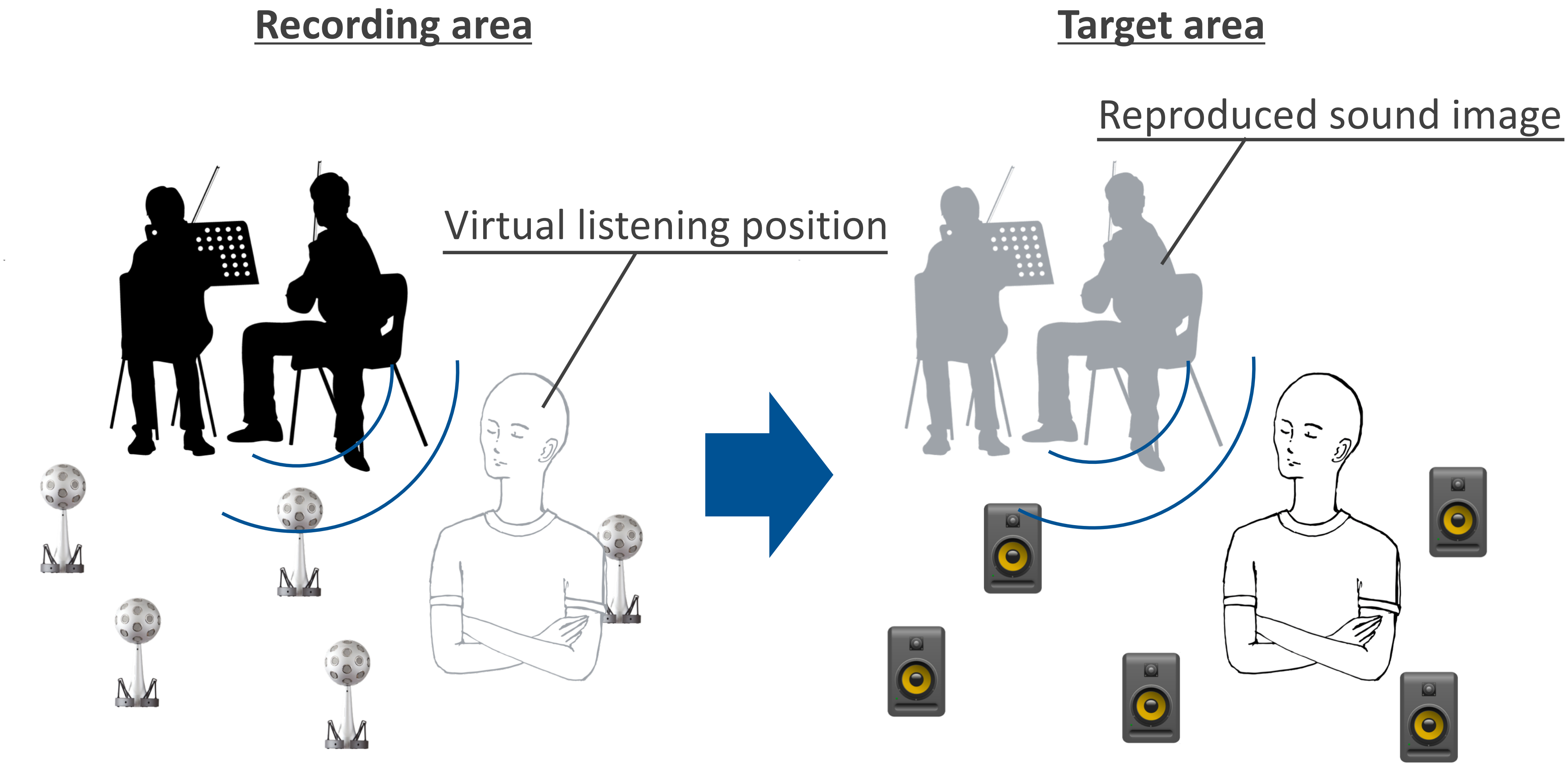
Shoichi Koyama

The University of Tokyo

JST, PRESTO

Spatial Sound Recording and Reproduction

How to capture and reproduce physically correct sound field?

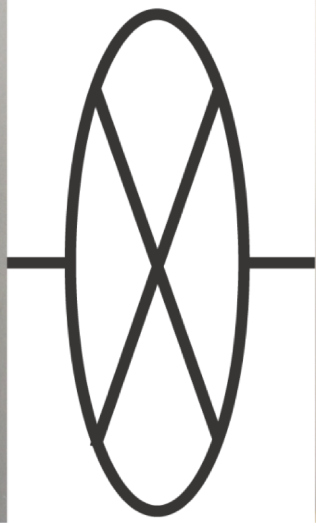


Real-time Sound Field Transmission System

System developed while I worked at NTT [Koyama+ IEICE Trans 2014]

Kanagawa

Network

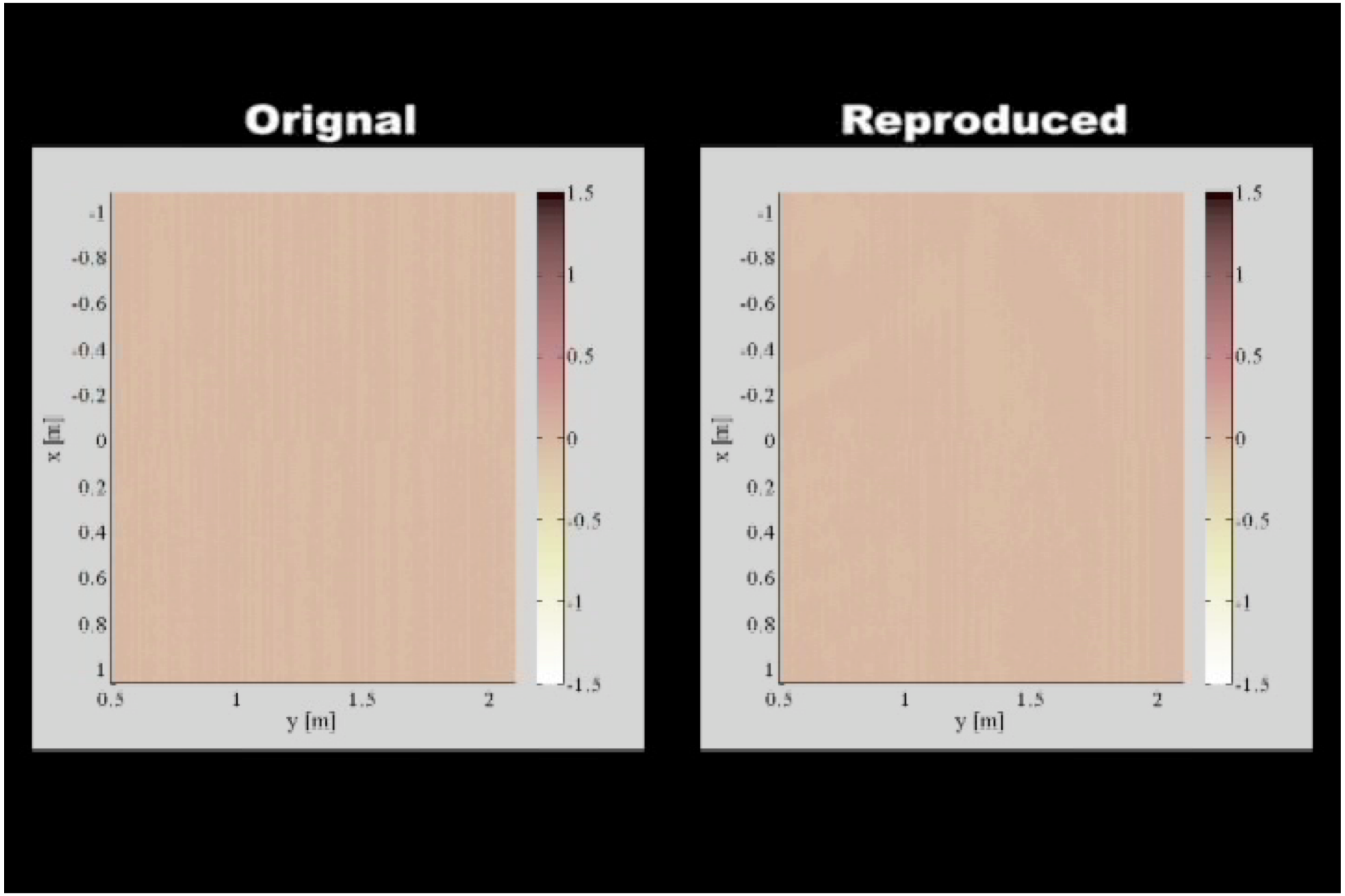


Tokyo

- Loudspeakers (for high freq.): 64, 6cm intervals
- Loudspeakers (for low freq.): 32, 12cm intervals
- Microphones: 64, 6cm intervals
- Array size: 3.84 m
- Sampling freq.: 48 kHz, Delay: 152 ms

Visualization of Reproduced Sound Field

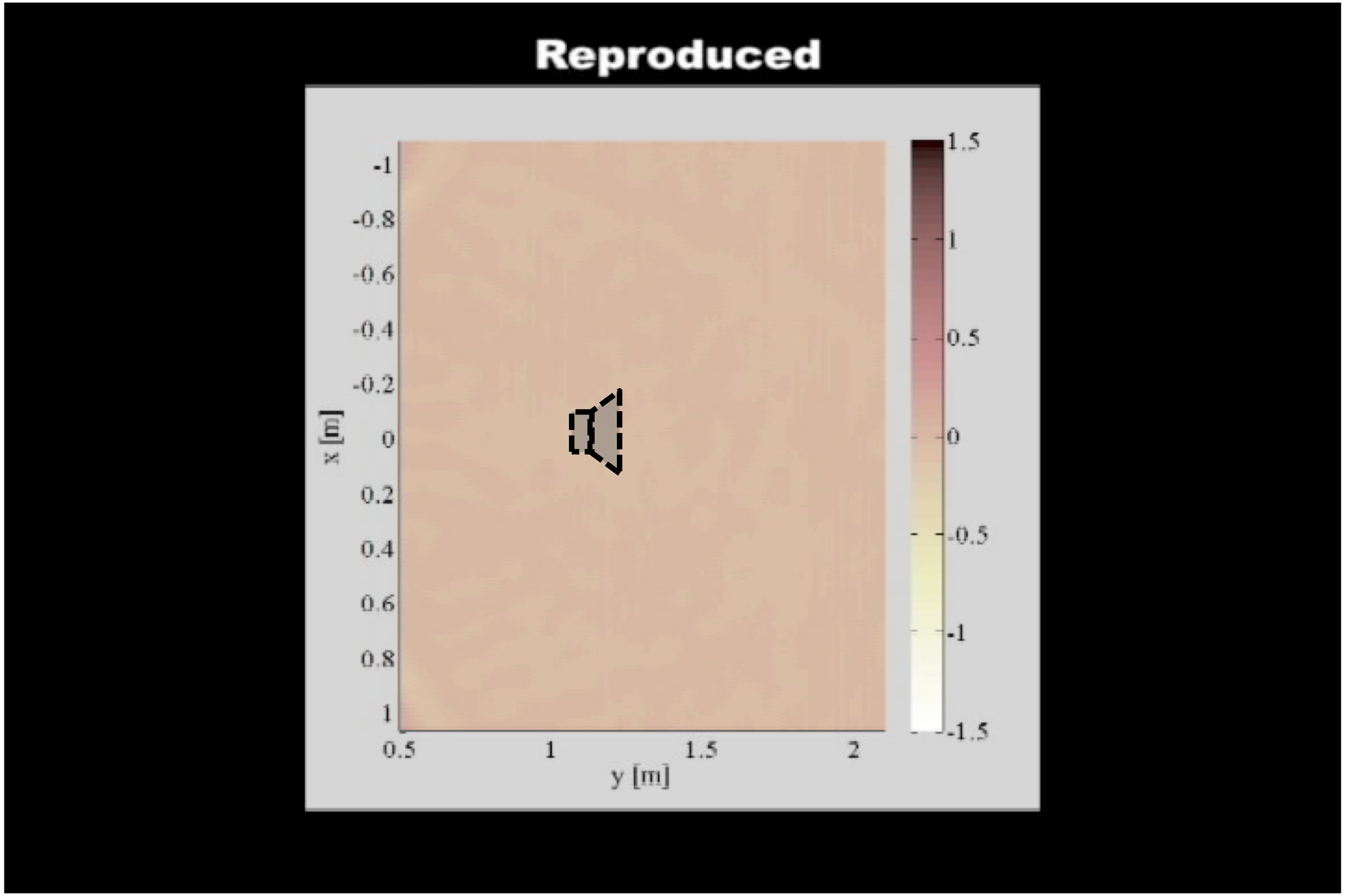
- Source signal: Low-passed pulse (0 – 2.6kHz)
- Source: Loudspeaker, Position: (-1.0, -1.0, 0.0) m



[Koyama+ IEEE TASLP 2013]

Visualization of Reproduced Sound Field

- Source signal: Low-passed pulse (0 – 2.6kHz)
- Source: Loudspeaker, Position: (0.0, -1.0, 0.0) m, 2.0 m forward shift

