ICASSP Tutorial T1: Personalising sound over loudspeakers

Dr Philip Jackson, CVSSP, U. Surrey, UK

Dr Filippo Fazi, ISVR, U. Southampton, UK

Dr Phil Coleman, IoSR/CVSSP, U. Surrey, UK

14:00-17:00 Sunday 12 May 2019, Syndicate 1, Brighton Conference Centre, Brighton, UK

Acknowledgements

The following people have made important contributions to the work presented in this tutorial:

Khan Baykaner, Søren Bech, Jon Francombe, Craig Henry, Russell Mason, Martin Møller, Marek Olik, Martin Olsen, Qiaoxi Zhu, Ferdinando Olivieri, Mincheol Shin, Falk-Martin Hoffmann, Eric Hamdan

Introduction to personalising sound over loudspeakers

Why sound?



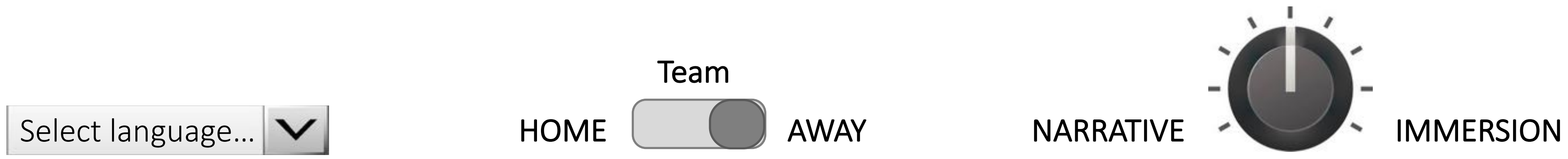
- In evolutionary terms, **sensing** of sound and vibration pre-dates vision
- Hearing provides **connection** to the world around us, through the grapevine, the jungle drum, the market hubbub, the buzz
- It is especially adept at detecting transient events, as with unexpected sounds that turn the head and grab **attention**
- As a truly panoramic sensing modality, we can interpret **immersive** environments


Moore, 2004. An Introduction to the Psychology of Hearing

Blauert., 1997. Spatial Hearing: The Psychophysics of Human Sound Localization

Why personal?

- Increasingly common to multi-task during **shared** experiences
- Even in company, **personal** life continues (alerts, notifications)
- We like **choice** to control our experiences, including selective enhancement, content filters or supplementary feeds
- Personalisation improves autonomy, intimacy, and sense of privacy



- **Object-based audio** (e.g., MPEG-H) offers many personalisation features

Bleidt et al., 2014. Object-based audio: Opportunities for improved listening..., SMPTE
ITU, 2015. ITU-R BS.2076-0: Audio Definition Model
Herre et al., 2015. MPEG-H 3D Audio - The new standard for coding of immersive spatial audio, JSTSP
Coleman et al., 2018. An audio-visual system for object-based audio: from recording to listening, TMM

Why loudspeakers?

- Loudspeakers can be incorporated into the built environment
    - Clean, safe, convenient
- Absence of wearables (c.f. headphones, earphones, headsets)
    - Not weighed down, good comfort and hygiene, free to move, open to alerts
- Creates space for shared experiences
    - Being together in the space, maintains open communication channel
- Provide consistent spatial and timbral quality



ICASSP 2019

What are the practical challenges?

- Sound is invisible!
- Sources spread sound out in all directions
- Air moves and changes temperature, pressure and humidity
- Environments reflect sound all over the place, can be noisy and vary
- Sound diffracts around obstacles and permeates barriers
- Transducers have noise, colouration, non-linear distortion and frequency/power limits
- Listeners are complex, whose perception changes with context

Kinsler et al., 2000. Fundamentals of Acoustics
Moore, 2004. An Introduction to the Psychology of Hearing

Vision for future ears-free personal sound

The promise of sound is in its power:

- To tell a story
- To transport you to another world
- To envelope you in it
- To transcend the mundane, to evoke emotion, passion and encounter empathy
- To connect with your loved ones and your tribe!
  grab our attention

To fulfil this, we need capability:

- To deliver you sound that is reliable, enveloping and tailored to your needs

What will we cover in this tutorial?

- **Formulation**
  - Definitions of personal sound zones and spatial audio
  - Key approaches to manipulate sound fields
  - Measures of performance

- **Engineering**
  - Design parameters
  - Regularization
  - Room effects
  - Practical sound zone filter design

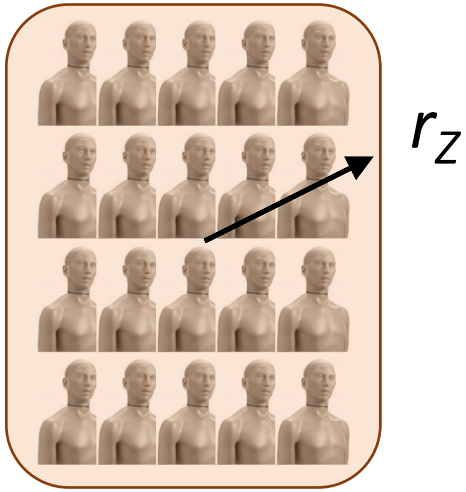
- **Experience**
  - Soundbar listening demonstration
  - Perceptual models
  - Alternative approaches and applications

- Summary of conclusions and perspectives

Definitions: Personal sound zone and spatial audio problems

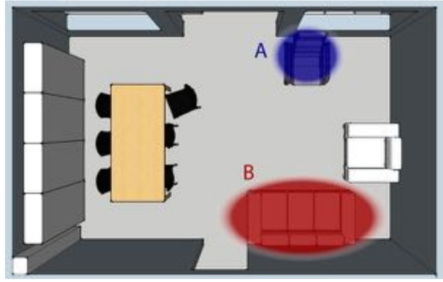
Zone size

Theatre (*auditorium*)



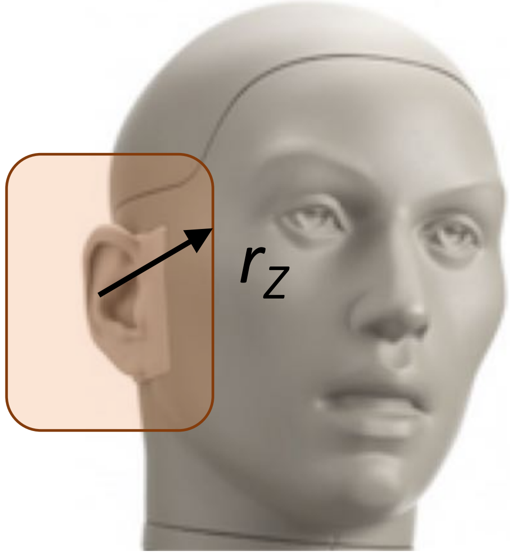
André et al., 2014, IJHCS

Lounge (*head*)



Olik et al., 2013, AES conf

Person (*ear*)



Hollebon et al., 2019, AES conf

Zone count

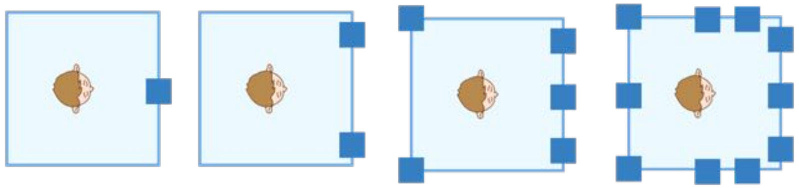
With Z zones, each bright zone has Z-1 dark zones:

Bright	Dark
{A}	
{A}	{B}
{A}	{BC}
	⋮
{A}	{$BCDEF$}

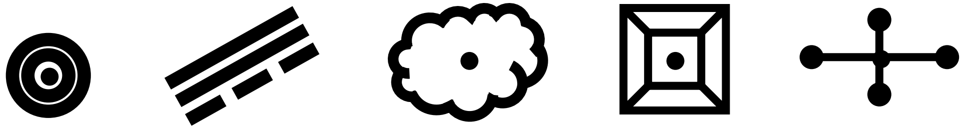
Poletti & Betlehem, 2014, INTER-NOISE

Spatial sound formats

- 1D (distance): mono
- 2D (distance+azimuth): stereo, 5.1 surround
- 3D (distance+azimuth+height): 9.1, 22.2, etc.



- Object-based audio

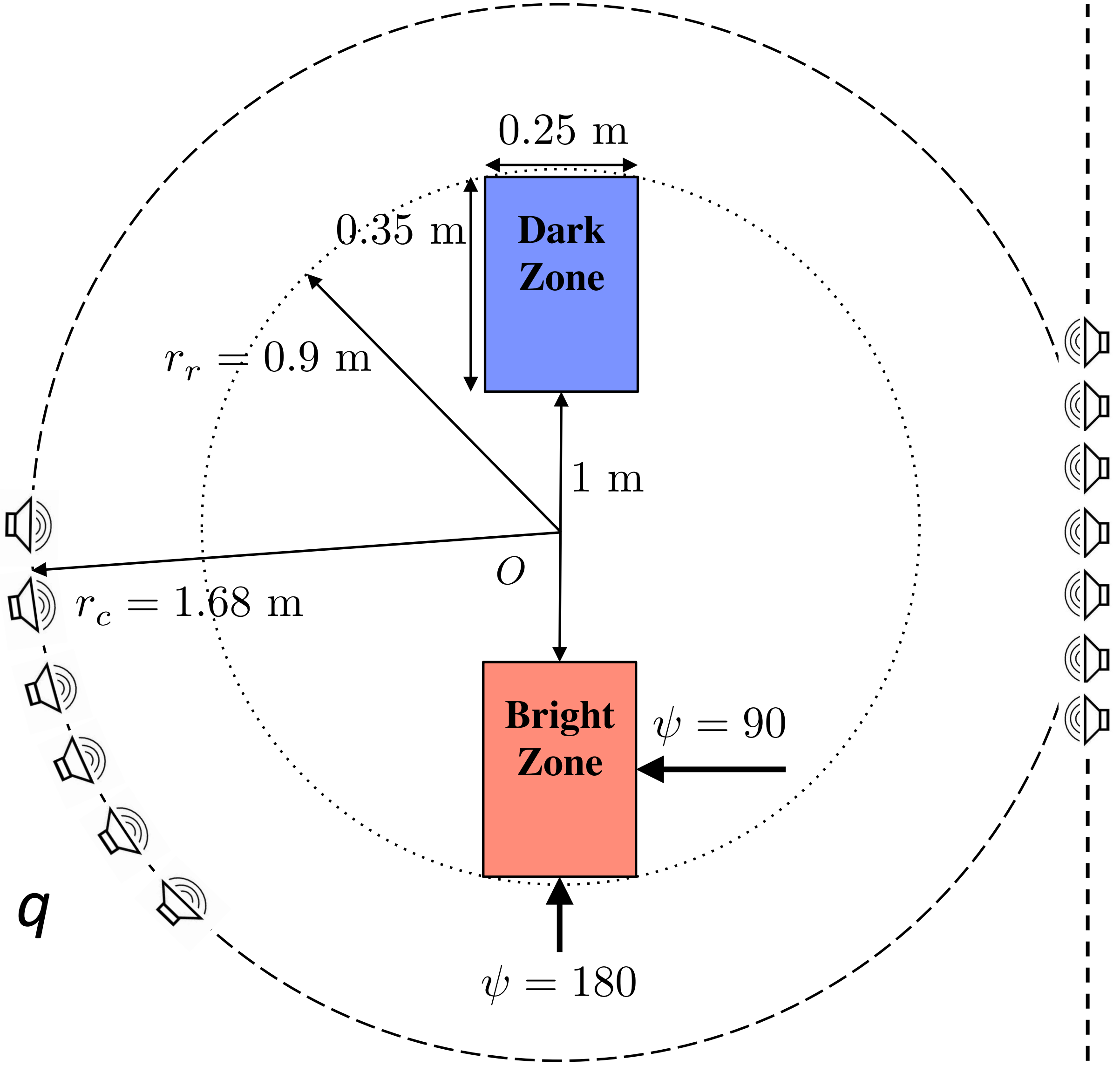


- Real vs. virtual channels

Rumsey, 2001. Spatial audio
Coleman et al., 2014. Stereophonic personal audio reproduction using planarity control optimization, ICSV
Coleman & Jackson, 2016. Planarity-based sound field optimization for multi-listener spatial audio, AES
ITU, 2015. ITU-R BS.2076-0: Audio Definition Model
Thresh & Kearney, 2017. A direct comparison of localisation performance when using first..., AES

Reproduction system

- Number of channels, L
- Loudspeaker arrangement:
  - Uniform line array
  - Uniform circular array
  - Arbitrary positions
- Assuming:
  - Calibrated gain
  - Flat full-range frequency response
  - Perfect synchronisation
- Ideal (plane or) monopole source, q



Fazi, 2010. Sound field reproduction. PhD thesis
Møller & Olsen, 2011. Sound zones, MSc thesis
Coleman, 2014. Loudspeaker array processing for personal sound zone reproduction. PhD thesis

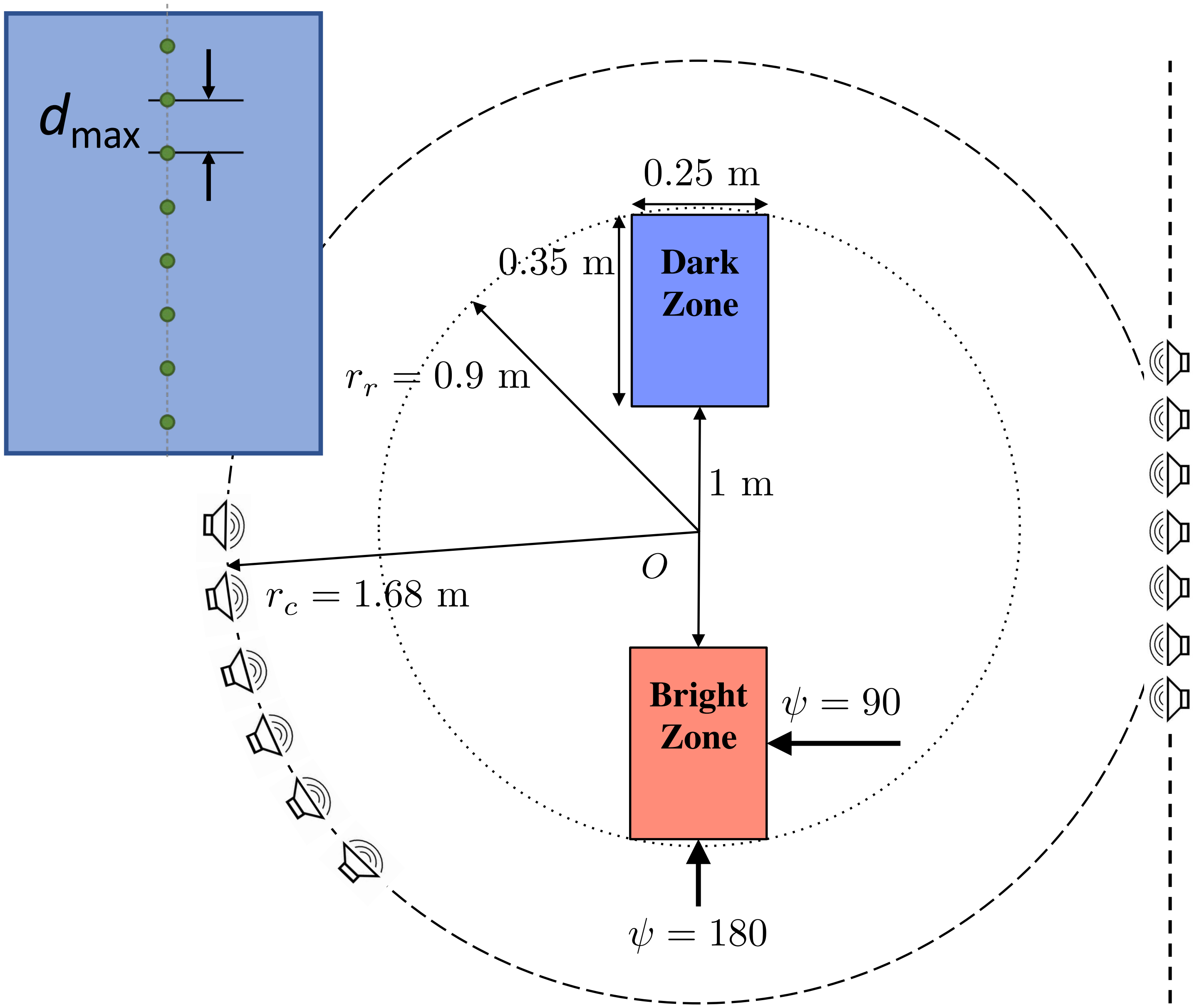
Zone sampling

Nyquist principle imposes maximum spacing at half the wavelength of the highest frequency:

$$d_{\max} = \frac{c}{2f}$$

e.g.,

$d_{\max}$ = 5 cm at 3.4kHz



Williams, 1999, Fourier Acoustics

Essential notation

Sources $\mathbf{q}(f)=[q_1,\ldots,q_l,\ldots,q_L]^T$

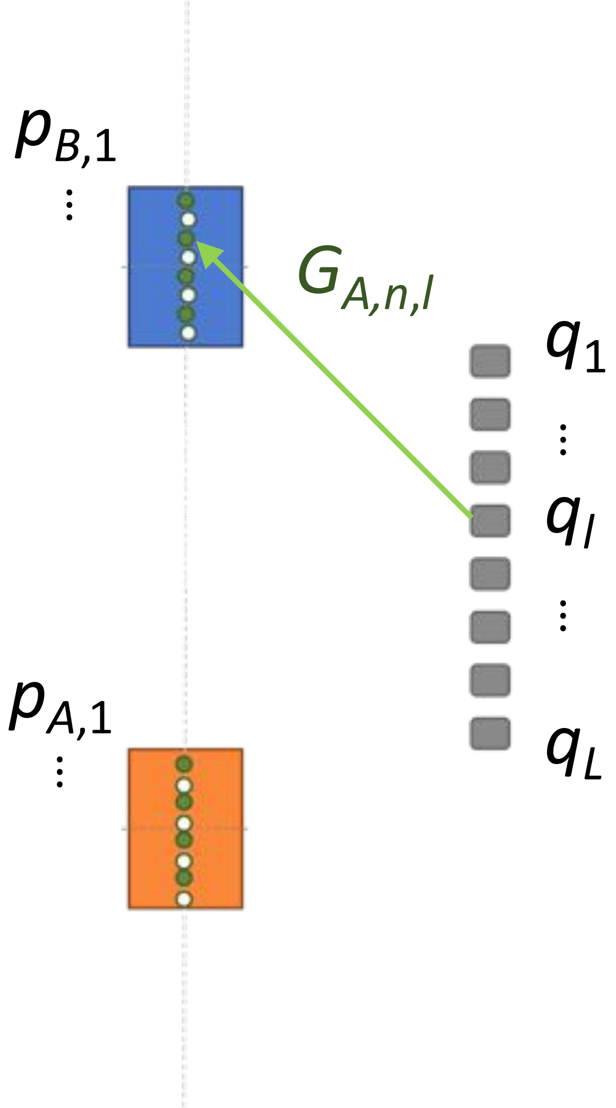
Control mics $\mathbf{p}_A(f)=[p_{A,1},\ldots,p_{A,n},\ldots,p_{A,N}]^T$

Transfer functions

$$\mathbf{G}_A(f) = \begin{bmatrix} G_{A,1,1} & \cdots & G_{A,1,L} \\ \vdots & \ddots & \vdots \\ G_{A,N,1} & \cdots & G_{A,N,L} \end{bmatrix}; \mathbf{G}_B(f) = \begin{bmatrix} G_{B,1,1} & \cdots & G_{B,1,L} \\ \vdots & \ddots & \vdots \\ G_{B,N,1} & \cdots & G_{B,N,L} \end{bmatrix}$$

Relations $\mathbf{p}_A=\mathbf{G}_A\mathbf{q}$

and $\mathbf{p}_B=\mathbf{G}_B\mathbf{q}$ or $\begin{bmatrix}\mathbf{p}_A \\ \mathbf{p}_B\end{bmatrix} = \begin{bmatrix}G_A \\ G_B\end{bmatrix}[\mathbf{q}]$



Coleman, 2014. Loudspeaker array processing for personal sound zone reproduction. PhD thesis

Definitions summary

- Number, size and spacing of zones
- Spatial audio objectives
- Reproduction setup
- Control and monitor microphones
- For monopole sources:

$$G_{n,l} = \frac{j\rho f}{2R} e^{j2\pi f R/c}, \qquad \text{where } R = |\mathbf{r}_{n,l}|$$

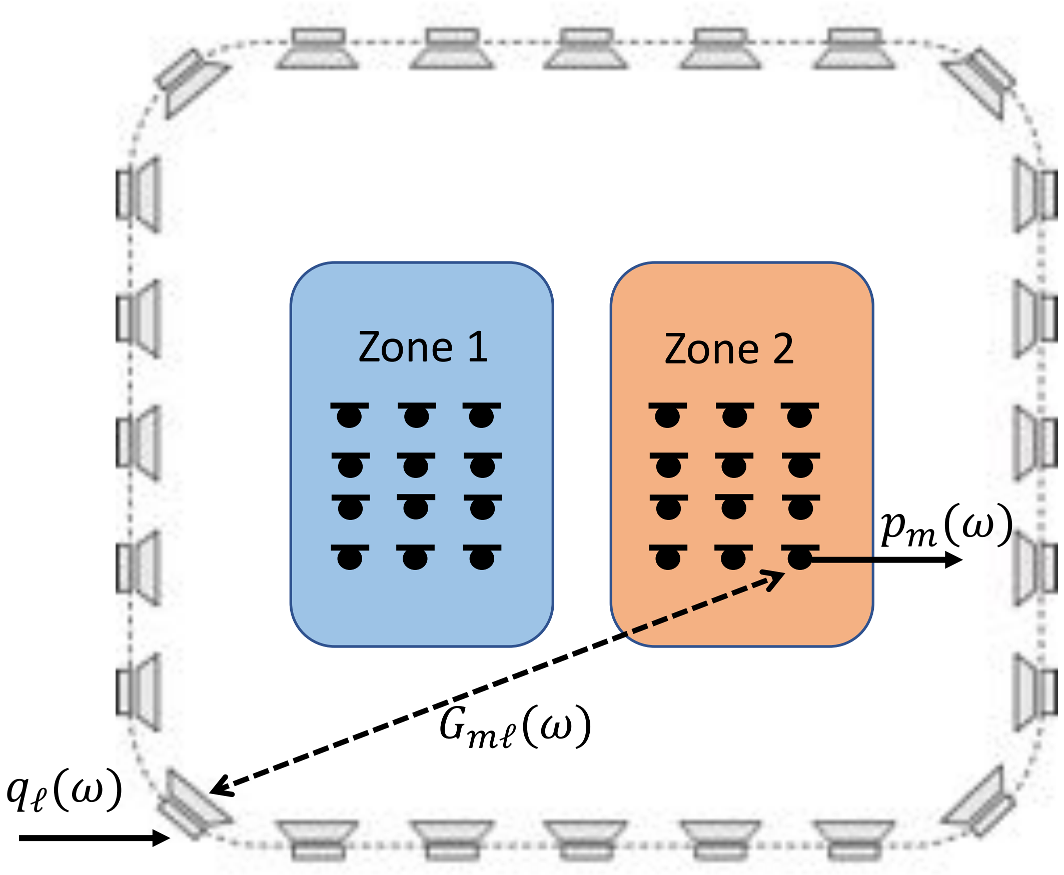
Nelson & Elliott, 1993. Active Control of Sound

Key approaches to filter design for personal sound zones

Notation

- $q_\ell(\omega)$ driving signal of the ℓ–th loudspeaker
- $p_m(\omega)$ signal of the ℓ–th microphone/control point
- $G_{m\ell}(\omega)$ electroacoustical transfer function between the ℓ–th speaker and the m–th control point

$$p_m(\omega) = \sum_{\ell=1}^{L} G_{m\ell}(\omega) q_\ell(\omega)$$



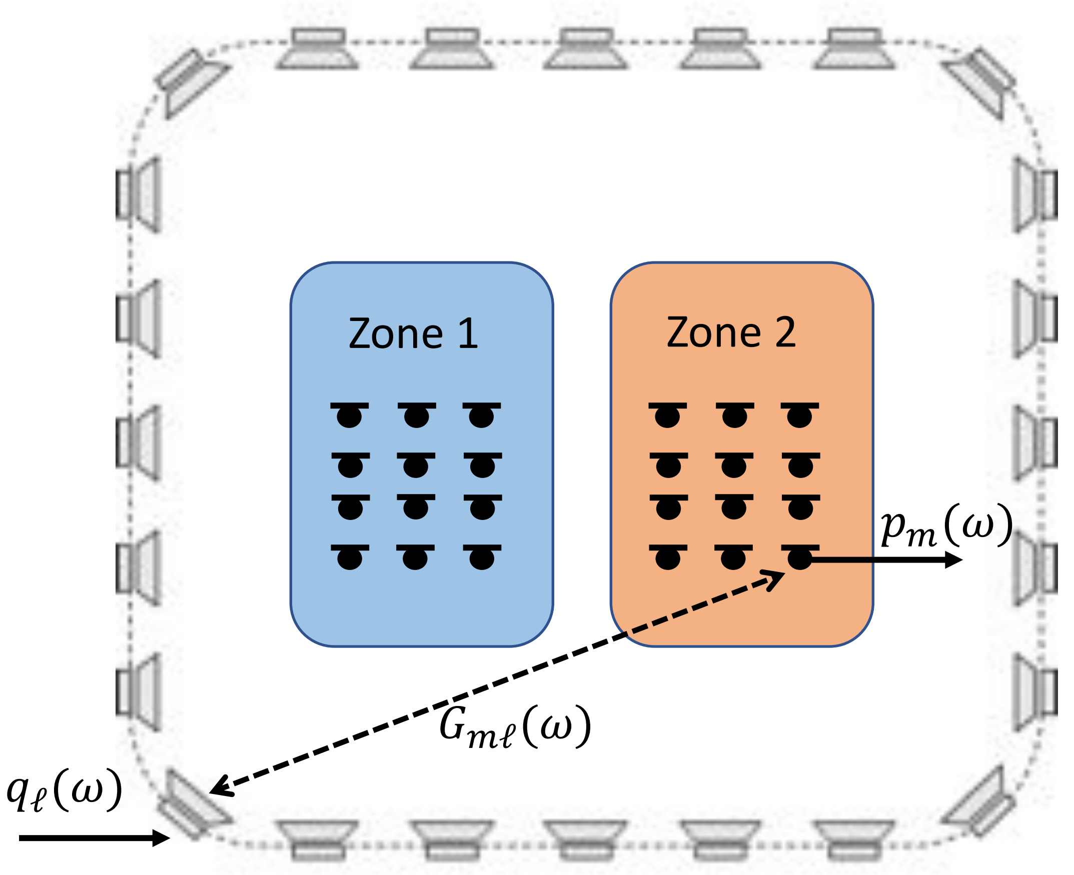
Notation

$$\mathbf{q} = \left[q_1(\omega), q_2(\omega), \ldots q_L(\omega)\right]^T$$

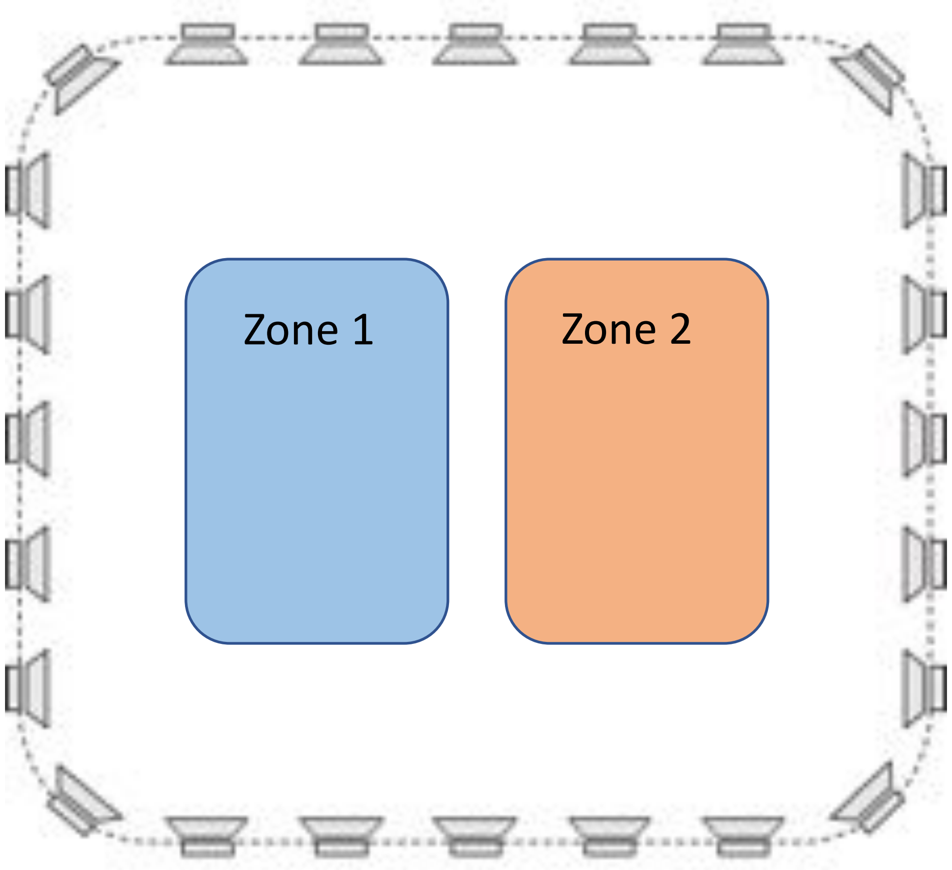
$$\mathbf{p} = \left[p_1(\omega), p_2(\omega), \ldots p_N(\omega)\right]^T$$

$$\mathbf{G} = \begin{bmatrix} G_{1,1}(\omega) & \cdots & G_{1,L}(\omega) \\ \vdots & \ddots & \vdots \\ G_{N,1}(\omega) & \cdots & G_{N,L}(\omega) \end{bmatrix}$$

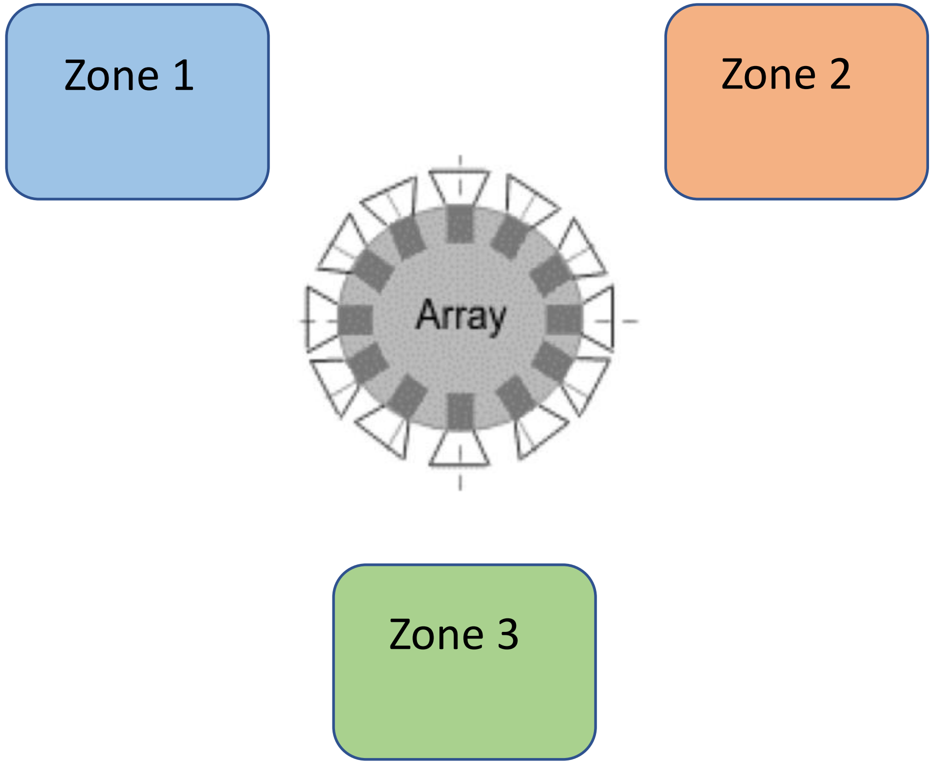
$$\mathbf{p} = \mathbf{G}\,\mathbf{q}$$



Interior problem



Exterior problem



Pressure matching

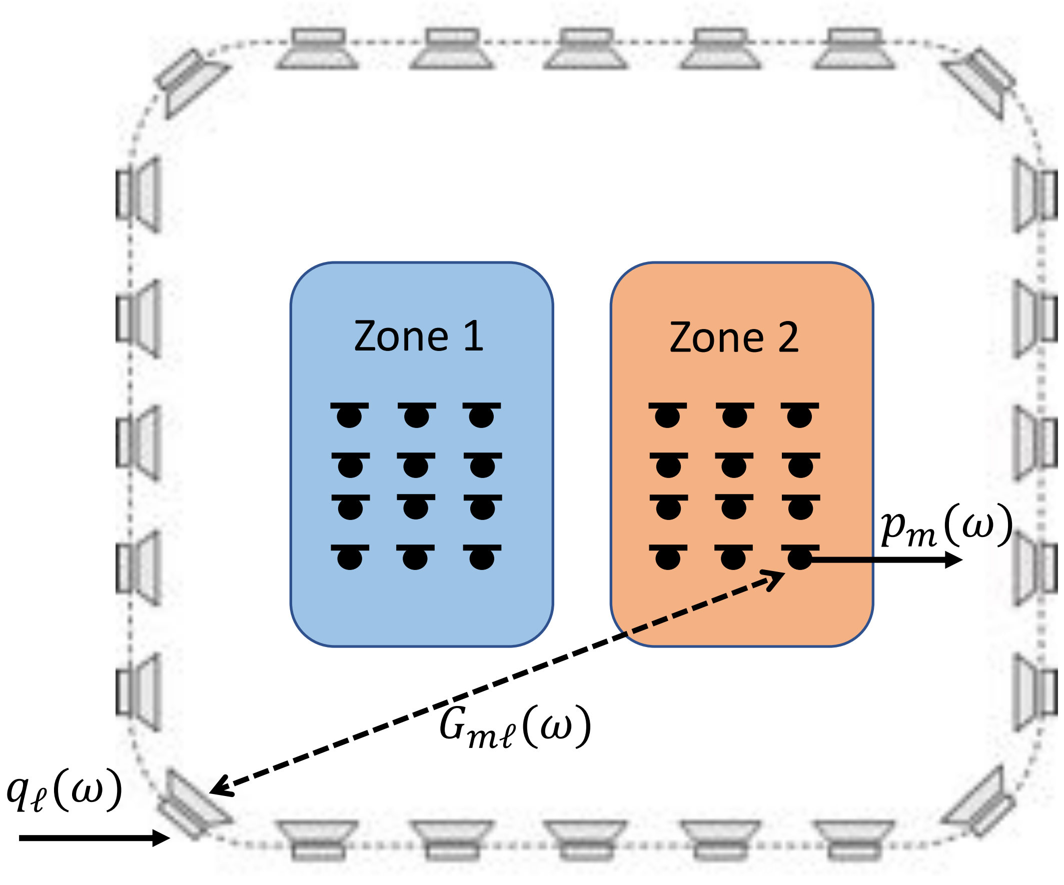
$\mathbf{p}_T$ is the target pressure vector in the bright zone

$$\mathbf{p}_T = [\underbrace{p_1, p_2, \ldots, p_{N_B}}_{\text{Zone 1}}, \underbrace{0, 0, \ldots, 0}_{\text{Zone 2}}]^T$$

$$N > L \rightarrow \text{overdetermined problem}$$

$$J = \|\mathbf{p} - \mathbf{p}_T\|^2 = \|\mathbf{G}\,\mathbf{q} - \mathbf{p}_T\|^2$$

$$\mathbf{q}_{opt} = \mathbf{G}^\dagger \mathbf{p}_T = \left(\mathbf{G}^H \mathbf{G}\right)^{-1} \mathbf{G}^H \mathbf{p}_T$$

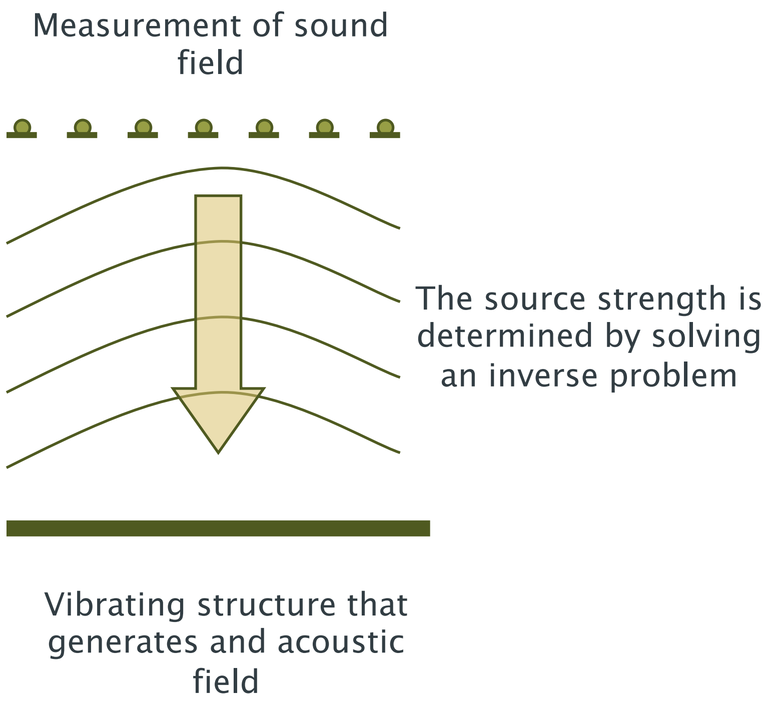


$$\mathbf{p} = \mathbf{G}\,\mathbf{q}$$

Kirkeby and Nelson, 1993. “Reproduction of plane wave sound fields”, JASA

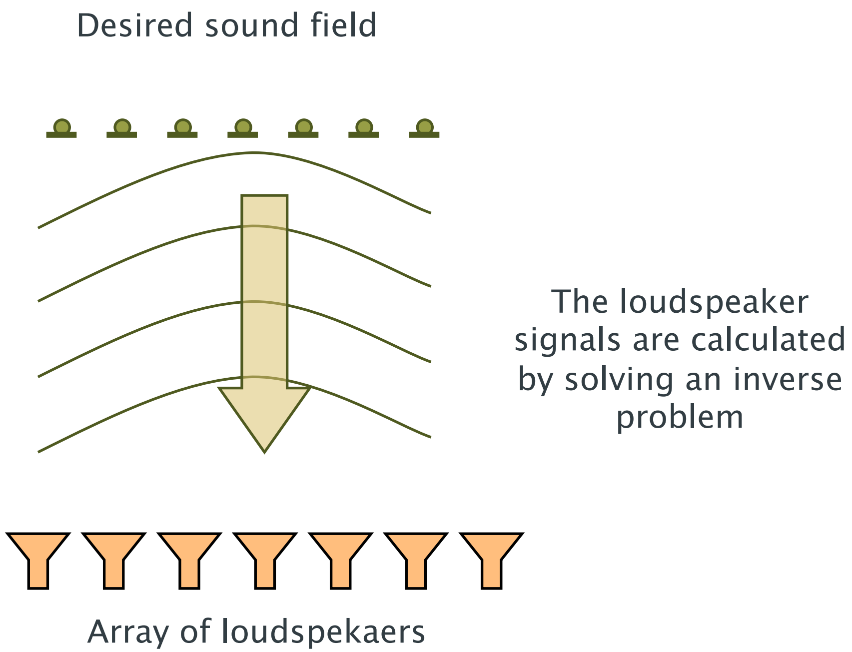
Relation to acoustical holography

ACOUSTICAL HOLOGRAPHY



- The (physical) solution exists

SOUND FIELD CONTROL



- The solution might not exist!

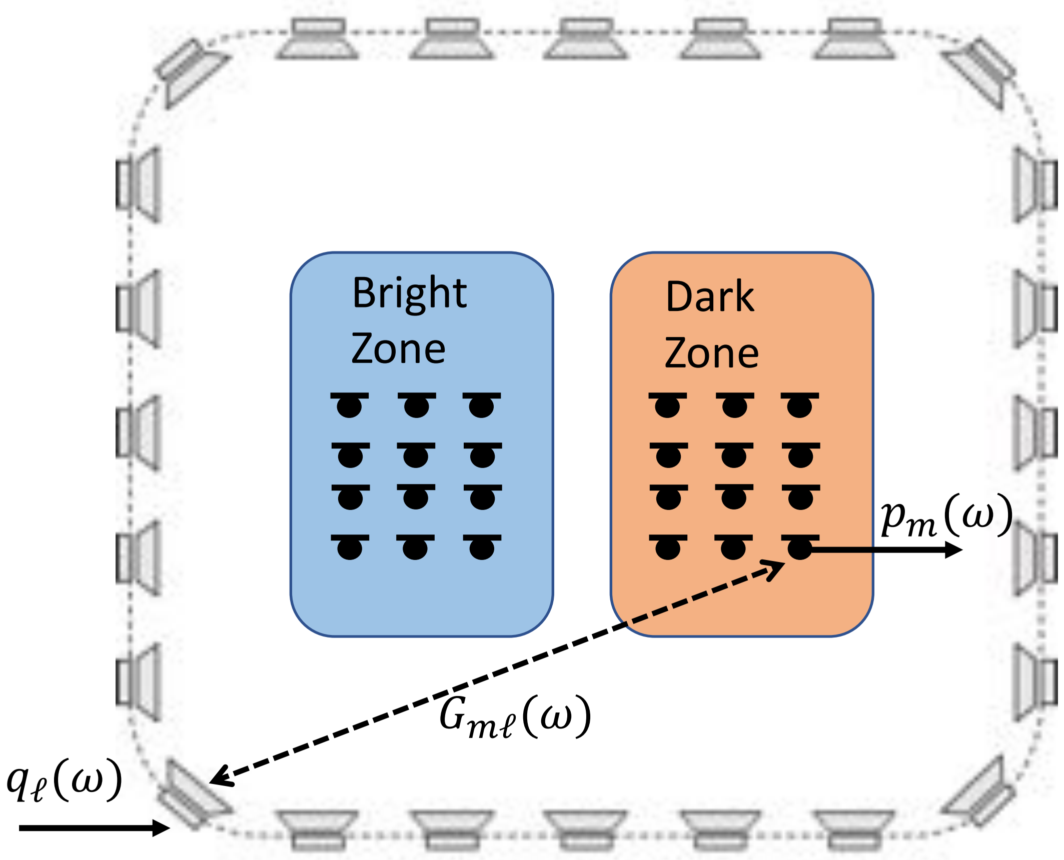
Ill-conditioning

$$\mathbf{q}_{opt} = \mathbf{G}^{\dagger}\mathbf{p}_T \qquad J = \|\mathbf{G}\,\mathbf{q} - \mathbf{p}_T\|^2$$

Matrix **G** can be ill-conditioned, especially at low frequencies (columns of **G** are almost identical).

