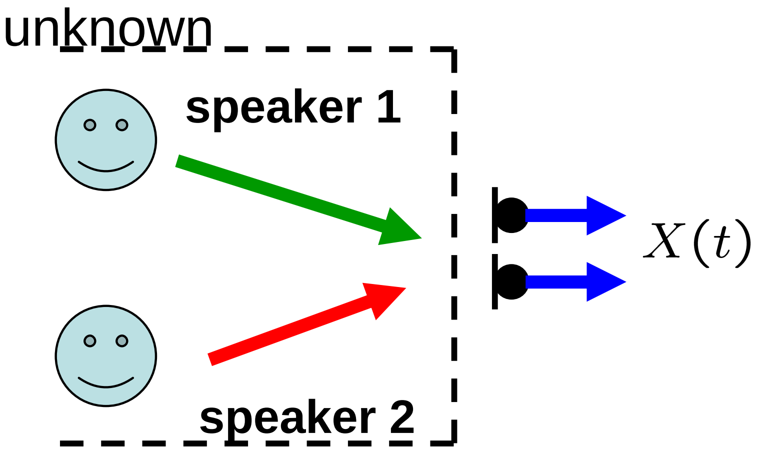
RESOLVING FD-BSS PERMUTATION FOR ARBITRARY ARRAY IN PRESENCE OF SPATIAL ALIASING

Jani Even, Norihiro Hagita
even@atr.jp

Introduction

Cocktail party (speech/speech)



- Blind signal separation (BSS)
- Frequency domain
- Permutation resolution
- DOA based approach
- Spatial aliasing

- Model for the spatial aliasing
- Sparse solution
- Permutation resolution

Outline

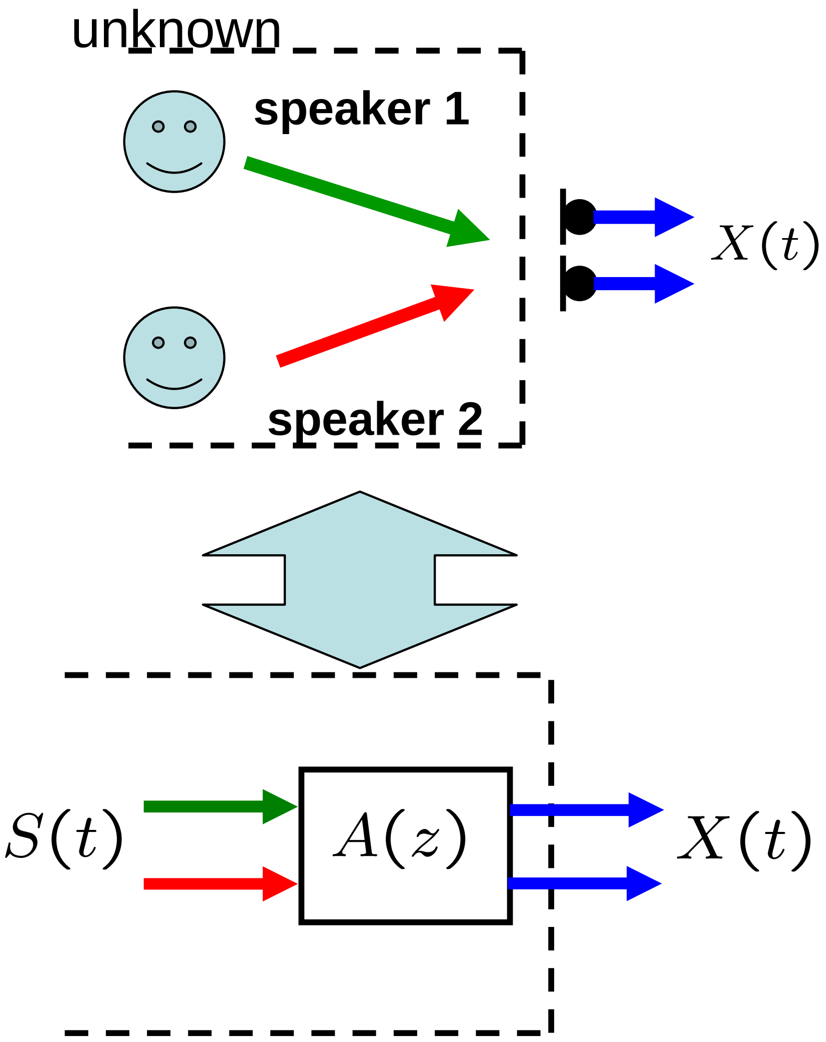
1. Preliminaries
   1. FD-BSS & Permutation
   2. Method based on DOA estimation
2. Proposed approach
   1. Spatial aliasing equation
   2. Finding the solution
   3. Permutation resolution
3. Simulation results

Outline

1. Preliminaries
    1. FD-BSS & Permutation
    2. Method based on DOA estimation
2. Proposed approach
    1. Spatial aliasing equation
    2. Finding the solution
    3. Permutation resolution
3. Simulation results

