MuDiS: An Audio-independent, Wide-angle, and Leak-free Multi-directional Speaker

MobiCom 2024

SJTU

Directional Speaker Applications

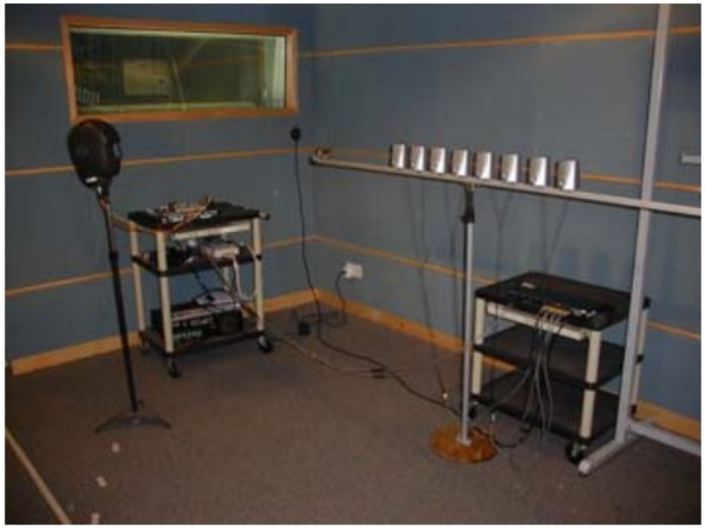
Museum

Gym

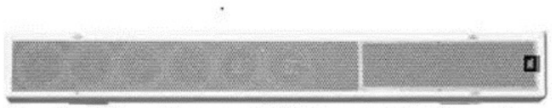
Advertisements

Directional speakers can direct sound waves only to specific areas, which can create customized sound zones for individuals

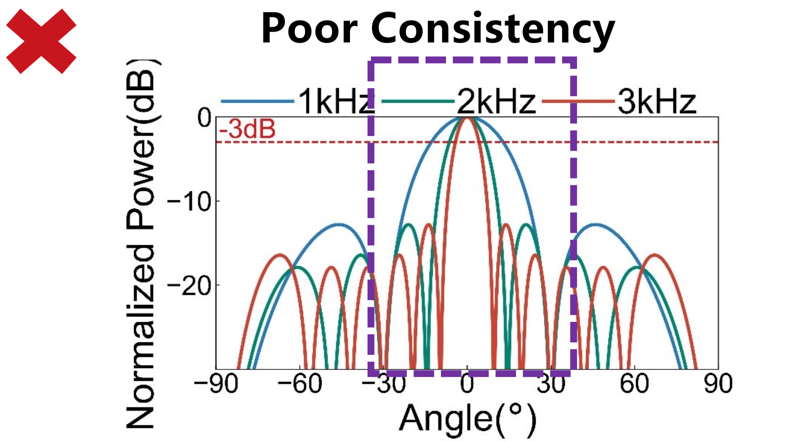
Existing Solutions – Speaker Phased Array

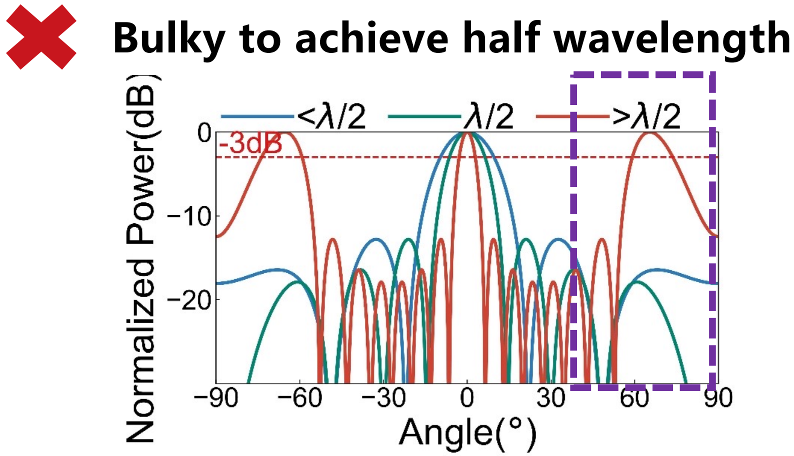


Conventional Speaker Phased Array



JBL Intellivox DS115





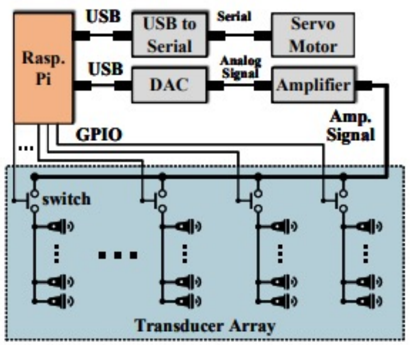
Parametric Array - Air Nonlinearity

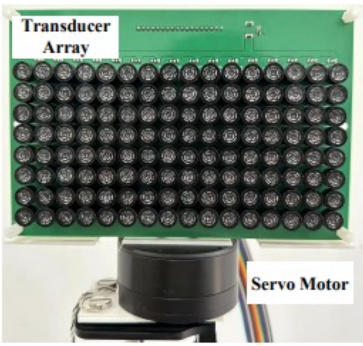
Air Nonlinearity

Primary High-Frequency Beam (f1,f2)

Air Nonlinearity can reproduce audible sounds from ultrasound.

Air Nonlinearity ➕ **Phased Array**





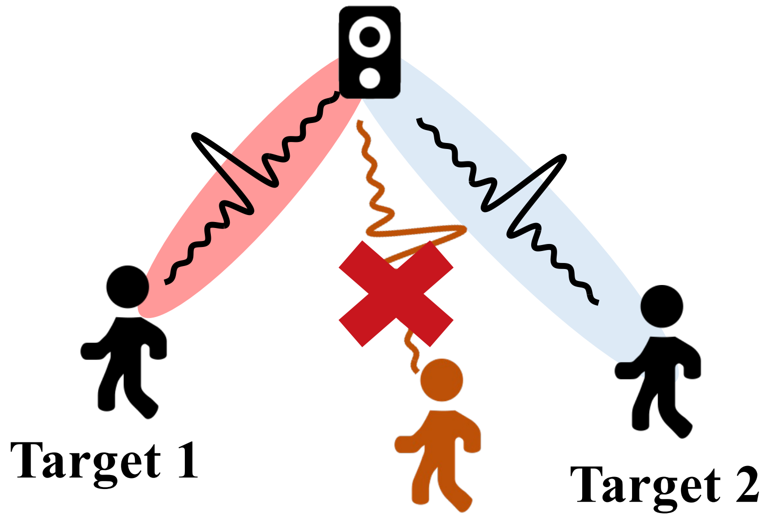
✓
- Compact Size
- Better consistency

✗
- Cannot support Multi-beam
- Mechanical steering

[Mobicom 24'] An Audio-independent, Wide-angle, and Leak-free Multi-directional Speaker

MuDiS – Multi-Directional Speaker

Multi-Directional Speaker



Multi-Beamforming

Design Goal

- **Independent Beams:**
  Diverse content must be concurrently delivered in distinct directions.
- **Wide-angle Steering:**
  The beam must be dynamically adjustable to encompass a broad coverage
- **Leakage suppression:**
  The sound should exclusively appear in the desired direction, preventing sound leakage in unintended directions.