Array effort $$E(\omega) \propto \|\mathbf{q}\|^2 = \sum_{\ell=1}^{L} |\,q_\ell(\omega)\,|^2$$

Effort my be very large if **G** is ill-conditioned, leading to unstable solutions.



$$\mathbf{p} = \mathbf{G}\,\mathbf{q}$$

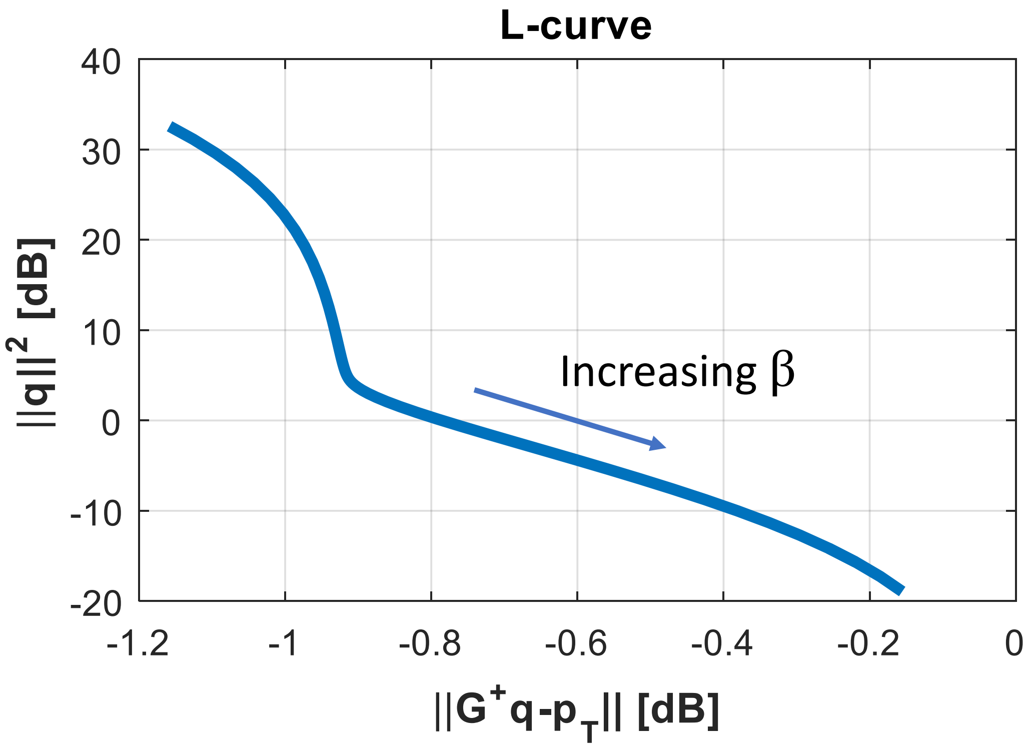
Tikhonov Regularization

Cost function with Tikhonov regularization

$$J = \left\| \mathbf{G}\,\mathbf{q} - \mathbf{p}_T \right\|^2 + \beta \left\| \mathbf{q} \right\|^2$$

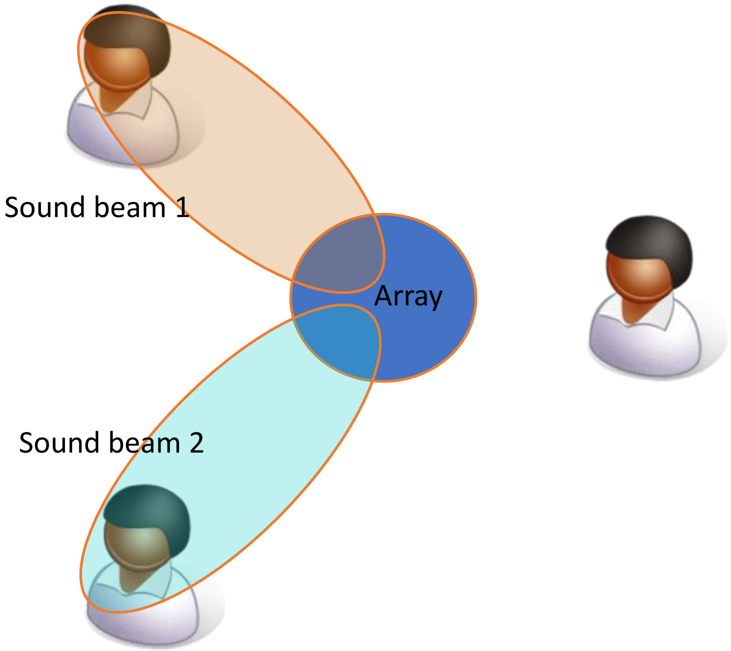
$$\mathbf{q}_{opt} = (\mathbf{G}^H \mathbf{G} + \beta \mathbf{I})^{-1} \mathbf{G}^H \mathbf{p}_T$$

Reduces of array effort, but increases error

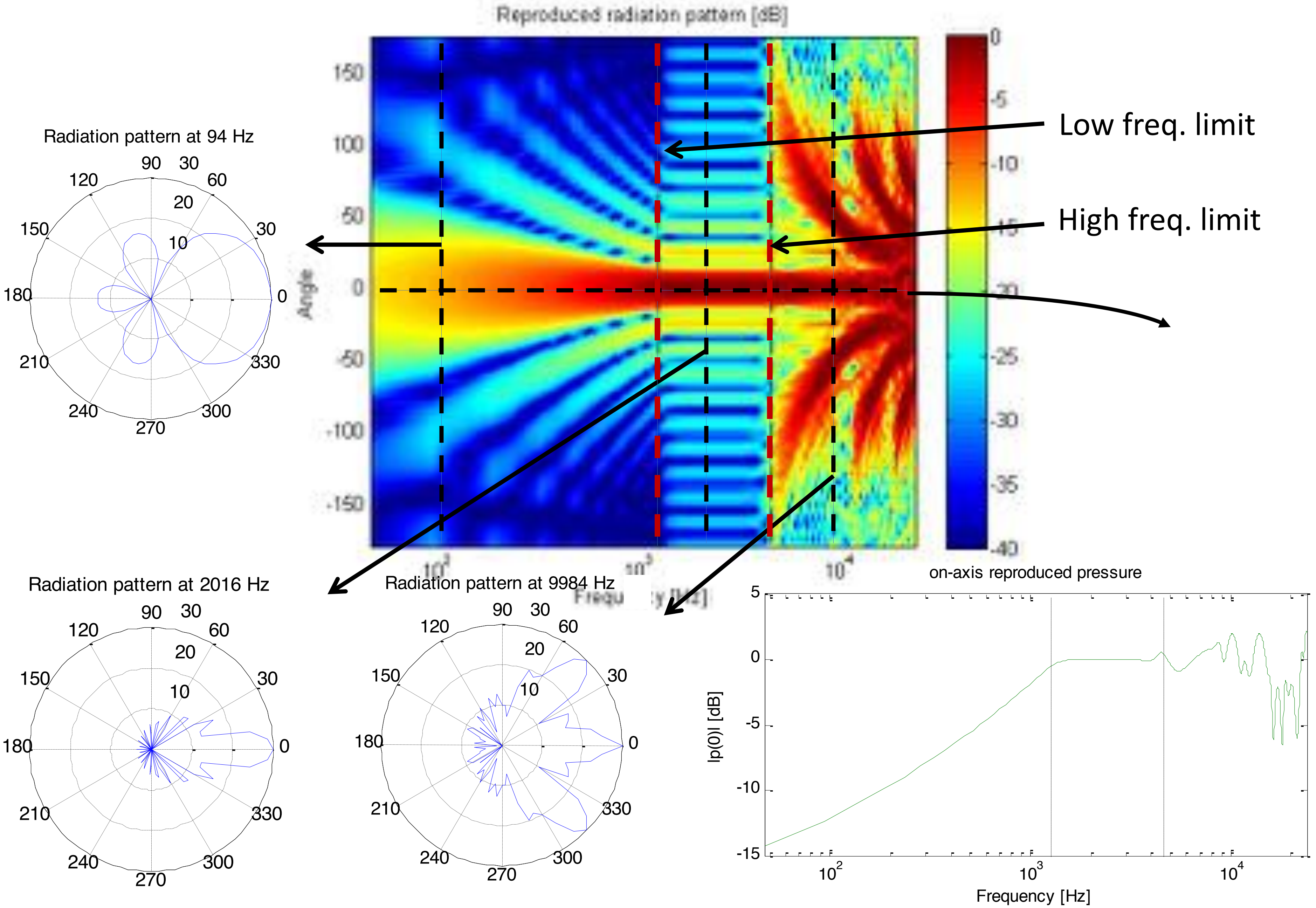


Kirkeby and Nelson, 1993. “Reproduction of plane wave sound fields”, JASA

Example: multi-zone with cylindrical array



Reproduced radiation pattern



Bright and dark zones

$$\mathbf{q} = \left[q_1(\omega), q_2(\omega), \ldots q_L(\omega)\right]^T$$

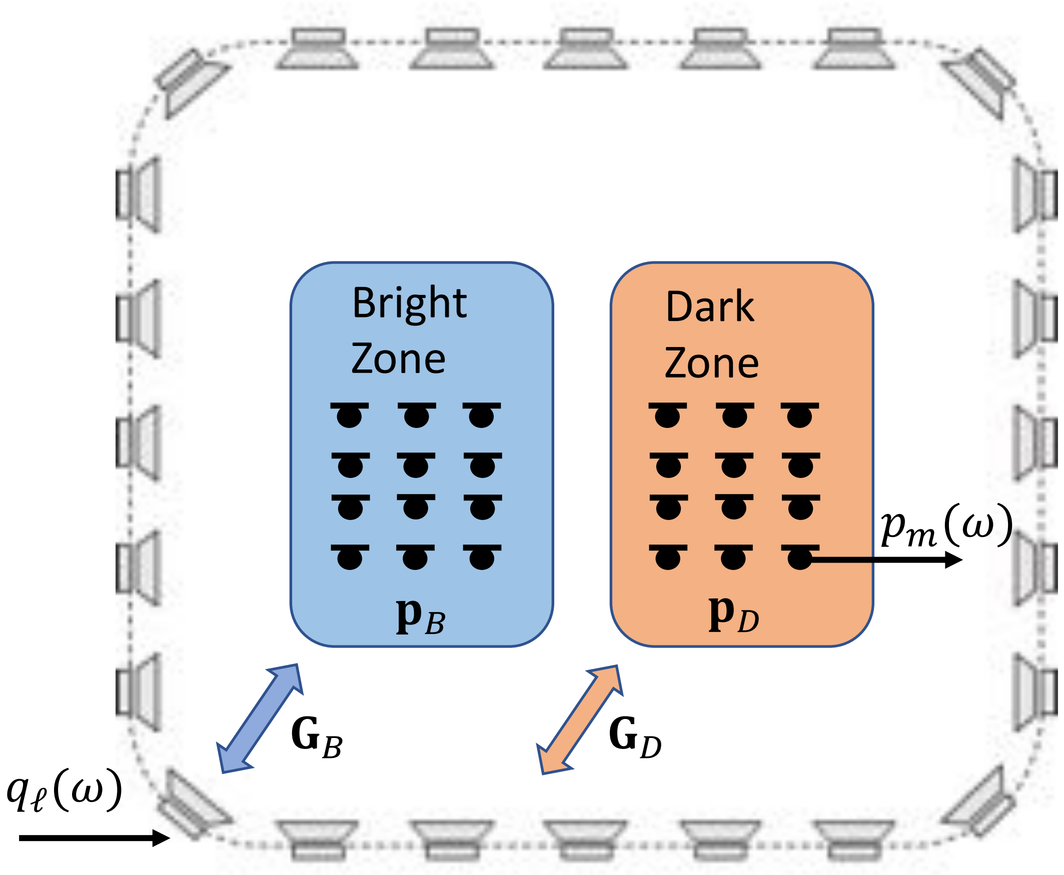
$$\mathbf{p}_B = \left[p_1^{(B)}(\omega), p_2^{(B)}(\omega), \ldots p_{N_B}^{(B)}(\omega) \right]^T$$

$$\mathbf{p}_D = \left[p_1^{(D)}(\omega), p_2^{(D)}(\omega), \ldots p_{N_D}^{(D)}(\omega) \right]^T$$

$$\mathbf{p}_B = \mathbf{G}_B \, \mathbf{q}$$

$$\mathbf{p}_D = \mathbf{G}_D \, \mathbf{q}$$

$$\mathbf{p} = \begin{vmatrix} \mathbf{p}_B \\ \hdashline \mathbf{p}_D \end{vmatrix}, \quad \mathbf{G} = \begin{vmatrix} \mathbf{G}_B \\ \hdashline \mathbf{G}_D \end{vmatrix}$$

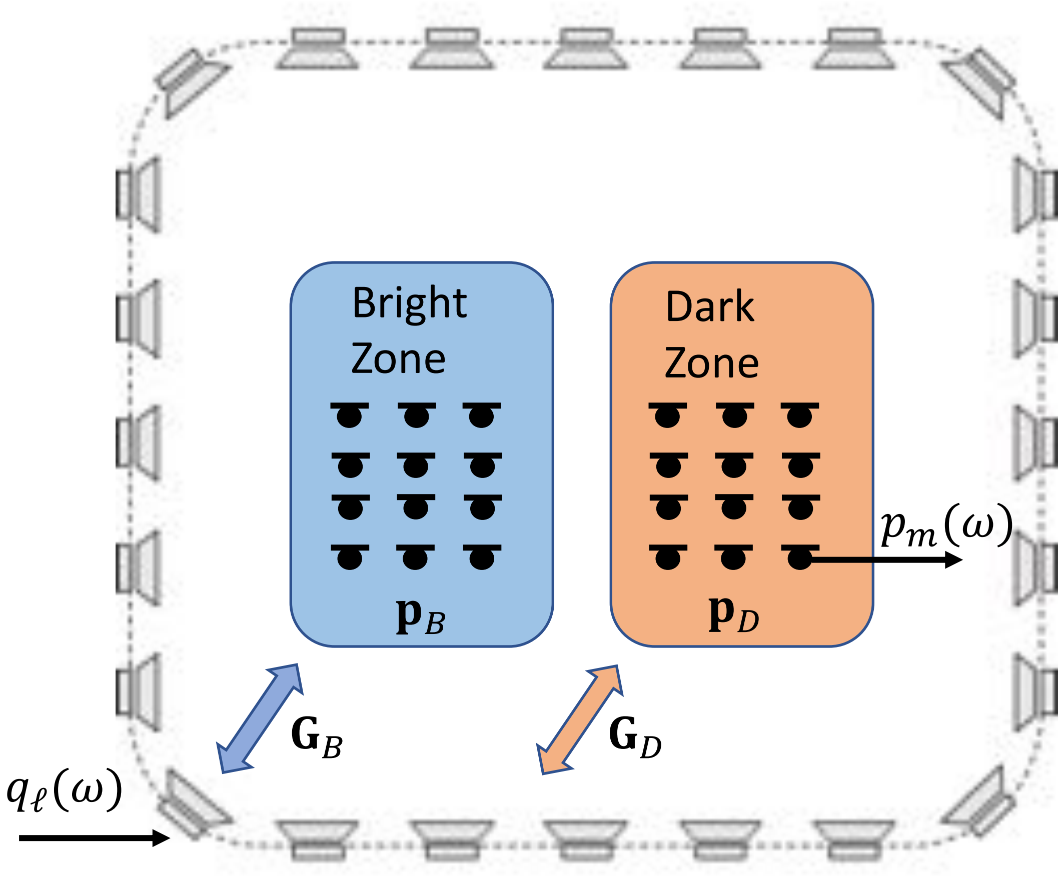


Acoustic contrast

It is the ratio of the average acoustic potential energy in two zones

$$C = \frac{<E_B>}{<E_D>} \approx \frac{\|\mathbf{p}_B\|^2 / N_B}{\|\mathbf{p}_D\|^2 / N_D}$$

$$= \frac{N_D}{N_B} \frac{\mathbf{q}^H \mathbf{G}_B{}^H \mathbf{G}_B \mathbf{q}}{\mathbf{q}^H \mathbf{G}_D{}^H \mathbf{G}_D \mathbf{q}}$$



$$\mathbf{p}_{B/D} = \mathbf{G}_{B/D}\, \mathbf{q}$$

Acoustic contrast maximisation

- Direct formulation: maximises the energy in the bright zone while keeping the energy in the dark zone to a constant value D and the effort below a given value E

$$\text{Maximise} \quad \|\mathbf{p}_B\|^2 \quad \text{s.t.} \quad \|\mathbf{p}_D\|^2 = D \quad \text{and} \quad \|\mathbf{q}\|^2 \leq E$$

- Indirect formulation: minimise the energy in the dark zone while keeping the energy in the bright zone to a constant value B and the effort below a given value E

$$\text{Minimise} \quad \|\mathbf{p}_D\|^2 \quad \text{s.t.} \quad \|\mathbf{p}_B\|^2 = B \quad \text{and} \quad \|\mathbf{q}\|^2 \leq E$$

Elliott, S.J., et al., 2012. "Robustness and regularization of personal audio systems". *TASLP*.

Acoustic contrast maximisation (direct)

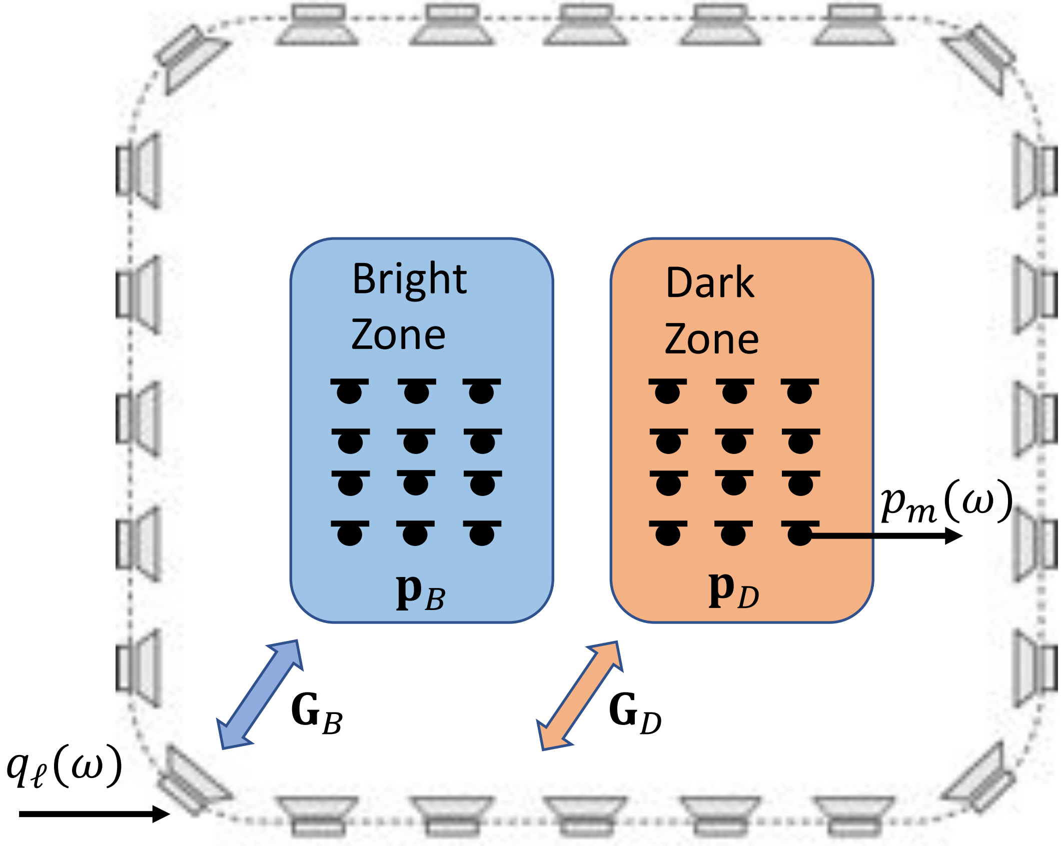
- Maximise $\|\mathbf{p}_B\|^2$ s.t. $\|\mathbf{p}_D\|^2 = D$ and $\|\mathbf{q}\|^2 \leq E$

$$L = \mathbf{q}^H \mathbf{G}_B^H \mathbf{G}_B \mathbf{q} - \lambda_1 (\mathbf{q}^H \mathbf{G}_D^H \mathbf{G}_D \mathbf{q} - D) - \lambda_2 (\mathbf{q}^H \mathbf{q} - E)$$

$$\lambda_1 \mathbf{q}_{opt} = \left(\mathbf{G}_D^H \mathbf{G}_D\right)^{-1} \left(\mathbf{G}_B^H \mathbf{G}_B - \lambda_2 \mathbf{I}\right) \mathbf{q}_{opt}$$

- The optimal solution $\mathbf{q}_{opt}$ is the eigenvector associated with the largest eigenvalue λ_1

$$\mathbf{p}_{B/D} = \mathbf{G}_{B/D}\, \mathbf{q}$$



Elliott, S.J., et al., 2012. “Robustness and regularization of personal audio systems”. *TASLP*.
Choi and Kim, 2002. “Generation of an acoustically bright zone with an illuminated region using multiple sources”. *JASA*

Acoustic contrast maximisation (indirect)

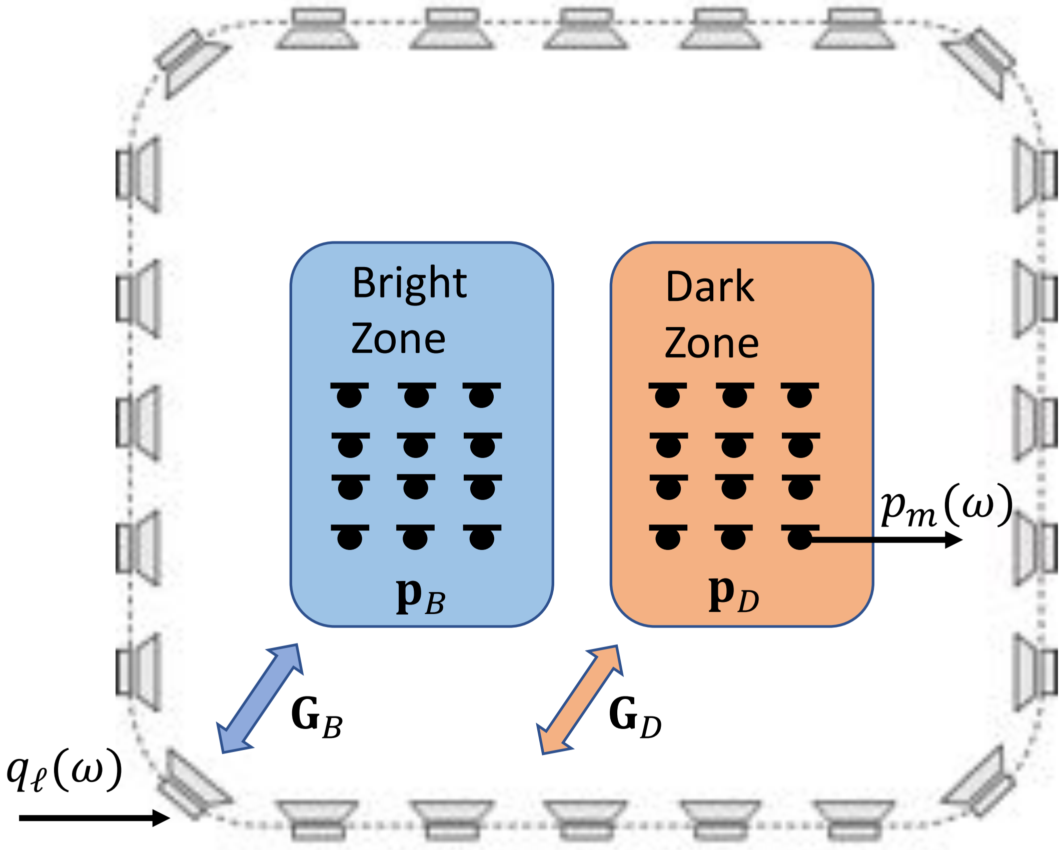
- Minimise $\|\mathbf{p}_D\|^2$ s.t. $\|\mathbf{p}_B\|^2 = B$ and $\|\mathbf{q}\|^2 \le E$

$$L = \mathbf{q}^H \mathbf{G}_D^H \mathbf{G}_D \mathbf{q} + \lambda_1 (\mathbf{q}^H \mathbf{G}_B^H \mathbf{G}_B \mathbf{q} - B) + \lambda_2 (\mathbf{q}^H \mathbf{q} - E)$$

$$\lambda \mathbf{q}_{opt} = \left(\mathbf{G}_D^H \mathbf{G}_D + \lambda_2 \mathbf{I}\right)^{-1} \left(\mathbf{G}_B^H \mathbf{G}_B\right) \mathbf{q}_{opt}$$

- The optimal solution $\mathbf{q}_{opt}$ is the eigenvector associated with the largest eigenvalue λ

$$\mathbf{p}_{B/D} = \mathbf{G}_{B/D}\, \mathbf{q}$$



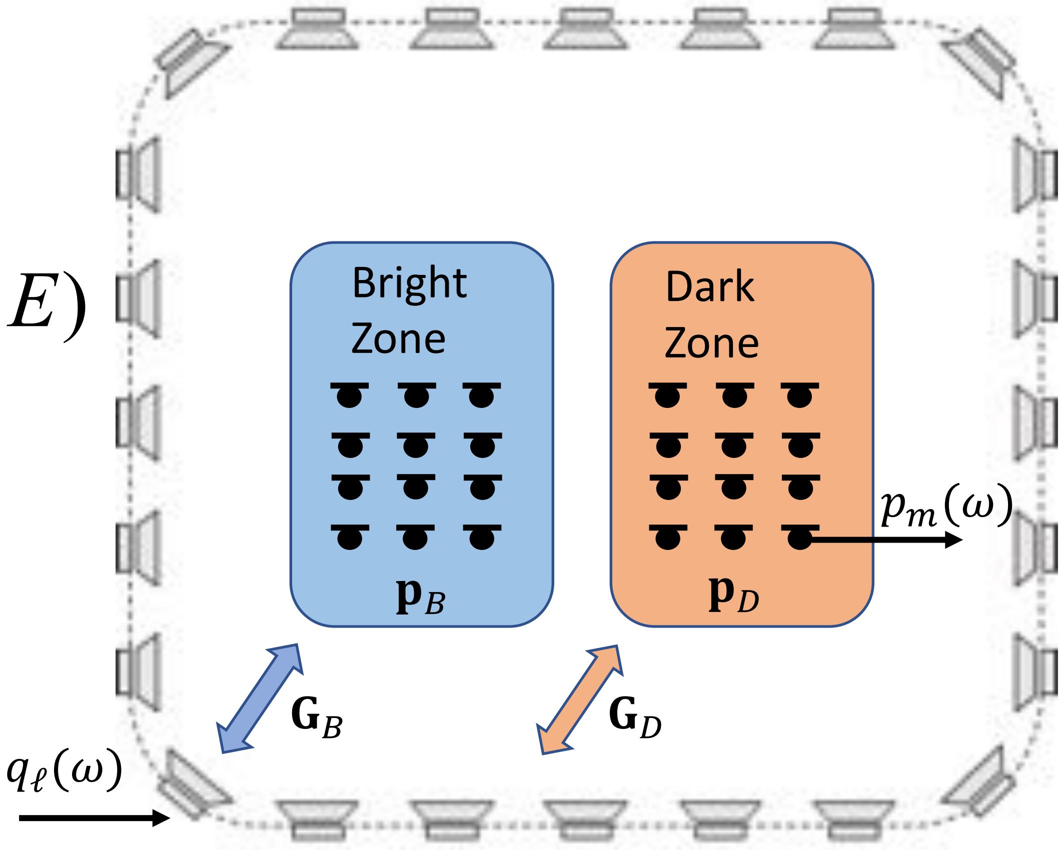
Elliott, S.J., et al., 2012. "Robustness and regularization of personal audio systems". *TASLP*.

Energy difference maximisation

- Maximise energy difference

$$\|\mathbf{p}_B\|^2 - \alpha\|\mathbf{p}_D\|^2 \quad \text{s.t. } \|\mathbf{q}\|^2 \leq E$$

$$\mathbf{p}_{B/D} = \mathbf{G}_{B/D}\,\mathbf{q}$$



$$L = \mathbf{q}^H\mathbf{G}_B^H\mathbf{G}_B\mathbf{q} - \alpha\,\mathbf{q}^H\mathbf{G}_D^H\mathbf{G}_D\mathbf{q} - \lambda_2(\mathbf{q}^H\mathbf{q} - E)$$

$$\lambda_2\mathbf{q}_{opt} = \left(\mathbf{G}_B^H\mathbf{G}_B - \alpha\mathbf{G}_D^H\mathbf{G}_D\right)\mathbf{q}_{opt}$$

- The optimal solution $\mathbf{q}_{opt}$ is the eigenvector associated with the largest eigenvalue λ_2

Shin et al., 2010. “Maximization of acoustic energy difference between two spaces”. JASA

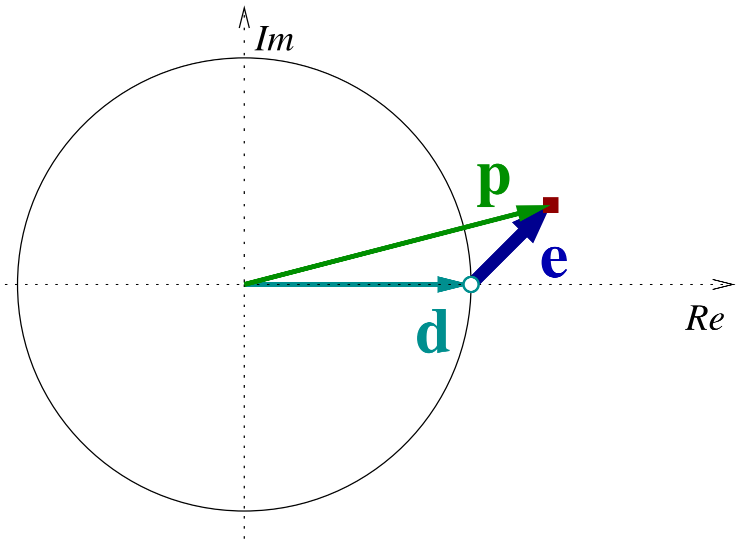
Performance measures

Measures of personal sound quality

- Interference
  - From sounds other than the intended personal sound
- Robustness
  - For any practical implementation
- Timbral quality
  - Quality of the received personal sound
- Spatial quality
  - Reproduction artefacts or spatial distortion

Emiya et al., 2011. Subjective and objective quality assessment of audio source separation, TASLP
Francombe et al., 2014. Elicitation of attributes for the evaluation of audio-on-audio interference, JASA
Conetta et al., 2014. Spatial Audio Quality Perception (Part 2): A linear regression model, JAES
George et al., 2010. Development and validation of an unintrusive model for predicting..., JAES
Jackson et al., 2008. QESTRAL (Part 3): System and metrics for spatial quality prediction, AES

Goodness of fit: MMSE

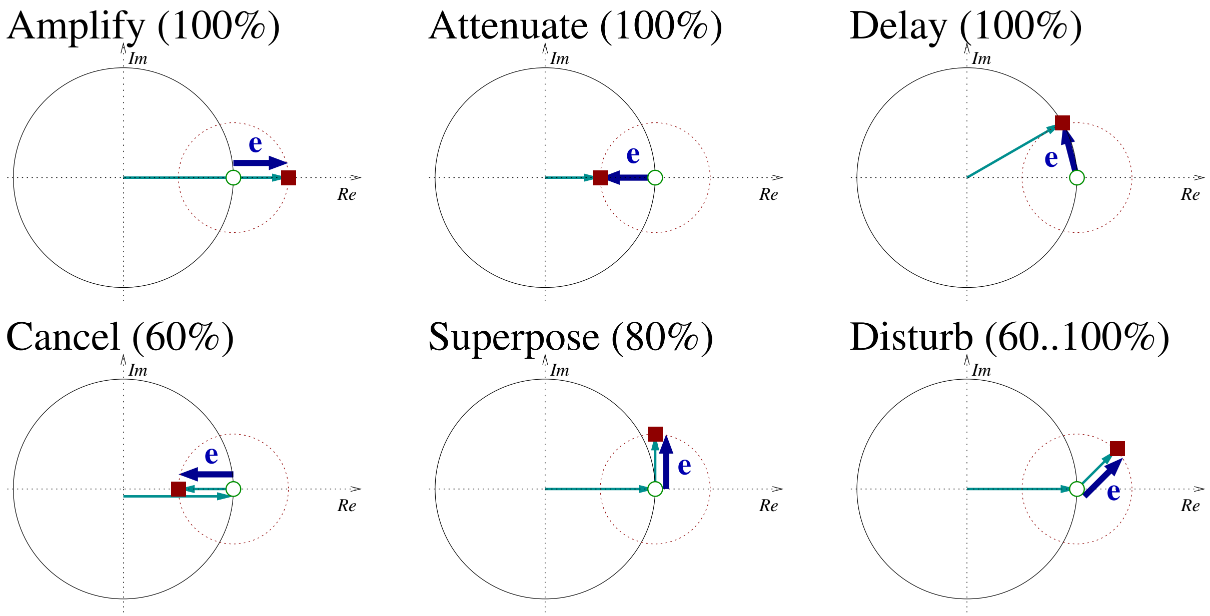


- Target sound field
- Given a target sound field, $\mathbf{d}_A(f)=[d_{A,1},\ldots,d_{A,m},\ldots,d_{A,M}]^T$, the mean squared error is

$$\mathrm{MSE} = \frac{1}{M}\sum_{m=1}^{M}\left|p_{A,m} - d_{A,m}\right|^2$$

or

$$\mathrm{MSE} = \frac{1}{M}(\mathbf{p}_A - \mathbf{d}_A)^H(\mathbf{p}_A - \mathbf{d}_A)$$



Kirkeby & Nelson, 1993. Reproduction of plane wave sound fields, JASA

Independent sampling points

Sources $\mathbf{q}(f)=[q_1,\ldots,q_l,\ldots,q_L]^T$

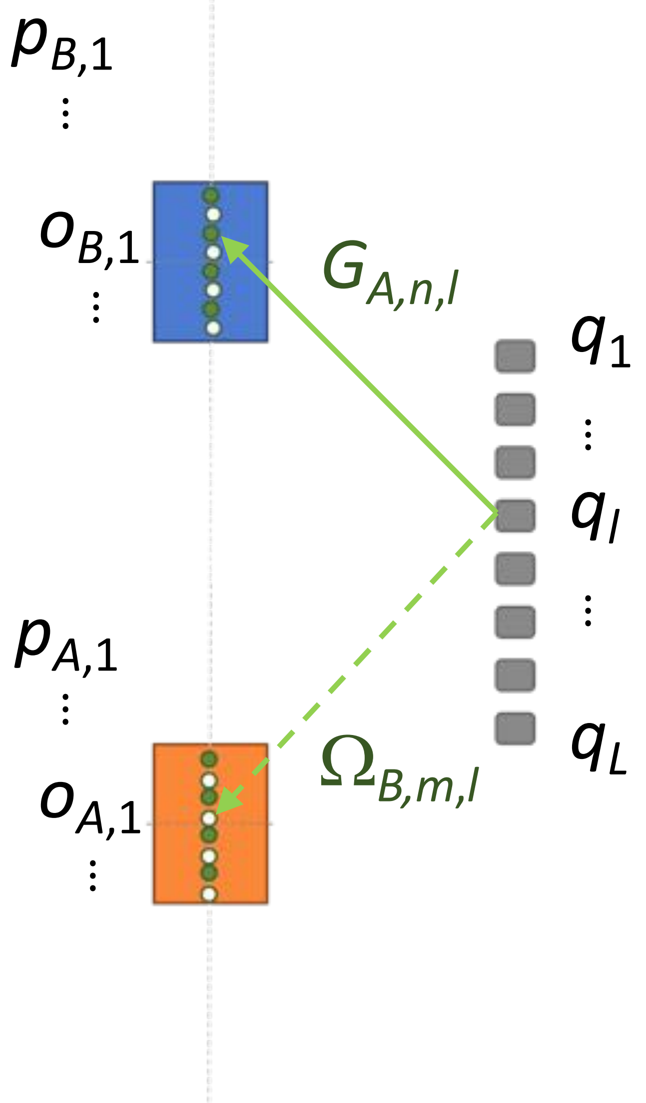
Control mics $\mathbf{p}_A(f)=[p_{A,1},\ldots,p_{A,n},\ldots,p_{A,N}]^T$

Monitor mics $\mathbf{o}_A(f)=[o_{A,1},\ldots,o_{A,m},\ldots,o_{A,M}]^T$

Transfer functions

$$\mathbf{G}_A(f) = \begin{bmatrix} G_{A,1,1} & \cdots & G_{A,1,L} \\ \vdots & \ddots & \vdots \\ G_{A,N,1} & \cdots & G_{A,N,L} \end{bmatrix}; \quad \mathbf{\Omega}_A(f) = \begin{bmatrix} \Omega_{A,1,1} & \cdots & \Omega_{A,1,L} \\ \vdots & \ddots & \vdots \\ \Omega_{A,M,1} & \cdots & \Omega_{A,M,L} \end{bmatrix}$$

Relations $\begin{bmatrix} \mathbf{p}_A \\ \mathbf{p}_B \end{bmatrix} = \begin{bmatrix} G_A \\ G_B \end{bmatrix} [\mathbf{q}]$ $\begin{bmatrix} \mathbf{o}_A \\ \mathbf{o}_B \end{bmatrix} = \begin{bmatrix} \Omega_A \\ \Omega_B \end{bmatrix} [\mathbf{q}]$



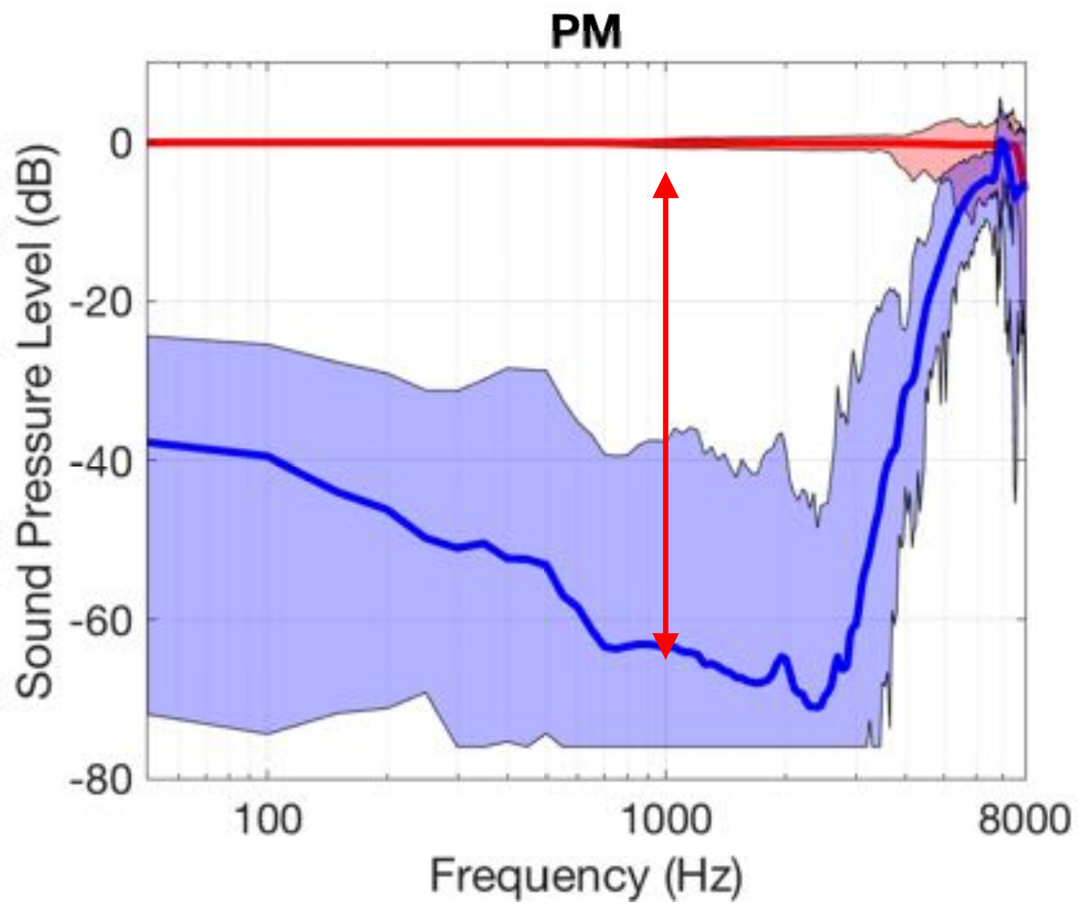
Coleman, 2014. Loudspeaker array processing for personal sound zone reproduction. PhD thesis

Contrast

- Contrast reflects the ratio of the sound pressure levels (SPLs) in the zones

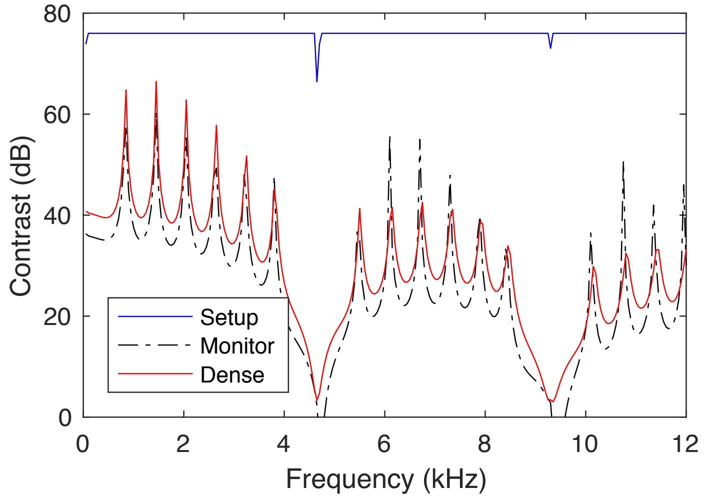
$$C = 10\log_{10}\frac{N_B \mathbf{p}_A^H \mathbf{p}_A}{N_A \mathbf{p}_B^H \mathbf{p}_B}$$

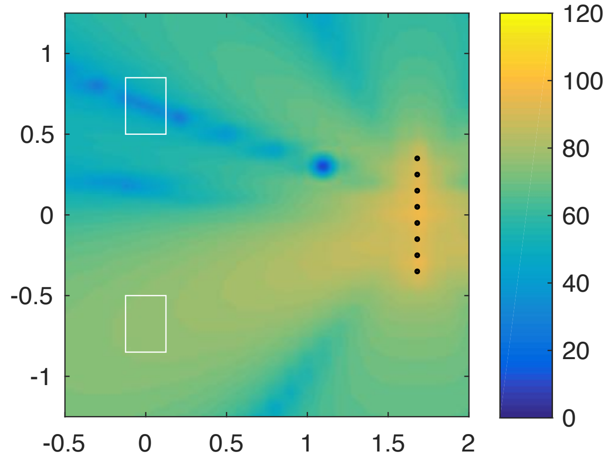
$$C = 10\log_{10}\frac{M_B \mathbf{o}_A^H \mathbf{o}_A}{M_A \mathbf{o}_B^H \mathbf{o}_B}$$

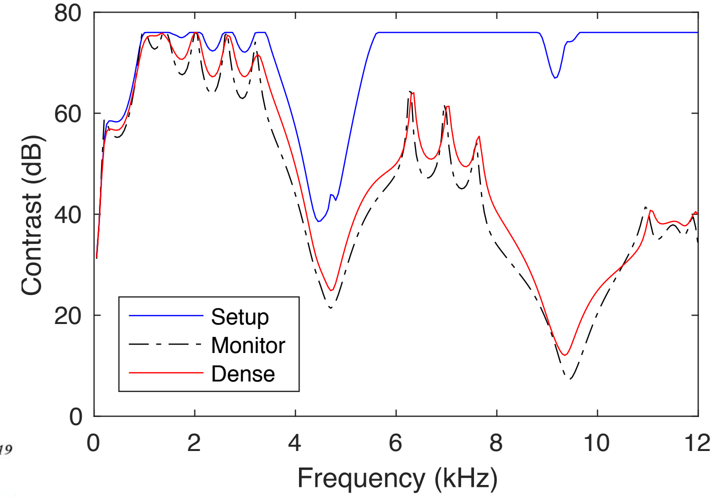


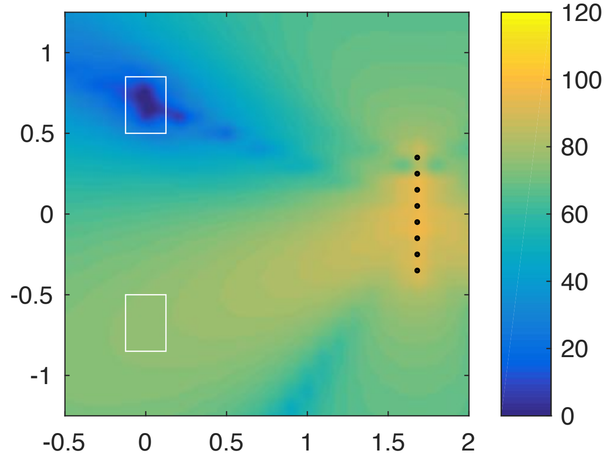
Choi & Kim, 2002. Generation of an acoustically bright zone with an illuminated region..., JASA
Coleman et al., 2014. Acoustic contrast, planarity and robustness of sound zone methods..., JASA

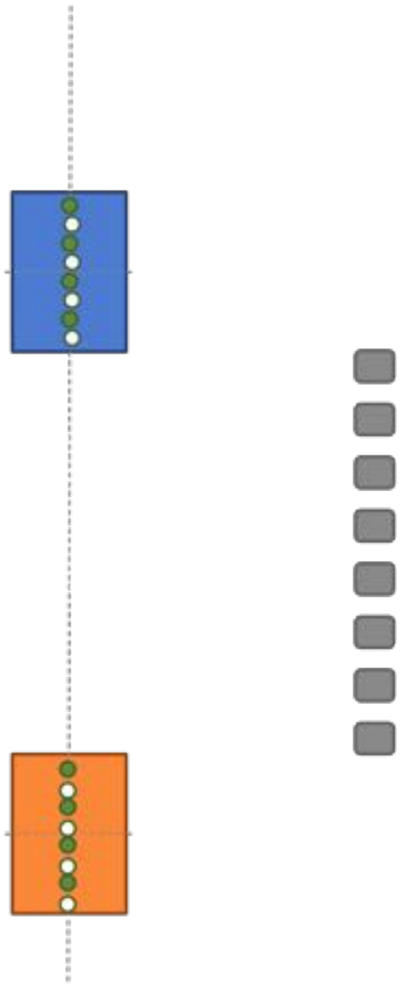
Calibration (control) and evaluation (monitor) microphones









