

# IMPACT OF MICROPHONE ARRAY CONFIGURATIONS ON ROBUST INDIRECT 3D ACOUSTIC SOURCE LOCALIZATION

**Elizabeth Vargas, Keith Brown, Kartic Subr**



*Heriot-Watt University*



*University of Edinburgh*

ICASSP 2018

# Acoustic Source Localization

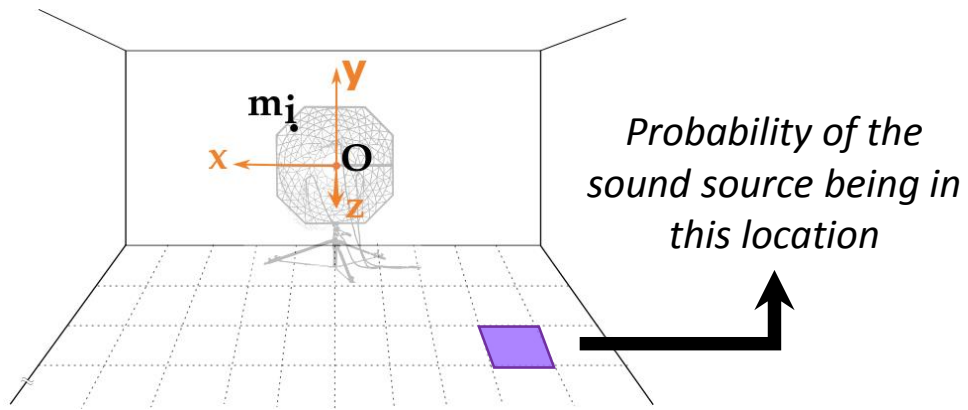


1. Record acoustic signals using a microphone array
2. Calculate time difference of arrivals (TDOA)
3. Calculate the position using geometric formulae (optimization)

# Locating The Source In 3D

Using a microphone array, we calculate Time Difference of Arrivals (TDOA) before we can estimate the acoustic source location

## Steered Response Power (SRP)

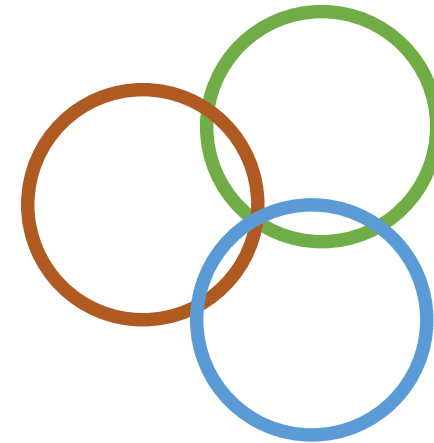


Most likely position amongst a grid of candidate locations

✓ Accurate

✗ Slow

## Multilateration



Infer the source position via least squares optimization

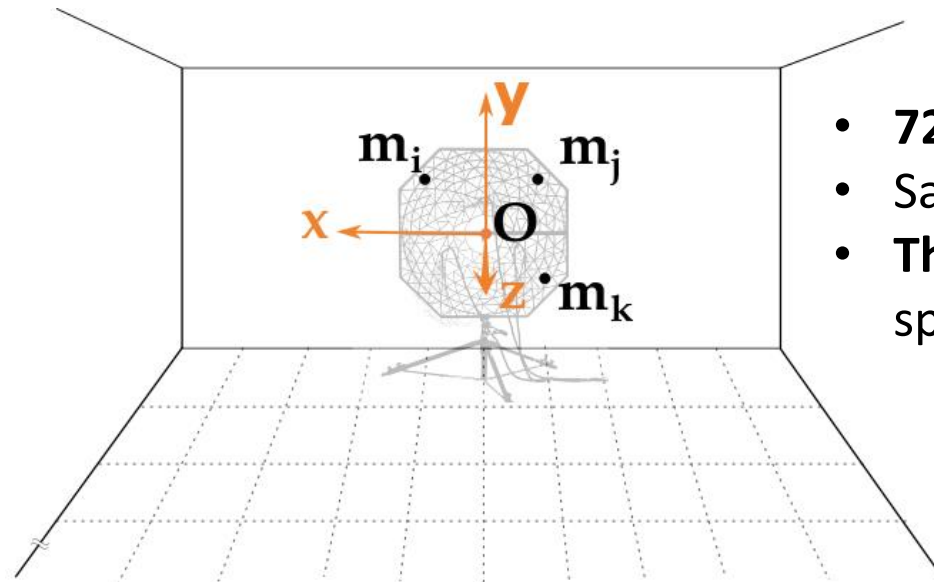
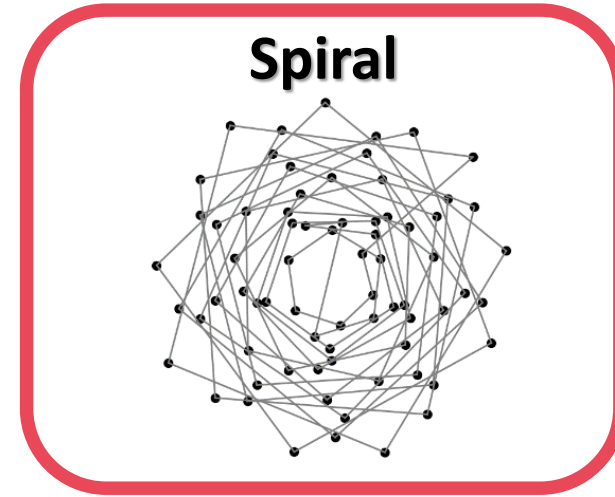
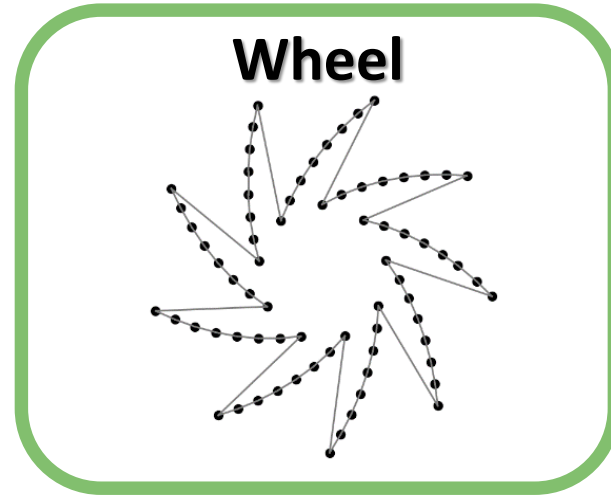
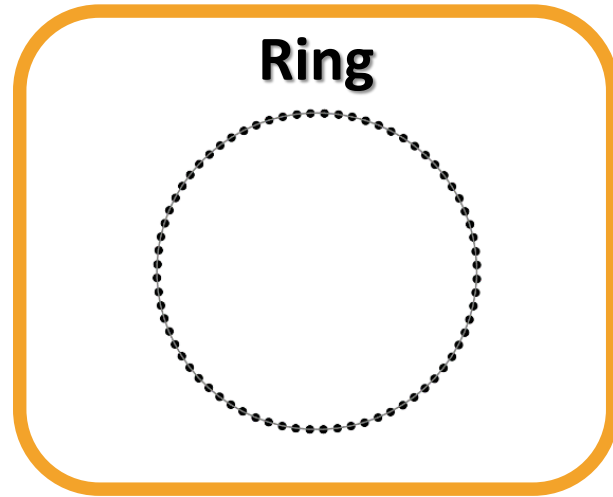
✓ Fast