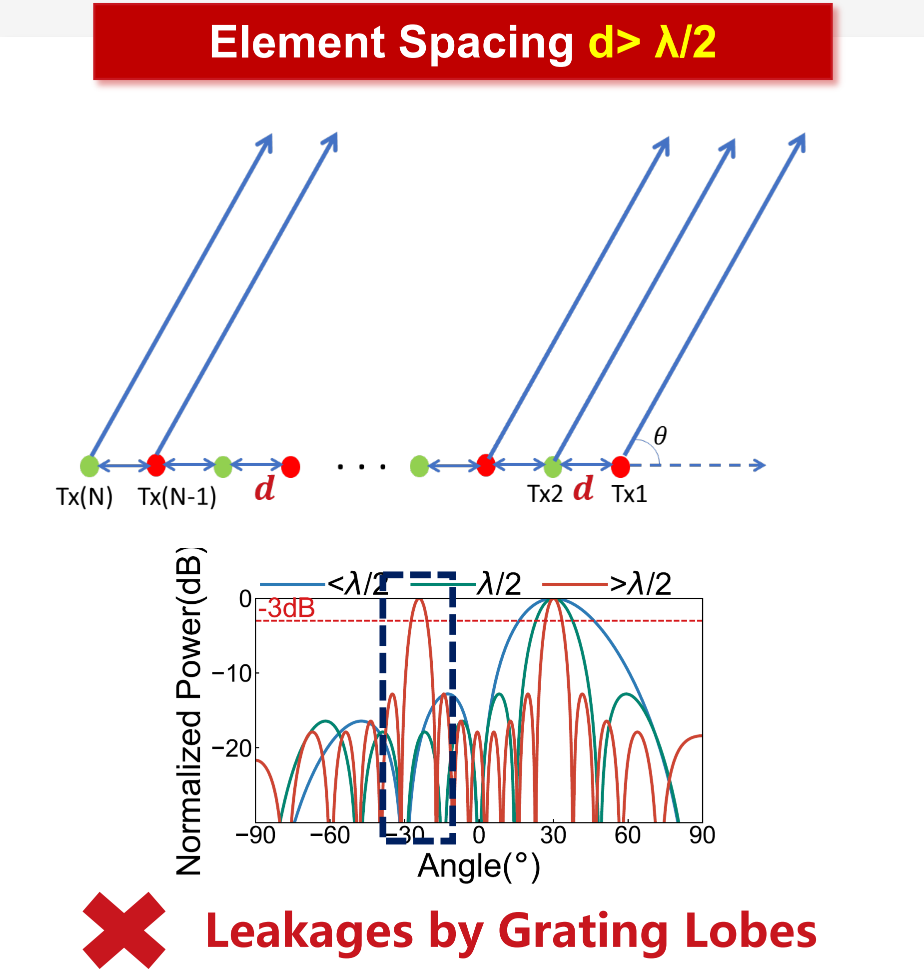
[Mobicom 24'] An Audio-independent, Wide-angle, and Leak-free Multi-directional Speaker

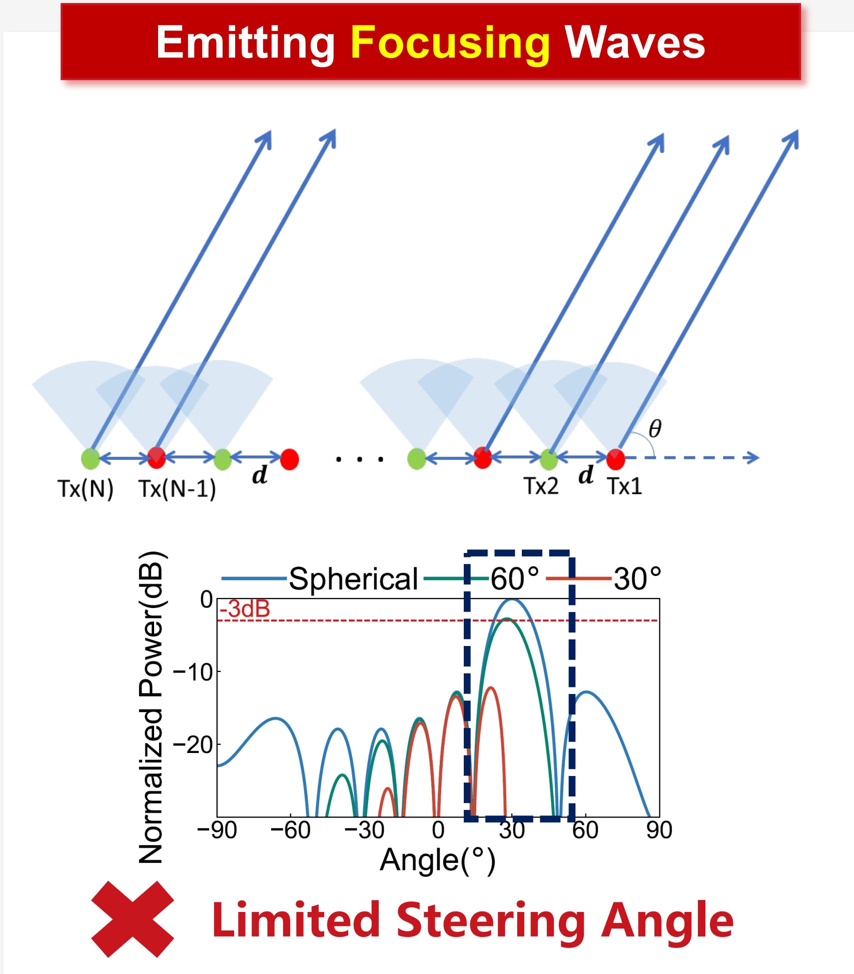
Challenges

1) **Physical Limitation of Transducer**

2) **Additional Leakage due to Nonlinear Effect**

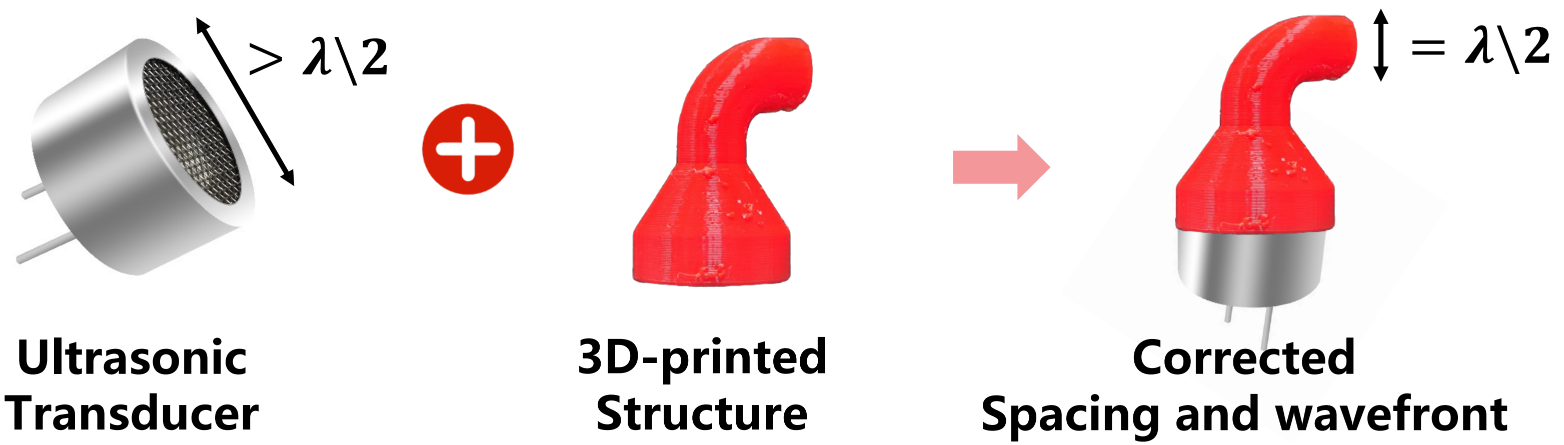
Challenge 1: Physical Limitations of Transducer





