－ 信号処理特論 / Advanced Signal Processing －

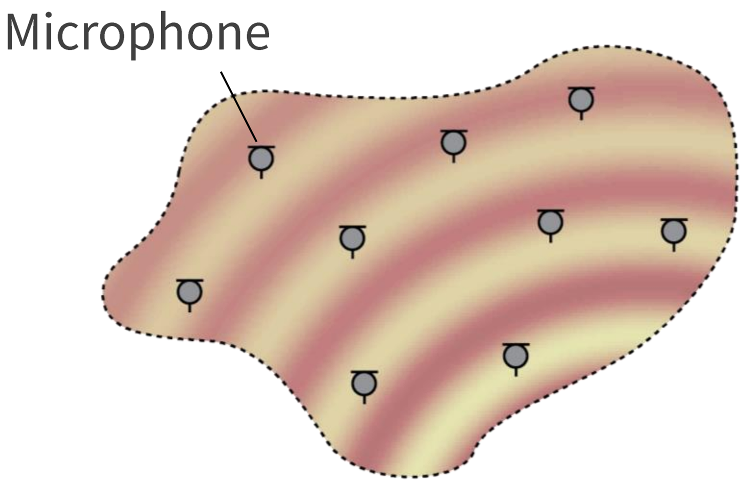
音場の計測と制御
Sound Field Analysis and Control

小山 翔一 / Shoichi Koyama

東京大学 大学院情報理工学系研究科
Graduate School of Information Science and Technology, UTokyo

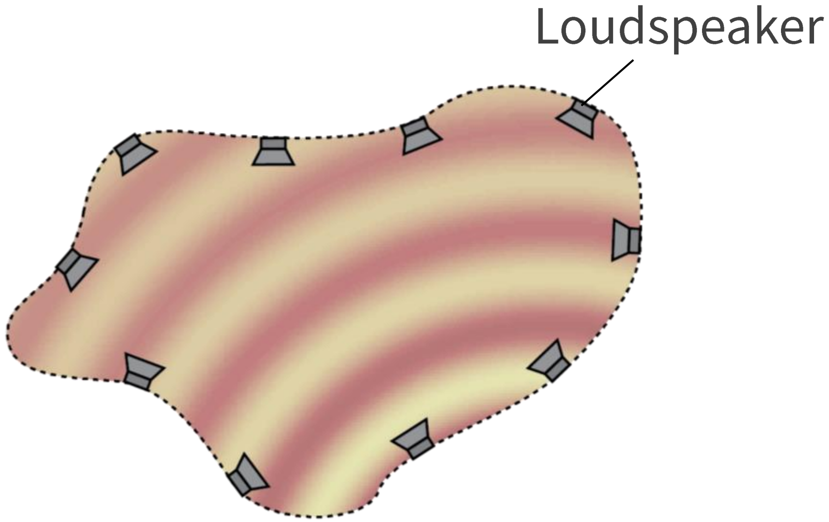
Analysis and control of acoustic field

Analysis



- Visualization and reconstruction of acoustic field
- Estimation of source locations and room-acoustic parameters
- Spatial sound field recording

Control

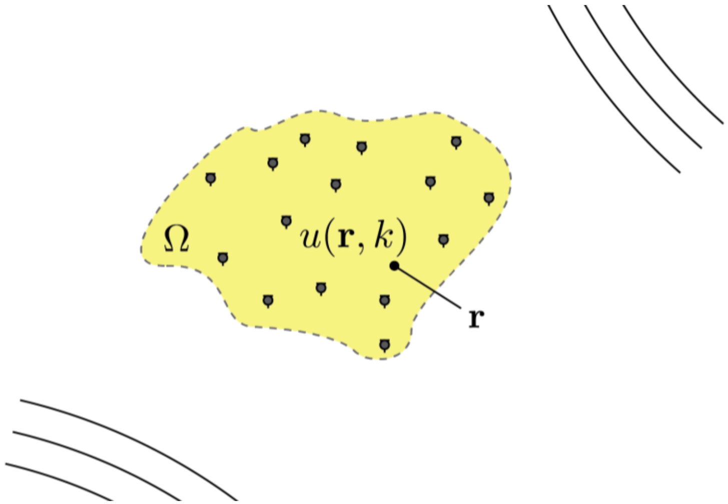


- High-fidelity spatial audio reproduction
- Directivity control and local reproduction
- Spatial active noise control

From theory to application of signal processing and inverse problems for acoustic fields

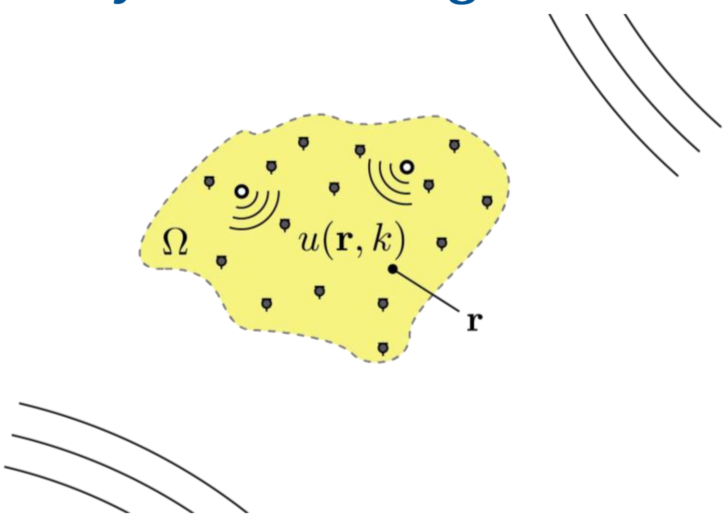
Summary of sound field analysis

Analysis inside region without sources



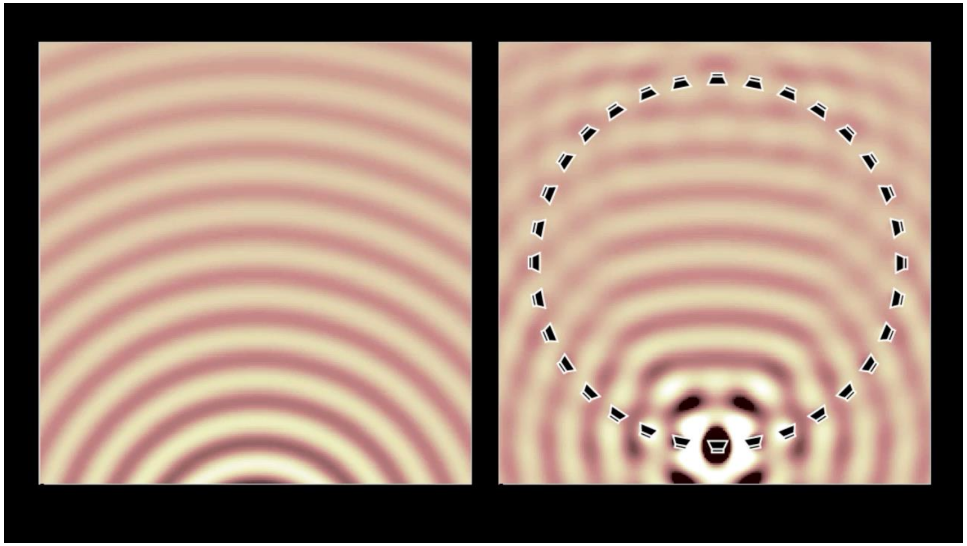
- Sound field reconstruction based on harmonic analysis of infinite order [Ueno+ IEEE SPL 2018]
- Optimization algorithm for sparse representation of acoustic field [Murata+ IEEE TSP 2018]

Analysis inside region including sources



- Sparse sound field decomposition in reverberant environment [Koyama+ IEEE JSTSP 2019]
- Estimation of source parameters based on Reciprocity Gap Functional [Takida+ Elsevier SP 2020]
- Separation of internal and external sound fields [Takida+ EUSIPCO 2018]

Summary of sound field control



- Sound field recording and reproduction in wave-number domain [Koyama+ IEEE(/ACM) TASLP 2013, 2014, JASA 2016]
- Super-resolution in recording and reproduction [Koyama+ IEEE JSTSP 2015, JASA 2018]
- Sound field control based on weighted mode-matching [Ueno+ IEEE/ACM TASLP 2019]
- Optimization of source and sensor placement for sound field control [Koyama+ IEEE/ACM TASLP 2020, IEEE ICASSP 2018]
- Spatial active noise control based on kernel interpolation [Ito+ IEEE ICASSP 2019, 2020]

Today's topics

- Sparse modeling and its application to acoustic signal processing
- Optimal source and sensor placement for sound field cotnrol

SPARSE MODELING AND ITS APPLICATION TO ACOUSTIC SIGNAL PROCESSING

Linear inverse problem

- Suppose that measurement $\mathbf{y} \in \mathbb{R}^M$ is modeled by linear equation with unknown variable $\mathbf{x} \in \mathbb{R}^N$ and sensing matrix $\mathbf{D} \in \mathbb{R}^{M \times N}$ as

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

  where $\mathbf{n} \in \mathbb{R}^M$ is additive noise.
- Estimation problem of $\mathbf{x}$ with given $\mathbf{y}$ and $\mathbf{D}$ is referred to as linear inverse problem.
- Consider the case of $N > M$, i.e., underdetermined problem. This type of problem (normally) has infinitely many solutions, which means preferable features should be imposed on the estimate.

Least-norm solution

- Typical approach to solve underdetermined linear inverse problem is least-norm solution (a.k.a. minimum-norm solution), where the following optimization problem is considered:

$$\underset{\mathbf{x}}{\text{minimize}} \, \|\mathbf{x}\|_2^2 \quad \text{subject to} \quad \mathbf{y} = \mathbf{D}\mathbf{x}$$

- This problem can be solved by the method of Lagrange multiplier as

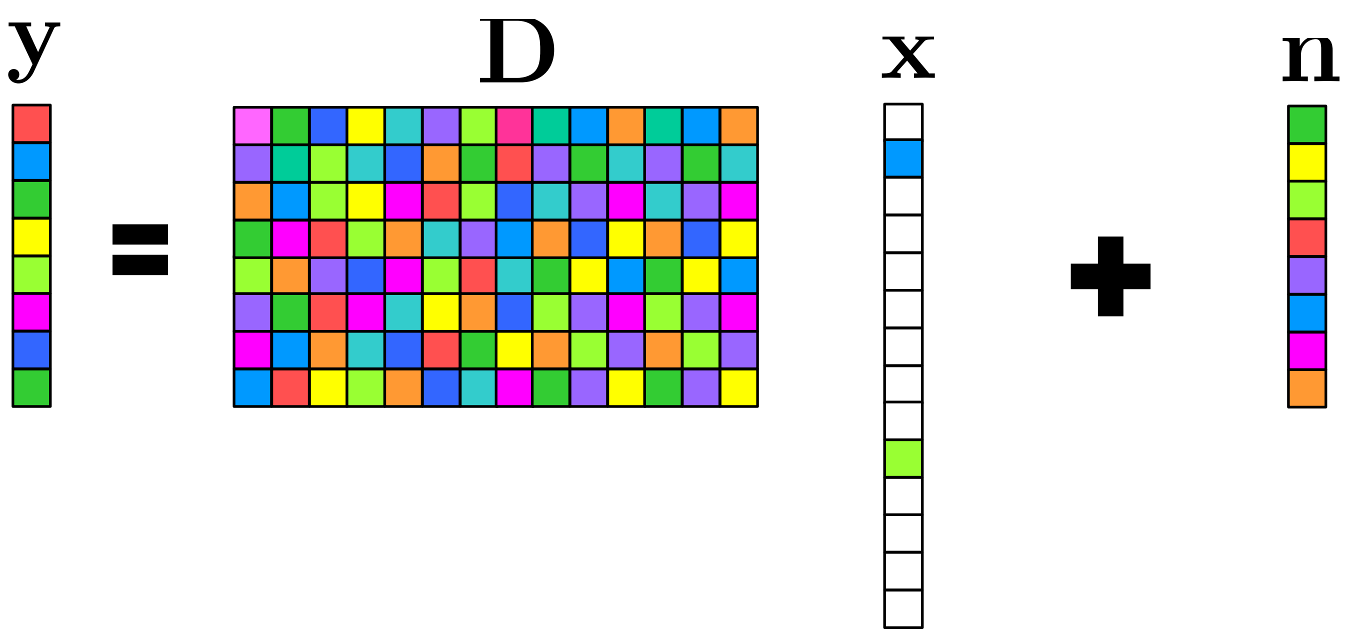
$$\hat{\mathbf{x}} = \mathbf{D}^\mathsf{T} \left(\mathbf{D}\mathbf{D}^\mathsf{T}\right)^{-1} \mathbf{y}$$

  – Regularized solution to increase robustness:

$$\hat{\mathbf{x}} = \mathbf{D}^\mathsf{T} \left(\mathbf{D}\mathbf{D}^\mathsf{T} + \lambda \mathbf{I}\right)^{-1} \mathbf{y}$$

Sparse modeling

- Sometimes we would like to impose sparsity on the estimate.



- *Occam's razor (law of parsimony):*
  - *Entities should not be multiplied without necessity.*

Example: images in wavelet domain

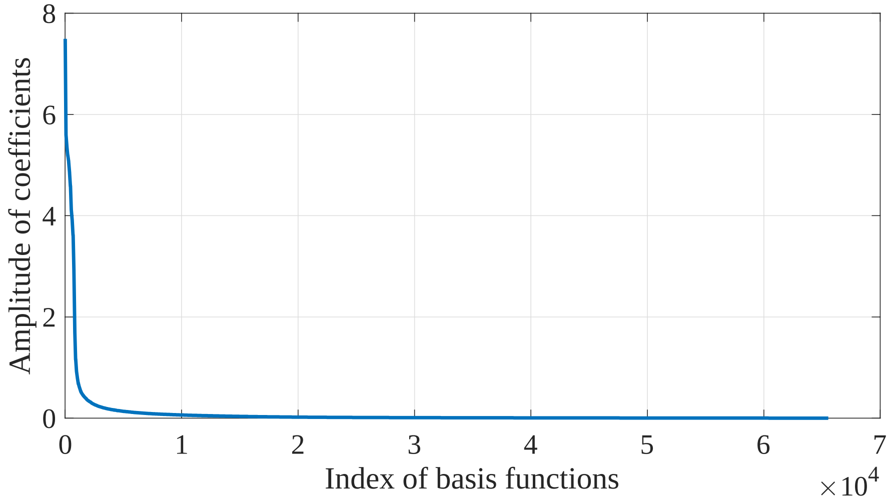
- Representing image with small number of coefficients

Original

10% of coefs

3% of coefs

- Distribution of coefficients is sparse:



Sparsity-inducing norm

- Least-norm solution:

$$\underset{\mathbf{x}}{\text{minimize}} \, \|\mathbf{x}\|_2^2 \quad \text{subject to} \quad \mathbf{y} = \mathbf{D}\mathbf{x}$$

- Solution with sparsity-inducing norm:

$$\underset{\mathbf{x}}{\text{minimize}} \, \|\mathbf{x}\|_p \quad \text{subject to} \quad \mathbf{y} = \mathbf{D}\mathbf{x} \qquad (0 \leq p \leq 1)$$