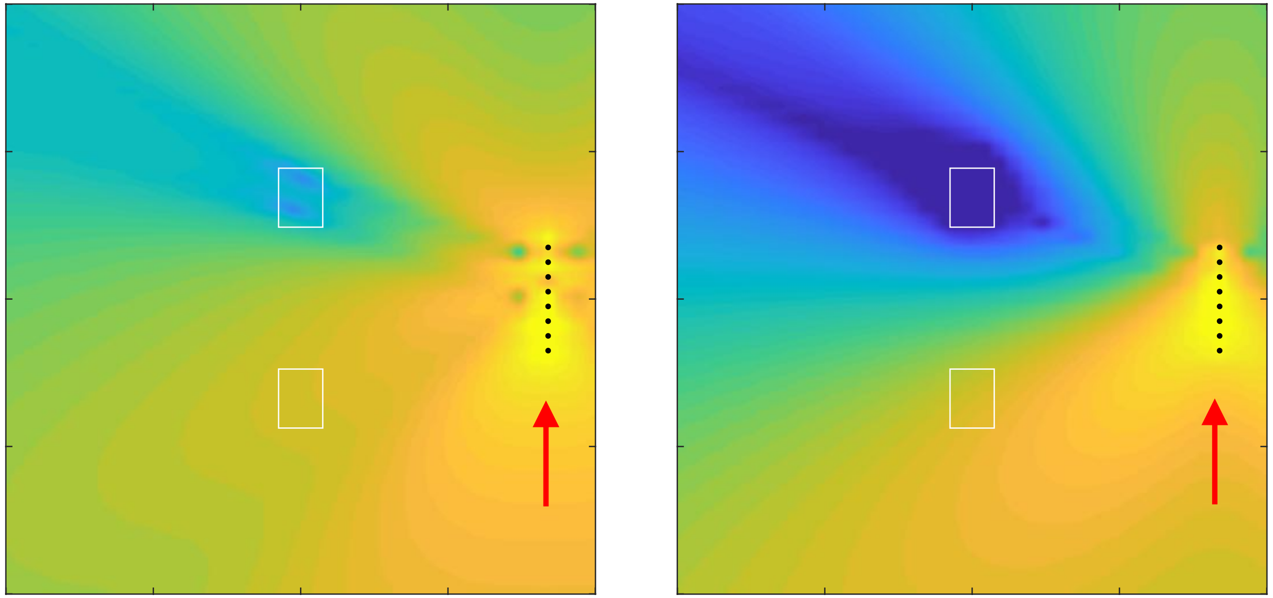


Jackson & Ross, 1996. Application of active noise control to corporate aircraft, ASME
Møller & Olsen, 2016. Sound zones: On performance prediction…, AES conf

Effort

- Efficiency
  - Total acoustical power
- Robustness
  - Errors in calibration
  - Influence of reflections
- Effort metric:

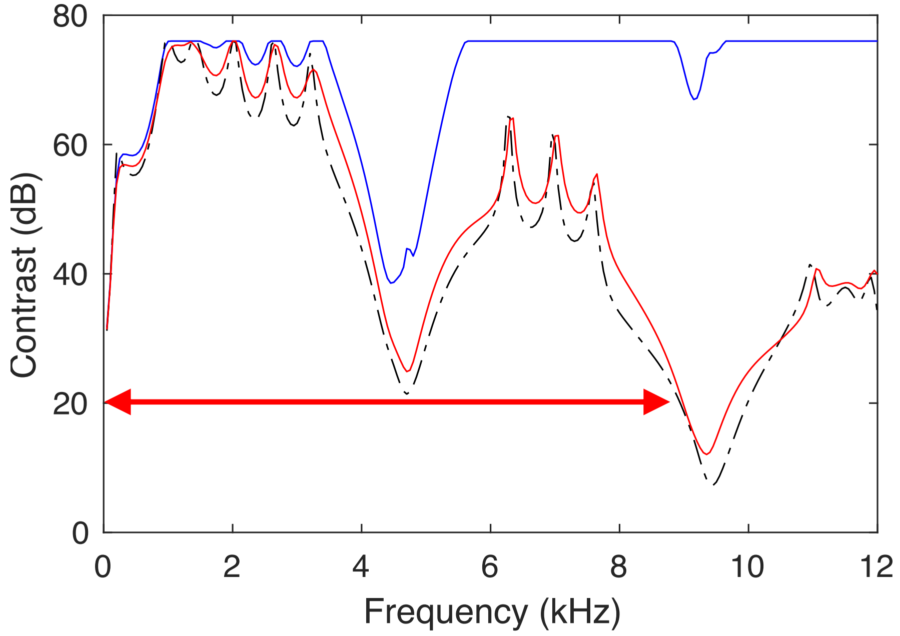
$$E = 10\log_{10}\frac{\mathbf{q}^H\mathbf{q}}{|q_r|^2}$$

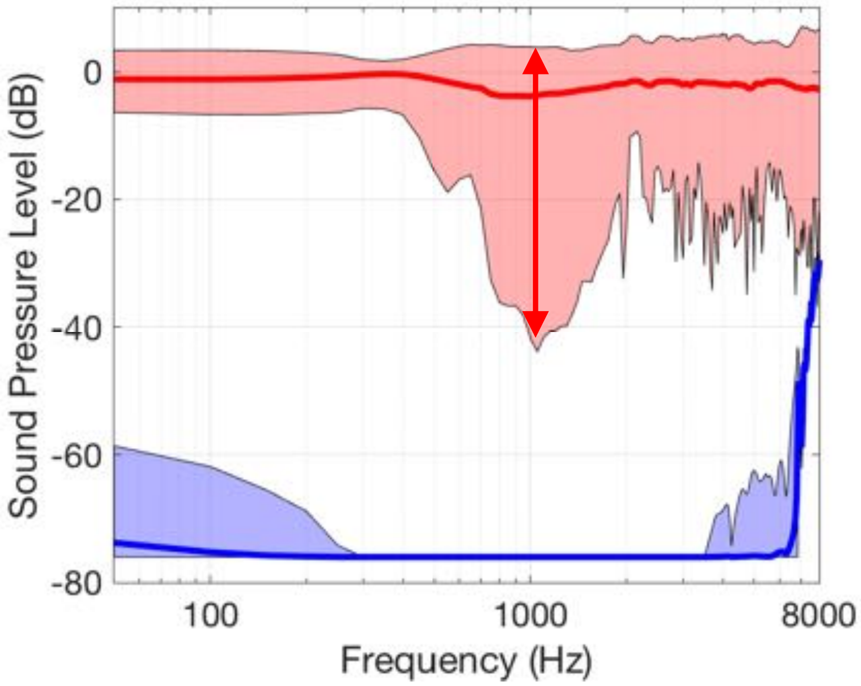


Elliott et al., 2012. Robustness and regularization of personal audio systems, TASLP
Coleman et al., 2014. Acoustic contrast, planarity and robustness of sound zone methods..., JASA

Frequency response

- Timbral quality
- Effective bandwidth
- Flatness across frequency
  - Tone, colouration
  - Comb filtering



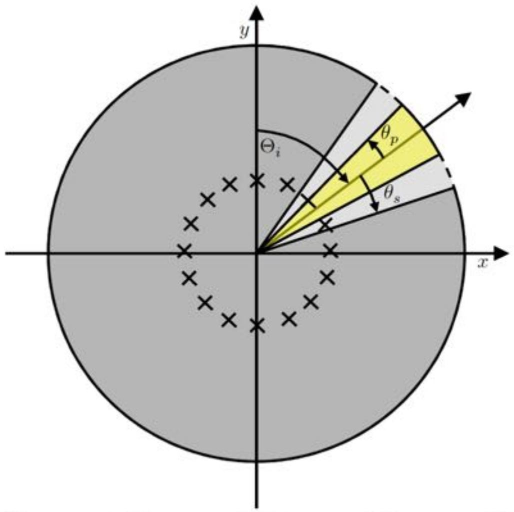


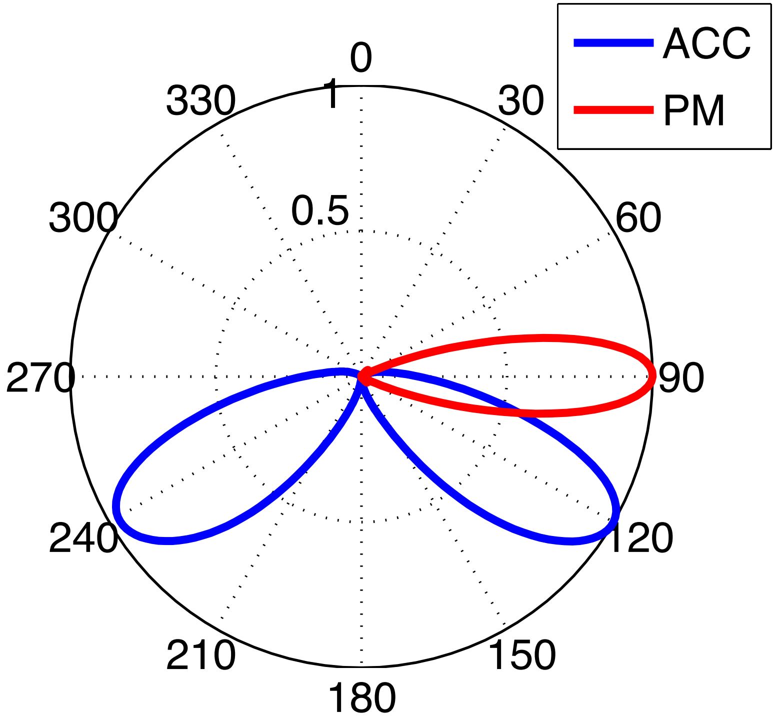
PEAQ

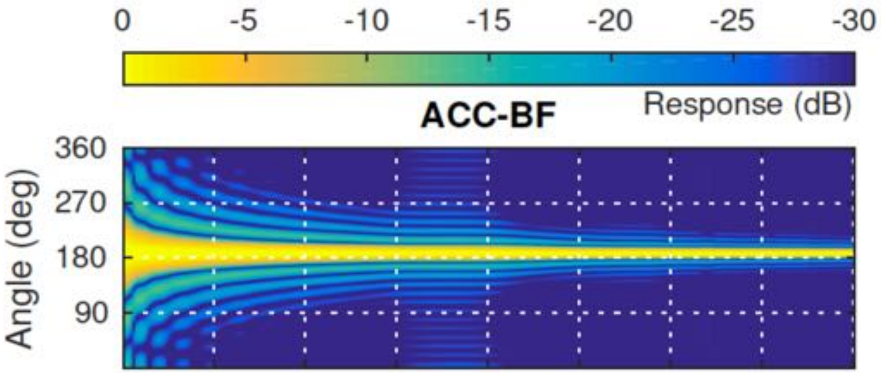
Baykaner et al., 2015. The relationship between target quality and interference in sound zone, JAES

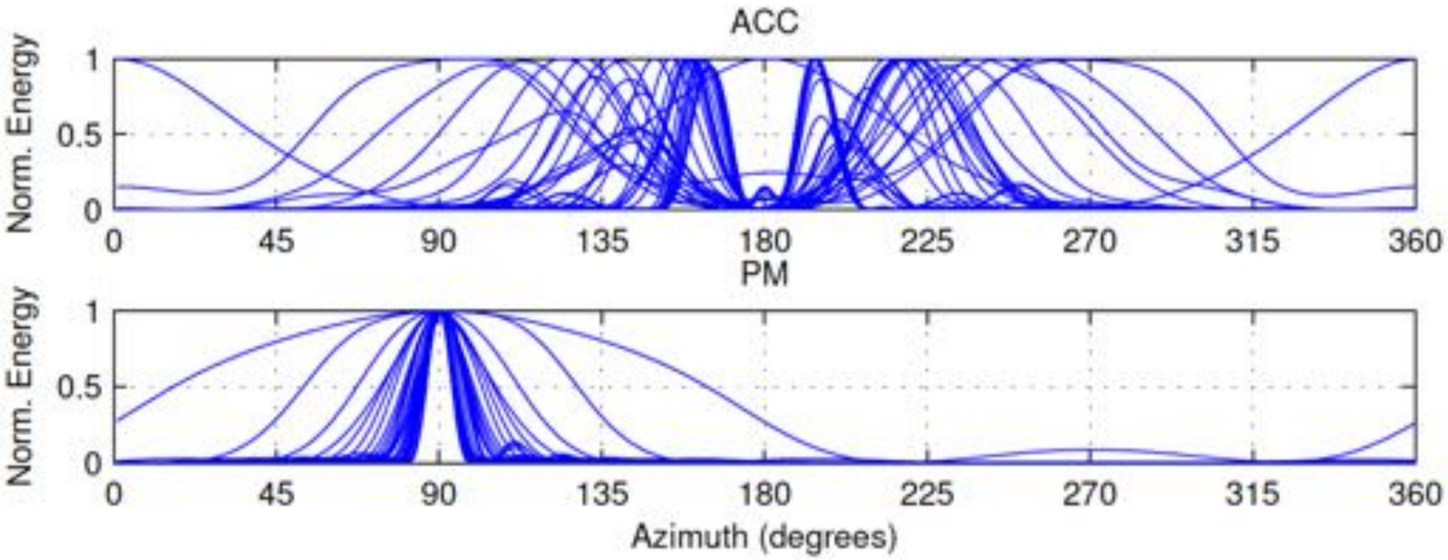
Spatial measures

- Estimated angular spectrum



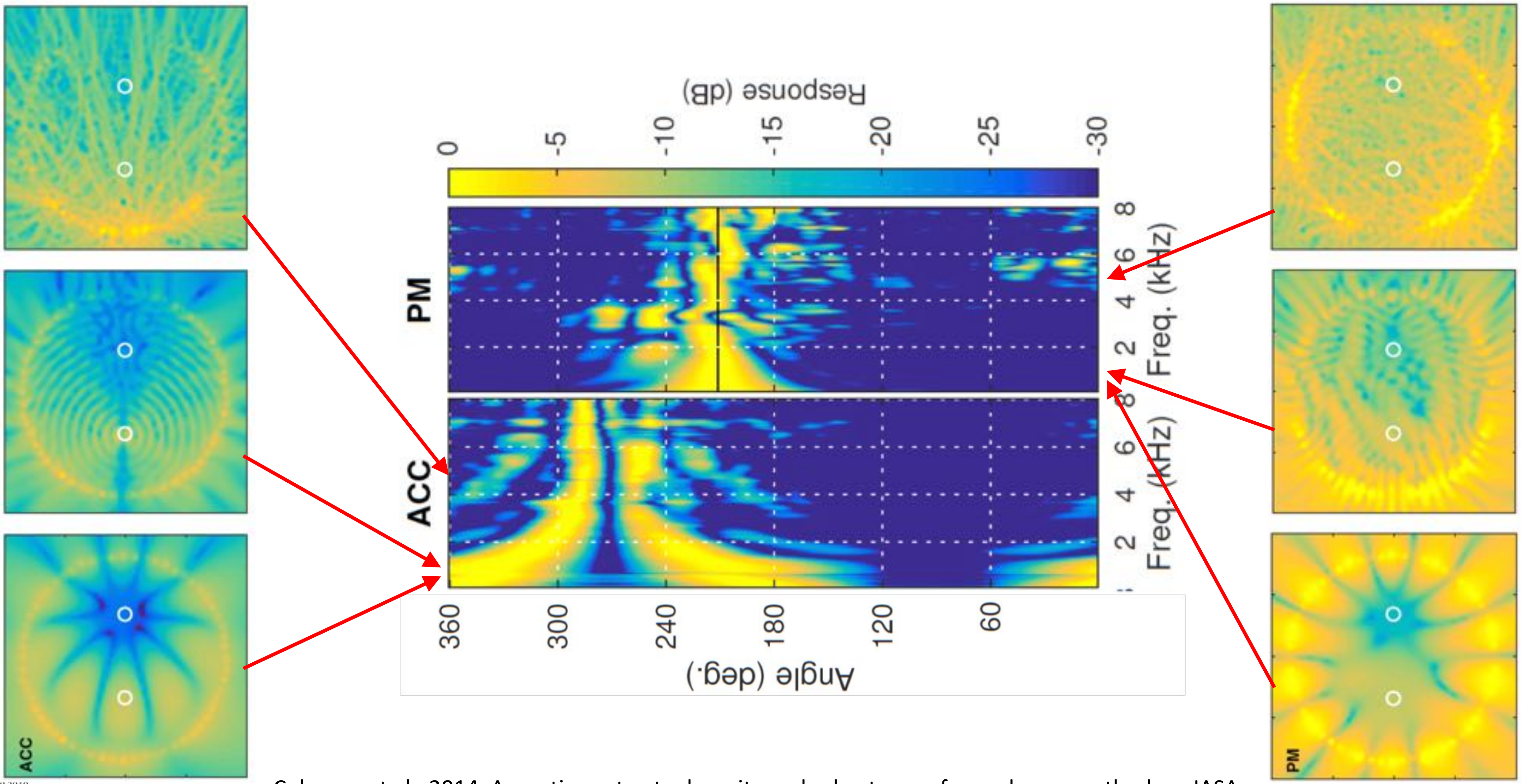






Williams, 1999, Fourier Acoustics

Spatial examples



Coleman et al., 2014. Acoustic contrast, planarity and robustness of sound zone methods..., JASA

Planarity

- Summary measure of sound field homogeneity, i.e., vs. plane wave

$$\eta = \frac{\sum_i w_i \mathbf{u}_i \cdot \mathbf{u}_\alpha}{\sum_i w_i}$$

where unit vector $\mathbf{u}_i$ points in the ith look direction and has weight w_i, and principal unit vector $\mathbf{u}_\alpha$ is defined as α=argmax$_i$ w_i.

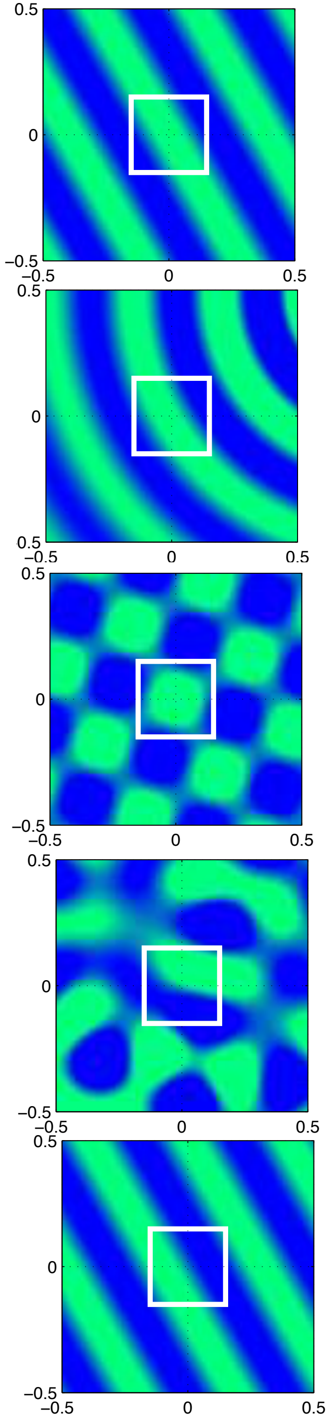
Plane wave
100%

Point source
~90%

Orthogonal
50%

Diffuse
~0%

Standing wave
0%



Merimaa, 2007. Energetic sound field analysis of loudspeaker reproduction, AES
Jackson et al., 2013. Sound field planarity characterized by superdirective beamforming, POMA

ICASSP 2019

Metrics summary

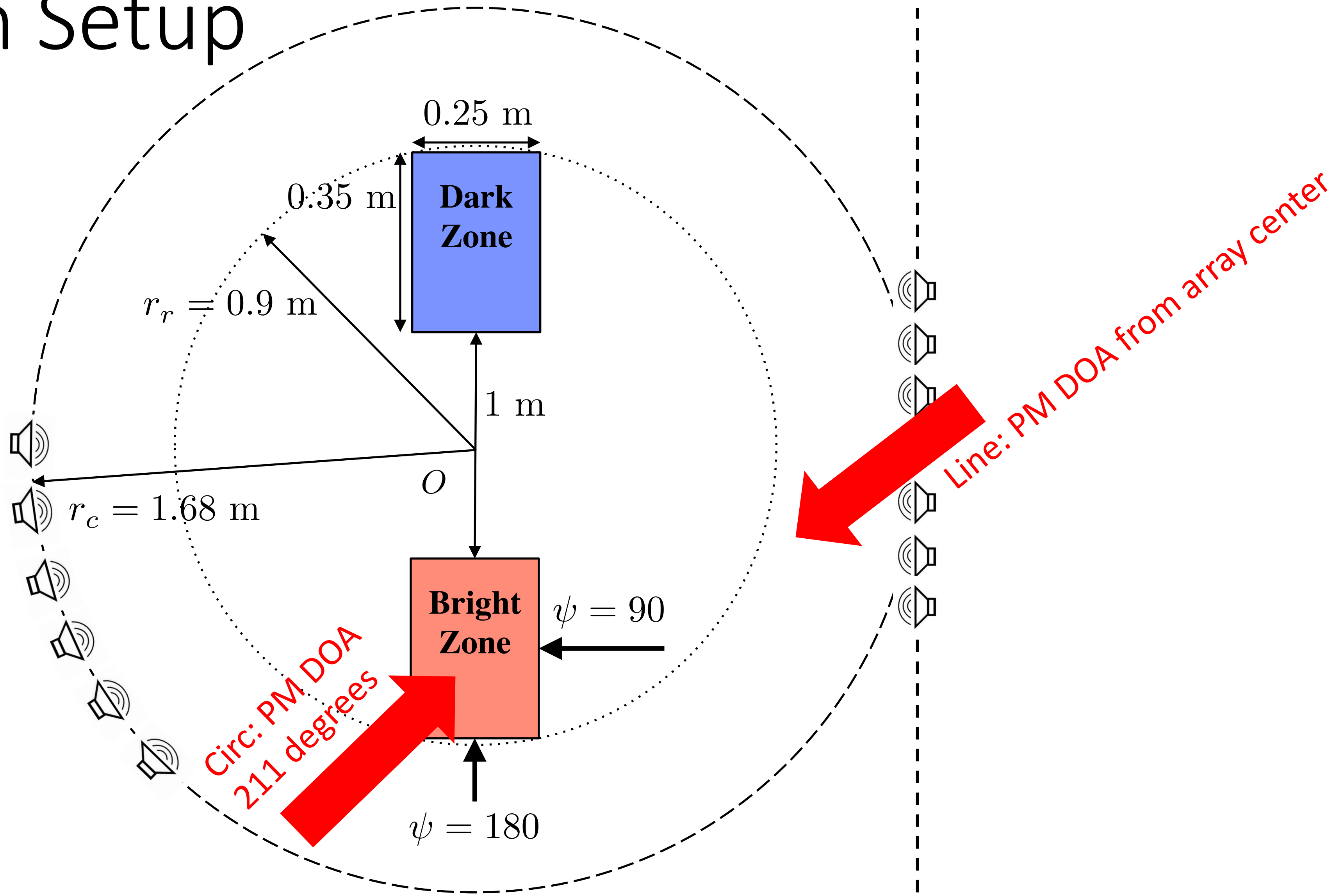
- Goodness of fit: **Mean squared error**, MSE
  for least-squares optimization formulation
- Interference: **Contrast**, C
  for assessing relative SPL or loudness
- Robustness: **Effort**, E
  as a strong indicator for practical performance
- Bandwidth: **Frequency range**, F
  for application-driven content coverage
- Directivity: **Planarity**, η
  for spatial coherence at the listener

Fundamental Relationships under Ideal Conditions: Cost Function and System Setup

Aims

- Demonstrate the characteristics of ACC and PM:
  - Sound fields reproduced
  - Frequency range of operation
- Demonstrate the influence of loudspeaker array design:
  - Circular array (fully enclosed zones, relatively sparse loudspeaker spacing)
  - Line array (compact, relatively close loudspeaker spacing)

Simulation Setup



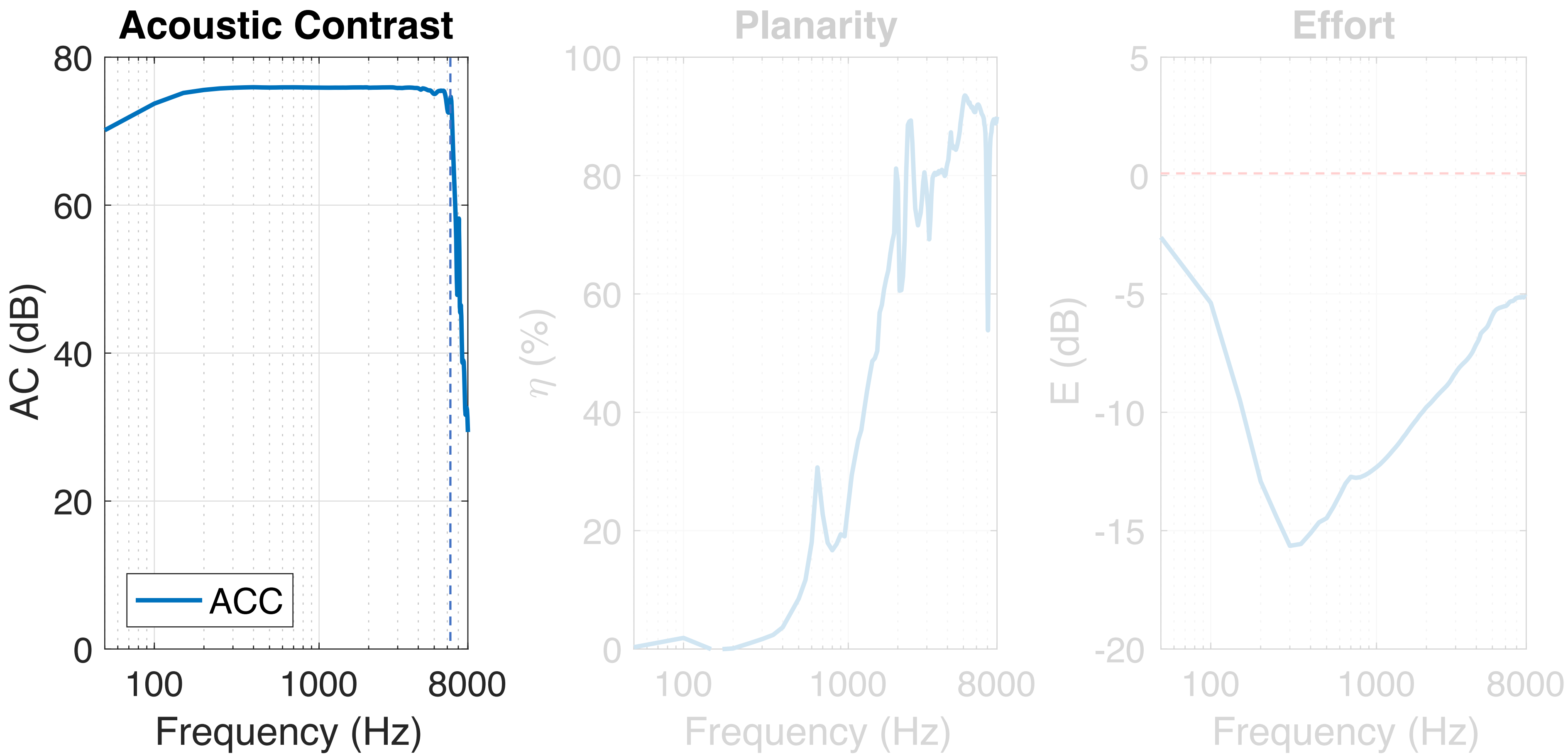
Coleman P., 2014. "Loudspeaker Array Processing for Personal Sound Zone Reproduction", PhD thesis, Univ. of Surrey, UK

Simulation Setup

- Two arrays: line and circle
  - Circle: 60 loudspeakers, 1.68 m radius
  - Line: 8 loudspeakers, 10 cm spacing
- Two zones
- Regularization: Max matrix condition number of 10^{10}, Array Effort limit of 0 dB
- Listening level 76 dB SPL
- Acoustic contrast capped at 76 dB (i.e. no audible sound in dark zone)

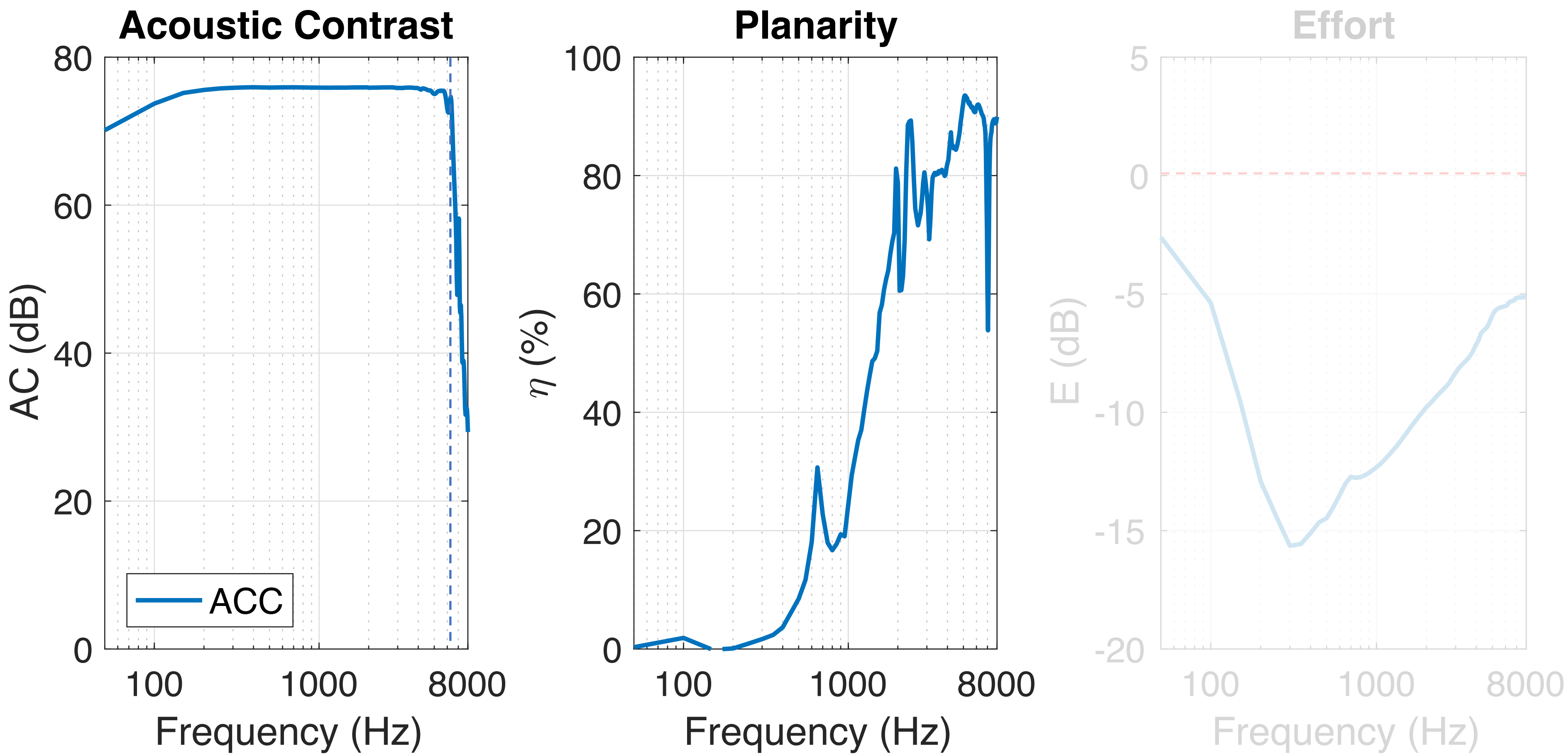
Acoustic Contrast Control: Circular Array

- Performance over frequency



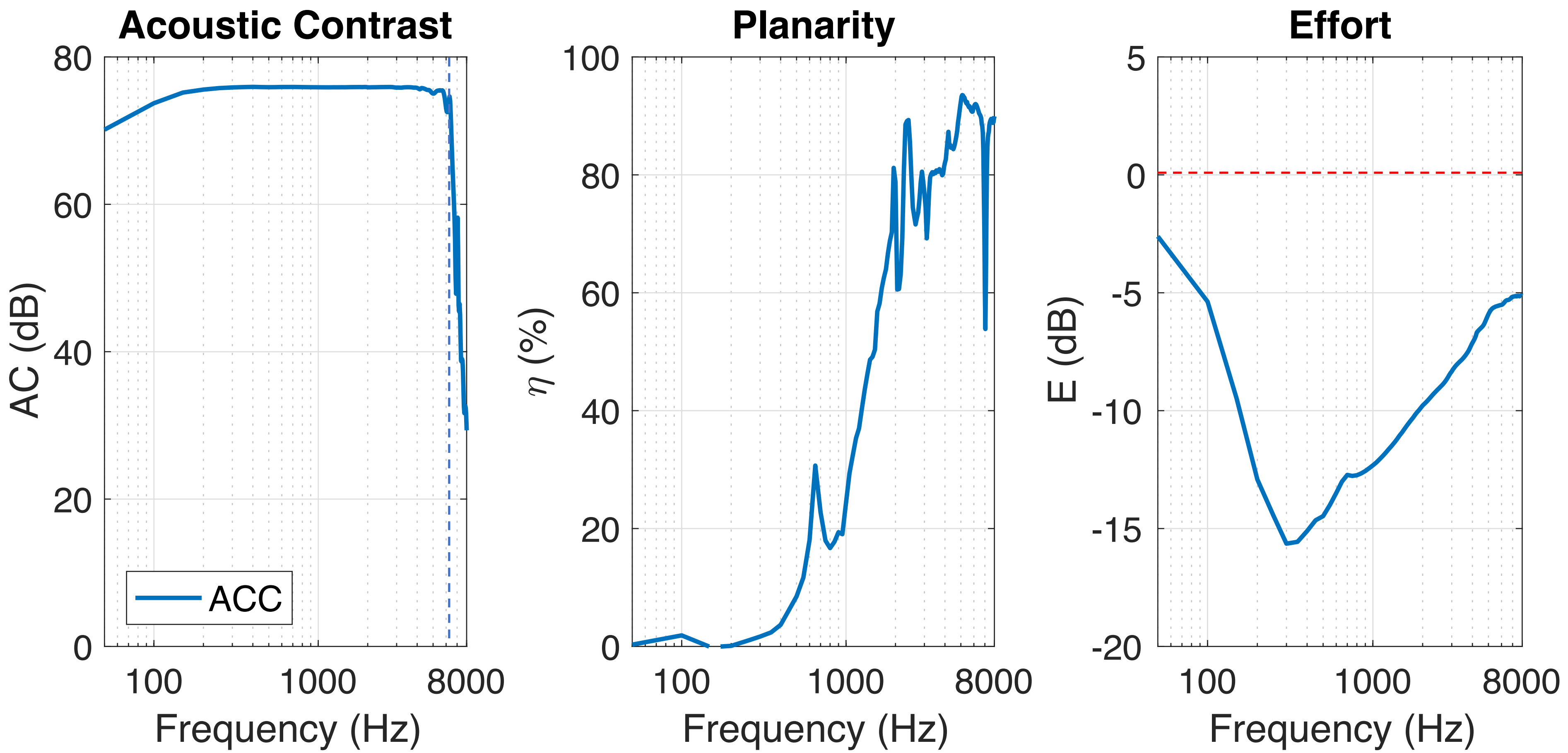
Acoustic Contrast Control: Circular Array

- Performance over frequency



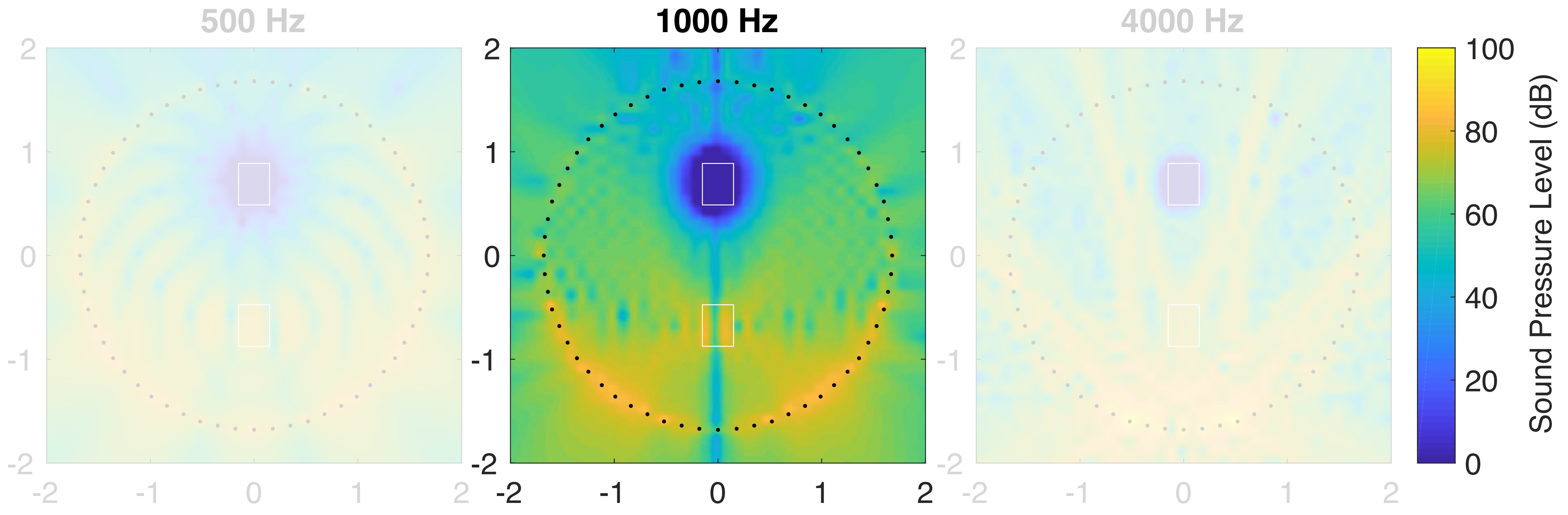
Acoustic Contrast Control: Circular Array

- Performance over frequency



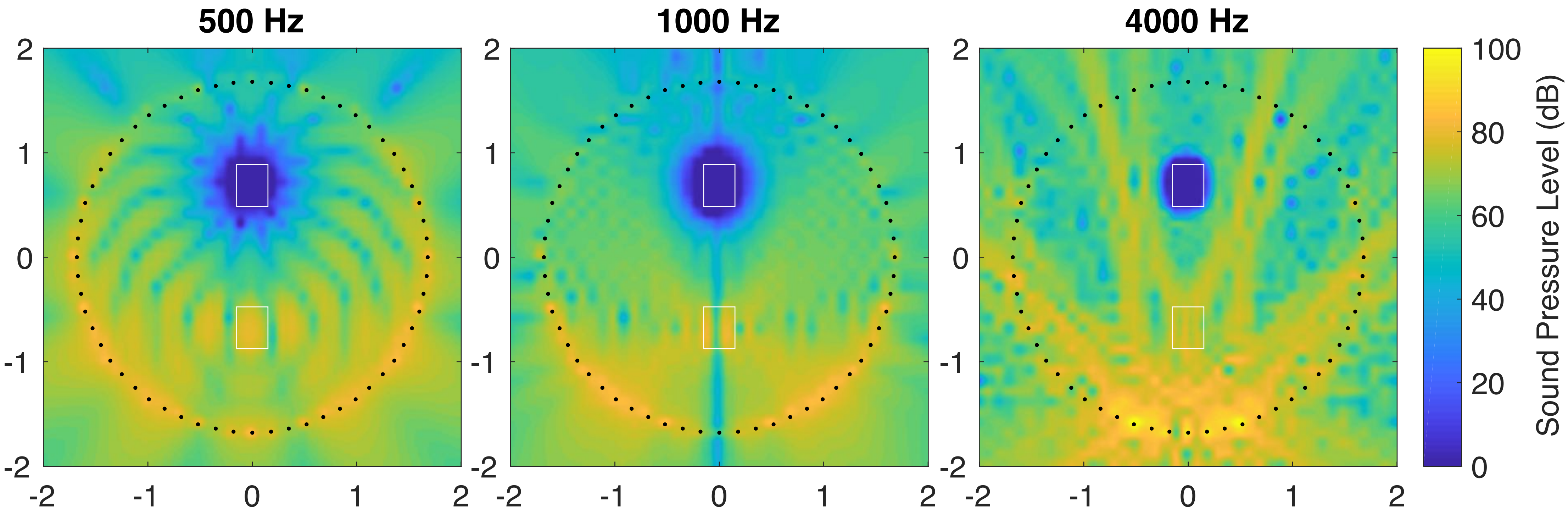
Acoustic Contrast Control: Circular Array

- Sound field: sound pressure level



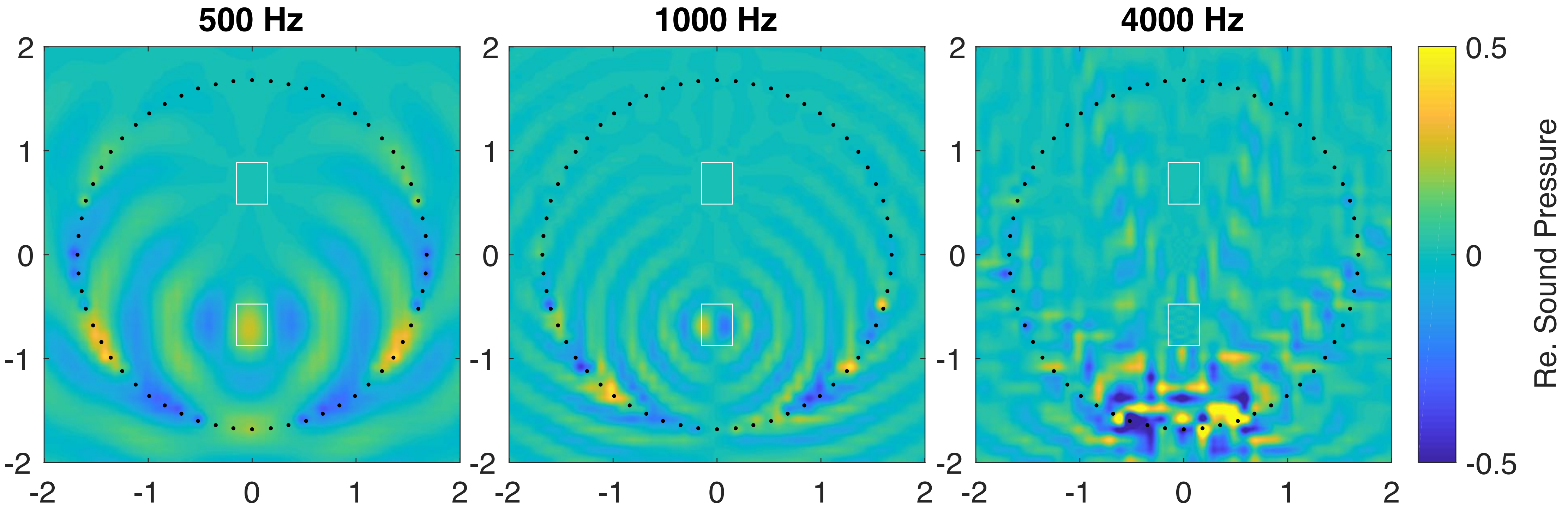
Acoustic Contrast Control: Circular Array

- Sound field: sound pressure level



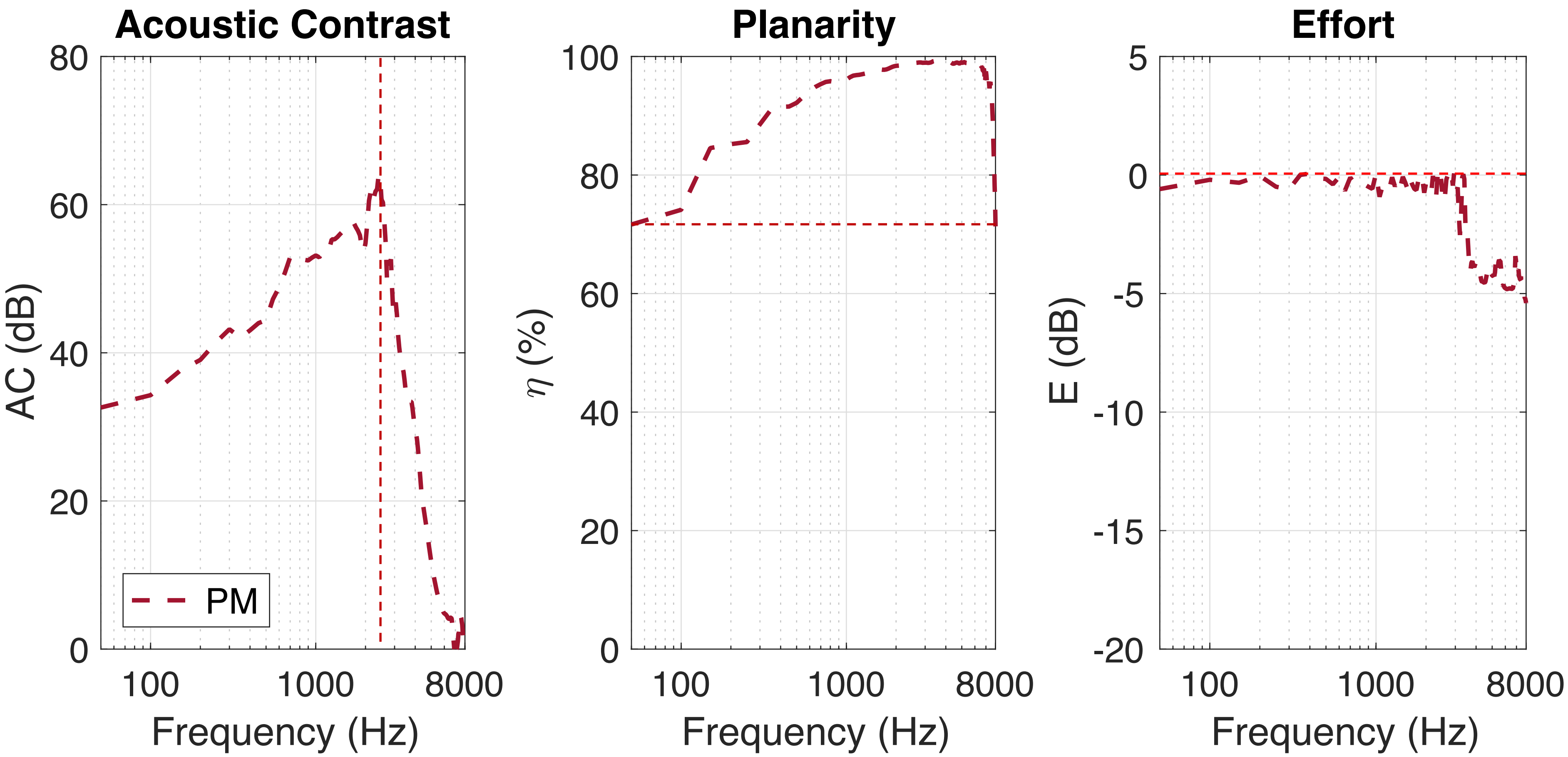
Acoustic Contrast Control: Circular Array