[Mobicom 24'] An Audio-independent, Wide-angle, and Leak-free Multi-directional Speaker

Key idea – Design a 3D-printed Structure

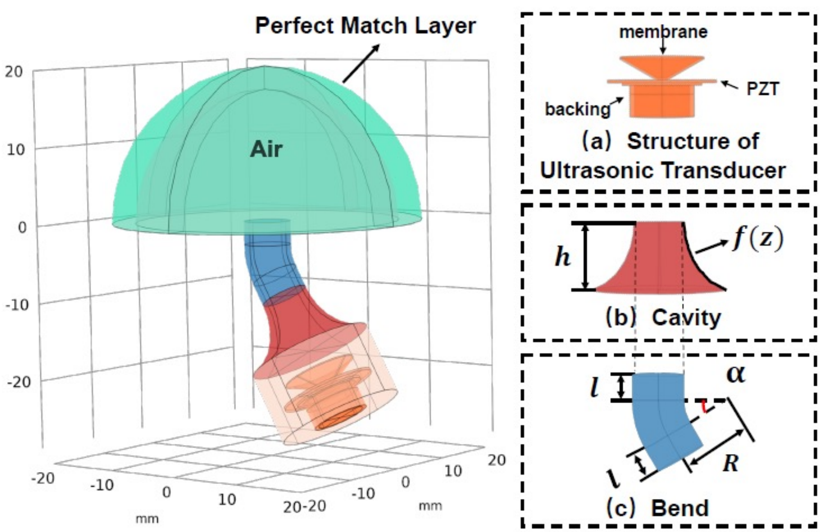


Design Goal

- Correct Element Spacing to $\lambda\backslash 2$
- Correct to approximate spherical wavefront
- Maximize the output energy

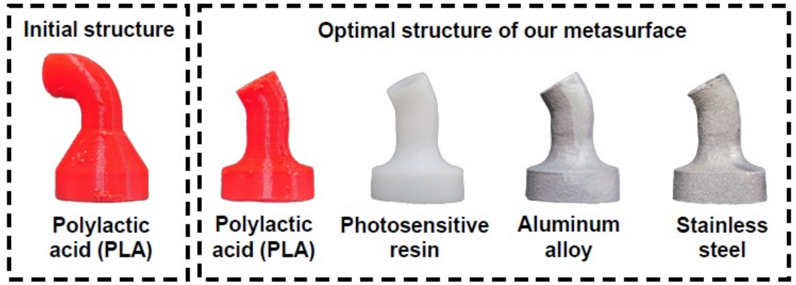
Empirical Acoustic Metasurface Design

Simulation in COMSOL

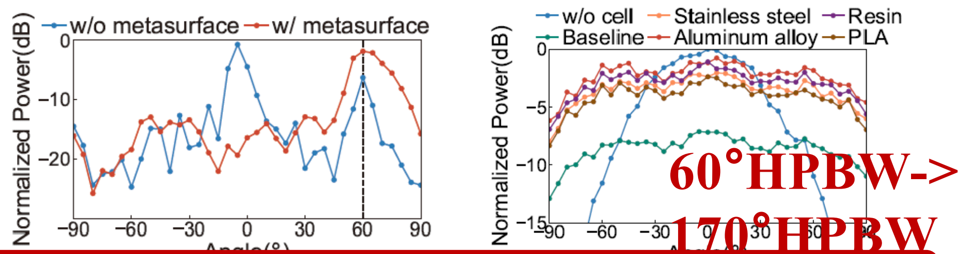


- Correct Element Spacing to $\lambda\backslash2$
- Correct to approximate **spherical** wavefront

Optimal 3D Printed Cells



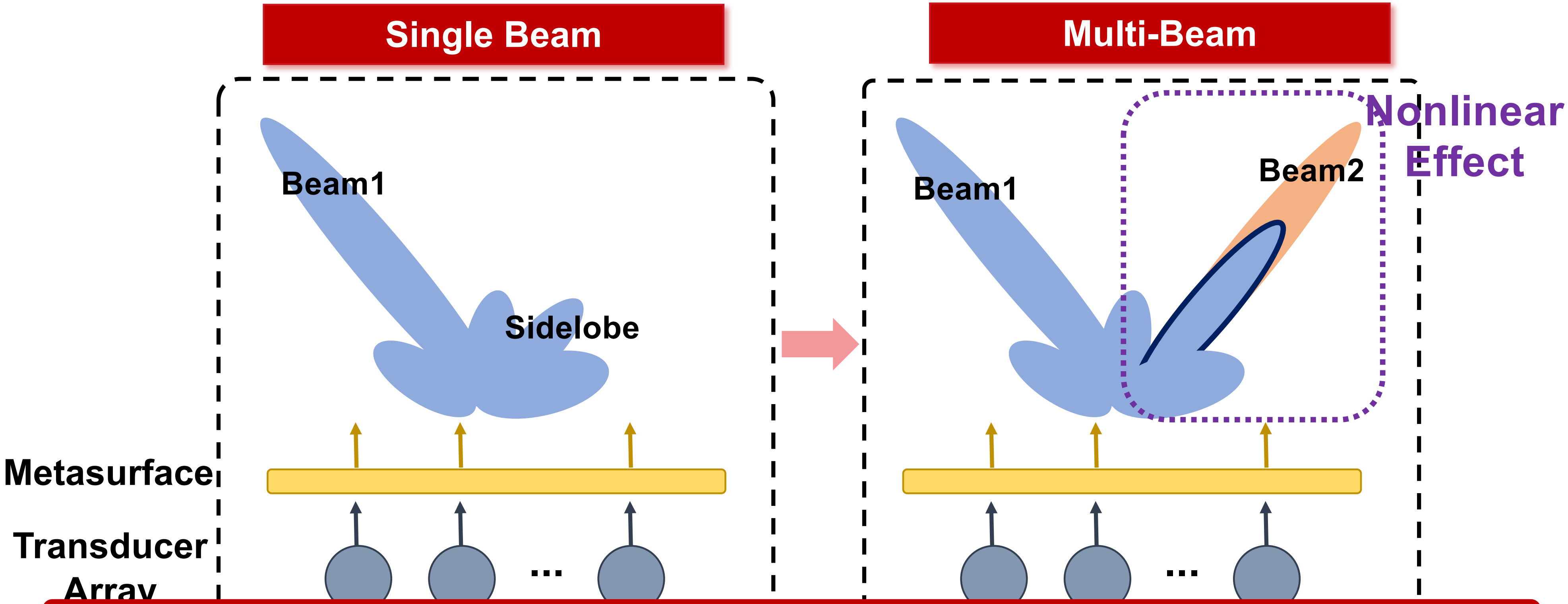
Performance



Metasurface helps correct element spacing and output wavefront

[Mobicom 24'] An Audio-independent, Wide-angle, and Leak-free Multi-directional Speaker

