AES EUROPE SPRING 2022
AUDIO ENGINEERING CONVENTION

The Performance of A Personal Sound Zone System with Generic and Individualized Binaural Room Transfer Functions

Yue Qiao, Edgar Choueiri

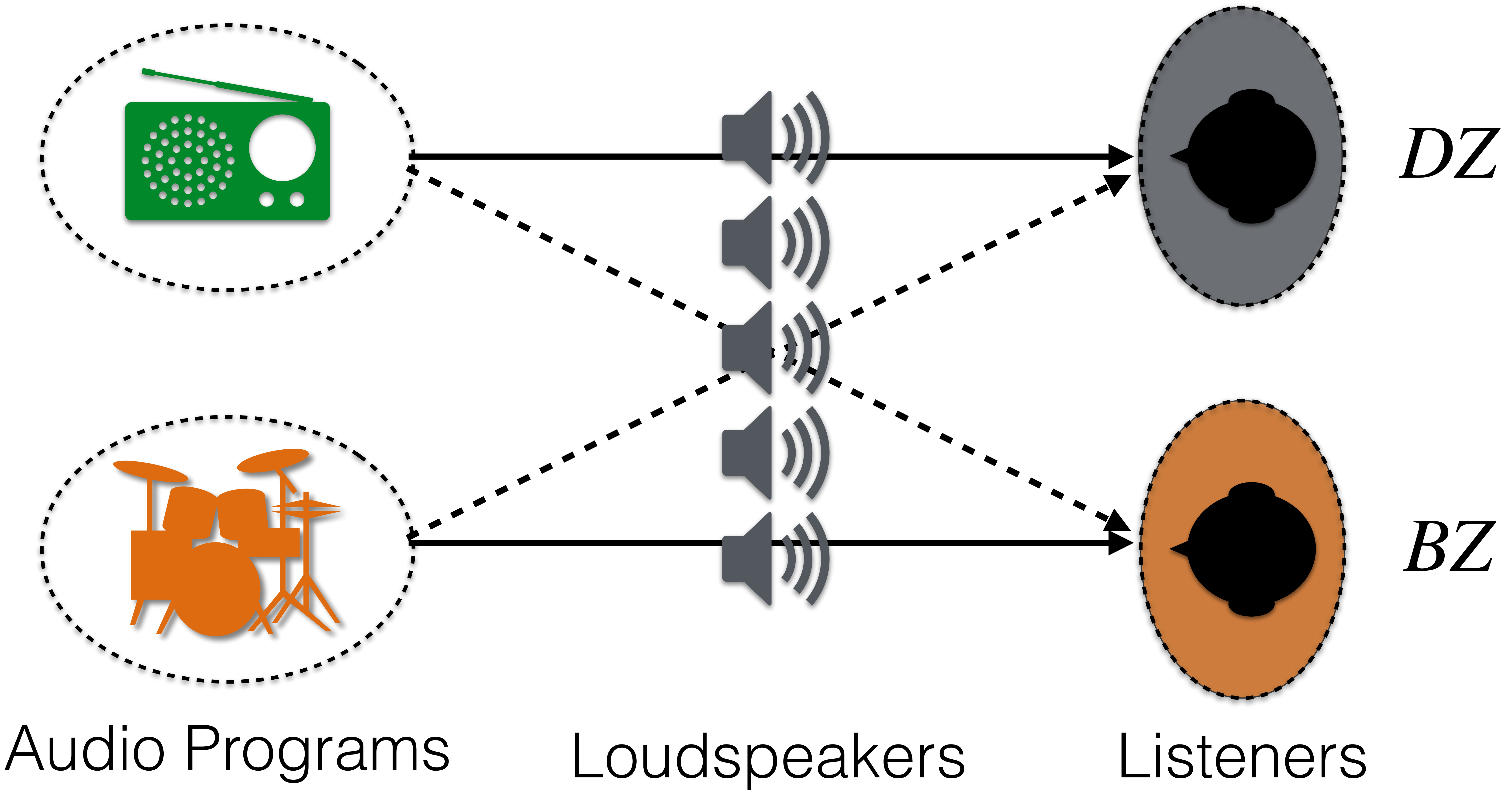
The Performance of A Personal Sound Zone System with Generic and Individualized Binaural Room Transfer Functions

Yue Qiao* (presenter) & Edgar Choueiri
3D Audio and Applied Acoustics (3D3A) Lab
Princeton University

Presented at the 152nd AES Convention
May 17, 2022

*E-mail: yqiao@princeton.edu

Concept of PSZ[1]



[1] Druyvesteyn and Garas, JAES, 1997

PSZ Setup

PSZ Playback

The Pressure Matching (PM)[2] Method

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

Mismatch

Difference

PSZ Filters

Target Pressure

Transfer Functions (TFs)

Actual Pressure

Mismatched TFs ⇒ Degraded PSZs

[2] Poletti, AES Conv. 125, 2008

Sources of TF Mismatches

Head-Related Transfer Functions (HRTFs)

Binaural Room Transfer Functions (BRTFs)

Sound speed, Ambient temperature, background noise…

Loudspeaker/Microphone Positions & Responses

Number of Passengers, Seat Positions…

BRTF Mismatch in PSZ

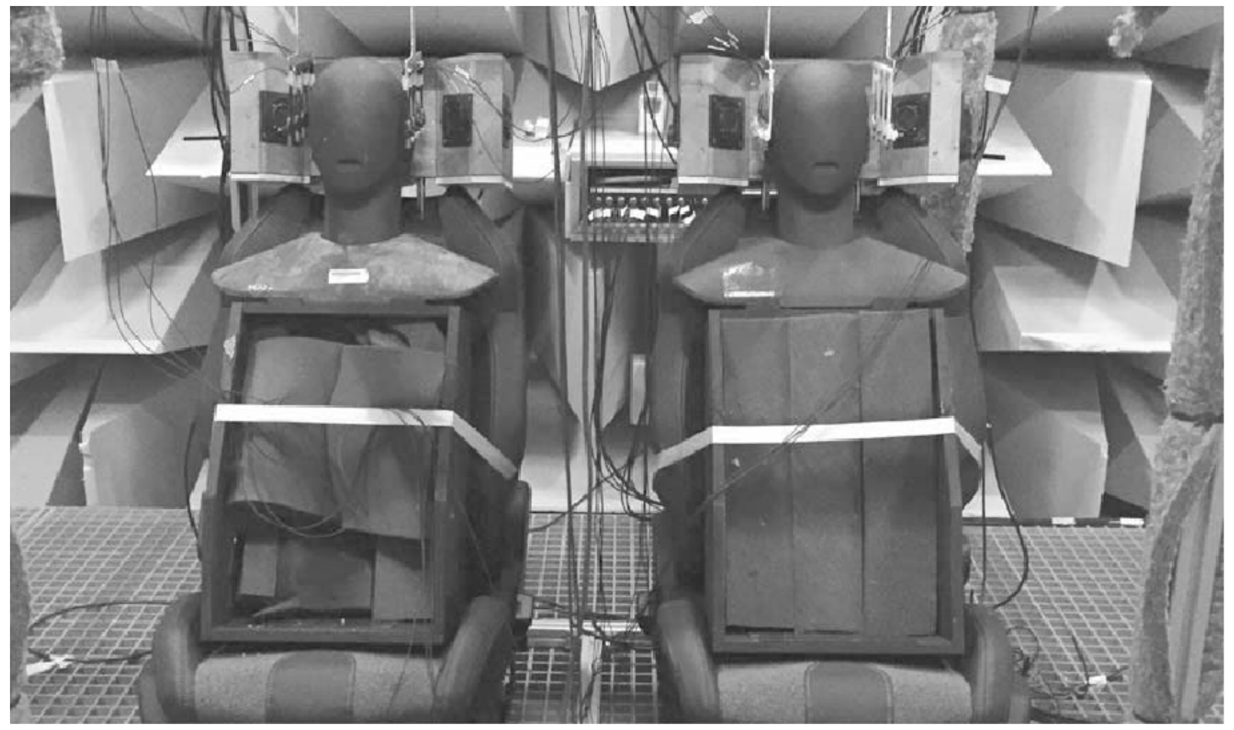
Head-Related Transfer Functions (HRTFs)

Binaural Room Transfer Functions (BRTFs)

Why not use *actual listeners'* BRTFs in both phases?

Setup Phase

Cho and Chang, ICA, 2019



Vindrola et al., JAES, 2020

Playback Phase

Ebri et al., AES Conv., 2020

Molés-Cases et al., JASA, 2022

Individualizing BRTFs

Benefits

- For binaural reproduction
  - Improves localization accuracy with headphone-based systems[3]
  - Improves performance for crosstalk cancellation systems[4]
- For PSZ systems?

Potential Barriers

- Time, cost, and equipment
- Additional measurements & signal processing
- Availability
- …

Does individualization for PSZ make a difference?

If so, is the improvement worth the efforts?