- Sound field: real part of the sound pressure



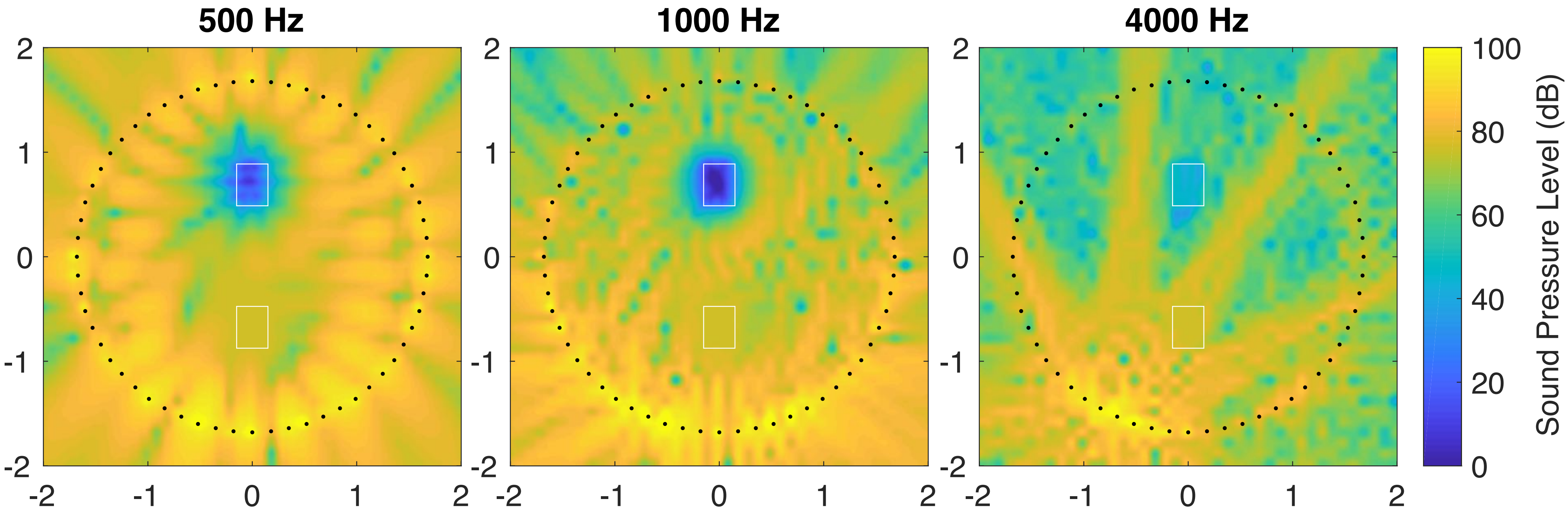
Pressure Matching: Circular Array

- Performance over frequency



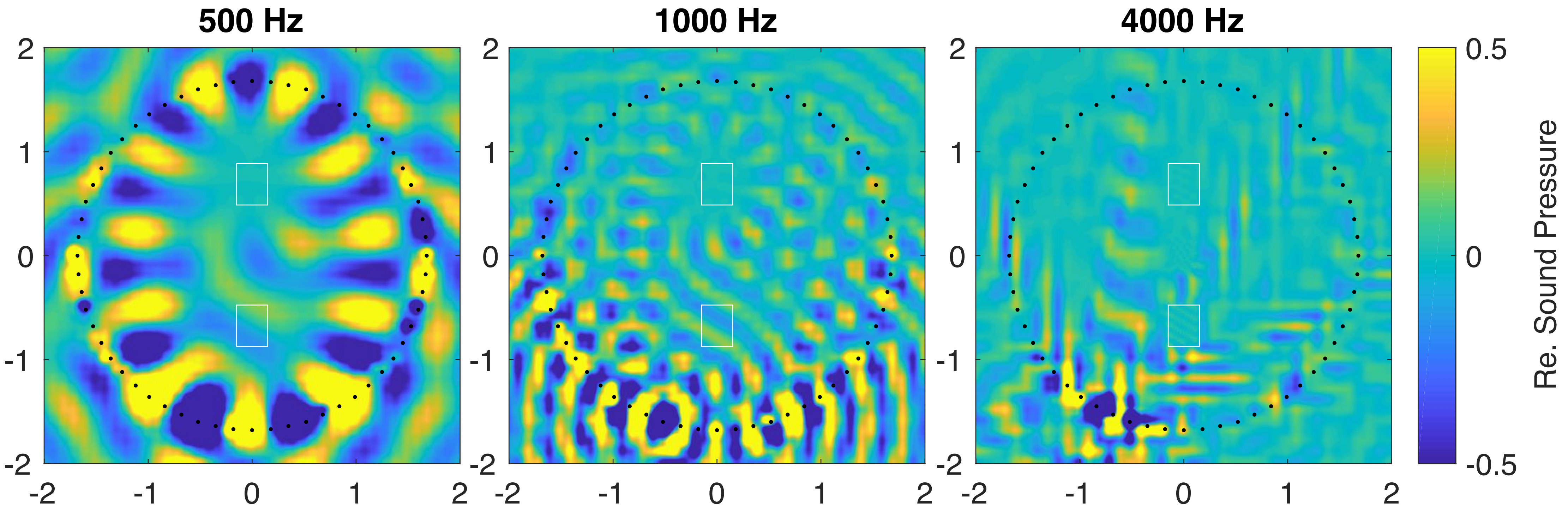
Pressure Matching: Circular Array

- Sound field: sound pressure level



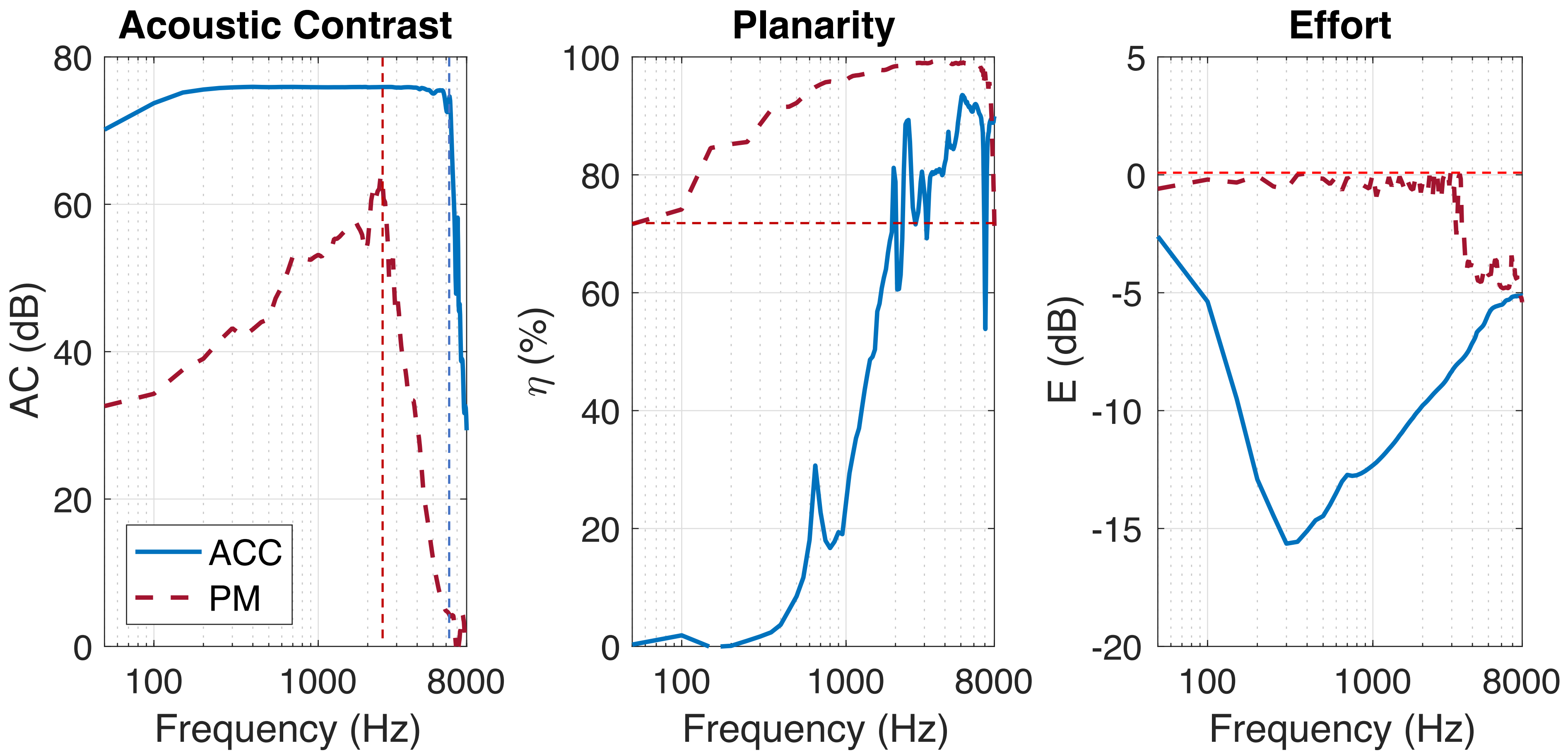
Pressure Matching: Circular Array

- Sound field: real part of the sound pressure



Performance Comparison: Circular Array

- Control of energy vs control of phase

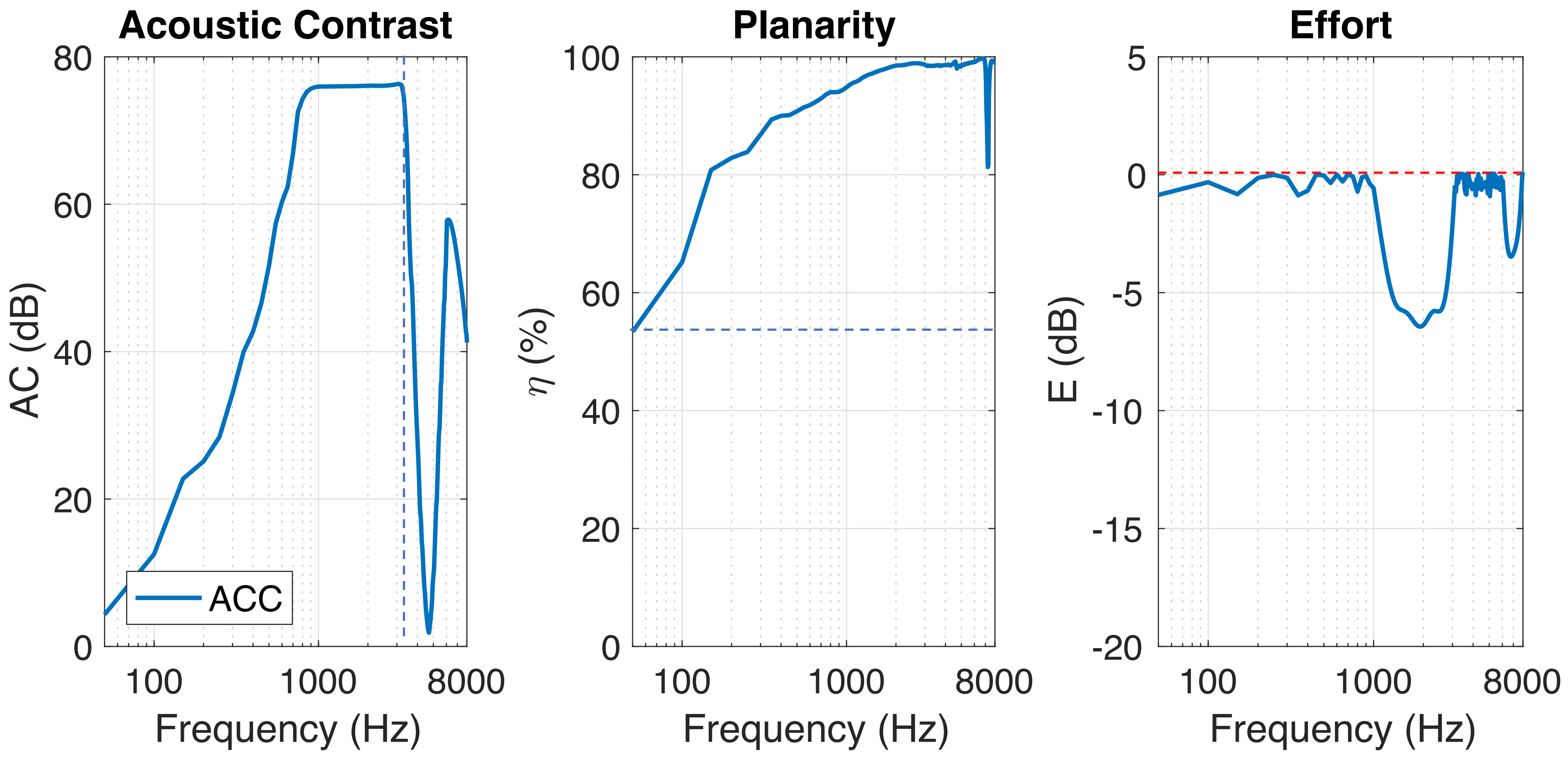


Summary: Circular Array

- ACC controls energy only; PM controls energy and phase
  - ACC gives excellent contrast
  - ACC controls contrast over a wide frequency range
  - PM gives excellent control of sound in the bright zone (homogenous, planar)
  - PM requires higher effort filters to achieve this control

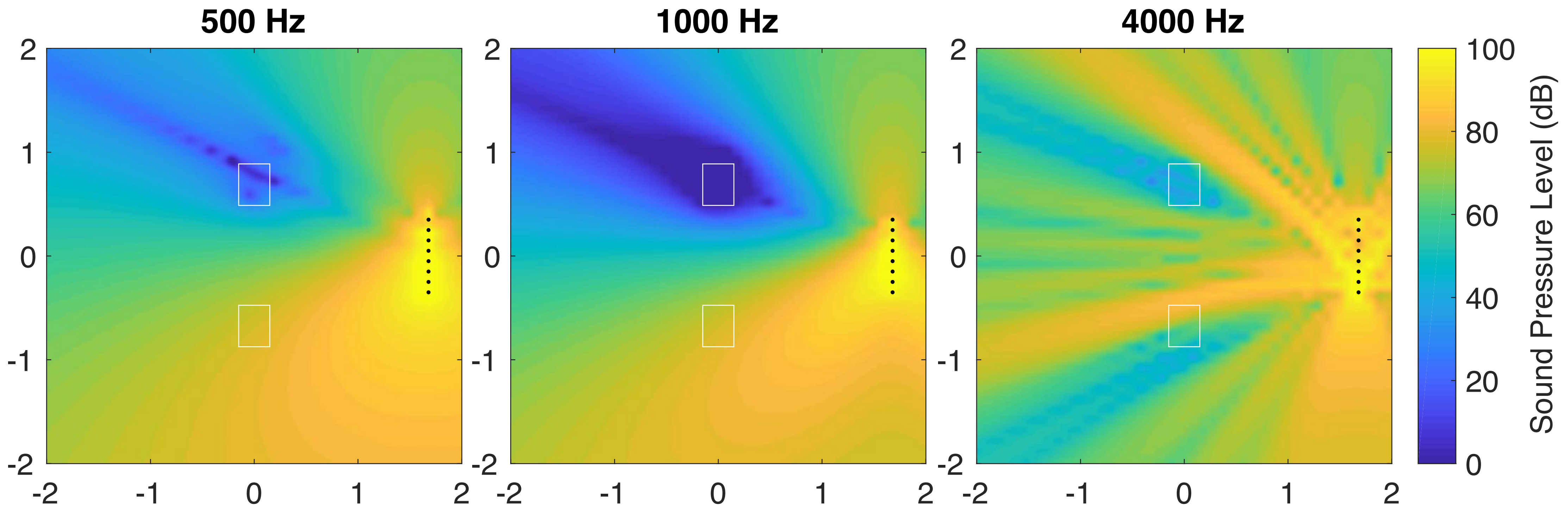
Acoustic Contrast Control: Line Array

- Performance over frequency



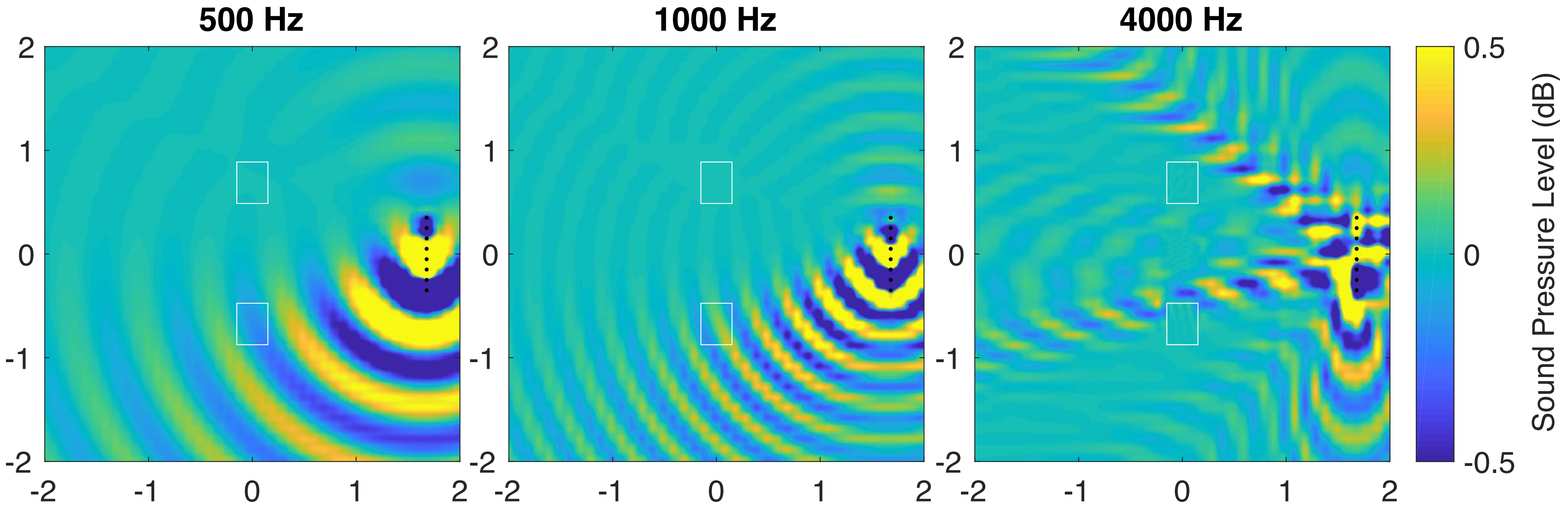
Acoustic Contrast Control: Line Array

- Sound field: sound pressure level



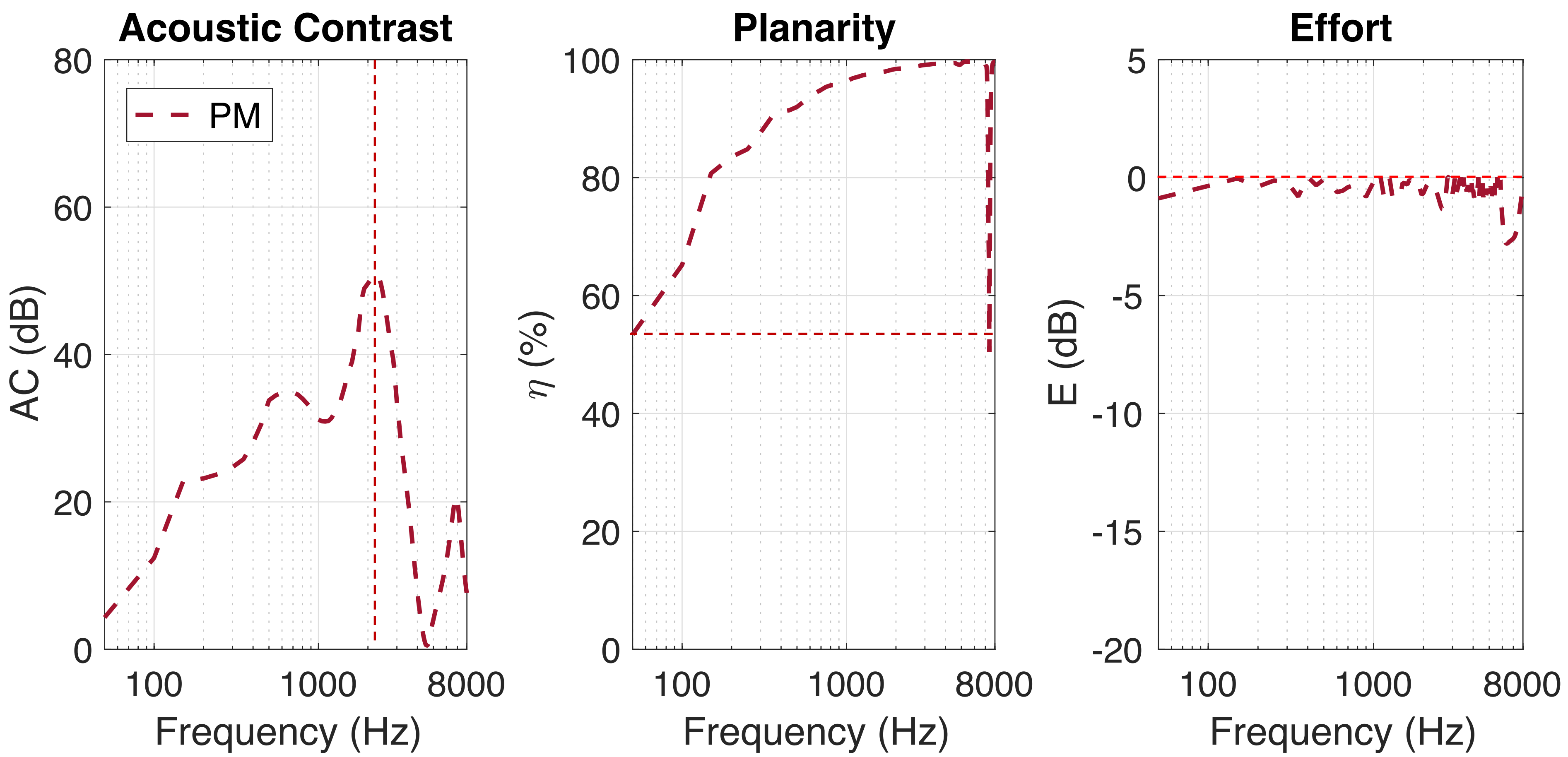
Acoustic Contrast Control: Line Array

- Sound field: real part of the sound pressure



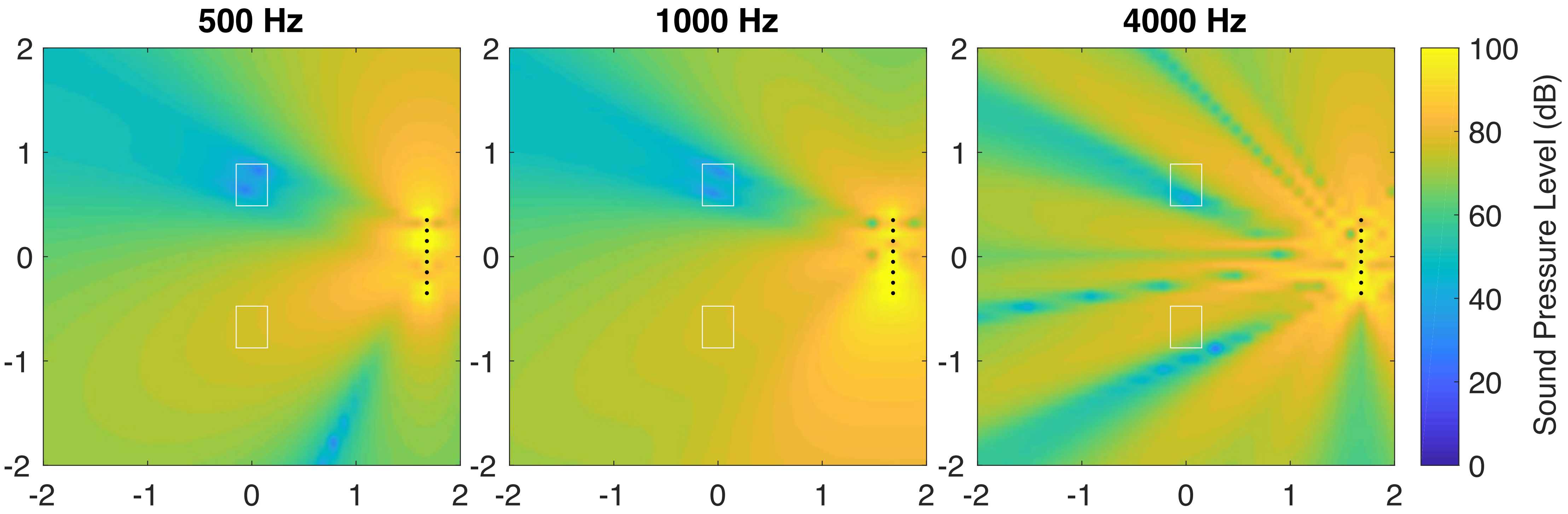
Pressure Matching: Line Array

- Performance over frequency



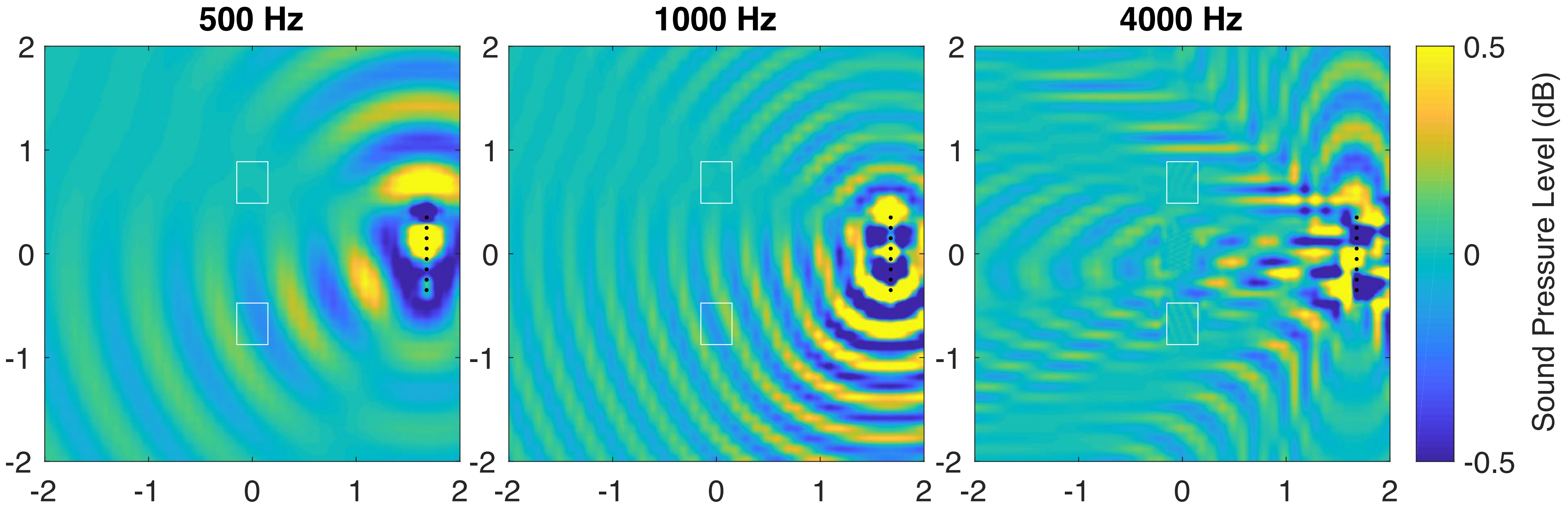
Pressure Matching: Line Array

- Sound field: sound pressure level



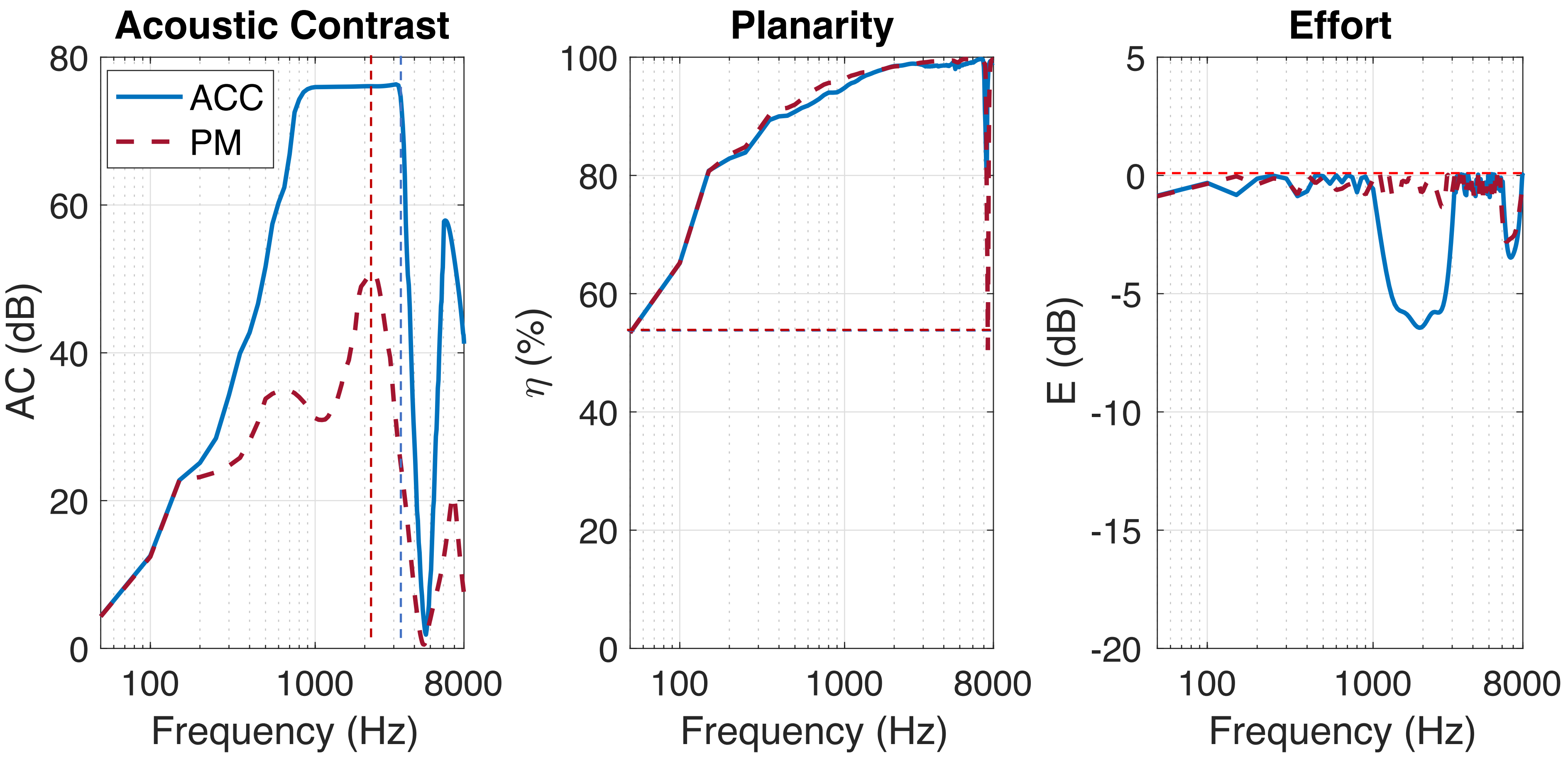
Pressure Matching: Line Array

- Sound field: real part of the sound pressure



Performance Comparison: Line Array

- Control of energy vs control of phase (within array capability)

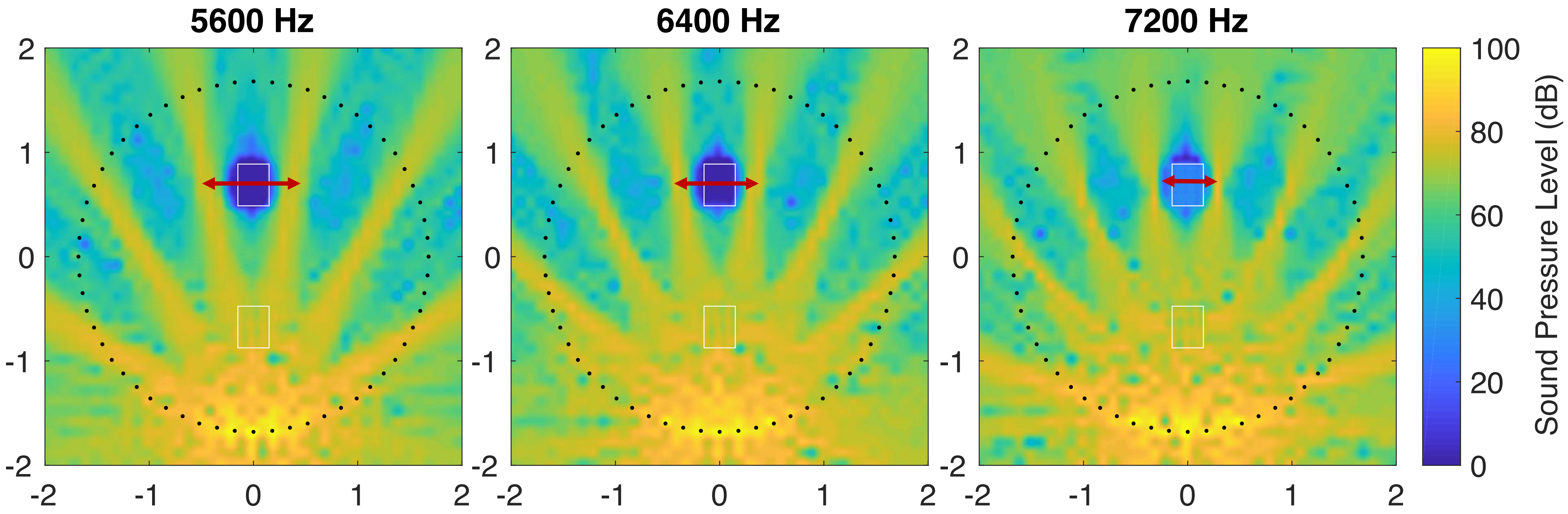


Summary: Line Array

- ACC controls energy only; PM controls energy and phase
  - ACC gives excellent contrast
    - But only reaches the maximum contrast over a relatively narrow frequency range
  - ACC controls contrast over a wide frequency range
    - Extra bandwidth compared to PM not as significant as for circular array
    - Low frequency contrast performance the same for both methods
  - PM gives excellent control of sound in the bright zone (homogenous, planar)
    - But ACC still planar due to limited possible DOAs into zone
  - PM requires higher effort filters to achieve this control
    - But ACC also high effort at most frequencies

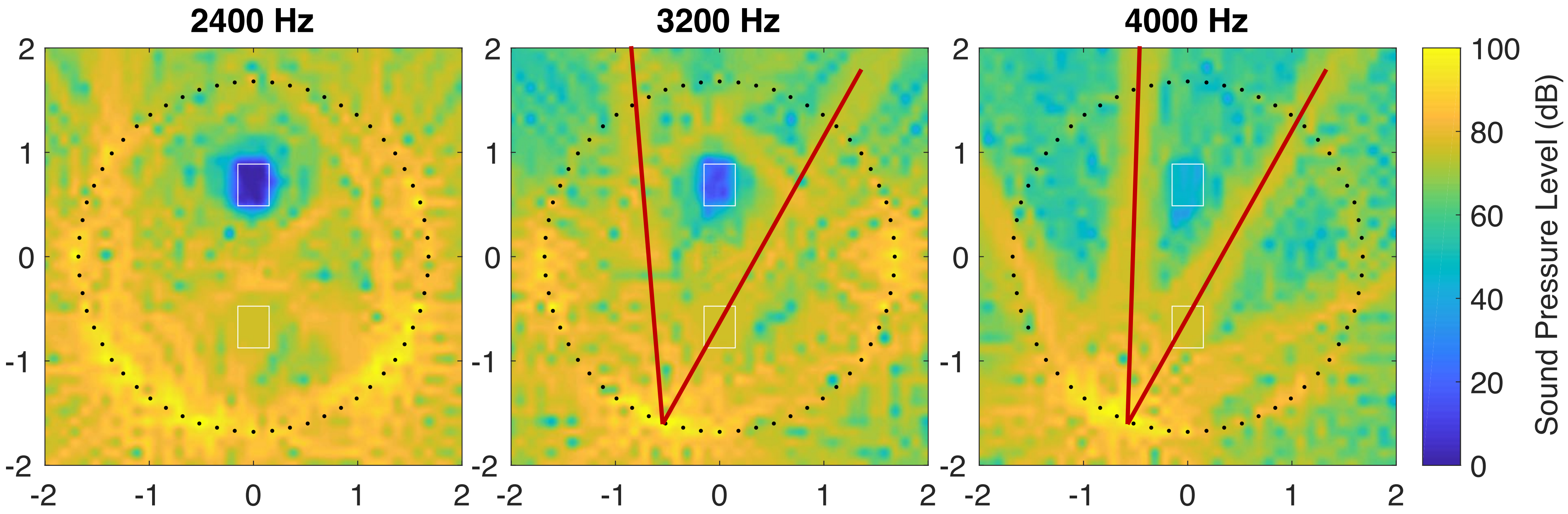
Loudspeaker Array Design

- Circular array degradation of acoustic contrast performance
  - ACC: onset when aliasing lobe separation equals dark zone diameter



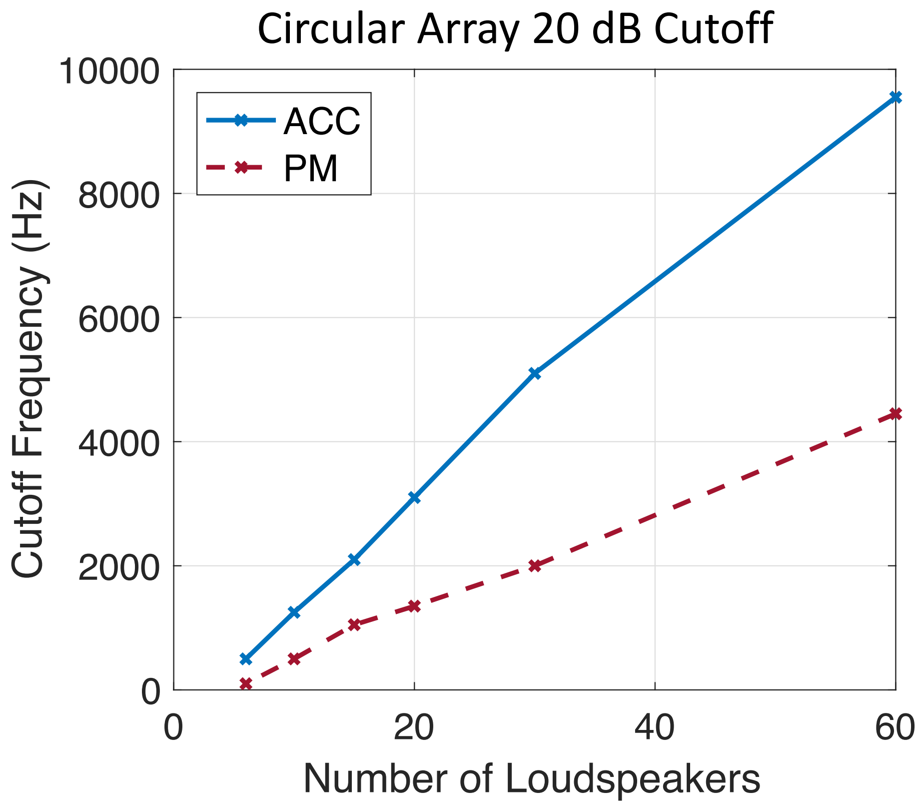
Loudspeaker Array Design

- Circular array degradation of acoustic contrast performance
  - PM: onset when aliasing lobe begins to impinge on dark zone



Loudspeaker Array Design

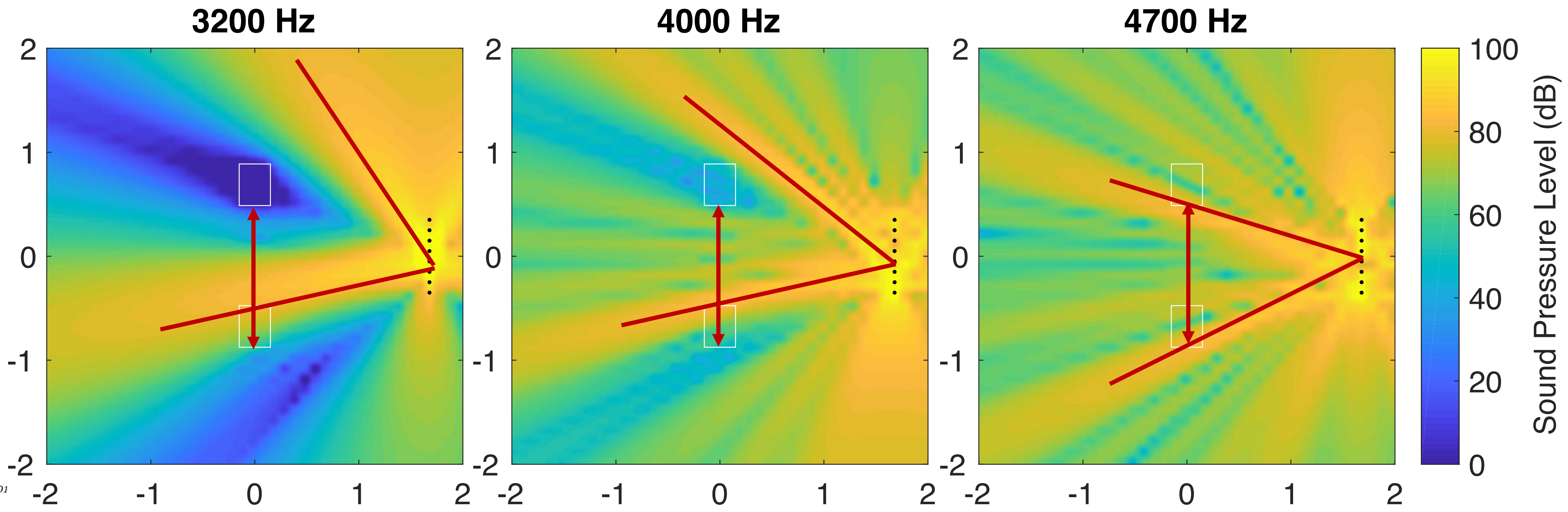
- Circular array:
  - ACC moves target beam direction to avoid aliasing for as long as possible
    - Aliasing depends on dark zone size
  - PM target direction specified in optimization
    - Aliasing depends on loudspeaker separation and overall control region size
  - To compare, consider the frequency at which contrast falls below 20 dB
- Recent work by Winter et al. to predict spatial aliasing for SFS



Coleman P., 2014. “Loudspeaker Array Processing for Personal Sound Zone Reproduction”, PhD thesis, Univ. of Surrey, UK
Winter, F., et al., 2019. A Geometric Model for Prediction of Spatial Aliasing in 2.5 D Sound Field Synthesis, TASLP.