  – where

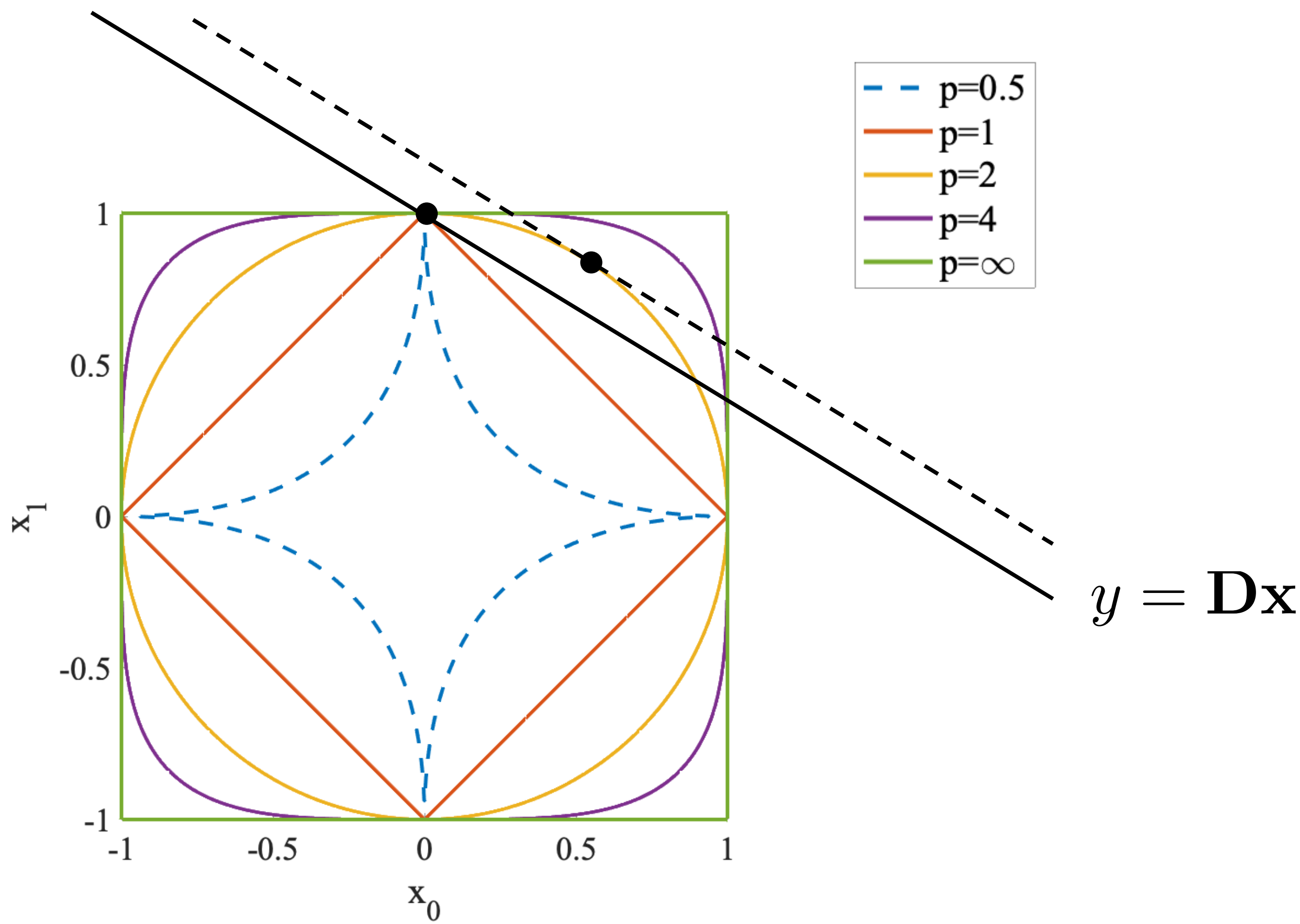
$$\|\mathbf{x}\|_p = \begin{cases} \left(\sum_{n=1}^{N} |x_n|^p \right)^{\frac{1}{p}}, & 0 < p \leq 1 \\ \lim_{p \to 0} \sum_{n=1}^{N} |x_n|^p, & p = 0 \end{cases}$$

Counting the number of nonzero elements

  – $\|\cdot\|_p$ is called ℓ_p -norm whereas axioms of norm is not valid for $0 \leq p < 1$

Sparsity-inducing norm

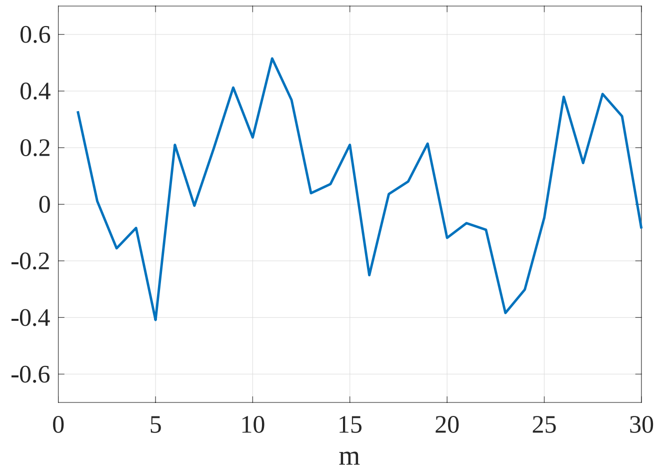
➢ Contour of $\|\mathbf{x}\|_p = c$ with constant c in 2D case.



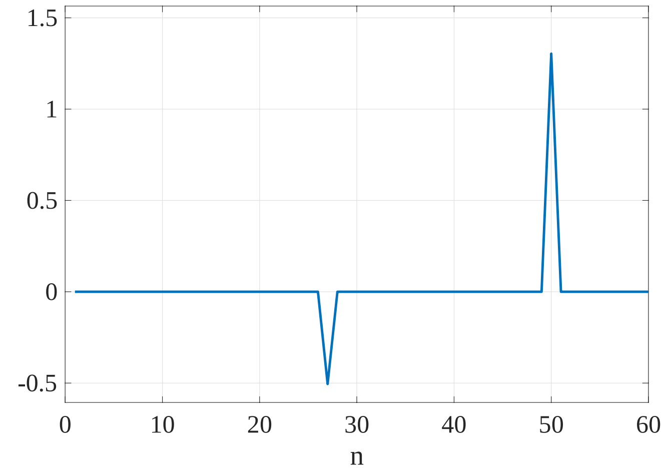
Example of sparse solution

- Basis pursuit for ℓ_1-norm minimization problem

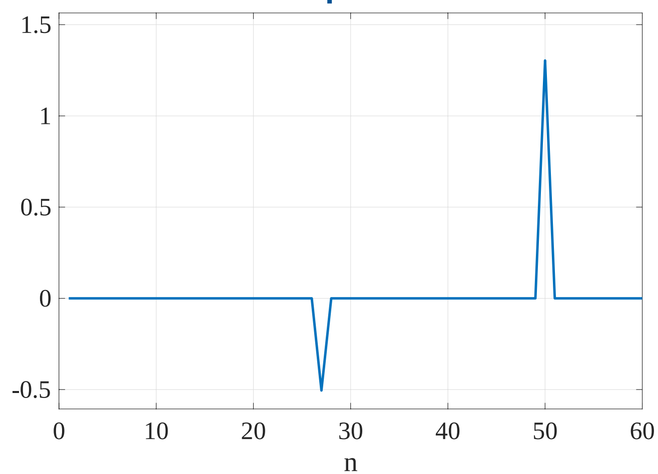
Observation



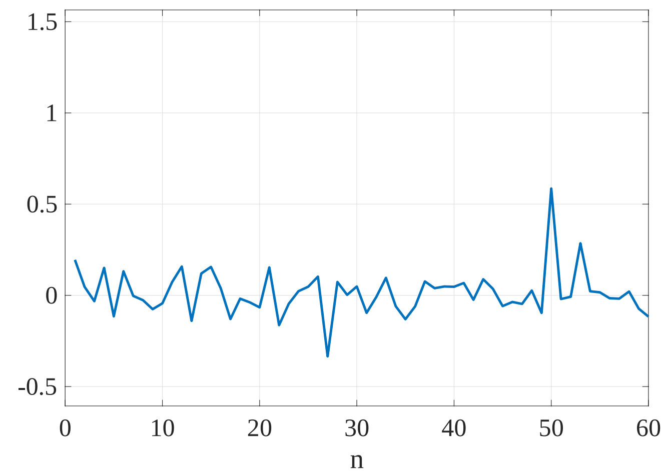
True



Basis pursuit



Least-norm



How to solve sparse optimization problem?

- Algorithms for solving sparse optimization problem can be classified into three categories:
  - Greedy algorithm
    - (Orthogonal) matching pursuit, etc···
  - Convex relaxation
    - Basis pursuit, (Accelerated) proximal gradient, etc···
  - Majorization-minimization algorithm
    - Iteratively-reweighted least squares, etc···
  - Probabilistic inference
    - Sparse Bayesian learning, etc···

Majorization-minimization algorithm

- Construct surrogate function $\mathcal{L}^{+}(\mathbf{x}, \boldsymbol{\xi})$ for (non-convex) objective function $\mathcal{L}(\mathbf{x})$

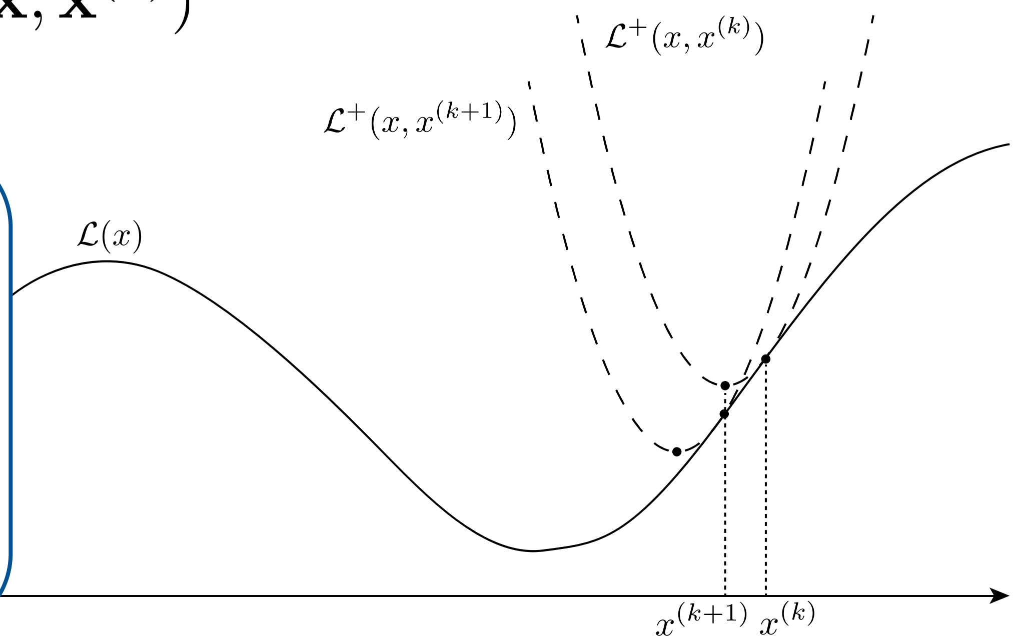
$$\mathcal{L}(\mathbf{x}) \leq \mathcal{L}^{+}(\mathbf{x}, \boldsymbol{\xi})$$
$$\mathcal{L}(\mathbf{x}) = \mathcal{L}^{+}(\mathbf{x}, \mathbf{x})$$

- Alternately updating parameter of surrogate function and parameter to be optimized

$$\mathbf{x}^{(k+1)} = \arg\min_{\mathbf{x}} \mathcal{L}^{+}(\mathbf{x}, \mathbf{x}^{(k)})$$

Monotonic non-increase of objective function is guaranteed

$$\begin{aligned}\mathcal{L}(\mathbf{x}^{(k+1)}) &\leq \mathcal{L}^{+}(\mathbf{x}^{(k+1)}, \mathbf{x}^{(k)}) \\ &\leq \mathcal{L}^{+}(\mathbf{x}^{(k)}, \mathbf{x}^{(k)}) \\ &= \mathcal{L}(\mathbf{x}^{(k)})\end{aligned}$$

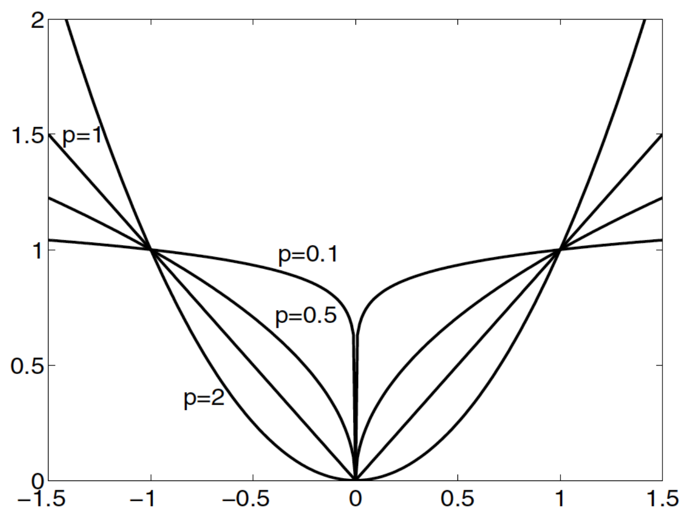


ℓ_p -regularization problem

- MM algorithm for ℓ_p-regularization problem

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

- Sparsity of $\mathbf{x}$ can be induced by the regularization term, but the objective function becomes non-convex for $0 < p < 1$

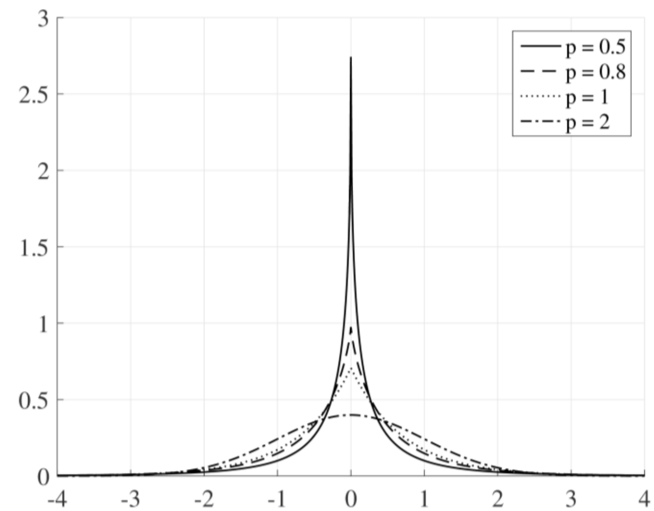


Plot of $|x|^p$ [Elad 2010]