[3] Wenzel et al., JASA, 1993 [4] Akeroyd et al., JASA, 2007

Robust, but not too robust

Regularized PM Solution

$$\tilde{\mathbf{g}}^* = (\mathbf{H}^H\mathbf{H} + \beta\mathbf{I})^{-1}\mathbf{H}^H\mathbf{p}_T$$

β ↓ Regularization

Regularization ↗ { Robustness ↗ ; Best Performance ↘ }

Optimize regularization s.t. PSZ filters are robust to

- small head misalignments of a single listener ✓
- BRTF mismatch between listeners ✗
- larger head movements, torso movements, etc. N/A

Robust, but not too robust

A Probabilistic approach[5]

$\mathbf{H}$ as random variables

$$H_{ml} = A_{ml} e^{i\phi_{ml}}$$

$$A_{ml} \sim N(\hat{A}_{ml}, \sigma^2_{A,ml})$$

$$\phi_{ml} \sim N(\hat{\phi}_{ml}, \sigma^2_{\phi,ml})$$

Modified Cost Function

$$J_{prob} = \mathbb{E}\{\|\mathbf{H}\mathbf{g} - \mathbf{p}_T\|^2\}$$

Optimal Solution

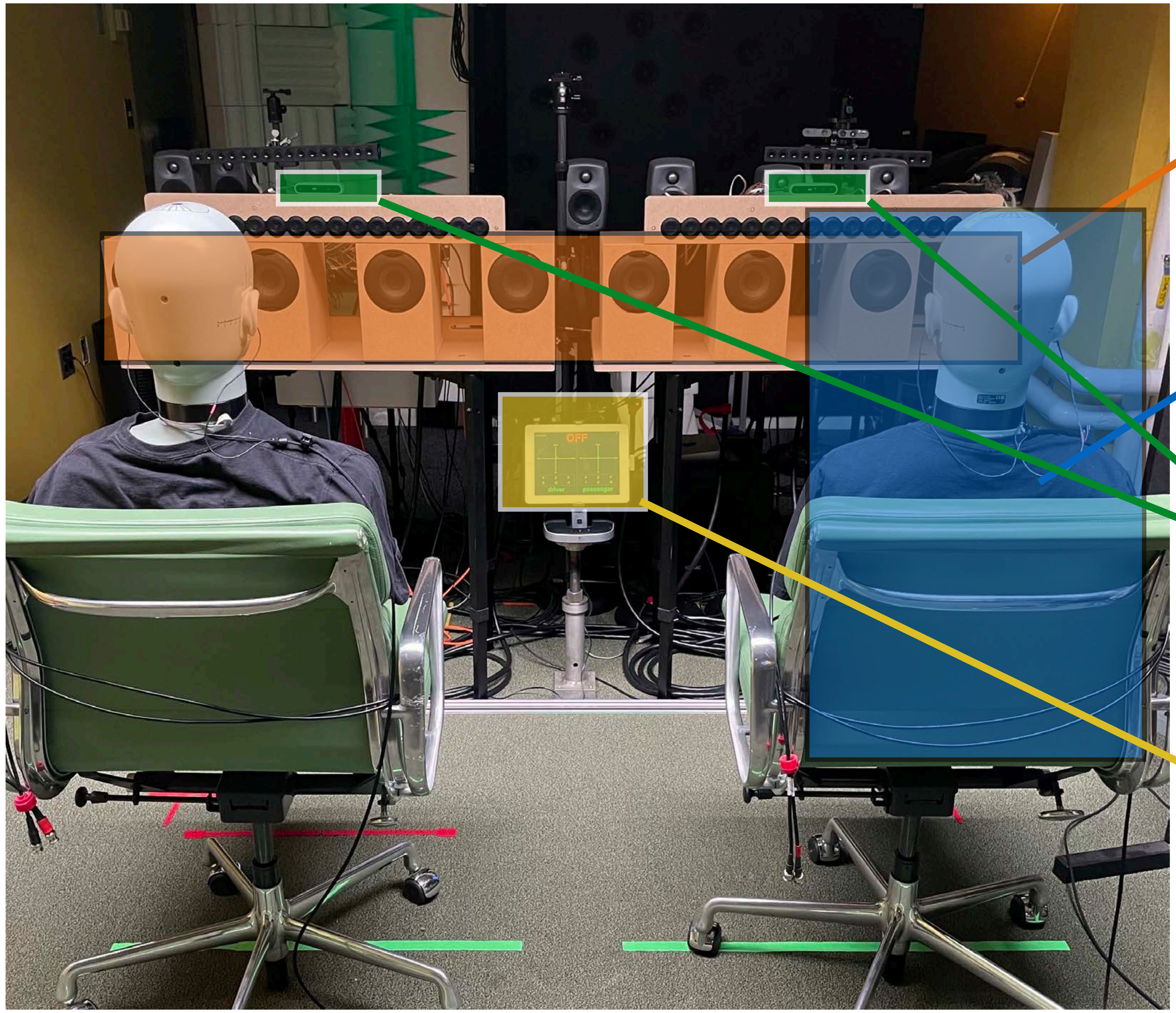
$$\boxed{\mathbf{g}^*_{prob} = (\hat{\mathbf{H}}^H \hat{\mathbf{H}} + \Sigma_{m=1}^{M} \Sigma_m)^{-1} \hat{\mathbf{H}}^H \mathbf{p}_T}$$

Variance matrix Σ_m: empirically derived

Only account for head misalignments

[5] Møller et al., ITASLP, 2019

PSZ system for evaluation

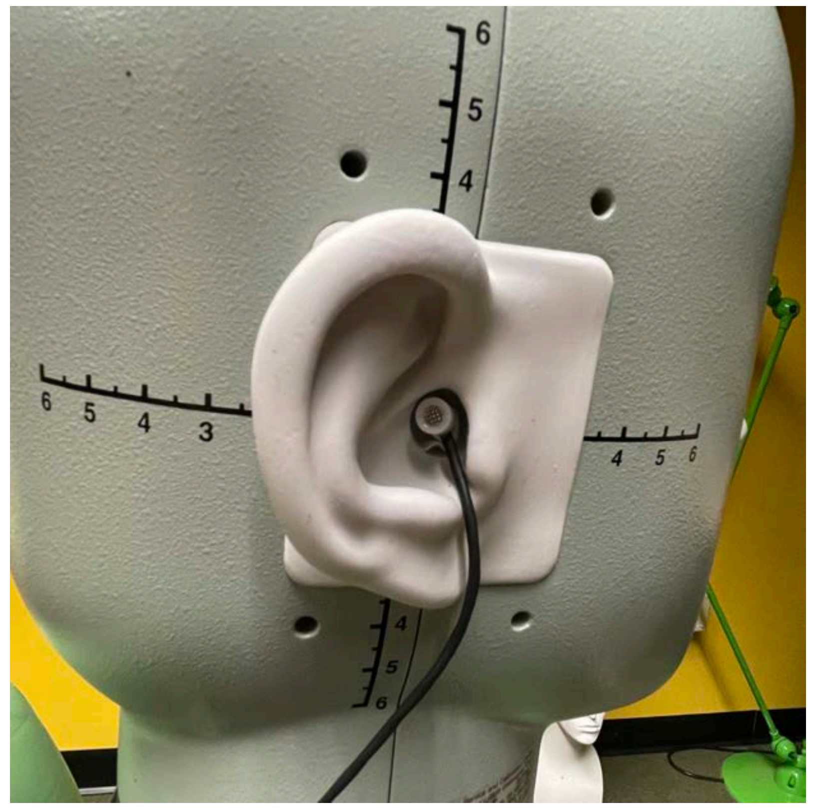


Loudspeaker Array
(200~7000Hz)

B&K HATS
dummy head

Head Tracker
(Infrared depth sensor)

