Optimal Spatial Sampling of Plant Transfer Functions for Head-Tracked Personal Sound Zones

Yue Qiao* (presenter) & Edgar Choueiri

3D Audio and Applied Acoustics (3D3A) Lab

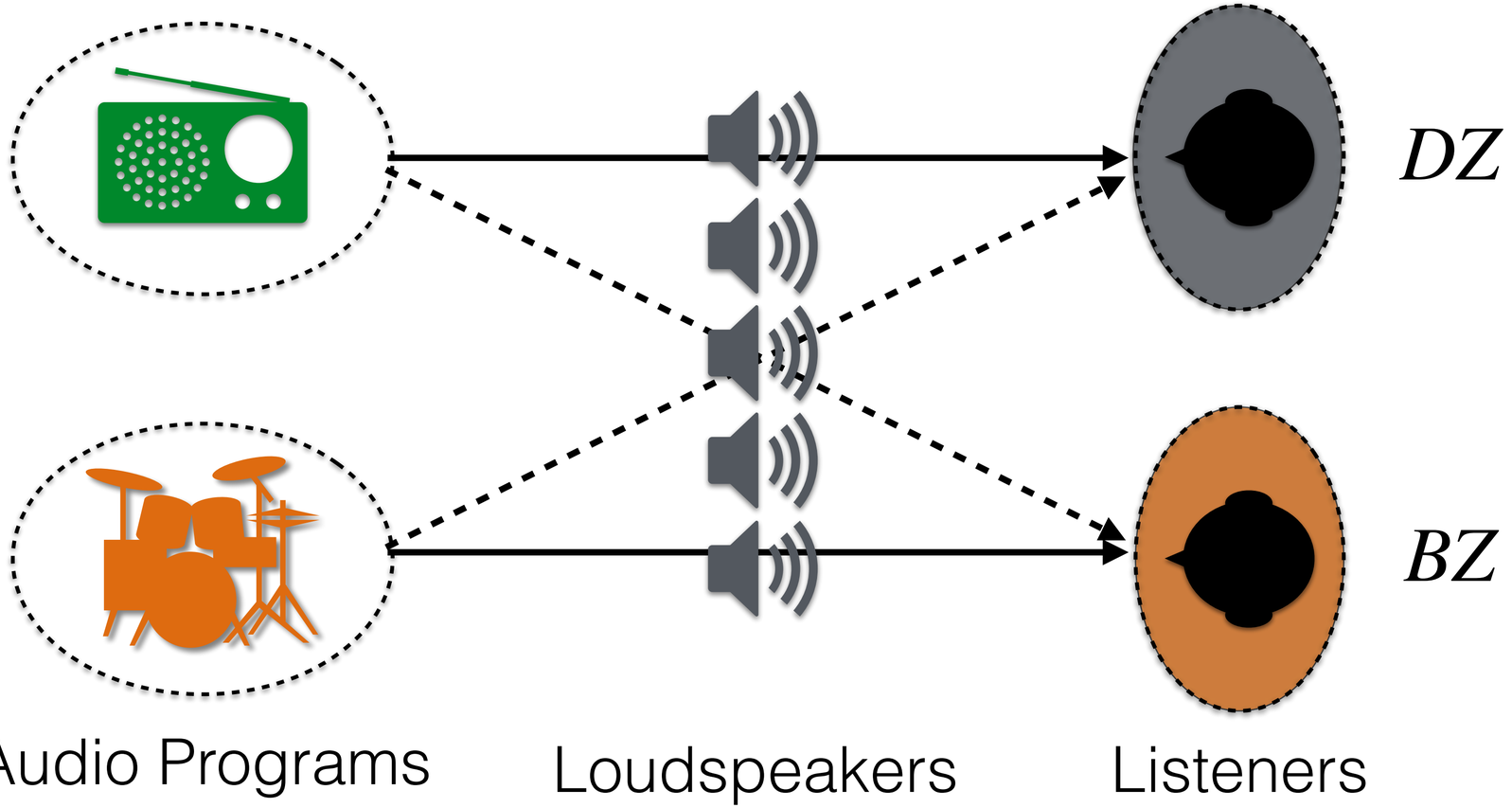
Princeton University

Presented at the 154th AES Convention

May 15, 2023

*E-mail: yqiao@princeton.edu

Personal Sound Zones[1]



[1] Druyvesteyn and Garas, JAES, 1997

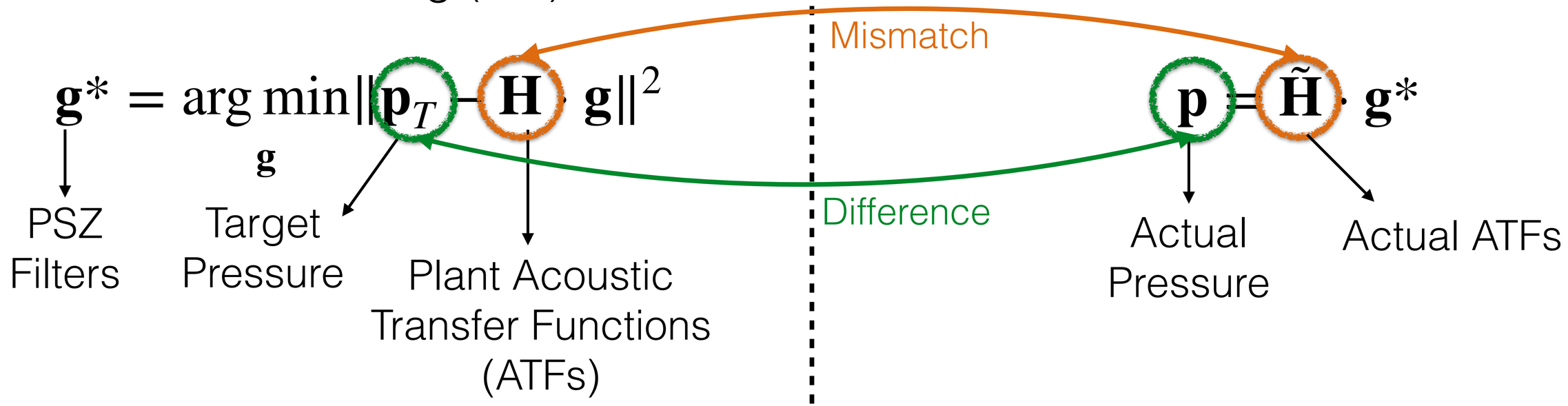
PSZ Setup

PSZ Playback

The Pressure Matching (PM)[2] Method

$$\mathbf{g}^* = \arg\min_{\mathbf{g}} \|\mathbf{p}_T - \mathbf{H} \cdot \mathbf{g}\|^2$$