✗ Non-convex function, local minima

## **QUESTION #1:**

Can localization be accurate and fast at the same time?

# Microphone Arrays For Acoustic Source Localization

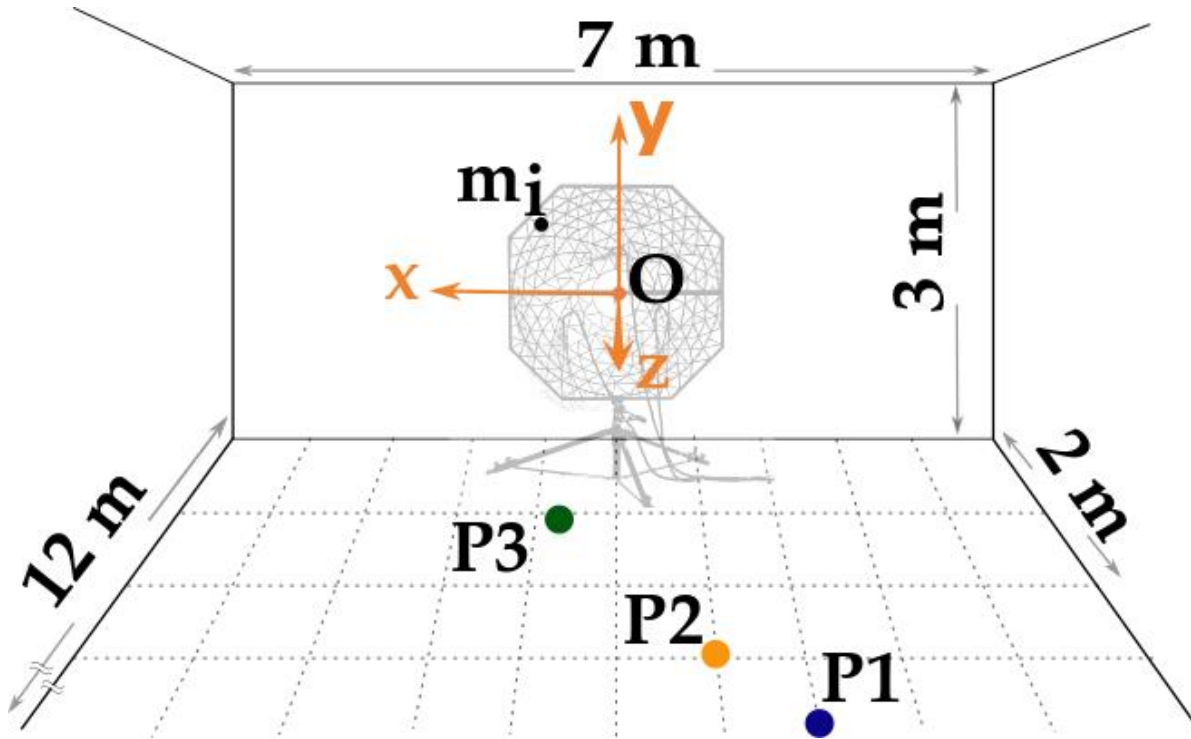


- **72** microphones
- Sampled at **192 kHz**
- **Three configurations** spanning the same area

## **QUESTION #2:**

How does the microphone configuration affect localization accuracy?