ℓ_p -regularization problem

- 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

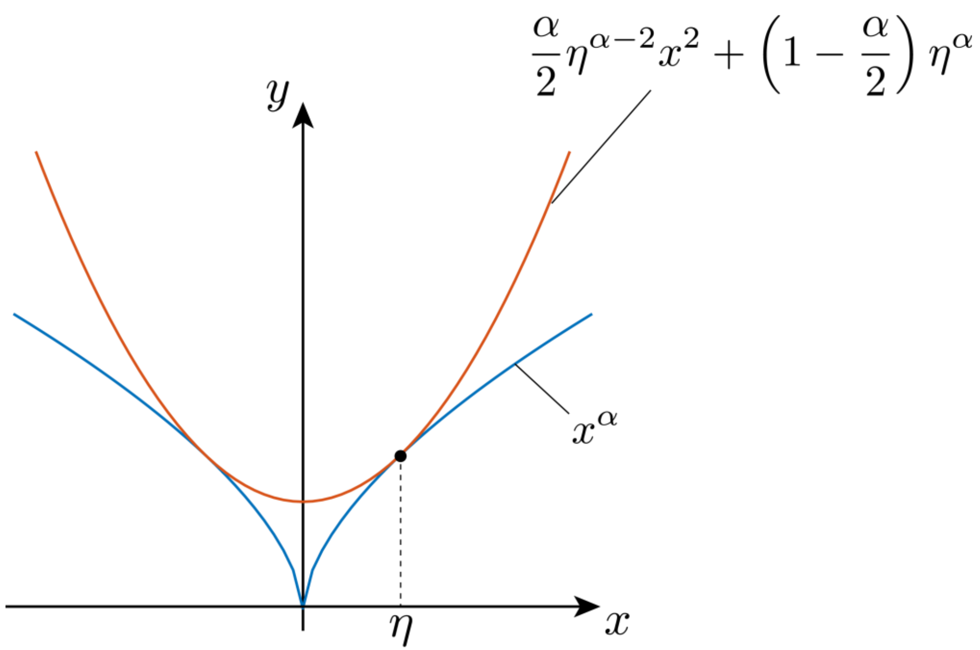
➡ $$\mathbf{x}_{\mathrm{MAP}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \frac{\sigma^2}{\beta^p}\sum_n |x_n|^p$$

Identical to ℓ_p-norm penalty

ℓ_p -regularization problem

- Consider to develop surrogate function of ℓ_p-regularization term for $0 < p \leq 1$
- Concave function x^α $(0 < \alpha \leq 1)$ lies below tangent quadratic function x^2 at η; therefore,

$$\|\mathbf{x}\|_p^p = \sum_{n=1}^{N} |x_n|^p \leq \sum_{n=1}^{N} \left\{ \frac{p}{2}\eta_n^{p-2} x_n^2 + \left(1 - \frac{p}{2}\eta_n^2\right) \right\}$$



ℓ_p -regularization problem

- Surrogate function is developed as

$$\begin{aligned}\mathcal{L}(\mathbf{x}) &= \frac{1}{2}\|\mathbf{y}-\mathbf{D}\mathbf{x}\|_2^2 + \lambda\sum_{n=1}^{N}|x_n|^p \\ &\leq \frac{1}{2}\|\mathbf{y}-\mathbf{D}\mathbf{x}\|_2^2 + \lambda\frac{p}{2}\sum_{n=1}^{N}\eta_n^{p-2}x_n^2 + C \\ &:= \mathcal{L}^+(\mathbf{x},\boldsymbol{\eta})\end{aligned}$$

  – C is variable not related to optimization, $\boldsymbol{\eta} = [\eta_1, \ldots, \eta_N]^\mathsf{T}$, and equality holds for $\mathbf{x} = \boldsymbol{\eta}$.

- Update rule of $\mathbf{x}$

$$\mathbf{x}^{(k+1)} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y}-\mathbf{D}\mathbf{x}\|_2^2 + \frac{1}{2}\lambda\mathbf{x}^\mathsf{T}\mathbf{P}^{(k)}\mathbf{x}$$

$$\left(\mathbf{P}^{(k)}\right)_{n,n'} = \begin{cases} p\left(x_n^{(k)}\right)^{p-2}, & n = n' \\ 0, & \text{otherwise}\end{cases}$$

ℓ_p -regularization problem

- This minimization problem is simply solved as weighted least-squares solution:

$$\mathbf{x}^{(k+1)} = \left(\mathbf{D}^{\mathsf{T}}\mathbf{D} + \lambda \mathbf{P}^{(k)}\right)^{-1} \mathbf{D}^{\mathsf{T}}\mathbf{y}$$

- Rewrite with $\mathbf{W}^{(k)} = (\mathbf{P}^{(k)})^{-1/2}$ and $\mathbf{A}^{(k)} = \mathbf{D}\mathbf{W}^{(k)}$ by using matrix inversion lemma as

$$\mathbf{x}^{(k+1)} = \mathbf{W}^{(k)}\mathbf{A}^{(k)\mathsf{T}}\left(\mathbf{A}^{(k)}\mathbf{A}^{(k)\mathsf{T}} + \lambda \mathbf{I}\right)^{-1} \mathbf{y}$$

  – Inverse of $N \times N$ matrix is turned into inverse of $M \times M$ matrix. Besides, elements of $\mathbf{W}^{(k)}$ are stably computed compared to $\mathbf{P}^{(k)}$.

Iteratively-reweighted least-squares algorithm

- MM algorithm for ℓ_p-regularization problem is called iteratively-reweighted least-squares or focal underdetermined system solver (FOCUSS)

[Gorodnitsky+ 1997, Figueiredo+ 2007]

- Summary of algorithm:

Set initial value $\mathbf{x}^{(0)}$, then repeat

- Update $\mathbf{W}^{(k)} = \mathrm{diag}\left(p\left(x_n^{(k)}\right)^{p-2}\right)$
- Update $\mathbf{A}^{(k)} = \mathbf{D}\mathbf{W}^{(k)}$
- Update $\mathbf{x}^{(k+1)} = \mathbf{W}^{(k)}\mathbf{A}^{(k)\mathsf{T}}\left(\mathbf{A}^{(k)}\mathbf{A}^{(k)\mathsf{T}} + \lambda\mathbf{I}\right)^{-1}\mathbf{y}$

Beyond sparsity

- Various signal structures other than sparsity can be induced by vector or matrix norms.
- Signal separation by using constraint of such norms is well studied in the context of convex optimization [McKoy+ 2014].

Structure	Atomic gauge
Sparse vector	ℓ^1-norm
Binary sign vector	ℓ^∞-norm
Low-rank matrix	Nuclear norm / Schatten 1-norm
Orthogonal matrix	Schatten ∞-norm
Row-sparse matrix	Row-ℓ^1-norm

Proximal gradient method

- Consider the following optimization problem:

$$\underset{\mathbf{x}}{\text{minimize}}\, f(\mathbf{x}) + g(\mathbf{x})$$

  - where $f, g : \mathbb{R}^p \to \mathbb{R} \cup \{\infty\}$ are proper convex lower semicontinuous function, which are denoted as $f, g \in \Gamma_0(\mathbb{R}^p)$, and f is differentiable.
  - This type of problem can be solved by proximal gradient method.

Set initial value $\mathbf{x}^{(0)},\ \gamma > 0$, then repeat

- Update $\mathbf{x}^{(k+1)} = \text{prox}_{\gamma g}\left(\mathbf{x}^{(k)} - \gamma \nabla f(\mathbf{x}^{(k)})\right)$

Proximal operator:

$$\text{prox}_{\gamma f}(\mathbf{v}) = \underset{\mathbf{x} \in \text{dom}(f)}{\arg\min} \left\{ f(\mathbf{x}) + \frac{1}{2\gamma}\|\mathbf{x} - \mathbf{v}\|_2^2 \right\}$$

- Consider the following optimization problem:

$$\underset{\mathbf{x},\mathbf{z}}{\text{minimize}}\, f(\mathbf{x}) + g(\mathbf{z}) \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{z}$$

  - where $f \in \Gamma_0(\mathbb{R}^p),\ g \in \Gamma_0(\mathbb{R}^q)$, and $\mathbf{A} \in \mathbb{R}^{M \times p},\ \mathbf{B} \in \mathbb{R}^{M \times q}$
  - This type of problem can be solved by alternating direction method of multipliers (ADMM) [Boyd+ 2011]

Set initial value $\mathbf{z}^{(0)}, \boldsymbol{\theta}^{(0)}$, then repeat

- Update $\mathbf{x}^{(k+1)} = \arg\min_{\mathbf{x}} \left\{ f(\mathbf{x}) + \frac{\rho}{2} \left\| \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{z}^{(k)} - \mathbf{y} + \boldsymbol{\theta}^{(k)} \right\|_2^2 \right\}$
- Update $\mathbf{z}^{(k+1)} = \arg\min_{\mathbf{z}} \left\{ g(\mathbf{x}) + \frac{\rho}{2} \left\| \mathbf{A}\mathbf{x}^{(k+1)} + \mathbf{B}\mathbf{z} - \mathbf{y} + \boldsymbol{\theta}^{(k)} \right\|_2^2 \right\}$
- Update $\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} + \rho \left(\mathbf{A}\mathbf{x}^{(k+1)} + \mathbf{B}\mathbf{z}^{(k+1)} - \mathbf{y} \right)$

Primal-dual splitting method

- Consider the following optimization problem:

$$\underset{\mathbf{x}}{\text{minimize}}\, f(\mathbf{x}) + g(\mathbf{x}) + h(\mathbf{A}\mathbf{x})$$

  - where $f, g \in \Gamma_0(\mathbb{R}^p)$ and $h \in \Gamma_0(\mathbb{R}^M)$, and f is differentiable.
  - This type of problem can be solved by primal-dual splitting method [Condat 2013, Vu 2013]

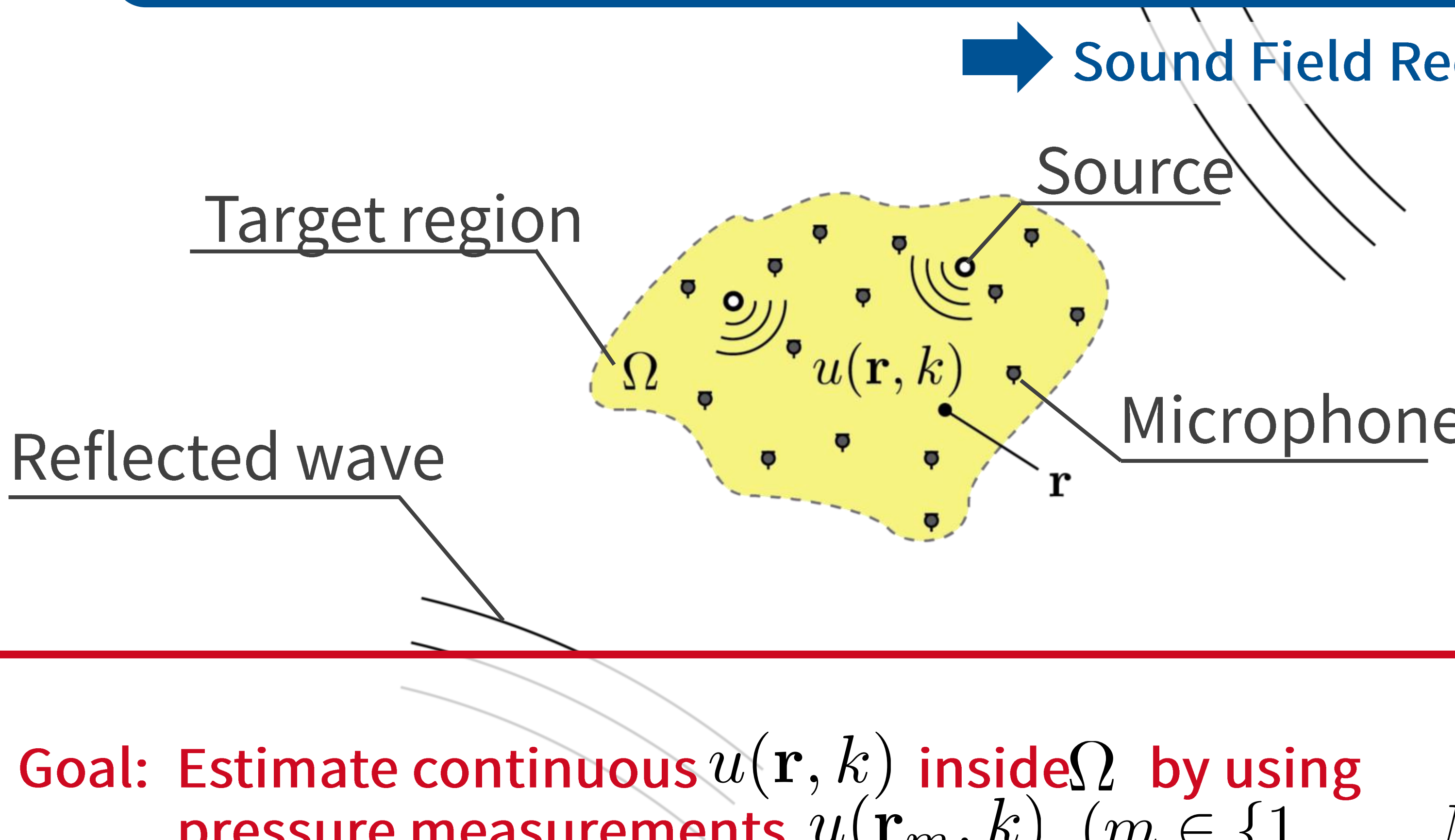
Set initial value $\mathbf{x}^{(0)}, \mathbf{z}^{(0)}, \gamma_1, \gamma_2 > 0$, then repeat

- Update $\mathbf{x}^{(k+1)} = \mathrm{prox}_{\gamma_1 g}\left(\mathbf{x}^{(k)} - \gamma_1\left(\nabla f(\mathbf{x}^{(k)}) + \mathbf{A}^\top \mathbf{z}^{(k)}\right)\right)$
- Update $\mathbf{z}^{(k+1)} = \mathrm{prox}_{\gamma_2 h^*}\left(\mathbf{z}^{(k)} + \gamma_2 \mathbf{A}\left(2\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\right)\right)$

Sound field reconstruction

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

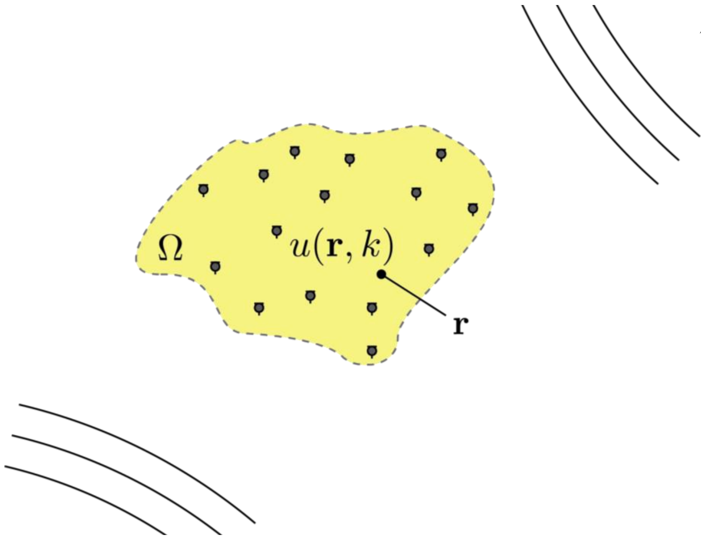
➡ Sound Field Reconstruction



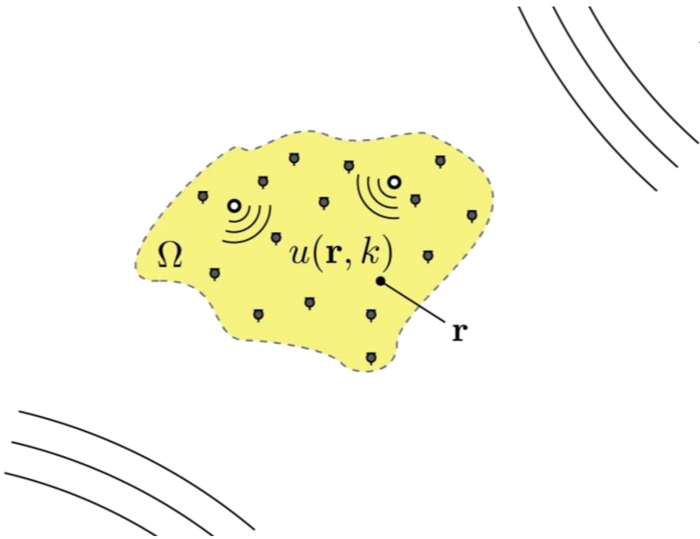
Goal: Estimate continuous $u(\mathbf{r}, k)$ inside Ω by using pressure measurements $u(\mathbf{r}_m, k)$ $(m \in \{1, \ldots, M\})$