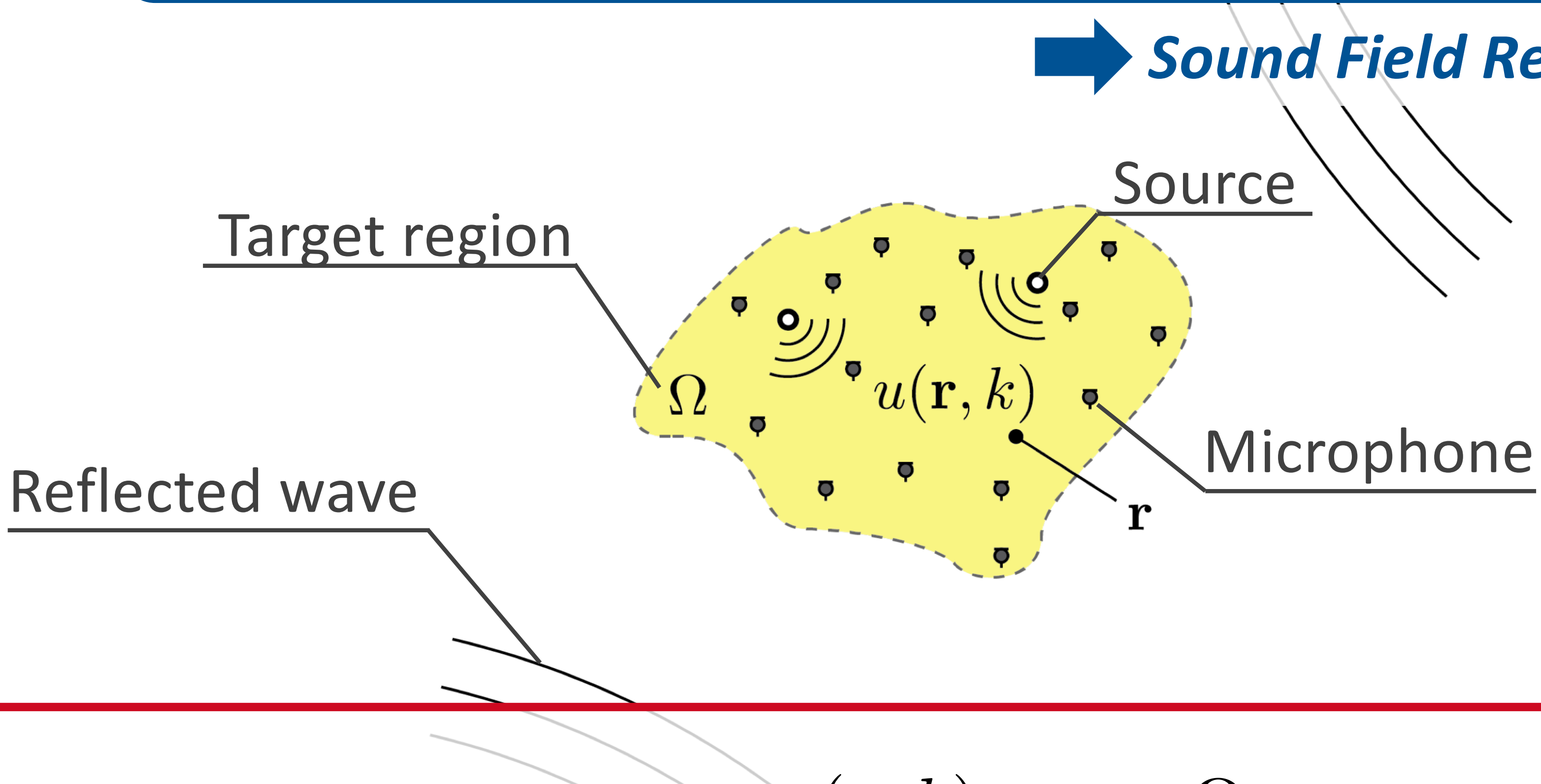


[Koyama+ IEEE TASLP 2013]

Today's Topic

How to estimate and interpolate continuous sound field from measurements of multiple microphones?

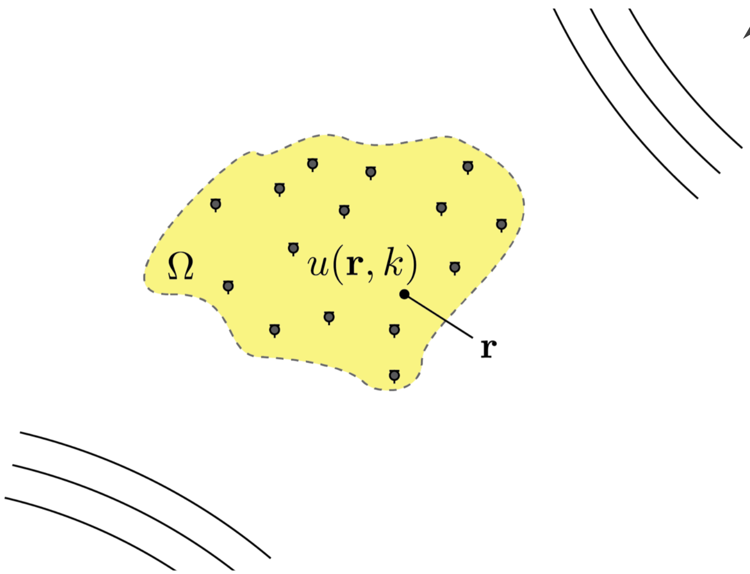
➡ ***Sound Field Reconstruction***



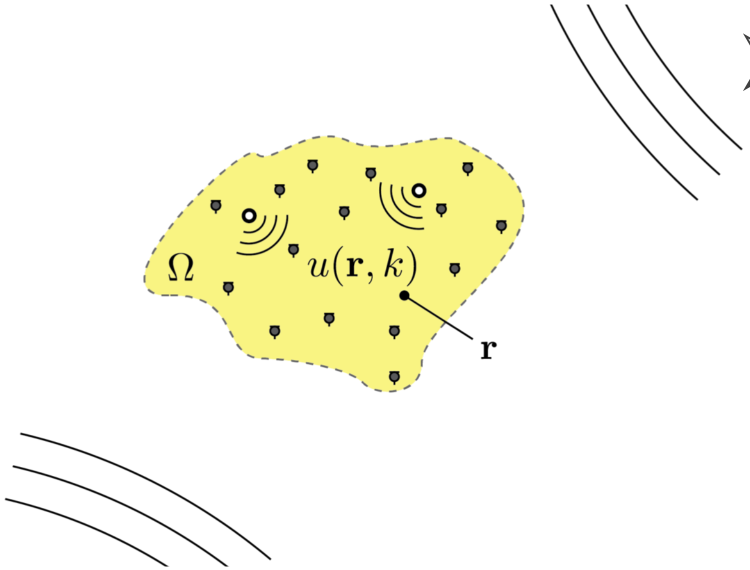
Goal: Estimate continuous $u(\mathbf{r}, k)$ inside Ω by using pressure measurements $u(\mathbf{r}_m, k)$ $(m \in \{1, \ldots, M\})$

➡ **Visualization, reproduction by loudspeakers/headphones etc…**

Sound Field Reconstruction

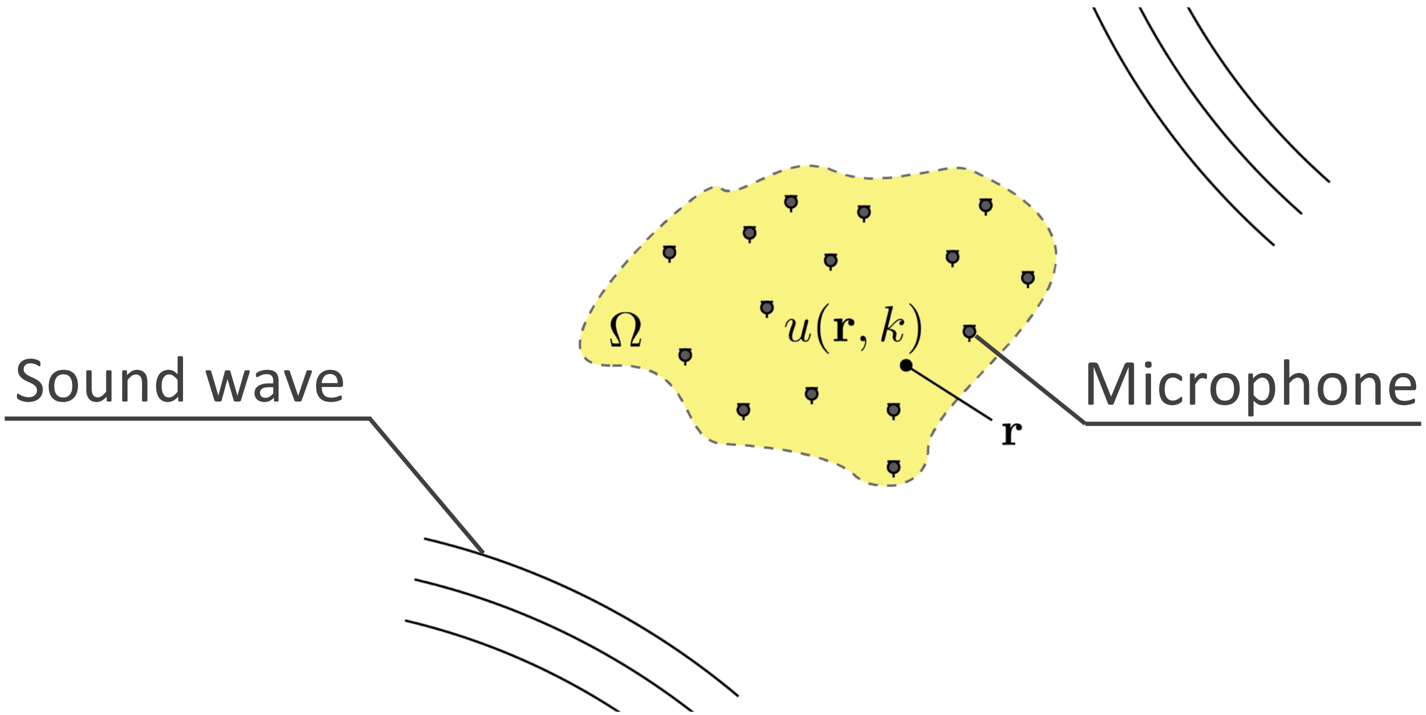


- **Target region does NOT include any sources**
  - Interpolation with constraint of homogeneous Helmholtz eq.
  - Decomposition of captured sound field into plane-wave or harmonic functions: ***sound field decomposition***



- **Target region includes some sources**
  - *ill-posed problem!*
  - Some assumptions must be imposed on source distribution

Homogeneous Sound Field Reconstruction



- Sound field inside source-free region

➡ $u(\mathbf{r}, k)$ **satisfies homogeneous Helmholtz eq.**

$$\begin{cases} (\nabla^2 + k^2) u(\mathbf{r}, k) = 0 \\ \text{Unknown boundary condition on room surface} \end{cases}$$

Homogeneous Sound Field Reconstruction

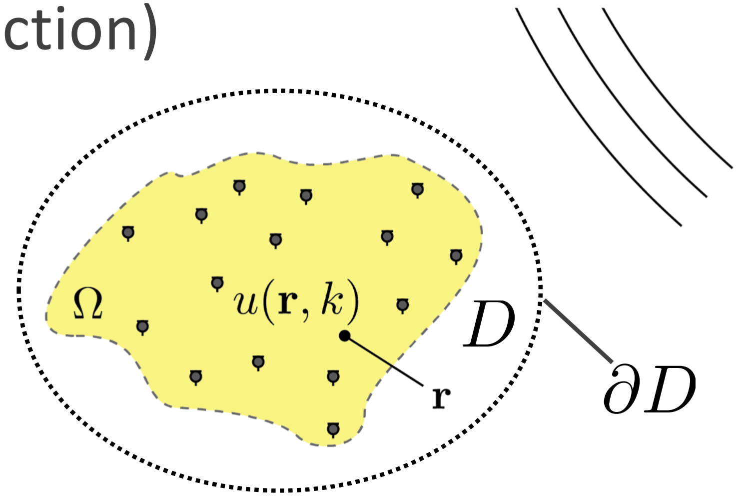
Decomposition into element solutions of Helmholtz eq.