Frequency domain mixture

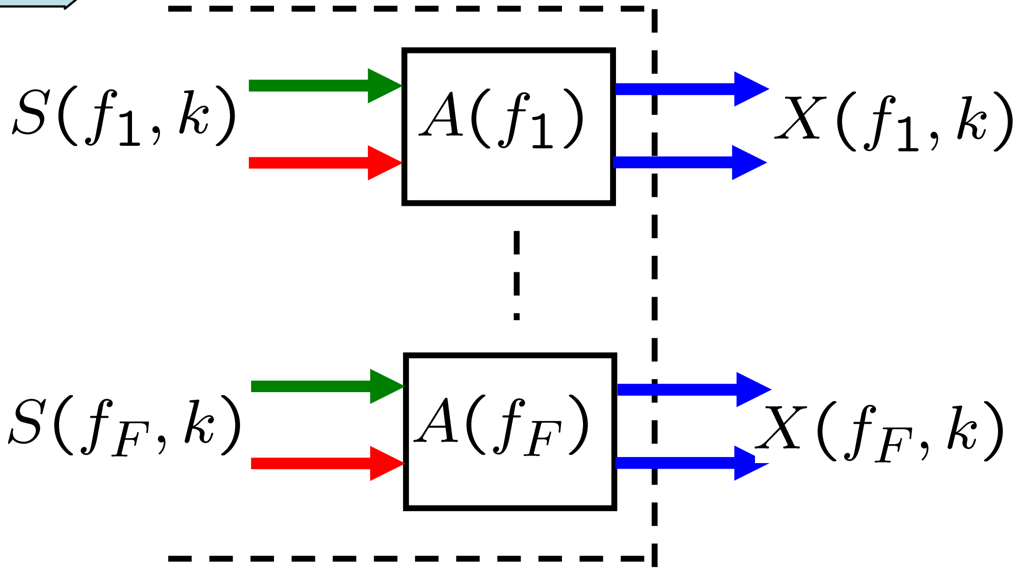
Time domain
Convolutive mixture



Unknown

STFT

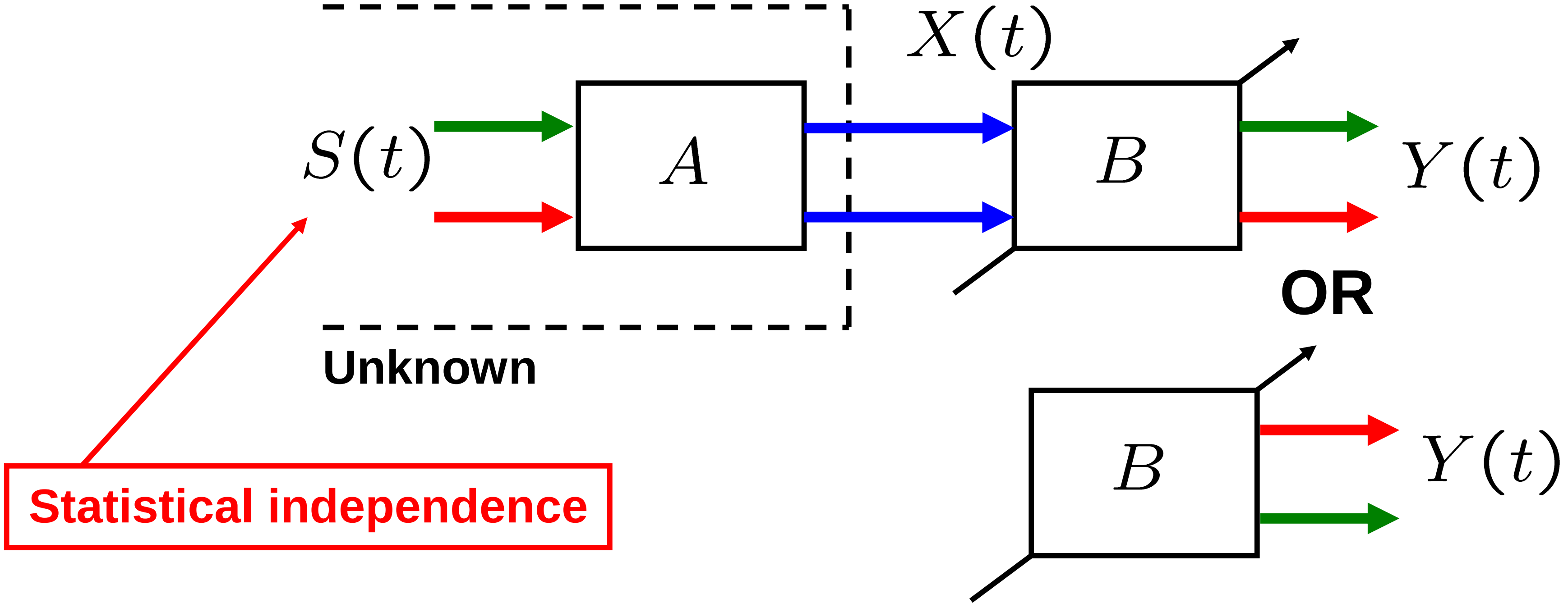
Frequency domain
instantaneous mixtures



Unknown

- **User 1**
- **User 2**
- **Mixtures**

Frequency domain BSS



(independent component analysis ICA)

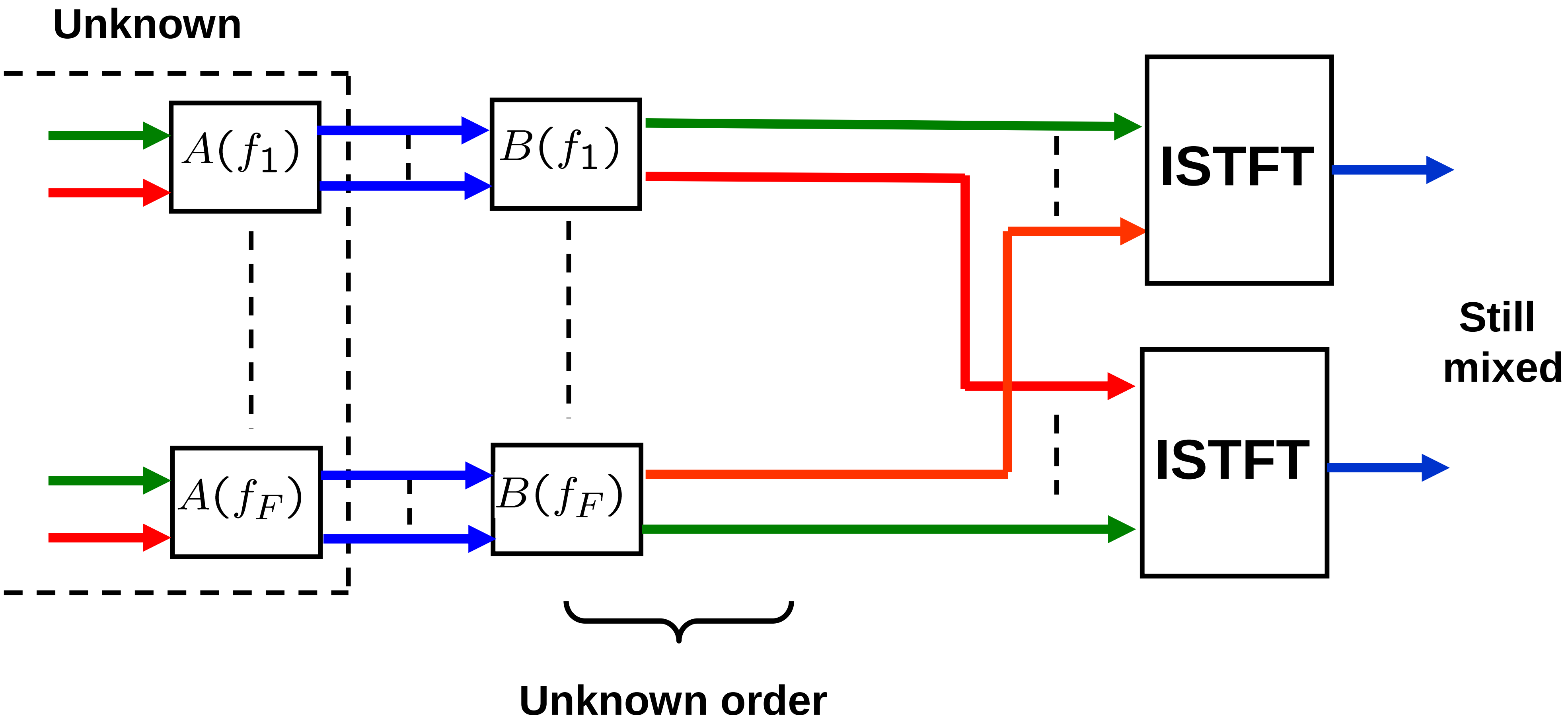
🟩 **Speech 1**
🟥 **Speech 2**
🟦 **Mixture**