➡ Visualization, reproduction by loudspeakers/headphones etc⋯

Sound field reconstruction

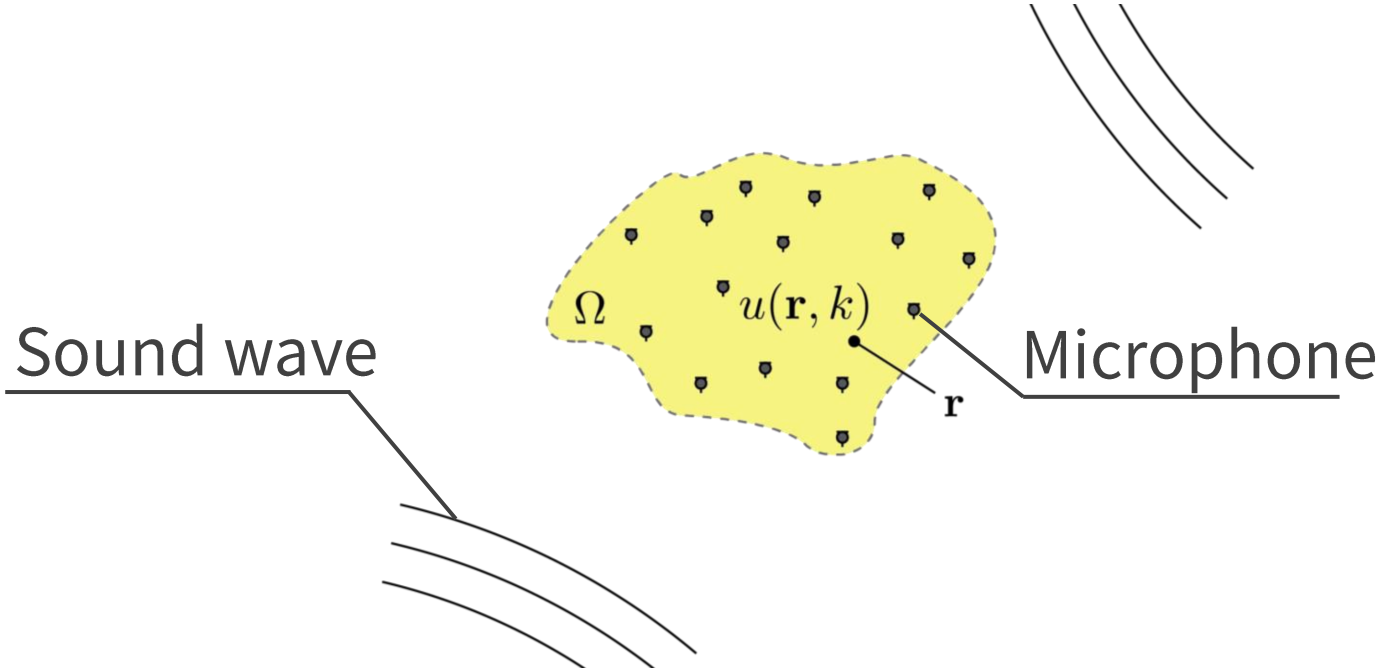


- **Target region does NOT include any sources**
  - Interpolation with constraint of homogeneous Helmholtz eq.
  - Decomposition of captured sound field into plane-wave or harmonic functions: **sound field decomposition**



- **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.**

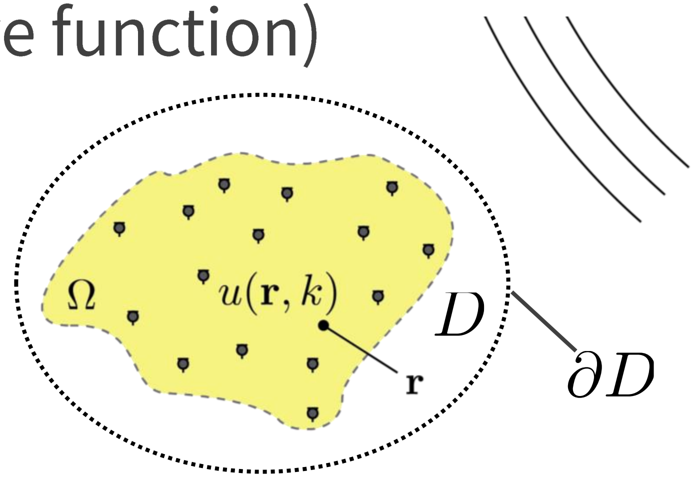
$$\begin{cases} (\nabla^2 + k^2) u(\mathbf{r}, k) = 0 \\ \text{Unknown boundary condition on room surface} \end{cases}$$

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]

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]

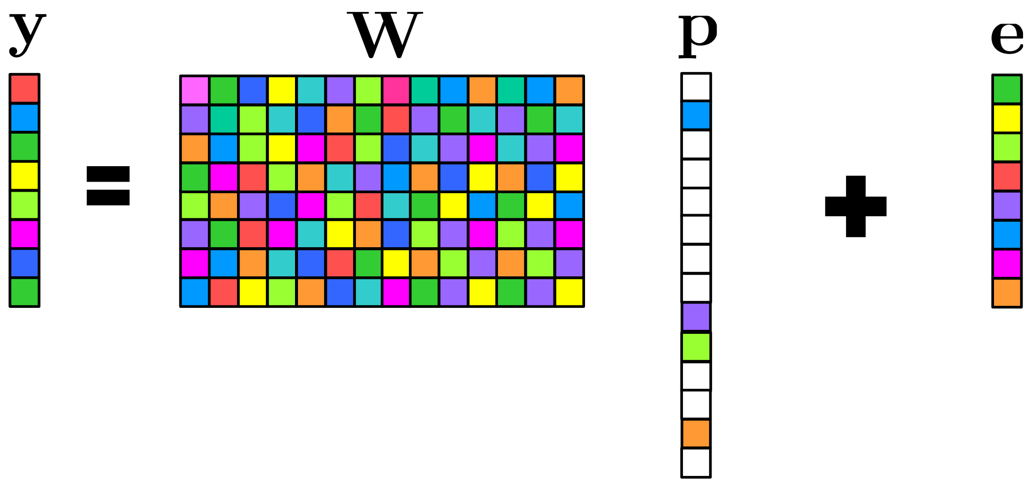
- Matrix form by using dictionary matrix $\mathbf{W} \in \mathbb{C}^{M \times L}$ consisting of plane-wave functions

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

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

Sparse plane-wave decomposition

- Sparse approximation by plane-wave dictionary matrix



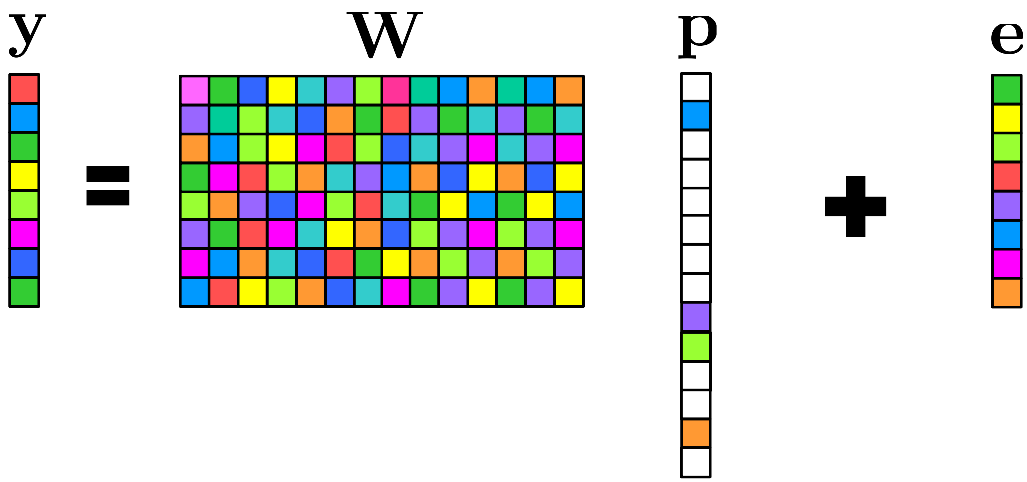
- Optimization problem for sparse approximation

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

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

Sparse plane-wave decomposition

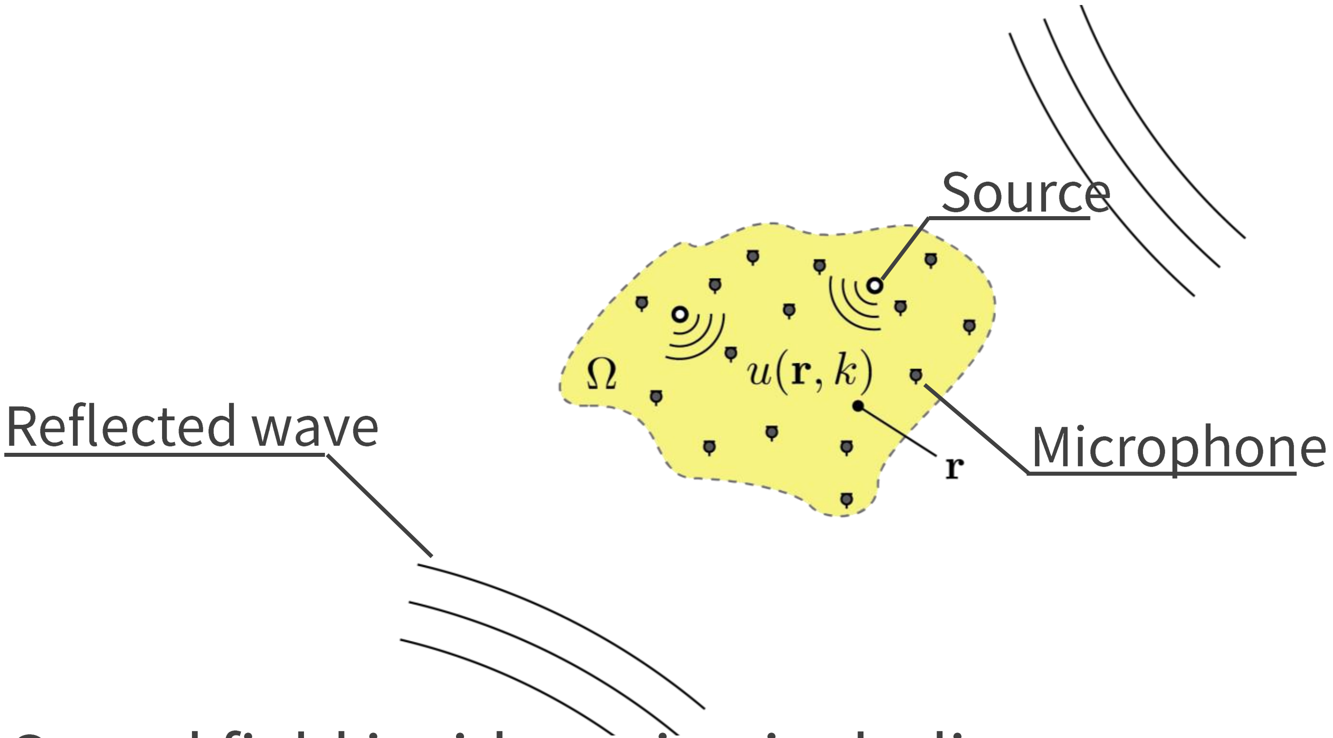
- Sparse approximation by plane-wave dictionary matrix



➡ **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.

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

Source distribution

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 $u(\mathbf{r})$:

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

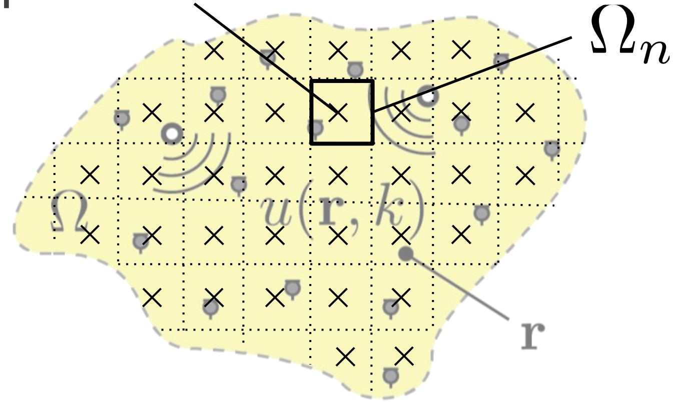
➡ Estimate $u(\mathbf{r})$ and $Q(\mathbf{r})$ from measurements $u(\mathbf{r}_m)$

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

Sparse sound field decomposition

[Koyama+ JASA 2018]

- Discretization of region Ω



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

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

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

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

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

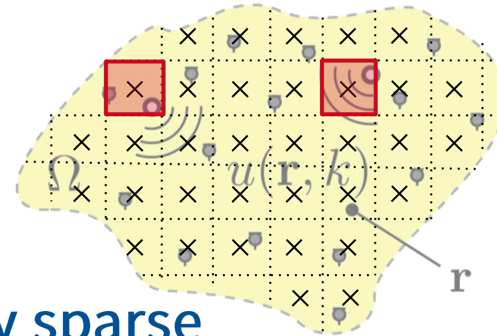
$$\begin{cases} \mathbf{y} = \left[u(\mathbf{r}_1), \ldots, u(\mathbf{r}_M)\right]^{\mathsf{T}} \\ \mathbf{x} = \left[\int_{\Omega_1} Q(\mathbf{r}') \mathrm{d}\mathbf{r}', \ldots, \int_{\Omega_N} Q(\mathbf{r}') \mathrm{d}\mathbf{r}'\right]^{\mathsf{T}} \end{cases}$$

Sparse sound field decomposition

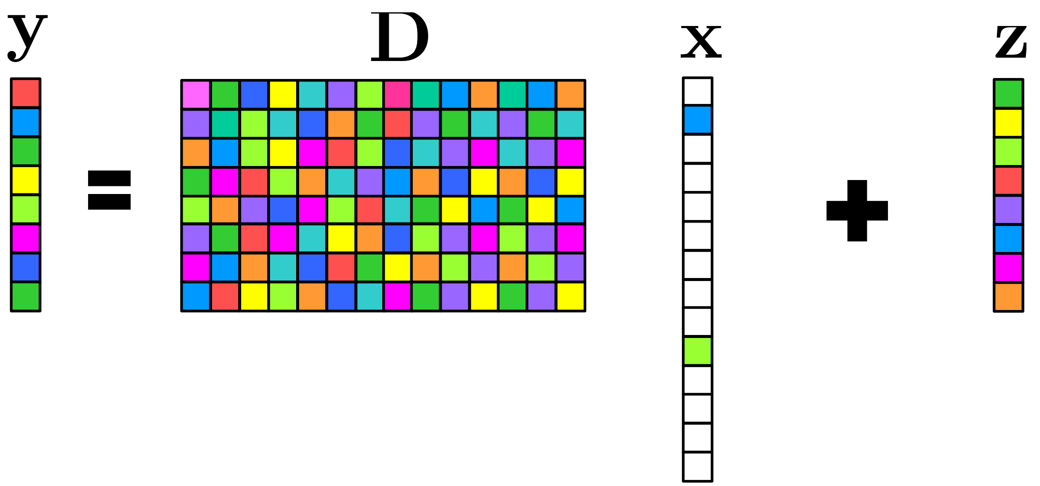
[Koyama+ JASA 2018]