$$\mathbf{p} = \tilde{\mathbf{H}} \cdot \mathbf{g}^*$$

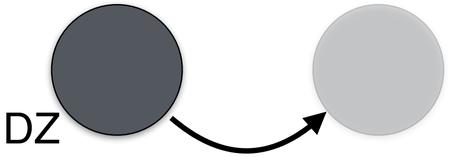


[2] Poletti, AES Conv. 125, 2008

The Importance of Head Tracking in PSZ

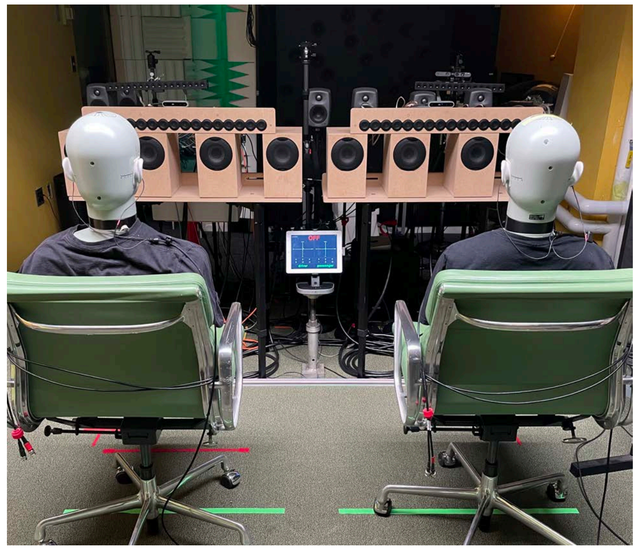
Head movements → Mismatched plant ATFs → Degraded PSZ isolation

$\mathbf{H} \quad \tilde{\mathbf{H}}$



Especially true for **ear-targeting** *PSZ systems*

[3] Vindrola et al., JAES, 2020



[4] Qiao and Choueiri, AES Conv. 152, 2022

Solutions for Head-Tracked Reproduction

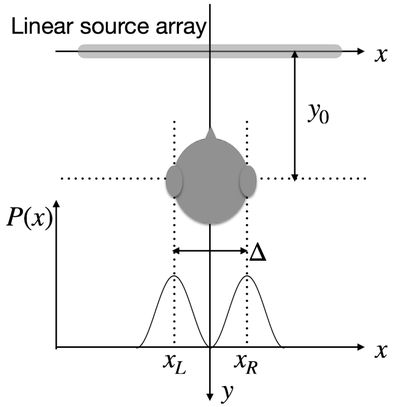
Approaches

- Dynamic loudspeaker beamforming[5]
- Adaptive filtering (Filtered-x Least-Mean-Square[6], Recursive Least Square)
- Filter cross-fading with plant spatial sampling[7]

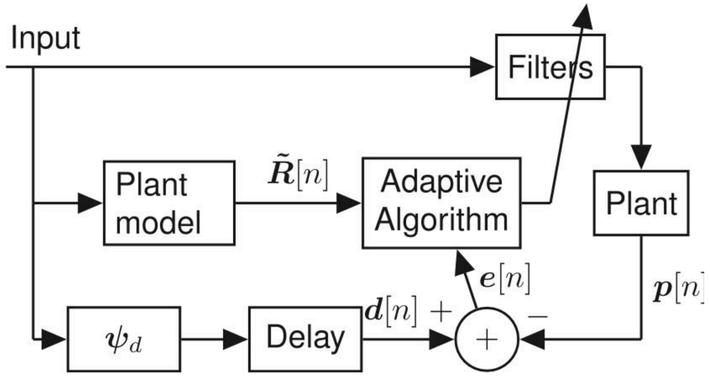
Modeling accuracy ↑

Isolation performance ↑

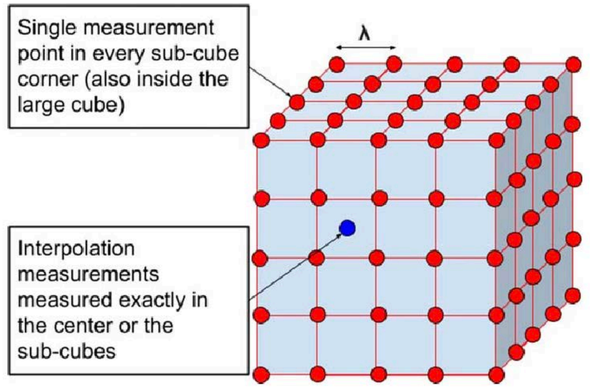
Implementation cost ↑



[5] Qiao and Choueiri, AES Conv. 151, 2021



[6] Vindrola et al., JASA, 2021



[7] Lindfors et al., JAES, 2022

Challenges with implementing head-tracked PSZ systems…

	Degrees-of-freedom in head movements	Number of plant ATF channels	Performance Requirement[8]
Single-listener Crosstalk Cancellation systems	6	2x2	~20.7 dB for envelopment
Two-listener PSZ systems	**6^2**	**4xN**	~**25.6** dB for non-distraction