# Simulated Source Locations



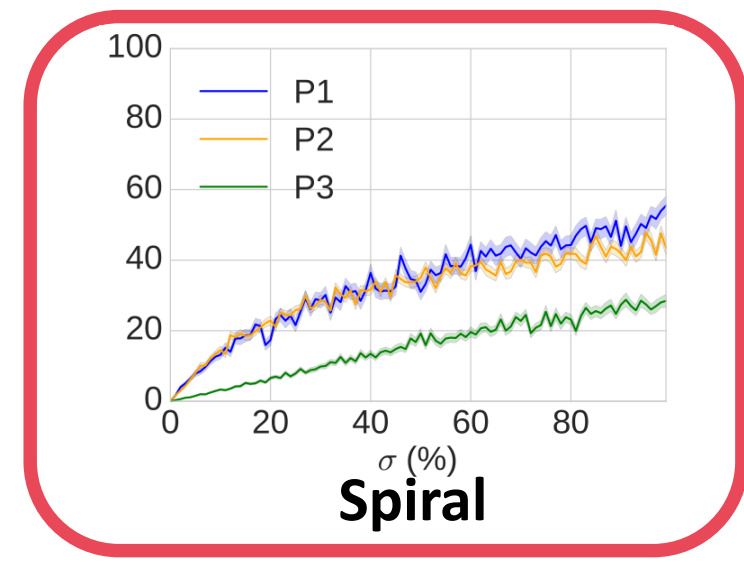
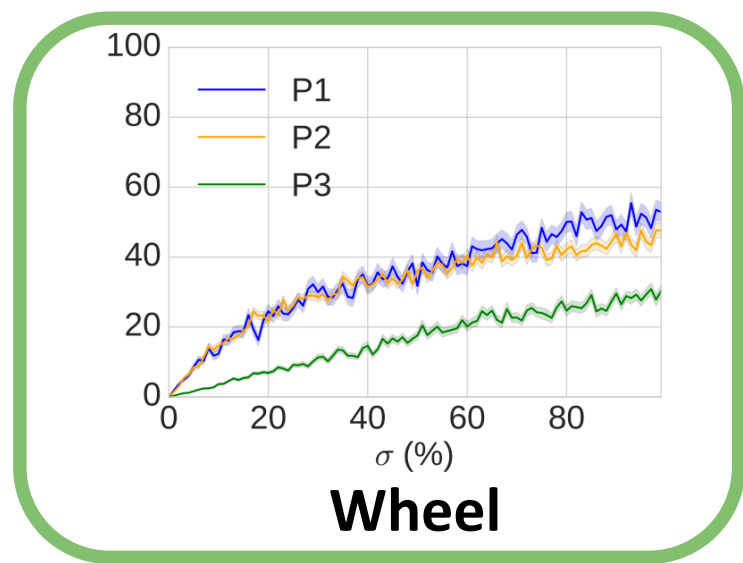
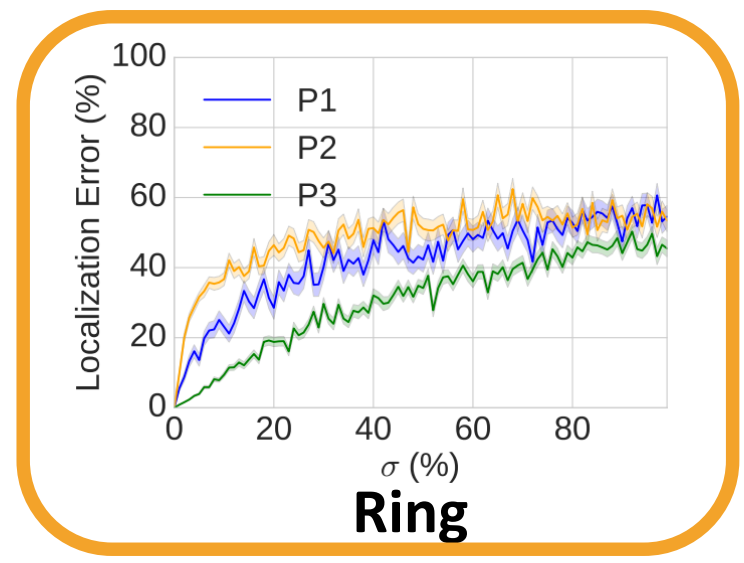
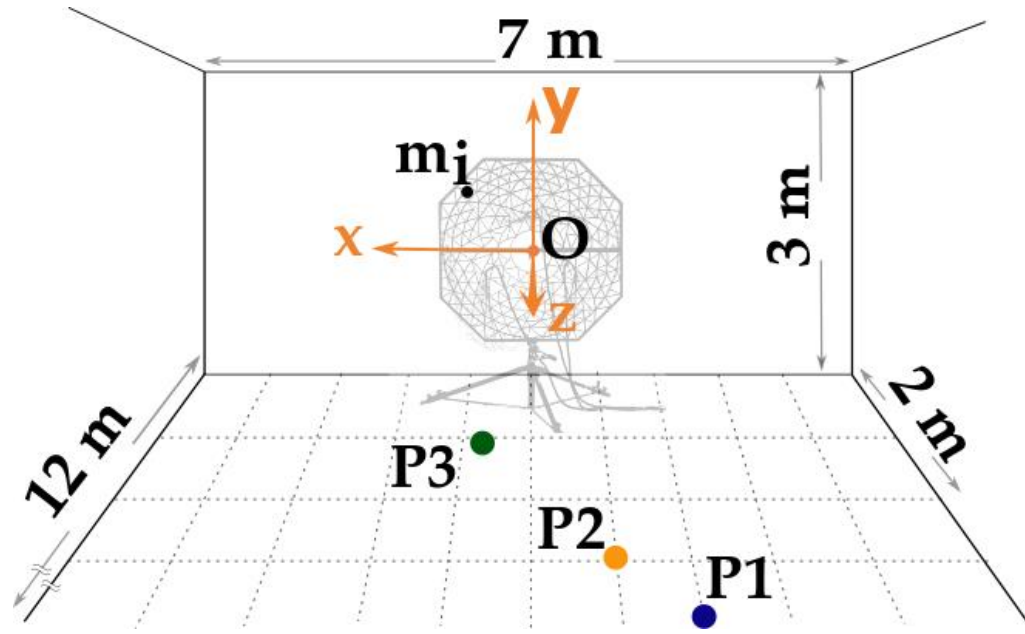
Noise added to Time Difference of Arrivals (TDOA)