- 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 inducing 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: $\begin{cases} t \in \{1, \ldots, T\} \\ f \in \{1, \ldots, F\} \end{cases}$

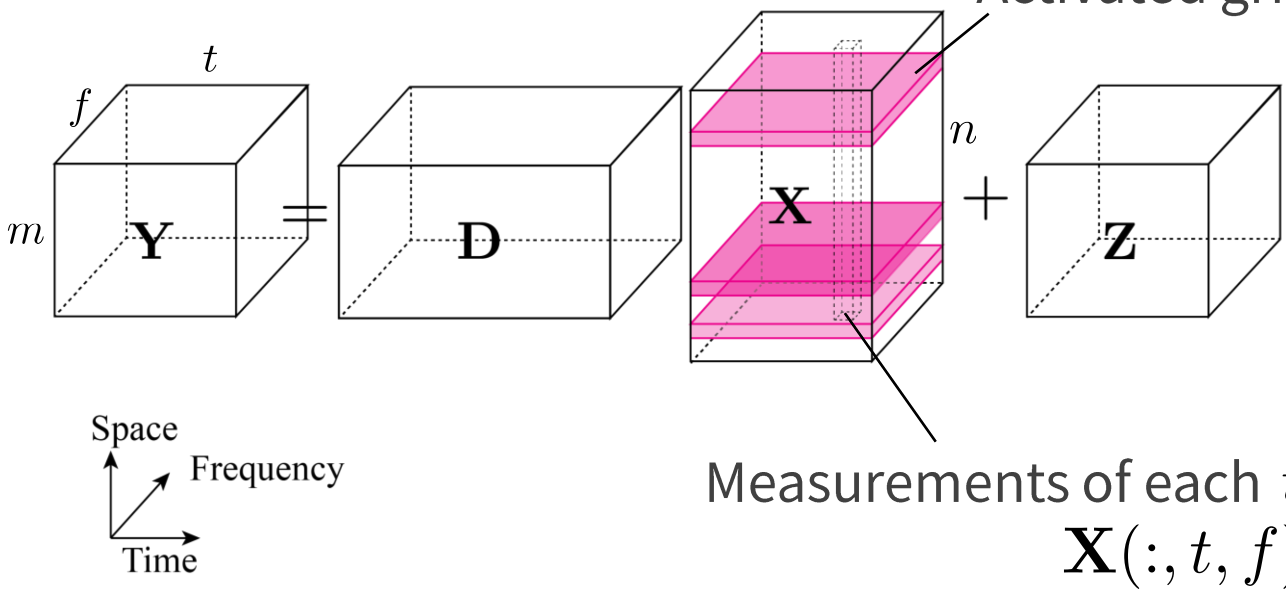
- Group sparsity for robust and accurate decomposition
  - Sound sources are static for several time frames
  - Acoustic source signals have a broad frequency band

Each $\mathbf{x}_{t,f}$ 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

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

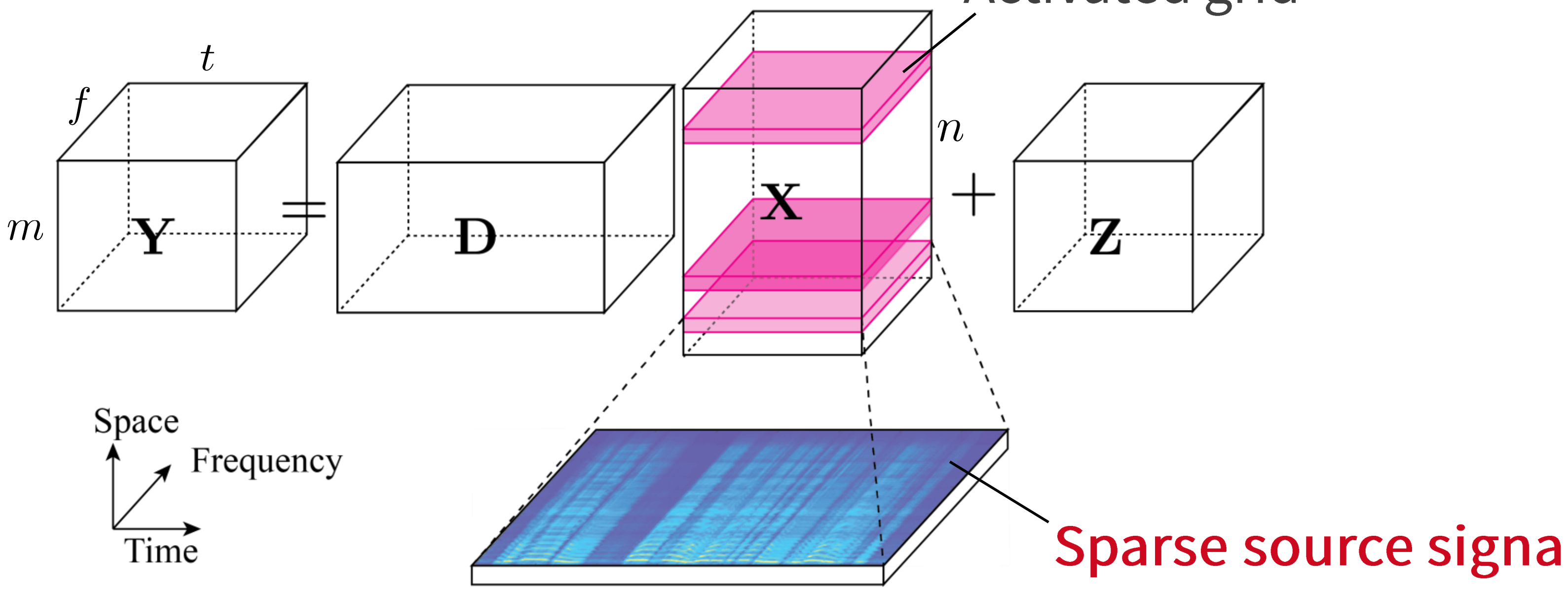
$$\mathcal{J}_{p,2,2}(\mathbf{X}) = \sum_{n} \left(\sum_{t,f} |\mathbf{X}(n,t,f)|^2 \right)^{\frac{p}{2}}$$

Penalty term for inducing 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]

- Surrogate func. for mixed-norm penalty term

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

$$\leq \sum_{n,t,f} \frac{p}{2} \eta_n^{\frac{p}{q}-1} \eta_{n,t,f}^{\frac{q}{2}-1} |\mathbf{X}(n,t,f)|^2 + C$$

$$= \mathcal{J}^{+}_{p,q,q}(\mathbf{X}|\boldsymbol{\Xi})$$

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

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

- Alternately update the parameters

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

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

➡ 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**
  $\mathbf{\Xi}^{(i)} = \mathbf{X}^{(i)}$
  $\eta_n^{(i)} = \sum_{t,f} \left|\mathbf{\Xi}^{(i)}(n,t,f)\right|^q$ for $\forall n$
  for $t = 1$ to T **do**
    for $f = 1$ to F **do**
      $\eta_{n,t,f}^{(i)} = \left|\mathbf{\Xi}^{(i)}(n,t,f)\right|^2$ 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

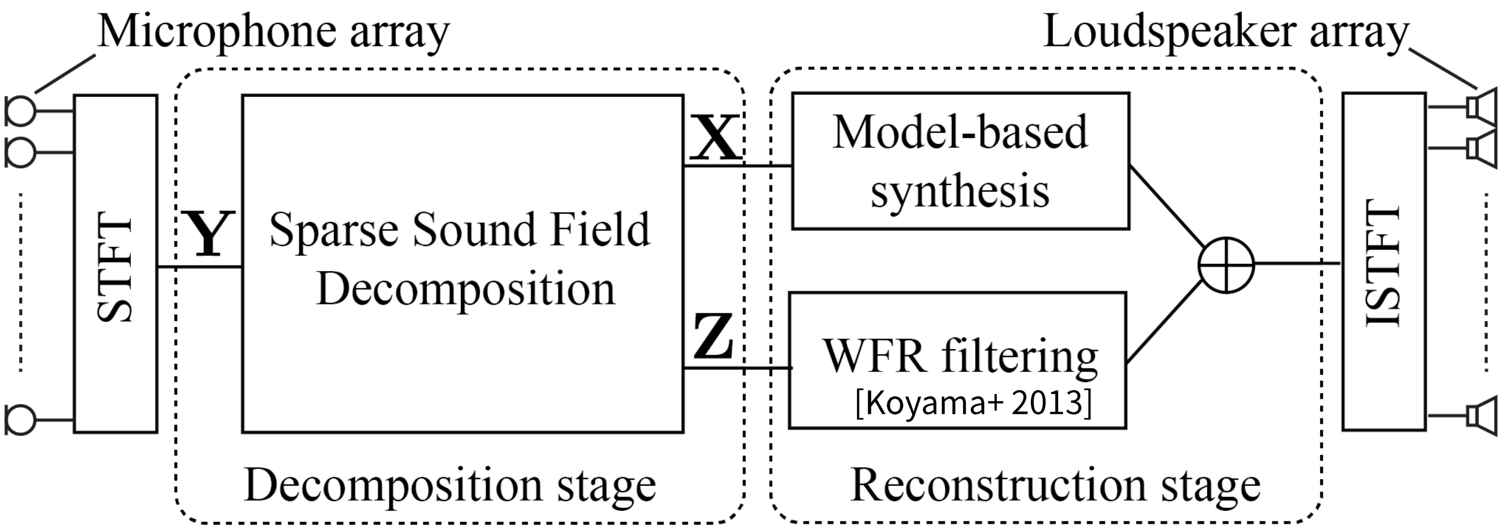
Several extensions of sparse decomposition

- Non-Gaussian reverberantion [Koyama+ IEEE JSTSP 2019]
  - 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+ Elsevier SP 2020]
  - Approximate sources as delta functions
  - Reciprocity gap functional in spherical harmonic domain
  - Closed-form solution using Hankel matrix

Application of sparse decomposition

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

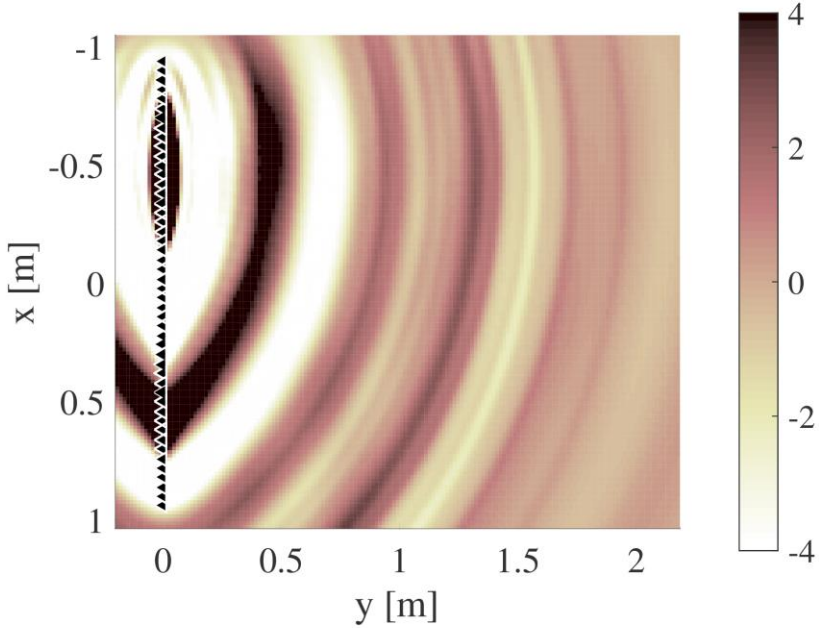


- **Decomposition stage:**
  - Group sparse decomposition of $\mathbf{Y}$ into $\mathbf{X}$ and $\mathbf{Z}$
- **Reconstruction stage:**
  - $\mathbf{X}$ and $\mathbf{Z}$ are separately converted into driving signals
  - Loudspeaker driving signals as sum of two components