Challenge 2: Leakage due to Nonlinear Effect



Nonlinear Effect will introduce additional leakages

[Mobicom 24'] An Audio-independent, Wide-angle, and Leak-free Multi-directional Speaker

Challenge 2: Leakage due to Nonlinear Effect

Assume the original signal are $h_1(t)$, $h_2(t)$, the modulated signal will be:

$$s_1(t) = h_1(t)\cos(2\pi f_c t) + \cos(2\pi f_c t)$$

$$s_2(t) = h_2(t)\cos(2\pi f_c t) + \cos(2\pi f_c t)$$

According to pre-set beam weights, when the target directions are θ_1, θ_2, the output signal will be $s_1(t)w_{\theta_1} + s_2(t)w_{\theta_2}$

Then the received high-frequency signal in θ_2 is:

$$R_{High}^{Multi}(\theta_2) = A\sum\left(s_1(t)w_{\theta_1} + s_2(t)w_{\theta_2}\right)e^{j2\pi\frac{d}{\lambda_i}(n-1)sin\theta_2}$$
$$= An(\alpha s_1(t)e^{j\sigma} + s_2(t))$$

Now the audible sound in θ_2 will be

$$R_{Low}^{Multi}(\theta_2) = LPF\left(A_2(R_{High}^{Multi}(\theta_2)^2)\right)$$
$$= A_2A^2n^2LPF[s_2(t)^2 + 2\alpha s_2(t)s_1(t)e^{j\sigma} + \alpha^2\left(s_1(t)e^{j\sigma}\right)^2]$$

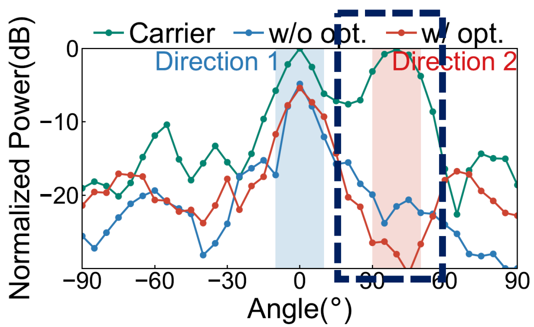
Expected **Leakages** Ignored

Challenge 2: Leakage due to Nonlinear Effect

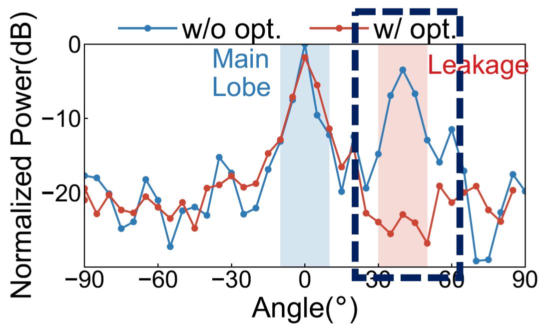
Assume the original signal are $h_1(t)$, $h_2(t)$, the modulated signal will be:

$$s_1(t) = h_1(t)\cos(2\pi f_c t) + \cos(2\pi f_c t)$$

Leakages in Unintended directions



High frequency beam pattern



Low frequency beam pattern

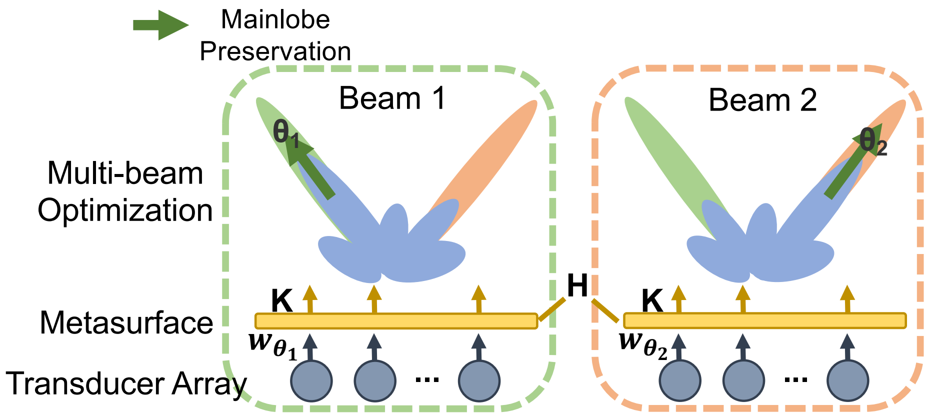
$$R_{Low}^{Multi}(\theta_2) = LPF\left(A_2\left(R_{High}^{Multi}(\theta_2)^2\right)\right)$$

$$= A_2 A^2 n^2 LPF\left[s_2(t)^2 + 2\alpha s_2(t) s_1(t) e^{j\sigma} + \alpha^2\left(s_1(t)e^{j\sigma}\right)^2\right]$$

Expected **Leakages** Ignored

Optimization-based Beamforming

Problem Formulation



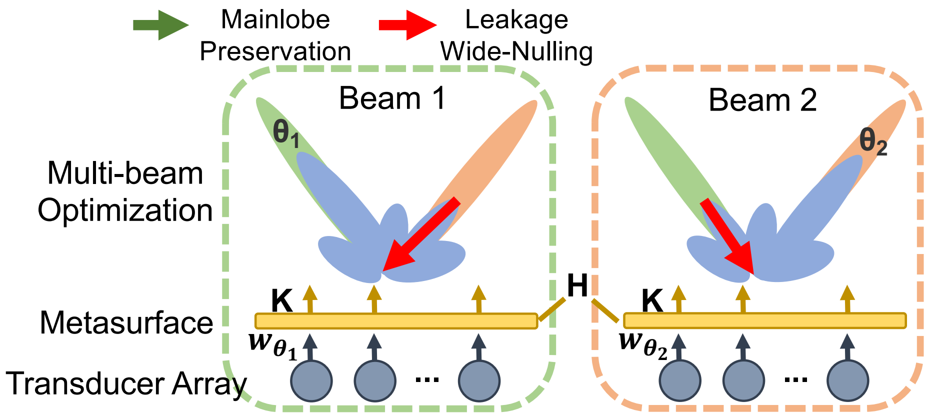
- **Mainlobe preservation**

$$L_{main} = \sum P \times mask([\theta - \alpha, \theta + \alpha])$$

[Mobicom 24'] An Audio-independent, Wide-angle, and Leak-free Multi-directional Speaker

Optimization-based Beamforming

Problem Formulation



- **Mainlobe preservation**

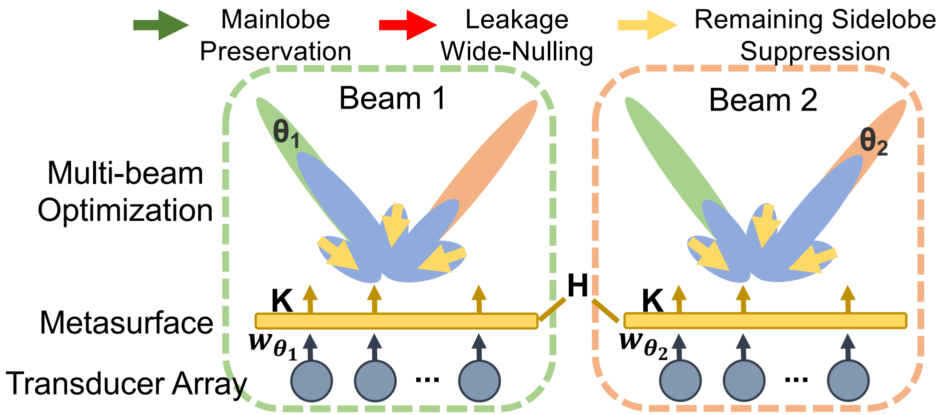
$$L_{main} = \sum P \times mask([\theta - \alpha, \theta + \alpha])$$