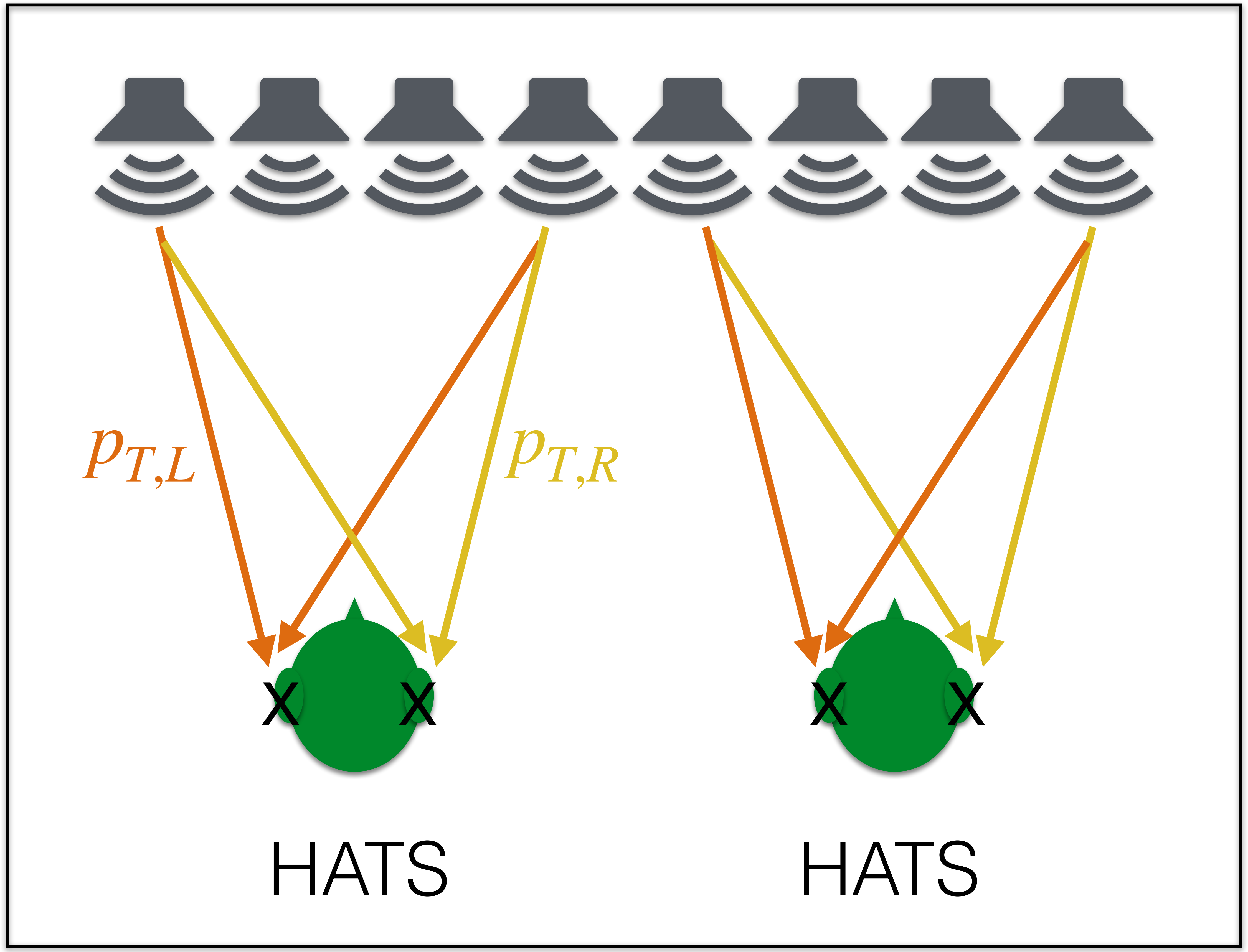
Head Position
Display



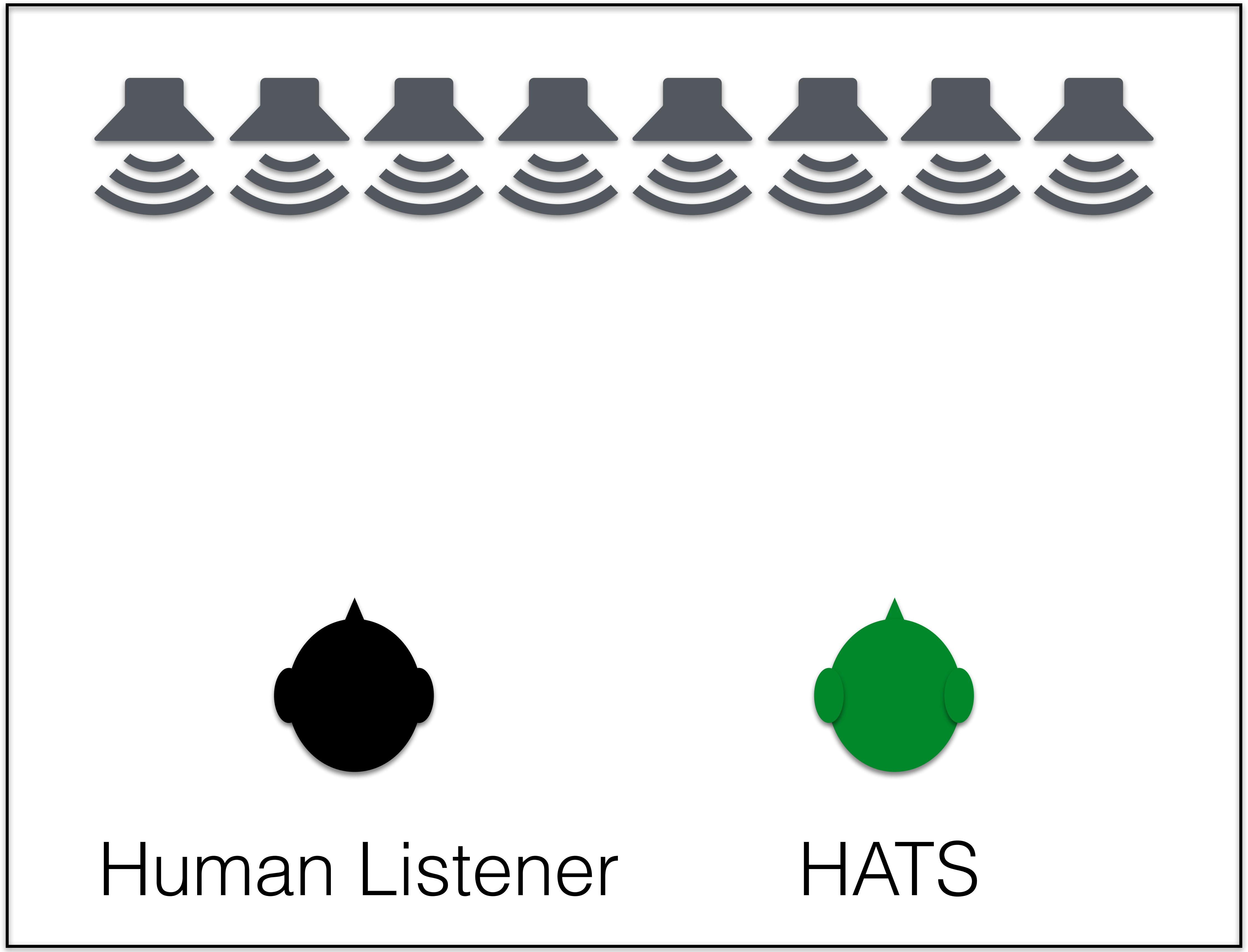
Binaural
Microphones

Allowable head
movement range:
1 cm/10 deg

2 Listener Setups

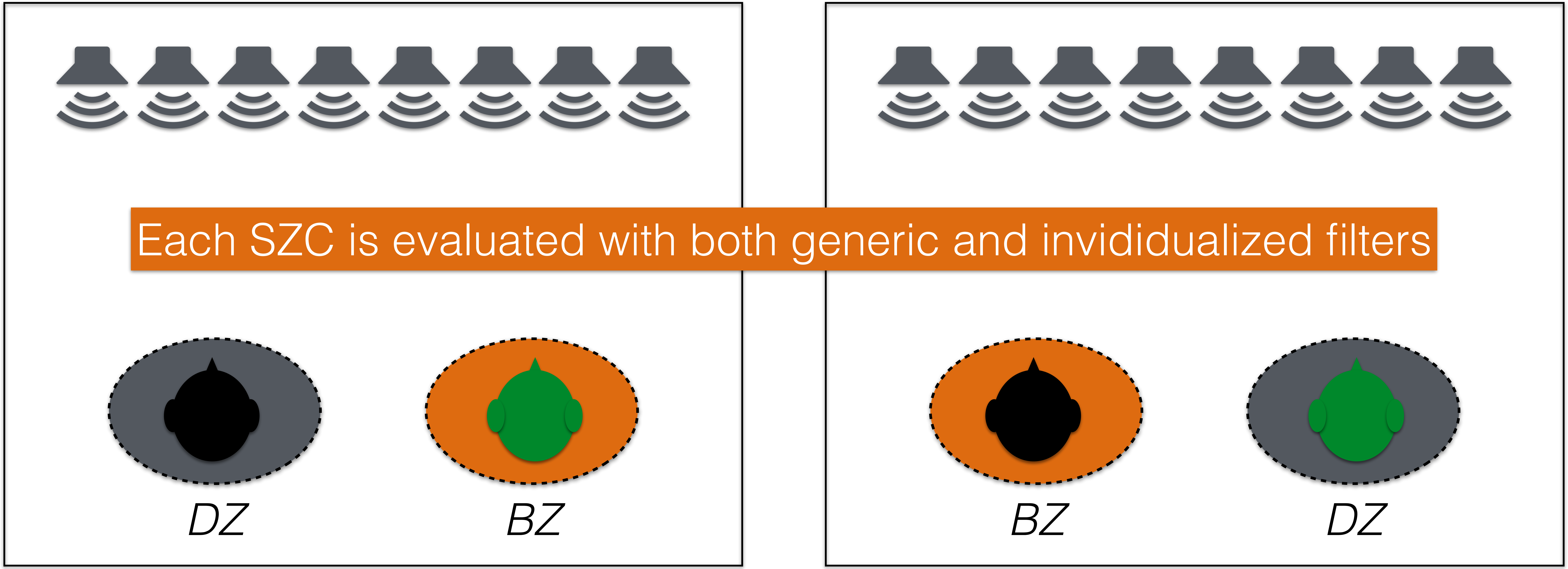


Generic Setup



Individualized Setup

2 Sound Zone Configurations (SZCs)



SZC 1

SZC 2

2 Types of Measurements

- BRTF Measurements
  - Prior to filter evaluation
  - Used for deriving TF variance & generating PSZ filters
- System TF Measurements (BRTFs*PSZ Filters)
  - After filter generated & loaded
  - *In situ*: minimal head movements
  - *Ex situ*: additional head misalignments
- Exponential sine sweeps used for TF measurements