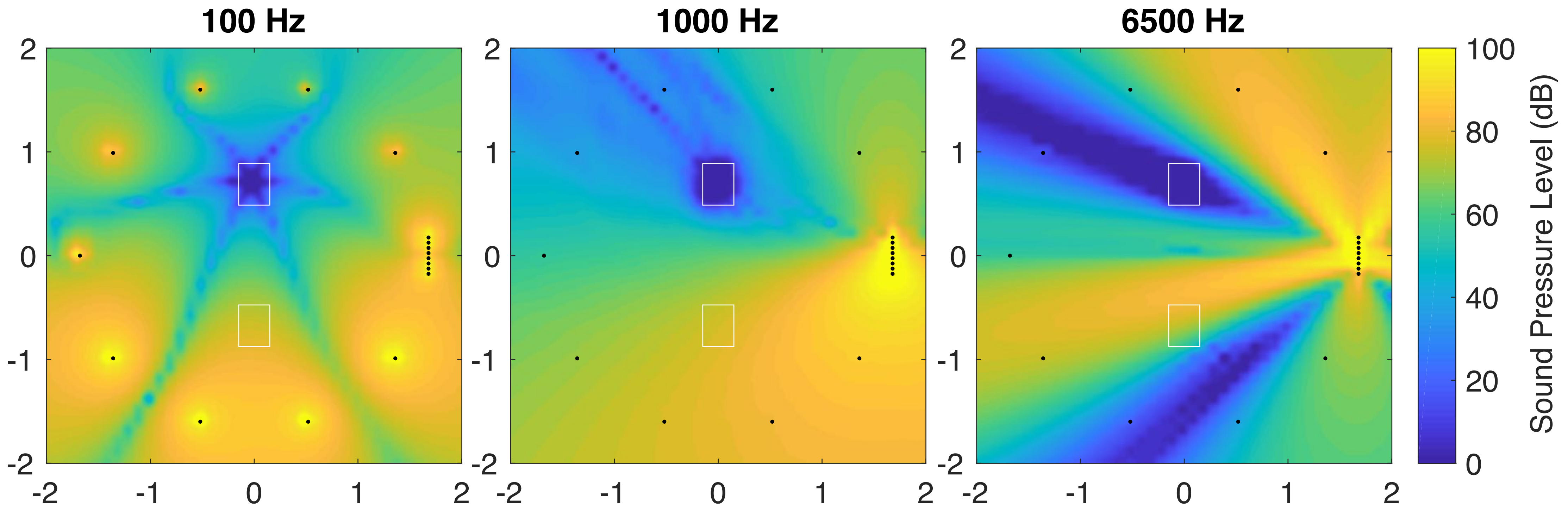
Loudspeaker Array Design

- Line array relationships with wavelength are well known:
  - Array aperture and regularization limit low frequency performance
  - Loudspeaker and zone spacing limit high frequency performance



Loudspeaker Array Design

- So we need a large aperture and close spacing 🤔



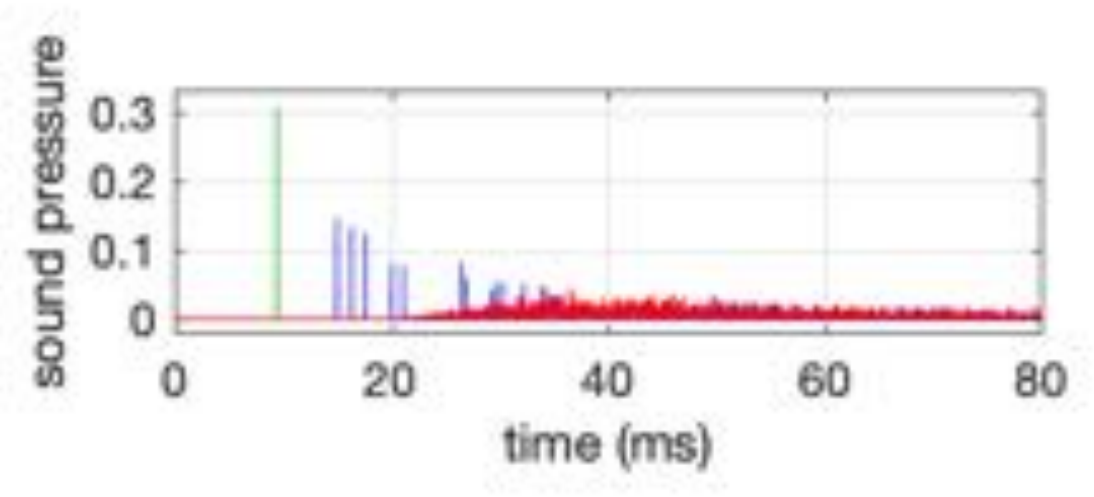
Summary: Performance Comparison

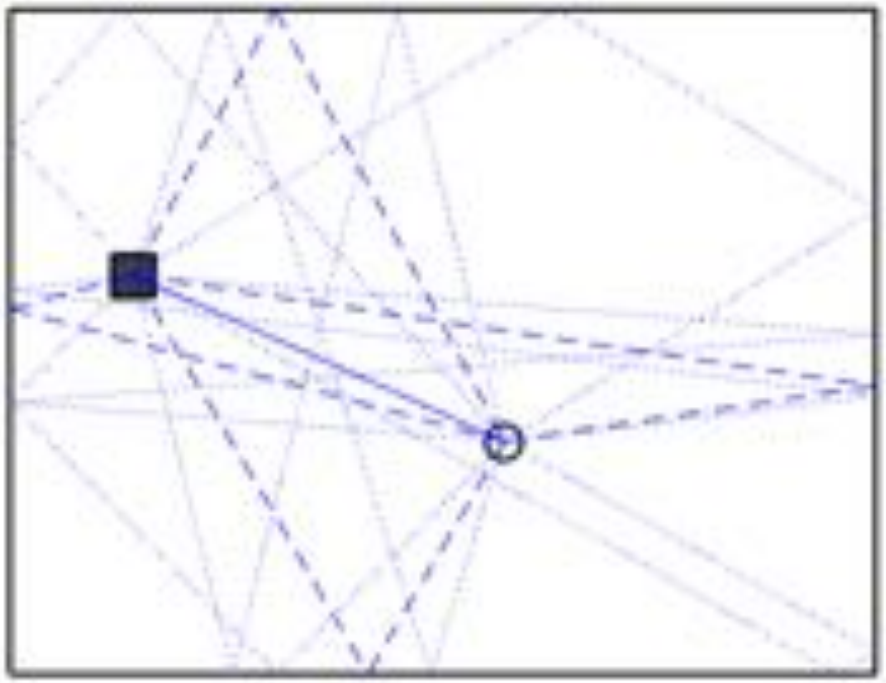
- ACC: creates maximal contrast between the zones, at the cost of having no control over the phase
  - High contrast
  - Low planarity (circular array)
  - Relatively efficient
- PM: aims to control the pressure amplitude and phase at each control microphone
  - Lower contrast
  - High planarity
  - Relatively inefficient
  - Weighted versions of PM have also been proposed
- Array geometry imposes fundamental limits on performance

Dealing with room reverberation

Structure of a room impulse response

- Direct sound
  - First, clean and narrow
  - From source direction
- Early reflections
  - Initially sparse, thickening
  - Absorbed and scattered
  - From all walls, up/down
- Late reverberation
  - Dense and diffuse
  - Directionless after mixing time

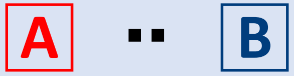




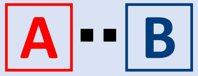
Kuttruff, 2009. Room Acoustics

Ironing out the earlies

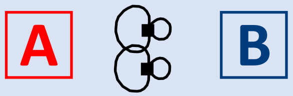
- Early reflections (coherent) colour the response
- Reflections add unwanted copies of loudspeakers as image sources



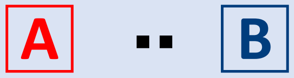
Reduced control effort [Elliott et al., 2012, TASLP] or acoustic power [Jones & Elliott, 2008, JASA]



Array close to zones [Chang et al., 2009, JASA]



Directive sources [Simón Gálvez et al., 2012, JASA]



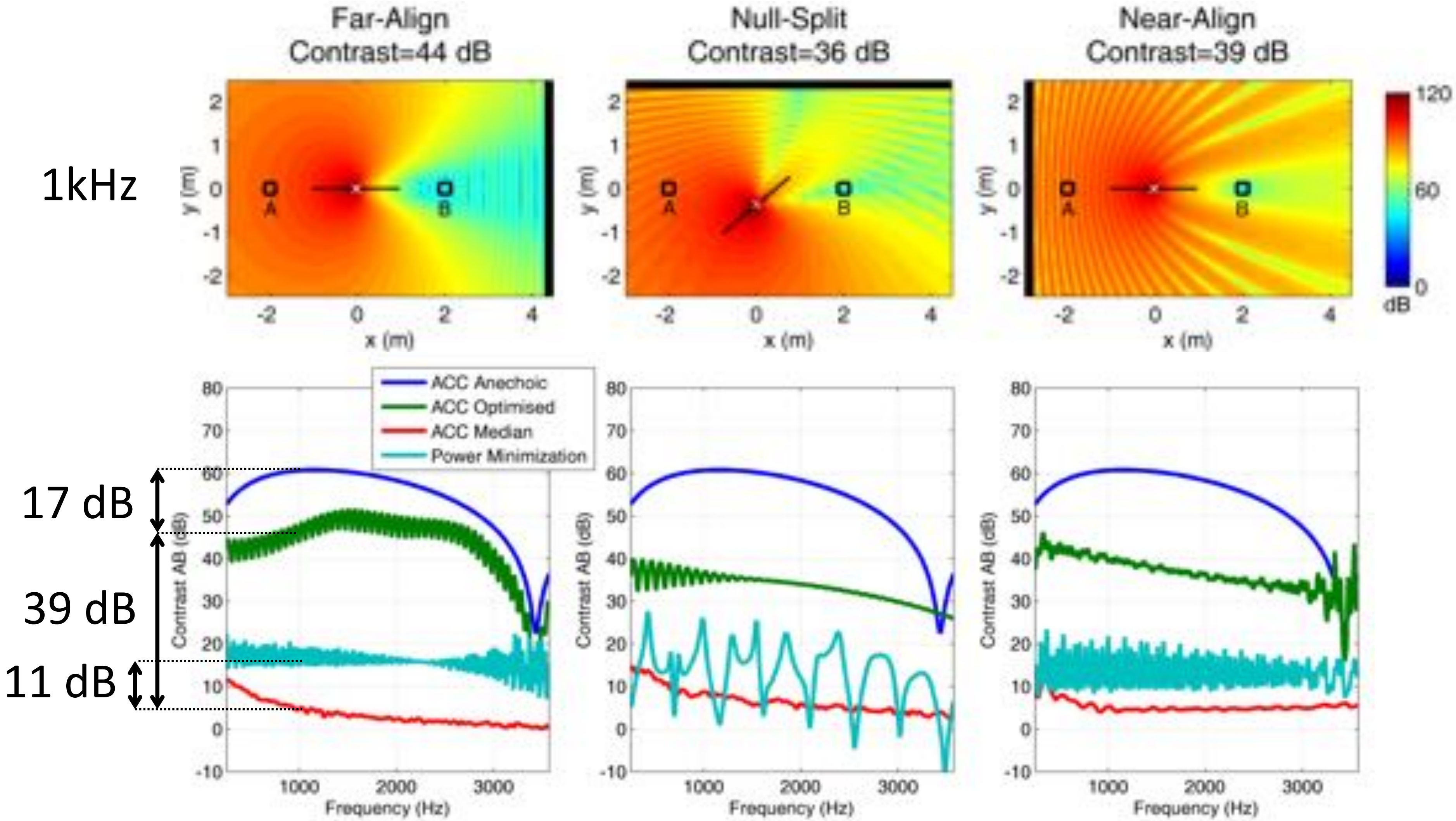
Superdirective beamforming [Elliott et al., 2010, JASA]

- For strong low order reflections, it may be beneficial to
  - Reduce radiation in specific directions
  - Maximize effectiveness of room equalization for specific reflection paths
- This can be achieved by geometrical optimization of source positions.

Olik, 2014. Personal sound zone reproduction with room reflections. PhD thesis

Geometrical considerations

- Arrangement of loudspeakers in relation to reflectors
- An unwanted reflection can be avoided, averted or cancelled
- Higher-order reflections use source modes



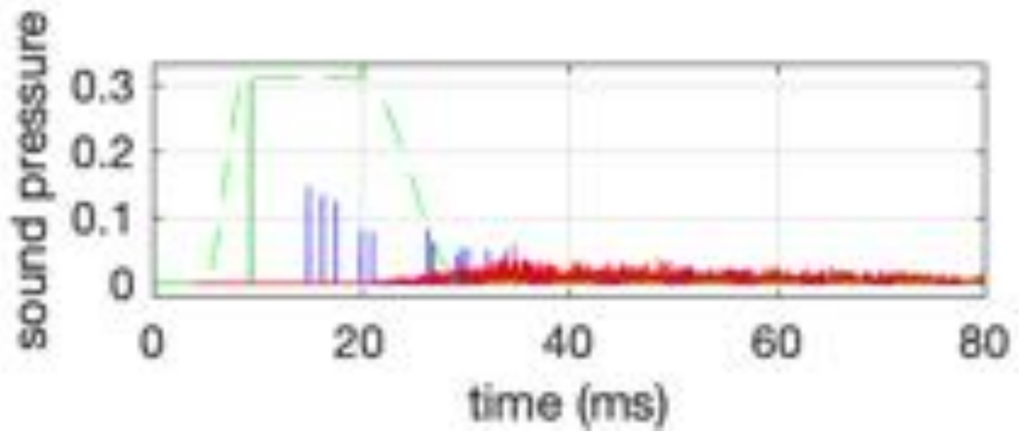
Olik et al., 2014. Optimum source geometry for sound zone reproduction with a single reflection, JASA

Late reverberation

There are two main strategies to mitigate effects of late reverberation:

- Truncation during calibration

  By shortening the calibration RIRs used to calculate the personal sound zone filters



- Suppression through regularization

  By applying regularization to make the filters less sensitive to convolutive noise during playback

Olik et al., 2013. A comparative performance study of sound zoning methods in a reflective..., AES

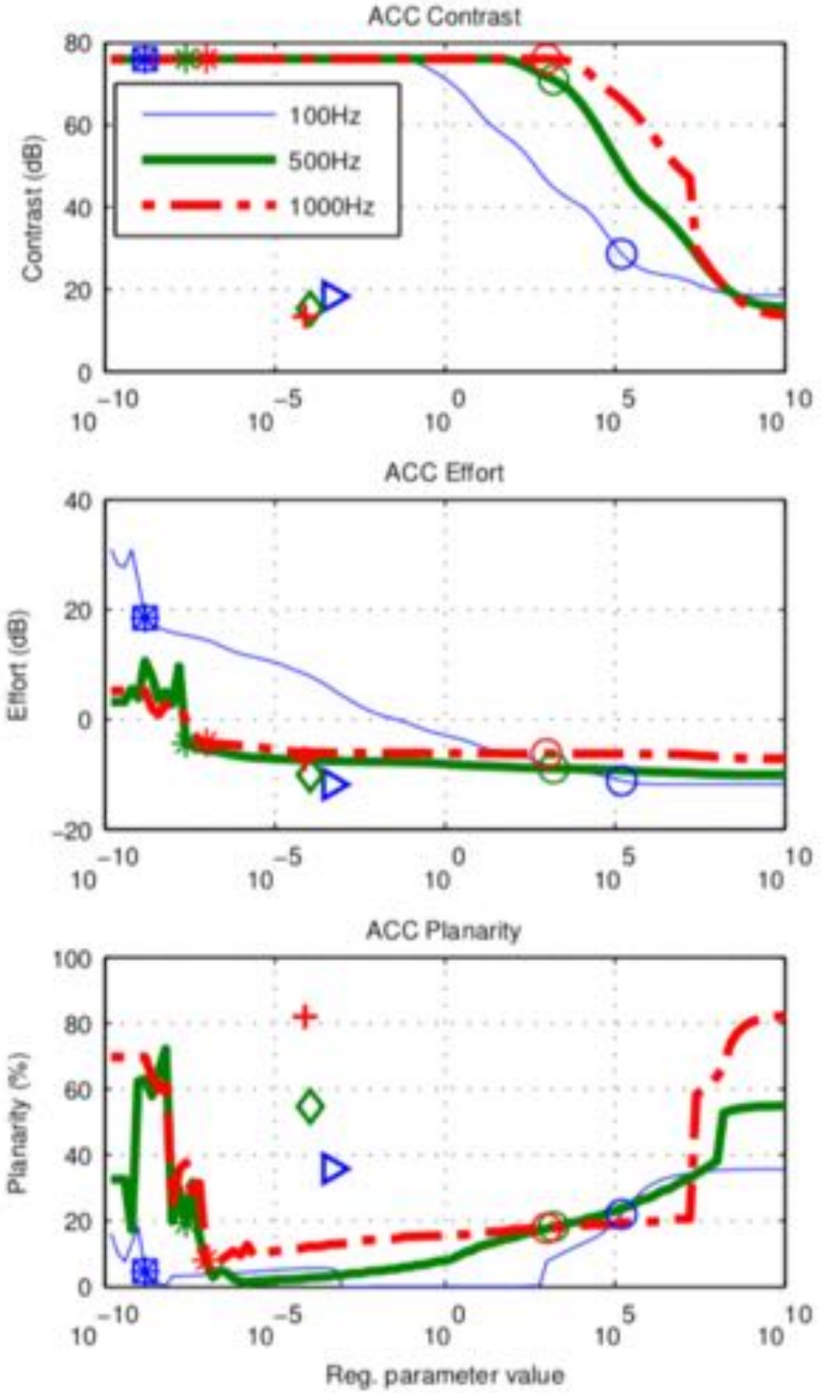
Room effect summary

- Acoustic environments add early reflections and late reverberation
- Left untreated, effects can severely degrade system performance
- Mitigation strategies can be deployed for partial recovery:
  1. Acoustically treat/select the room
  2. Arrange your setup to avoid, avert or cancel reflections
  3. Truncate reverberation from the RIRs in calibration
  4. Apply regularization wisely…

Fundamental Relationships under Non-Ideal Conditions: Robustness and Regularization

Regularization: Fundamentals

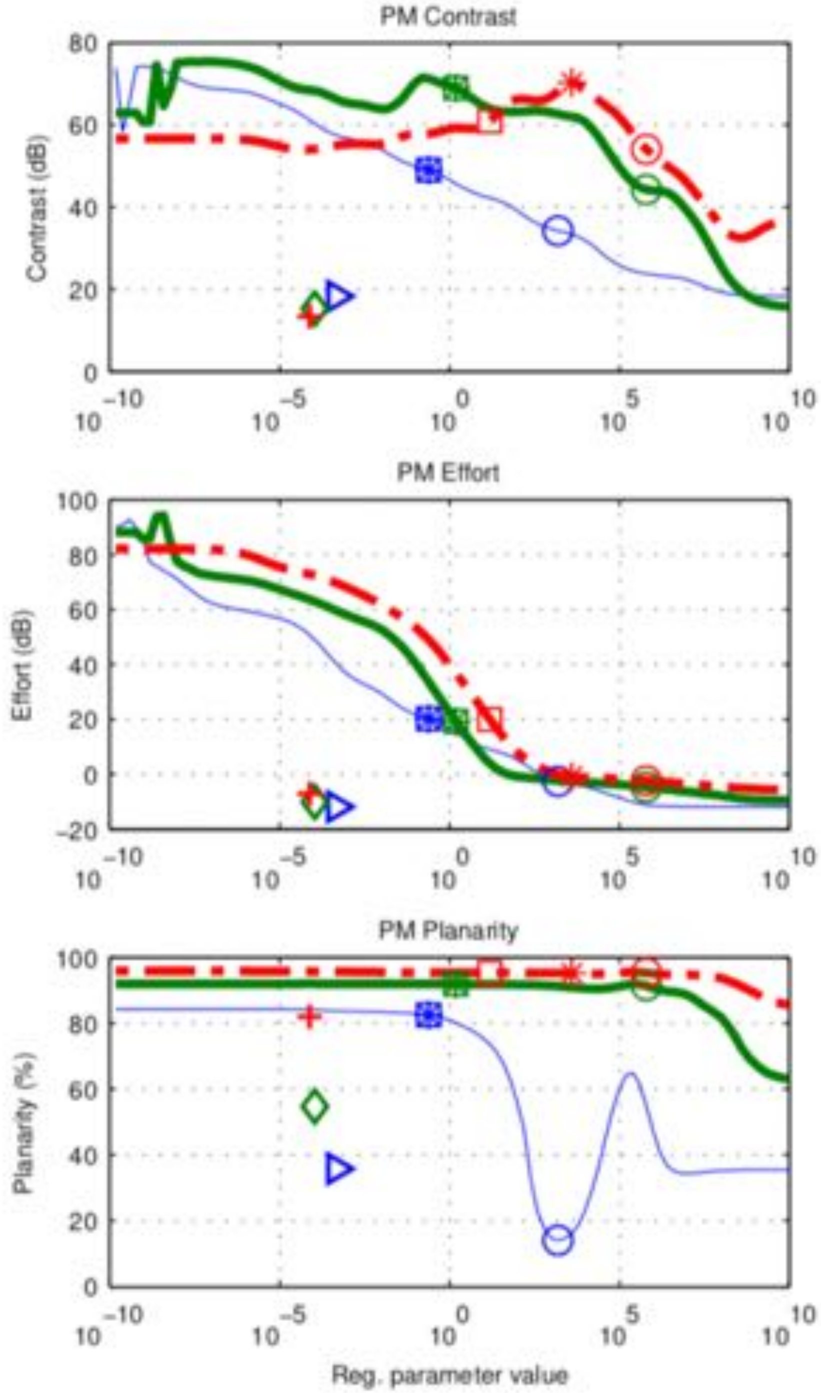
- Regularization parameter vs performance (ideal conditions)
- Acoustic Contrast Control:
  - **More regularizaton = lower effort**
  - Contrast and planarity fairly robust unless over-regularized



Coleman, P., et al., 2013, “The influence of regularization on anechoic performance and robustness of sound zone methods” *Proc. ICA2013*.

Regularization: Fundamentals

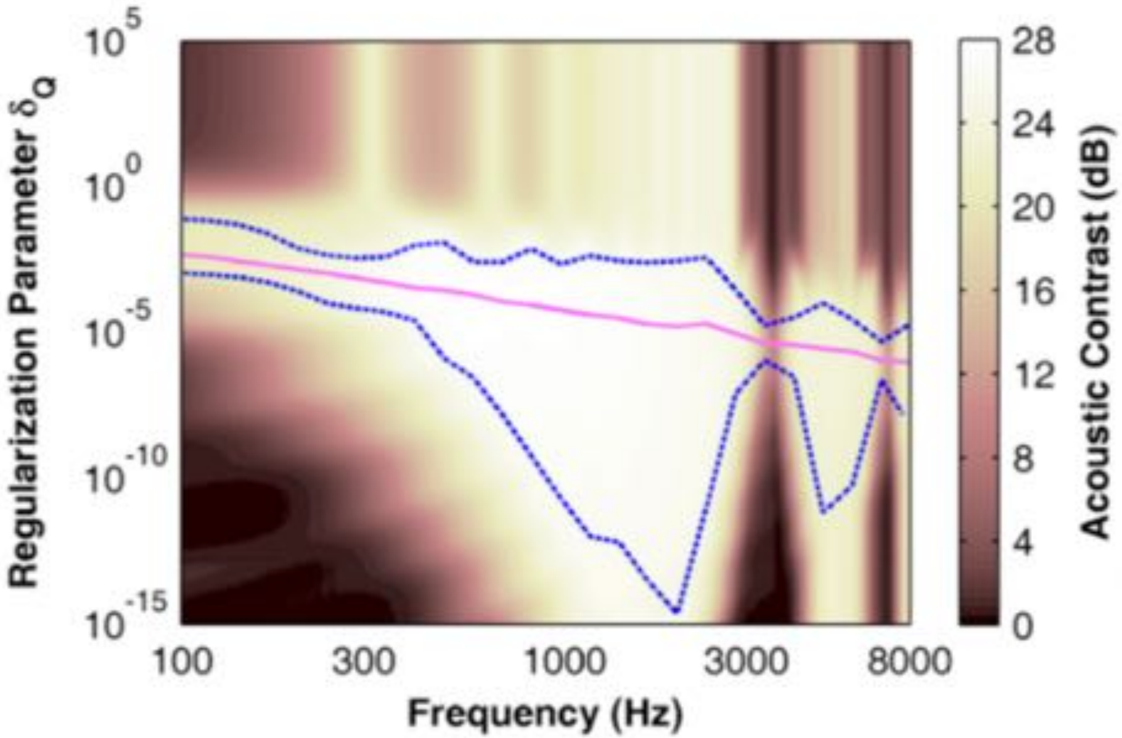
- Regularization parameter vs performance (ideal conditions)
- Pressure Matching:
  - **More regularizaton = lower effort**
  - Planarity fairly robust unless over-regularized
  - Some local optima in contrast



Coleman, P., et al., 2013, "The influence of regularization on anechoic performance and robustness of sound zone methods" *Proc. ICA2013*.

Regularization: Non-ideal Conditions

- Design challenge: choose optimal regularizaton parameter
- Some frequencies are more sensitive than others



Zhu, Q, et al., 2017, "Robust Acoustic Contrast Control with Reduced In-situ Measurement by Acoustic Modelling". JAES.

Regularization: Non-ideal Conditions

- Design challenge: choose optimal regularization parameter

	Approach	Example references
Matrix Perspective	Choose diagonal loading regularization parameter based on condition number / singular values of matrix	Shin, M., et al., 2014. "Controlled sound field with a dual layer loudspeaker array". *J. Sound Vib.*
Array Effort Perspective	Choose diagonal loading regularization parameter to meet a maximum control effort constraint	Elliott, S.J., et al., 2012. "Robustness and regularization of personal audio systems". *TASLP.* Coleman, P., et al., 2014. "Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array". *JASA.*
Error Modelling Perspective	Choose regularization parameters (matrix) based on predicted system errors	Doclo, S. & Moonen, M., 2003. "Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics". *TSP.* Zhu, Q, et al., 2017, "Robust reproduction of sound zones with local sound orientation". JASA-EL. Zhu, Q, et al., 2017, "Robust Acoustic Contrast Control with Reduced In-situ Measurement by Acoustic Modelling". JAES.

Regularization: Non-ideal Conditions

- Design challenge: choose optimal regularization parameter

	Approach	Example references
Matrix Perspective	Choose diagonal loading regularization parameter based on condition number / singular values of matrix	Shin, M., et al., 2014. "Controlled sound field with a dual layer loudspeaker array". *J. Sound Vib.*
Array Effort Perspective	Choose diagonal loading regularization parameter to meet a maximum control effort constraint	Elliott, S.J., et al., 2012. "Robustness and regularization of personal audio systems". *TASLP*. Coleman, P., et al., 2014. "Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array". *JASA*.
Error Modelling Perspective	Choose regularization parameters (matrix) based on predicted system errors	Doclo, S. & Moonen, M., 2003. "Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics". *TSP*. Zhu, Q, et al., 2017, "Robust reproduction of sound zones with local sound orientation". JASA-EL. Zhu, Q, et al., 2017, "Robust Acoustic Contrast Control with Reduced In-situ Measurement by Acoustic Modelling". JAES.

ICASSP 2019

Regularization: Non-ideal Conditions

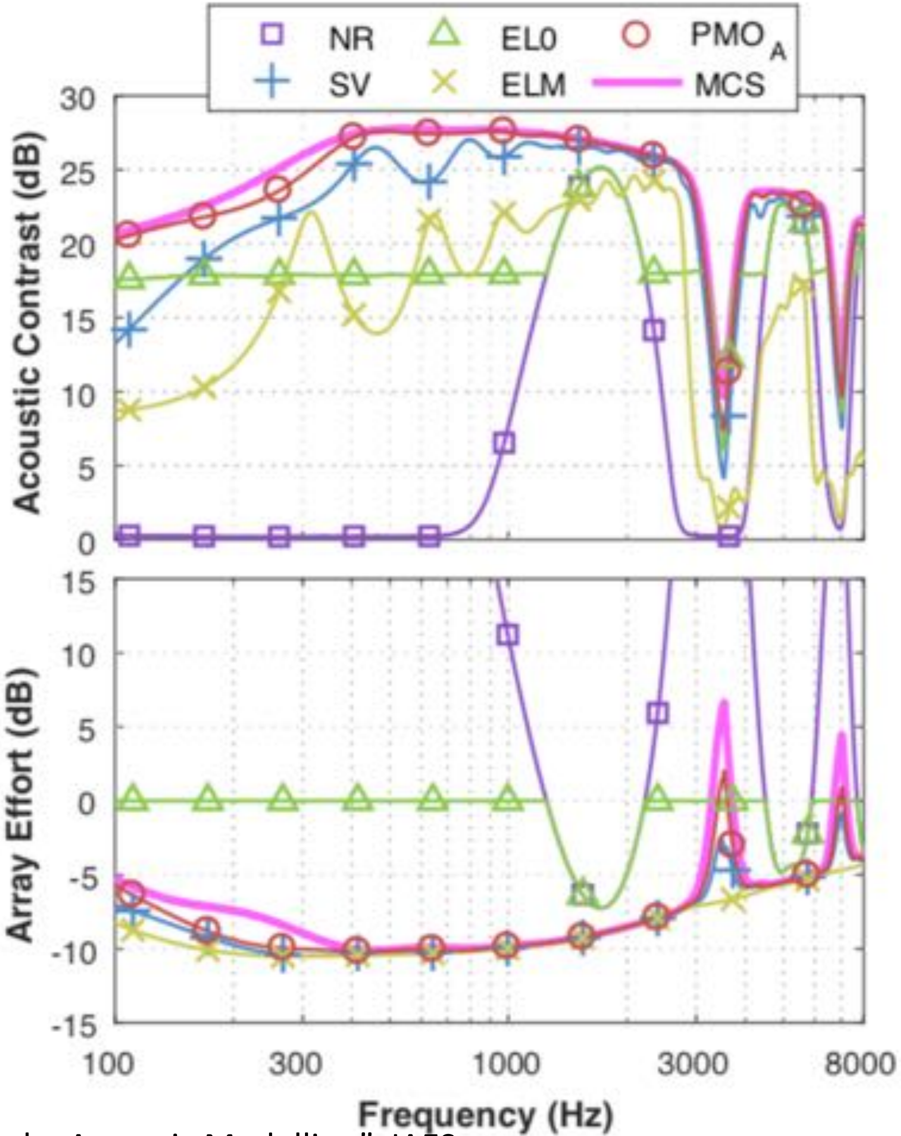
- Design challenge: choose optimal regularization parameter

	Approach	Example references
Matrix Perspective	Choose diagonal loading regularization parameter based on condition number / singular values of matrix	Shin, M., et al., 2014. “Controlled sound field with a dual layer loudspeaker array”. *J. Sound Vib.*
Array Effort Perspective	Choose diagonal loading regularization parameter to meet a maximum control effort constraint	Elliott, S.J., et al., 2012. “Robustness and regularization of personal audio systems”. *TASLP*. Coleman, P., et al., 2014. “Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array”. *JASA*.
Error Modelling Perspective	Choose regularization parameters (matrix) based on predicted system errors	Doclo, S. & Moonen, M., 2003. “Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics”. *TSP*. Zhu, Q, et al., 2017, ”Robust reproduction of sound zones with local sound orientation”. JASA-EL. Zhu, Q, et al., 2017, ”Robust Acoustic Contrast Control with Reduced In-situ Measurement by Acoustic Modelling”. JAES.

Regularization: Non-ideal Conditions

- Optimal regularizaton parameter?

	Approach
Matrix Perspective + SV	Choose diagonal loading regularizaton parameter based on condition number / singular values of matrix
Array Effort Perspective △ EL0 × ELM	Choose diagonal loading regularizaton parameter to meet a maximum control effort constraint
Error Modelling Perspective ○ PMO_A	Choose regularizaton parameters (matrix) based on predicted system errors

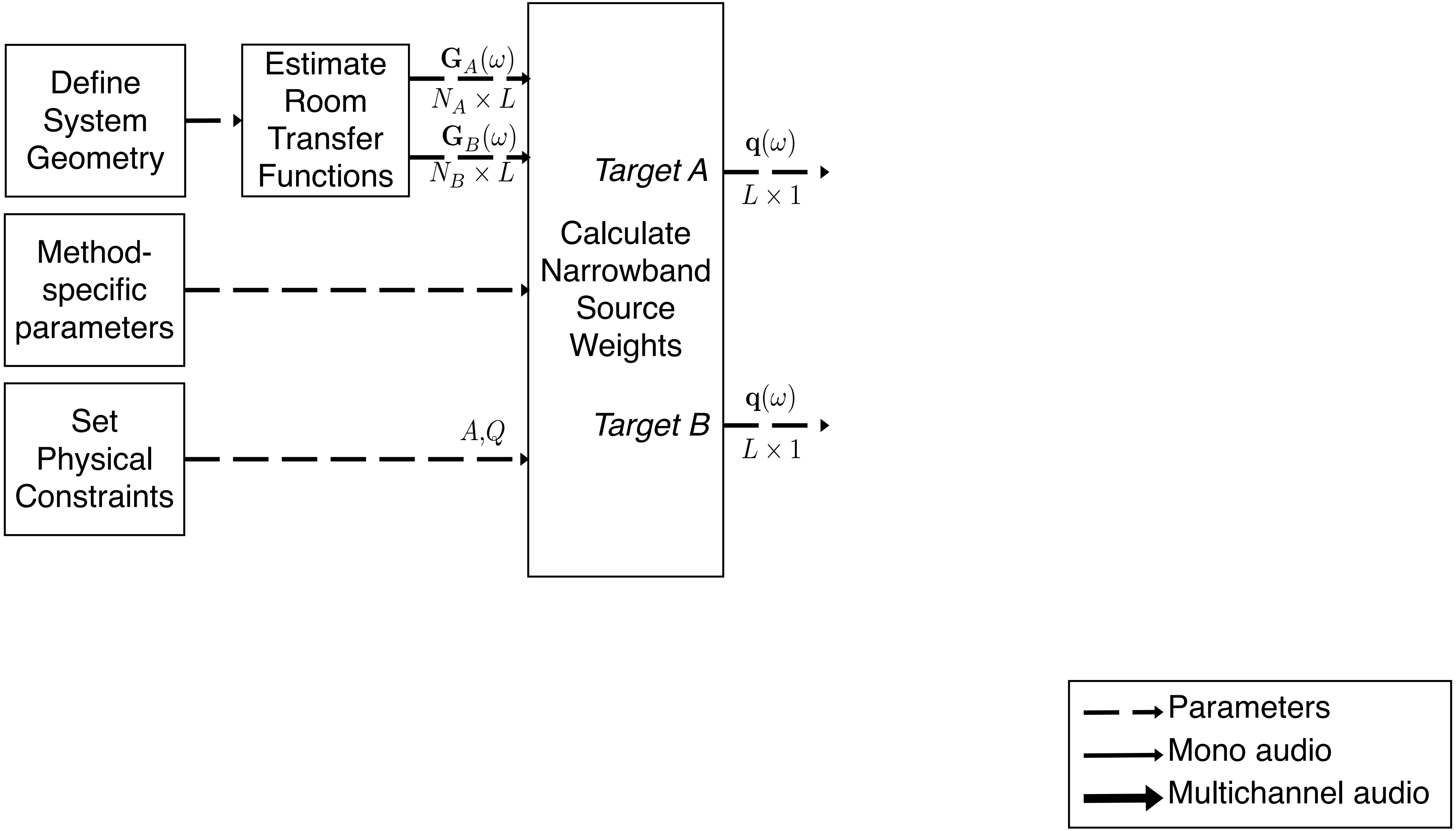


Zhu, Q, et al., 2017, ”Robust Acoustic Contrast Control with Reduced In-situ Measurement by Acoustic Modelling”. JAES.

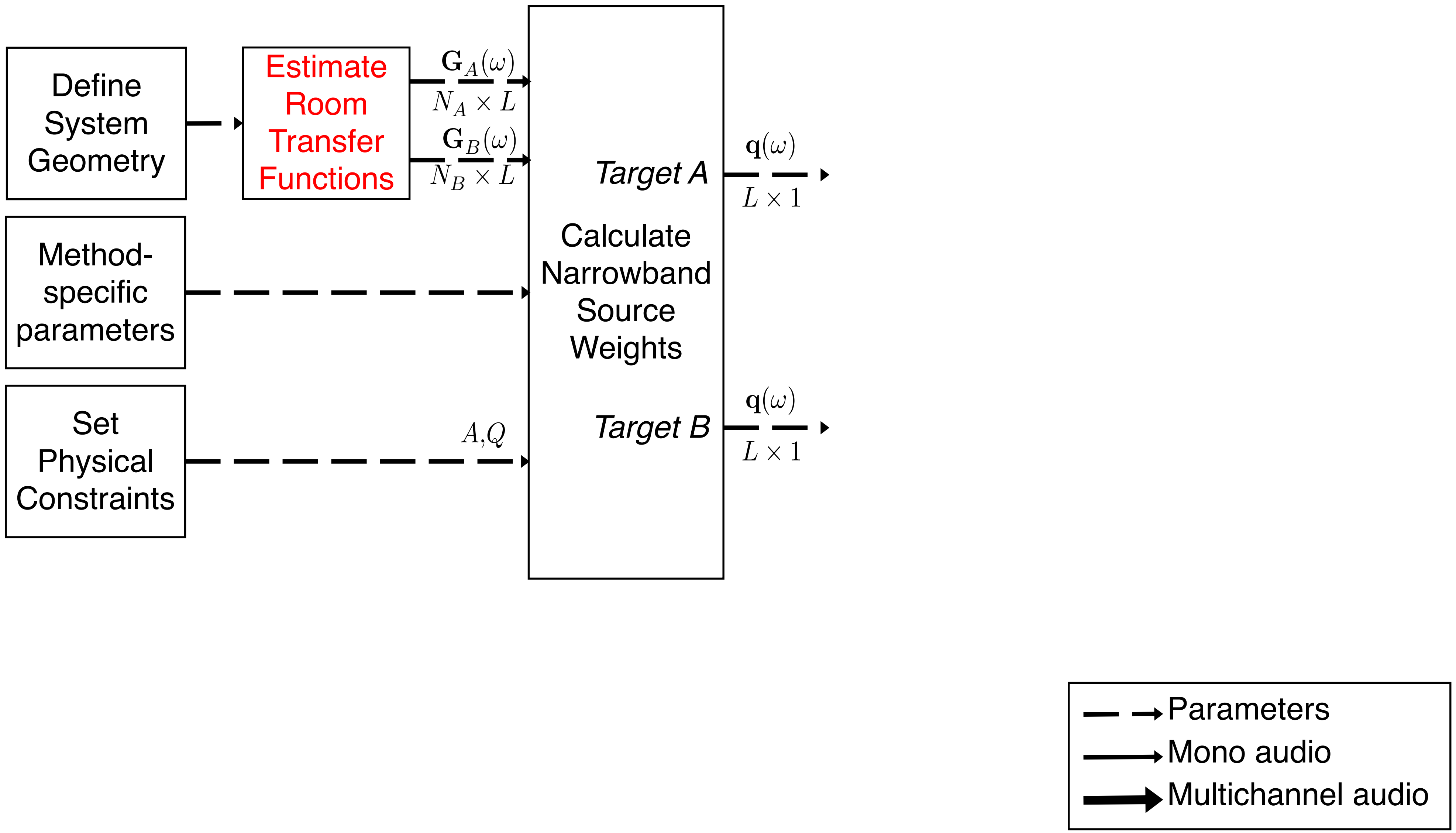
Regularization: Summary

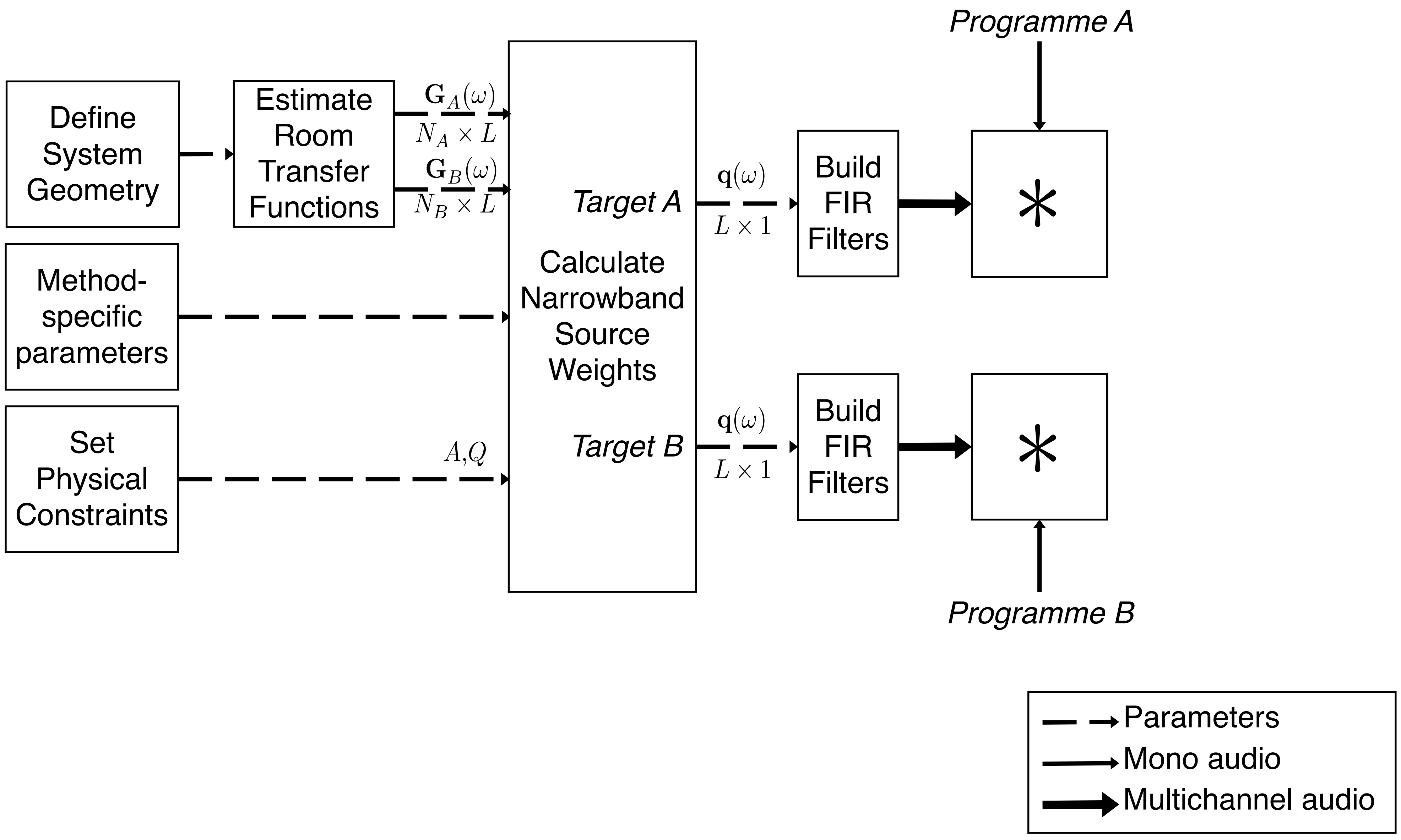
- Under real-world conditions, regularization is critical:
  - Too little regularization: not robust
  - Too much regularization: no contrast
- Three representative ways to find parameter(s) introduced:
  - Look at the matrix to be inverted
  - Fix the array effort
  - Consider the predicted kind/magnitude of errors in the whole system