- **Leakage wide-nulling**

$$L_{null} = \sum P \times mask([\Theta_l - \beta, \Theta_l + \beta])$$

Optimization-based Beamforming

Problem Formulation



- **Mainlobe preservation**

$$L_{main} = \sum P \times mask([\theta - \alpha, \theta + \alpha])$$

- **Leakage wide-nulling**

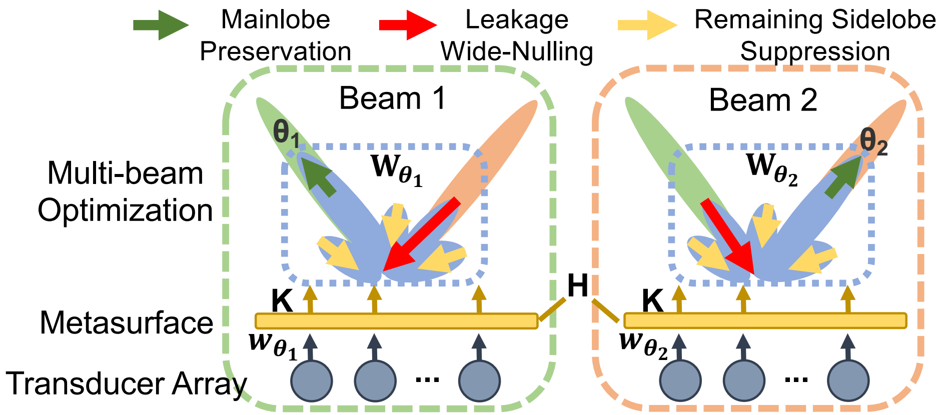
$$L_{null} = \sum P \times mask([\Theta_l - \beta, \Theta_l + \beta])$$

- **Remain sidelobe suppression**

$$L_{remain} = \sum peak(P \times mask(\Theta_m))$$

Optimization-based Beamforming

Problem Formulation



- **Mainlobe preservation**

$$L_{main} = \sum P \times mask([\theta - \alpha, \theta + \alpha])$$

- **Leakage wide-nulling**

$$L_{null} = \sum P \times mask([\Theta_l - \beta, \Theta_l + \beta])$$

- **Remain sidelobe suppression**

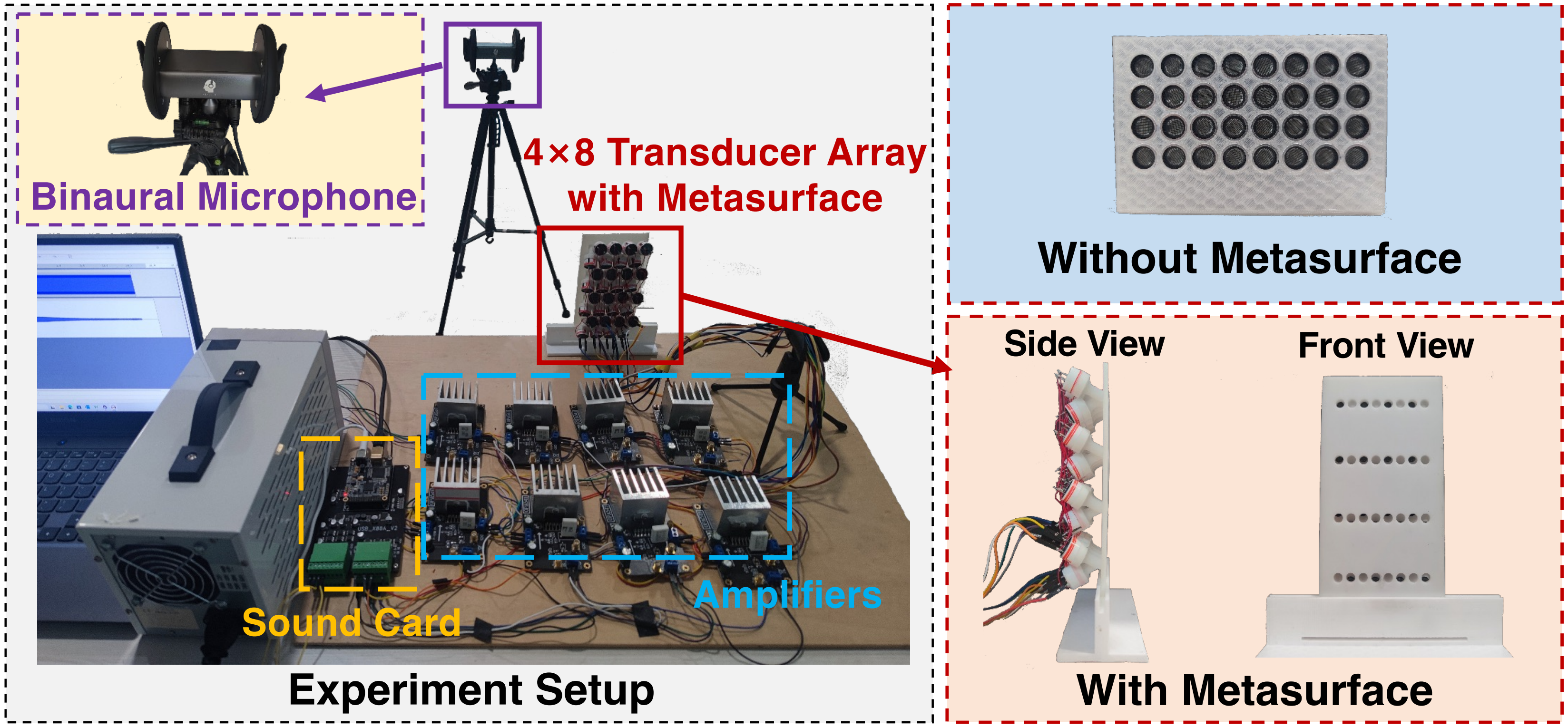
$$L_{remain} = \sum peak(P \times mask(\Theta_m))$$