$$\eta \sim \mathcal{N} \left( 0, \frac{\sigma}{100} \frac{\|s - O\|}{c} \right)$$

Localization Relative Error

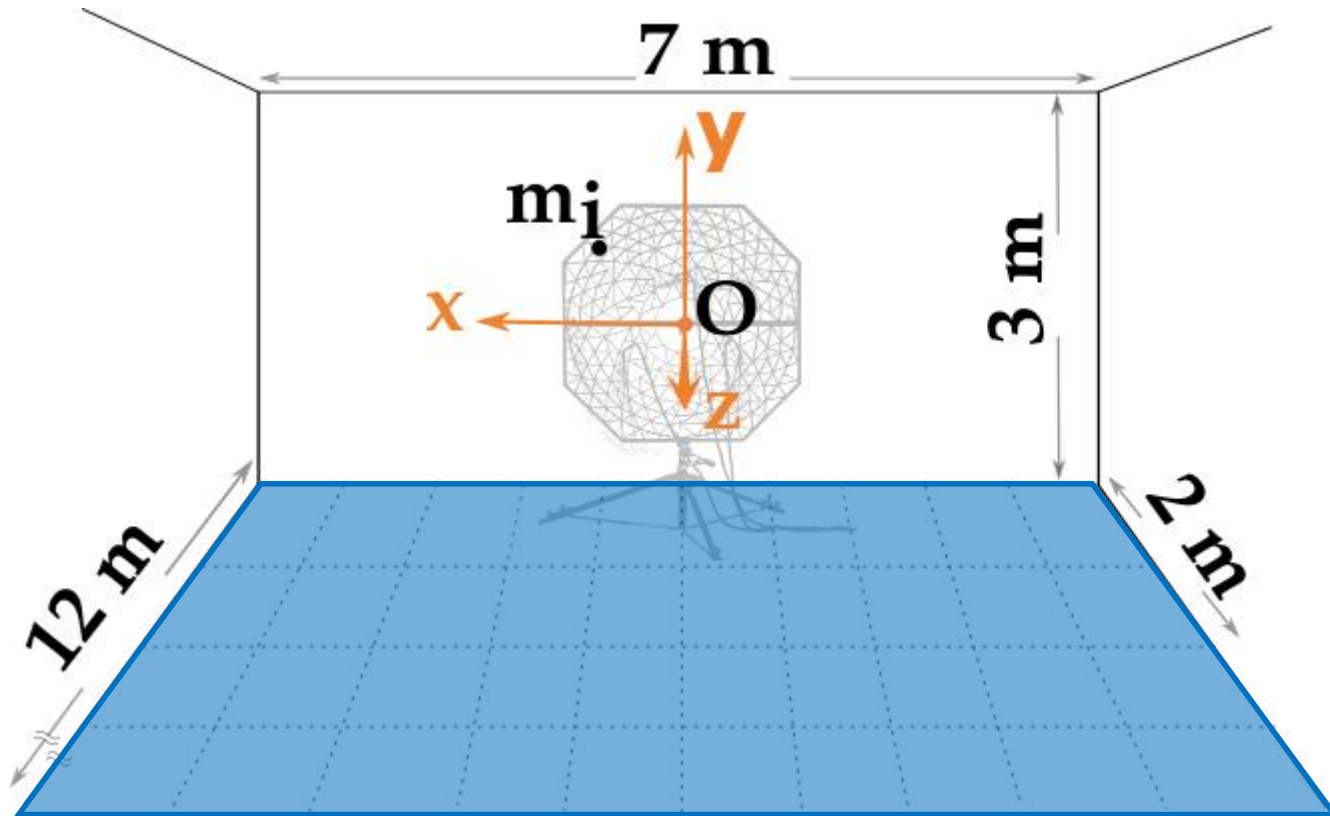
$$\text{error}(\%) = \frac{\|s - \tilde{s}\|}{\|s - O\|} * 100$$

# Wheel and Spiral Configurations Are More Robust





# Simulated Source Locations



Noise added to Time Difference of Arrivals (TDOA)