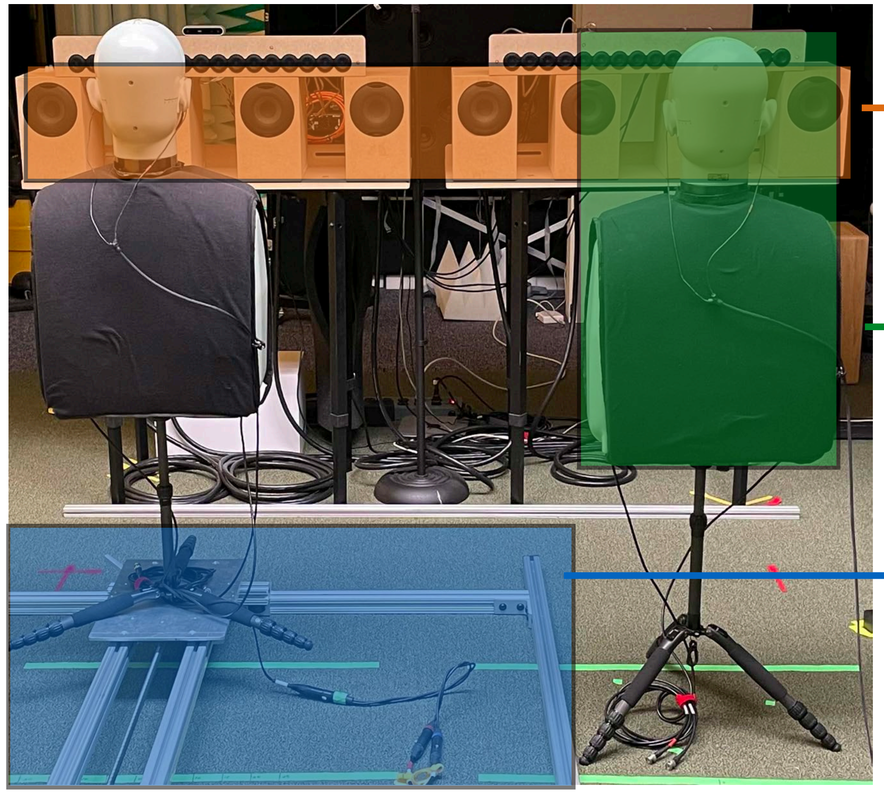
The implementation of head-tracked PSZ seems practically impossible!

What is the minimum required spatial sampling resolution?

What are the rules for optimizing the spatial sampling process?

[8] Canter and Coleman, AES Conv. 150, 2021

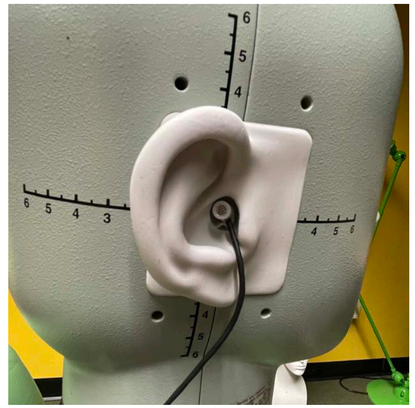
Experimental system setup



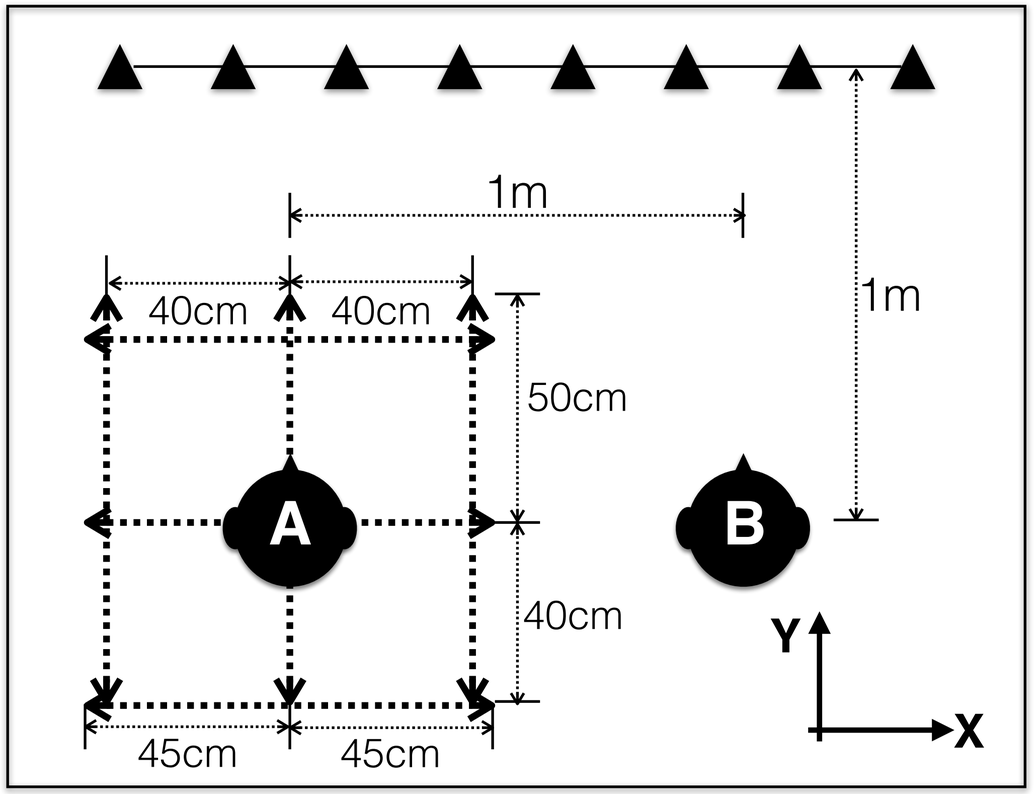
Loudspeaker array
(200~7000Hz)