$$Y(t) = P\Lambda S(t)$$

Λ **diagonal matrix**

P **permutation matrix**

Need for permutation resolution


Unknown

$A(f_1)$

$A(f_F)$

$B(f_1)$

$B(f_F)$

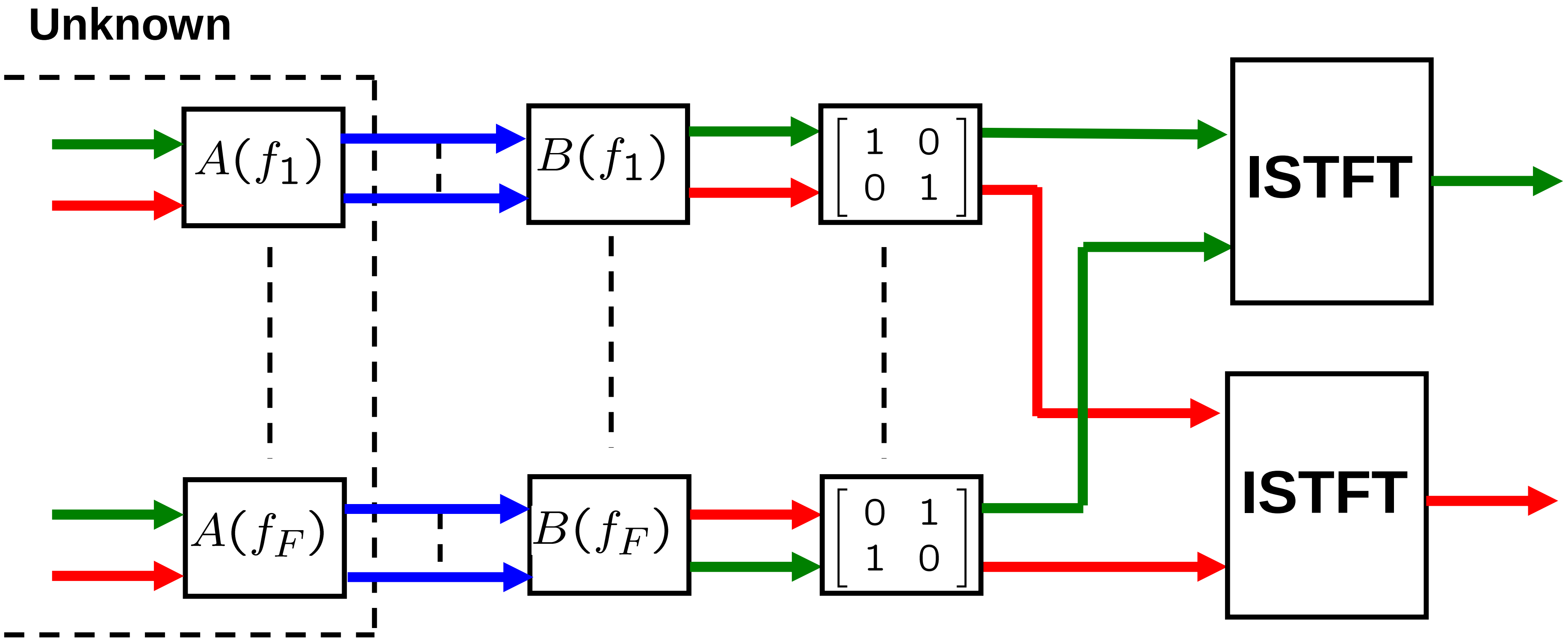
ISTFT

ISTFT

Still mixed

Unknown order

Set of permutation matrices



Permutation resolution ⟺ **Determine *F* permutation matrices**

Permutation resolution methods

See review paper [Pedersen, Larsen, Kjems, Parra 2007]

1. **Consistency of the spectrum of the recovered signals**
    1. **Amplitude modulation**
    2. **Cross correlation**
    3. **Source distribution**
    4. **…**
2. **Consistency of the filter coefficients**
    1. **Initialization**
    2. **Smooth spectrum**
    3. **Directivity pattern**
    4. **DOA**

Results in presence of spatial aliasing:

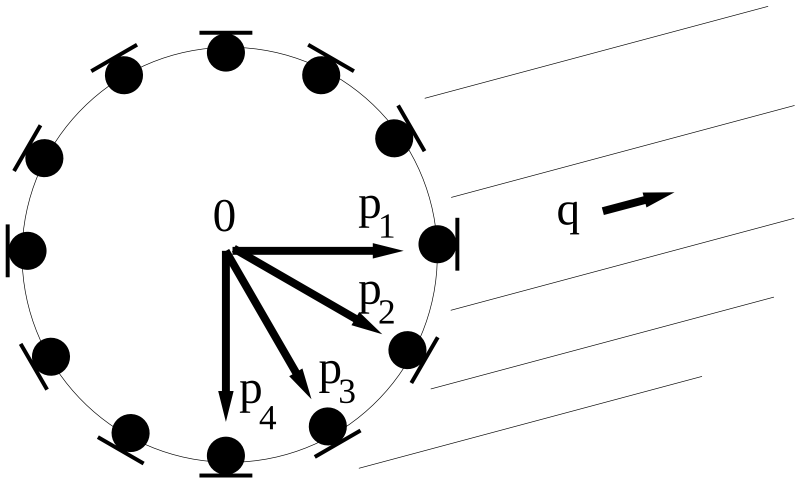
[Morgan, Ikram 2002]
[Sawada,Araki,Mukai,Makino, 2006]
[Nesta, Omologo,2010]

Outline

1. Preliminaries
    1. FD-BSS & Permutation
    2. Method based on DOA estimation
2. Proposed approach
    1. Spatial aliasing equation
    2. Finding the solution
    3. Permutation resolution
3. Simulation results

Steering vector model

Far field assumption



Attenuation and delay model

$$\mathbf{A}(f) \approx \begin{bmatrix} \frac{1}{\lambda_1} e^{-j2\pi \frac{f}{c}\mathbf{p}_1\cdot\mathbf{q}_1} \dots & \frac{1}{\lambda_n} e^{-j2\pi \frac{f}{c}\mathbf{p}_1\cdot\mathbf{q}_n} \\ \vdots & \vdots \\ \frac{1}{\lambda_1} e^{-j2\pi \frac{f}{c}\mathbf{p}_n\cdot\mathbf{q}_1} \dots & \frac{1}{\lambda_n} e^{-j2\pi \frac{f}{c}\mathbf{p}_n\cdot\mathbf{q}_n} \end{bmatrix}$$

$$\mathbf{q}_1 \text{ - - - - - - - - } \mathbf{q}_n$$

Absolute DOAs

Estimated mixing matrix approximation from BSS

$$\widehat{\mathbf{A}(f)} = \mathrm{B}(f)^{-1} \approx \begin{bmatrix} \frac{1}{\lambda_1} e^{-j2\pi \frac{f}{c}\mathbf{p}_1\cdot\mathbf{q}_{\sigma(f,1)}} \dots & \frac{1}{\lambda_n} e^{-j2\pi \frac{f}{c}\mathbf{p}_1\cdot\mathbf{q}_{\sigma(f,n)}} \\ \vdots & \vdots \\ \frac{1}{\lambda_1} e^{-j2\pi \frac{f}{c}\mathbf{p}_n\cdot\mathbf{q}_{\sigma(f,1)}} \dots & \frac{1}{\lambda_n} e^{-j2\pi \frac{f}{c}\mathbf{p}_n\cdot\mathbf{q}_{\sigma(f,n)}} \end{bmatrix}$$

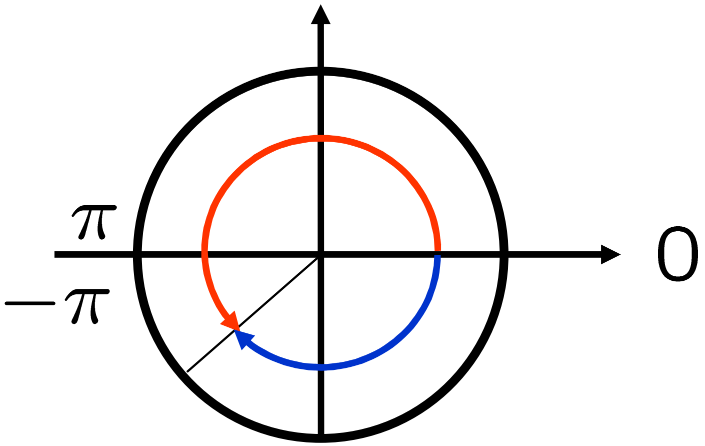
$$\mathbf{q}_{\sigma(f,1)} \text{ - - - - - - - - } \mathbf{q}_{\sigma(f,n)}$$