Objective Function

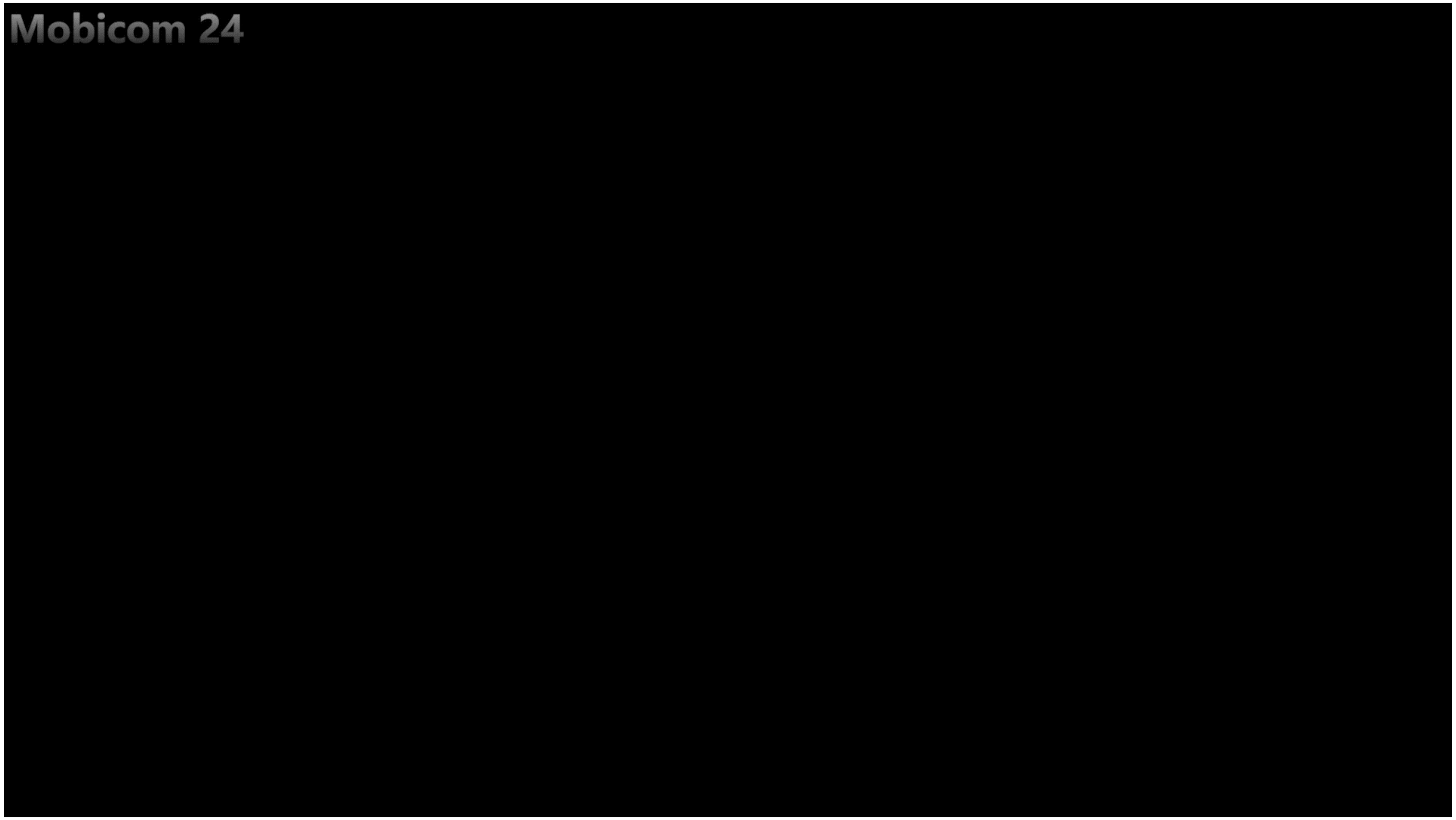
$$\max_{\mathbf{w}_\theta} L_{main} - \delta L_{null} - \eta L_{remain}$$

$$s.t. \ |\mathbf{w}_\theta| \leq 1 \quad (\theta \in \Theta)$$

Experiment Setup



Demo



https://acousticmudis.github.io/

Performance

MCD: Mel-cepstral Distance **SWPR:** Signal-to-Worst-Point Ratio

Table 1: The performance comparison of MuDiS under different schemes.

		Beam Num.	Methods					Metrics			
			Beam-forming	Meta-surface	Leakage Cancellation	Phase Robust	Distortion Reduction	SNR	PESQ	MCD	SWPR
MuDiS	2-beams	2	✔	✔	✔	✔	✔	**6.56**	**3.25**	**4.31**	**10.97**
	3-beams	3	✔	✔	✔	✔	✔	5.85	2.88	5.26	8.62
	4-beams	4	✔	✔	✔	✔	✔	5.66	2.67	6.03	6.93
	1-beam	1	✔	✔	✘	✘	✔	6.96	3.42	4.14	18.95
COTS Directional Speaker		1	✔	-	-	-	-	6.07	3.40	4.70	15.37
Smartphone Speaker		-	✘	-	-	-	-	8.52	4.02	3.04	0.86

The system can support 4 beams playing simultaneously

Performance

MCD: Mel-cepstral Distance **SWPR:** Signal-to-Worst-Point Ratio

Table 1: The performance comparison of MUDIS under different schemes.

		Beam Num.	Methods					Metrics			
			Beam-forming	Meta-surface	Leakage Cancellation	Phase Robust	Distortion Reduction	SNR	PESQ	MCD	SWPR
MUDIS	2-beams	2	✔	✔	✔	✔	✔	**6.56**	**3.25**	**4.31**	**10.97**
	3-beams	3	✔	✔	✔	✔	✔	5.85	2.88	5.26	8.62
	4-beams	4	✔	✔	✔	✔	✔	5.66	2.67	6.03	6.93
	1-beam	1	✔	✔	✘	✘	✔	6.96	3.42	4.14	18.95
COTS Directional Speaker		1	✔	-	-	-	-	6.07	3.40	4.70	15.37
Smartphone Speaker		-	✘	-	-	-	-	8.52	4.02	3.04	0.86

Close Performance to COTS directional speaker when playing multiple beams

Conclusion

- Implemented a **muti-directional speaker**, facilitating **broad angles** for delivering diverse audio content along various directions
- Designed a metasurface to **overcome physical limitation** of ultrasonic transducers
- Proposed an optimization-based beamforming scheme to **suppress unintended leakages**

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

Thank You

饮 水 思 源　爱 国 荣 校

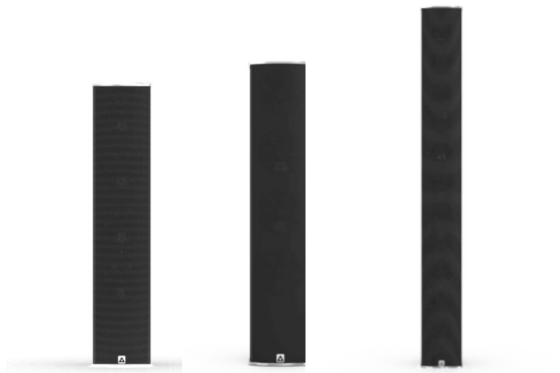
Other COTS Directional Speaker

Sonos Play + RocketFish Mount

❌ Mechanical Steering

SXS Multi Directional Audio Pods

❌ Limited Steering Angle (forward and back)