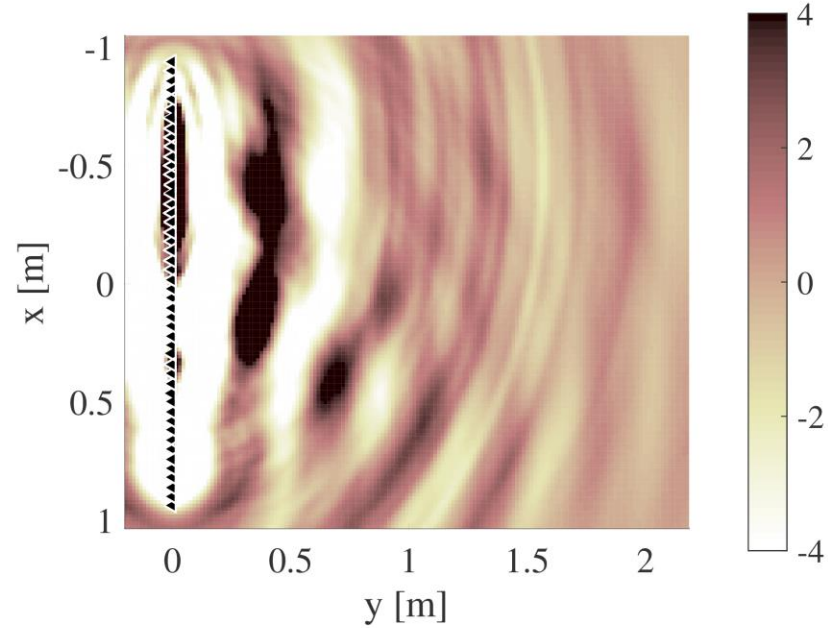
Reproduced pressure distribution

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

Proposed



Plane-wave-decomposition-based method

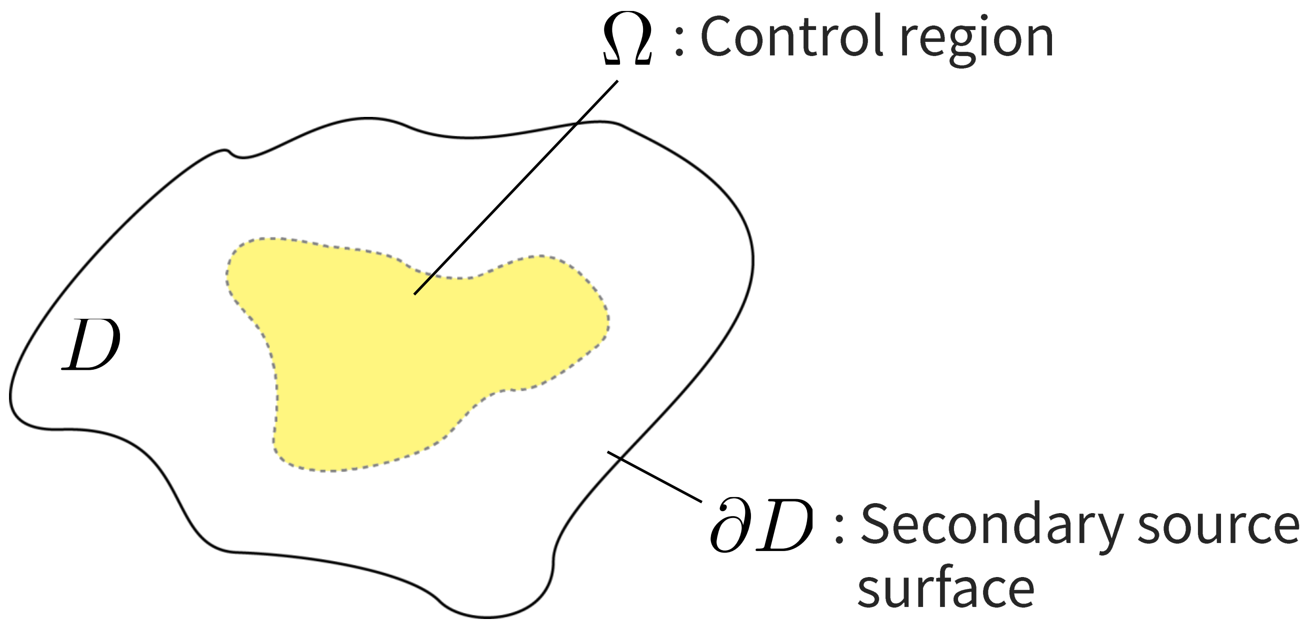


Spatial aliasing artifacts are reduced by proposed method

OPTIMAL SOURCE AND SENSOR PLACEMENT FOR SOUND FIELD CONTROL

Sound Field Control

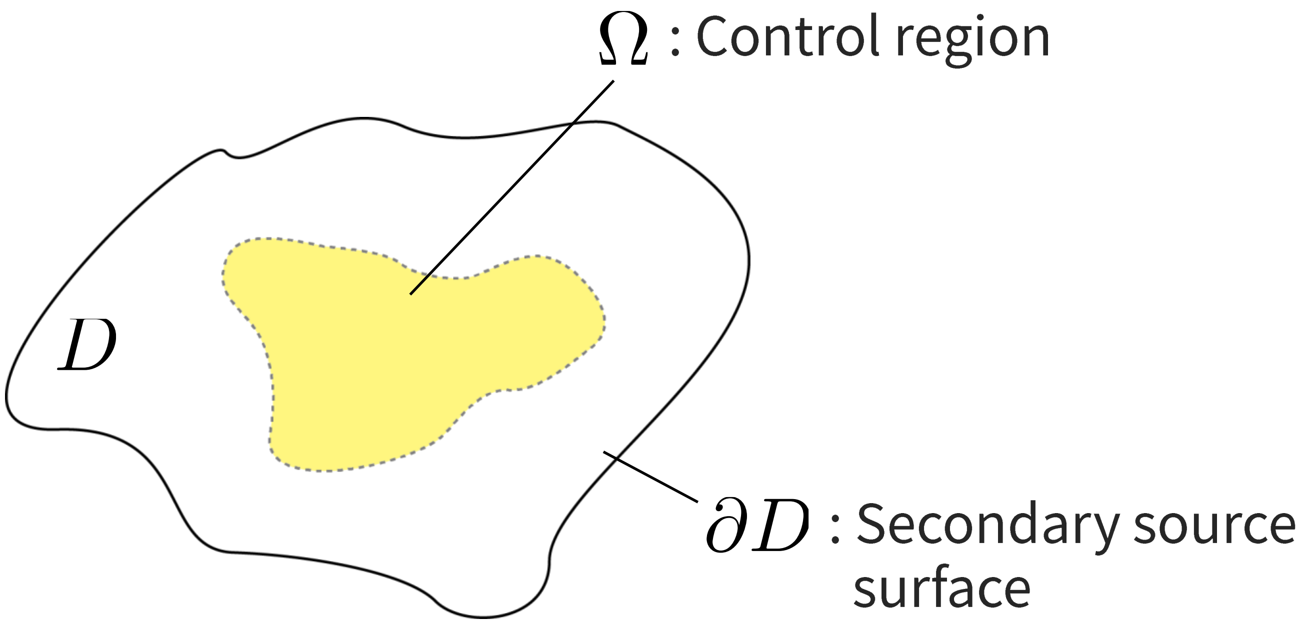
Synthesize desired sound field inside Ω by using secondary sources



- **High fidelity audio system**: synthesizing desired sound field
- **Spatial noise cancellation**: cancelling incoming noise

Source and sensor placement

Synthesize desired sound field inside Ω by using secondary sources



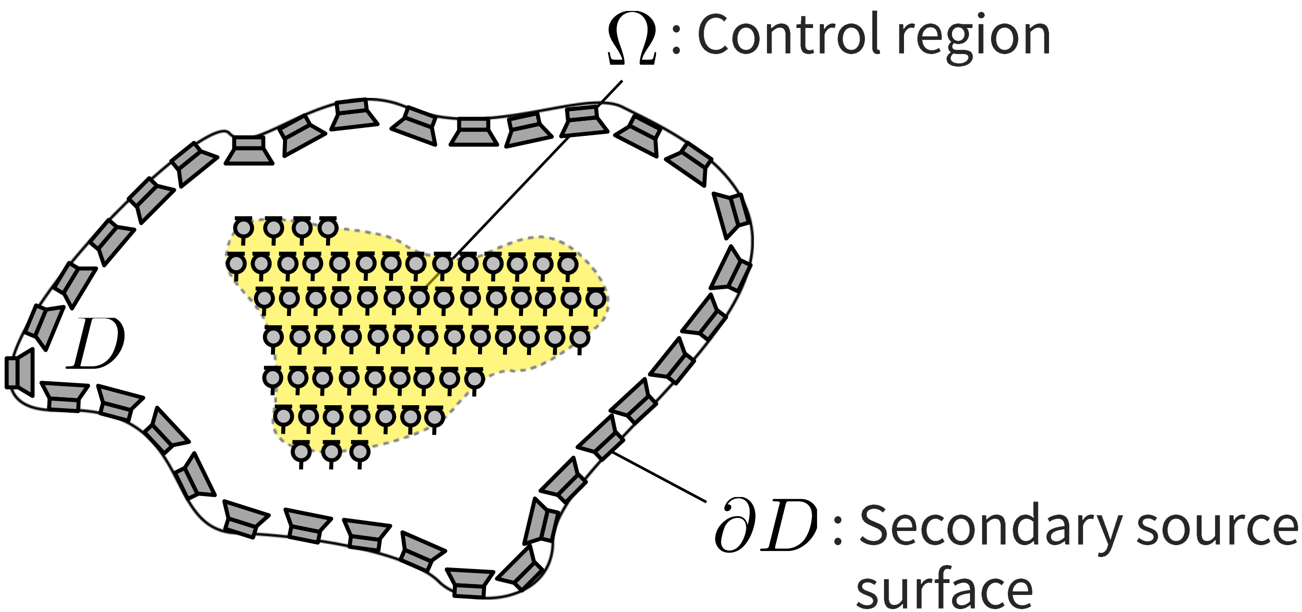
- Representation by single layer potential $\Rightarrow$ Inverse filter design

$$u(\mathbf{r},\omega) = \int_{\mathbf{r}'\in\partial D} \varphi(\mathbf{r}',\omega) G(\mathbf{r}|\mathbf{r}',\omega) d\mathbf{r}' \qquad (\mathbf{r}\in D)$$

Sound pressure — Driving signals — Monopole

Source and sensor placement

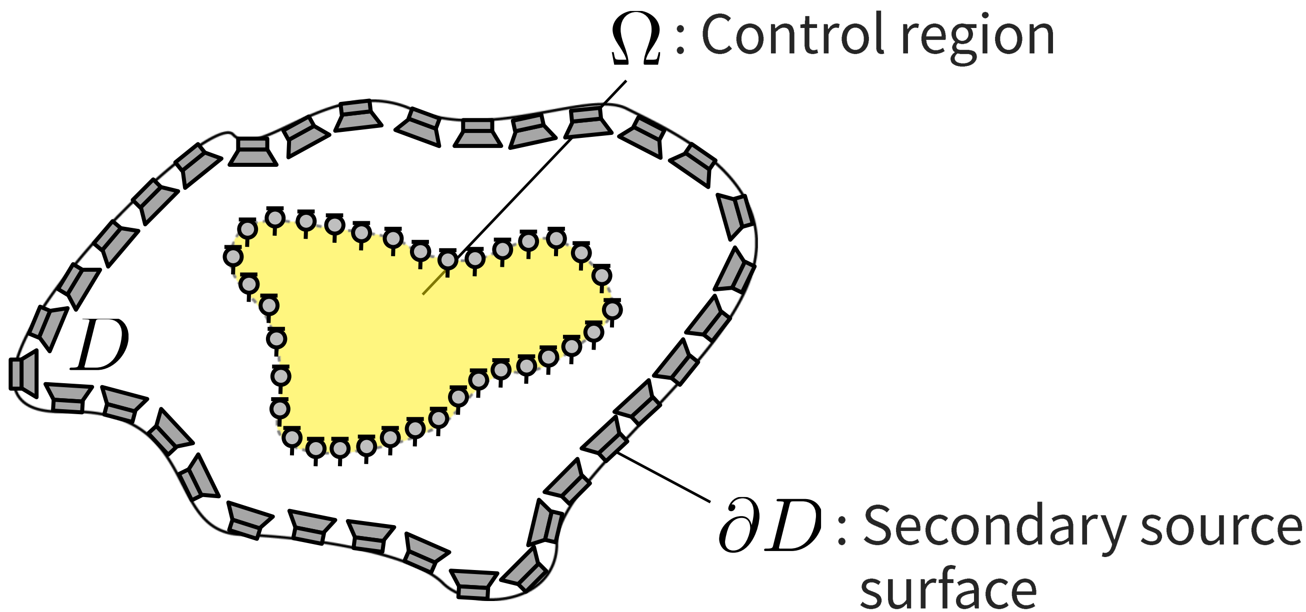
What is the best placement of secondary sources (loudspeakers) and sensors (control points/microphones)?



- Dense sampling over the region
  - Too many loudspeakers and microphones to measure transfer function in advance
  - Unstable inverse filter due to high correlation between transfer functions

Source and sensor placement

What is the best placement of secondary sources (loudspeakers) and sensors (control points/microphones)?



- Sampling only on boundary of Ω
  - Significant degradation of control accuracy at several frequencies (**forbidden frequency problem**)

Source and sensor placement

What is the best placement of secondary sources (loudspeakers) and sensors (control points/microphones)?

- Current method – secondary source placement
  - Method based on Gram—Schmidt orthogonalization [Asano+ 1999]
  - Sparse-approximation-based method [Khalilian+ 2016]

  ➡ Most algorithms depend on desired sound field
- Current method – sensor placement
  - Avoid forbidden frequency problem by introducing rigid baffle, directional microphones, and two layer array of microphones

  ➡ Most methods can be basically applied to simple array geometry

  ➡ Source and sensor placements are independently determined

A method for jointly determining the best placement of secondary sources and sensors for region of arbitrary geometry

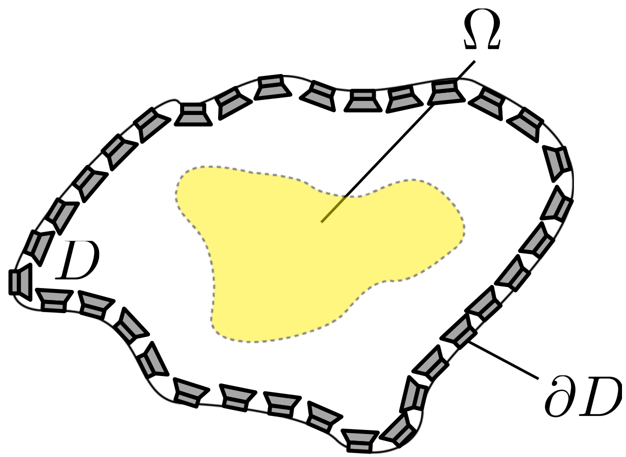
Sensor placement in machine learning

- **Cost function:**
  - Measures on Gram matrix $\mathbf{T}$ used in experimental design
    - Trace of $\mathbf{T}^{-1}$ [Liu+ 2016]
    - Log determinant of $\mathbf{T}^{-1}$ [Joshi+ 2009]
  - Information-theoretic measures:
    - Entropy [Wang+ 2004]
    - Mutual information [Krause 2008]
  - Frame potential [Ranieri+ 2014]
- **Algorithms:**
  - Greedy algorithm
  - Convex relaxation
  - Heuristics

Not applicable for joint source and sensor placement

Further investigations are given in our overview paper [Koyama+ IEEE/ACM TASLP 2020]

Problem statement



- Synthesized sound field by L loudspeakers

$$u_{\text{syn}}(\mathbf{r},\omega) = \sum_{l=1}^{L} d_l(\omega) g_l(\mathbf{r},\omega)$$

Driving signal — Transfer function

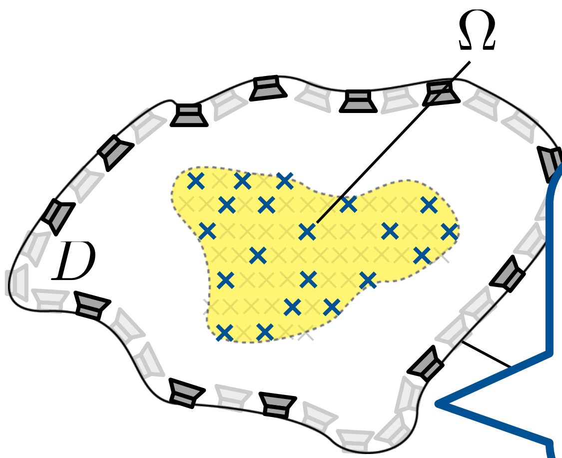
- Minimize squared error between synthesized and desired sound fields

$$\underset{d_l(\omega)}{\text{minimize}}\, \mathcal{J} = \int_{\mathbf{r}\in\Omega} \left| \sum_{l=1}^{L} d_l(\omega) g_l(\mathbf{r}) - u_{\text{des}}(\mathbf{r},\omega) \right|^2 d\mathbf{r}$$

Desired sound field

➡ Difficult to solve due to domain integral

Problem statement



Choose the best loudspeaker and control-point positions from candidates w.r.t. control accuracy and filter stability

- Linear equation by discretizing region

$$\mathbf{u}^{\mathrm{des}} = \mathbf{G}\mathbf{d}$$

Driving signal

Desired pressure

Transfer function matrix

- Driving signal by using Moore—Penrose pseudo inverse of $\mathbf{G}$

$$\mathbf{d} = \mathbf{G}^{\dagger}\mathbf{u}^{\mathrm{des}}$$

Moore—Penrose pseudo inverse

Idea

- Empirical Interpolation Method (EIM):
  - Proposed in the context of numerical analysis of partial differential equation [Barrault+ 2004]
  - Given functional space $\mathcal{V}$ defined on Ω, choose the best interpolation function and sampling points on Ω to approximate any function $v \in \mathcal{V}$ with greedy algorithm

- Apply EIM to source and sensor placement [Koyama+ 2018]
  - Regarding transfer function of each loudspeaker as interpolation function and control points as sampling points
  - Greedy algorithm for choosing source / sensor positions using transfer functions between candidate locations

Empirical Interpolation Method (EIM)

- Determine initial interpolation function and sampling point, and repeat the following procedure until interpolation error becomes smaller than threshold

1. Compute interpolation $I_Q(v)$ for $v \in \mathcal{V}$ by using interpolation functions h_q and sampling points x_q identified so far

$$I_Q(v) = \sum_{q=1}^{Q} c_q h_q$$

(c_q is solution of the following linear eq.