Permuted absolute DOAs

DOA estimation constraints

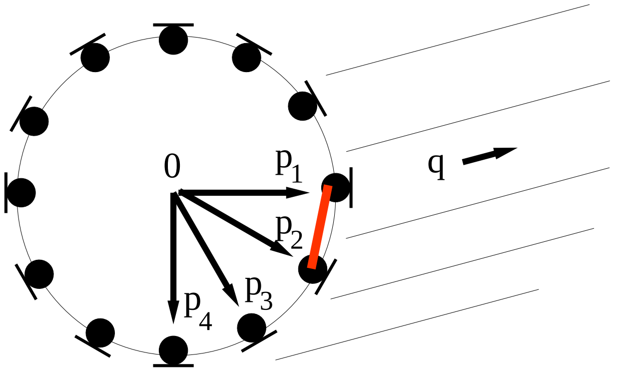
Ratio of coefficients

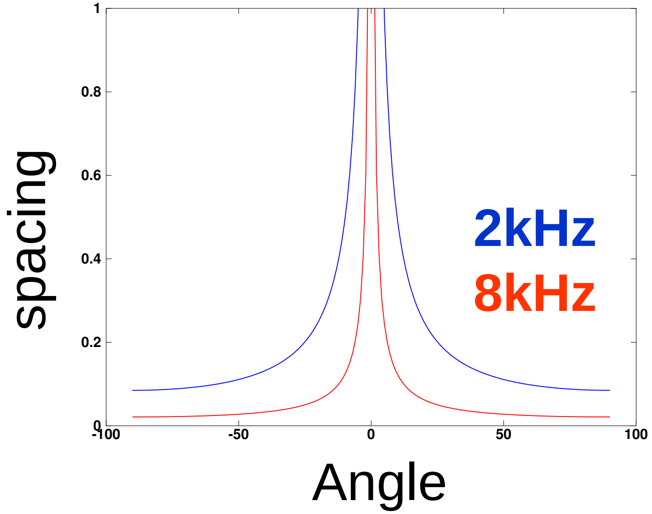
$$\arg\left(\frac{\widehat{\mathbf{A}}_{ij}(f)}{\widehat{\mathbf{A}}_{kj}(f)}\right) = -2\pi\frac{f}{c}(\mathbf{p}_i - \mathbf{p}_k)\cdot\mathbf{q}_{\sigma(f,j)},$$

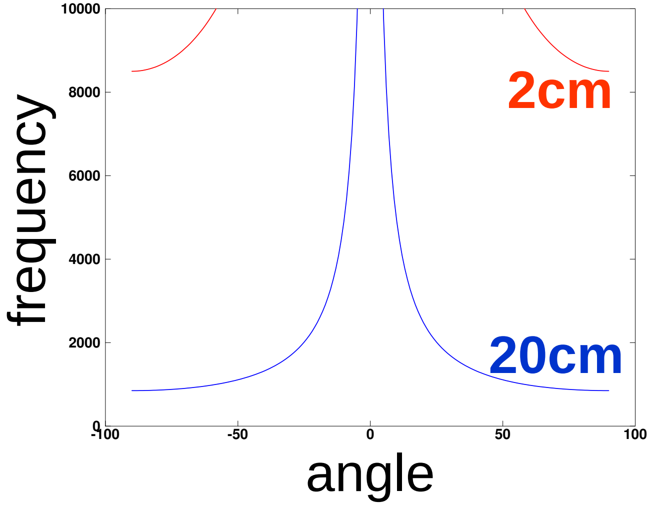


Limited to domain where arg is bijective

$$-\pi < 2\pi\frac{f}{c}(\mathbf{p}_i - \mathbf{p}_k)\cdot\mathbf{q}_{\sigma(f,j)} \le \pi$$







Use all sensors pairs

$$\arg\left(\frac{\widehat{\mathbf{A}}_{ij}(f)}{\widehat{\mathbf{A}}_{kj}(f)}\right) = -2\pi\frac{f}{c}(\mathbf{p}_i - \mathbf{p}_k)\cdot\mathbf{q}_{\sigma(f,j)},$$

Stack for all microphone pairs

$$L = \frac{n(n-1)}{2}$$

$$\mathbf{R}_j(f) = \begin{bmatrix} \arg\left(\frac{\widehat{\mathbf{A}}_{1j}(f)}{\widehat{\mathbf{A}}_{2j}(f)}\right) \\ \arg\left(\frac{\widehat{\mathbf{A}}_{2j}(f)}{\widehat{\mathbf{A}}_{3j}(f)}\right) \\ \vdots \\ \arg\left(\frac{\widehat{\mathbf{A}}_{n-1j}(f)}{\widehat{\mathbf{A}}_{nj}(f)}\right) \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} \mathbf{p}_1 - \mathbf{p}_2 \\ \mathbf{p}_2 - \mathbf{p}_3 \\ \vdots \\ \mathbf{p}_{n-1} - \mathbf{p}_n \end{bmatrix} \updownarrow L$$

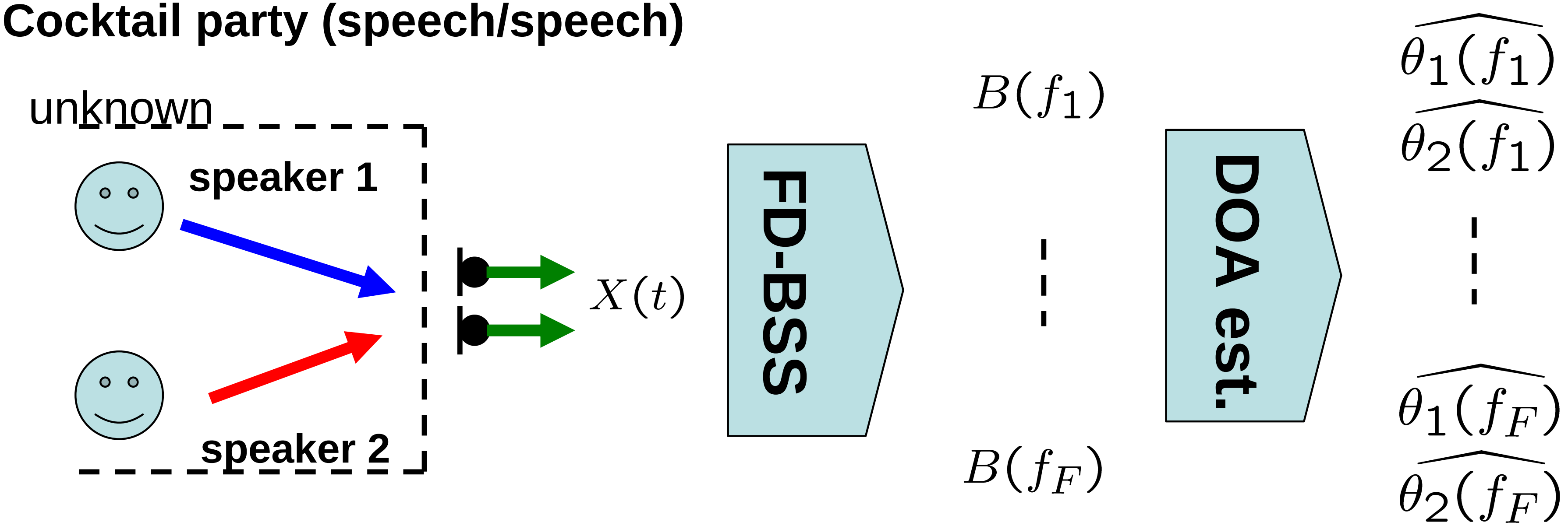
Least square solution

$$\mathbf{R}_j(f) = -2\pi\frac{f}{c}\mathbf{P}\boxed{\mathbf{q}_{\sigma(f,j)}} \Rightarrow \mathbf{q}_{\sigma(f,j)} = -\frac{c}{2\pi f}\mathbf{P}^{+}\mathbf{R}_j(f),$$