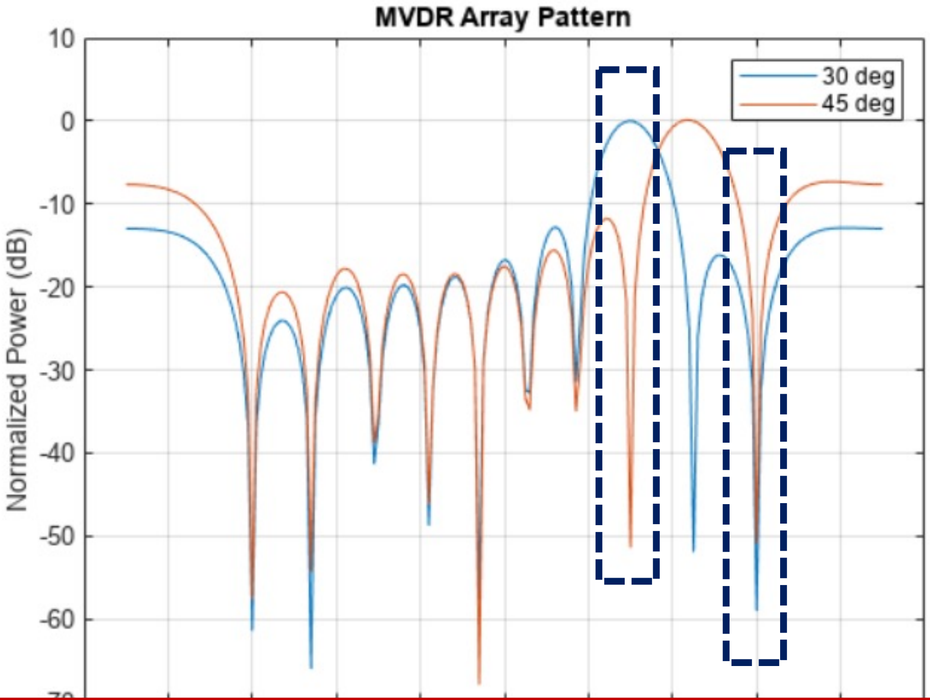


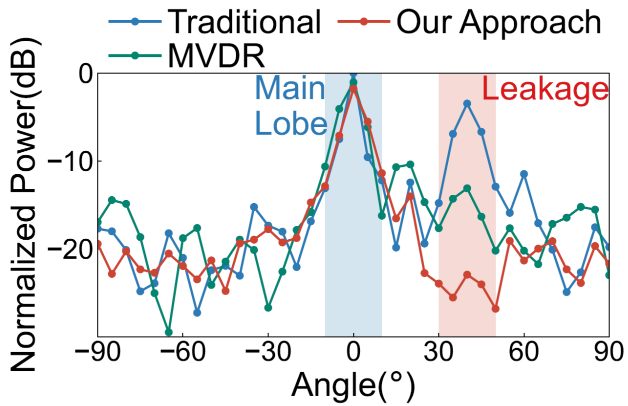
PAN Acoustics

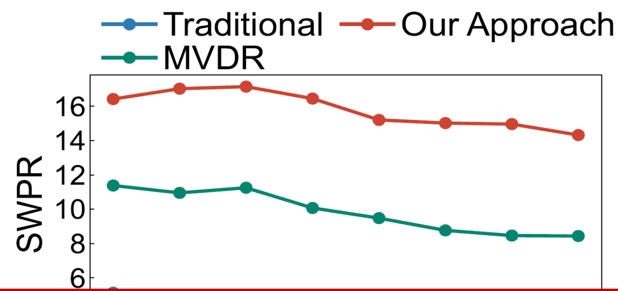
❌ Identical Audio Beams

Minimum Variance Distortionless Response

Minimum Variance Distortionless Response (MVDR)







MVDR can only suppress one angle point, thus the performance is worse

[Mobicom 24'] An Audio-independent, Wide-angle, and Leak-free Multi-directional Speaker

Metrics

- **SNR:** Signal-to-Noise Ratio

$$SNR(x,\hat{x}) = 10log_{10}(\frac{||x||_2^2}{||x-\hat{x}||_2^2})$$

- **PESQ:** Perceptual evaluation of speech quality

  assess speech quality and generate a score ranging from -0.5 to 4.5, which is based on s psychoacoustic model that simulates human auditory system's response

- **MCD:** Mel-cepstral Distance

  measures the similarity between sounds accounting for the ear's critical bandwidth as a function of frequency

- **SWPR:** Signal-to-Worst-Point Ratio

  assess the maximum interference in other directions apart from the intended one. It is defined as the difference between the peak of the main lobe and the peak in other directions.