$$v(x_q) = \sum_{q'=1}^{Q} c_{q'} h_{q'}(x_q)$$
)

2. v that maximizes L_∞-norm of error between v and its in interpolation $I_Q(v)$ is taken as h_{Q+1}
3. Point of maximal absolute value of error between $v(x)$ and its interpolation $I_Q(v)$ is taken as x_{Q+1}

➡ **Given function is guaranteed to be stably approximated below target error**

Proposed algorithm

➢ Applying EIM by regarding functional space $\mathcal{V}$ as transfer functions between candidate locations

➢ Input: Candidate locations of loudspeakers $\mathbf{r}_l$ ($l \in \{1, \dots, L\}$) and control points $\mathbf{r}_m$ ($m \in \{1, \dots, M\}$), transfer function matrix $\mathbf{G} \in \mathbb{C}^{M \times L}$, tolerance error ϵ_{tol}

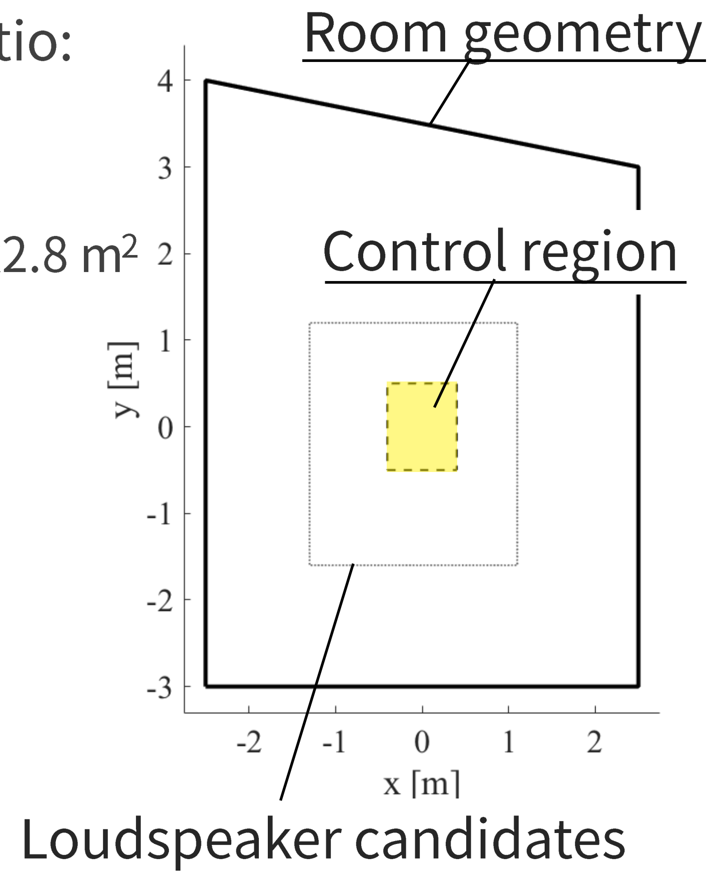
➢ Output: Set of indexes of loudspeakers and control points

1. Set Q=1
2. **while** $\epsilon > \epsilon_{\mathrm{tol}}$ **do**
3. Choose loudspeaker index
$$l_Q = \mathop{\arg\max}_{l=1,\dots,L} \left\| \mathbf{G}_{\cdot,l} - I_{Q-1}(\mathbf{G}_{\mathbf{m}_{Q-1},l}) \right\|_\infty$$
4. Choose control-point index
$$m_Q = \mathop{\arg\max}_{m=1,\dots,M} \left| \mathbf{G}_{m,l_Q} - (I_{Q-1}(\mathbf{G}_{\mathbf{m}_{Q-1},l_Q}))_m \right|$$
5. Compute error
$$\epsilon = \max_{l=1,\dots,L} \left\| \mathbf{G}_{\cdot,l} - I_{Q-1}(\mathbf{G}_{\mathbf{m}_{Q-1},l}) \right\|_2$$
6. Set Q=Q+1
7. **end while**

Approximation of $\mathbf{G}$ below ϵ_{tol} and inverse filter stability are guaranteed

Numerical simulations

- Experiments in 2D sound field
  - Transfer functions simulated by finite element method (FEM) (absorption ratio: 0.10)
  - Loudspeaker candidates:
    - Boundary of rectangular region of 2.4x2.8 m^2
    - Regularly discretized into 256 points
  - Control-point candidates:
    - Rectangular region of 0.8x1.0 m^2
    - Discretized every 0.04 m
  - Comparison:
    - Proposed method (**Proposed**)
    - Random (**Rand**)
    - Regular + Regular (**Reg-Reg**)
    - Regular + 2 layer (**Reg-2L**)
  - Desired field: plane wave field (every 10 deg)

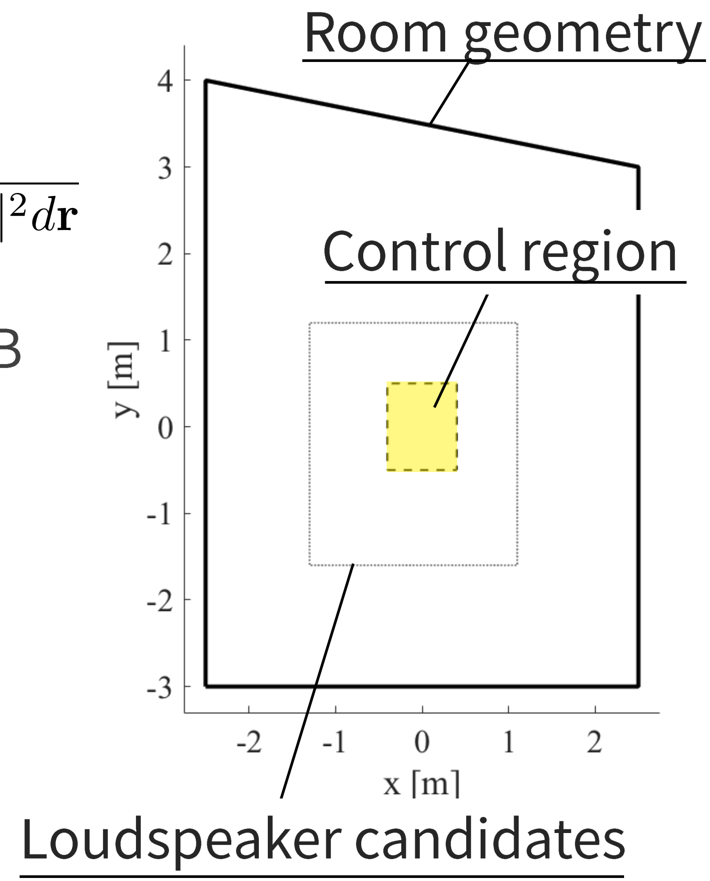


- Experiments in 2D sound field
  - Control accuracy: Signal-to-Distortion Ratio (SDR)

$$\mathrm{SDR}(\omega) = 10\log_{10}\frac{\int_{\Omega}|u_{\mathrm{des}}(\mathbf{r},\omega)|^2 d\mathbf{r}}{\int_{\Omega}|u_{\mathrm{syn}}(\mathbf{r},\omega) - u_{\mathrm{des}}(\mathbf{r},\omega)|^2 d\mathbf{r}}$$

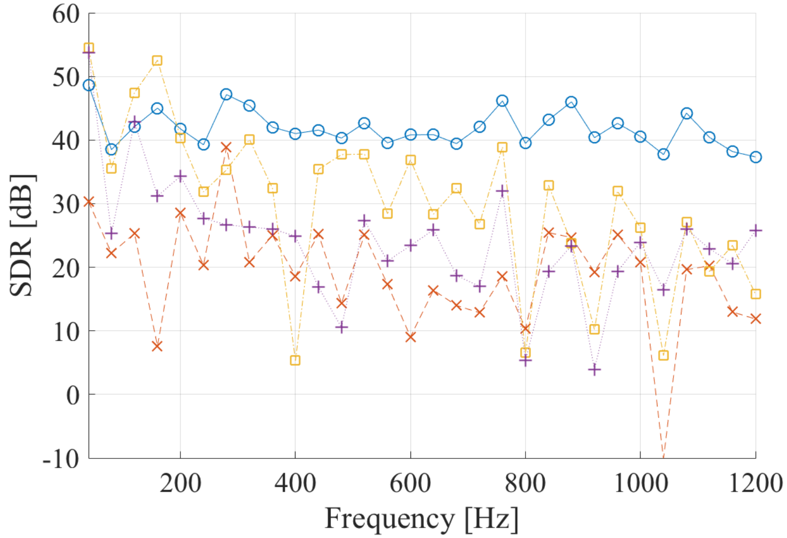
  - Filter stability: Condition number in dB

$$\kappa(\mathbf{G}) = 10\log_{10}\frac{\sigma^2_{\max}(\mathbf{G})}{\sigma^2_{\min}(\mathbf{G})}$$