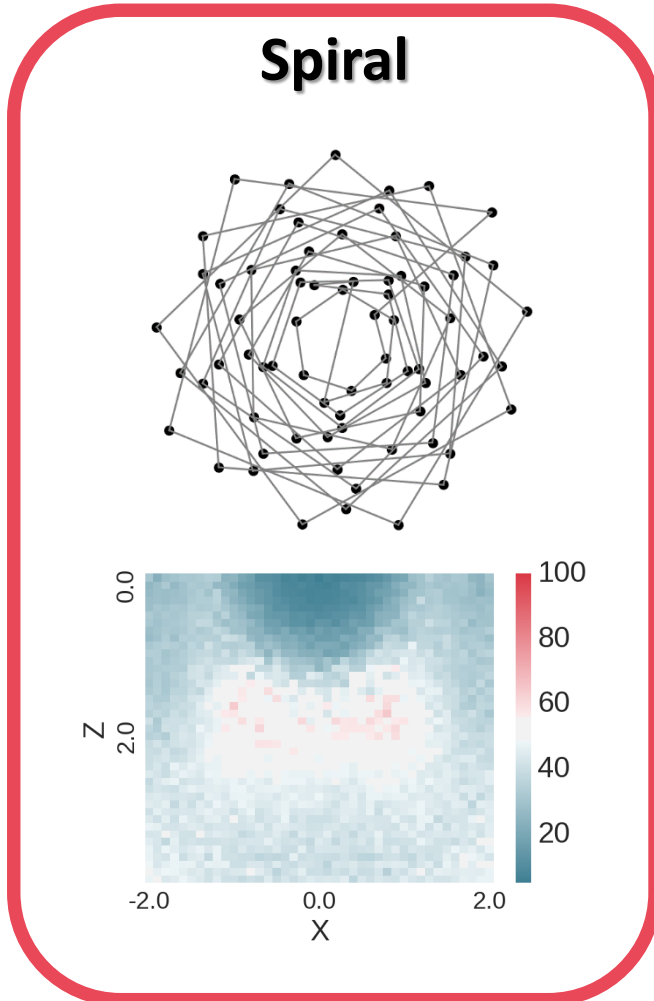
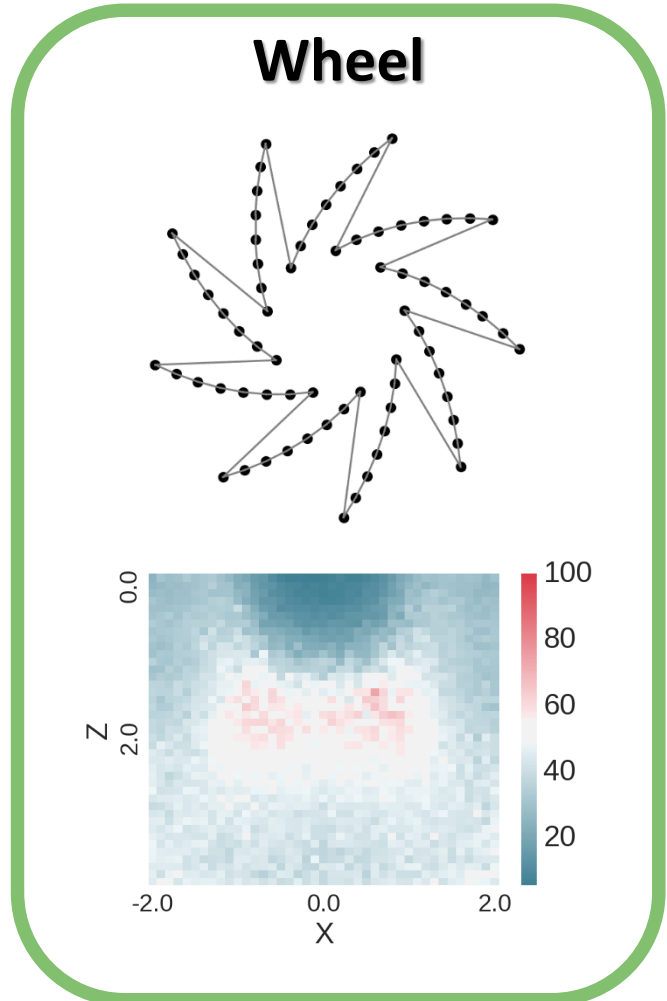
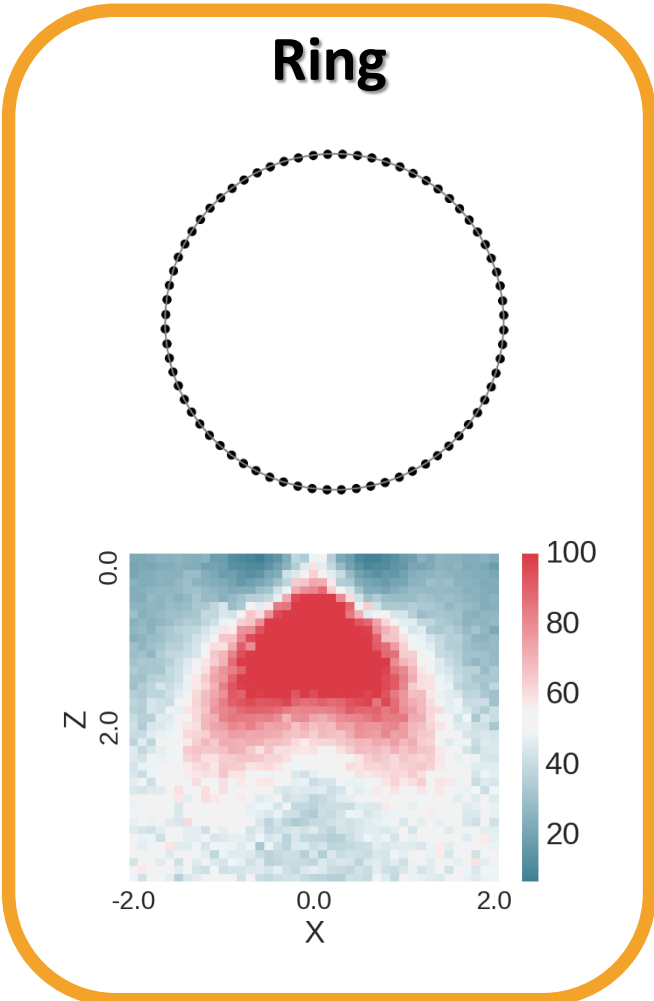
$$\eta \sim \mathcal{N} \left( 0, \frac{\sigma}{100} \frac{\|\mathbf{s} - \mathbf{O}\|}{c} \right)$$