Unknown

Permutation alignment

Cocktail party (speech/speech)



unknown

speaker 1

speaker 2

$X(t)$

FD-BSS

$B(f_1)$

$B(f_F)$

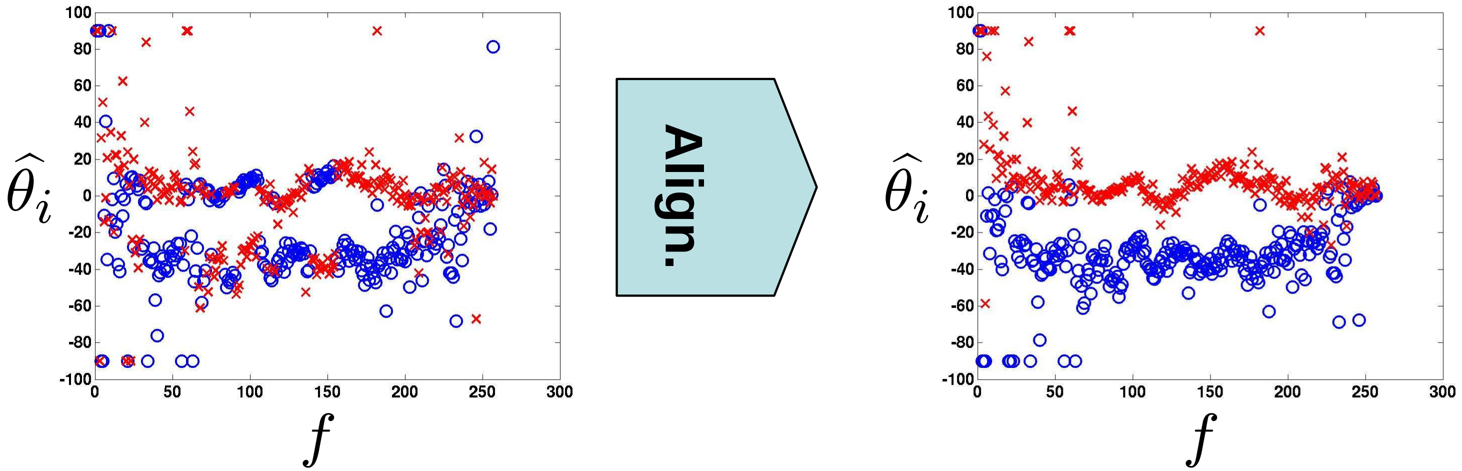
DOA est.

$\widehat{\theta_1(f_1)}$

$\widehat{\theta_2(f_1)}$

$\widehat{\theta_1(f_F)}$

$\widehat{\theta_2(f_F)}$



Outline

1. Preliminaries
    1. FD-BSS & Permutation
    2. Method based on DOA estimation
2. Proposed approach
    1. Spatial aliasing equation
    2. Finding the solution
    3. Permutation resolution
3. Simulation results

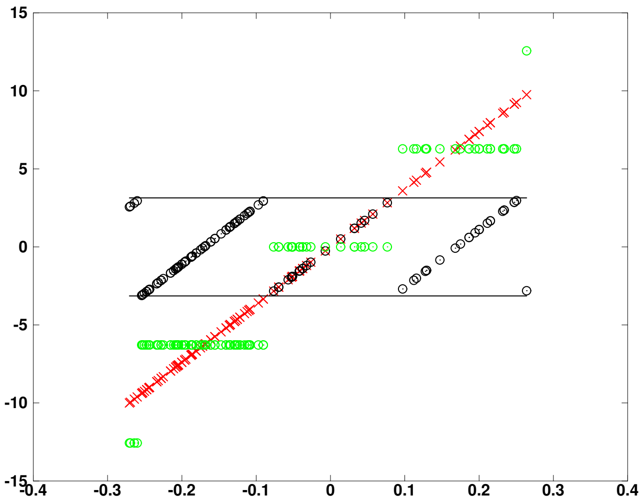
Including spatial aliasing

$$-\pi < 2\pi\frac{f}{c}(\mathbf{p}_i - \mathbf{p}_k)\cdot\mathbf{q}_{\sigma(f,j)} \leq \pi$$ **Not longer verified**

$$-\pi+\delta_{ijkf}2\pi < 2\pi\frac{f}{c}(\mathbf{p}_i - \mathbf{p}_k)\cdot\mathbf{q}_{\sigma(f,j)} \leq \pi+\delta_{ijkf}2\pi.$$

δ_{ijkf} **integers**

$$f = 2000\text{Hz}$$



$(\mathbf{p}_i - \mathbf{p}_k)\cdot\mathbf{q}_{\sigma(f,j)}$

🟥 $2\pi\frac{f}{c}(\mathbf{p}_i - \mathbf{p}_k)\cdot\mathbf{q}_{\sigma(f,j)}$

🟩 $\delta_{ijkf}2\pi$

⬛ $2\pi\frac{f}{c}(\mathbf{p}_i - \mathbf{p}_k)\cdot\mathbf{q}_{\sigma(f,j)} - \delta_{ijkf}2\pi$

Use all microphone pairs

$$\arg\left(\frac{\widehat{\mathbf{A}}_{ij}(f)}{\widehat{\mathbf{A}}_{kj}(f)}\right) = -2\pi\frac{f}{c}(\mathbf{p}_i - \mathbf{p}_k)\cdot\mathbf{q}_{\sigma(f,j)} + 2\pi\delta_{ijkf}.$$

Stack column ratios, sensor differences and aliasing compensations

$$\mathbf{R}_j(f) = \begin{bmatrix} \arg\left(\frac{\widehat{\mathbf{A}}_{1j}(f)}{\widehat{\mathbf{A}}_{2j}(f)}\right) \\ \arg\left(\frac{\widehat{\mathbf{A}}_{2j}(f)}{\widehat{\mathbf{A}}_{3j}(f)}\right) \\ \vdots \\ \arg\left(\frac{\widehat{\mathbf{A}}_{n-1j}(f)}{\widehat{\mathbf{A}}_{nj}(f)}\right) \end{bmatrix} \quad \mathbf{P} = \begin{bmatrix} \mathbf{p}_1 - \mathbf{p}_2 \\ \mathbf{p}_2 - \mathbf{p}_3 \\ \vdots \\ \mathbf{p}_{n-1} - \mathbf{p}_n \end{bmatrix} \qquad \mathbf{\Delta}_j(f) = \begin{bmatrix} \delta_{12jf} \\ \delta_{23jf} \\ \vdots \\ \delta_{nn-1jf} \end{bmatrix}$$

$$\mathbf{R}_j(f) = -2\pi\frac{f}{c}\mathbf{P}\boxed{\mathbf{q}_{\sigma(f,j)}} + 2\pi\boxed{\mathbf{\Delta}_j(f)}$$