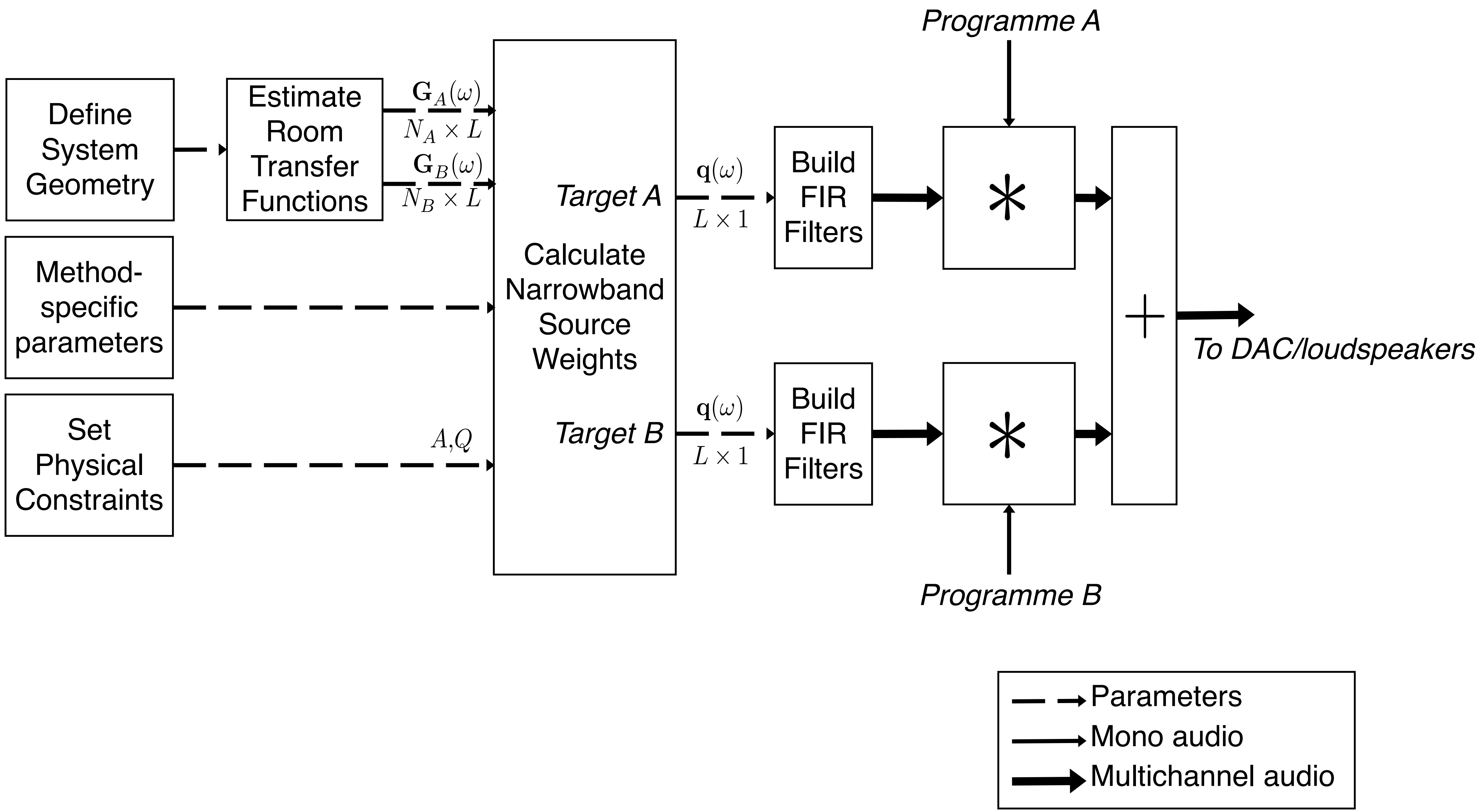
Making Sound Zone Filters

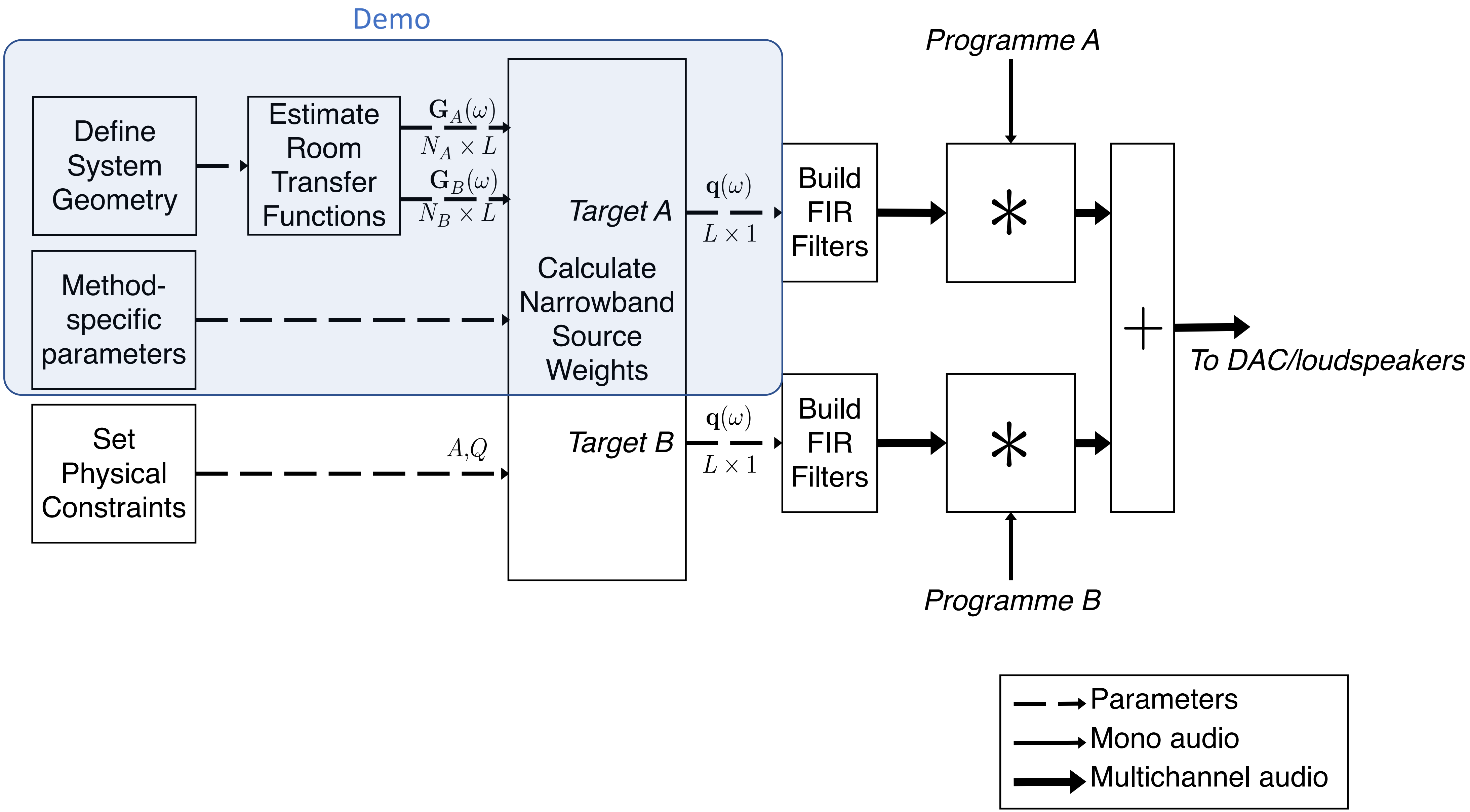


Define System Geometry
Estimate Room Transfer Functions
$\mathbf{G}_A(\omega)$
$N_A \times L$
$\mathbf{G}_B(\omega)$
$N_B \times L$
Method-specific parameters
Set Physical Constraints
A,Q
Target A
Calculate Narrowband Source Weights
Target B
$\mathbf{q}(\omega)$
$L \times 1$
$\mathbf{q}(\omega)$
$L \times 1$
Parameters
Mono audio
Multichannel audio


Define System Geometry
Estimate Room Transfer Functions
$\mathbf{G}_A(\omega)$
$N_A \times L$
$\mathbf{G}_B(\omega)$
$N_B \times L$
Method-specific parameters
Set Physical Constraints
A,Q
Target A
Calculate Narrowband Source Weights
Target B
$\mathbf{q}(\omega)$
$L \times 1$
$\mathbf{q}(\omega)$
$L \times 1$
Parameters
Mono audio
Multichannel audio


Programme A
Define System Geometry
Estimate Room Transfer Functions
$\mathbf{G}_A(\omega)$
$N_A \times L$
$\mathbf{G}_B(\omega)$
$N_B \times L$
Target A
Calculate Narrowband Source Weights
Target B
$\mathbf{q}(\omega)$
$L \times 1$
Build FIR Filters
*
Method-specific parameters
Set Physical Constraints
A,Q
$\mathbf{q}(\omega)$
$L \times 1$
Build FIR Filters
*
Programme B
Parameters
Mono audio
Multichannel audio
ICASSP 2019


Define System Geometry
Estimate Room Transfer Functions
$\mathbf{G}_A(\omega)$
$N_A \times L$
$\mathbf{G}_B(\omega)$
$N_B \times L$
Method-specific parameters
Set Physical Constraints
A,Q
Target A
Calculate Narrowband Source Weights
Target B
$\mathbf{q}(\omega)$
$L \times 1$
$\mathbf{q}(\omega)$
$L \times 1$
Build FIR Filters
Build FIR Filters
Programme A
Programme B
*
*
+
To DAC/loudspeakers
Parameters
Mono audio
Multichannel audio



Demo
Define System Geometry
Estimate Room Transfer Functions
$\mathbf{G}_A(\omega)$
$N_A \times L$
$\mathbf{G}_B(\omega)$
$N_B \times L$
Method-specific parameters
Set Physical Constraints
A,Q
Target A
Calculate Narrowband Source Weights
Target B
$\mathbf{q}(\omega)$
$L \times 1$
$\mathbf{q}(\omega)$
$L \times 1$
Build FIR Filters
Build FIR Filters
Programme A
Programme B
To DAC/loudspeakers
Parameters
Mono audio
Multichannel audio

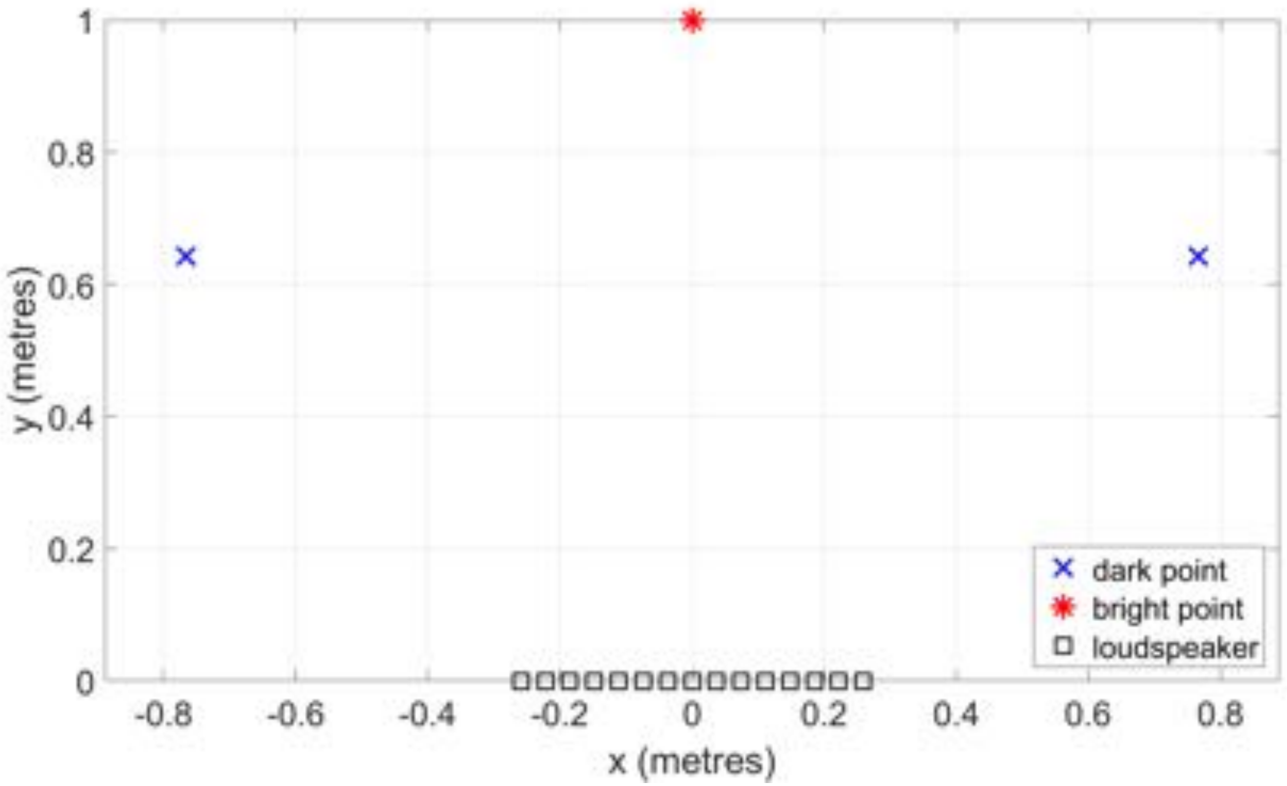
Demo: PyZones Toolbox for Sound Zone Filter Design

https://github.com/IoSR-Surrey/PyZones

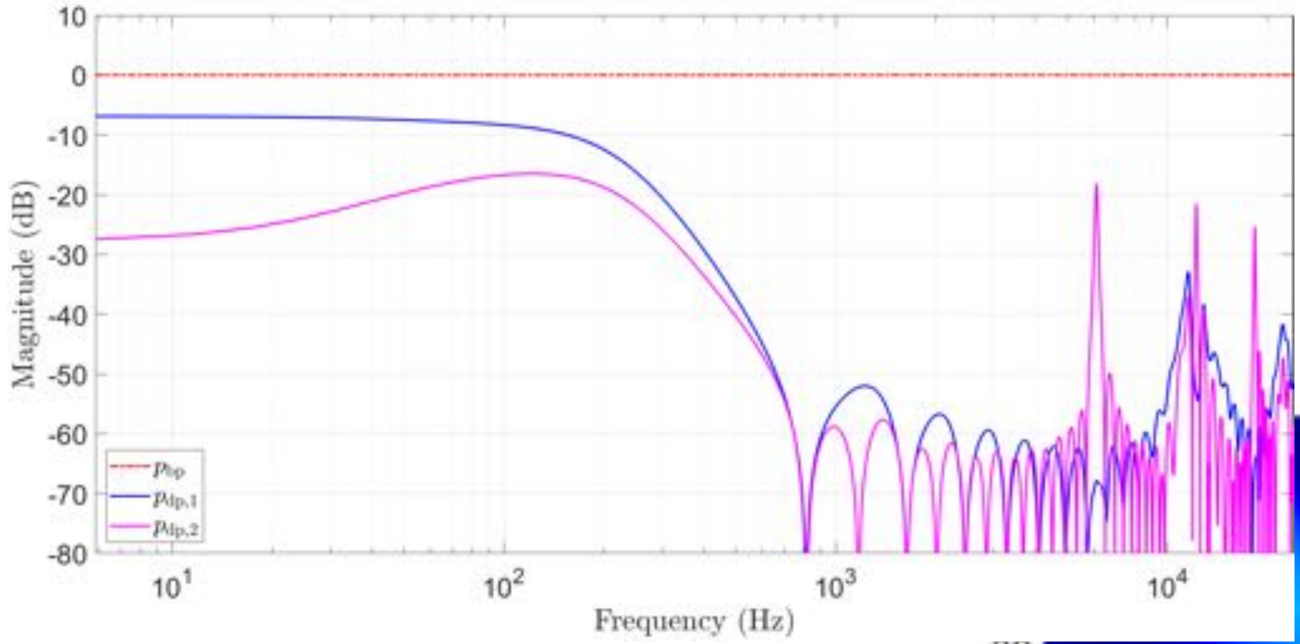
Demo: multi-zone audio delivery with a 15 channel linear loudspeaker array

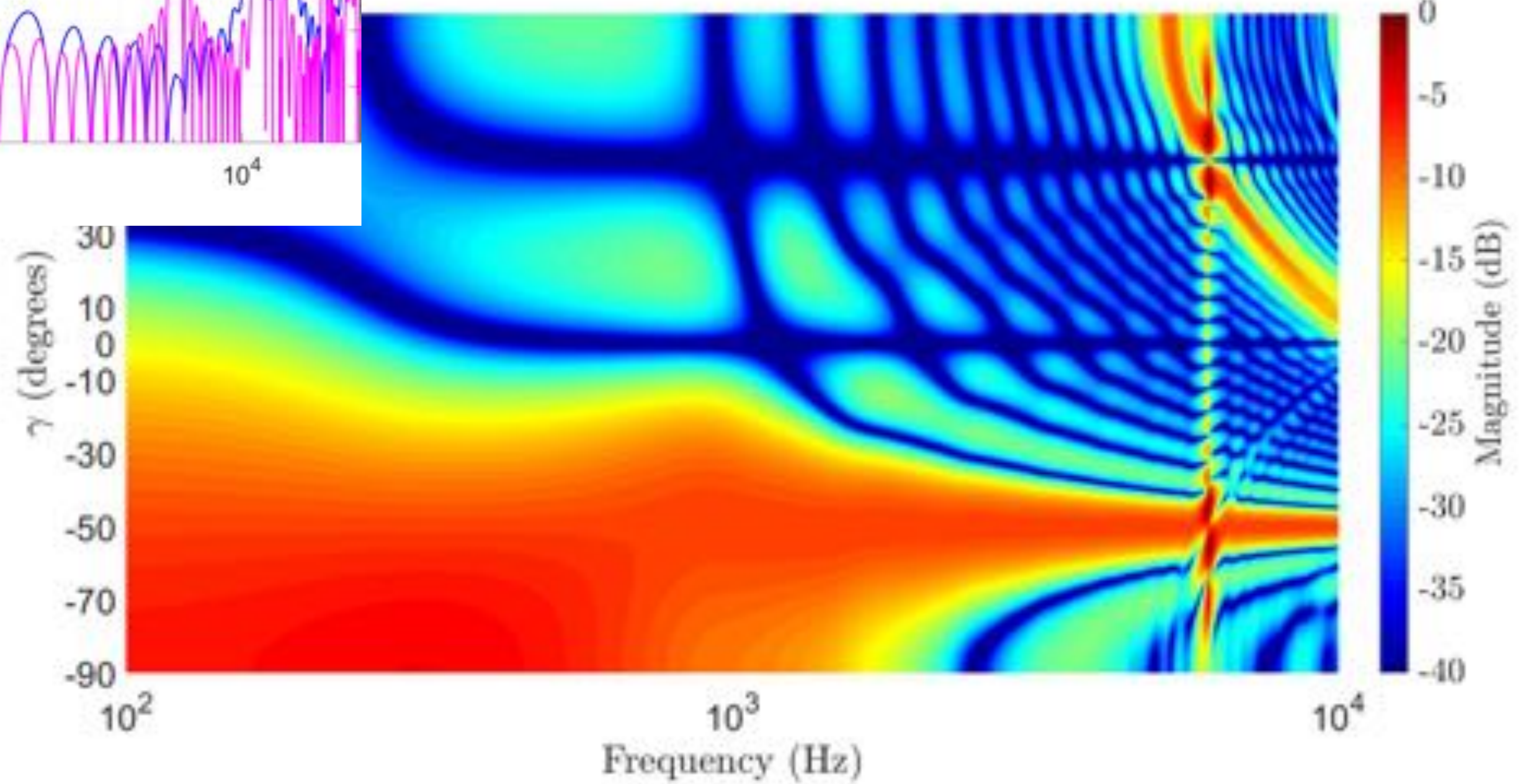
With contribution by Eric Hamdan

System geometry

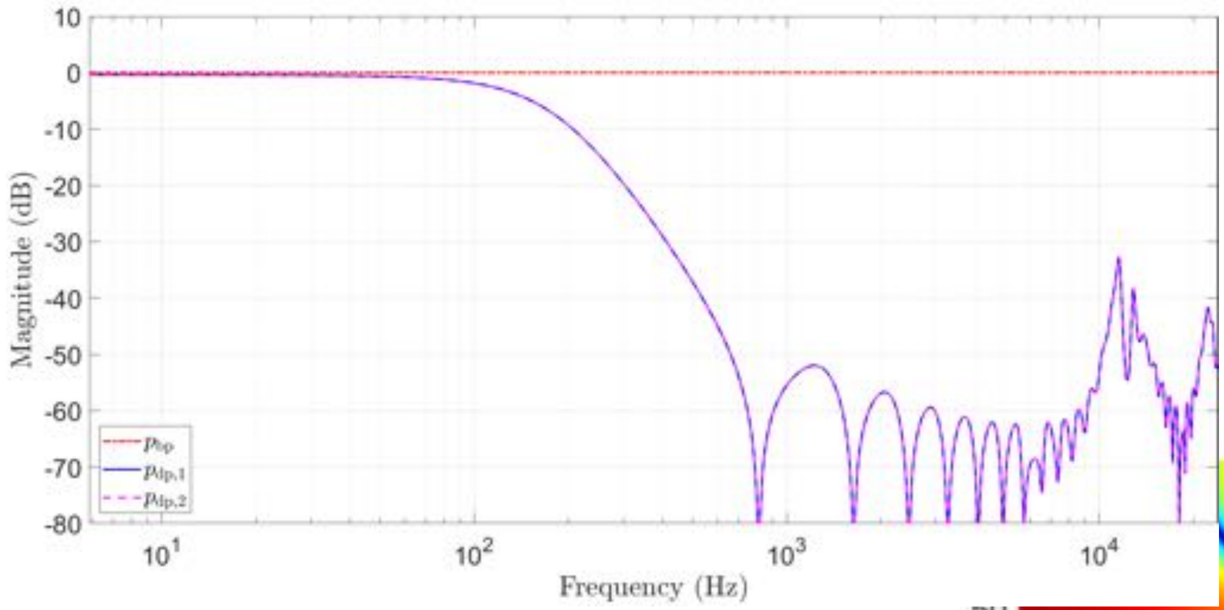


S3A

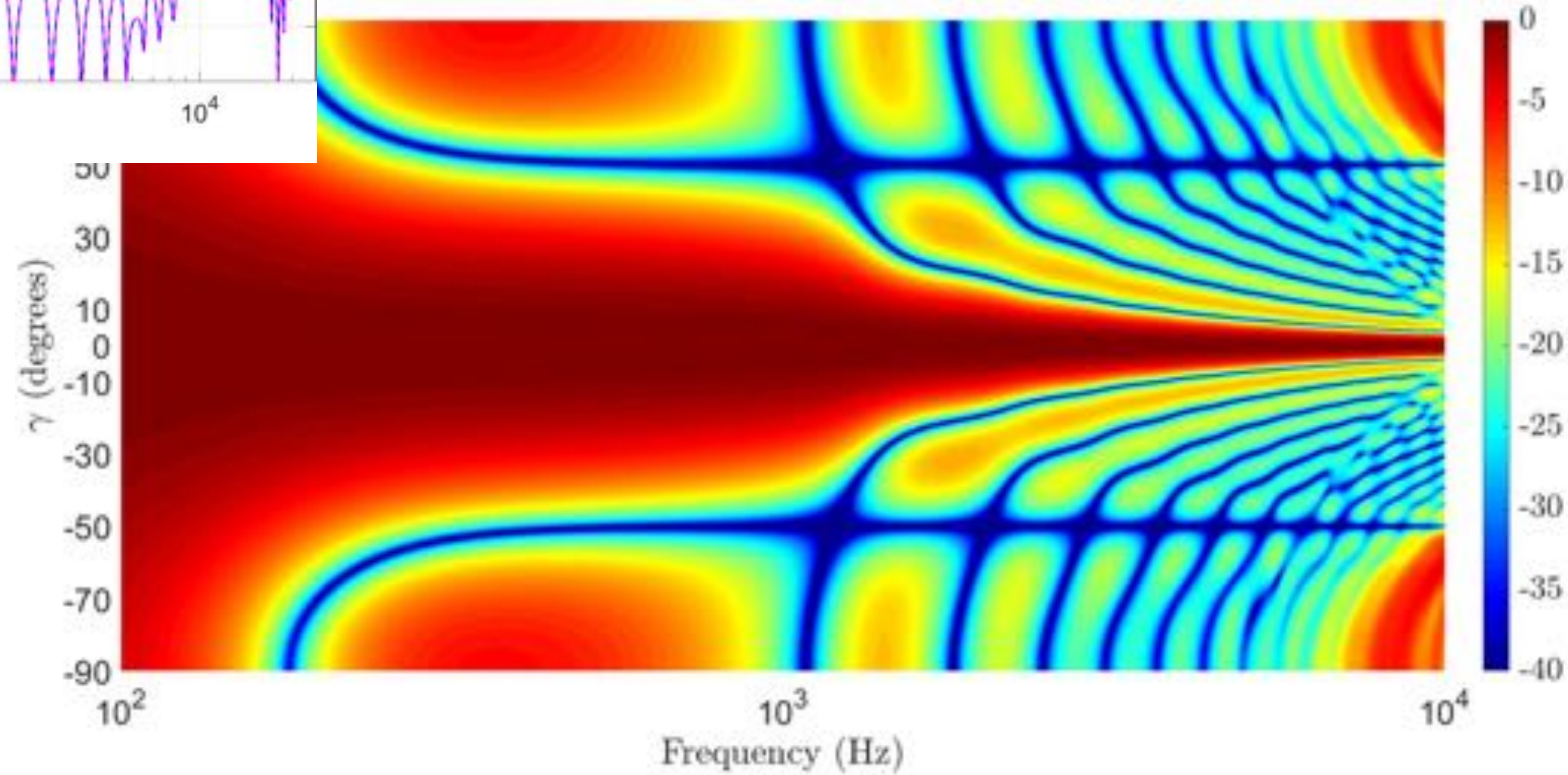




ICASSP 2019



Magnitude (dB)
Frequency (Hz)
p_{bp}
$p_{\text{dp},1}$
$p_{\text{dp},2}$



γ (degrees)
Frequency (Hz)

Perceptual Measures of Sound Zone Performance

Does it sound any good then?

- Two main factors contribute:
  - Interference suppression
  - Target sound quality
- Interference from the other zone is affected by:
  - Acoustic contrast (cost function design)
  - Relative loudness / spectra of the programme content (not accounted for in design)
- Sound quality in the target zone is affected by:
  - Flatness of frequency response (constraint, cost function design)
  - Ringing in filters (regularization, cost function design)
  - Bright zone sound propagation (cost function design)

Interference

- Understand the perceptual effects of audio interference
- First need to describe interference in a perceptually relevant way
- Attribute elicitation process led to **distraction** as the most perceptually relevant descriptor

Francombe, J. et al., 2014. "Elicitation of attributes for the evaluation of audio-on-audio interference". *JASA*.

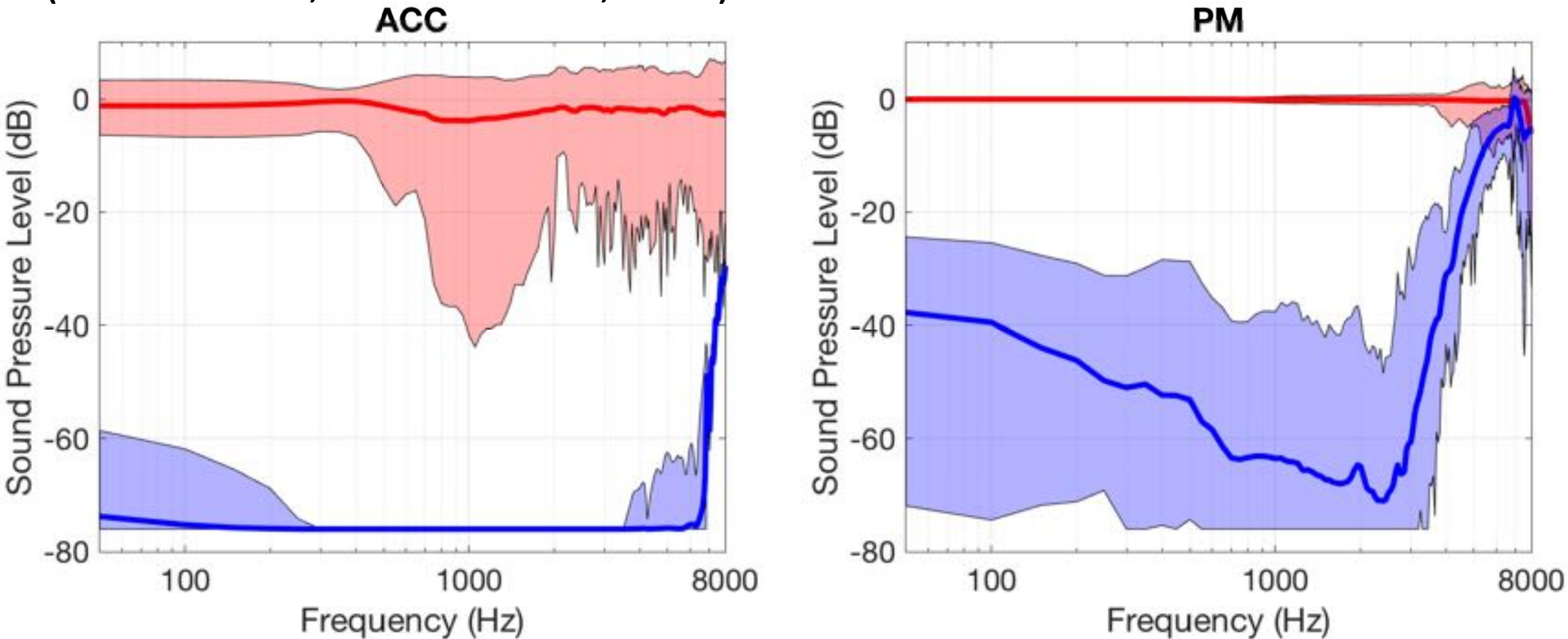
Modelling Distraction

- A regression model to predict distraction was developed at Surrey
- Main physical correlates: overall loudness, **loudness ratio**, interference rejection, and frequency content of the interferer
- Loudness ratio better correlation than (unweighted) target to interferer ratio / acoustic contrast
- Recent work has developed and validated a real-time implementation

Francombe, J., 2014. *Perceptual evaluation of audio-on-audio interference in a personal sound zone system* (Doctoral dissertation, University of Surrey).
Francombe, J. et al., 2015. "A model of distraction in an audio-on-audio interference situation with music program material". *JAES.*
Rämö, J., et al., 2017. "Real-time perceptual model for distraction in interfering audio-on-audio scenarios". *IEEE Sig. Proc. Letters.*
Rämö, J., et al., 2018. "Validating a real-time perceptual model predicting distraction caused by audio-on-audio interference". *JASA*.

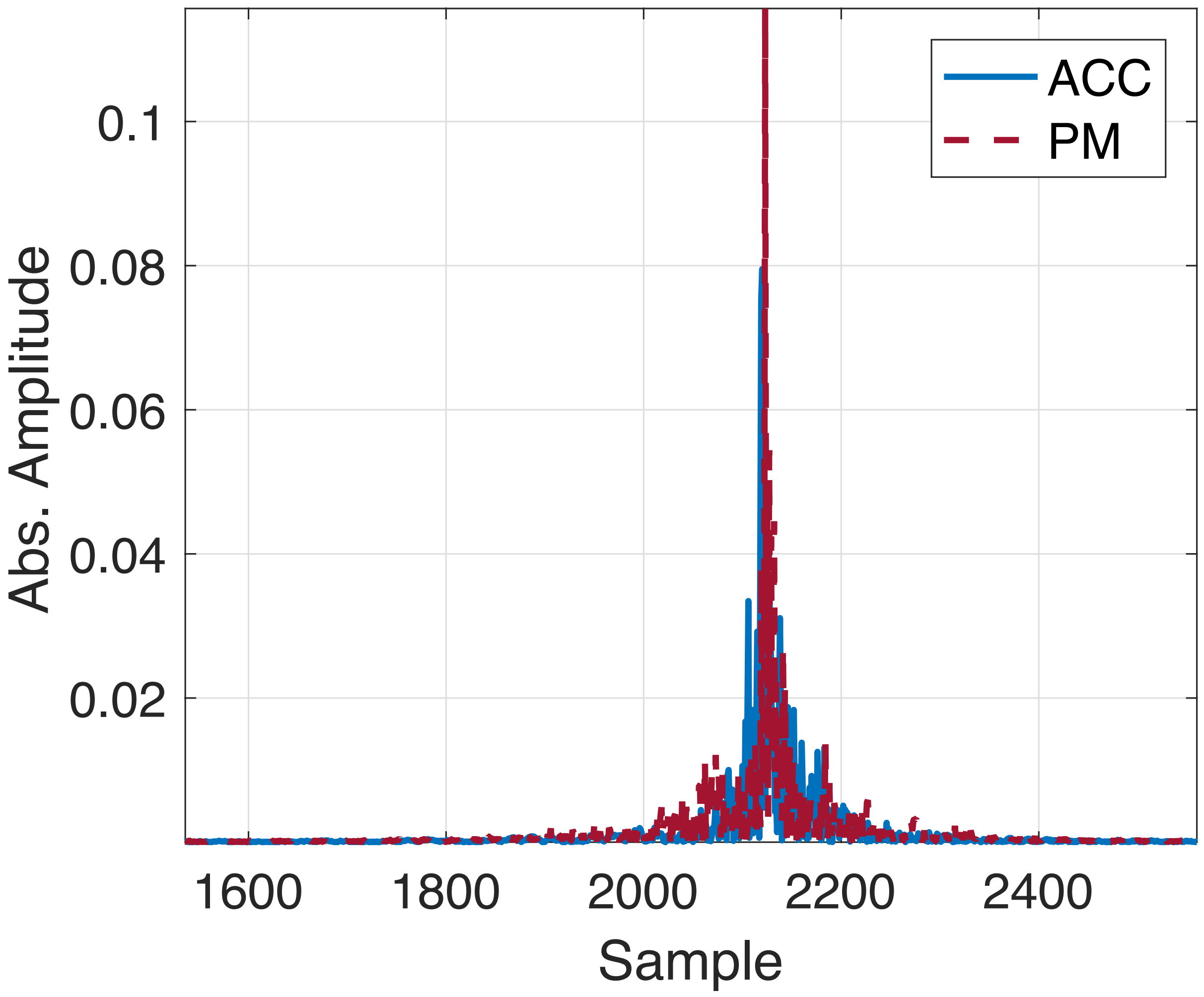
Sound Quality

- Frequency response across zone microphones (solid: mean; shaded: min, max)



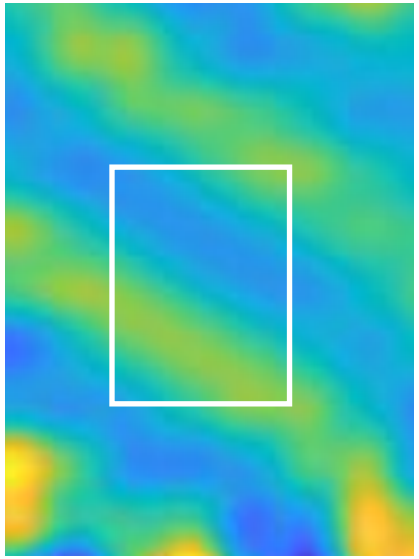
Sound Quality

- Ringing in filters
- Illustrated with simulated time response at central microphone in zone A

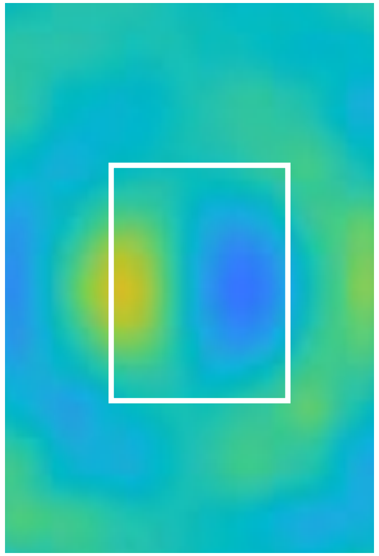


Sound Quality

- Spatial properties of the bright zone



PM: plane wave



ACC: ?

Planarity Control

- PC Cost function

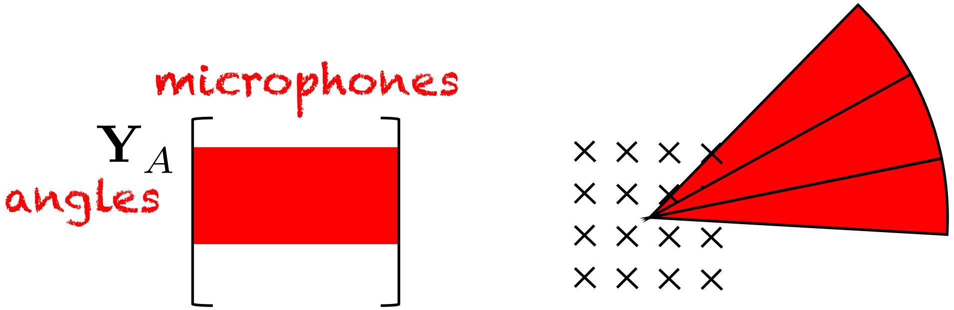
constraint on bright zone pressure
is now spatially weighted

$$\text{Min. } J_{PC} = \mathbf{p}_B^H \mathbf{p}_B + \mu(\mathbf{p}_A^H \mathbf{Y}_A^H \boldsymbol{\Gamma} \mathbf{Y}_A \mathbf{p}_A - A) + \lambda(\mathbf{q}^H \mathbf{q} - Q)$$

pressure in the dark zone
(like ACC)

constraint on effort

- Spatial steering matrix
  - Populated by superdirective beamforming

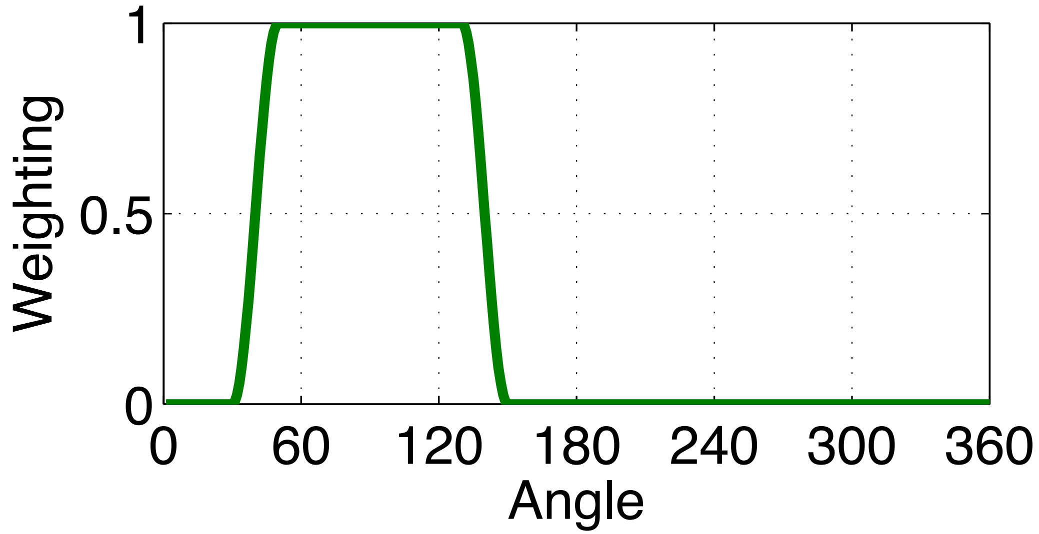


- Specify pass range of directions by means of Gamma term $\boldsymbol{\Gamma} = \text{diag}[\gamma_1, \gamma_2, \ldots, \gamma_I]$

Coleman P., 2014. *Loudspeaker array processing for personal sound zone reproduction* (Doctoral dissertation, University of Surrey).
Coleman, P., et al., 2014. “Personal audio with a planar bright zone”. *JASA.*

Angular pass range design

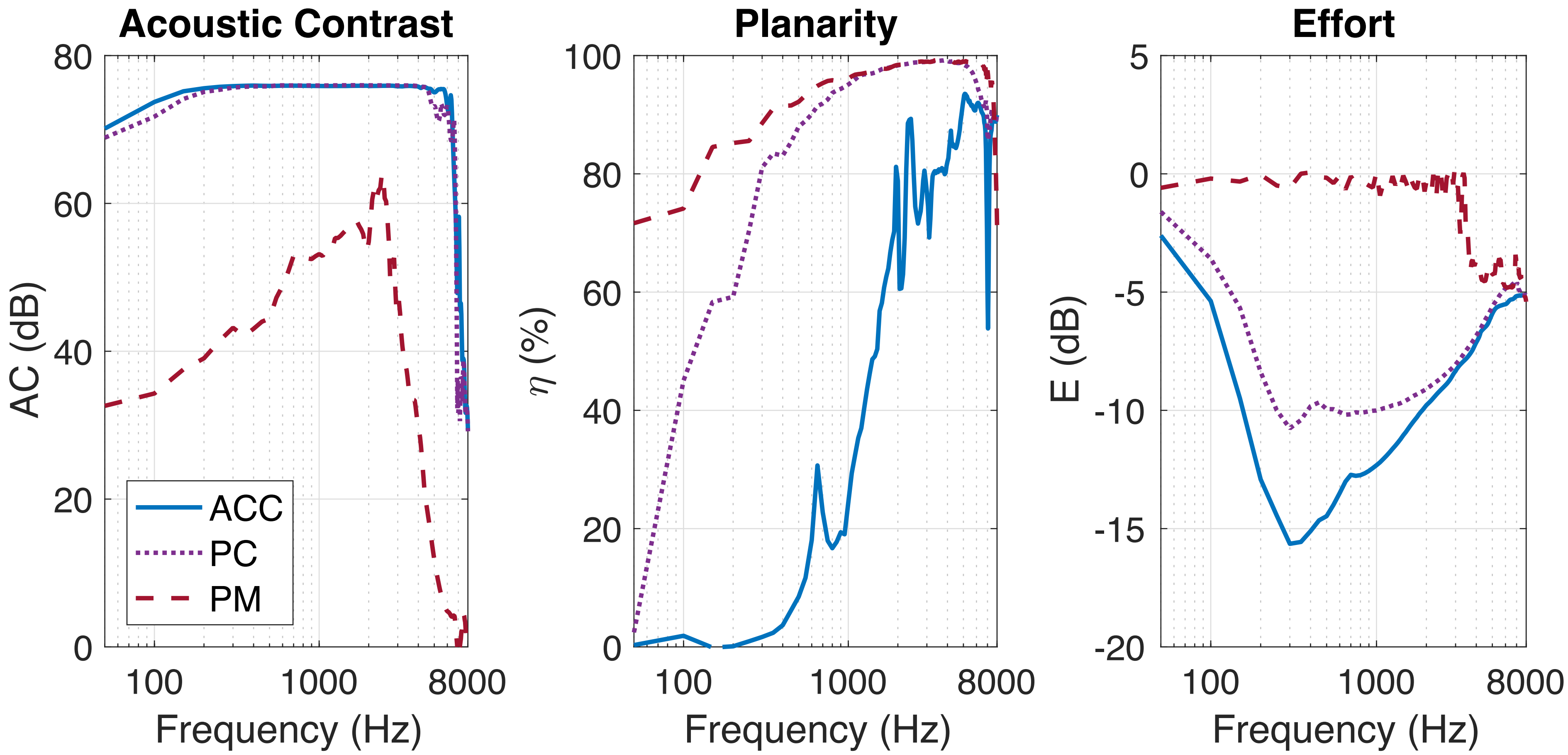
- Direction of impinging energy at bright zone
  - ACC not constrained: can leads to spatial problems
  - PM narrowly constrained
  - PC constrained to a range of angles



Coleman P., 2014. *Loudspeaker array processing for personal sound zone reproduction* (Doctoral dissertation, University of Surrey).
Coleman, P., et al., 2014. "Personal audio with a planar bright zone". *JASA.*