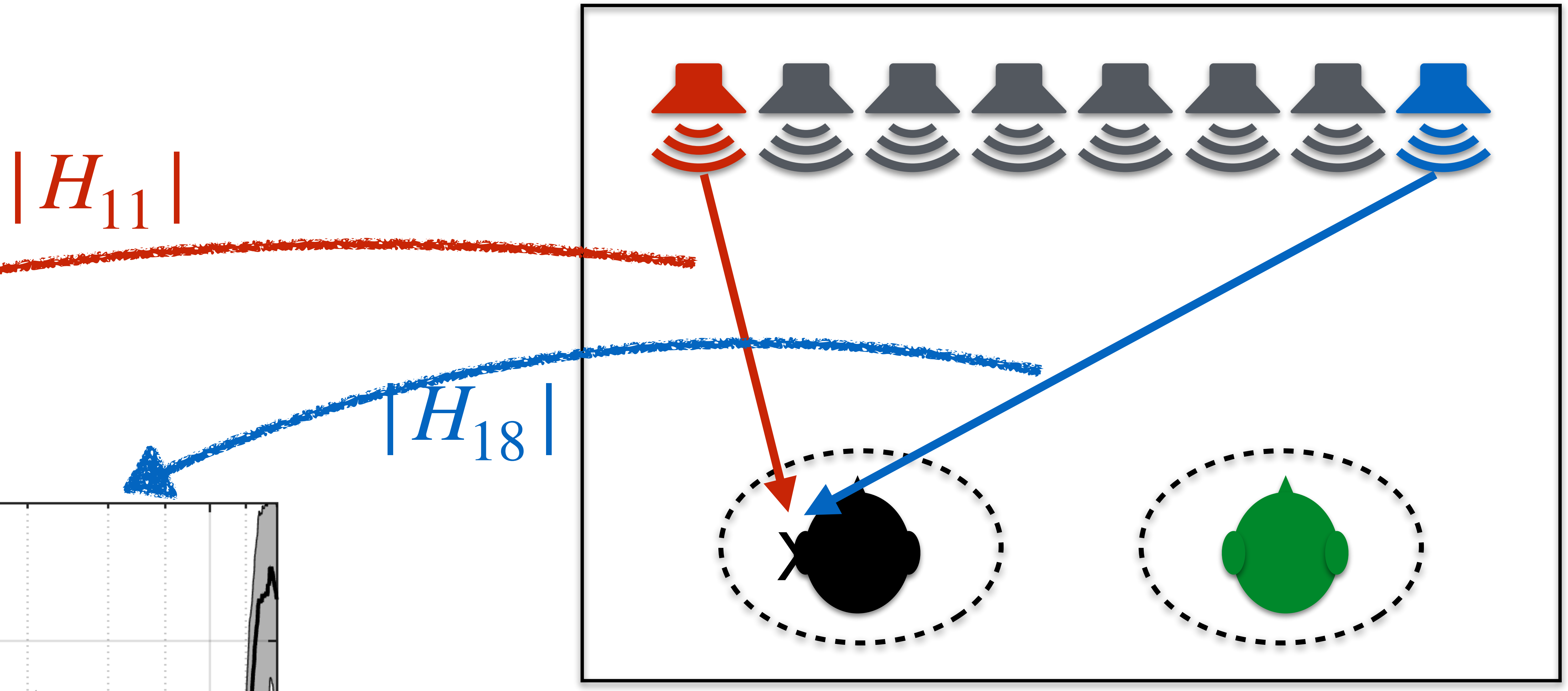
Performance Evaluation Metrics

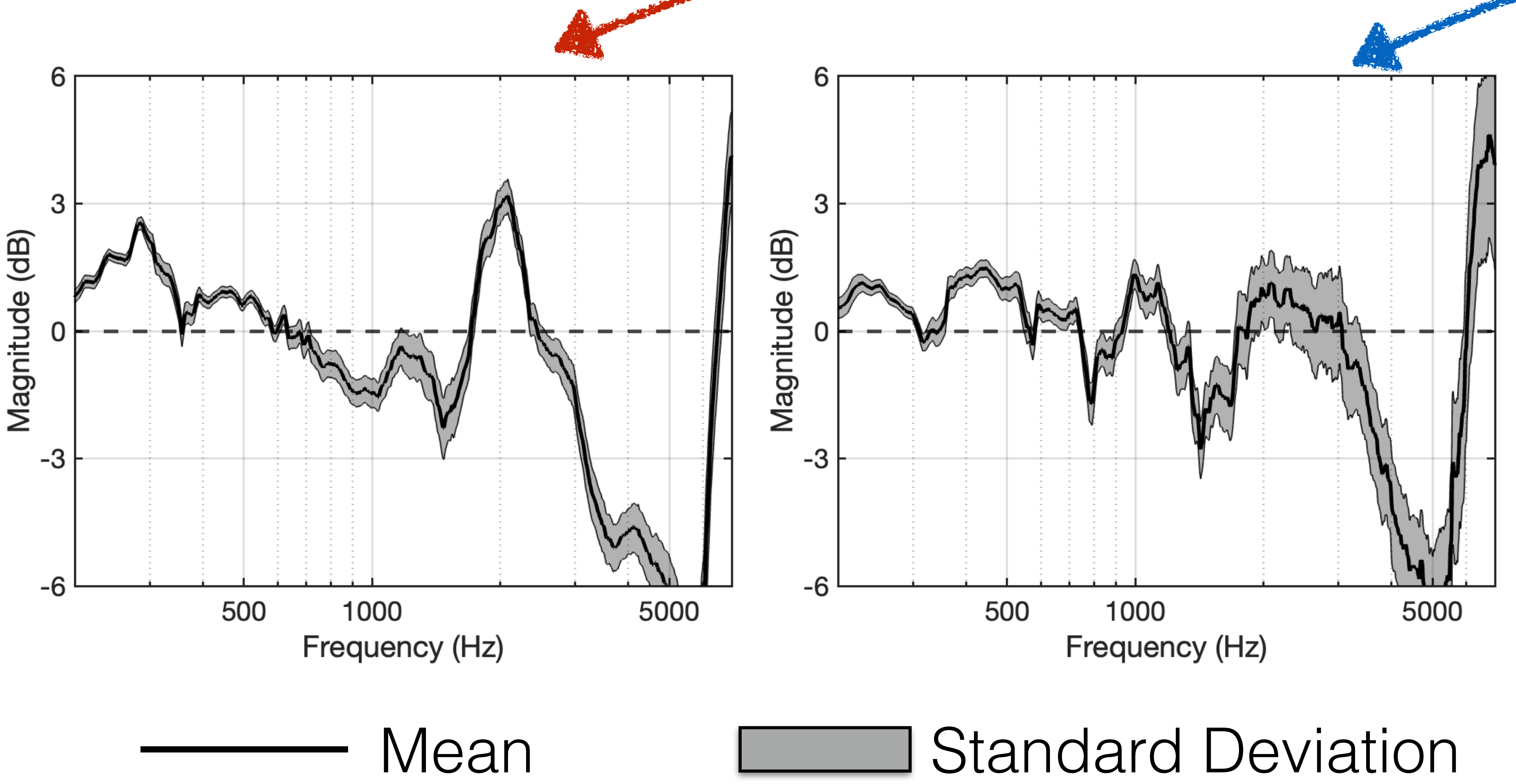
- Acoustic Contrast (AC)[6]
  - Quantifies the isolation level between two PSZs
  - $$AC = \frac{\mathbf{g}^H \mathbf{H}_B^H \mathbf{H}_B \mathbf{g}}{\mathbf{g}^H \mathbf{H}_D^H \mathbf{H}_D \mathbf{g}}$$
- Robustness against head misalignments
  - Variation in AC level from multiple *ex situ* measurements
- Other metrics (e.g., Array Effort, Reproduction Error) not adopted