B&K HATS
dummy head

Mechanical
translation stage



Binaural
Microphones



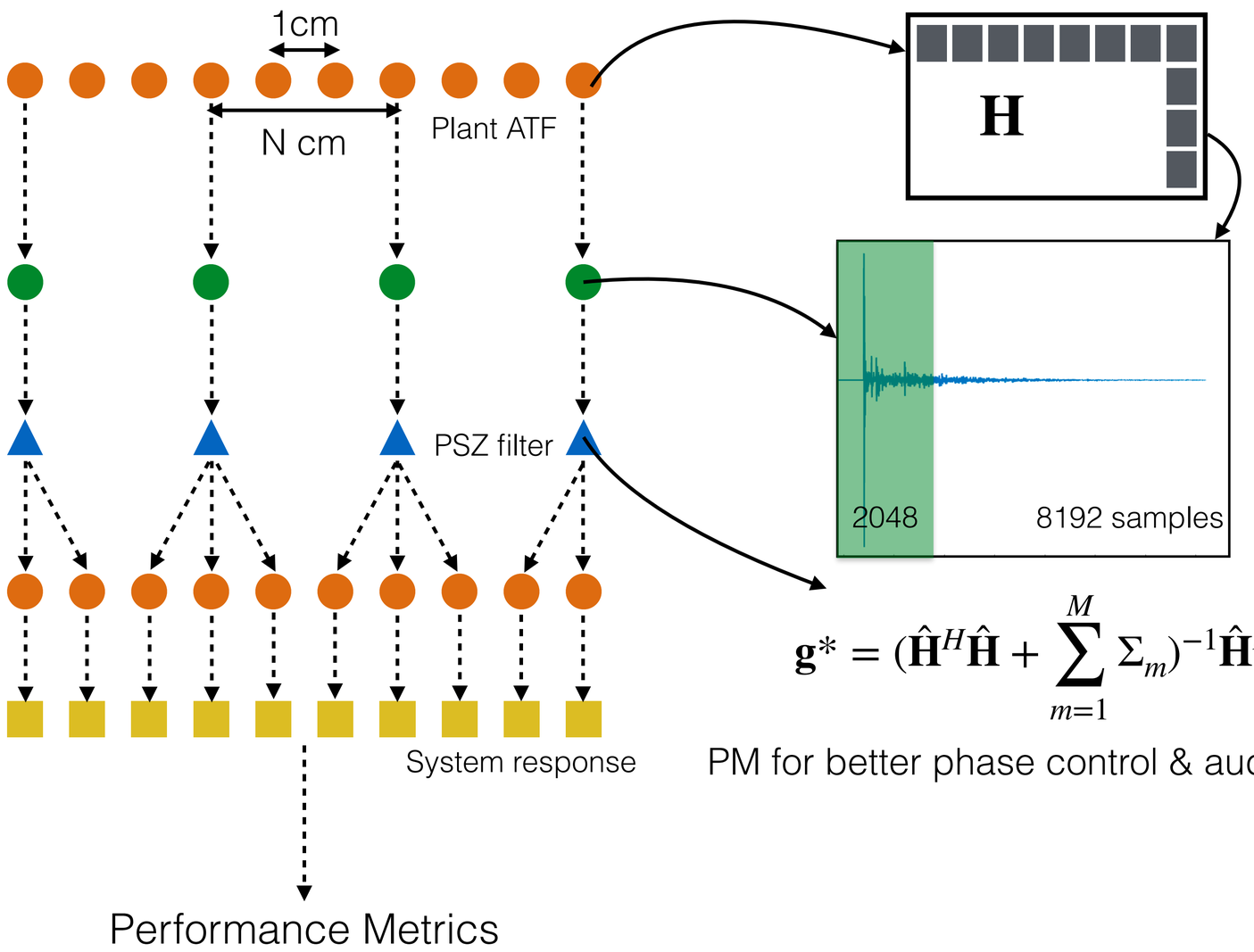
Plant sampling resolution: $\Delta x = \Delta y = 1cm$

Evaluation Procedure

Step I.
Plant spatial sampling

Step II.
PSZ filter generation

Step III.
Performance evaluation



$$\mathbf{g}^* = (\hat{\mathbf{H}}^H\hat{\mathbf{H}} + \sum_{m=1}^{M}\Sigma_m)^{-1}\hat{\mathbf{H}}^H\mathbf{p}_T$$

PM for better phase control & audio quality

Evaluation Metrics

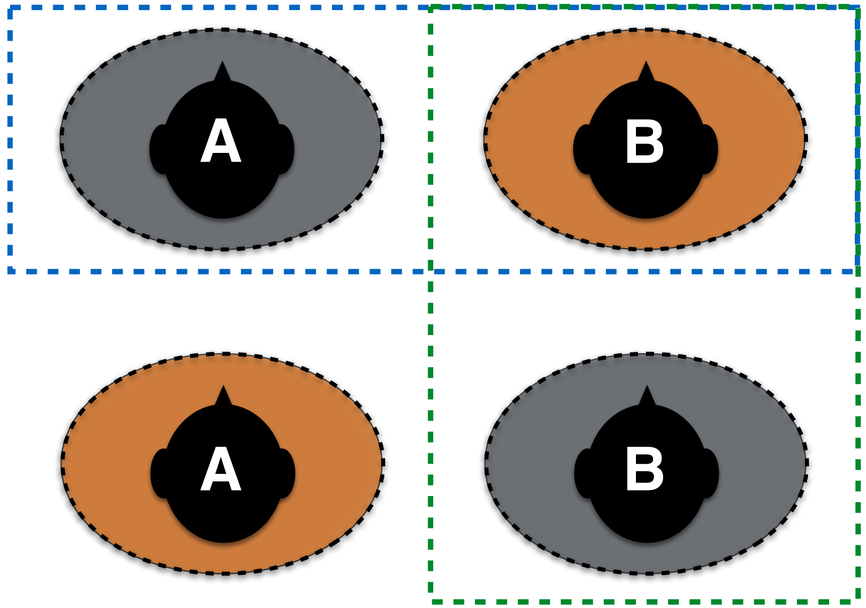
Two aspects of isolation[9]

Between BZ and DZ:
Inter-Zone Isolation (IZI)

$$IZI_2 = \frac{\|\mathbf{H}_2\mathbf{g}_2^*\|^2}{\|\mathbf{H}_1\mathbf{g}_2^*\|^2}$$

Between target and interfering programs:
Inter-Program Isolation (IPI)

$$IPI_2 = \frac{\|\mathbf{H}_2\mathbf{g}_2^*\|^2}{\|\mathbf{H}_2\mathbf{g}_1^*\|^2}$$