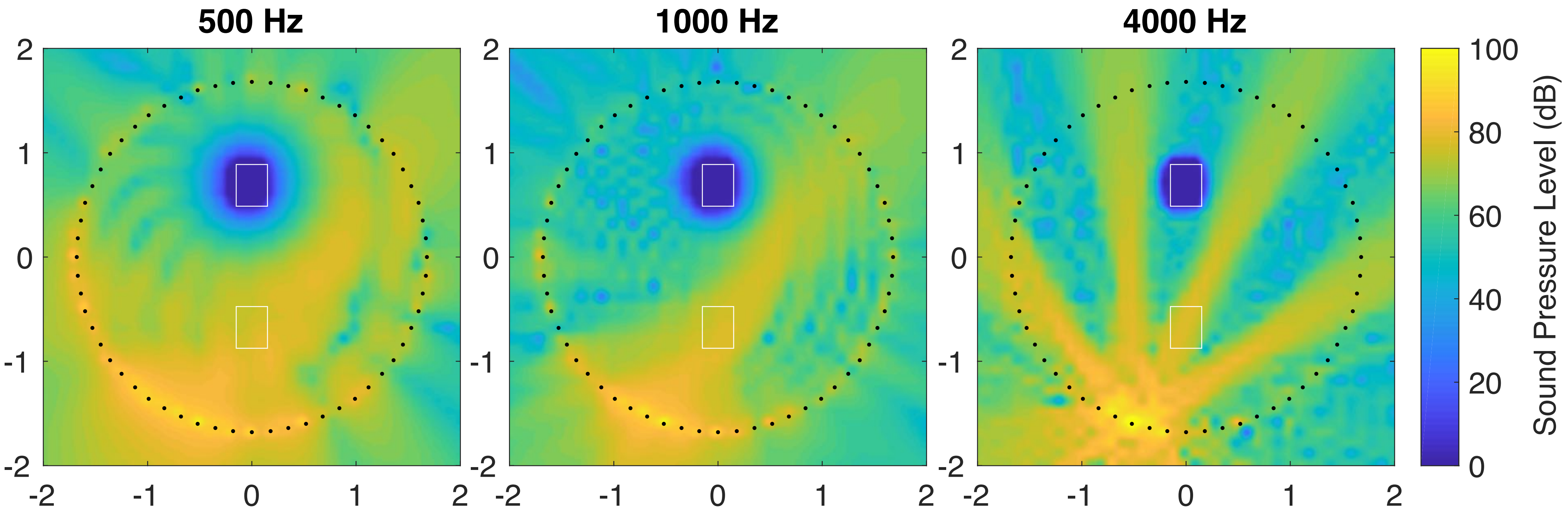
Planarity Control

- Performance over frequency



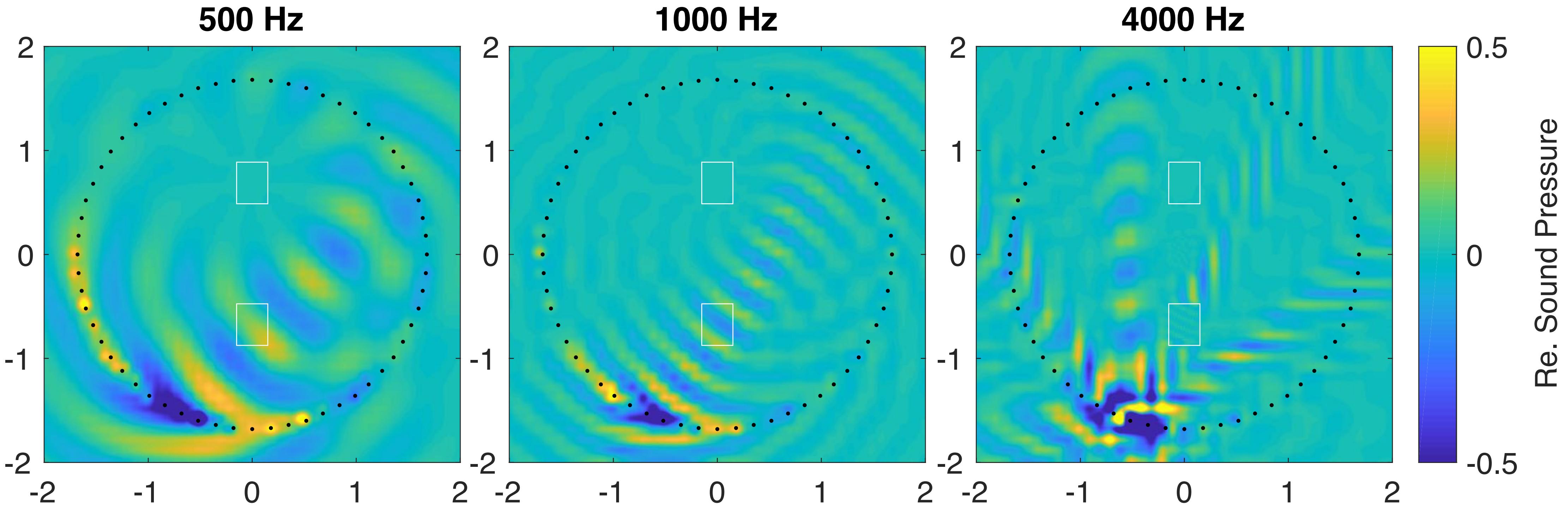
Planarity Control

- Sound field: sound pressure level



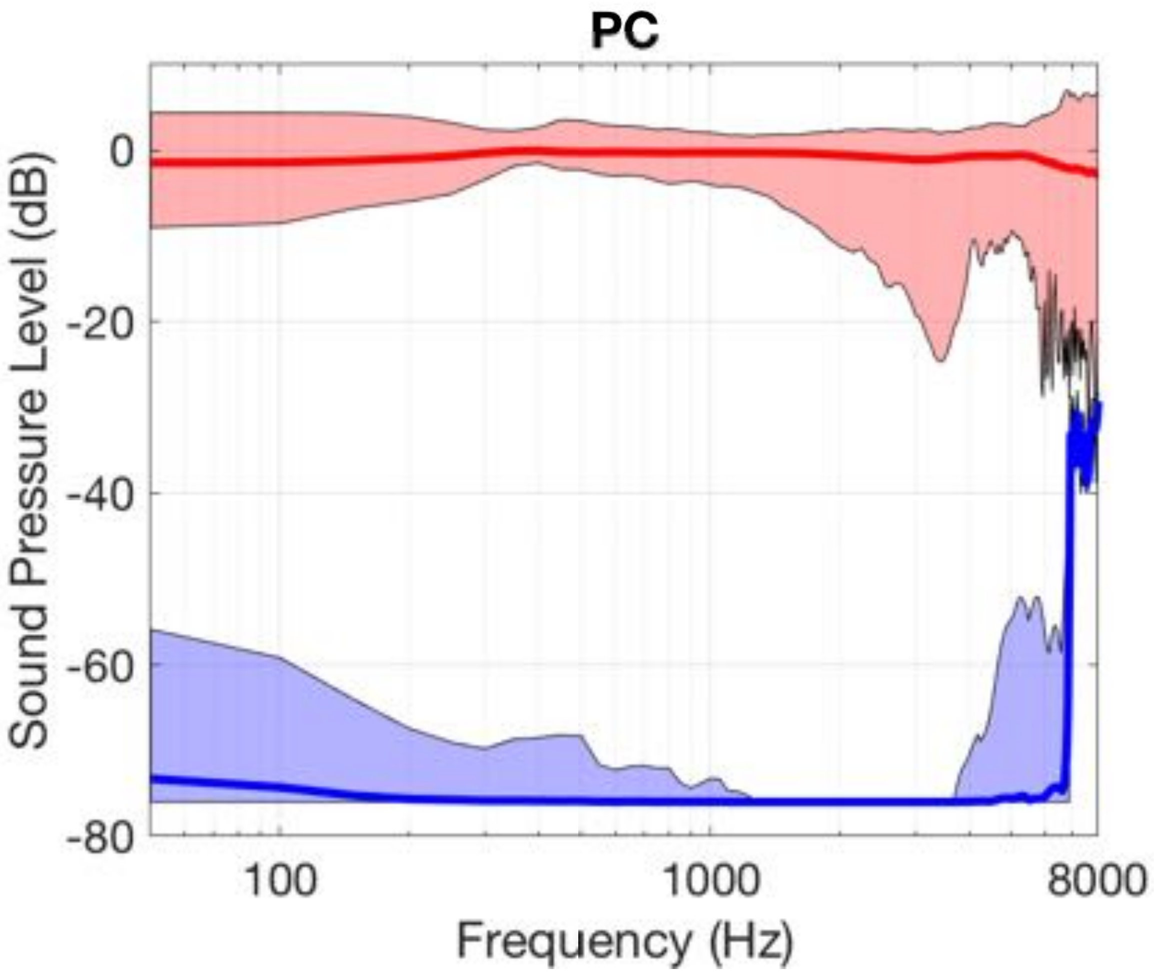
Planarity Control

- Sound field: real part of the sound pressure



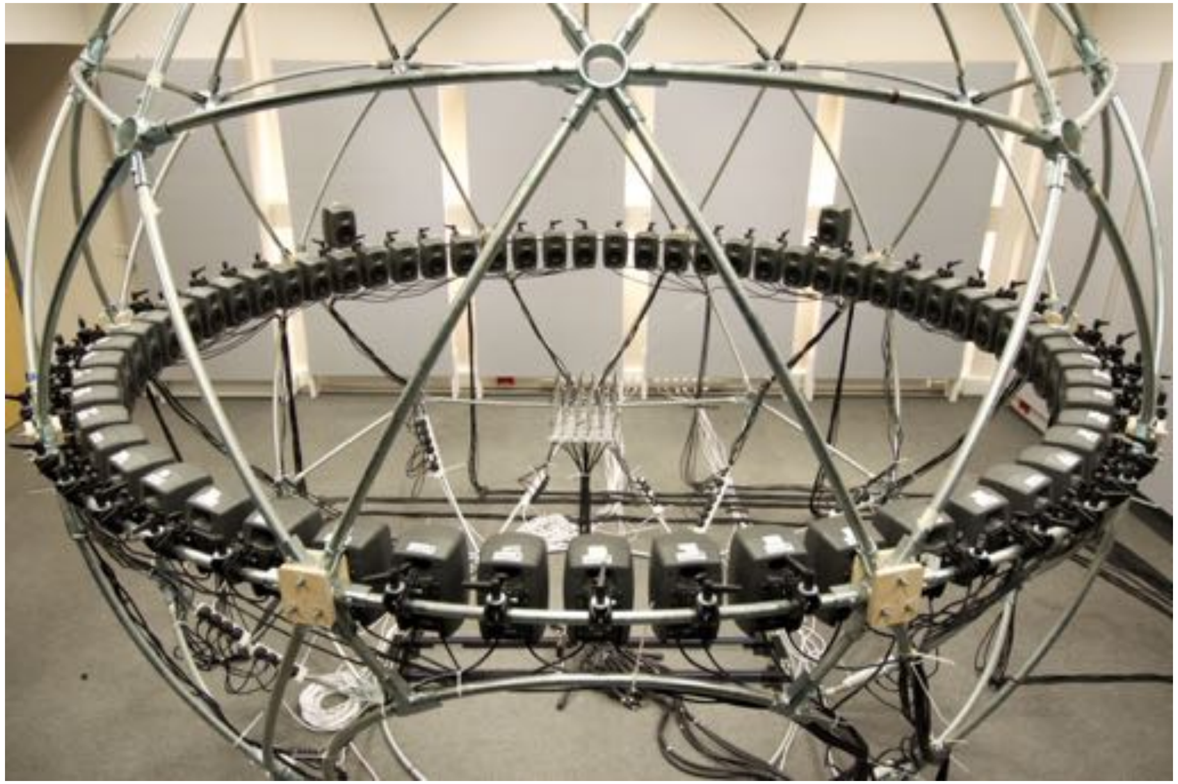
Planarity Control

- Frequency response across zone microphones
  (solid: mean; shaded: min, max)



Perceptual Evaluation

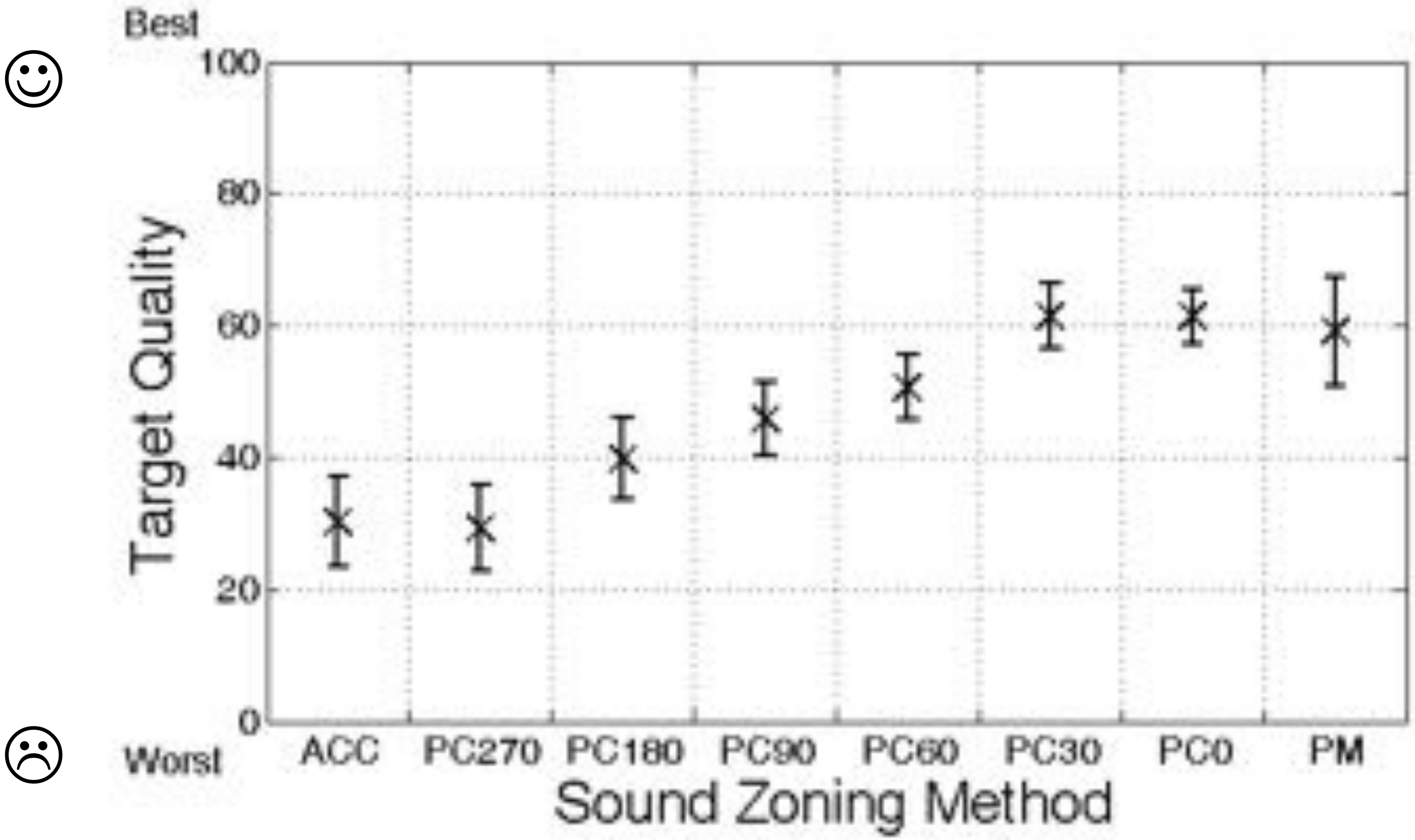
- We ran listening tests to compare ACC, PC and PM
  - Including some variants of PC
- Implemented in the Surrey Sound Sphere (nothing simulated!)
- Listeners separately rated:
  - Target quality
  - Distraction
  - Overall quality



Baykaner, K. et al., 2015. "The relationship between target quality and interference in sound zones". *JAES*.

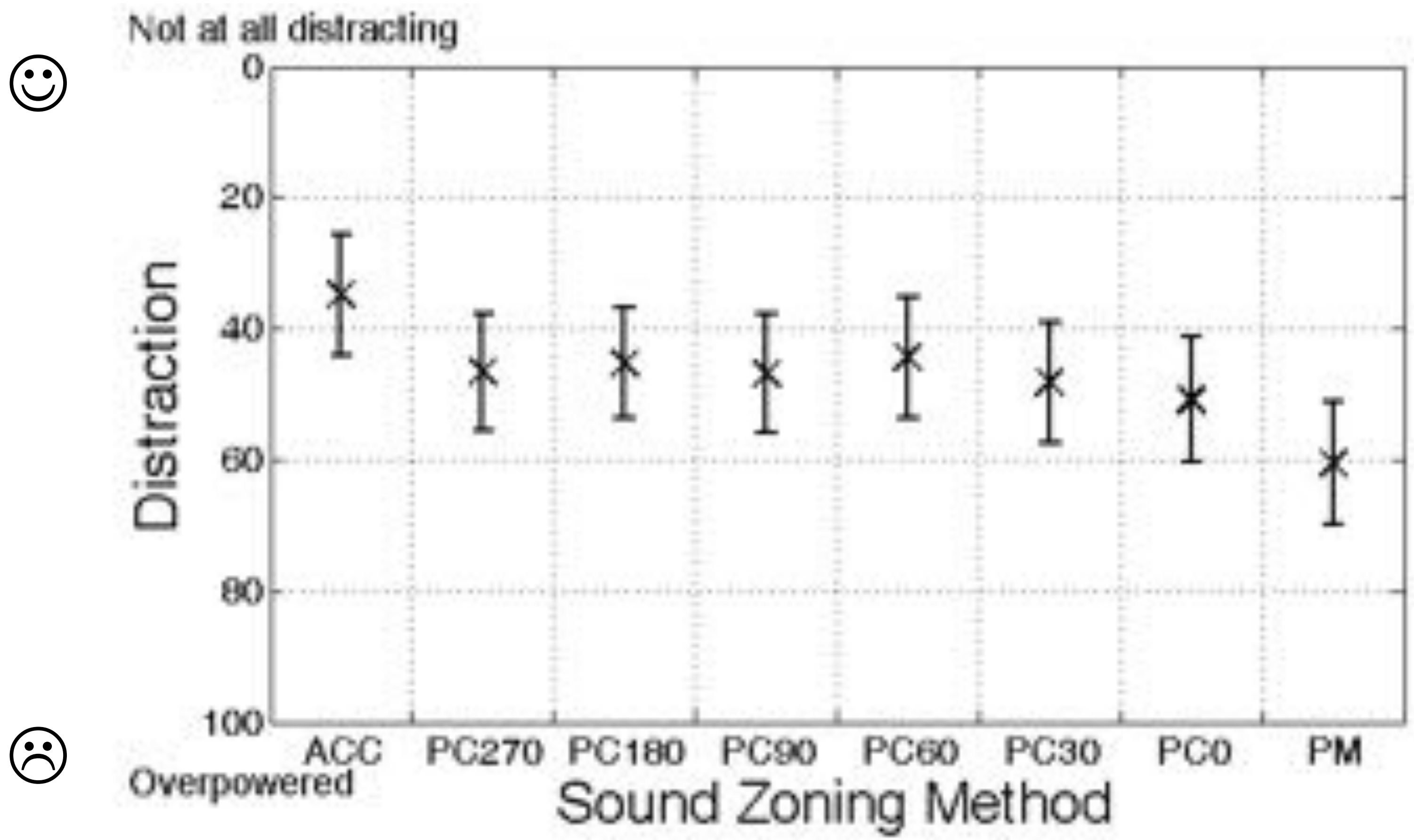
Perceptual Evaluation

- Target Quality



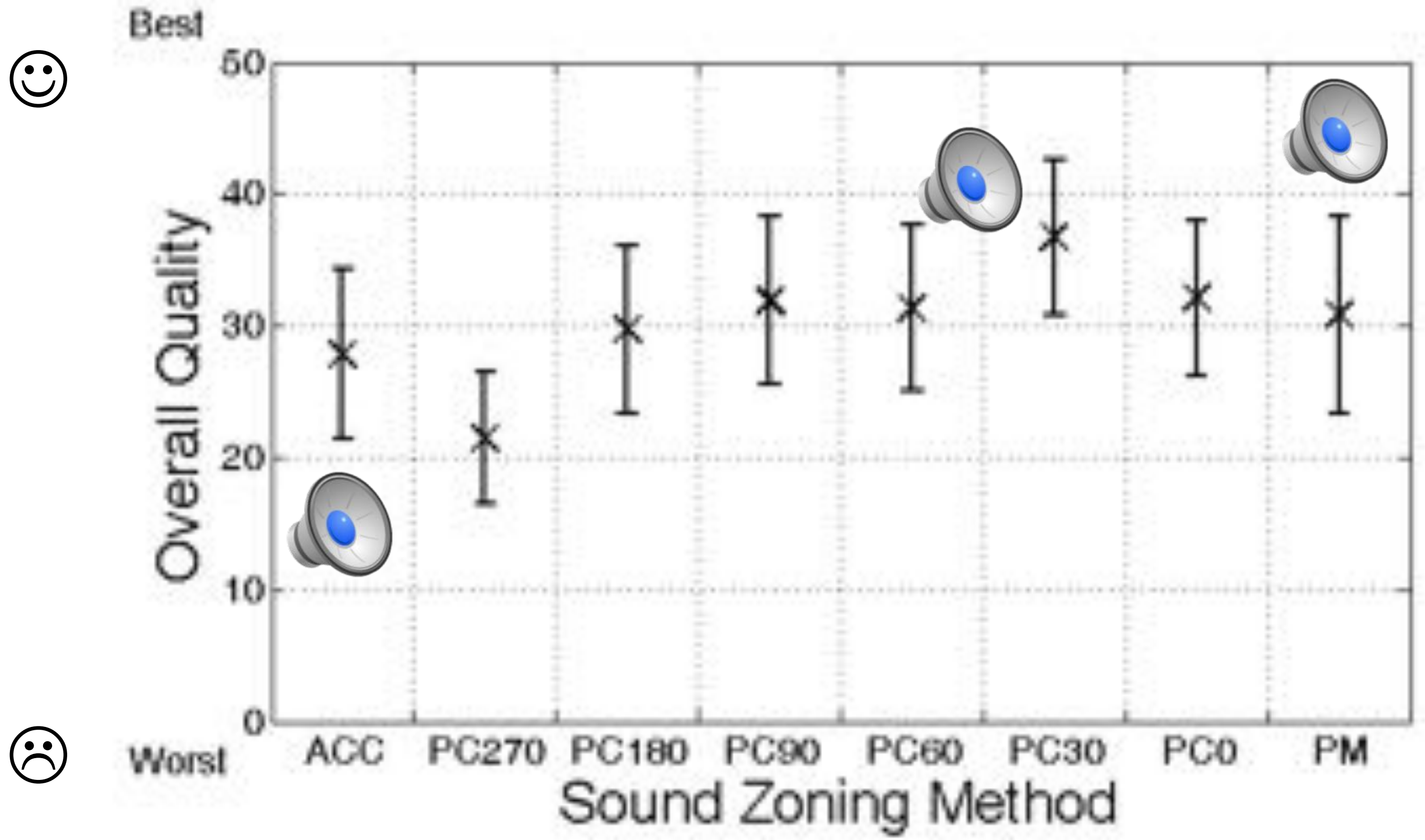
Perceptual Evaluation

- Distraction



Perceptual Evaluation

- Overall Quality



Perceptual Evaluation

- Conclusions:
  - In choosing a cost function, there is a trade-off between target quality and interference
  - The spatial energy distribution and phase in the bright zone was the main difference between methods as tested
  - Planarity control, with a soft DOA constraint, gives a good trade-off between the characteristics of the baseline methods (ACC and PM)
  - Other aspects (frequency response, ringing in filters) have not been explored to date in the context of sound zones

Alternative approaches to personal sound zones

Spherical Harmonic Representation of Sound Fields

- A sound field that is solution the homogeneous Helmholtz equation can be represented as

Basis functions

$$p(r,\theta,\varphi,\omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} B_{nm}(\omega)\, i^n j_n(kr)\, Y_n^m(\theta,\varphi)$$

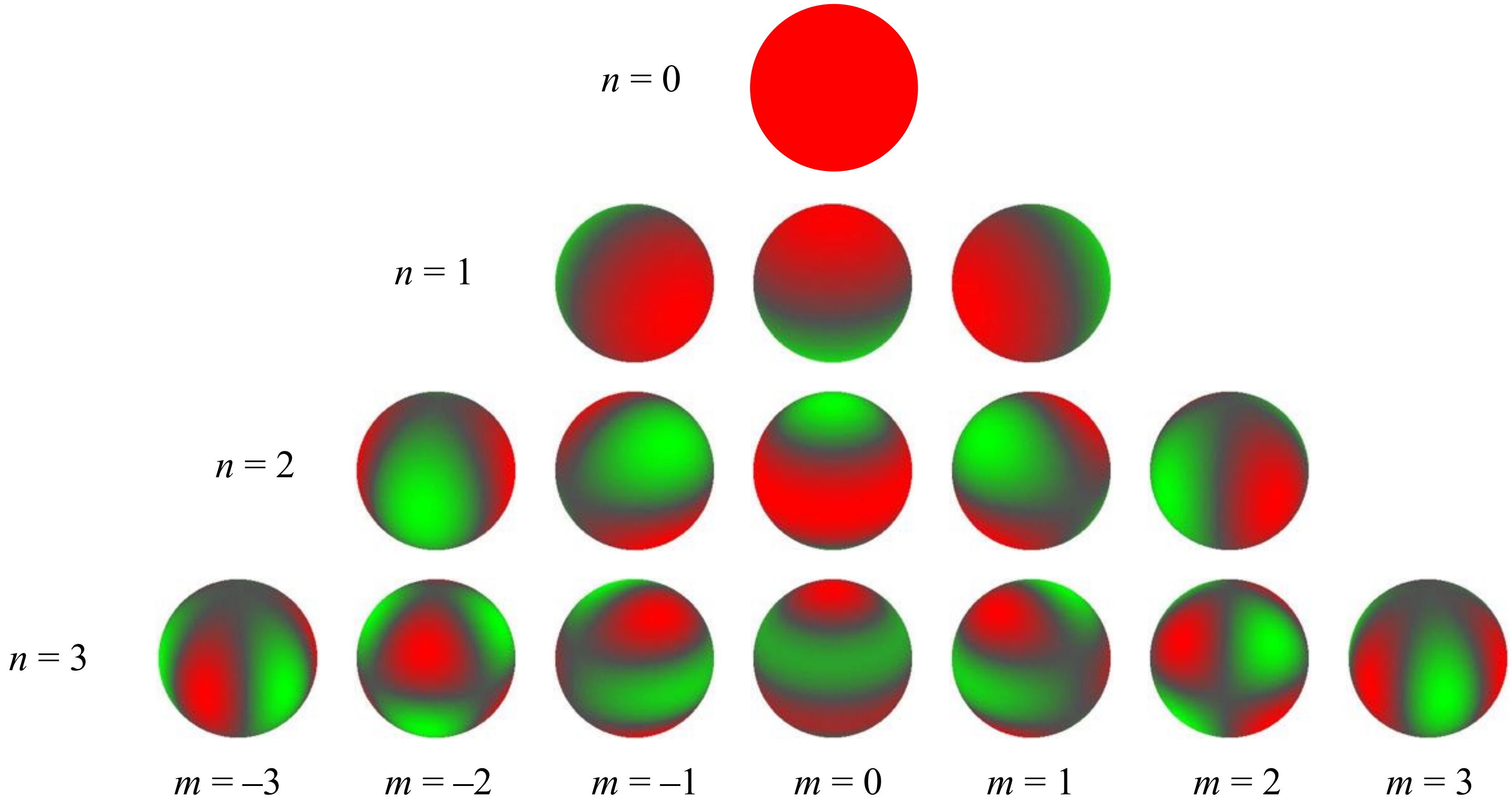
Sound field

Series coefficients

Spherical Bessel functions

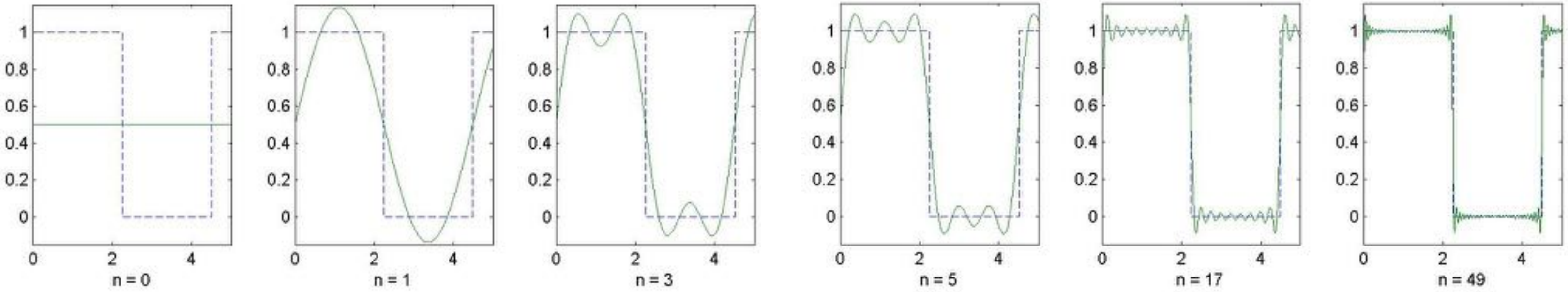
Spherical harmonics

The Spherical Harmonics

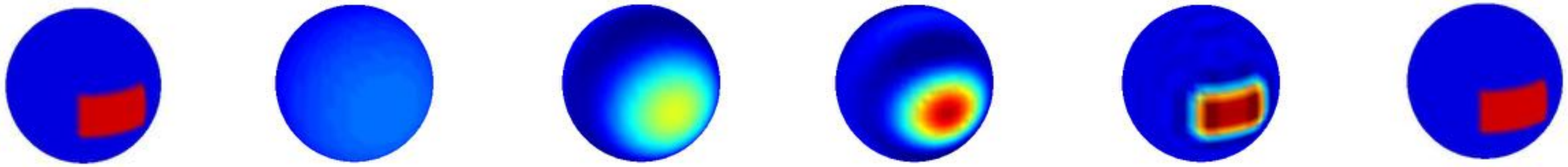


Spherical Harmonics and Fourier Series

- Fourier series



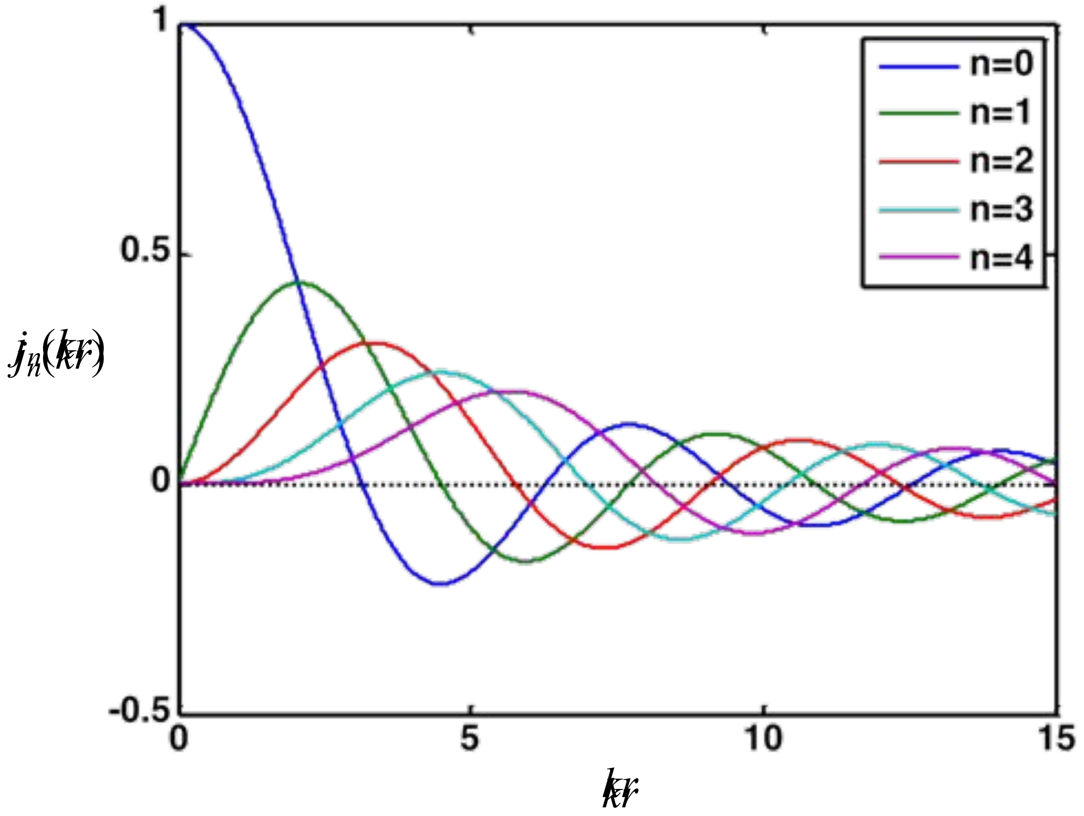
- Spherical harmonic expansion (generalised Fourier series)

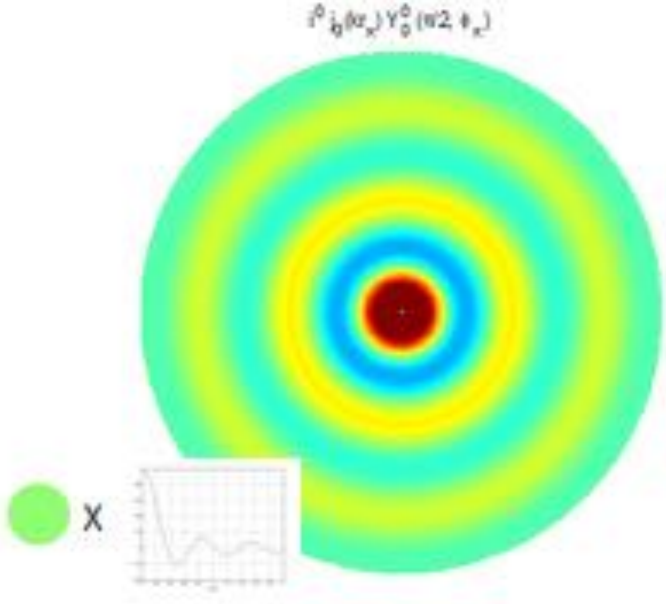


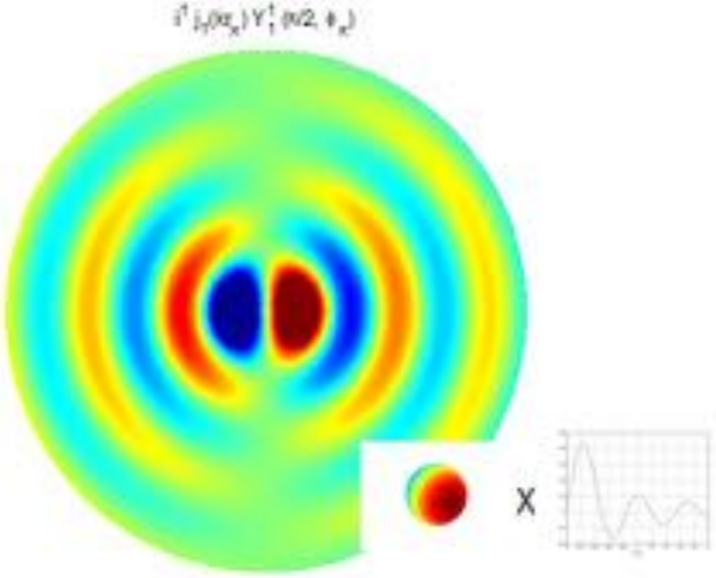
Target n=1 n=3 n=5 n=17 n=49

The Spherical Bessel Functions

- The radial dependence of a sound field is expressed by the spherical Bessel functions

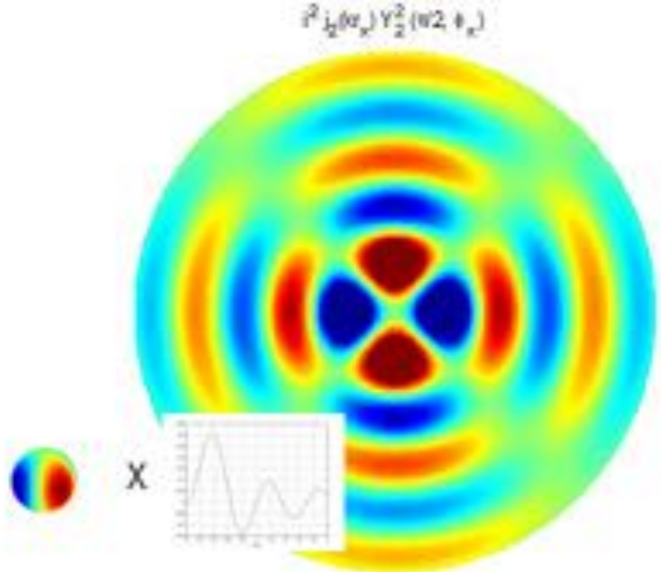


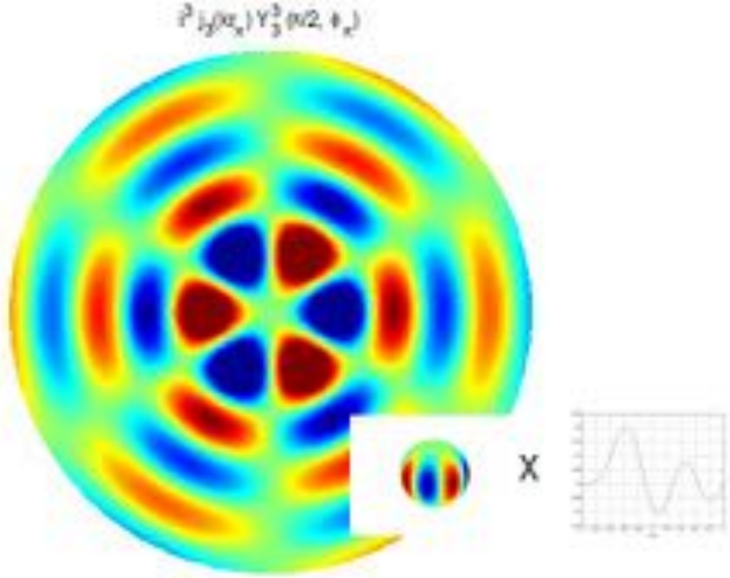




Basis functions

$$i^n\, j_n(kr)\; Y_n^m(\theta,\varphi)$$

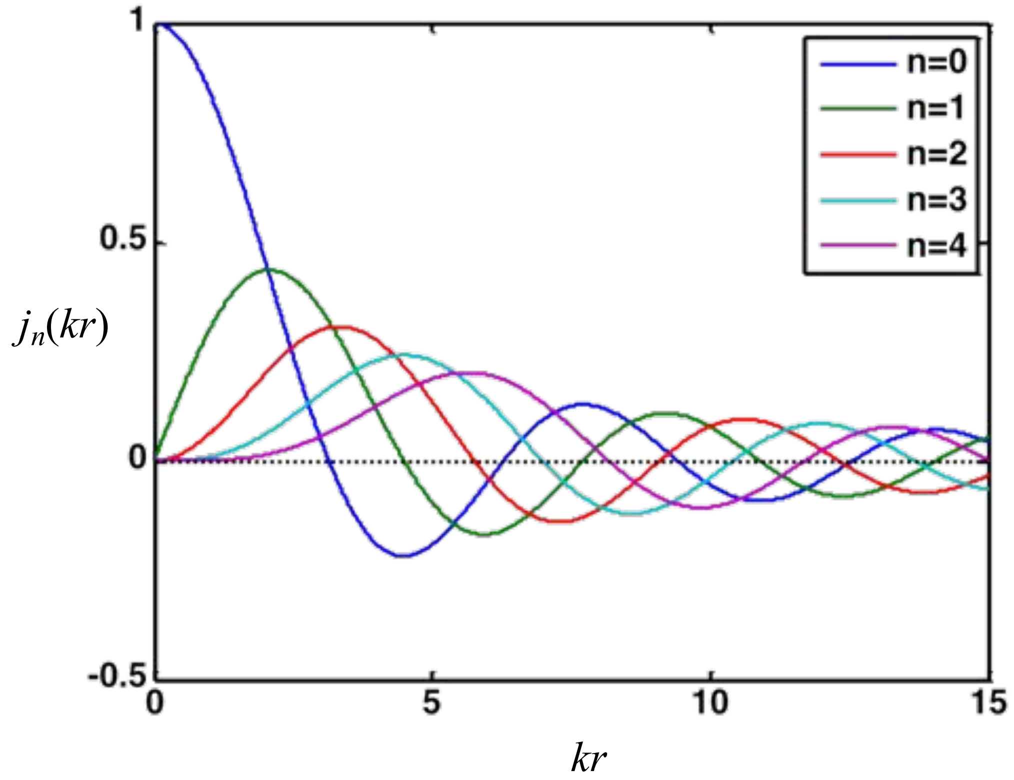


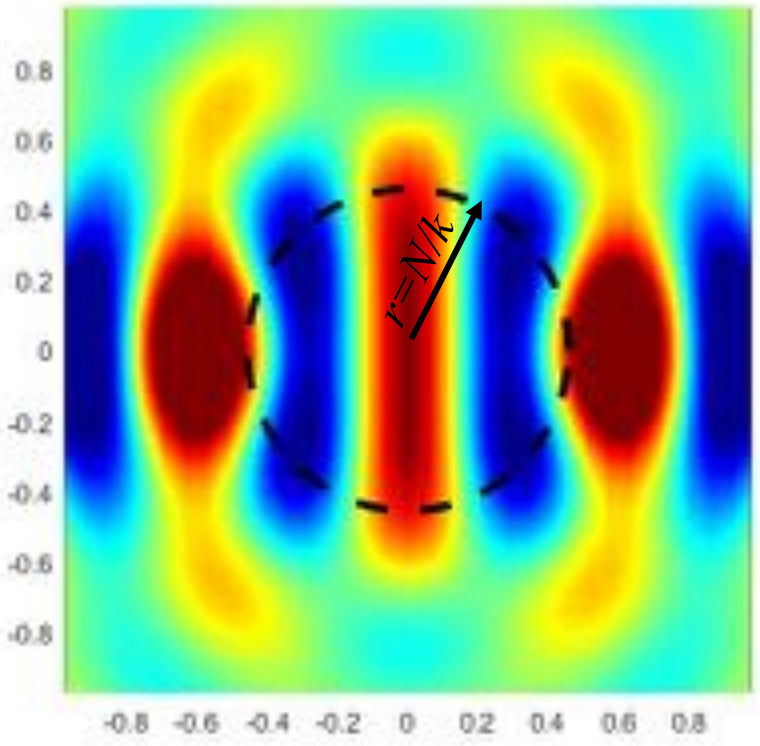


Series truncation

$$p(r,\theta,\varphi) \approx \sum_{n=0}^{N} \sum_{m=-n}^{n} B_{nm}\, i^n\, j_n(kr)\, Y_n^m(\theta,\varphi)$$

$$N \geq kr$$





Mode matching

Target sound field

$$p_T(r,\theta,\varphi) \approx \sum_{n=0}^{N}\sum_{m=-n}^{n} B_{nm}\, i^n j_n(kr)\, Y_n^m(\theta,\varphi)$$

Reproduced sound field

$$p(r,\theta,\varphi) \approx \sum_{n=0}^{N}\sum_{m=-n}^{n}\sum_{\ell=1}^{L} H_{nm\ell}\, q_\ell\, i^n j_n(kr)\, Y_n^m(\theta,\varphi)$$

Basis functions

Least squares minimisation

$$J = \| p - p_T \|^2 \;\Rightarrow\; J = \sum_{n=0}^{N}\sum_{m=-n}^{n} \left| \sum_{\ell=1}^{L} H_{nm\ell}\, q_\ell - B_{nm} \right|^2 \;\Rightarrow\; \mathbf{J} = \| \mathbf{Hq} - \mathbf{b} \|^2$$

Loudspeaker signals ($\mathbf{q}$)

Loudspeaker field coeff. matrix ($\mathbf{H}$)

Target field coefficients ($\mathbf{b}$)

Multi-zone mode matching

$$\mathbf{b}_{M\times 1} = \begin{vmatrix} \mathbf{b}_A \\ \cdots \\ \mathbf{b}_B \end{vmatrix}, \quad \mathbf{H}_{M\times L} = \begin{vmatrix} \mathbf{H}_A \\ \cdots \\ \mathbf{H}_B \end{vmatrix}$$

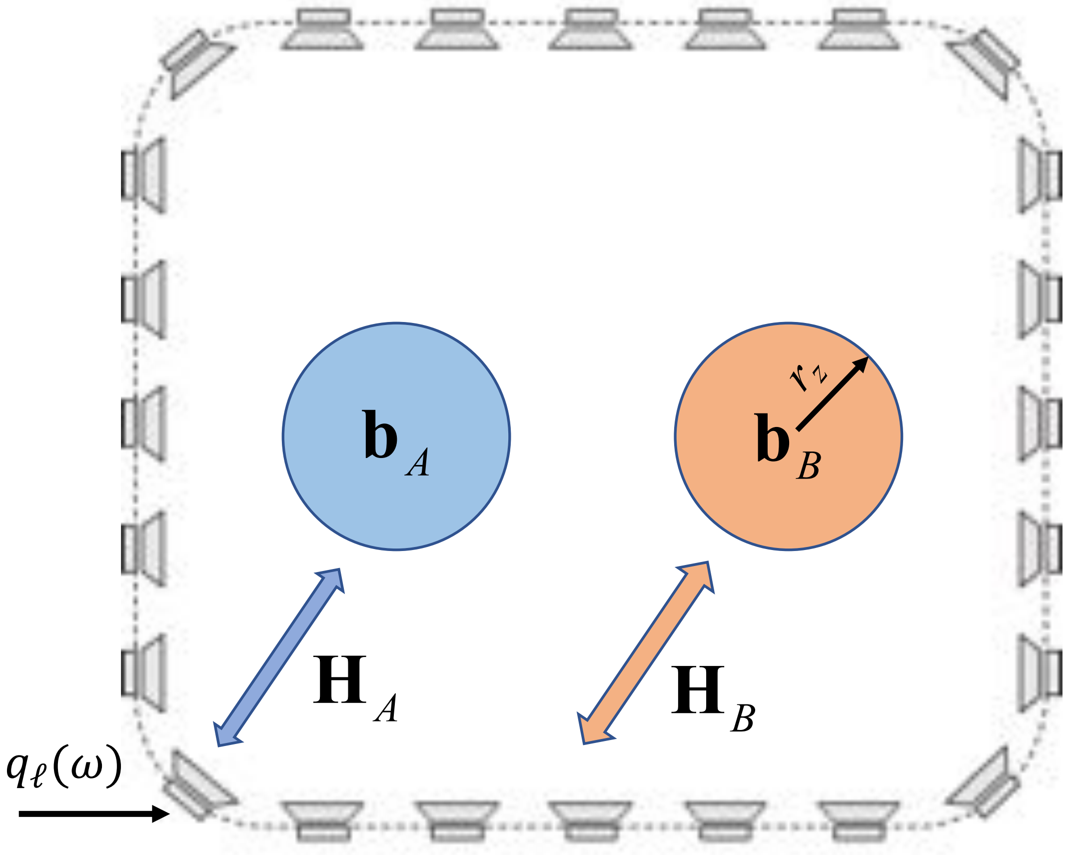
$$M \geq N_z (kr_z + 1)^2$$

- Least squares minimisation

$$J = \left\| \mathbf{H}\,\mathbf{q} - \mathbf{b} \right\|^2$$

$$\mathbf{q}_{opt} = \mathbf{H}^{\dagger}\mathbf{b} = \left(\mathbf{H}^H\mathbf{H} + \beta\mathbf{I}\right)^{-1}\mathbf{H}^H\mathbf{b}$$

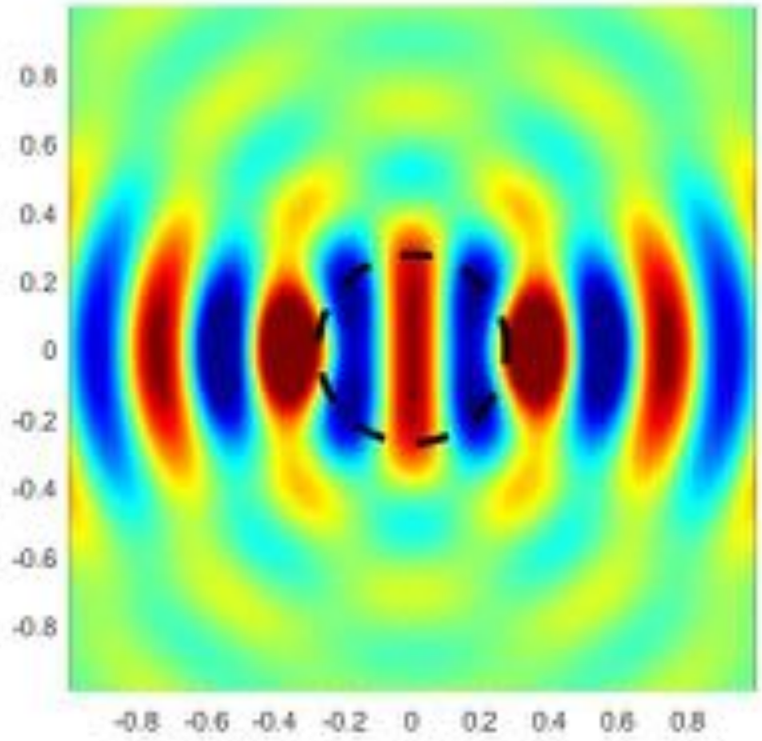
- Conceptually similar to pressure matching