- Plane-wave function (Herglotz wave function)

$$u(\mathbf{r}) = \int_{\boldsymbol{\eta}\in\mathbb{S}^2} \gamma(\boldsymbol{\eta}) \underline{e^{jk\langle \mathbf{r},\boldsymbol{\eta}\rangle}} \mathrm{d}\boldsymbol{\eta}$$

- Spherical wave function

$$u(\mathbf{r}) = \sum_{\nu=0}^{\infty} \sum_{\nu=-\mu}^{\mu} \alpha_\nu^\mu \underline{j_\nu(kr) Y_\nu^\mu(\theta,\phi)}$$



- Equivalent source method [Koopmann+ 1989]

$$u(\mathbf{r}) = \int_{\mathbf{r}'\in\partial D} \psi(\mathbf{r}') \underline{G(\mathbf{r}|\mathbf{r}')} \mathrm{d}\mathbf{r}'$$: single layer potential

Free-field Green's func.: $G(\mathbf{r}|\mathbf{r}') = \dfrac{e^{jk\|\mathbf{r}-\mathbf{r}'\|_2}}{4\pi\|\mathbf{r}-\mathbf{r}'\|_2}$

[Colton+ 2013]

Harmonic Analysis of Infinite Orders

[Ueno+ IEEE SPL 2018]

- Spherical wave function expansion

$$u(\mathbf{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \alpha_\nu^\mu(\mathbf{r}_0) \varphi_\nu^\mu(\mathbf{r} - \mathbf{r}_0)$$

Expansion center

$$\varphi_\nu^\mu(\mathbf{r}) = \sqrt{4\pi} j_\nu(kr) Y_\nu^\mu(\theta, \phi)$$

Spherical harmonic function

- Representation by infinite vectors

$$u(\mathbf{r}) = \boldsymbol{\alpha}(\boldsymbol{r}_0)^{\mathsf{T}} \boldsymbol{\varphi}(\boldsymbol{r} - \boldsymbol{r}_0)$$

– Coefficient vector: $\boldsymbol{\alpha}(\mathbf{r}_0) \in \mathbb{C}^\infty$

– Basis-function vector: $\boldsymbol{\varphi}(\mathbf{r} - \mathbf{r}_0) \in \mathbb{C}^\infty$

Harmonic Analysis of Infinite Orders

[Ueno+ IEEE SPL 2018]

- Measurement by m th microphone at $\mathbf{r}_m$ with directivity of $c_m(\theta, \phi)$ is represented as

Measurement noise

$$s_m = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} c_{m,\nu}^{\mu *} \alpha_\nu^\mu(\mathbf{r}_0) + \epsilon_m$$

Expansion coefficient of $c_m(\theta, \phi)$

$$= \mathbf{c}_m^{\mathsf{H}} \boldsymbol{\alpha}(\mathbf{r}_m) + \epsilon_m$$: Representation by infinite vectors

$$= \mathbf{c}_m^{\mathsf{H}} \mathbf{T}(\mathbf{r}_m - \mathbf{r}_0) \boldsymbol{\alpha}(\mathbf{r}_0) + \epsilon_m$$

Translation matrix from $\mathbf{r}_0$ to $\mathbf{r}_m$ [Martin 2006]

- Stacking M measurements: $\mathbf{s} = [s_1, \ldots, s_M]^{\mathsf{T}}, \ \boldsymbol{\epsilon} = [\epsilon_1, \ldots, \epsilon_M]^{\mathsf{T}}$

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

$$\boldsymbol{\Xi}(\mathbf{r}_0) = [(\mathbf{c}_1^{\mathsf{H}} \mathbf{T}(\mathbf{r}_1 - \mathbf{r}_0))^{\mathsf{H}}, \ldots, (\mathbf{c}_M^{\mathsf{H}} \mathbf{T}(\mathbf{r}_M - \mathbf{r}_0))^{\mathsf{H}}]$$
$$= [\mathbf{T}(\mathbf{r}_1 - \mathbf{r}_0)\mathbf{c}_1, \ldots, \mathbf{T}(\mathbf{r}_M - \mathbf{r}_0)\mathbf{c}_M]$$

Harmonic Analysis of Infinite Orders

[Ueno+ IEEE SPL 2018]

- Expansion coef. at $\mathbf{r}_0$ is estimated as

$$\hat{\boldsymbol{\alpha}}(\mathbf{r}_0) = \boldsymbol{\Xi}(\mathbf{r}_0)(\boldsymbol{\Psi} + \lambda\mathbf{I})^{-1}\mathbf{s}$$

$$\begin{aligned}(\boldsymbol{\Psi})_{m,m'} &= \left(\boldsymbol{\Xi}(\mathbf{r}_0)^{\mathsf{H}}\boldsymbol{\Xi}(\mathbf{r}_0)\right)_{m,m'} \\ &= \mathbf{c}_m^{\mathsf{H}}\mathbf{T}(\mathbf{r}_m - \mathbf{r}_0)\mathbf{T}(\mathbf{r}_0 - \mathbf{r}_{m'})\mathbf{c}_{m'} \\ &= \mathbf{c}_m^{\mathsf{H}}\mathbf{T}(\mathbf{r}_m - \mathbf{r}_{m'})\mathbf{c}_{m'}\end{aligned}$$

➡ **Dependency on expansion center $\mathbf{r}_0$ is removed**

- Expansion coef. at arbitrary position $\mathbf{r}$:

$$\underline{\hat{\boldsymbol{\alpha}}(\mathbf{r}) = \boldsymbol{\Xi}(\mathbf{r})(\boldsymbol{\Psi} + \lambda\mathbf{I})^{-1}\mathbf{s}}$$

Expansion coef. at arbitrary position can be estimated independently of truncation and expansion center

Conventional Harmonic Analysis

[Laborie+ 2003, Samarasinghe+ 2014]

- Spherical wave function expansion w/ truncation

$$u(\mathbf{r}) \approx \sum_{\nu=0}^{N} \sum_{\mu=-\nu}^{\nu} \alpha_\nu^\mu(\mathbf{r}_0)\varphi_\nu^\mu(\mathbf{r}-\mathbf{r}_0)$$: approx. by truncation

- Microphone measurements

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

$$\begin{cases} \bar{\boldsymbol{\alpha}}(\mathbf{r}_0) \in \mathbb{C}^{(N+1)^2} \\ \bar{\mathbf{\Xi}}(\mathbf{r}_0) \in \mathbb{C}^{(N+1)^2 \times M} \end{cases}$$

- Estimate of expansion coef. at $\mathbf{r}$

$$\hat{\boldsymbol{\alpha}}(\mathbf{r}) = \bar{\mathbf{T}}(\mathbf{r}-\mathbf{r}_0)\underbrace{\bar{\mathbf{\Xi}}(\mathbf{r}_0)\left(\bar{\mathbf{\Xi}}(\mathbf{r}_0)^{\mathsf{H}}\bar{\mathbf{\Xi}}(\mathbf{r}_0) + \lambda\mathbf{I}\right)^{-1}\mathbf{s}}_{\text{Estimate of } \bar{\boldsymbol{\alpha}}(\mathbf{r}_0)}$$

Setting of appropriate truncation order and expansion center is necessary

Relation to kernel ridge regression

[Ueno+ IEEE SPL 2018, IWAENC 2018]

- Harmonic analysis of infinite orders for pressure microphone case ($c^{\mu}_{m,\nu} = \delta_{\nu,0}\delta_{\mu,0}$)

$$\hat{u}(\mathbf{r}) = \boldsymbol{\varphi}(\mathbf{r}-\mathbf{r}_0)^{\mathsf{T}} \boldsymbol{\Xi}(\mathbf{r}_0)(\boldsymbol{\Psi}+\lambda\mathbf{I})^{-1}\mathbf{s}$$

$$= \sum_{m=1}^{M} \left((\boldsymbol{\Psi}+\lambda\mathbf{I})^{-1}\mathbf{s}\right)_m j_0(k\|\mathbf{r}-\mathbf{r}_m\|_2)$$

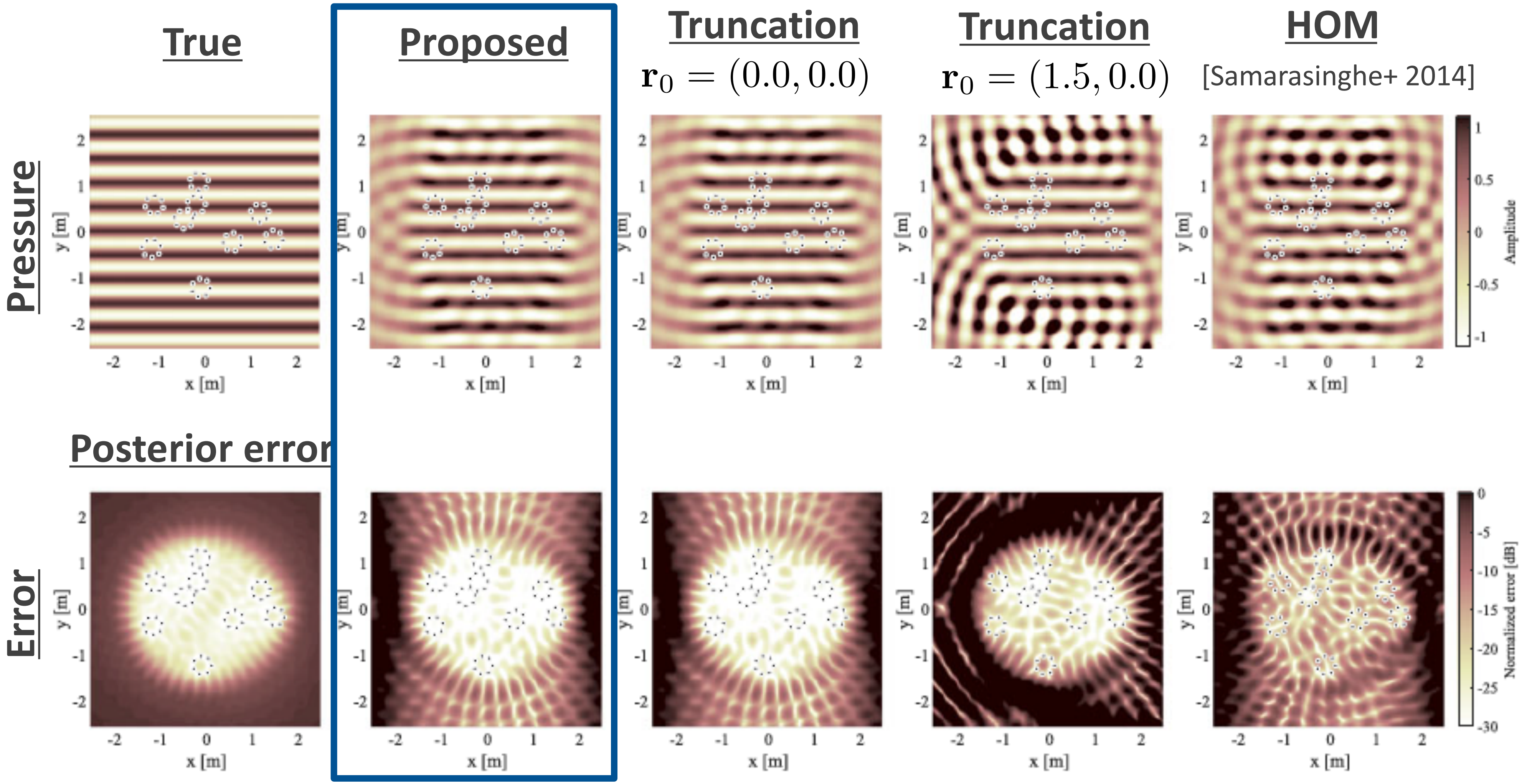
$$\boldsymbol{\Psi} = \begin{bmatrix} j_0(k\|\mathbf{r}_1-\mathbf{r}_1\|) & \cdots & j_0(k\|\mathbf{r}_1-\mathbf{r}_M\|) \\ \vdots & \ddots & \vdots \\ j_0(k\|\mathbf{r}_M-\mathbf{r}_1\|) & \cdots & j_0(k\|\mathbf{r}_M-\mathbf{r}_M\|) \end{bmatrix}$$

Correspond to kernel ridge regression with kernel function of 0th-order spherical Bessel function

Harmonic Analysis of Infinite Orders

➢ **Simulation in 2D for estimating plane wave field (650 Hz)**

[Ueno+ IEEE SPL 2018]



High reconstruction accuracy comparable to the conventional method w/ optimal truncation order and expansion center is achieved

Sparse Plane-wave Decomposition

- Representation by overcomplete plane-wave basis functions ($L \gg M$)

$$u(\mathbf{r}) \approx \sum_{l=1}^{L} \gamma_l e^{j\mathbf{k}_l^{\mathsf{T}}\mathbf{r}}$$

($\mathbf{k}_l$: wave vector of l th plane wave)

➡ **A limited number of nonzero γ_l is sufficient for approximation**

> Sound field in a certain star-shaped region can be well approximated by a limited number of plane waves [Moiola+ 2011]