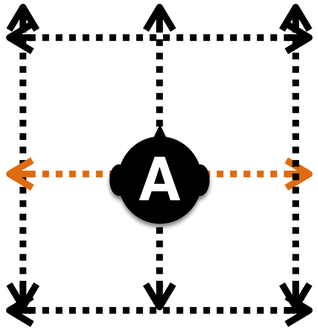
Only consider right listener being in *BZ*

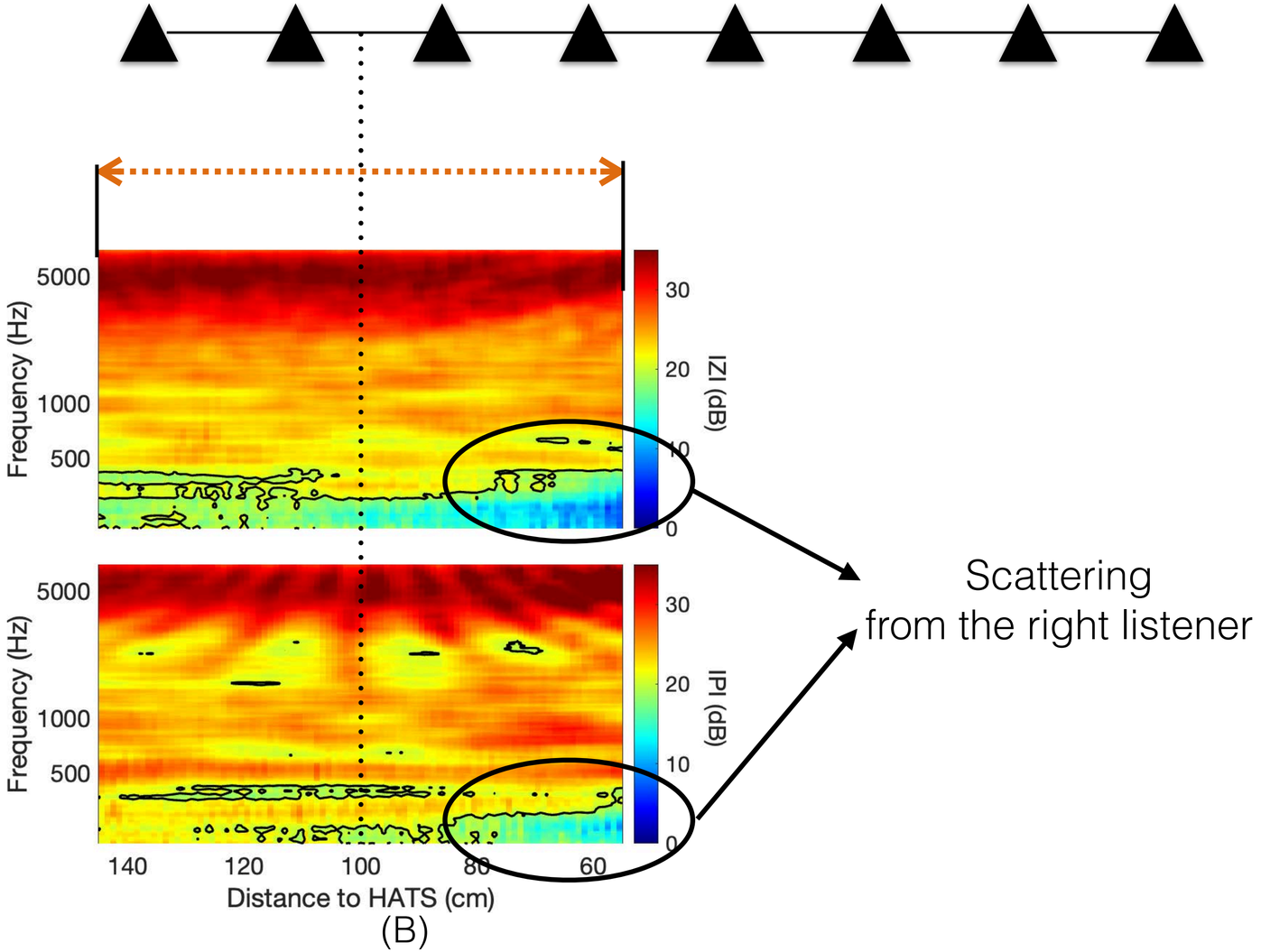
For the left moving listener: IZI ~ moving *DZ*, IPI ~ moving *BZ*