[6] Coleman et al., JASA, 2014

Results - BRTF Differences

$$\frac{|\text{Individualized BRTF}|}{|\text{Generic BRTF}|}$$

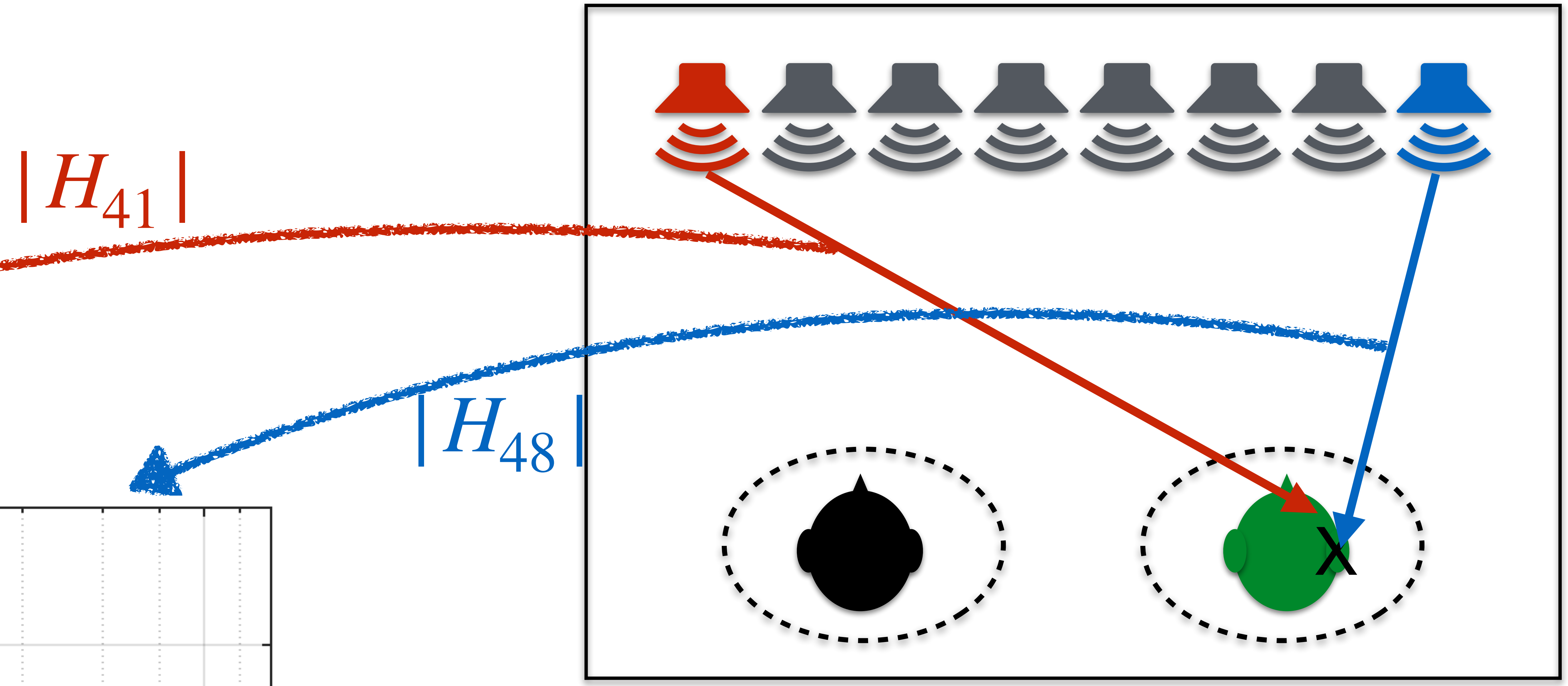


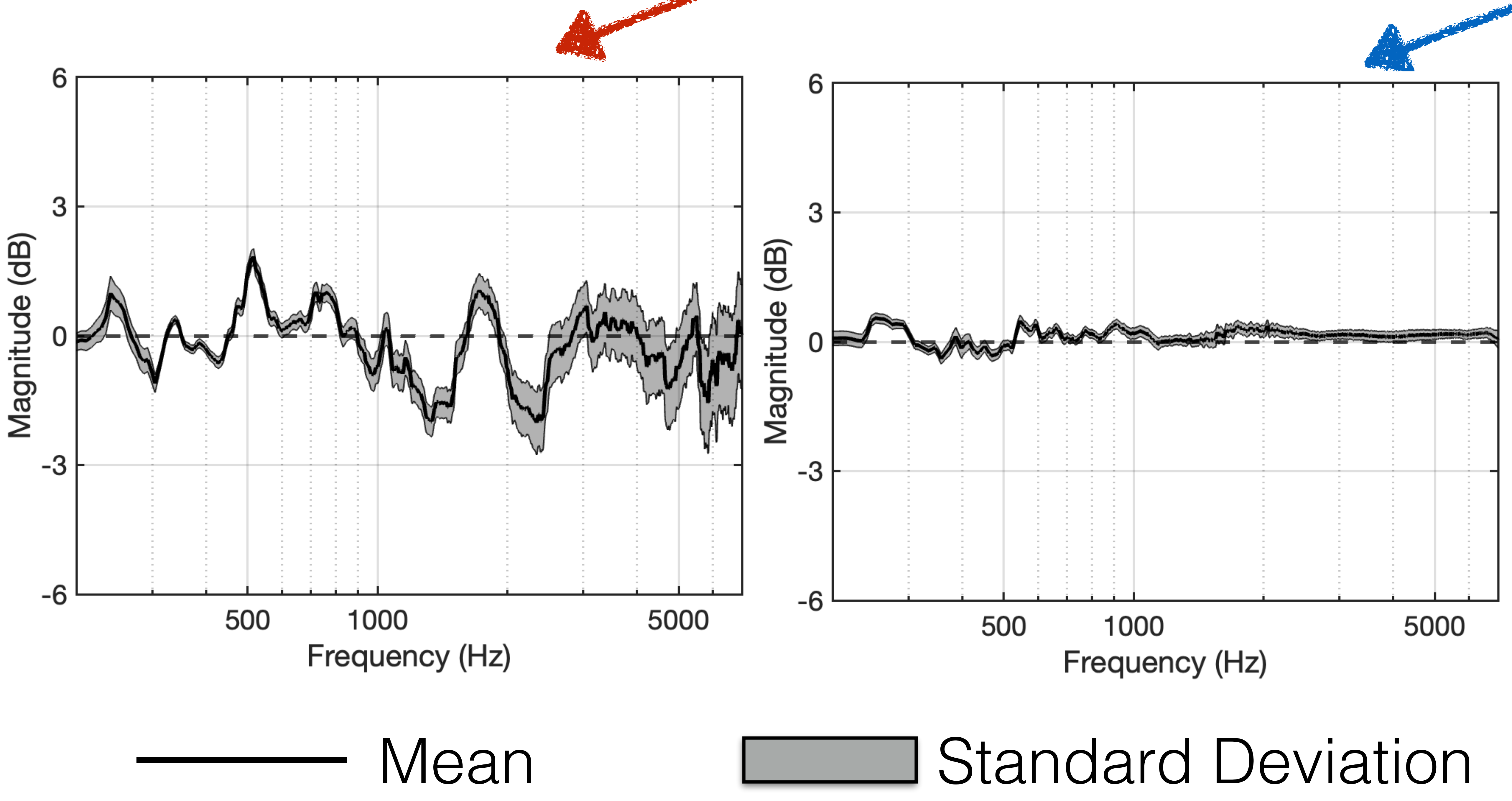


- Difference between the HATS and human listener's BRTFs
- Difference increases with frequency