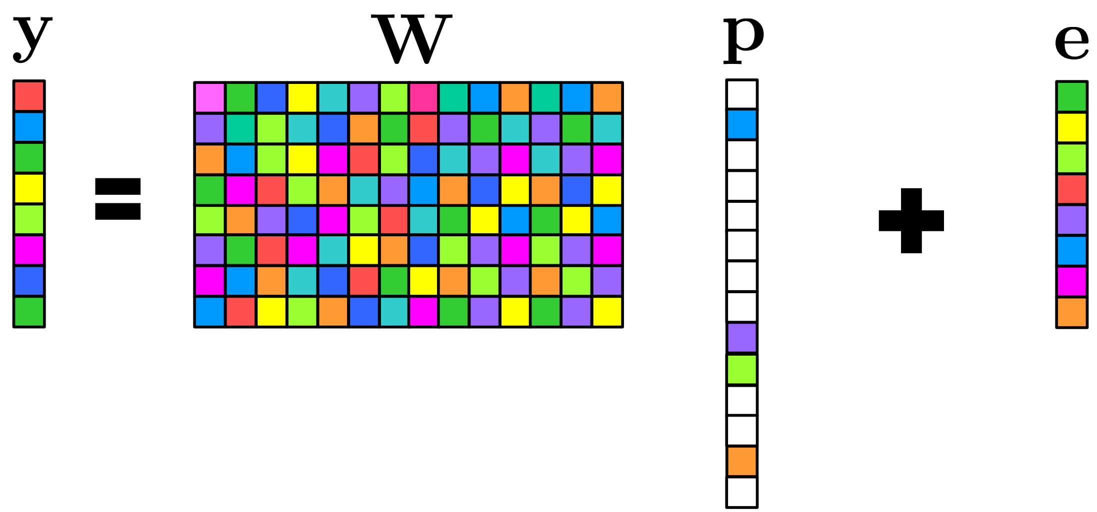
- Matrix form by using dictionary matrix $\mathbf{W} \in \mathbb{C}^{M \times L}$ consisting of plane-wave functions

$$\mathbf{y} = \mathbf{W}\mathbf{p} + \mathbf{e}$$

$$\begin{cases} \mathbf{y} = [u(\mathbf{r}_1), \ldots, u(\mathbf{r}_M)]^{\mathsf{T}} \\ \mathbf{p} = [\gamma_1, \ldots, \gamma_L]^{\mathsf{T}} \end{cases}$$

Sparse Plane-wave Decomposition

- Sparse approximation by plane-wave dictionary matrix



- Optimization problem for sparse approximation

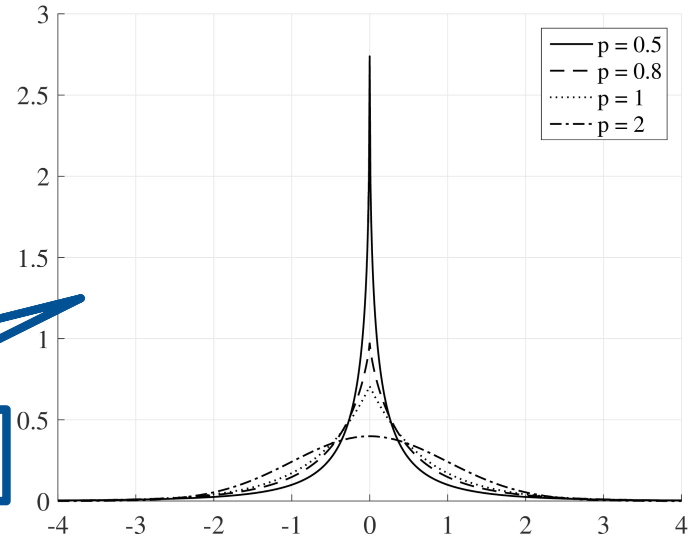
$$\underset{\mathbf{p}}{\text{minimize}}\ \frac{1}{2}\|\mathbf{y}-\mathbf{W}\mathbf{p}\|_2^2 + \lambda\|\mathbf{p}\|_p^p \qquad (0<p\le 1)$$

Penalty term of ℓ_p-(quasi) norm for promoting sparsity of $\mathbf{p}$

Sparse Plane-wave Decomposition

- Generalized Gaussian distribution (GGD)
  - P.d.f. of GGD

$$f(u;p,\beta) = \frac{p}{2\sqrt[p]{2}\beta\Gamma\left(\frac{1}{p}\right)} e^{-\frac{|u|^p}{2\beta^p}}$$



p controls the shape of p.d.f.

- MAP estimation w/ prior distribution of GGD

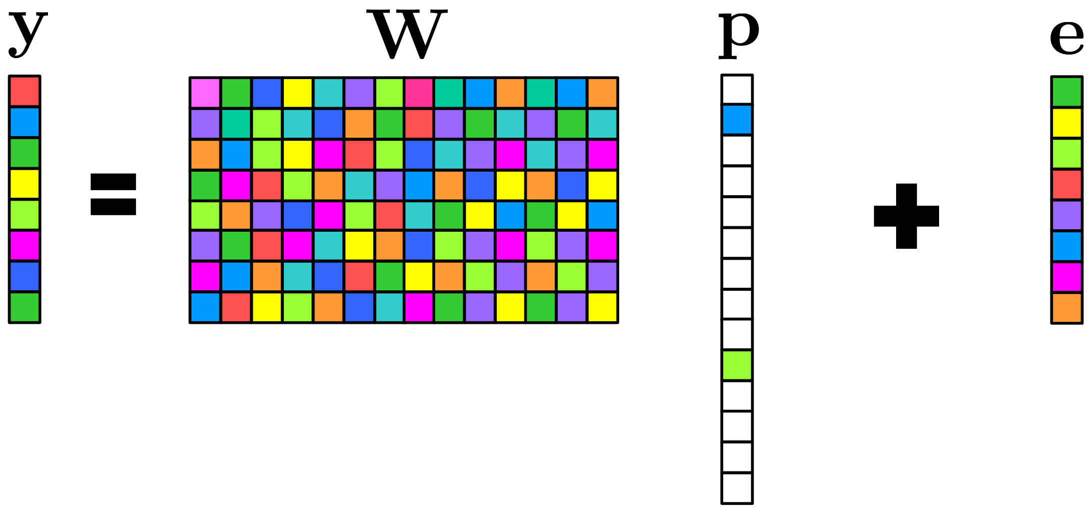
$$p(\mathbf{x}) = \left(\frac{p}{2\sqrt[p]{2}\beta\Gamma\left(\frac{1}{p}\right)}\right)^N \exp\left(-\frac{1}{2\beta^p}\sum_n |x_n|^p\right)$$ **: Prior distribution**

$$\Rightarrow \mathbf{x}_{\mathrm{MAP}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \frac{\sigma^2}{\beta^p}\sum_n |x_n|^p$$

Identical to ℓ_p -norm penalty

Sparse Plane-wave Decomposition

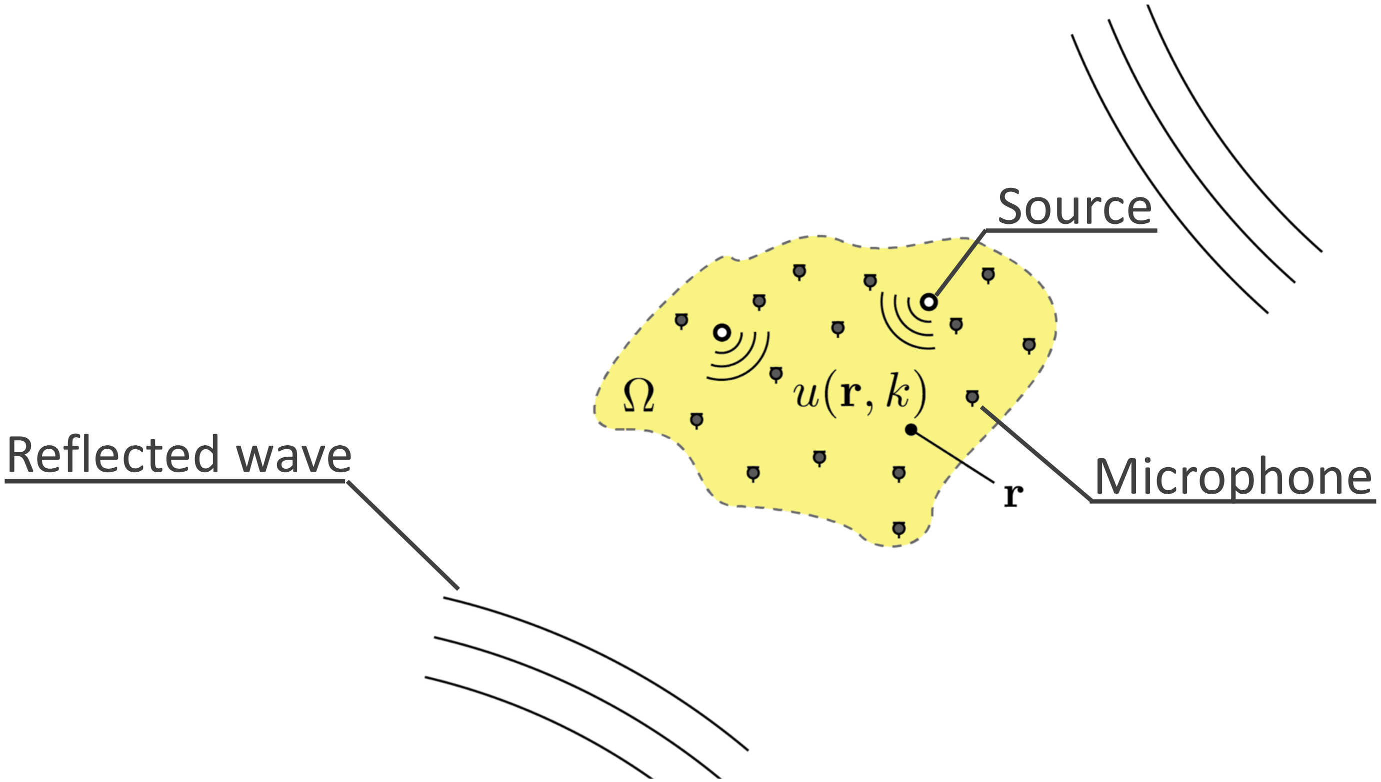
- ➢ Sparse approximation by plane-wave dictionary matrix



➡ **Improve spatial resolution in sound field reconstruction**

- ➢ Application of sparse plane-wave decomposition
  - DOA estimation [Malioutov+ 2005]
  - Nearfield acoustic holography [Chardon+ 2012]
  - Estimation of acoustic transfer functions [Mignot+ 2013]
  - Upscaling of ambisonics coefficients [Wabnitz+ 2013]
  - Multizone sound field control [Jin+ 2015]
  - Exterior and interior sound field separation [Takida+ 2018]

Inhomogeneous Sound Field Reconstruction



- Sound field inside region including sources

➡ $u(\mathbf{r}, k)$ **satisfies inhomogeneous Helmholtz eq.**

$$\left\{ \begin{array}{l} (\nabla^2 + k^2) u(\mathbf{r}, k) = -\underbrace{Q(\mathbf{r}, k)}_{\text{Source distribution}} \\ \text{Unknown boundary condition on room surface} \end{array} \right.$$

Inhomogeneous Sound Field Reconstruction

- $u(\mathbf{r})$ is represented by the sum of particular and homogeneous solutions:

$$u(\mathbf{r}) = u_{\mathrm{P}}(\mathbf{r}) + u_{\mathrm{H}}(\mathbf{r})$$

- $u_{\mathrm{P}}(\mathbf{r})$ can be obtained by convolution of source distribution and free-field Green's func.

$$u_{\mathrm{P}}(\mathbf{r}) = \int_{\mathbf{r}' \in \Omega} Q(\mathbf{r}') G(\mathbf{r}|\mathbf{r}') \mathrm{d}\mathbf{r}' \qquad \left(G(\mathbf{r}|\mathbf{r}') = \frac{e^{jk\|\mathbf{r}-\mathbf{r}'\|_2}}{4\pi\|\mathbf{r}-\mathbf{r}'\|_2} \right)$$

- Integral form of $u(\mathbf{r})$:

$$u(\mathbf{r}) = \int_{\mathbf{r}' \in \Omega} Q(\mathbf{r}') G(\mathbf{r}|\mathbf{r}') \mathrm{d}\mathbf{r}' + u_{\mathrm{H}}(\mathbf{r})$$

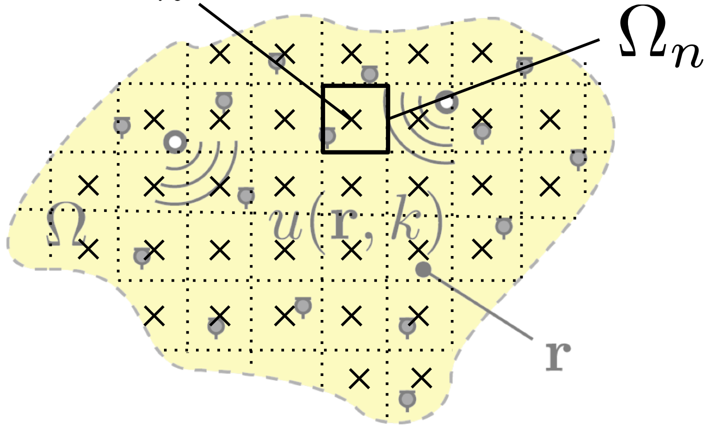
➡ **Estimate $u(\mathbf{r})$ and $Q(\mathbf{r})$ from measurements $u(\mathbf{r}_m)$**

Some constraints on source distribution is required to make this problem solvable

Sparse Sound Field Decomposition

[Koyama+ JASA 2018]

- Discretization of region Ω



$$u_{\mathrm{P}}(\mathbf{r}) = \sum_{n=1}^{N} \int_{\mathbf{r}' \in \Omega_n} Q(\mathbf{r}') G(\mathbf{r}|\mathbf{r}') \mathrm{d}\mathbf{r}'$$

$$\approx \sum_{n=1}^{N} G(\mathbf{r}|\mathbf{r}_n) \int_{\mathbf{r}' \in \Omega_n} Q(\mathbf{r}') \mathrm{d}\mathbf{r}'$$

$$\Rightarrow \quad u(\mathbf{r}) \approx \sum_{n=1}^{N} G(\mathbf{r}|\mathbf{r}_n) \int_{\mathbf{r}' \in \Omega_n} Q(\mathbf{r}') \mathrm{d}\mathbf{r}' + u_{\mathrm{H}}(\mathbf{r})$$