[9] Qiao et al., JASA Express Lett., 2022

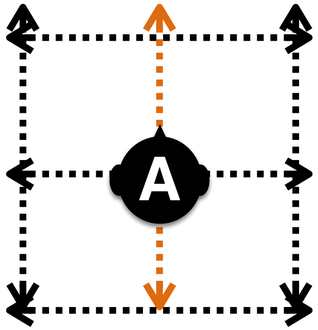
Results — Full Resolution

- X movements

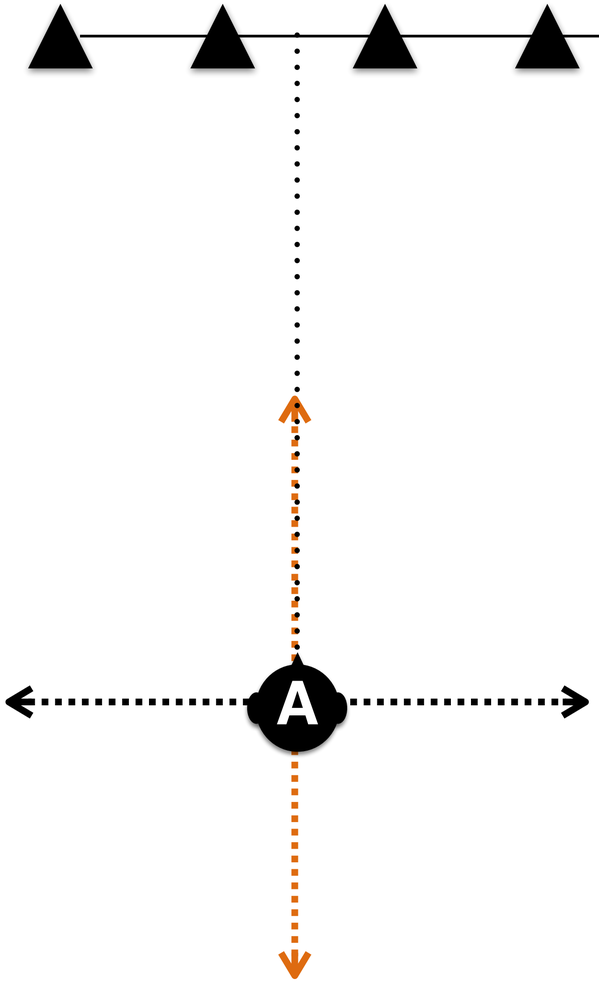


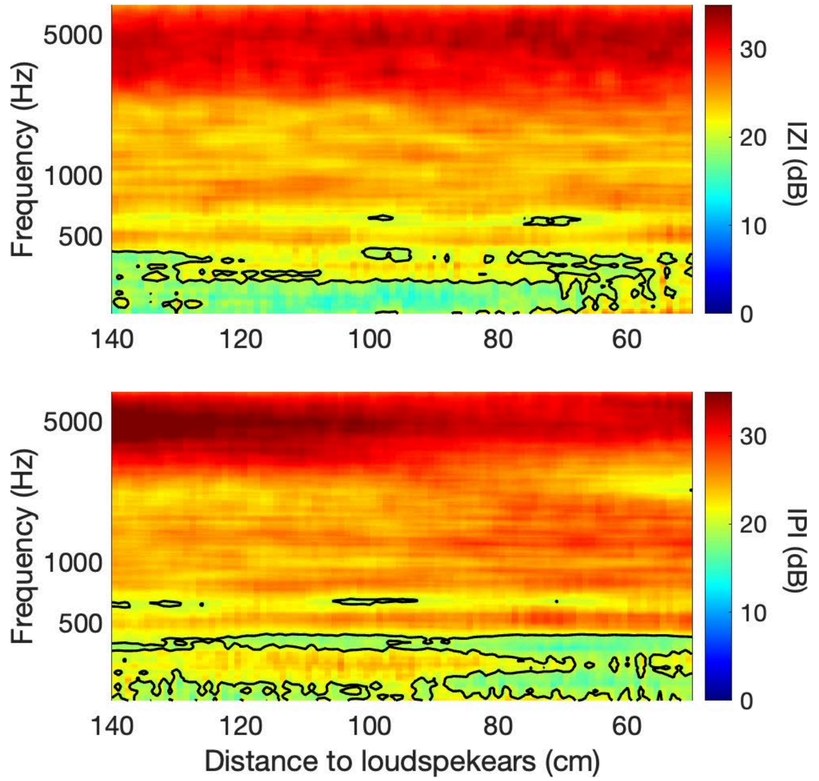


- Y movements


A


B


A

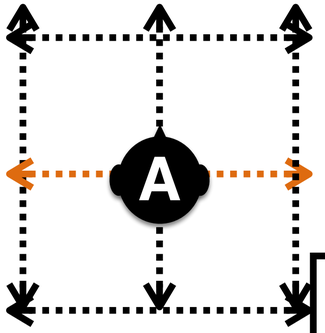

5000
1000
500
Frequency (Hz)
140
120
100
80
60
30
20
10
0
IZI (dB)
5000
1000
500
Frequency (Hz)
140
120
100
80
60
30
20
10
0
IPI (dB)
Distance to loudspekears (cm)

Takeaways

- In the best case scenario, high isolation can be preserved over a large area
- Isolation is inherently lower at low frequencies due to room modes, setup limitations, etc.

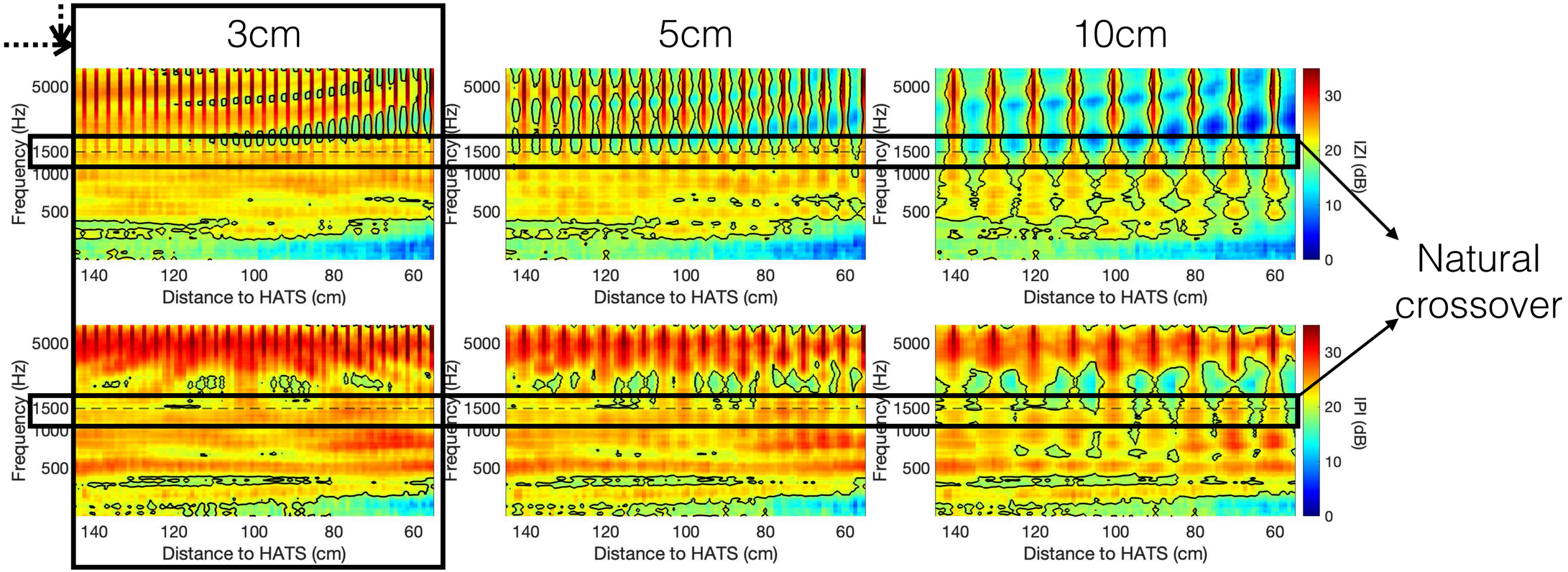
Results — Spatial Downsampling

- X movements

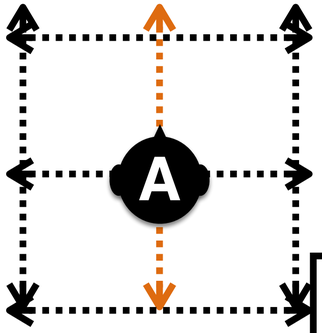


IPI (moving *BZ*) is more robust than IZI (moving *DZ*)