Localization Relative Error

$$\text{error}(\%) = \frac{\|\mathbf{s} - \tilde{\mathbf{s}}\|}{\|\mathbf{s} - \mathbf{O}\|} * 100$$

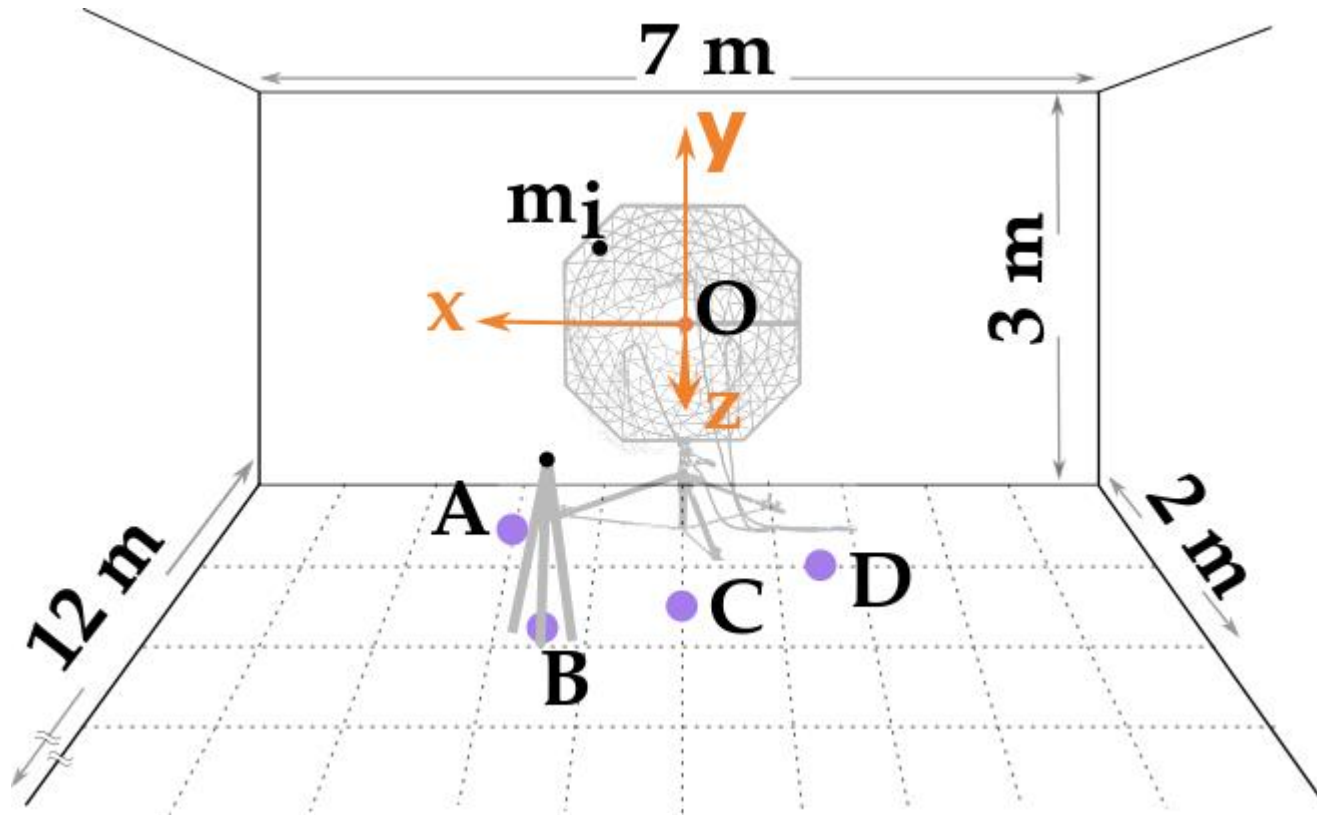
# Higher Errors Observed With Ring Configuration

100% noise for a 2m by 2m room with 3 different configurations spanning the same area

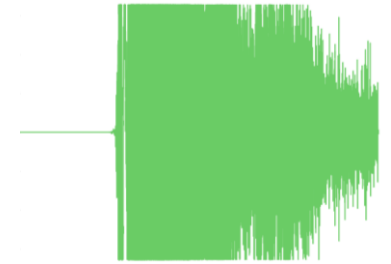


# Using Real Data

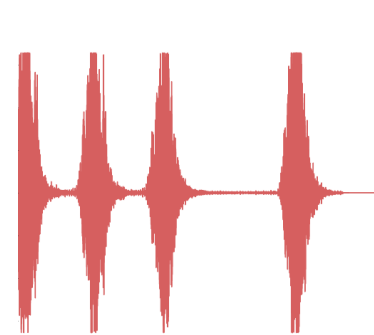
We tested both Multilateration and Steered Response Power (SRP)