Unknown **Unknown**

Equation verified by $\Delta_j(f)$

$$\text{(I)} \quad \mathbf{R}_j(f) = -2\pi\frac{f}{c}\mathbf{P}\mathbf{q}_{\sigma(f,j)} + 2\pi\boldsymbol{\Delta}_j(f)$$

$$\text{(II)} \leftarrow \mathbf{P}\mathbf{P}^{+} \times \text{(I)}$$

left pseudo-inverse

$$\mathbf{P}^{+}\mathbf{P} = \mathbf{I}$$

$$\begin{aligned}\text{(II)} \quad \mathbf{P}\mathbf{P}^{+}\mathbf{R}_j(f) &= -2\pi\frac{f}{c}\mathbf{P}\mathbf{P}^{+}\mathbf{P}\mathbf{q}_{\sigma(f,j)} + 2\pi\mathbf{P}\mathbf{P}^{+}\boldsymbol{\Delta}_j(f)\\ &= -2\pi\frac{f}{c}\mathbf{P}\mathbf{q}_{\sigma(f,j)} + 2\pi\mathbf{P}\mathbf{P}^{+}\boldsymbol{\Delta}_j(f)\end{aligned}$$

$$\text{(II)} - \text{(I)}$$

$$\left(\mathbf{P}\mathbf{P}^{+} - \mathbf{I}\right)\mathbf{R}_j(f) = 2\pi\left(\mathbf{P}\mathbf{P}^{+} - \mathbf{I}\right)\boldsymbol{\Delta}_j(f).$$

$$\left\{\begin{aligned}&\mathbf{D}_j(f) = \frac{1}{2\pi}\left(\mathbf{P}\mathbf{P}^{+} - \mathbf{I}\right)\mathbf{R}_j(f)\\ &\mathbf{C} = \mathbf{P}\mathbf{P}^{+} - \mathbf{I}\\ &\mathbf{X} = \boldsymbol{\Delta}_j(f).\end{aligned}\right.$$

$$\mathbf{C}\mathbf{X} = \mathbf{D}_j(f)$$

Outline

1. Preliminaries
   1. FD-BSS & Permutation
   2. Method based on DOA estimation
2. Proposed approach
   1. Spatial aliasing equation
   2. Finding the solution
   3. Permutation resolution
3. Simulation results

Solution of the equation

$$\mathrm{CX} = \mathrm{D}_j(f)$$

$$\mathbf{C} = \mathbf{P}\mathbf{P}^{+} - \mathbf{I}$$

$$L \times L$$

$$\mathrm{rank}(C) = L - \mathrm{rank}(P)$$

$$\left(\begin{array}{r} \mathbf{v} \in \ker(\mathbf{C}) \rightarrow \mathbf{v} = \mathbf{P}\mathbf{u} \\ \mathbf{u} = \mathbf{P}^{+}\mathbf{v} \end{array} \right)$$

$\mathrm{rank}(P)$ **Depends of array configuration**

➡ **Infinite number of real solutions**

Real solution with minimal norm $\mathrm{X} = \mathrm{C}^{+}\mathrm{D}_j(f)$

Finding integer solution $\mathrm{X} = \Delta_j(f)$ **?**

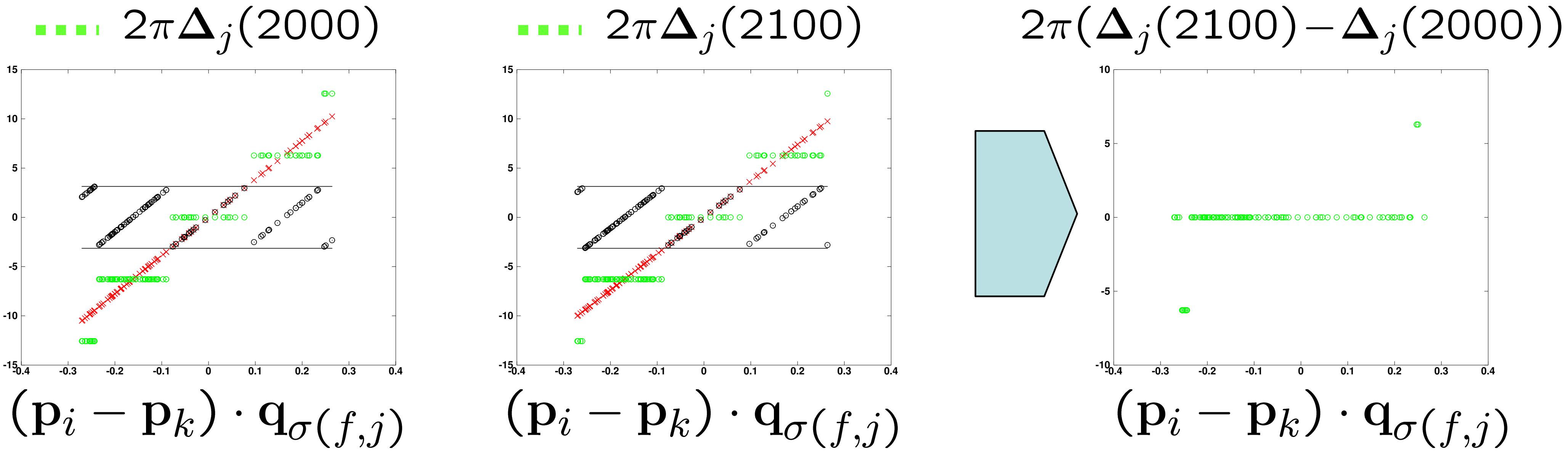
Sparse solution

$\mathbf{\Delta}_j(f)$ **null entries for row corresponding to microphone pair without aliasing**

"Good" initial guess $\widehat{\mathbf{\Delta}}_j(f)$

$\mathbf{\Delta}_j(f) - \widehat{\mathbf{\Delta}}_j(f)$ **sparser than** $\mathbf{\Delta}_j(f)$

Solution of $\mathbf{CX} = \mathbf{D}_j(f) - \mathbf{C}\widehat{\mathbf{\Delta}}_j(f)$



Find the solution

Use previous bin as initial guess $\widehat{\mathbf{\Delta}}_j(f) = \mathbf{\Delta}_j(f-1)$

Start from lower frequencies

$$\begin{cases} \mathbf{X} = \mathbf{C}^{+}\left(\mathbf{D}_j(f) - \mathbf{C}\mathbf{\Delta}_j(f-1)\right) + \mathbf{\Delta}_j(f-1) \\ \mathbf{\Delta}_j(f) = \lceil \mathbf{X} \rfloor \end{cases}$$

Goal: having $\mathbf{X}$ **close to** $\mathbf{\Delta}_j(f)$

Residual $\|\mathbf{C}\left(\mathbf{\Delta}_j(f) - \mathbf{D}_j(f)\right)\|$

Outline

1. Preliminaries
   1. FD-BSS & Permutation
   2. Method based on DOA estimation
2. Proposed approach
   1. Spatial aliasing equation
   2. Finding the solution
   3. Permutation resolution
3. Simulation results

Permutation resolution

$\widehat{\mathbf{A}}(f-1)$

permutation
$k \rightleftharpoons j$

$\widehat{\mathbf{A}}(f)$

Equation in bin f **for row** k $\mathbf{CX} = \mathbf{D}_k(f) - \mathbf{C}\widehat{\mathbf{\Delta}}_j(f-1).$

$\widehat{\mathbf{\Delta}}_j(f-1) \neq \widehat{\mathbf{\Delta}}_k(f-1)$ **Close integer solution not found**

Residual $\left\|\mathbf{C}\left(\mathbf{\Delta}_j(f) - \mathbf{D}_j(f)\right)\right\|$ **large**

- **Compare to a threshold (hard to fix threshold with noise)**
- **Compare residual for different rows (more robust)**

$\widehat{\mathbf{\Delta}}_j(f-1) \approx \widehat{\mathbf{\Delta}}_k(f-1)$ **Close integer solution found**