Results - BRTF Differences

$$\frac{|\text{Individualized BRTF}|}{|\text{Generic BRTF}|}$$

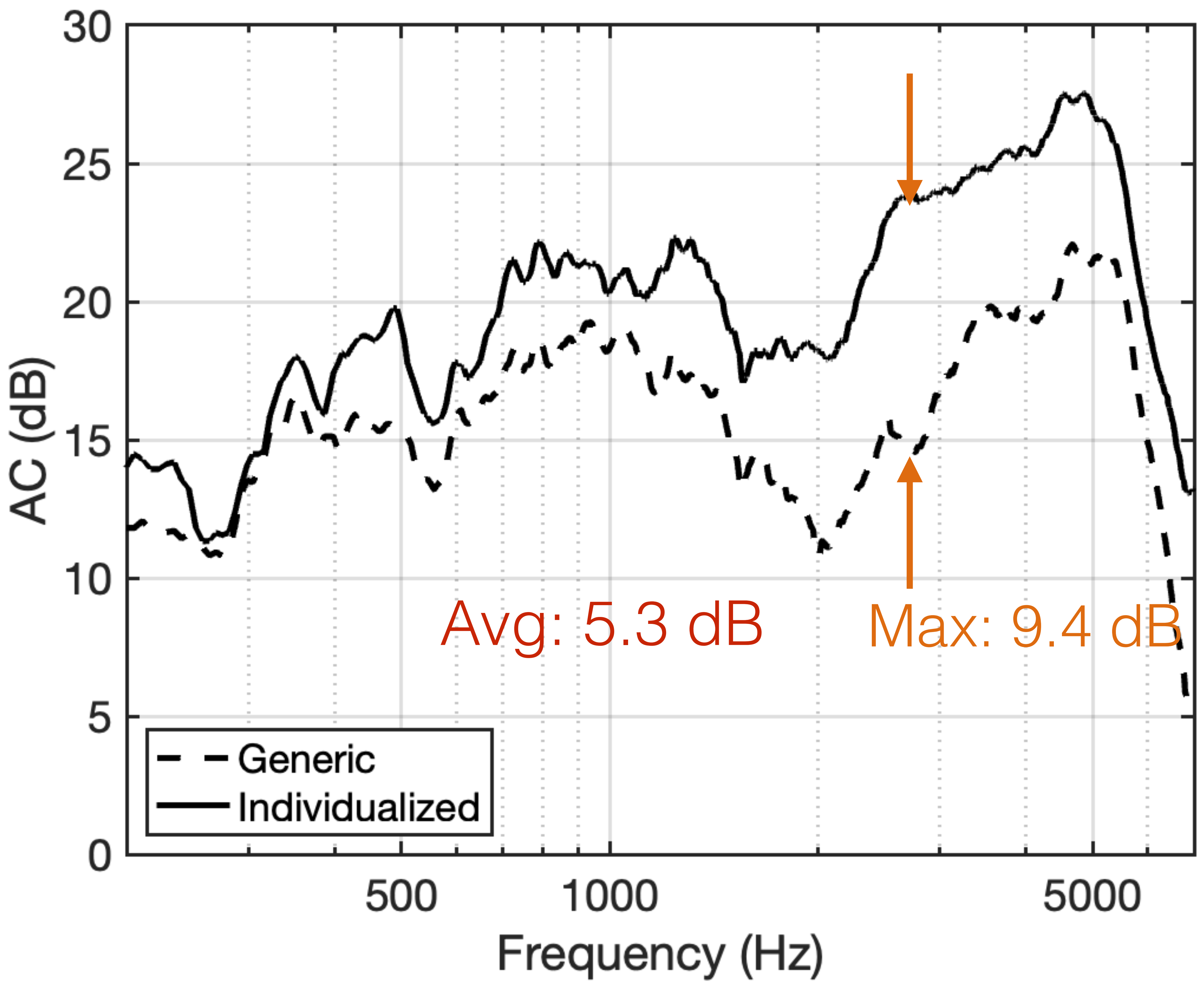


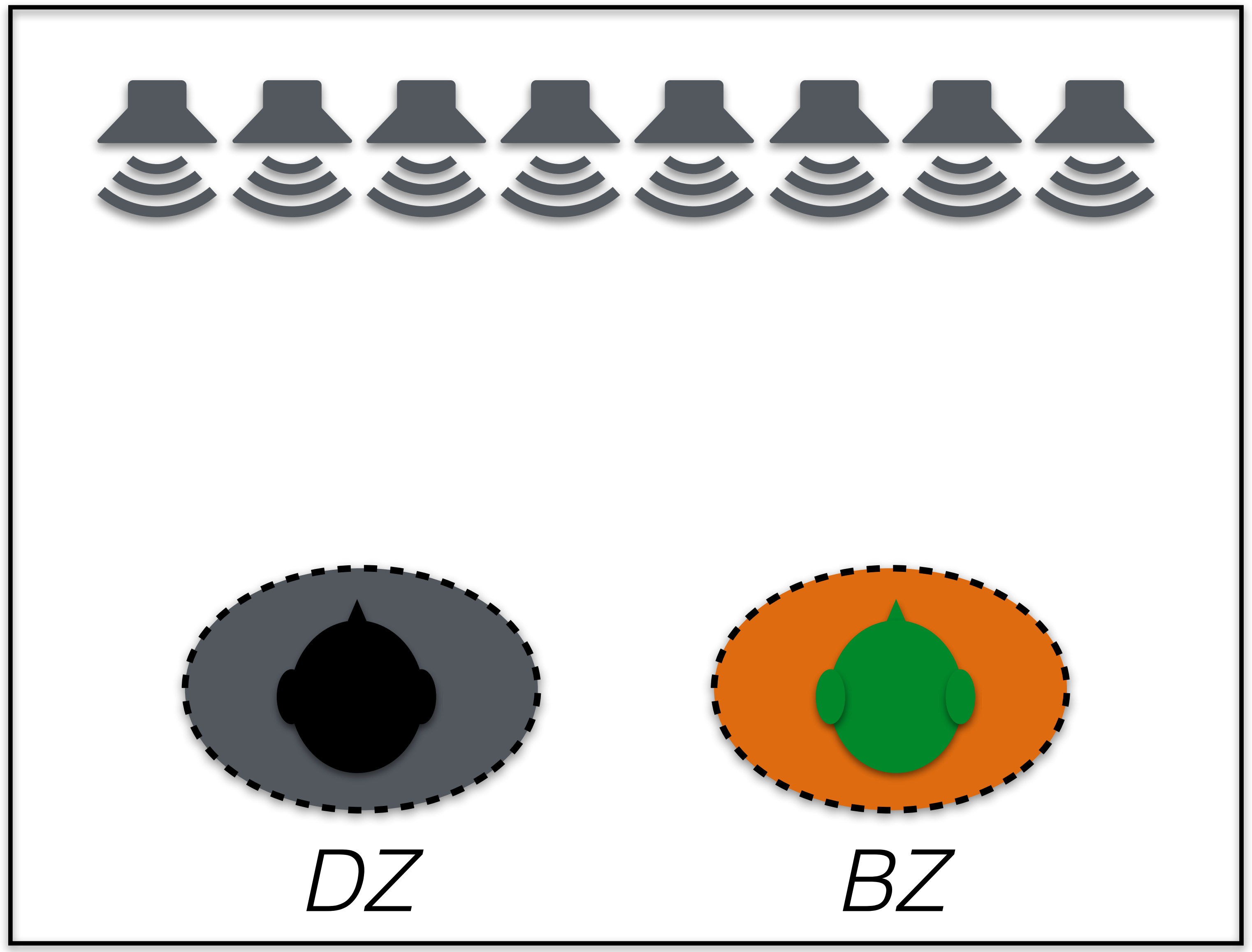


- Difference still exists even though the HATS is fixed
- Potentially due to change in head scattering