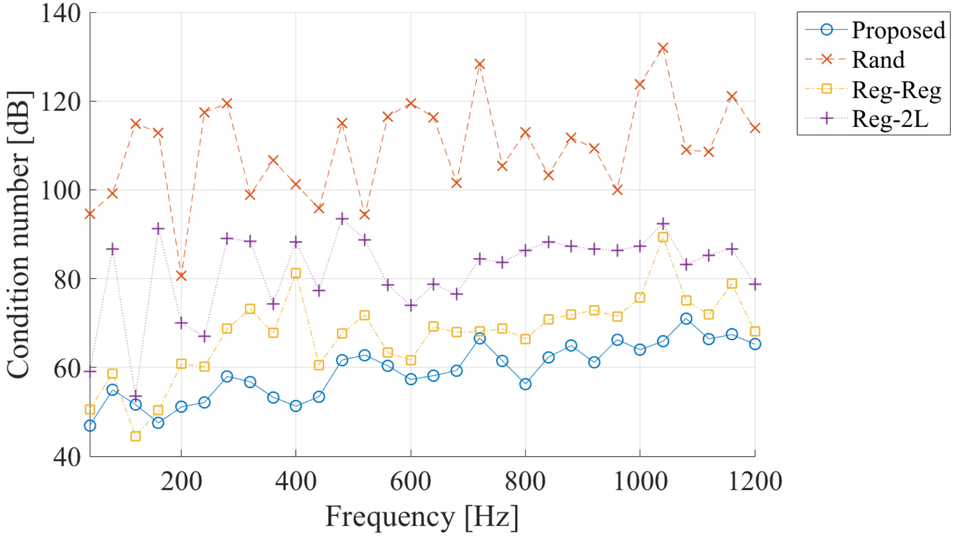


Results – single frequency case

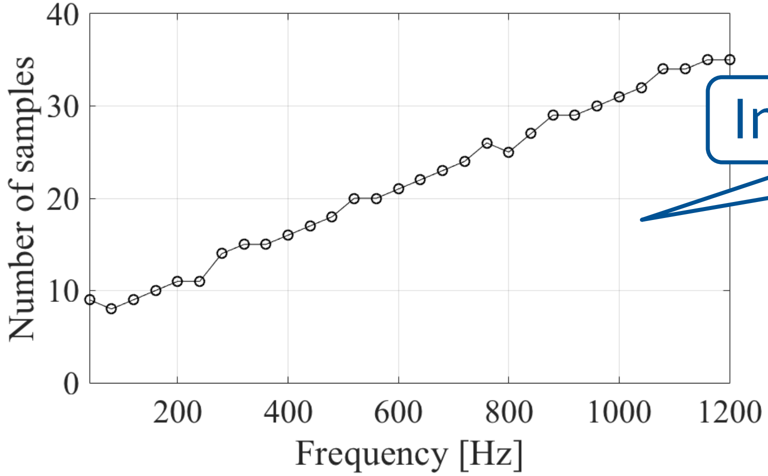
SDR



Condition number



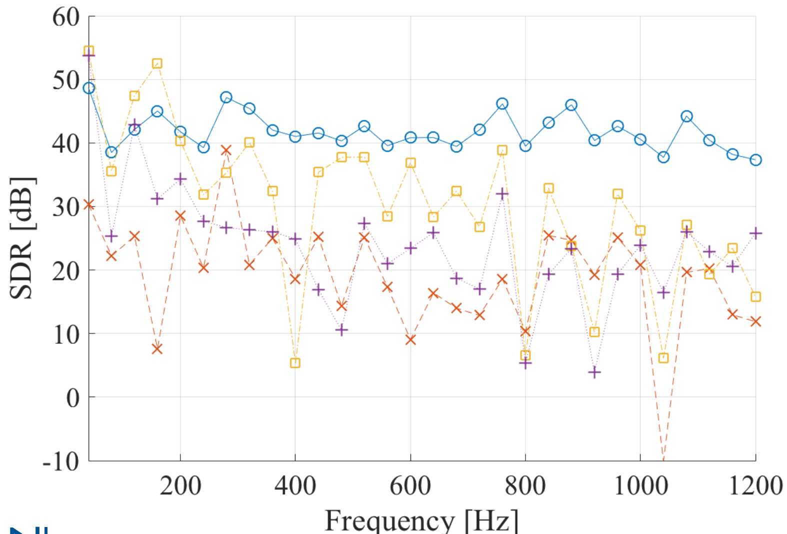
Number of loudspeakers/control points



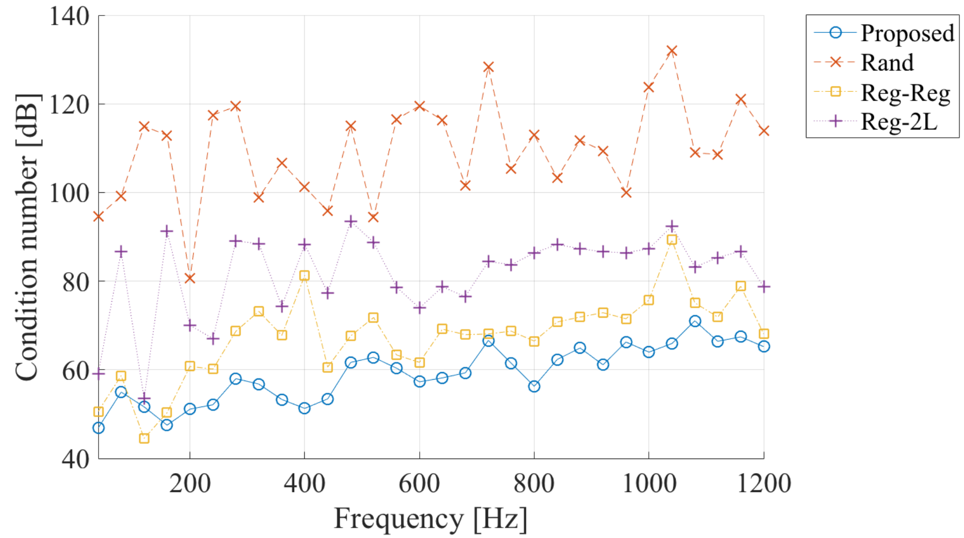
Increase with frequency

Results – single frequency case

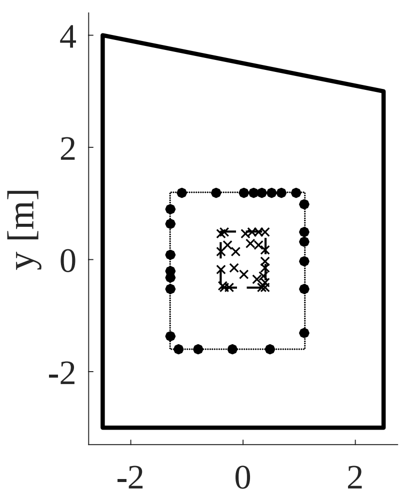
SDR



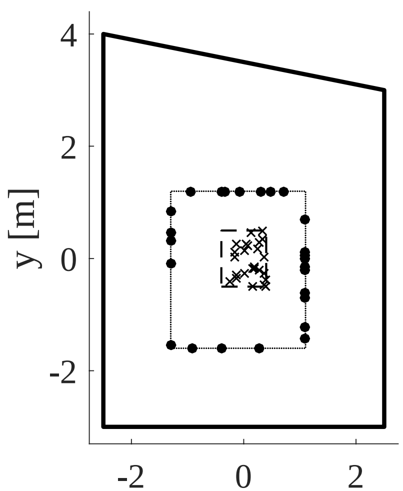
Condition number



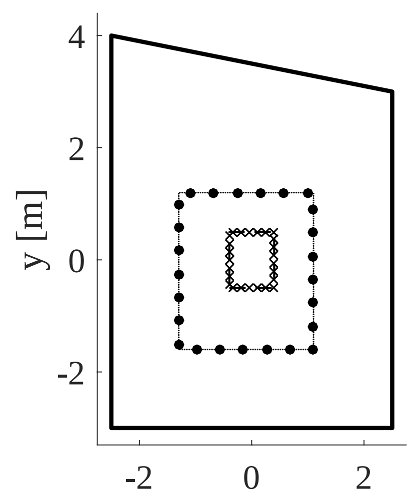
Locations at 800 Hz



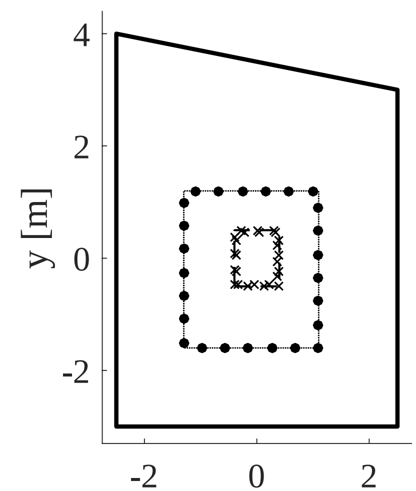
Proposed



Rand

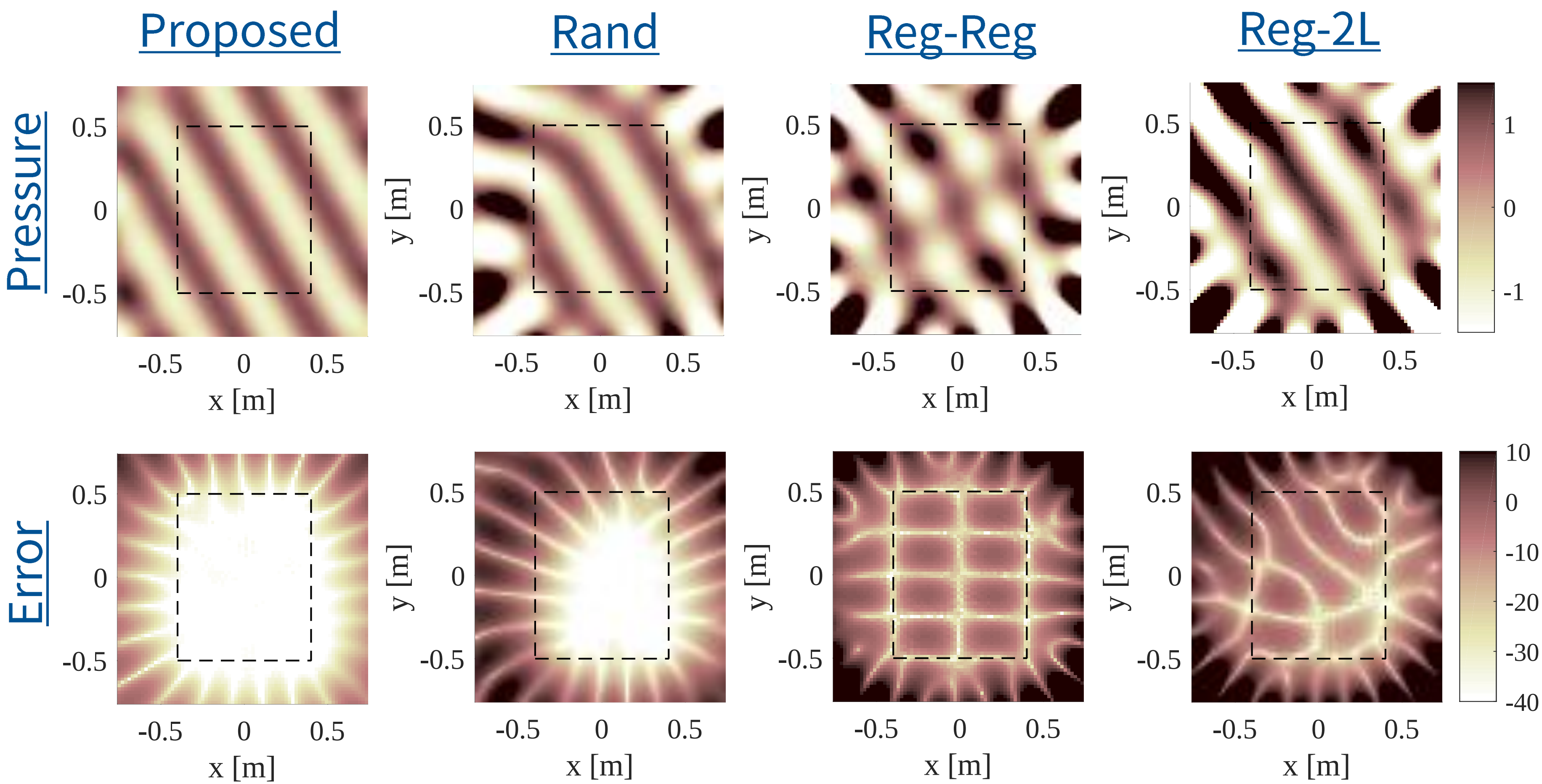


Reg-Reg



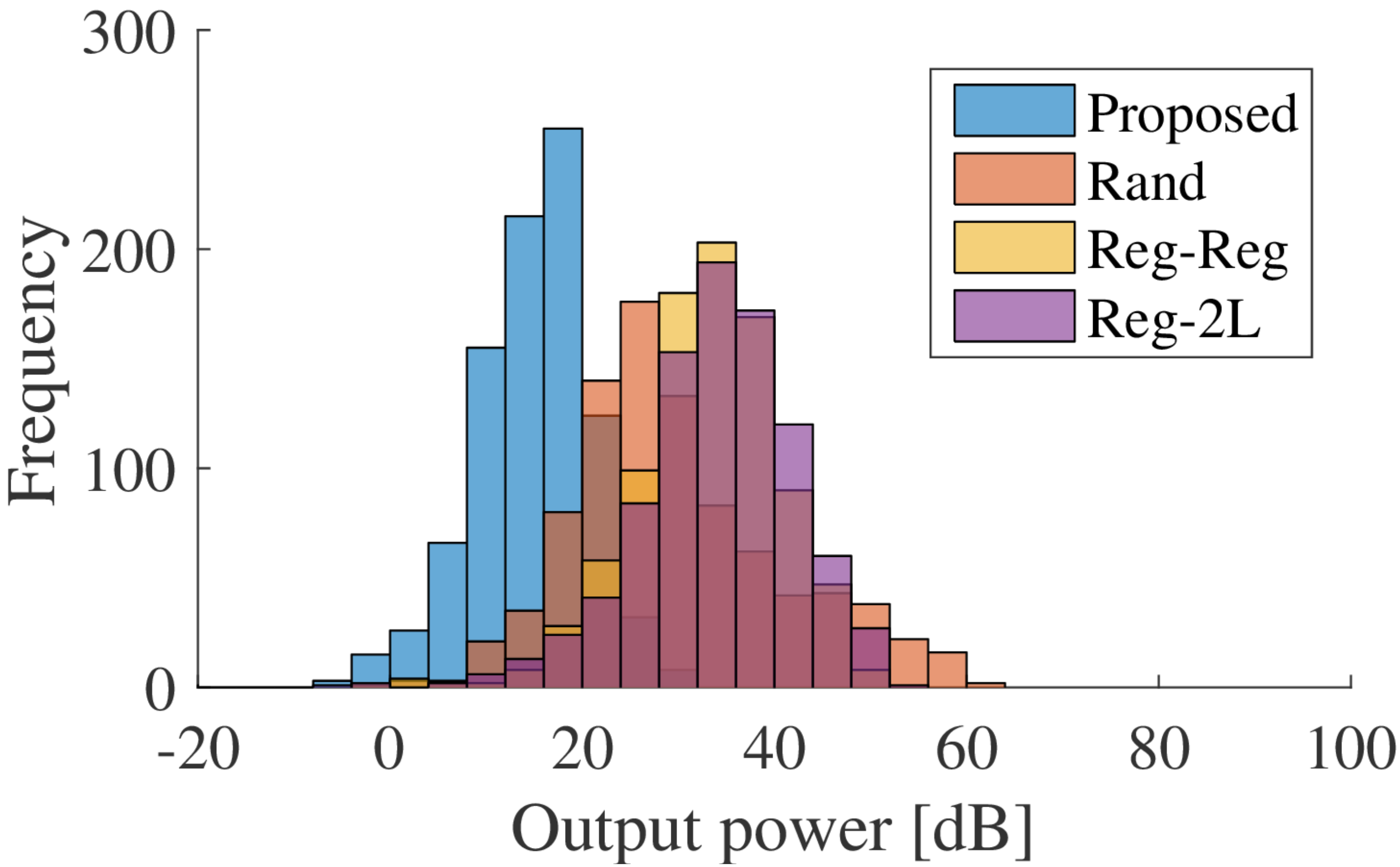
Reg-2L

- Synthesized pressure and error distributions at 800 Hz

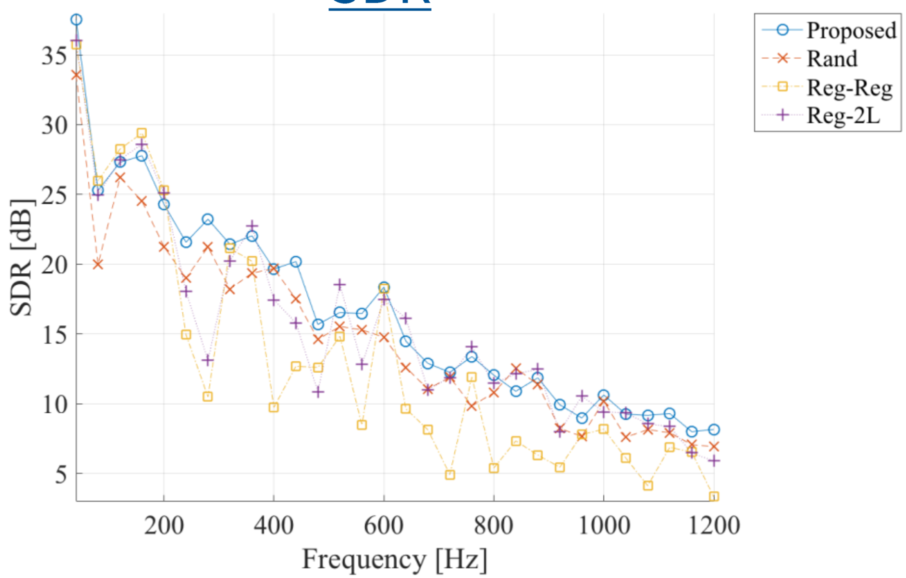


Results – single frequency case

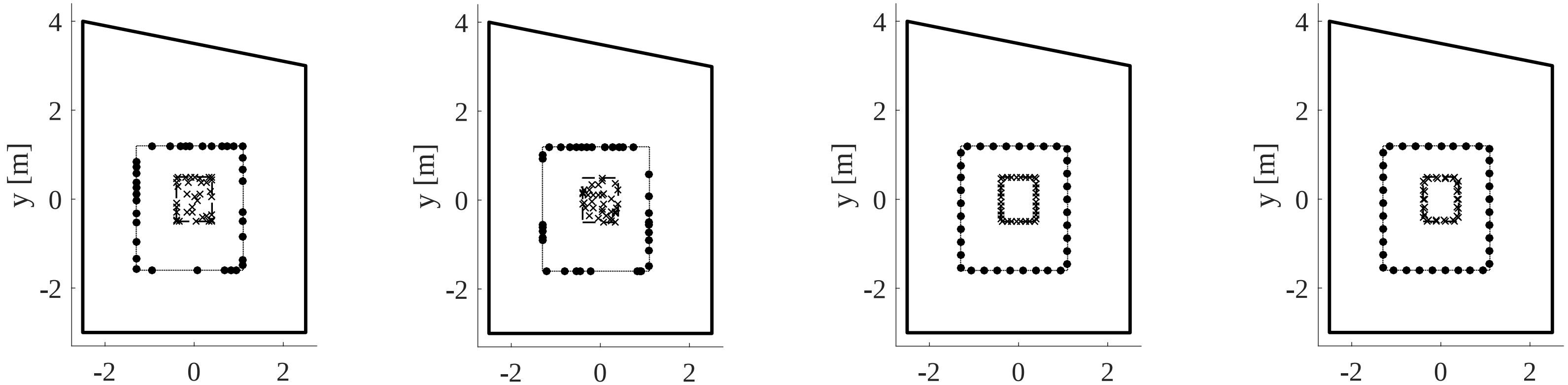
- Output power of loudspeakers at 800 Hz



SDR



Locations



Proposed Rand Reg-Reg Reg-2L

Conclusion

- Sparse modeling and its application to acoustic signal processing
  - Source-free region: Harmonic analysis of infinite orders / Sparse plane wave decomposition
  - Region including sources: Sound field decomposition based on sparsity of source distribution
  - Application to recording and reproduction

- Optimal source and sensor placement for sound field control
  - Optimal placement loudspeaker and control points
  - Empirical interpolation method by regarding sound field control problem as function interpolation
  - High reproduction accuracy and filter stability with preventing forbidden frequency problem

Related publications

- S. Koyama, *et al.* "Optimizing source and sensor placement for sound field control: an overview," *IEEE/ACM Trans. ASLP*, 2020.
- Y. Takida, S. Koyama, *et al.* "Reciprocity gap functional in spherical harmonic domain for gridless sound field decomposition," Elsevier Signal Process., 2020.
- N. Ueno, S. Koyama, and H. Saruwatari, "Three-dimensional sound field reproduction based on weighted mode-matching method," *IEEE/ACM Trans. ASLP*, 2019.
- S. Koyama and L. Daudet. "Sparse representation of a spatial sound field in a reverberant environment," *IEEE J. STSP*, 2019.
- H. Ito, S. Koyama, *et al.* "Feedforward spatial active noise control based on kernel interpolation of sound field," *Proc. IEEE ICASSP*, 2019.
- S. Koyama, *et al.* "Sparse sound field decomposition for super-resolution in recording and reproduction," *JASA*, 2018.
- S. Koyama, *et al.* "Joint source and sensor placement for sound field control based on empirical interpolation method," *Proc. IEEE ICASSP*, 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, *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. STSP*, 2015.
- S. Koyama, *et al.* "Sparse sound field representation in recording and reproduction for reducing spatial aliasing artifacts," *Proc. IEEE ICASSP*, 2014.