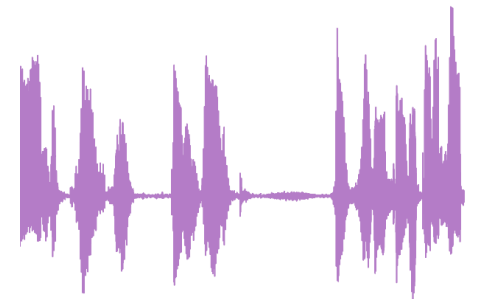
chirp



gunshot



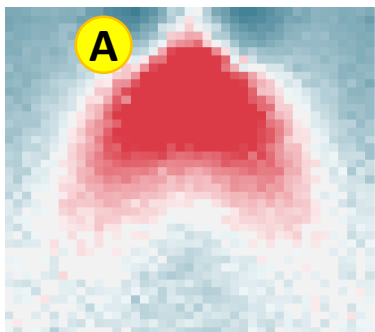
dogbark



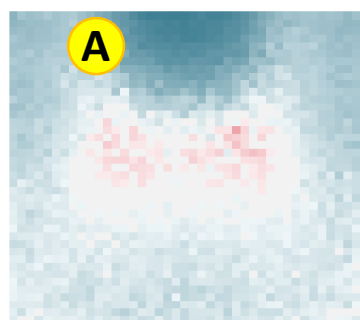
speech

# Short Range Localization

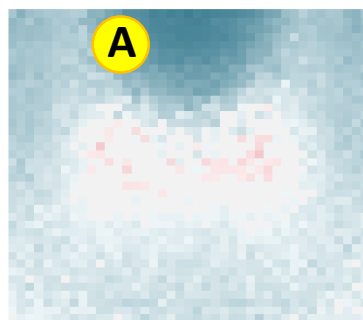
Ring



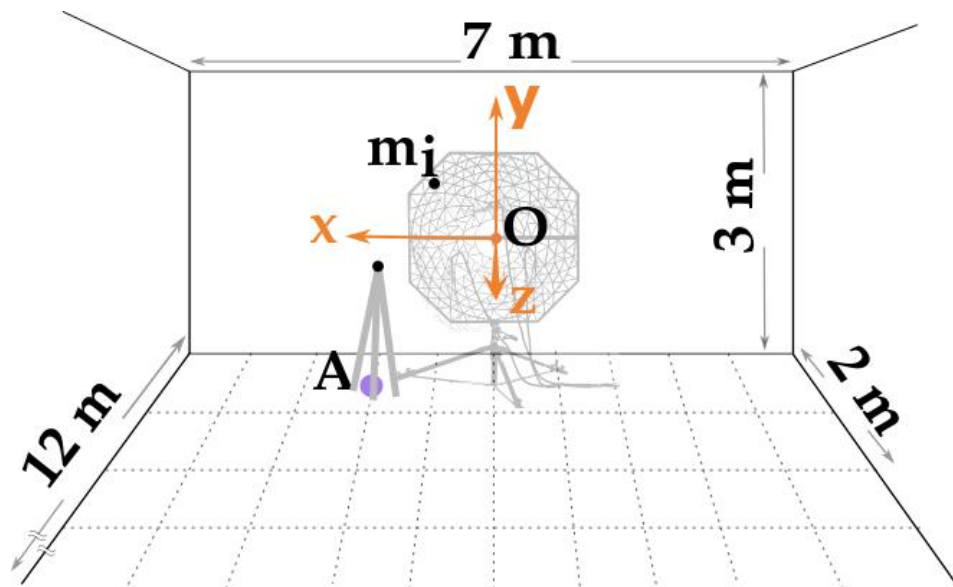
Wheel



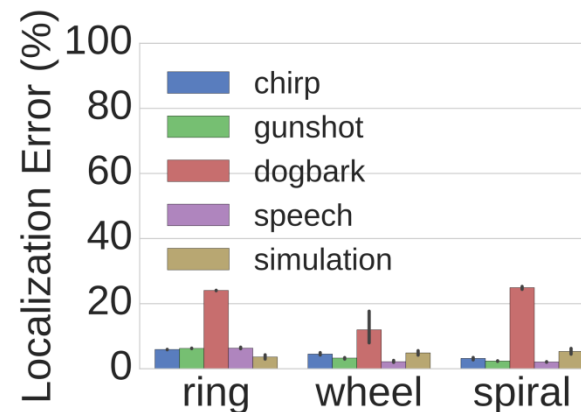