- Matrix form by using dictionary matrix $\mathbf{D} \in \mathbb{C}^{M \times N}$ consisting of free-field Green's func. (i.e., monopoles)

$$\mathbf{y} = \mathbf{D}\mathbf{x} + \mathbf{z}$$

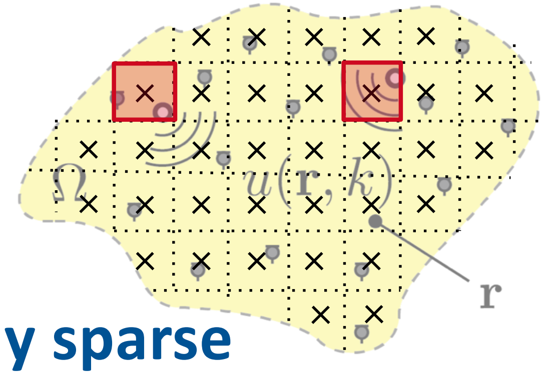
$$\begin{cases} \mathbf{y} = [u(\mathbf{r}_1), \ldots, u(\mathbf{r}_M)]^{\mathsf{T}} \\ \mathbf{x} = \left[\int_{\Omega_1} Q(\mathbf{r}') \mathrm{d}\mathbf{r}', \ldots, \int_{\Omega_N} Q(\mathbf{r}') \mathrm{d}\mathbf{r}' \right]^{\mathsf{T}} \end{cases}$$

Sparse Sound Field Decomposition

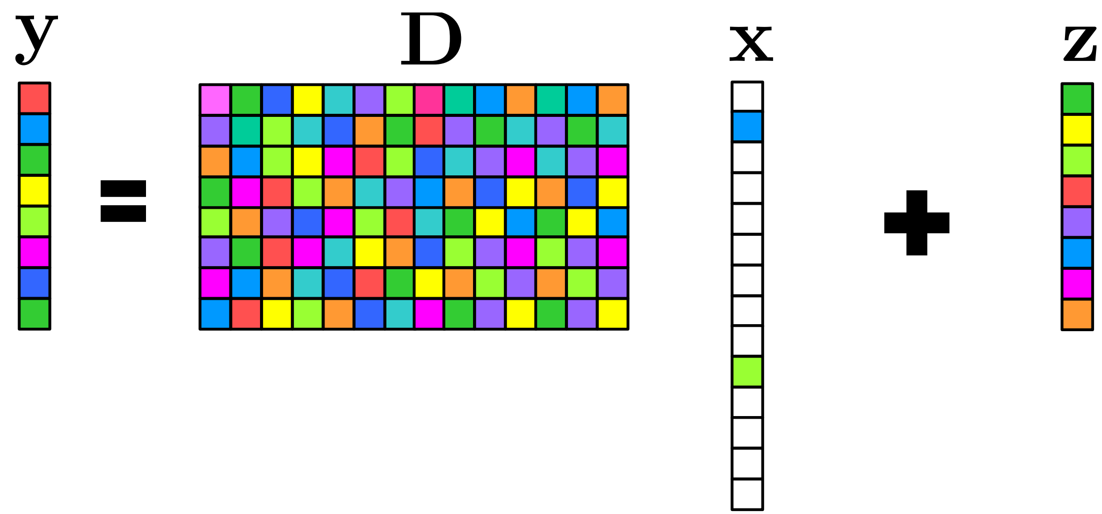
[Koyama+ JASA 2018]

- Linear eq. of measurement model

$$\mathbf{y} = \underbrace{\mathbf{D}\mathbf{x}}_{\text{Direct source component}} + \underbrace{\mathbf{z}}_{\text{Reverberant component}}$$



Assume that source distribution is spatially sparse



- Optimization problem for *sparse sound field decomposition*

$$\underset{\mathbf{x}}{\text{minimize}} \ \frac{1}{2}\|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_p^p \qquad (0 < p \leq 1)$$

Sparsity promoting penalty term

Mixed-norm Penalty for Group Sparsity

[Murata+ IEEE TSP 2018]

- Measurement for each time-frequency bin

$$\mathbf{y}_{t,f} = \mathbf{D}_f \mathbf{x}_{t,f} + \mathbf{z}_{t,f}$$

Indexes of time-frequency bins: $\begin{cases} t \in \{1, \ldots, T\} \\ f \in \{1, \ldots, F\} \end{cases}$

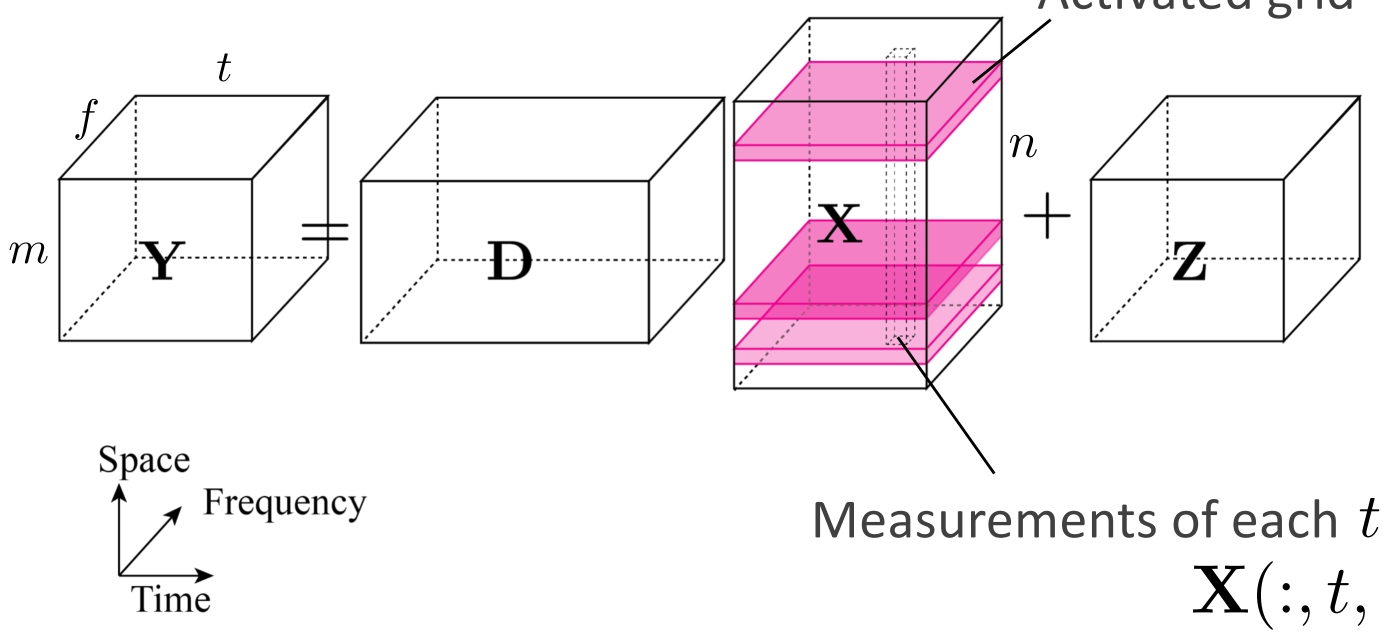
- Group sparsity for robust and accurate decomposition
  - **Sound sources are static for several time frames**
  - **Acoustic source signals have a broad frequency band**

$\boxed{\textbf{Each } \mathbf{x}_{t,f} \textbf{ will have same sparsity pattern}}$

Mixed-norm Penalty for Group Sparsity

[Murata+ IEEE TSP 2018]

- Tensor-form measurement model



Measurements of each t and f :

$$\mathbf{X}(:, t, f) = \mathbf{x}_{t,f}$$

- Optimization problem for group sparse decomposition

$$\underset{\mathbf{X}}{\text{minimize}} \ \frac{1}{2} \sum_{t,f} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \lambda \mathcal{J}_{p,2,2}(\mathbf{X}) \qquad (0 < p \leq 1)$$

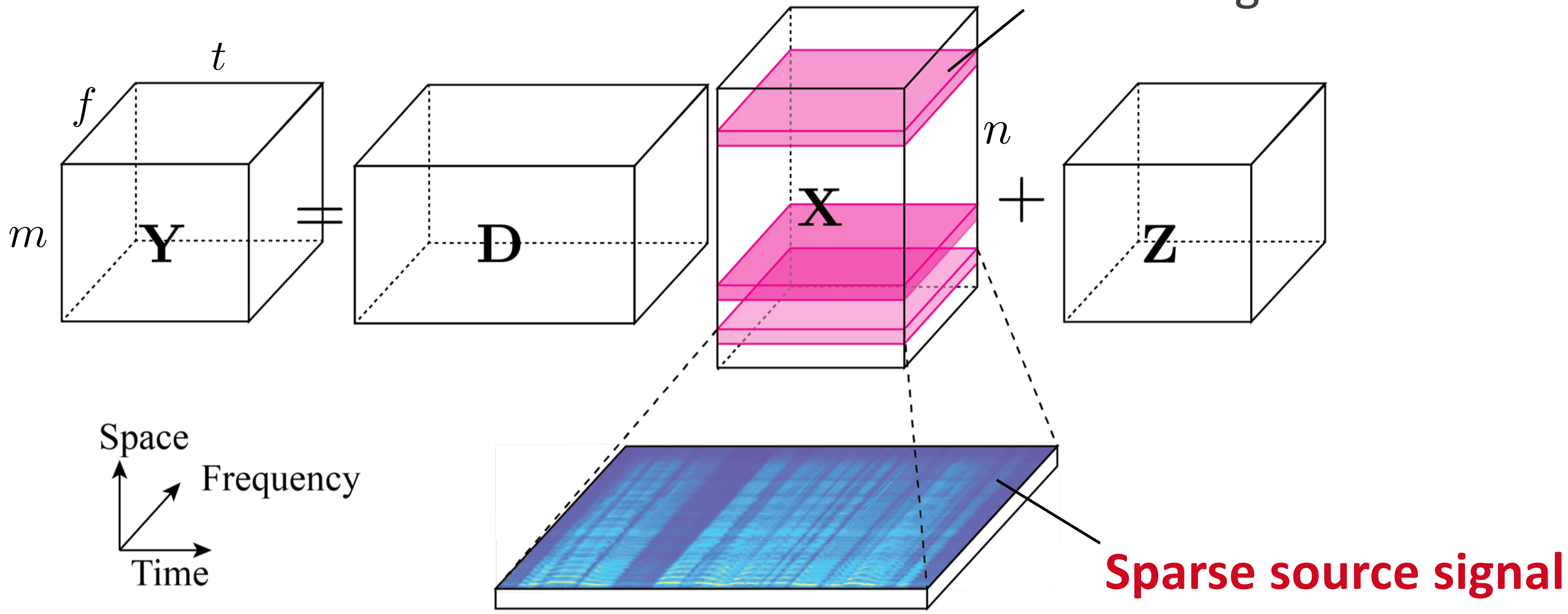
$$\mathcal{J}_{p,2,2}(\mathbf{X}) = \sum_n \left(\sum_{t,f} |\mathbf{X}(n,t,f)|^2 \right)^{\frac{p}{2}}$$

Penalty term for promoting group sparsity ($\ell_{p,2,2}$-norm)

Mixed-norm Penalty for Group Sparsity

[Murata+ IEEE TSP 2018]

- Tensor-form measurement model



- Optimization problem for multidimensional sparsity

$$\underset{\mathbf{X}}{\text{minimize}} \frac{1}{2} \sum_{t,f} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \lambda \mathcal{J}_{p,q,q}(\mathbf{X}) \qquad (0 < p \leq q \leq 1)$$

$$\mathcal{J}_{p,q,q}(\mathbf{X}) = \sum_{n} \left(\sum_{t,f} |\mathbf{X}(n,t,f)|^q \right)^{\frac{p}{q}}$$

Multidimensional mixed-norm penalty term ($\ell_{p,q,q}$-norm)

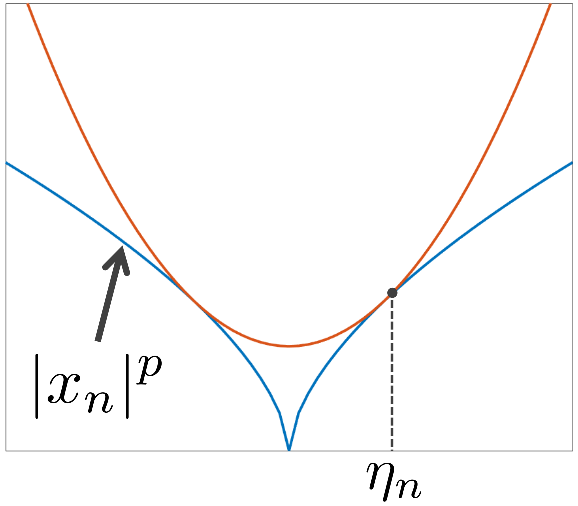
Mixed-norm Penalty for Group Sparsity

[Murata+ IEEE TSP 2018]