PSZ Performance in Best-Case Scenario

AC Spectrum from the *in situ* case



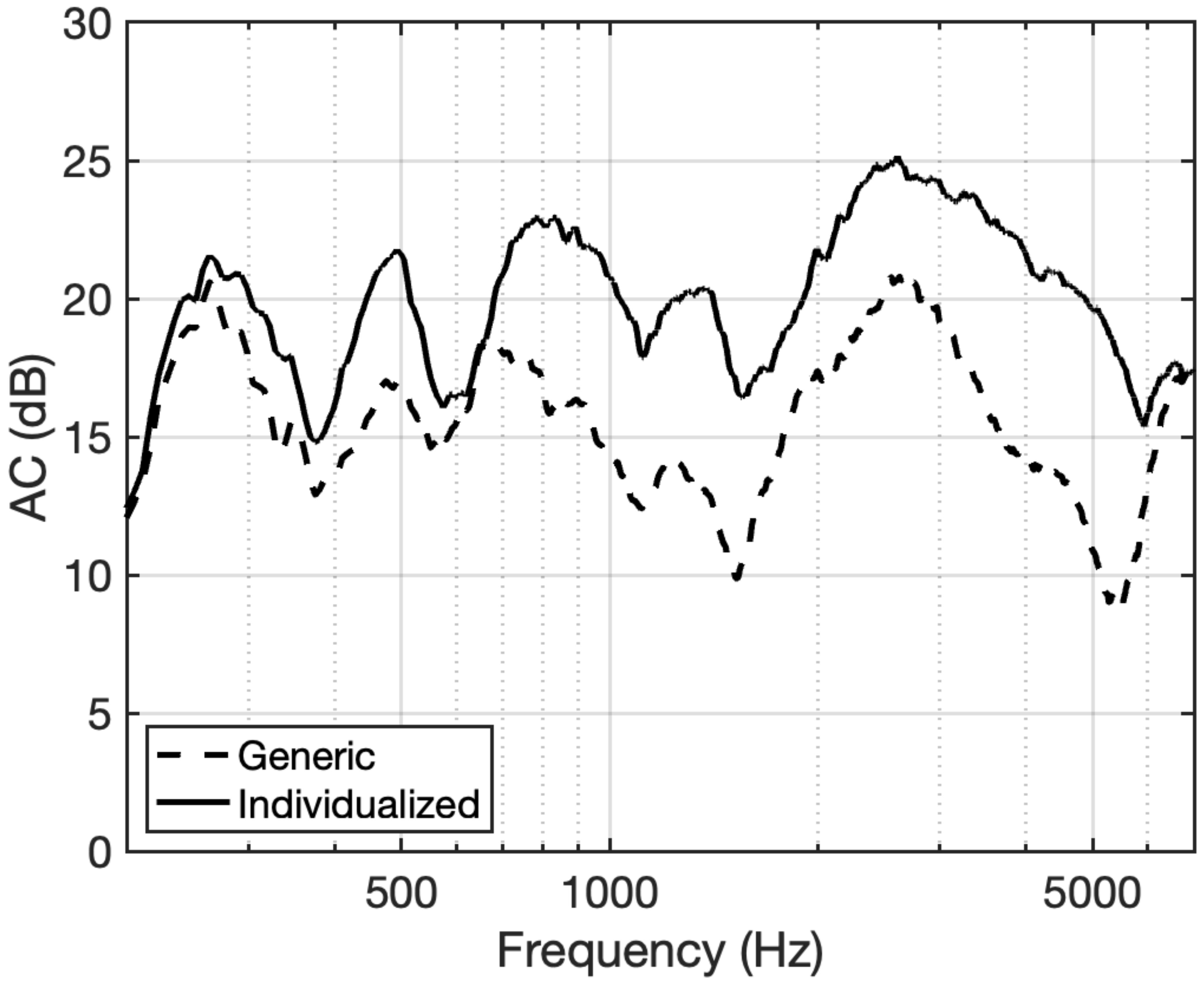


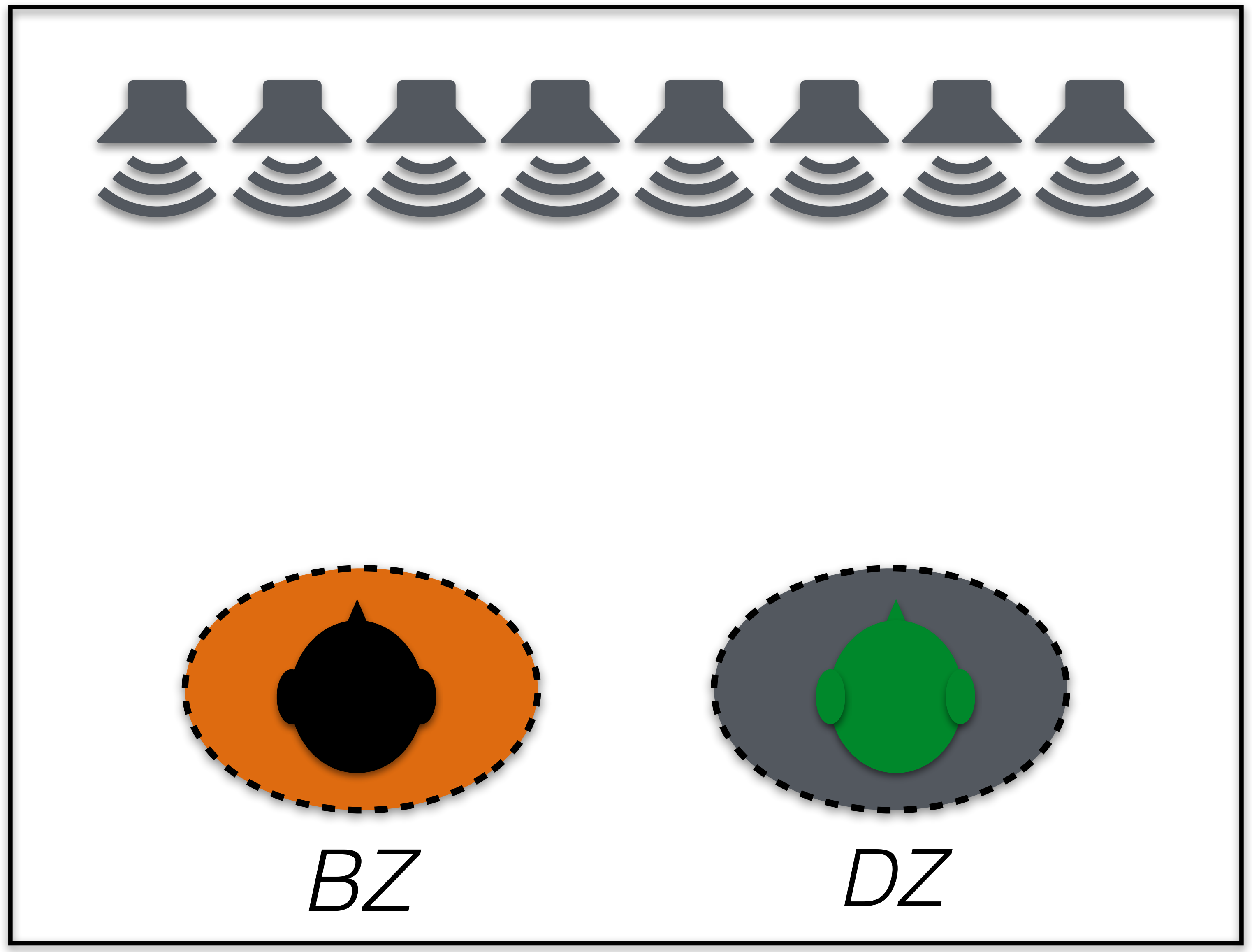
SZC 1

- AC increases with individualized filters
- More noticeable above 2kHz

PSZ Performance in Best-Case Scenario

AC Spectrum from the *in situ* case





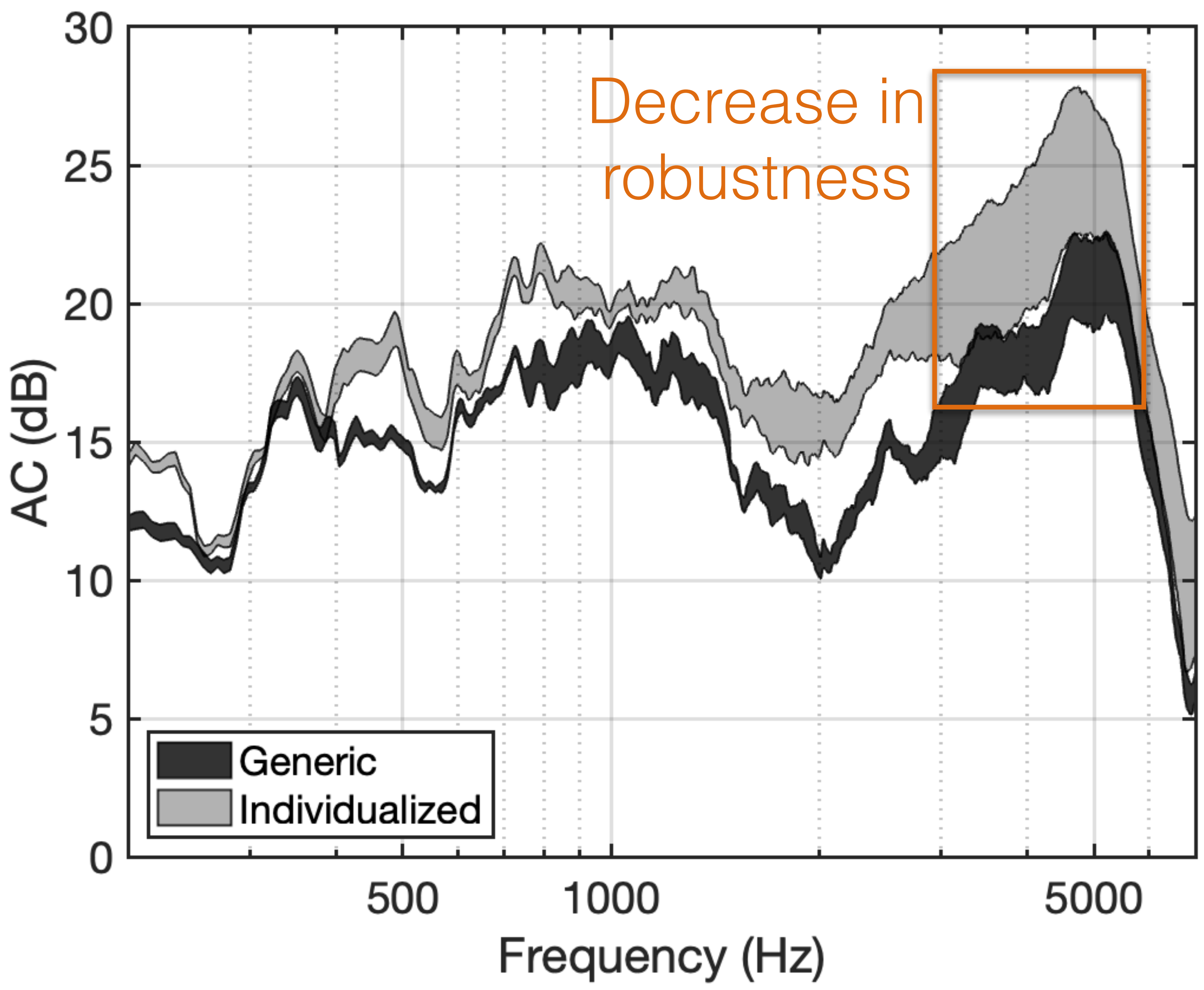
SZC 2

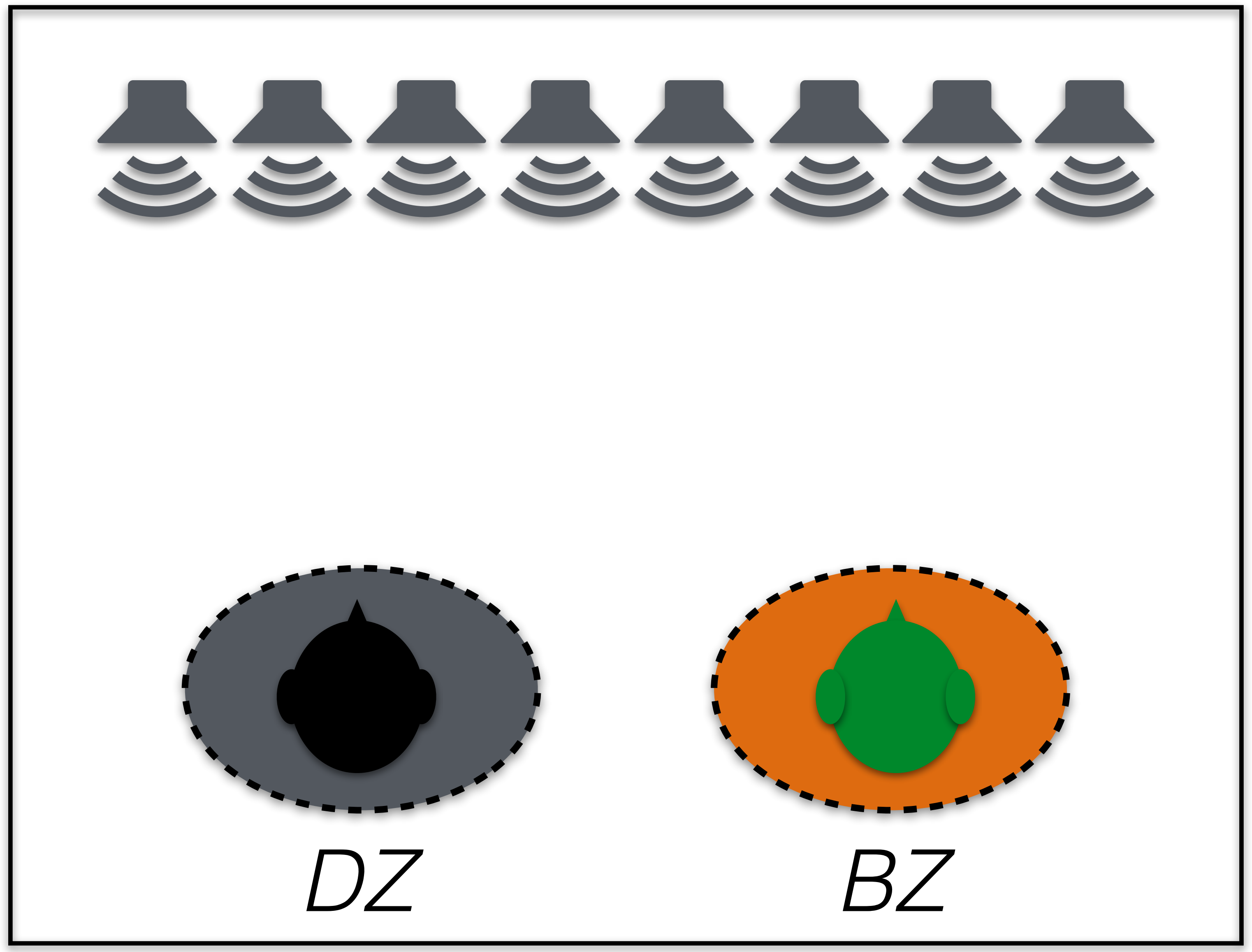
- *DZ* for HATS degraded by replacing the other listener
- Change in one => degradation in both!