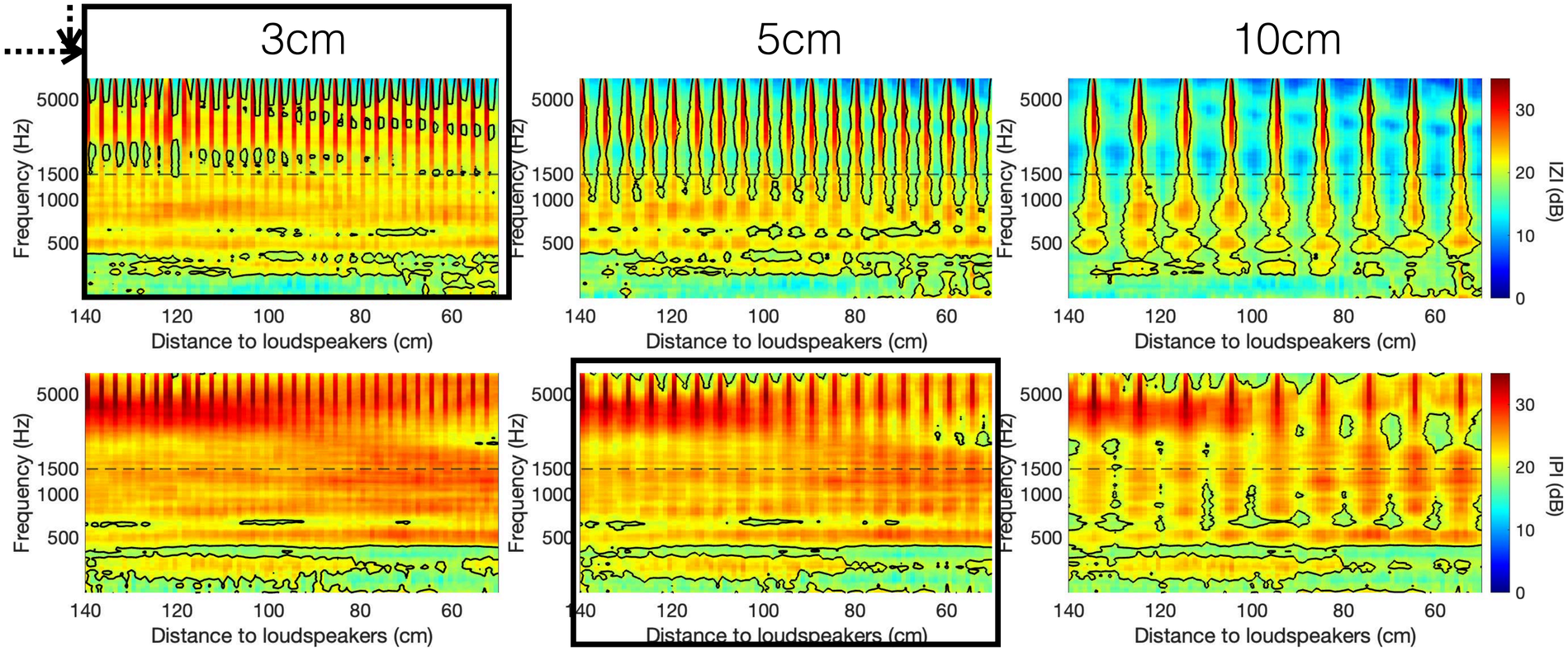
3cm for both IZI and IPI



- Y movements

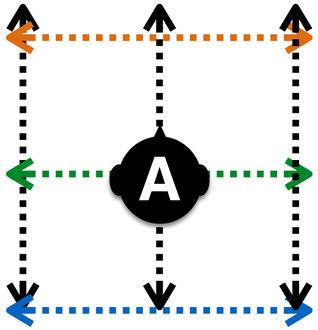


3cm for IZI and 5cm for IPI



Results — Shifting Positions

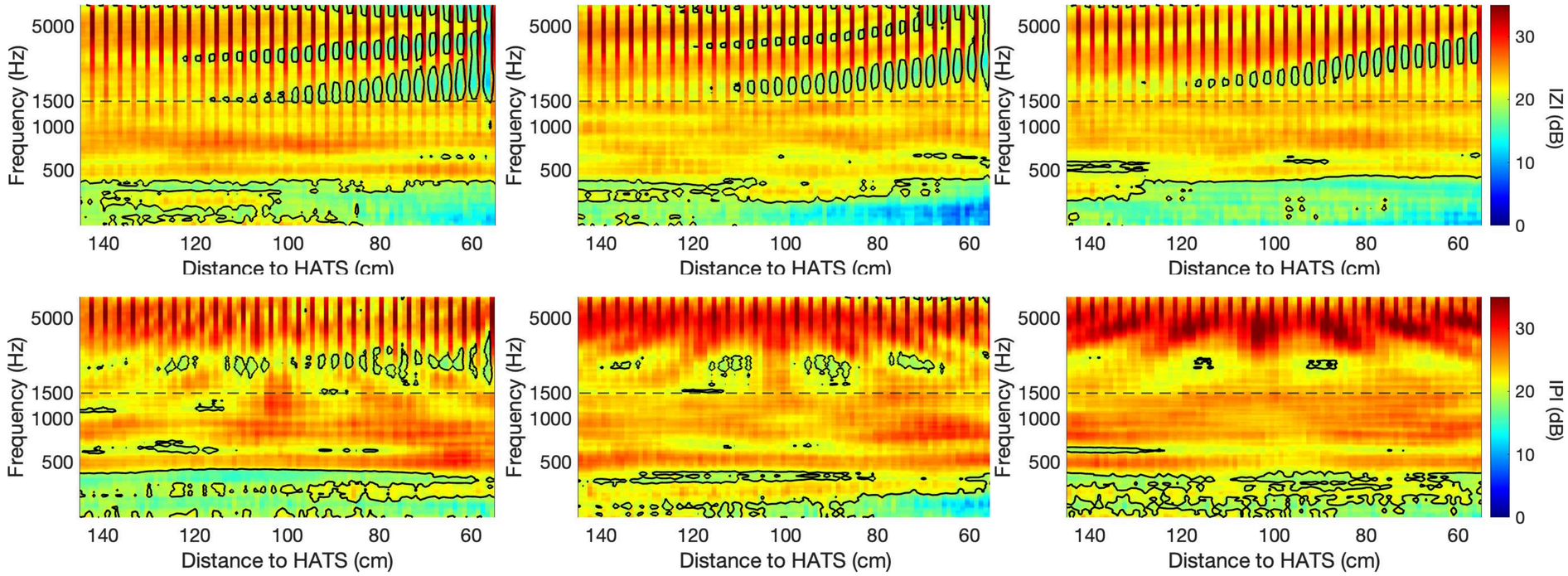
- X movements, shifting in Y



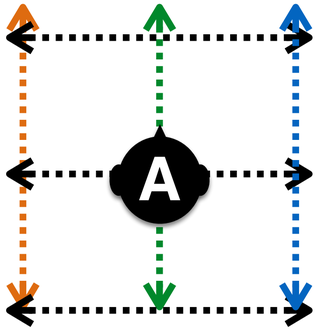
Filter robustness

Near → Far from the array

Front **Mid** **Back**



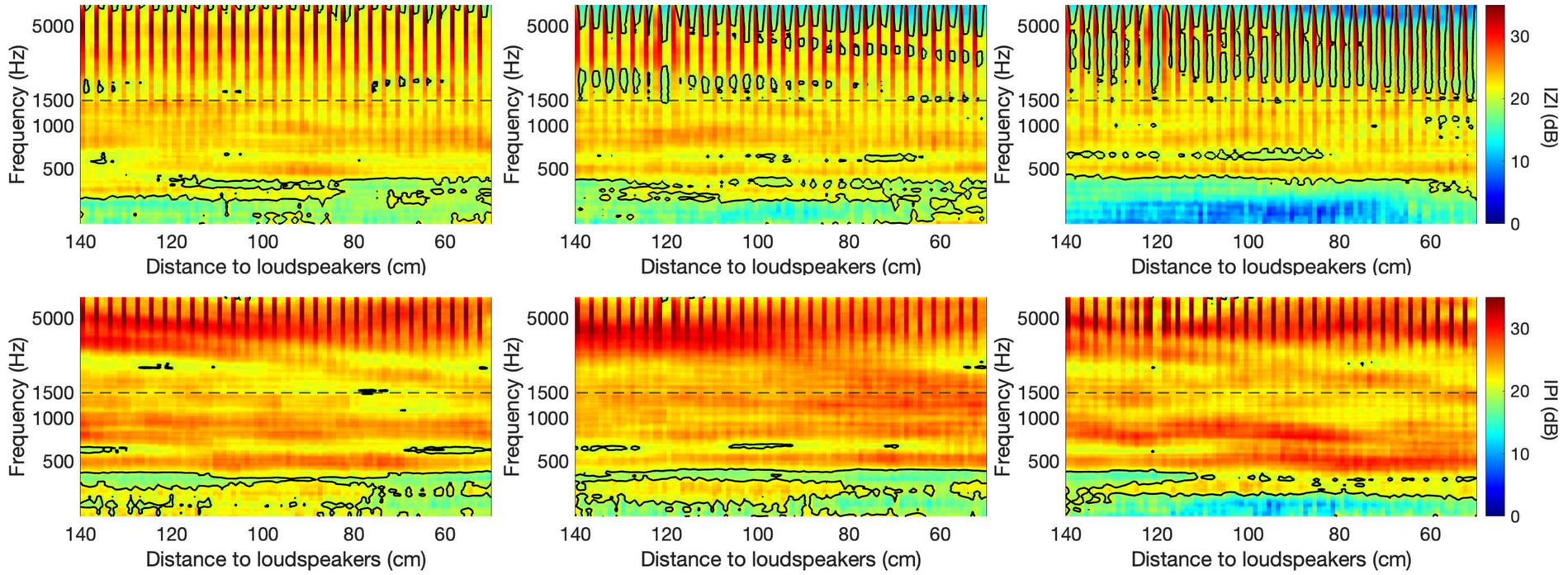
- Y movements, shifting in X



Filter robustness ↓

Far from → Near Listener B

Left **Mid** **Right**



What are the rules for optimizing the spatial sampling process?

- 2 Distances
  - the distance between **two listeners** ↓ , sampling resolution ↑
  - the distance between **the listener and the array** ↓ , sampling resolution ↑
- Temporal frequency: frequency ↑ , sampling resolution ↑
- *BZ*/*DZ*: for **moving DZ**, sampling resolution ↑ ; for **moving BZ**, resolution ↓

Discussion and Future Work

- The qualitative rules are generalizable to other PSZ systems of similar dimensions
- The findings suggest a sampling of *BZ* and *DZ* at different resolutions
- The findings are also insightful for implementing interpolation/adaptive filtering
- The observed crossover frequency (1500 Hz) can be used for splitting approaches
- Future work: investigating spatial sampling of head rotations

Optimal Spatial Sampling of Plant Transfer Functions for Head-Tracked Personal Sound Zones

Yue Qiao* (presenter) & Edgar Choueiri

3D Audio and Applied Acoustics (3D3A) Lab

Princeton University

Presented at the 154th AES Convention

May 15, 2023

*E-mail: yqiao@princeton.edu