- Optimization problem using ℓ_p-norm penalty term

$$\underset{\mathbf{x}}{\text{minimize}} \frac{1}{2}\|\mathbf{y}-\mathbf{D}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_p^p$$

Non-convex problem for $p < 1$



- Majorization-minimization (MM) algorithm
  - Construct surrogate function and alternately update the parameters

$$\|\mathbf{x}\|_p^p = \sum_n |x_n|^p$$

Upper-bounded by quadratic func.

$$\leq \frac{1}{2}\sum_n \{p\eta_n^{p-2}x_n^2 + (2-p)\eta_n^p\}$$

(Equality holds for $x_n = \eta_n$)

➡ **Iteratively reweighted least-squares algorithm**

[Gorodnitsky+ 1997]

Mixed-norm Penalty for Group Sparsity

[Murata+ IEEE TSP 2018]

- Surrogate func. for mixed-norm penalty term

$$\mathcal{J}_{p,q,q}(\mathbf{X}) = \sum_{n} \left(\sum_{t,f} \left(|\mathbf{X}(n,t,f)|^2 \right)^{\frac{q}{2}} \right)^{\frac{p}{q}}$$
$$\leq \sum_{n,t,f} \frac{p}{2} \eta_n^{\frac{p}{q}-1} \eta_{n,t,f}^{\frac{q}{2}-1} |\mathbf{X}(n,t,f)|^2 + C$$
$$= \mathcal{J}^{+}_{p,q,q}(\mathbf{X}|\mathbf{\Xi})$$

(Equality holds for $\mathbf{X} = \mathbf{\Xi}$)

$$\begin{cases} \eta_n = \sum_{t,f} |\mathbf{\Xi}(n,t,f)|^q \\ \eta_{n,t,f} = |\mathbf{\Xi}(n,t,f)|^2 \end{cases}$$

- Alternately update the parameters

$$\begin{cases} \mathbf{x}_{t,f}^{(i+1)} = \underset{\mathbf{x}_{t,f}}{\arg\min} \frac{1}{2} \sum_{t,f} \|\mathbf{y}_{t,f} - \mathbf{D}_f \mathbf{x}_{t,f}\|_2^2 + \frac{1}{2} \lambda \mathbf{x}_{t,f}^{\mathsf{H}} \mathbf{P}_{t,f}^{(i)} \mathbf{x}_{t,f} \\ \mathbf{\Xi}^{(i)} = \mathbf{X}^{(i)} \end{cases}$$

$$\left(\mathbf{P}_{t,f}^{(i)} \right)_{n,n'} = \begin{cases} p \left(\eta_n^{(i)} \right)^{\frac{p}{q}-1} \left(\eta_{n,t,f}^{(i)} \right)^{\frac{p}{2}-1}, & n = n' \\ 0, & n \neq n' \end{cases}$$

➡ **Iteratively reweighted least-squares algorithm**

Mixed-norm Penalty for Group Sparsity

[Murata+ IEEE TSP 2018]

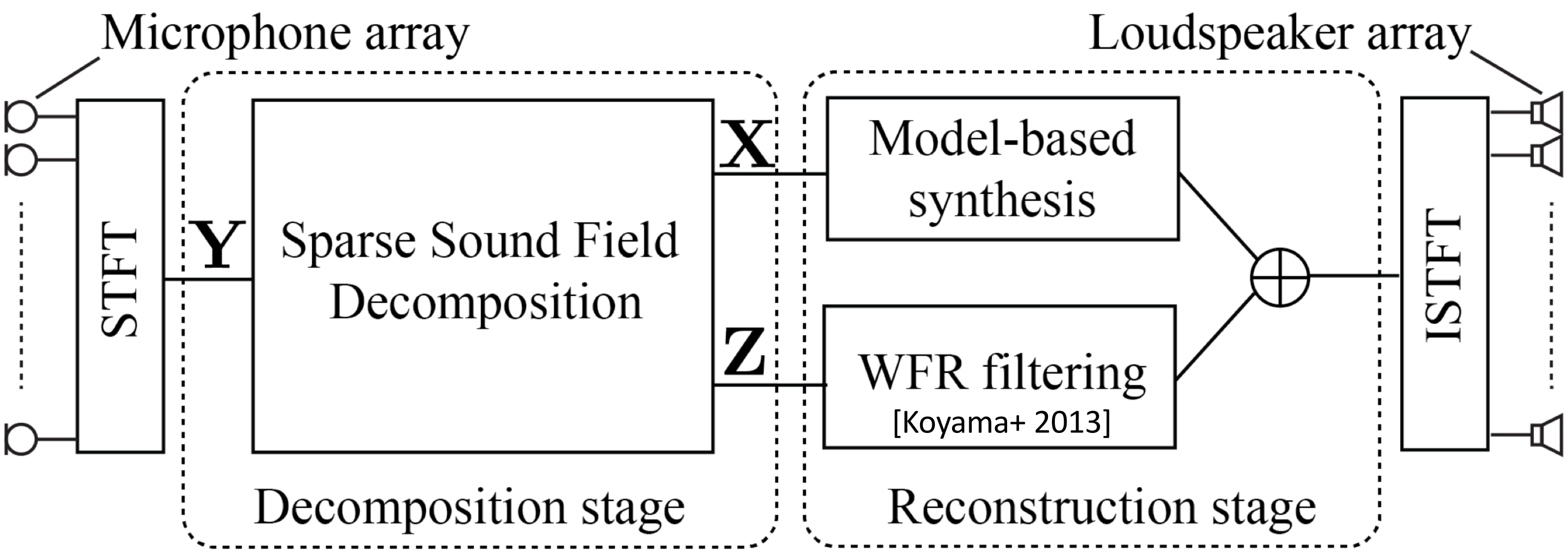
Algorithm 1 Sparse sound field decomposition algorithm using $\ell_{p,q,q}$-norm penalty.

Initialize $\mathbf{X}^{(0)}$, $i = 0$
while loop $\neq 0$ **do**
  $\Xi^{(i)} = \mathbf{X}^{(i)}$
  $\eta_n^{(i)} = \sum_{t,f} \left|\Xi^{(i)}(n,t,f)\right|^q$ for $\forall n$
  for $t = 1$ to T **do**
    for $f = 1$ to F **do**
      $\eta_{n,t,f}^{(i)} = \left|\Xi^{(i)}(n,t,f)\right|^2$ for $\forall n$
      $\mathbf{W}_{t,f}^{(i)} = \mathrm{diag}\left(\sqrt{p^{-1}\left(\eta_n^{(i)}\right)^{1-p/q}\left(\eta_{n,t,f}^{(i)}\right)^{1-q/2}}\right)$
      $\mathbf{A}_{t,f}^{(i)} \leftarrow \mathbf{D}_f \mathbf{W}_{t,f}^{(i)}$
      $\mathbf{x}_{t,f}^{(i+1)}$
        $\leftarrow \mathbf{W}_{t,f}^{(i)}(\mathbf{A}_{t,f}^{(i)})^{\mathsf{H}}\left(\mathbf{A}_{t,f}^{(i)}(\mathbf{A}_{t,f}^{(i)})^{\mathsf{H}} + \lambda\mathbf{I}\right)^{-1}\mathbf{y}_{t,f}$
    end for
  end for
  $i \leftarrow i + 1$
  if stopping condition is satisfied **then**
    loop $= 0$
  end if
end while

Monotonic non-increase of objective func. is guaranteed

Application of Sparse Decomposition

Sparse decomposition for recording and reproduction [Koyama+ JASA 2018]



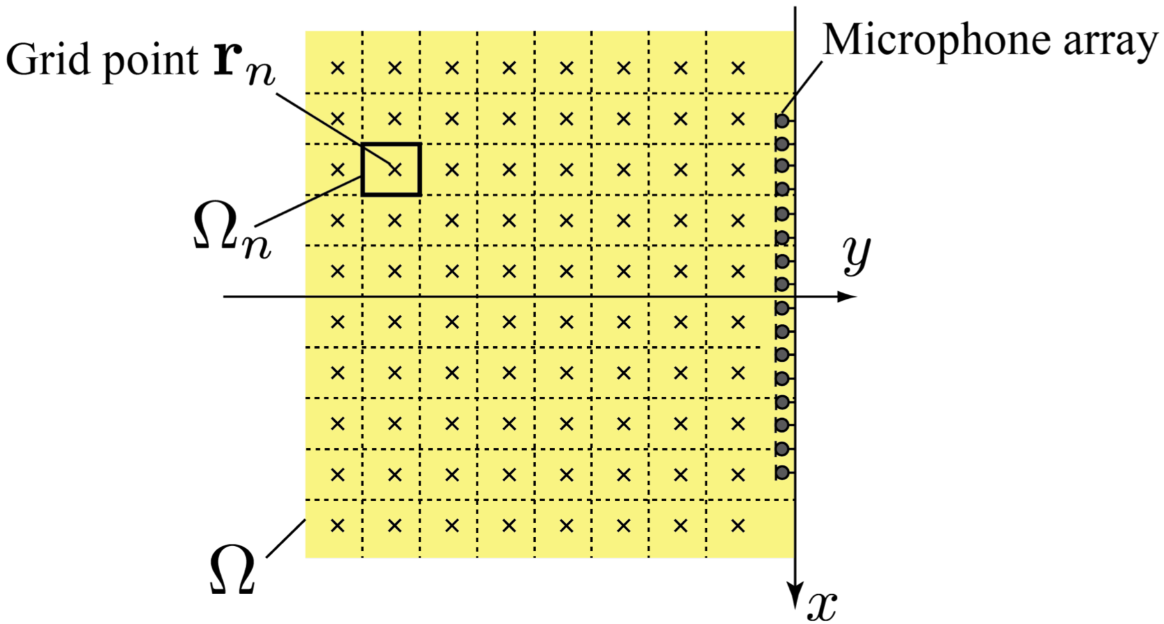
- **Decomposition stage:**
  - Group sparse decomposition of $\mathbf{Y}$ into $\mathbf{X}$ and $\mathbf{Z}$
- **Reconstruction stage:**
  - $\mathbf{X}$ and $\mathbf{Z}$ are separately converted into driving signals
  - Loudspeaker driving signals as sum of two components

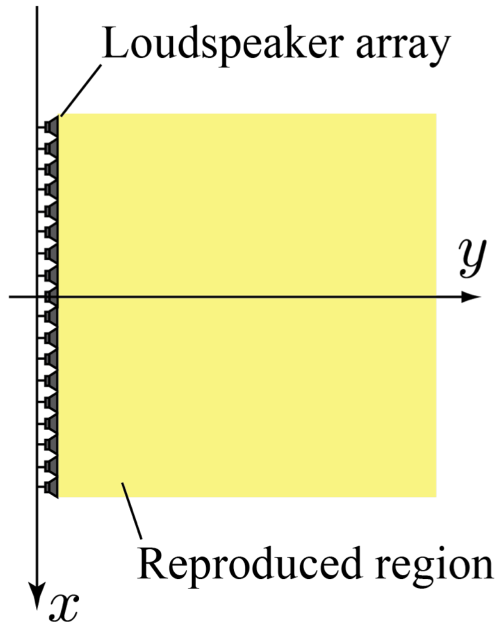
Several Extensions of Sparse Decomposition

- Non-Gaussian reverberantion [Koyama+ IEEE WASPAA 2017]
  - Explicit modeling of reverberant component such as sparsity in plane-wave domain and low-rankness
  - ADMM algorithm for solving joint optimization

- Gridless sound field decomposition [Takida+ IEEE SAM 2018]
  - Approximate sources as delta functions
  - Reciprocity gap functional in spherical harmonic domain
  - Closed-form solution using Hankel matrix

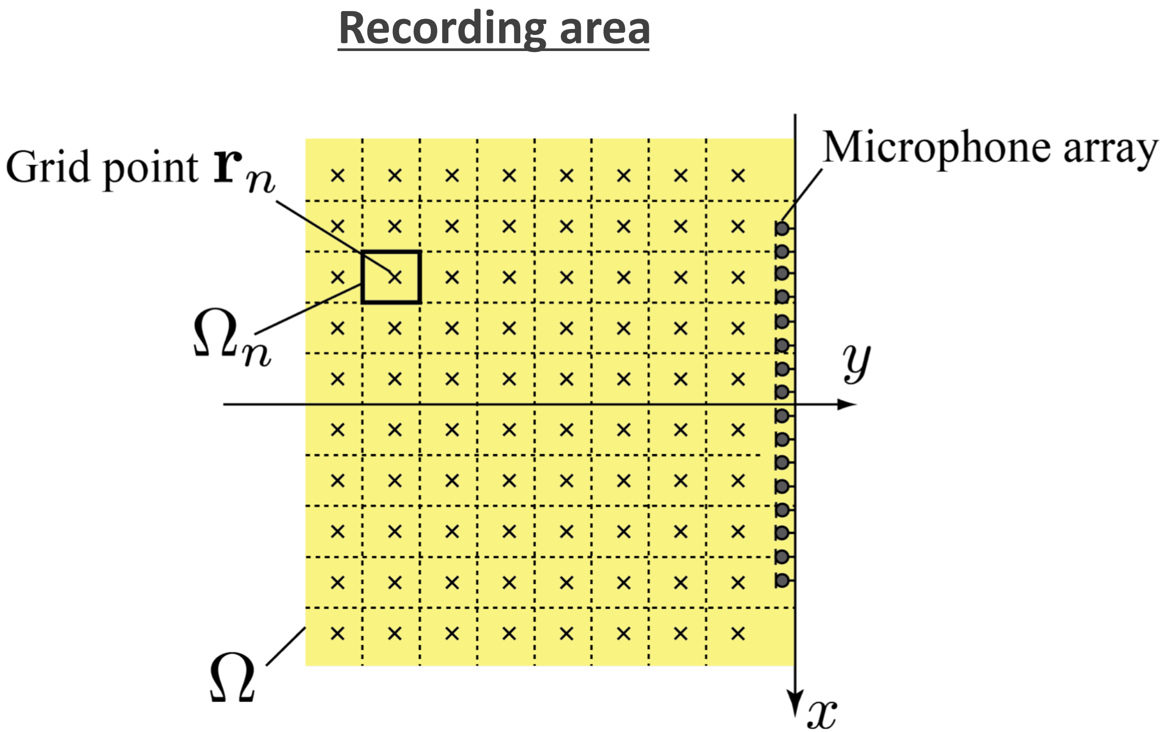
Experiments in Recording and Reproduction