Poletti and Betlehem, 2014. “Creation of a single sound field for multiple listeners”. INTER-NOISE
Wu and Abhayapala, 2011 . “Spatial Multizone Soundfield Reproduction: Theory and Design”. *IEEE*

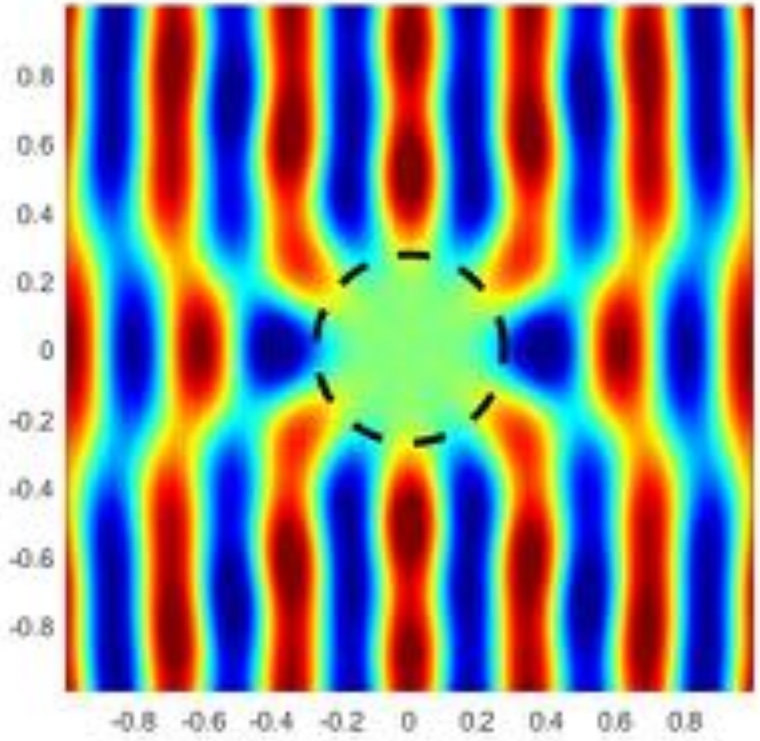
Zone of silence

Truncated field



$$p_N = \sum_{n=0}^{N} \sum_{m=-n}^{n} B_{nm}\, i^n j_n(kr)\, Y_n^m(\theta,\varphi)$$

Complementary field



$$p_N^{(C)} = \sum_{n=N+1}^{\infty} \sum_{m=-n}^{n} B_{nm}\, i^n j_n(kr)\, Y_n^m(\theta,\varphi)$$

ICASSP 2019

Zone of silence

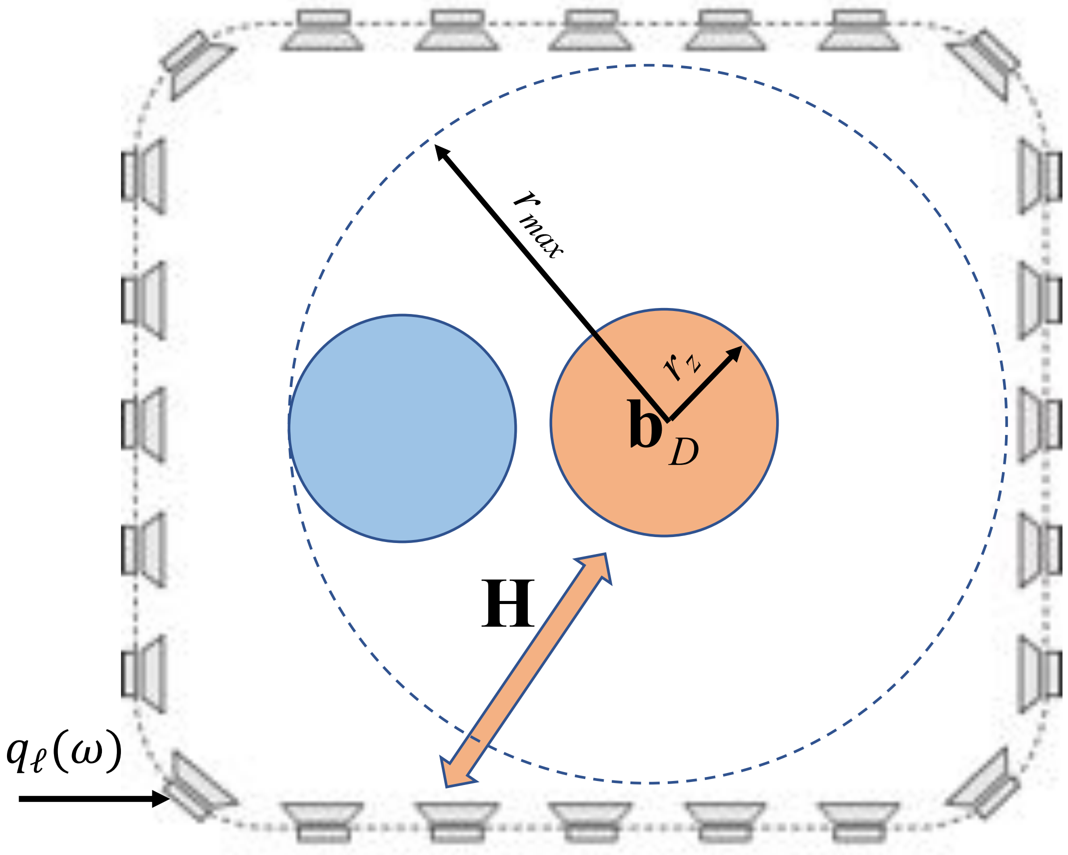
$$\mathbf{b}_D = [0, 0, 0, \ldots, B_{M+1,0}, \ldots, B_{NN}]^T$$

$$M = kr_z, \quad N \geq kr_{\max}$$

- Least squares minimisation

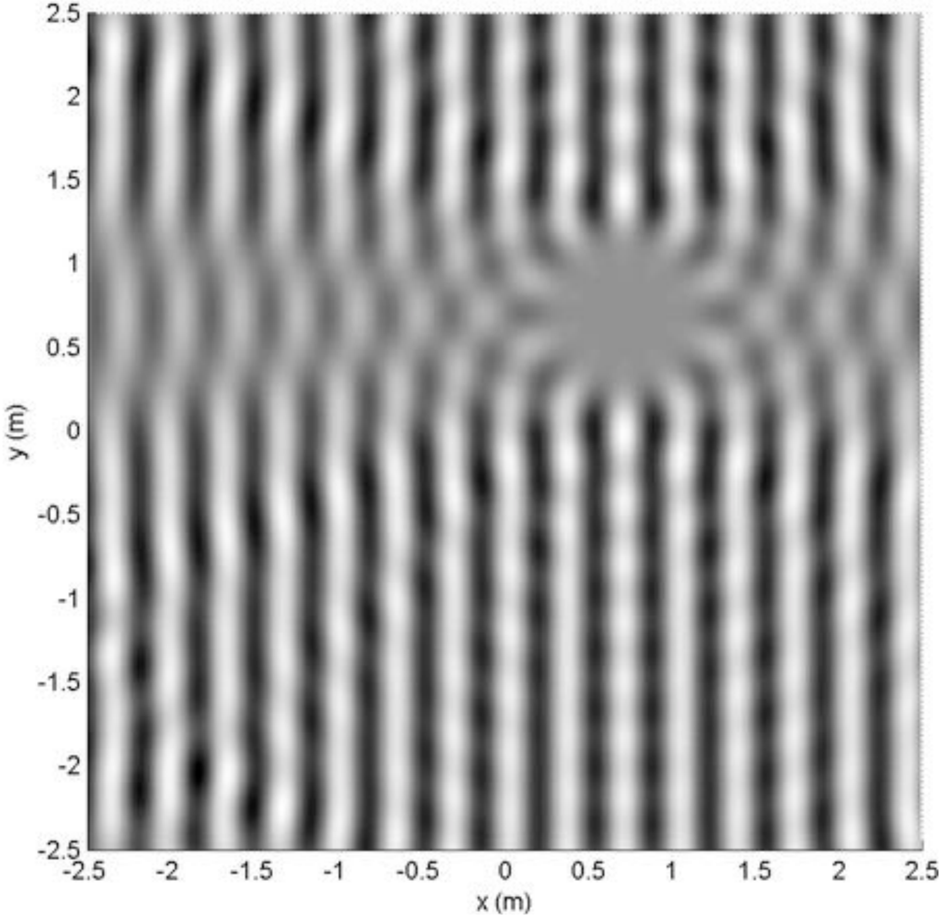
$$J = \left\| \mathbf{H}\,\mathbf{q} - \mathbf{b} \right\|^2$$

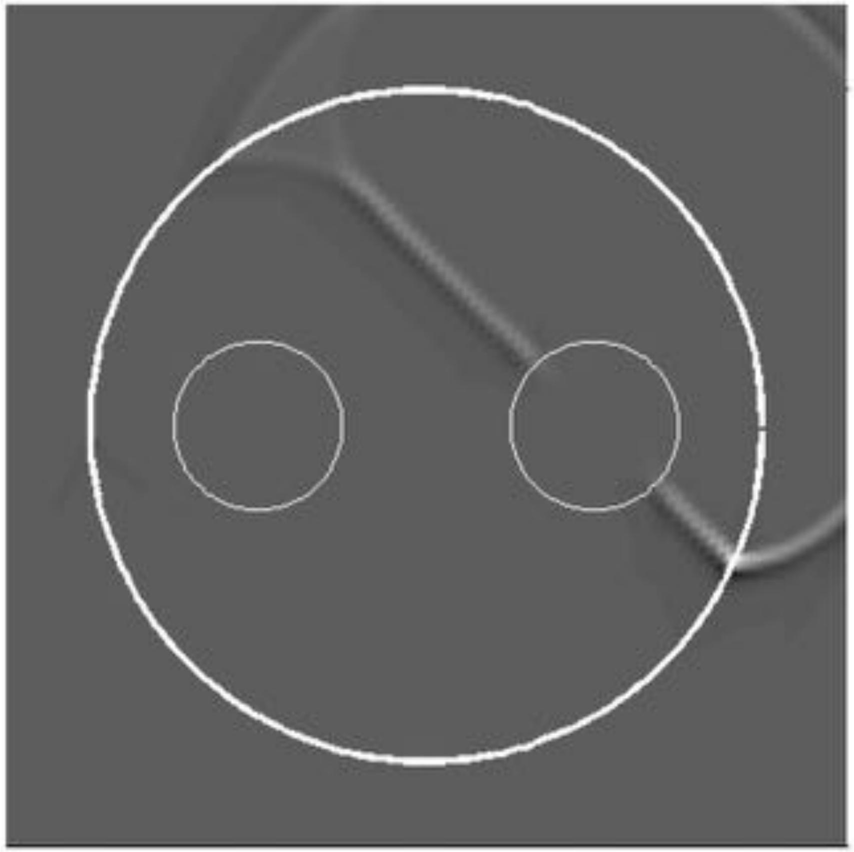
$$\mathbf{q}_{opt} = \mathbf{H}^{\dagger}\mathbf{b} = \left(\mathbf{H}^H \mathbf{H} + \beta \mathbf{I} \right)^{-1} \mathbf{H}^H \mathbf{b}$$



Poletti and Fazi, 2015. "An approach to generating two zones of silence with application to personal sound systems". JASA

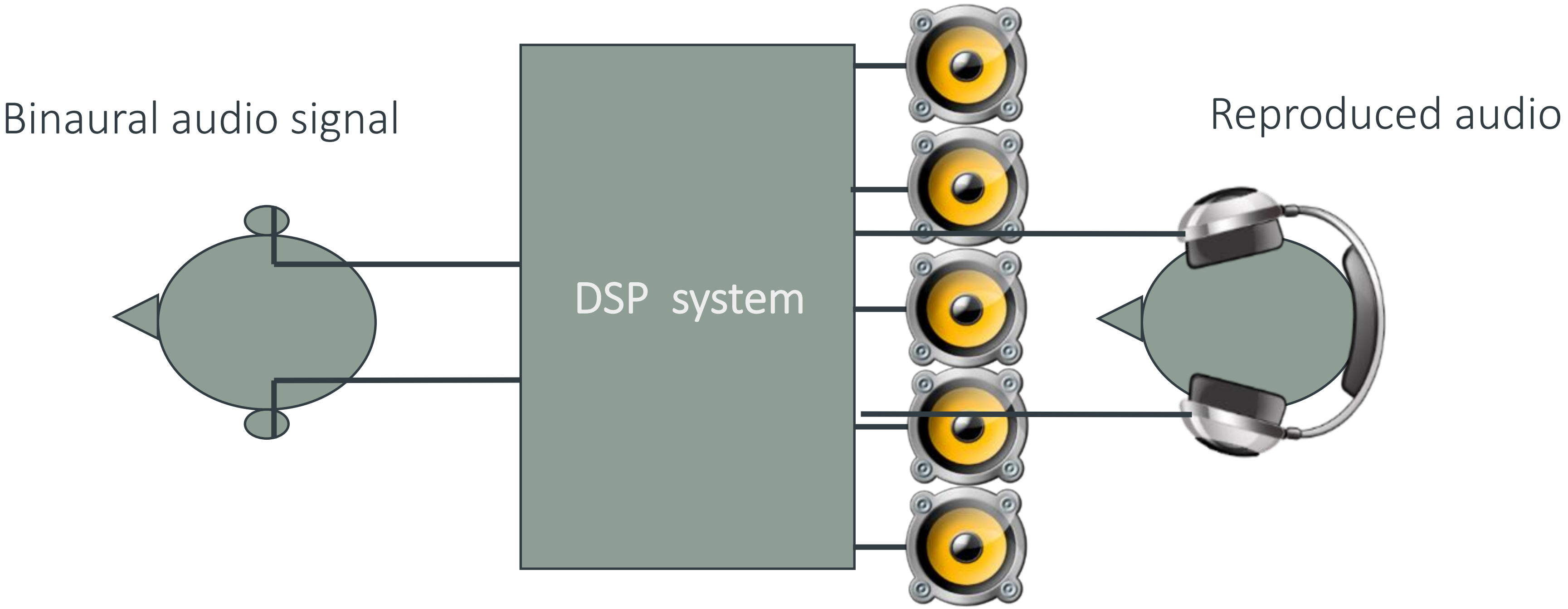
Zone of silence





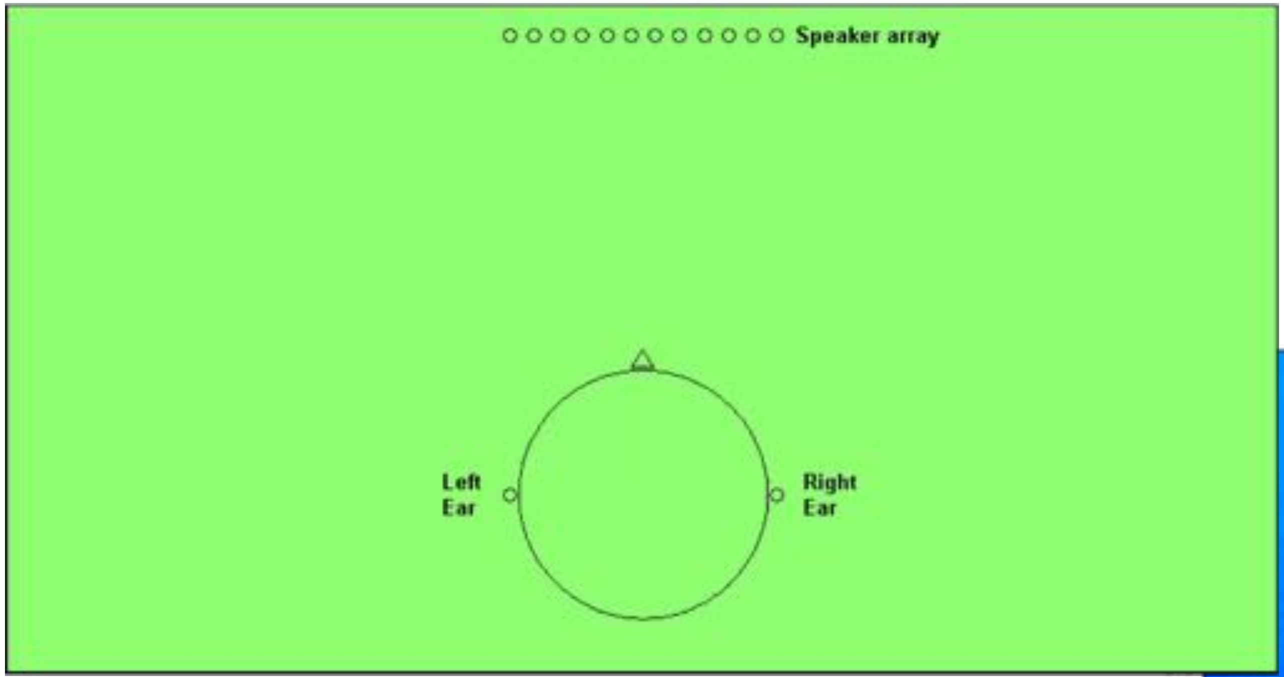
Poletti and Fazi, 2015. “An approach to generating two zones of silence with application to personal sound systems”. JASA

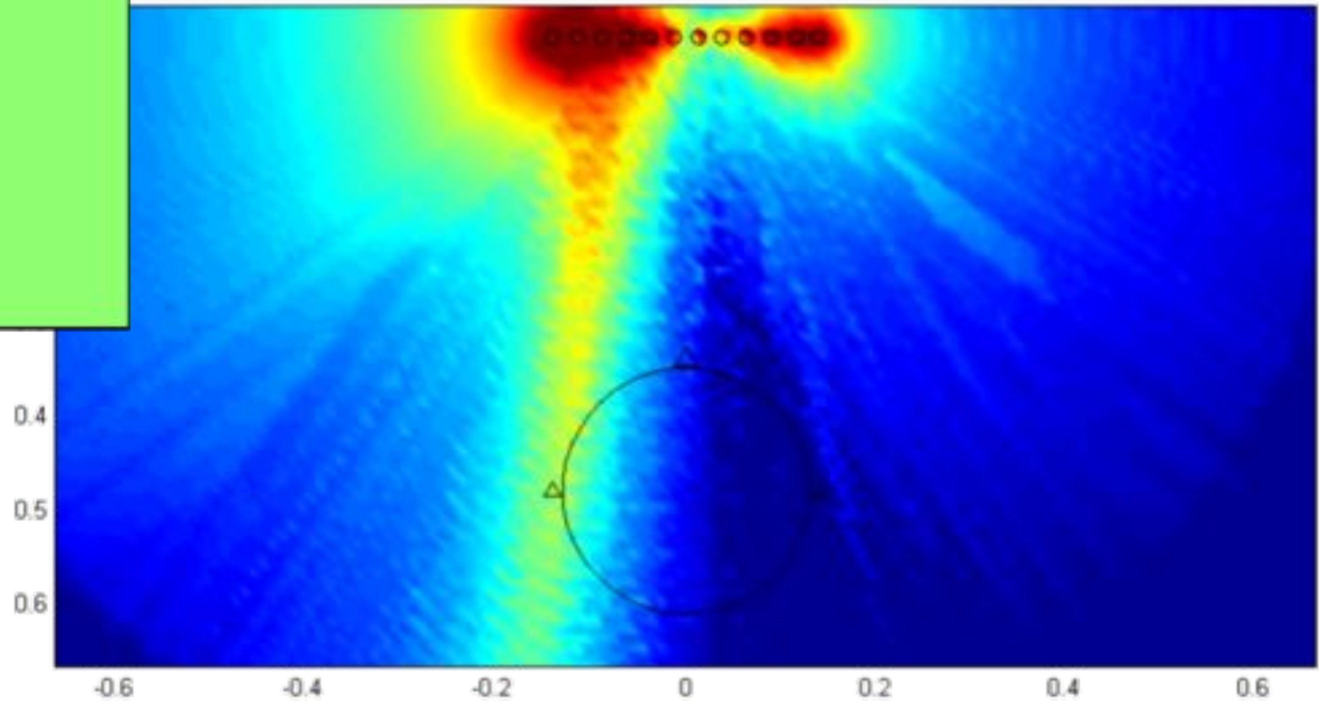
Cross-Talk Cancellation



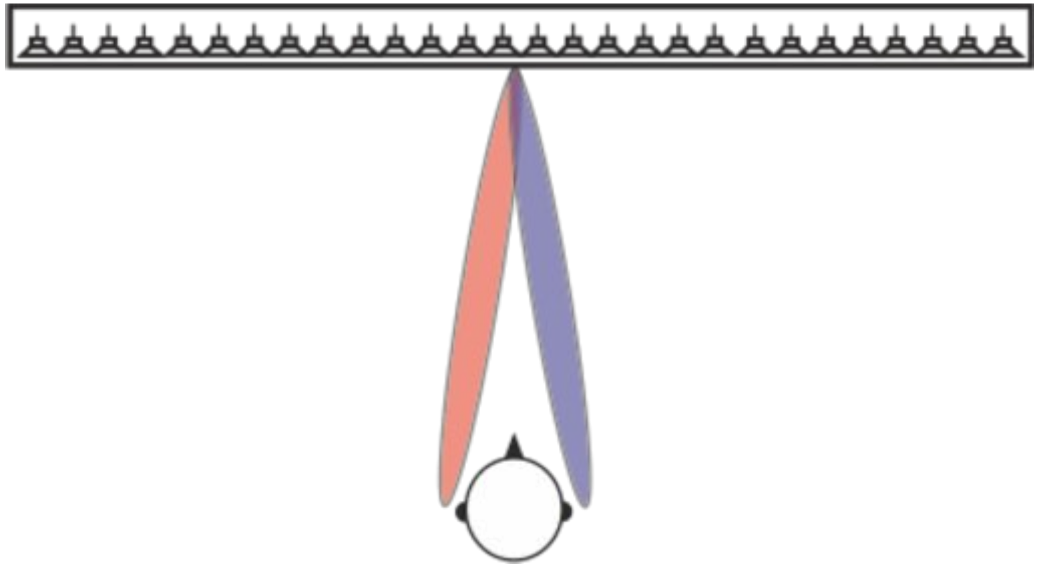
Cross-Talk Cancellation allows for the delivery of independent signals to the two ears of the listener

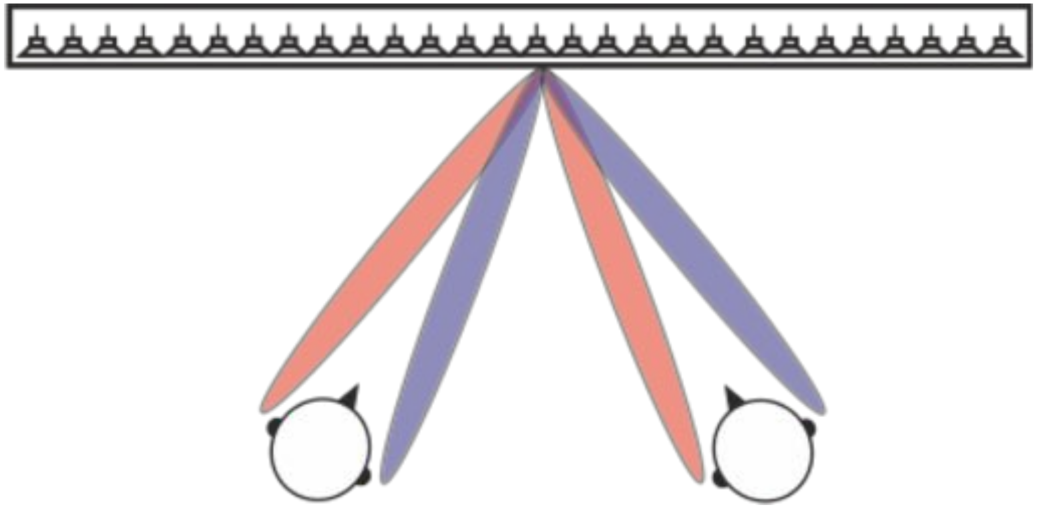
Cross-Talk Cancellation with a loudspeaker array





Multi-listener Cross-Talk Cancellation

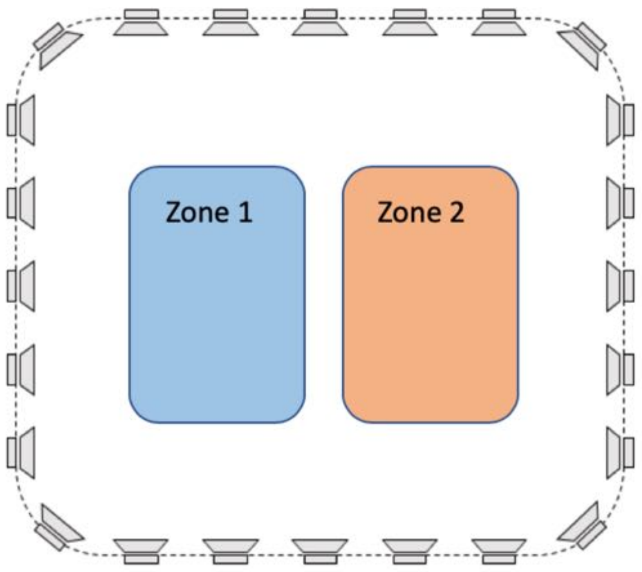


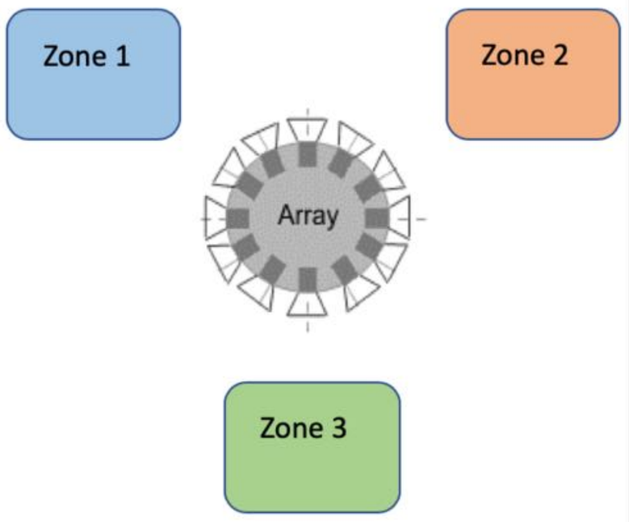


Putting it all together: summary & conclusions

Formulation

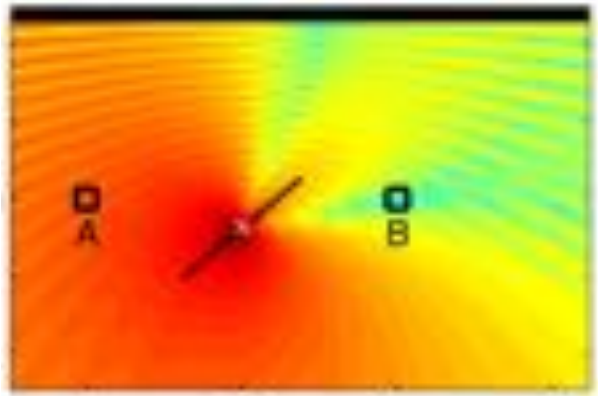
- Why? For convenient shared experiences
- Design choices
  - Zones, target experience, system setup, calibration & evaluation
- Key approaches:
  - pressure matching (PM), acoustic contrast control (ACC)
- Performance
  - Goodness of fit, contract, effort, bandwidth, planarity





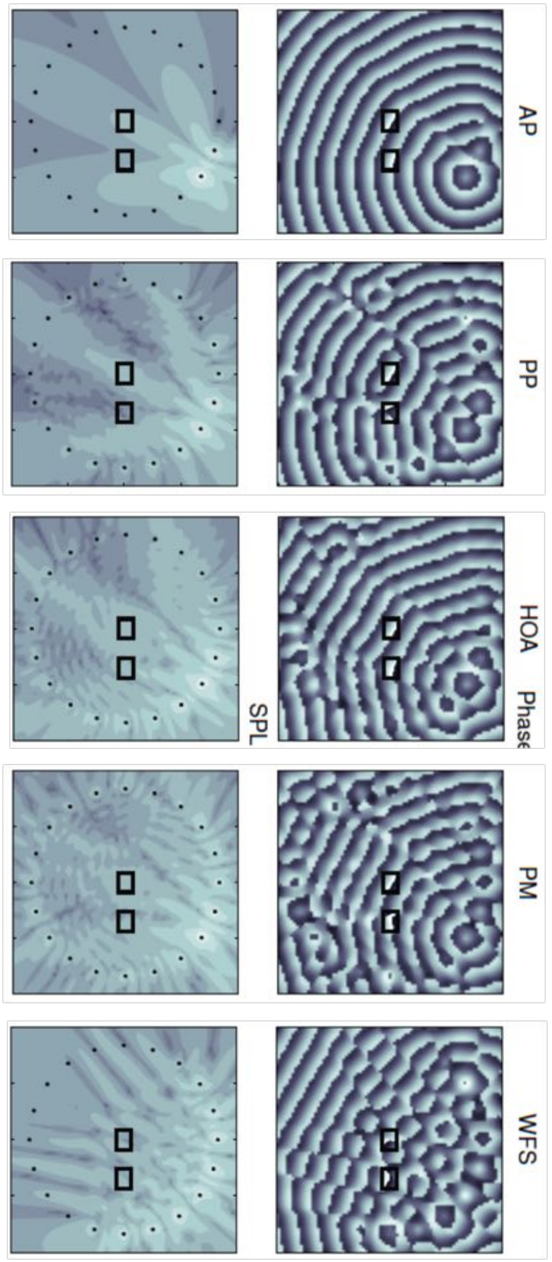
Engineering

- Design-performance axioms
  - Contrast vs. target quality (ACC vs. PM), # sources → bandwidth
- Robust design
  - Frequency-dependent regularization, conditioning/effort/error-model methods, performance-robustness trade off
- Room effects
  - Design & calibrate for the earlies
  - Truncate & regularize for the late
- PyZones interactive sound zone filter design



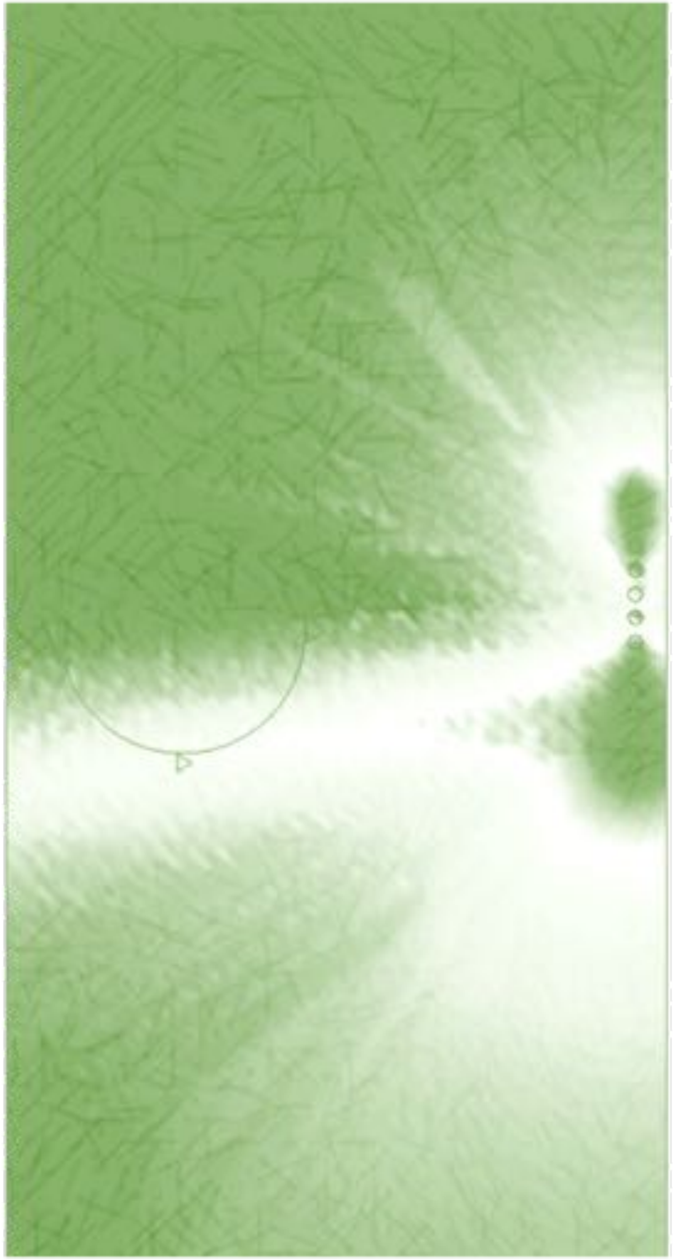
User experience

- Line-array sound zone listening demonstration
- Perceptual models of performance
  - Sound quality
  - Interference
- Alternative approaches
  - Perceptually-inspired planarity control (PC)
  - Spherical harmonics domain methods



Challenges & opportunities

- Perceptual models
  - Timbral & spatial target quality
  - Distraction & interference
  - Roles of context & of listener goals (attention)
- Adaptive systems
  - Moving listeners
  - Large & dynamic setups
  - Slick & robust calibration and synchronisation
- Paradigm
  - General framework for multi-listener spatial sound
  - Deployment of personalisable content



ICASSP Tutorial: Personalising sound over loudspeakers

Dr Philip Jackson, p.jackson@surrey.ac.uk, CVSSP, U. Surrey, UK

Dr Filippo Fazi, filippo.fazi@soton.ac.uk, ISVR, U. Southampton, UK

Dr Phil Coleman, p.d.coleman@surrey.ac.uk, IoSR/CVSSP, U. Surrey, UK

Thank you for your engagement!

Any questions?

S3A: Future Spatial Audio for an Immersive Listener Experience at Home
EPSRC Programme Grant (EP/L000539/1) and BBC R&D Audio Research Partnership

References

André, C., Corteel, E., Embrechts, J., Verly, J., and Katz, B.F.G., 2014. Subjective evaluation of the audiovisual spatial congruence in the case of stereoscopic-3D video and Wave Field Synthesis. International Journal of Human-Computer Studies, Elsevier, 72(1):23-32

Baykaner, K., Coleman, P., Mason, R., Jackson, P.J., Francombe, J., Olik, M. and Bech, S., 2015. The relationship between target quality and interference in sound zone. *Journal of the Audio Engineering Society*, *63*(1/2):78-89

Blauert, J., 1997. Spatial Hearing: The Psychophysics of Human Sound Localization, MIT Press: Cambridge, MA

Bleidt, R., Borsum, A., Fuchs, H., and Weiss, S.M., 2014. Object-based audio: Opportunities for improved listening experience and increased listener involvement, SMPTE Conference

Chang, J.-H., Lee, C.-H., Park, J.-Y., and Kim, Y.-H., 2009. A realization of sound focused personal audio system using acoustic contrast control. The Journal of the Acoustical Society of America, 125(4):2091-2097

Choi, J.W. and Kim Y.H., 2002. Generation of an acoustically bright zone with an illuminated region using multiple sources. *The Journal of the Acoustical Society of America,* 111:1695-1700

Coleman, P., Jackson, P.J., Olik, M., Olsen, M., Møller, M. and Pedersen, J.A., 2013, June. The influence of regularization on anechoic performance and robustness of sound zone methods. In *Proceedings of Meetings on Acoustics ICA2013.*

Coleman P., 2014. *Loudspeaker array processing for personal sound zone reproduction.* PhD thesis, Univ. of Surrey, UK

Coleman, P.D., Franck, A., Francombe, J., et al., 2018. An audio-visual system for object-based audio: from recording to listening, IEEE Transactions on Multimedia, 20(8):1919-1931

Coleman, P., and Jackson, P.J.B., 2014. Planarity panning for listener-centered spatial audio, Audio Engineering Society 55th International Conference on Spatial Audio, Helsinki

Coleman, P., and Jackson, P.J.B., 2016. Planarity-based sound field optimization for multi-listener spatial audio, Audio Engineering Society International Conference on Sound Field Control, Guildford, UK

Coleman, P., Jackson, P.J., Olik, M., Møller, M., Olsen, M. and Abildgaard Pedersen, J., 2014. Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array. *The Journal of the Acoustical Society of America*, *135*(4): 1929-1940

Coleman, P., Jackson, P.J., Olik, M. and Abildgaard Pedersen, J., 2014. Personal audio with a planar bright zone. *The Journal of the Acoustical Society of America*, *136*(4): 1725-1735

Coleman, P., Jackson, P.J.B., Olik, M., and Pedersen, J.A., 2014. Stereophonic personal audio reproduction using planarity control optimization (invited), 21st International Conference on Sound & Vibration, Beijing

Conetta, R., Brookes, T., Rumsey, F., et al., 2014. Spatial Audio Quality Perception (Part 2): A linear regression model, Journal of the Audio Engineering Society, 62(12):847-860

Doclo, S., and Moonen, M., 2003. Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics. *IEEE Transactions on Signal Processing*, *51*(10):2511-2526

Elliott, S.J., Cheer, J., Choi, J.W. and Kim, Y., 2012. Robustness and regularization of personal audio systems. *IEEE Transactions on Audio, Speech, and Language Processing*, *20*(7):2123-2133

Elliott, S.J., Cheer, J., Murfet, H., and Holland, K.R., 2010. Minimally radiating sources for personal audio. The Journal of the Acoustical Society of America, 128(4):1721-1728

Emiya, V., Vincent, E., Harlander, N., and Hohmann, V., 2011. Subjective and objective quality assessment of audio source separation, IEEE Transactions on Audio, Speech, and Language Processing, 19(7): 2046–2057

Fazi, F.M., 2010. Sound field reproduction. PhD thesis , Univ. of Southampton, UK

References (continued)

Francombe, J., 2014. *Perceptual evaluation of audio-on-audio interference in a personal sound zone system.* PhD thesis, Univ. of Surrey, UK

Francombe, J., Mason, R., Dewhirst, M. and Bech, S., 2014. Elicitation of attributes for the evaluation of audio-on-audio interference. *The Journal of the Acoustical Society of America*, *136*(5): 2630-2641

Francombe, J., Mason, R., Dewhirst, M. and Bech, S., 2015. A model of distraction in an audio-on-audio interference situation with music program material. *Journal of the Audio Engineering Society*, *63*(1/2): 63-77

George, S., Zielinski, S., Rumsey, F., et al., 2010. Development and validation of an unintrusive model for predicting the sensation of envelopment arising from surround sound recordings, Journal of the Audio Engineering Society, 58(12): 1013-1031

Hamasaki, K., et al., 2008. A 22.2 Multichannel Sound System for Ultrahigh-Definition TV (UHDTV), SMPTE Motion Imaging Journal

Herre, J., Hilpert, J., Kuntz, A., and Plogsties, J., 2015. MPEG-H 3D Audio - The new standard for coding of immersive spatial audio, IEEE J. Sel. Topics Signal Process., 9(5): 770–779; doi:10.1109/JSTSP.2015.2411578

Hollebon, J., Simón Gálvez, M.F., and Fazi, F.M., 2019. Multiple listener crosstalk cancellation using linear loudspeaker arrays for binaural cinematic audio, Audio Engineering Society International Conference on Immersive and Interactive Audio, York, UK

ITU, 2015. ITU-R BS.2076-0: Audio Definition Model, Recommendation, Geneva, Switzerland

Jackson, P.J.B., Dewhirst, M., Conetta, R., et al., 2008. QESTRAL (Part 3): System and metrics for spatial quality prediction, 125th Audio Engineering Society Conv., San Francisco CA, Preprint 7597

Jackson, P.J.B., Jacobsen, F., Coleman, P. and Pedersen, J.A., 2013. Sound field planarity characterized by superdirective beamforming. Proc. Of Mtgs. Acoust. 19, 055056; doi: 10.1121/1.4800877

Jackson, P.J.B., and Ross, C.F., 1996. Application of active noise control to corporate aircraft. Amer. Soc. Mech. Eng. (ASME 1996), DE93: 19-24, Atlanta, GA

Jones, M., and Elliott, S.J., 2008. Personal audio with multiple dark zones. The Journal of the Acoustical Society of America, 124(6): 3497-3506

Kinsler, L.E., Frey, A.R., Coppens, A.B., and Sanders, J.V., 2000. Fundamentals of Acoustics, John Wiley & Sons, 4th ed., ISBN-10: 0471847895

Kirkeby, O., and Nelson, P.A., 1993. Reproduction of plane wave sound fields. *The Journal of the Acoustical Society of America*, 94(5): 2992-3000

Kuttruff, H., 2009. Room Acoustics, Spon Press, Abingdon, UK, 5th ed.

Merimaa, J., 2007. Energetic sound field analysis of loudspeaker reproduction. Audio Engineering Society Conv., New York

Møller, M.B., and Olsen, M., 2011. Sound zones, MSc thesis, Tech. Univ. Denmark

Møller, M.B., Olsen, M., and Jacobsen, F., 2012. A hybrid method combining synthesis of a sound field and control of acoustic contrast, AES Conv., paper 8627

Møller, M.B., and Olsen, M., 2016. Sound zones: On performance prediction of contrast control methods, Audio Engineering Society International Conference on Sound Field Control, Guildford, UK

Moore, B.C.J., 2004. An Introduction to the Psychology of Hearing, 5th ed. Elsevier

Morse, P.M., and Ingard, K.U., 1987. Theoretical Acoustics, Princeton Univ. Press, ISBN 0691024014

Nelson, P.A., and Elliott, S.J., 1993. Active Control of Sound, Academic Press; ISBN 0125154267

Olik, M., 2014. Personal sound zone reproduction with room reflections. PhD thesis, Univ. of Surrey, UK

References (continued)

Olik, M., Francombe, J., Jackson, P.J.B., et al., 2013. A comparative performance study of sound zoning methods in a reflective environment, Audio Engineering Society 52nd International Conference on Sound Field Control, pp. 218-227, Guildford, UK

Olik, M., Jackson, P.J.B., Coleman, P., and Pedersen, J.A., 2014. Optimum source geometry for sound zone reproduction with a single reflection, The Journal of the Acoustical Society of America, 136(6): 3085-3096

Poletti, M.A., and Betlehem, T., 2014. Creation of a single sound field for multiple listeners. *INTER-NOISE and NOISE-CON Congress and Conference Proceedings,* 249(6): 2033-2040.

Poletti, M.A., and Fazi, F.M., 2015. An approach to generating two zones of silence with application to personal sound systems. *The Journal of the Acoustical Society of America,* 137(2): 598-605

Rämö, J., Bech, S. and Jensen, S.H., 2017. Real-time perceptual model for distraction in interfering audio-on-audio scenarios. *IEEE Signal Processing Letters*, *24*(10): 1448-1452.

Rämö, J., Bech, S. and Jensen, S.H., 2018. Validating a real-time perceptual model predicting distraction caused by audio-on-audio interference. *The Journal of the Acoustical Society of America*, *144*(1): 153-163.

Rumsey, F., 2001. Spatial audio, Focal Press; ISBN 0240516230

Shin, M., Lee, S.Q. , Fazi, F.M., Nelson, P.A., Kim, D., Wang, S., Park, K.H., Seo, J., 2010. Maximization of acoustic energy difference between two spaces. *The Journal of the Acoustical Society of America,* 128(1): 121-131

Shin, M., Fazi, F.M., Nelson, P.A. and Hirono, F.C., 2014. Controlled sound field with a dual layer loudspeaker array. *Journal of Sound and Vibration*, *333*(16): 3794-3817

Simón Gálvez, M.F., Elliott, S.J., and Cheer, J., 2012. A superdirective array of phase shift sources. The Journal of the Acoustical Society of America, 132(2): 746-756

Thresh, L., and Kearney, G.C., 2017. A direct comparison of localisation performance when using first, third and fifth order ambisonics for real loudspeaker and virtual loudspeaker rendering, 143rd AES Conv.

Williams, E.G., 1999, Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography. Academic Press, ISBN 0127539603

Winter, F., Schultz, F., Firtha, G. and Spors, S., 2019. A Geometric Model for Prediction of Spatial Aliasing in 2.5 D Sound Field Synthesis. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*.

Wu, J., and Abhayapala, T.D., 2011. Spatial Multizone Soundfield Reproduction: Theory and Design. *IEEE Transactions on Audio, Speech, and Language Processing,* 19(6): 1711-1720

Zhu, Q., Coleman, P., Wu, M. and Yang, J., 2017. *Robust reproduction of sound zones with local sound orientation.* The Journal of the Acoustical Society of America, 142(1): EL118-EL122

Zhu, Q., Coleman, P., Wu, M. and Yang, J., 2017. *Robust acoustic contrast control with reduced in-situ measurement by acoustic modelling.* Journal of the Audio Engineering Society, 65(6): 460-473

Zhu, Q., Coleman, P., Qiu, X., Wu, M., Yang, J., and Burnett, I., 2018. Robust personal audio geometry optimization in the SVD-based modal domain, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27(3): 610-620