Overall Performance

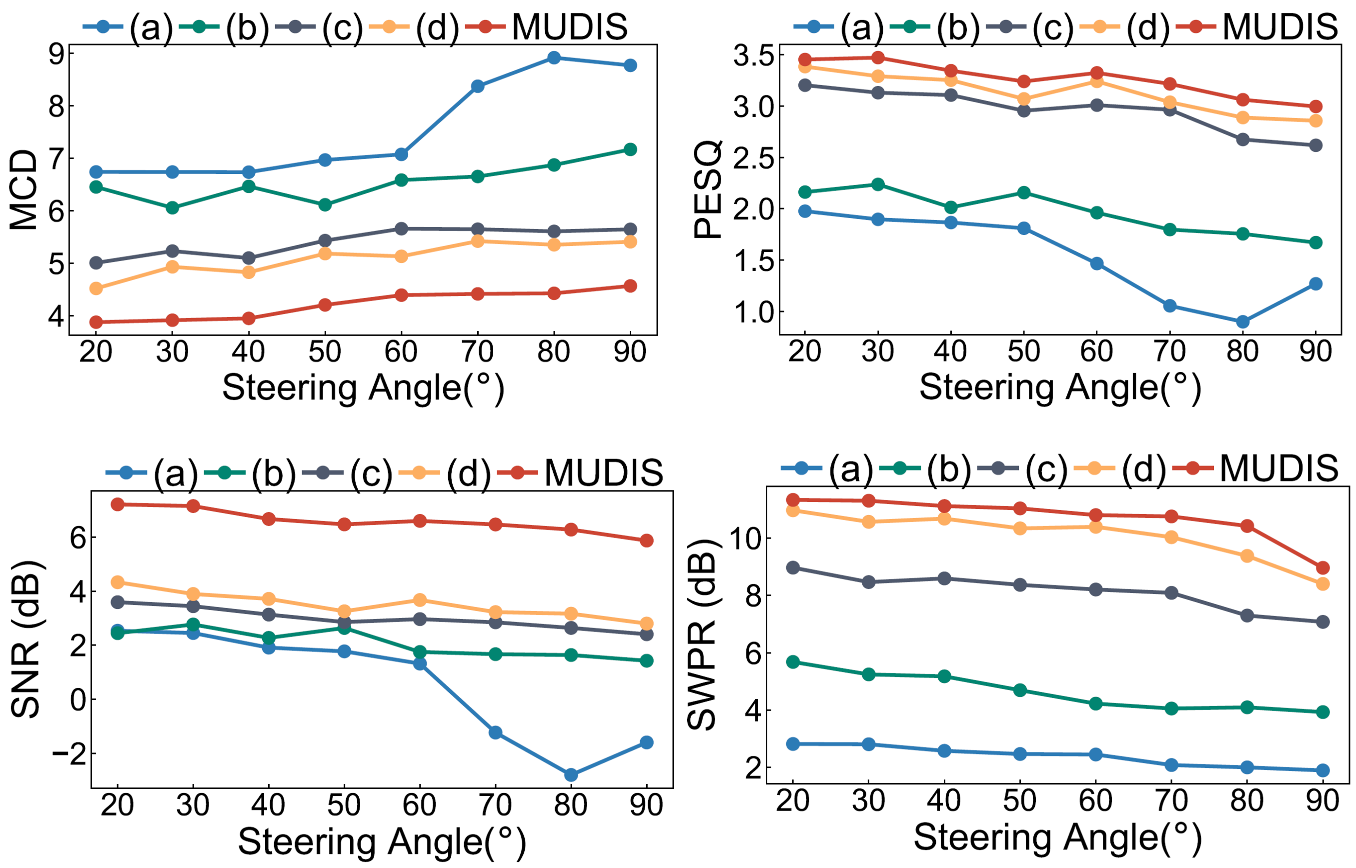
(a) Traditional beamforming

(b) Traditional beamforming + Metasurface

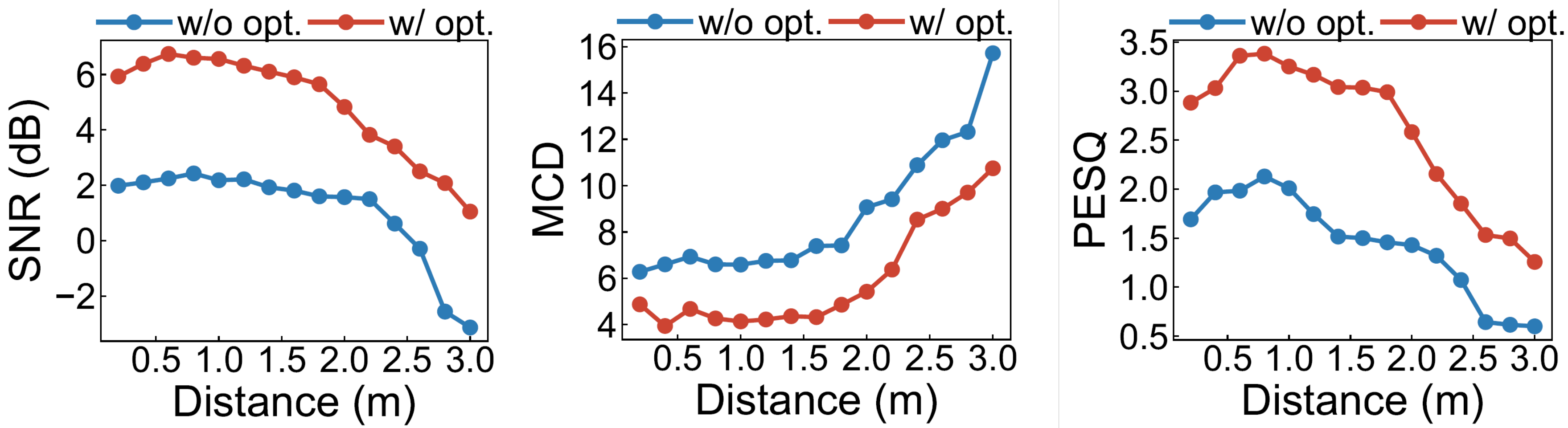
(c) Traditional beamforming + Metasurface + Optimization

(d) Traditional beamforming + Metasurface + Optimization+ Phase Robust

(e) MUDIS：Traditional beamforming + Metasurface + Optimization+ Phase Robust + Distortion Reduction

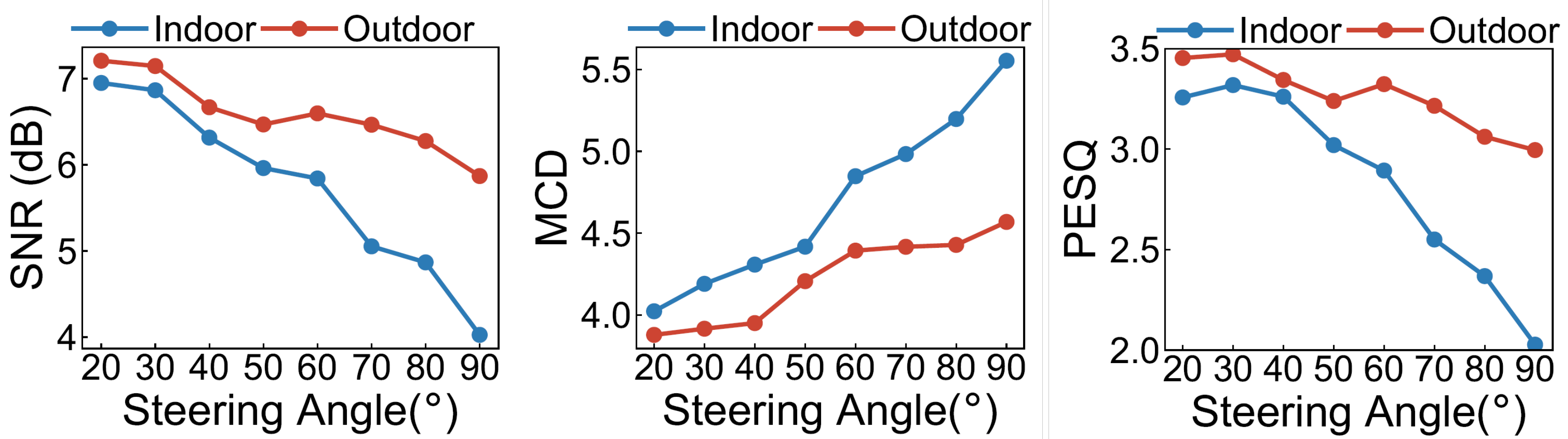


Impact of distance



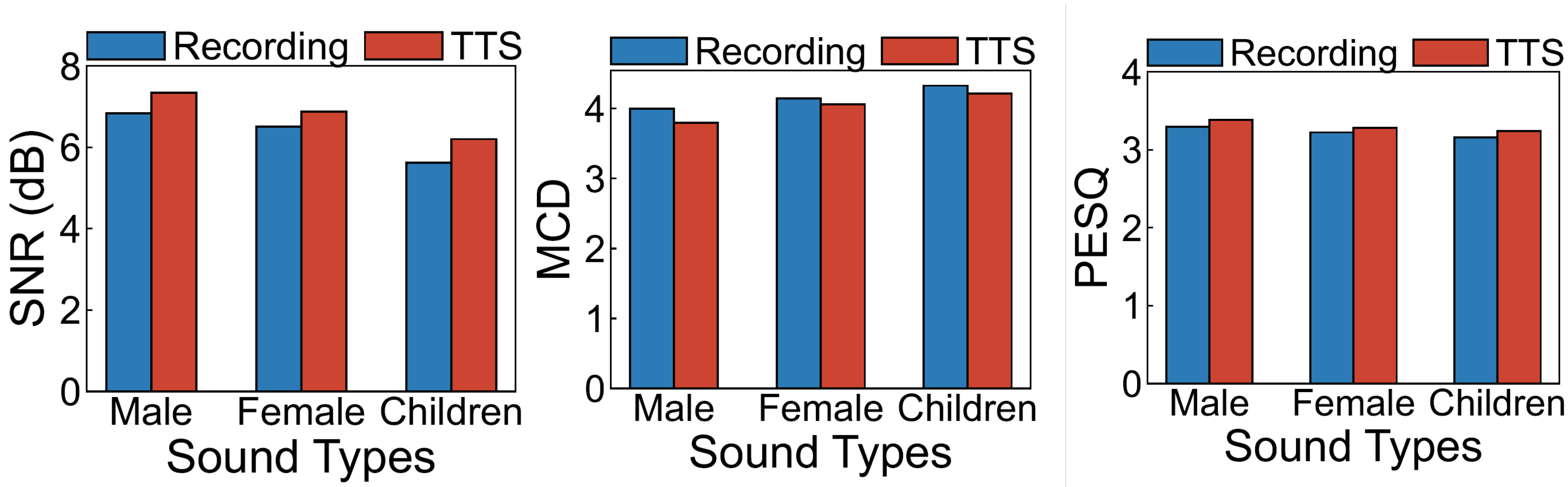
Our system can support 2m distance

Impact of Environment



Our system can work well in both indoor and outdoor environment

Impact of Sound Types



Our system can work well for all sound types