Robustness Against Head Misalignments

AC from 20 *ex situ* takes

Robustness ~ width of the curve





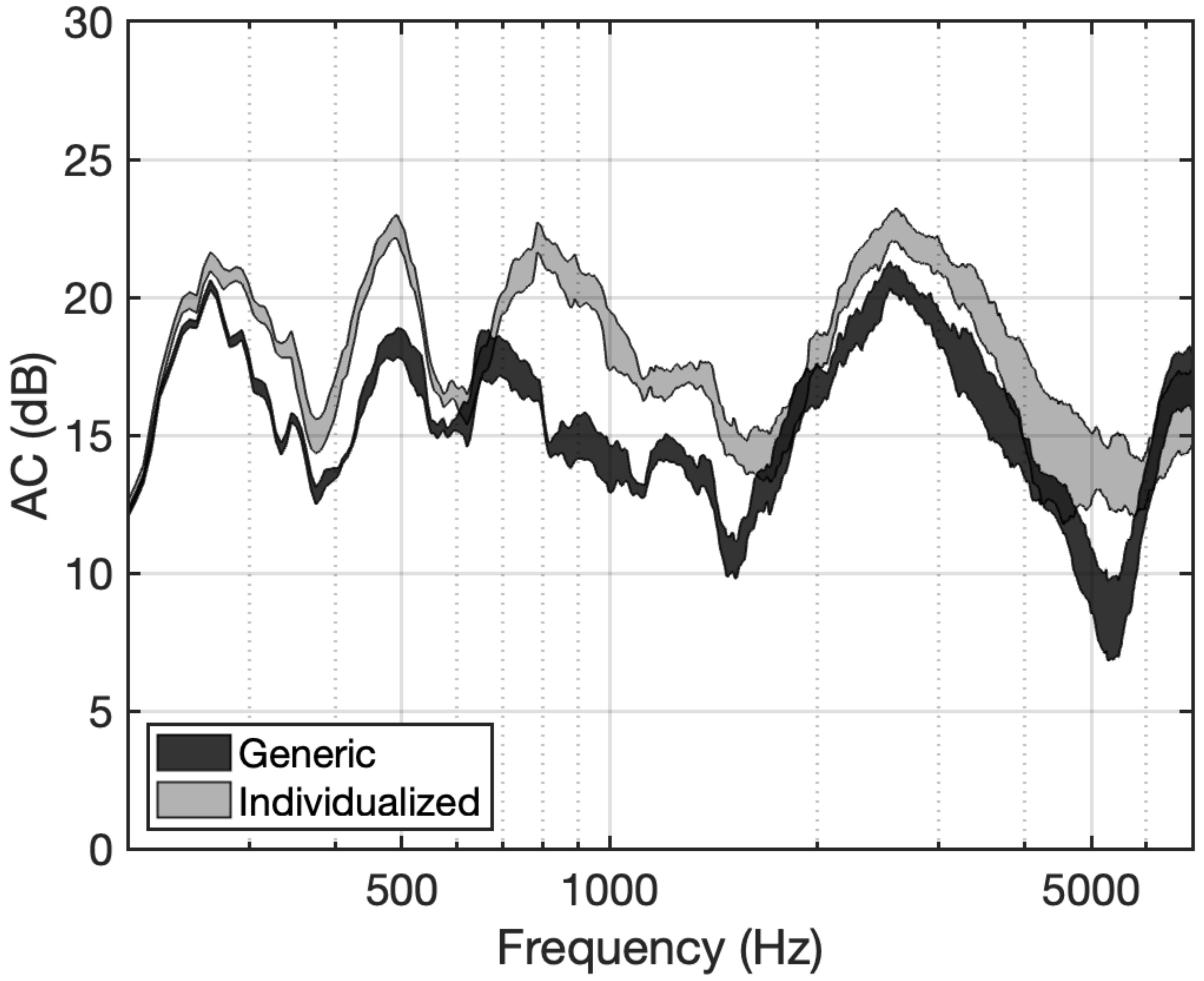
SZC 1

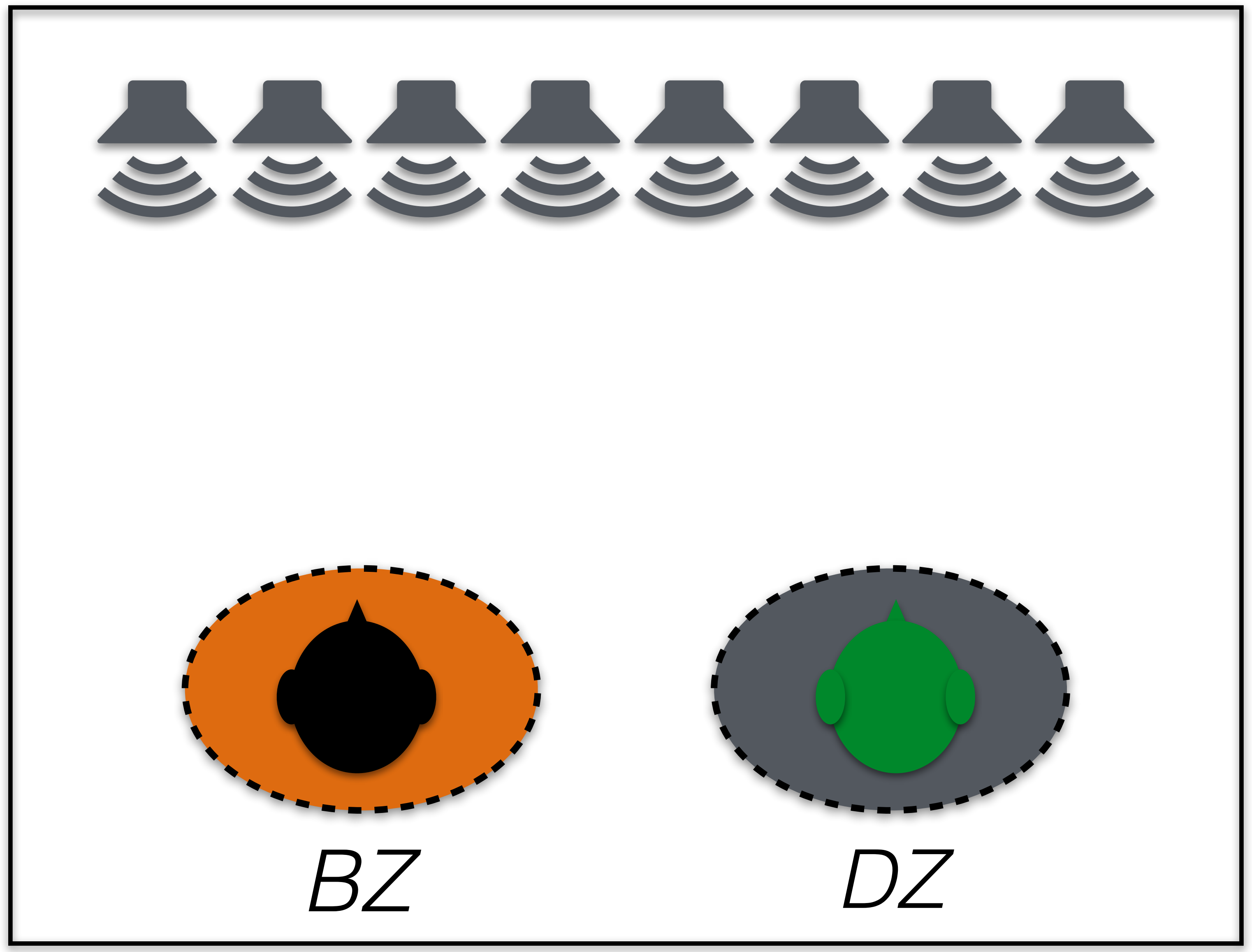
- Robustness decrease for individualized filters above 3kHz
- Individualized is still better
- AC decrease for individualized filters compared to the *in situ* case

Robustness Against Head Misalignments

AC from 20 *ex situ* takes

Robustness ~ width of the curve





SZC 2

- Similar robustness for both filters
- Same AC decrease for individualized filters as in SZC 1

Conclusions

- Does individualization for PSZ make a difference?
  - Yes, it does!
  - AC improvement at all frequencies, both *in situ* and *ex situ*
  - Slight decrease in robustness at high frequencies
- What else?
  - Replacing one listener can affect performance for both listeners
  - Best performance may be hard to retain… (without dynamic reproduction)
- Is the improvement worth the efforts?
  - Not known yet…
  - subjective evaluation required (to be presented in a future work)

The Performance of A Personal Sound Zone System with Generic and Individualized Binaural Room Transfer Functions

Yue Qiao* (presenter) & Edgar Choueiri
3D Audio and Applied Acoustics (3D3A) Lab
Princeton University

Presented at the 152nd AES Convention
May 17, 2022

*E-mail: yqiao@princeton.edu