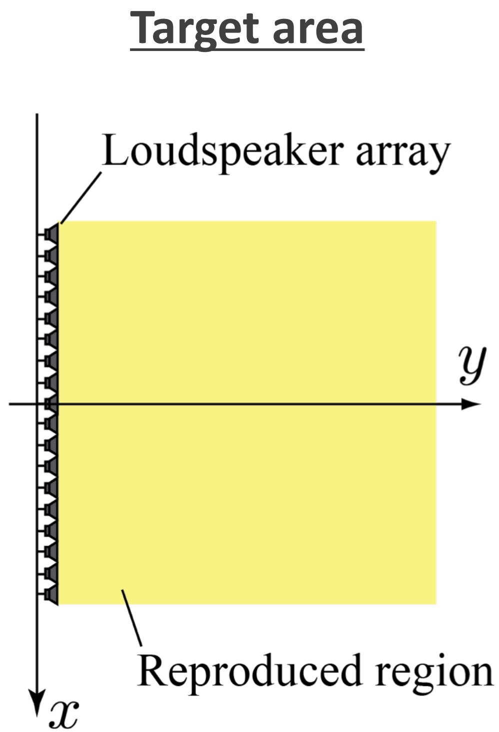




- Evaluation in sound field recording and reproduction using linear arrays
- Proposed method (**Proposed**) was compared with plane-wave-decomposition-based method (**WFR**) [Koyama+ 2013]
- 32 microphones (0.06 m intervals） and 48 loudspeakers (0.04 m intervals)
- Ω : Rectangular region of 2.4x2.4 m, Grid points: (0.01 m, 0.02 m) intervals

Experiments in Recording and Reproduction





- **Signal-to-distortion ratio of reproduction (SDRR)**

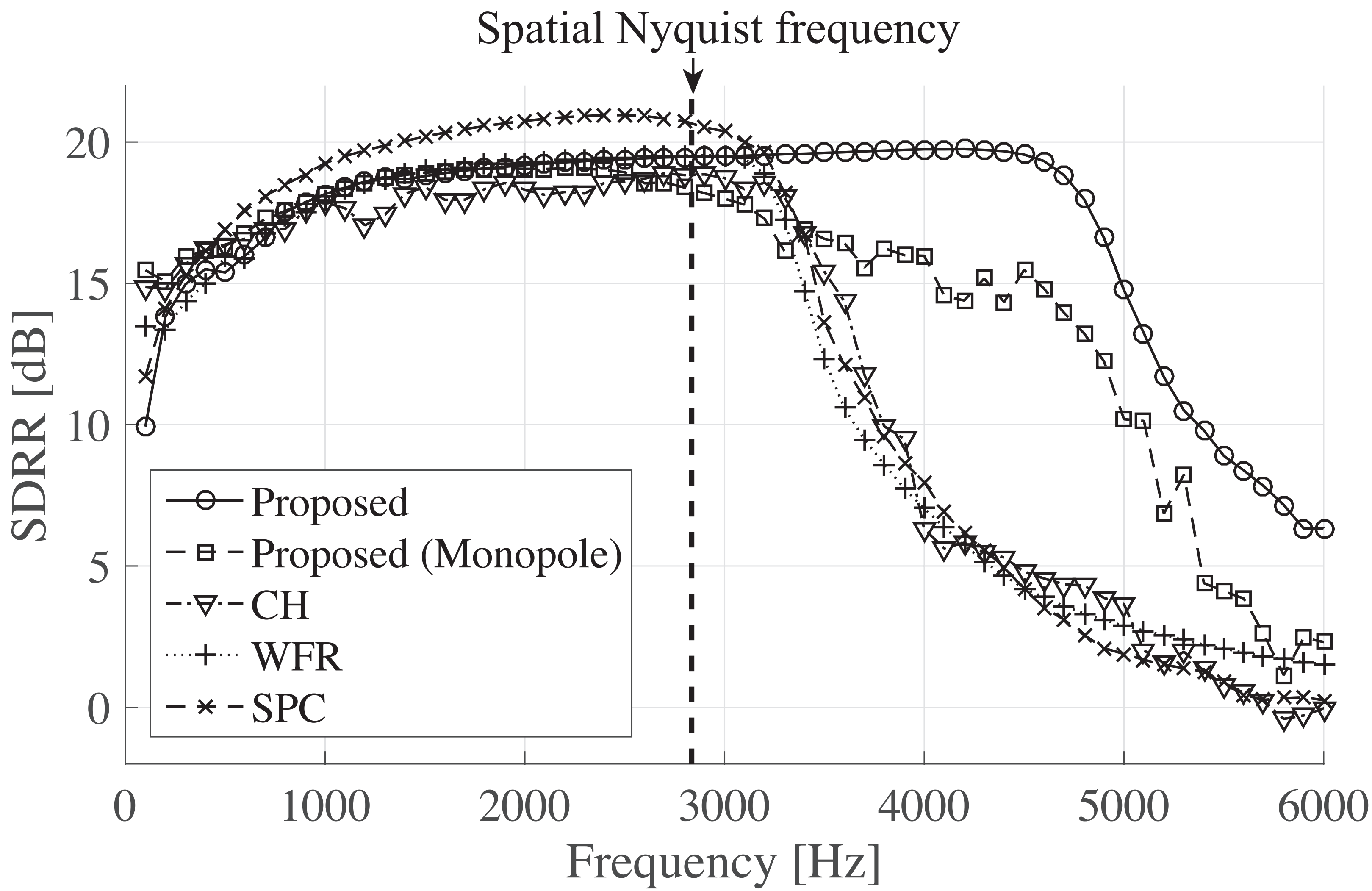
Desired pressure distribution

$$\mathrm{SDRR} = 10 \log_{10} \frac{\int |u_{\mathrm{des}}(\mathbf{r}, k)|^2 \, \mathrm{d}\mathbf{r}}{\int |u_{\mathrm{des}}(\mathbf{r}, k) - u_{\mathrm{syn}}(\mathbf{r}, k)|^2 \, \mathrm{d}\mathbf{r}}$$

Synthesized pressure distribution

Frequency vs. SDR

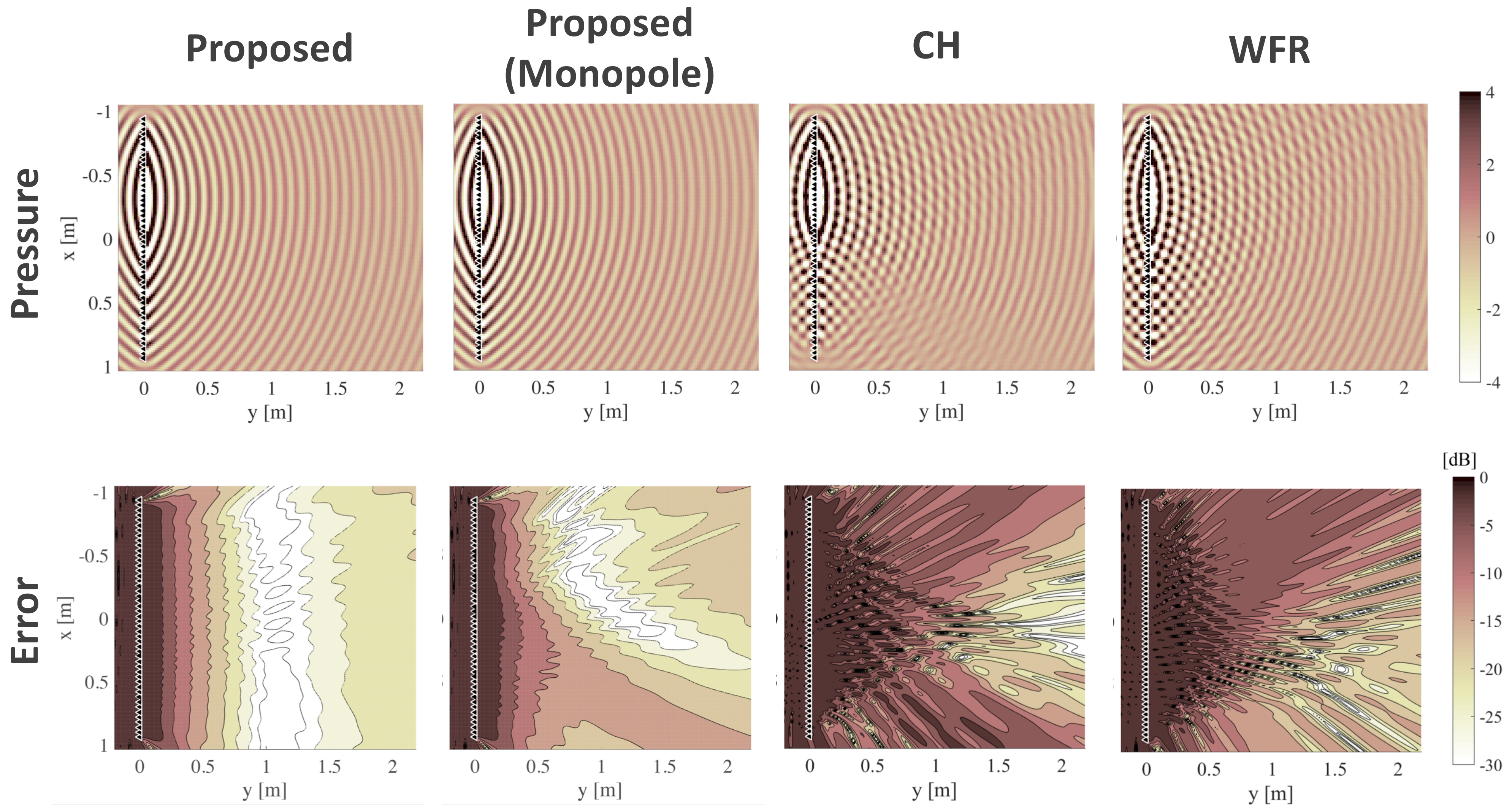
- Source location: (-0.32, -0.84, 0.0) m



SDRRs above spatial Nyquist frequency were improved

Reproduced Pressure Distribution (4.0 kHz)