Residual $\left\|\mathbf{C}\left(\mathbf{\Delta}_j(f) - \mathbf{D}_j(f)\right)\right\|$ **small**

Need to estimate $\mathbf{q}_{\sigma(f,j)} = -\frac{1}{2\pi f}\mathbf{P}^{+}(\mathbf{R}_j(f) + 2\pi\widehat{\mathbf{\Delta}}_j(f)).$

Post processing

Consider all frequency bins with all rows having small residuals

↓

Estimate the absolute DOAs using these frequency bins

↓

Average the DOAs along these frequency bins

↓

For all frequency bins and all rows,
compute the distances between the DOA and these averaged absolute DOAs

↓

cluster the rows according to these distances

Outline

1. Preliminaries
    1. FD-BSS & Permutation
    2. Method based on DOA estimation
2. Proposed approach
    1. Spatial aliasing equation
    2. Finding the solution
    3. Permutation resolution
3. Simulation results

Simulation model

16 microphones uniform circular array (diam. 31cm, minimum spacing 3.12cm)

16kHz sampling frequency and 512 point fft

Model for the estimated mixing matrix $\widehat{\mathbf{A}}_{ij}(f) = \frac{1}{\lambda_1} e^{-j2\pi \frac{f}{c} \mathbf{p}_i \cdot \tilde{\mathbf{q}}_j} + \epsilon,$

$$\tilde{\mathbf{q}}_j = \begin{bmatrix} \cos(\theta)\sin(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\phi) \end{bmatrix}$$

Angle errors

$\{\theta + \epsilon_2, \phi + \epsilon_3\}$

$\{\epsilon_2, \epsilon_3\}$ **Uniform in** $[-\gamma, \gamma]$

Additive noise ϵ **Zero mean Gaussian with variance** β

Randomly permutation of two columns for a percentage *d* of the frequency bins

Measure of performance

r Percentage of frequency bins with adequate permutation

e Mean square error of the absolute DOA (in dB)

Experiment 1 Influence of $\epsilon, \ \gamma, \ d$

$$\tilde{\mathrm{q}}_1 \ \{40, 30\}$$

$$\tilde{\mathrm{q}}_2 \ \{-10, 25\}$$

Average results on 100 independent runs

Experiment 2 Influence of angle between $\tilde{\mathrm{q}}_1, \ \tilde{\mathrm{q}}_1$

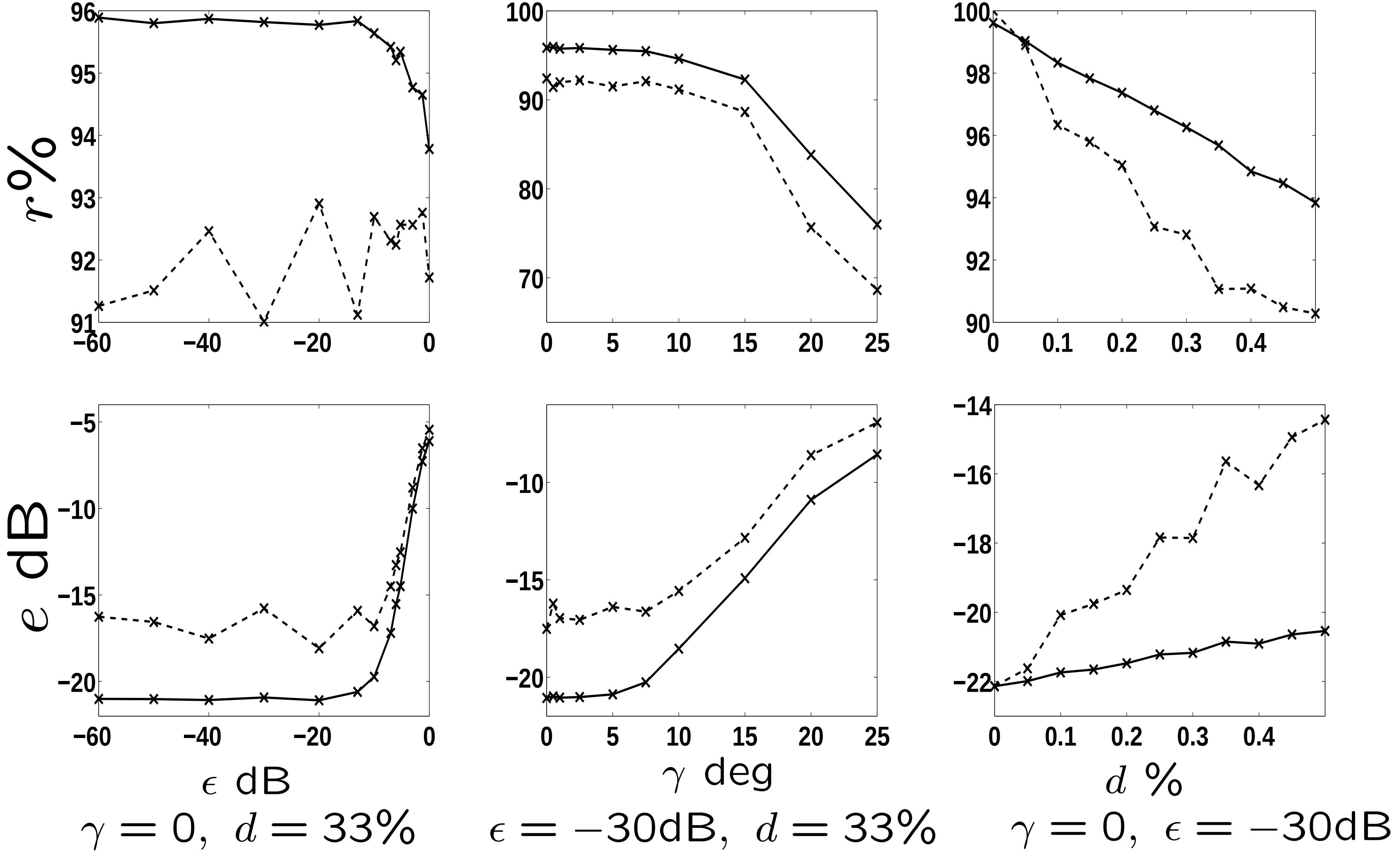
$$\epsilon = -30\mathrm{dB}, \ \gamma = 0, \ d = 33\%$$