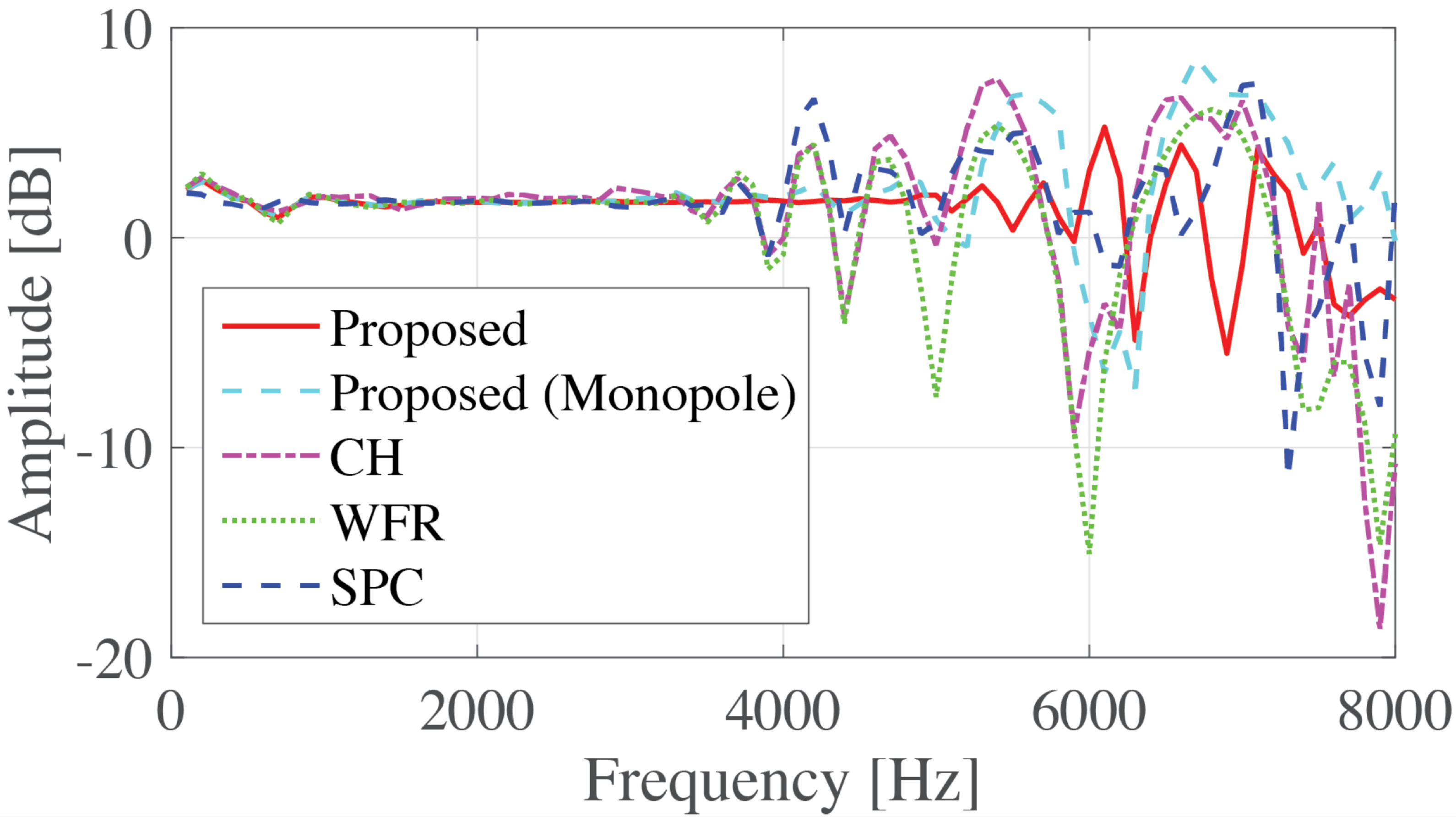
- Source location: (-0.32, -0.84, 0.0) m



SDRR: 19.7 dB 16.0 dB 6.3 dB 7.0 dB

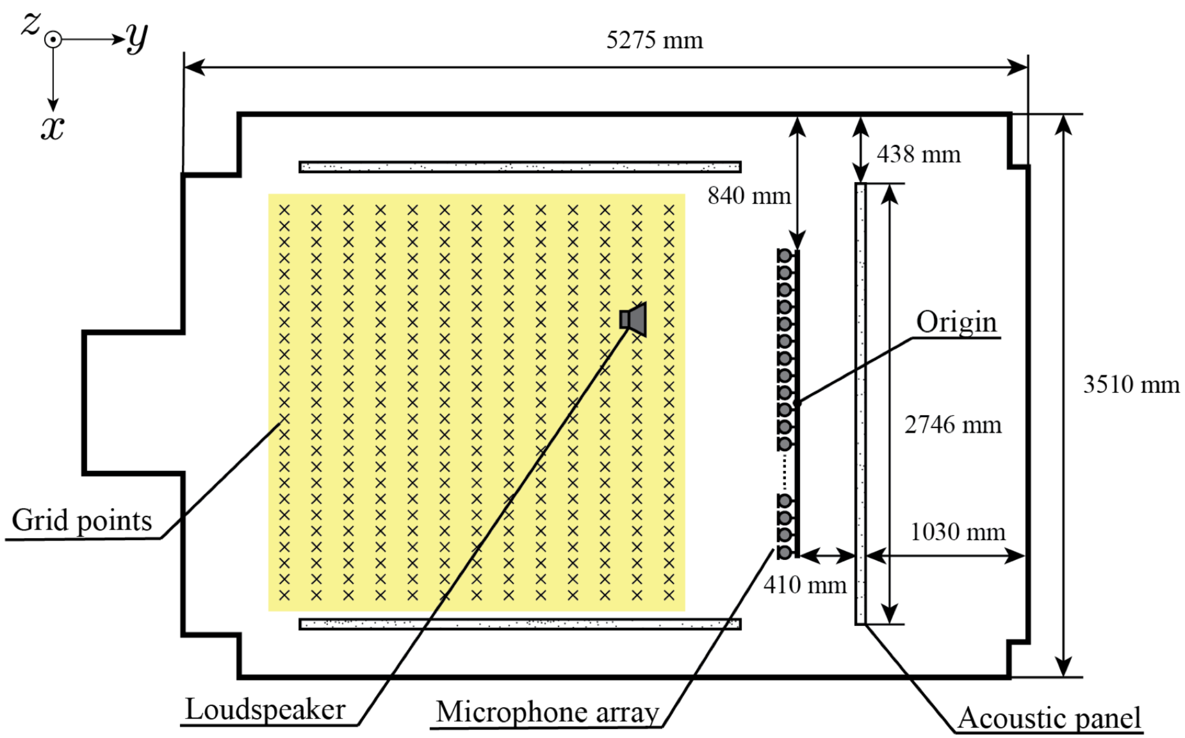
Frequency Response of Reproduced Sound Field

➢ Frequency response at (0.0, 1.0, 0.0) m



Reproduced frequency response was improved

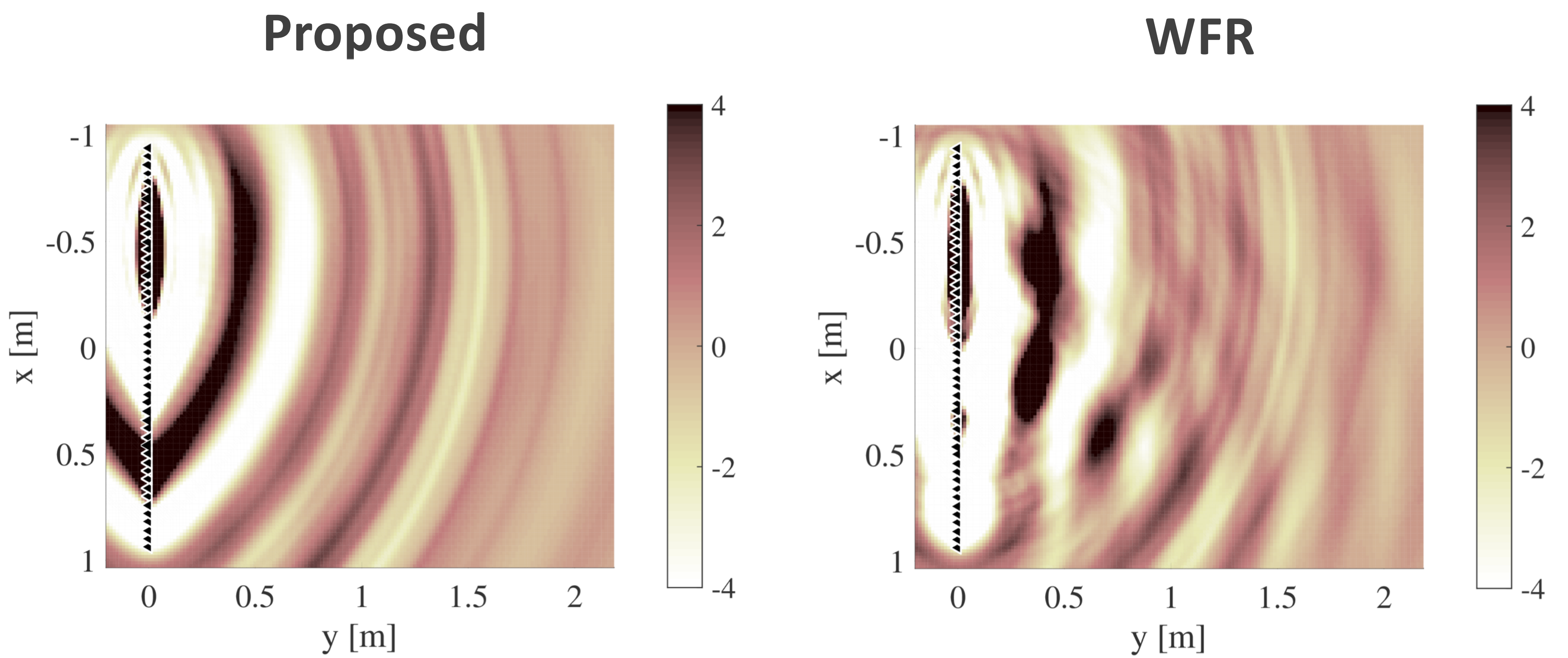
Experiments Using Real Data



- Recording area was real environment using loudspeaker as primary source
- Target area was simulated as free field
- Source signal: speech

Reproduced Pressure Distribution

- Loudspeaker at (-0.5, -1.0, 0.0) m, speech signal



Spatial aliasing artifacts are reduced by proposed method

Conclusion

- Sound field reconstruction inside source-free region
  - Decomposition into element solutions of Helmholtz eq.
  - Harmonic analysis of infinite orders
  - Sparse plane wave decomposition
- Sound field reconstruction inside region including sources
  - ill-posed problem - some constraints on source distribution is necessary
  - Sound field decomposition based on spatial sparsity of source distribution
  - Application to recording and reproduction and its experimental results

Thank you for your attention!

Related publications (1/2)

- N. Ueno, S. Koyama, and H. Saruwatari, “Kernel ridge regression with constraint of Helmholtz equation for sound field interpolation,” *Proc. IWAENC*, 2018.
- Y. Takida, S. Koyama, and H. Saruwatari. “Exterior and interior sound field separation using convex optimization: comparison of signal models,” *Proc. EUSIPCO*, 2018.
- Y. Takida, S. Koyama, et al. “Gridless sound field decomposition based on reciprocity gap functional in spherical harmonic domain,” *Proc. IEEE SAM*, 2018.
- N. Ueno, S. Koyama, and H. Saruwatari, “Sound field reproduction with exterior radiation cancellation using analytical weighting of harmonic coefficients,” *Proc. IEEE ICASSP*, 2018.
- S. Koyama, *et al.* “Sparse sound field decomposition for super-resolution in recording and reproduction,” *JASA*, 2018.
- N. Murata, S. Koyama, *et al.* “Sparse representation using multidimensional mixed-norm penalty with application to sound field decomposition,” *IEEE Trans. Signal Process.*, 2018.
- N. Ueno, S. Koyama, and H. Saruwatari, “Sound field recording using distributed microphones based on harmonic analysis of infinite order,” *IEEE Signal Process. Letters*, 2018.
- S. Koyama and L. Daudet, “Comparison of reverberation models for sparse sound field decomposition,” *Proc. IEEE WASPAA*, 2017.
- S. Koyama, et al. “Effect of multipole dictionary in sparse sound field decomposition for super-resolution in recording and reproduction,” *Proc. ICSV*, 2017.
- N. Ueno, S. Koyama, and H. Saruwatari, “Listening-area-informed sound field reproduction with Gaussian prior based on circular harmonic expansion,” *Proc. HSCMA*, 2017.
- N. Murata, S. Koyama, *et al.* “Spatio-temopral sparse sound field decomposition considering acoustic source signal characteristics,” *Proc. IEEE ICASSP*, 2017.
- N. Ueno, S. Koyama, and H. Saruwatari, “Listening-area-informed sound field reproduction based on circular harmonic expansion,” *Proc. IEEE ICASSP*, 2017.

Related publications (2/2)

- S. Koyama and H. Saruwatari, "Sound field decomposition in reverberant environment using sparse and low-rank signal models," *Proc. IEEE ICASSP*, 2016.
- N. Murata, S. Koyama, *et al.* "Sparse sound field decomposition with multichannel extension of complex NMF," *Proc. IEEE ICASSP*, 2016.
- S. Koyama, *et al*. "Structured sparse signal models and decomposition algorithm for super-resolution in sound field recording and reproduction," *Proc. IEEE ICASSP*, 2015.
- S. Koyama, *et al.* "Sparse sound field representation in recording and reproduction for reducing spatial aliasing artifacts," *Proc. IEEE ICASSP*, 2014.
- S. Koyama, *et al*. "Analytical approach to transforming filter design for sound field recording and reproduction using circular arrays with a spherical baffle," *JASA*, 2016.
- S. Koyama, *et al*. "Source-location-informed sound field recording and reproduction," *IEEE J. Sel. Topics Signal Process.*, 2015.
- J. Trevino, S. Koyama, *et al*. "Mixed-order ambisonics encoding of cylindrical microphone array signals," *Acoust. Sci. Tech.*, *Acoust. Letter*, 2014.
- S. Koyama, *et al*. "Wave field reconstruction filtering in cylindrical harmonic domain for with-height recording and reproduction," *IEEE/ACM Trans. ASLP*, 2014.
- S. Koyama, *et al*. "Real-time sound field transmission system by using wave field reconstruction filter and its evaluation," *IEICE Trans. Fundam.*, 2014.
- S. Koyama, *et al*. "Analytical approach to wave field reconstruction filtering in spatio-temporal frequency domain," *IEEE Trans. ASLP*, 2013.
- S. Koyama, *et al*. "Reproducing virtual sound sources in front of a loudspeaker array using inverse wave propagator," *IEEE Trans. ASLP*, 2012.