Average results on 100 independent runs

Results 1

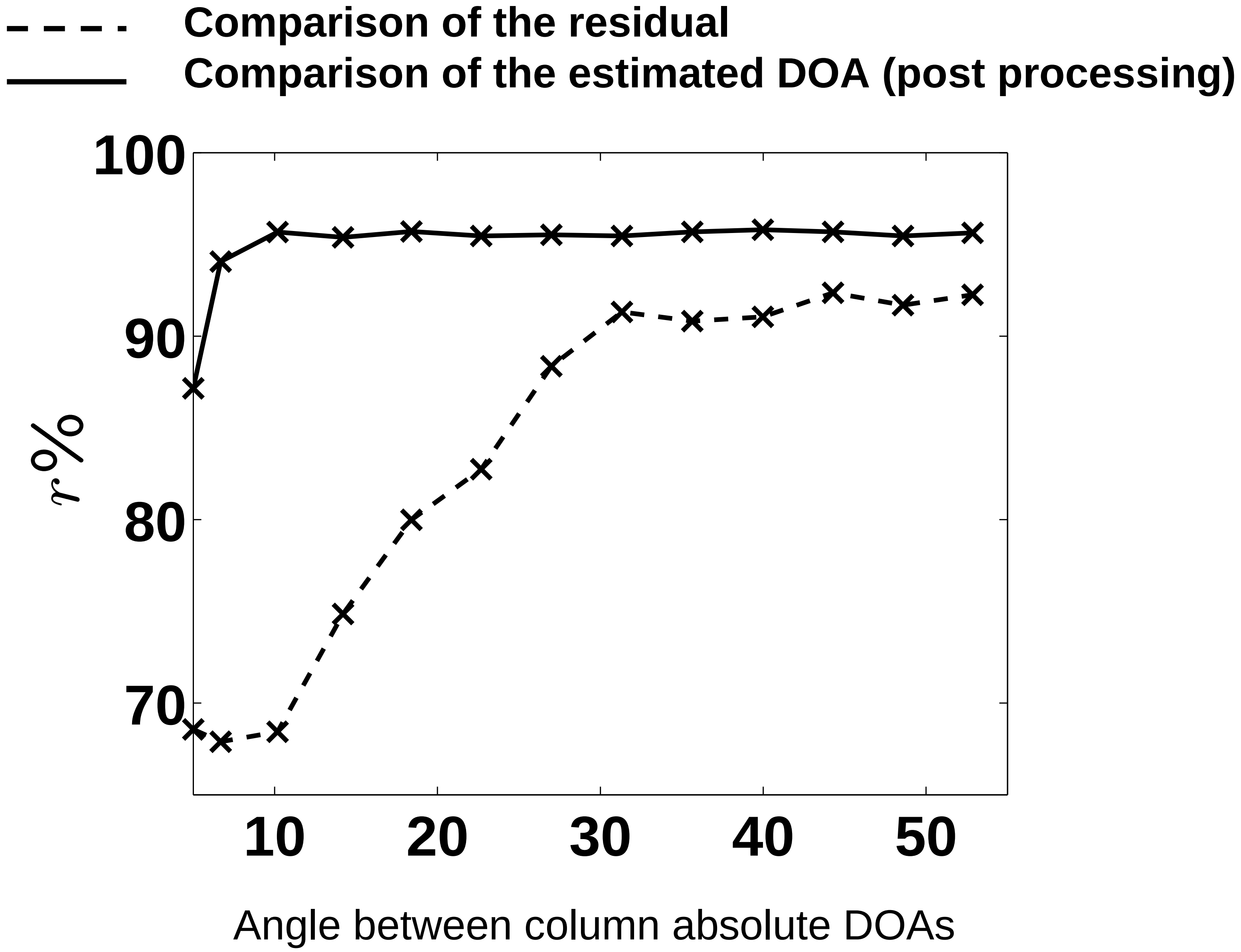
- - - - Comparison of the residual

—— Comparison of the estimated DOA (post processing)



$\gamma = 0,\ d = 33\%$ $\epsilon = -30\text{dB},\ d = 33\%$ $\gamma = 0,\ \epsilon = -30\text{dB}$

Results 2



Comparison of the residual
Comparison of the estimated DOA (post processing)
100
90
80
70
r%
10
20
30
40
50
Angle between column absolute DOAs

Outline

1. Preliminaries
    1. FD-BSS & Permutation
    2. Method based on DOA estimation
2. Proposed approach
    1. Spatial aliasing equation
    2. Finding the solution
    3. Permutation resolution
3. Simulation results

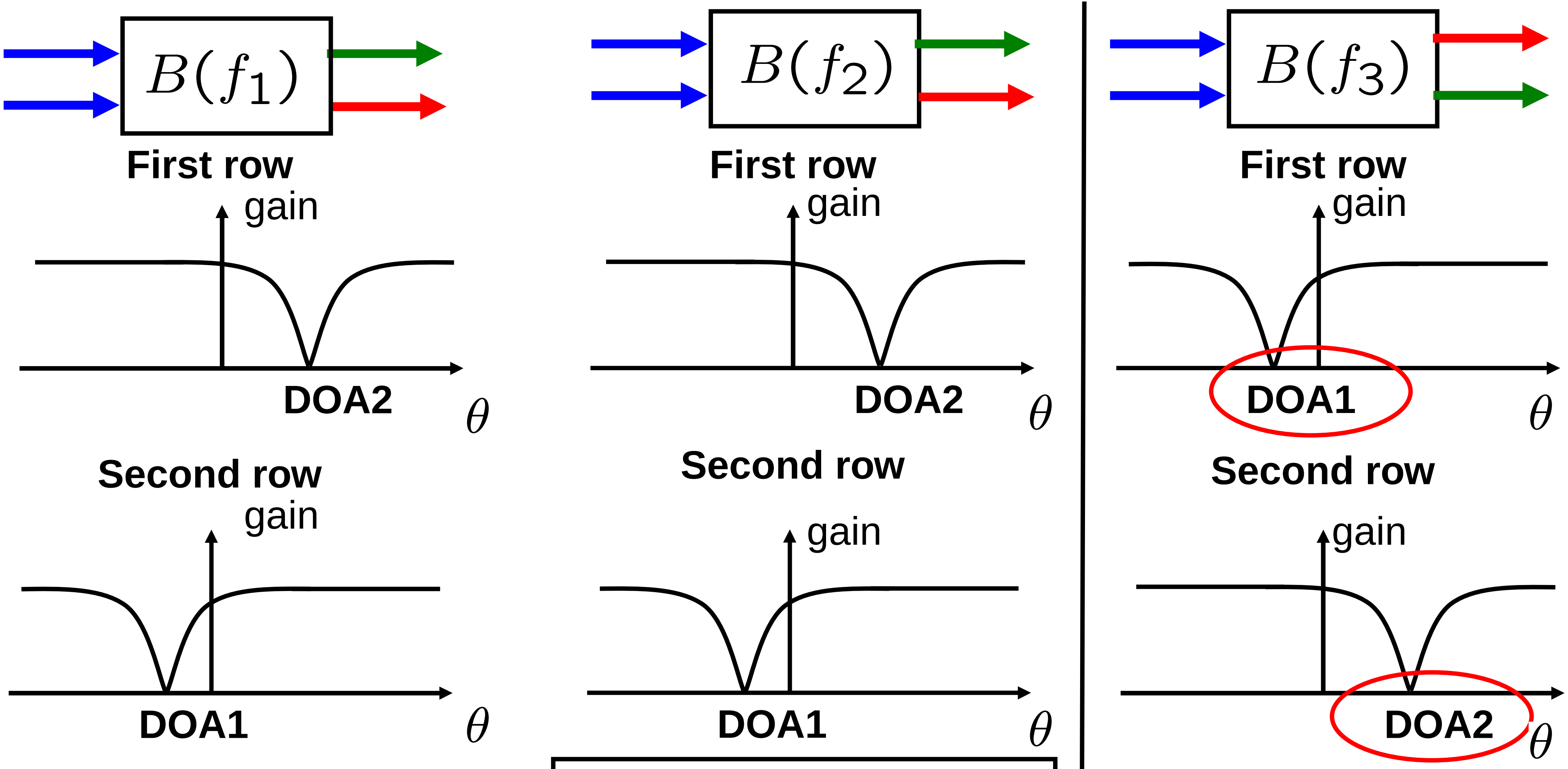
Conclusion

- Blind signal separation (BSS)
- Frequency domain
- Permutation resolution
- DOA based approach
- Spatial aliasing

- Model for the spatial aliasing
- Sparse solution
- Permutation resolution

- Better resolution of the equation
- Apply to real data

Use of directivity pattern



Efficient suppression of Point sources

- 🟩 → **Speech 1 DOA1**
- 🟥 → **Speech 2 DOA 2**
- 🟦 → **Mixture**

Permuted