Spiral



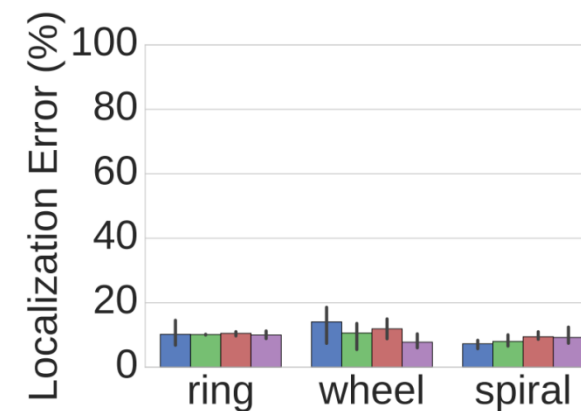
**A: (2.0,-0.32,0.5)**



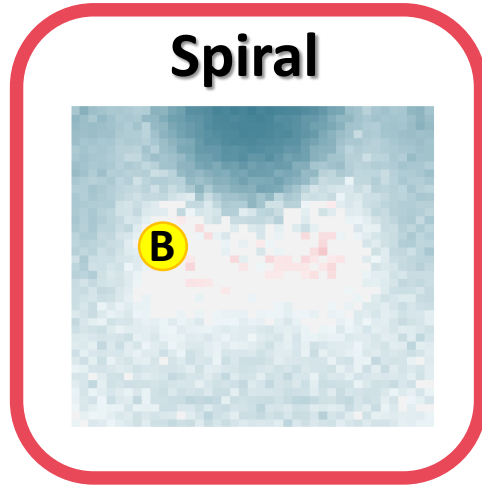
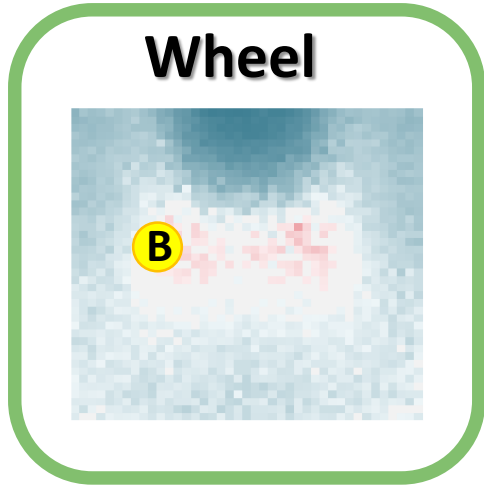
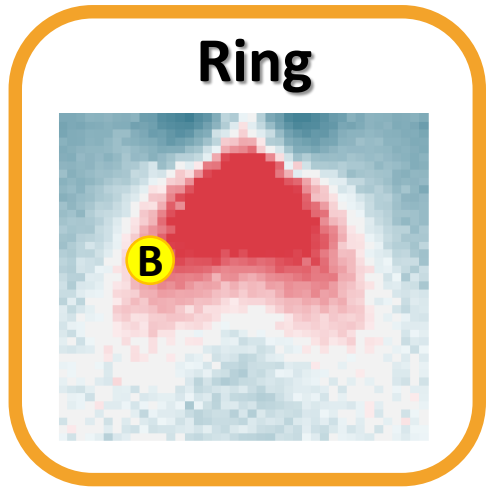
Multilateration



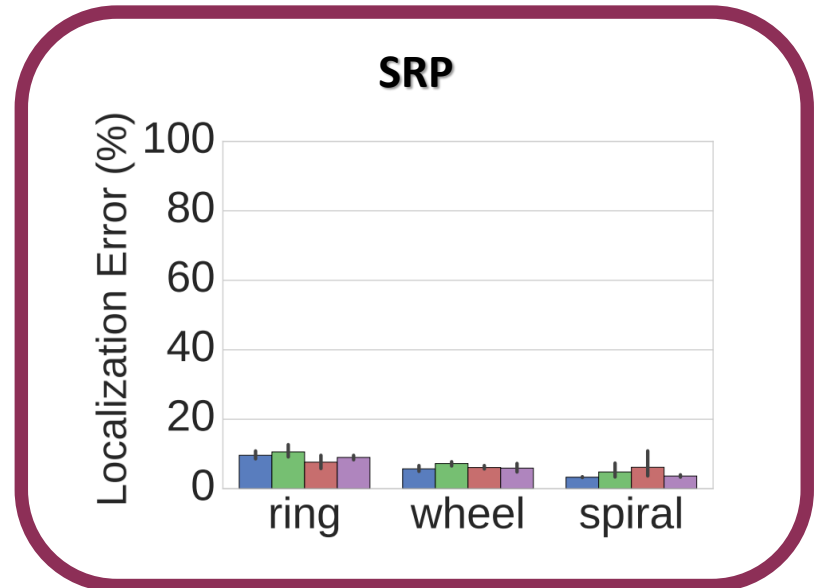
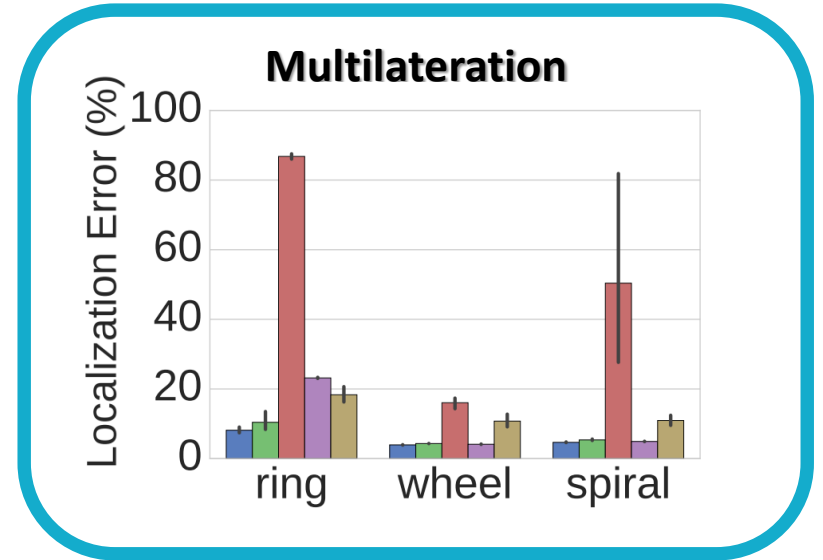
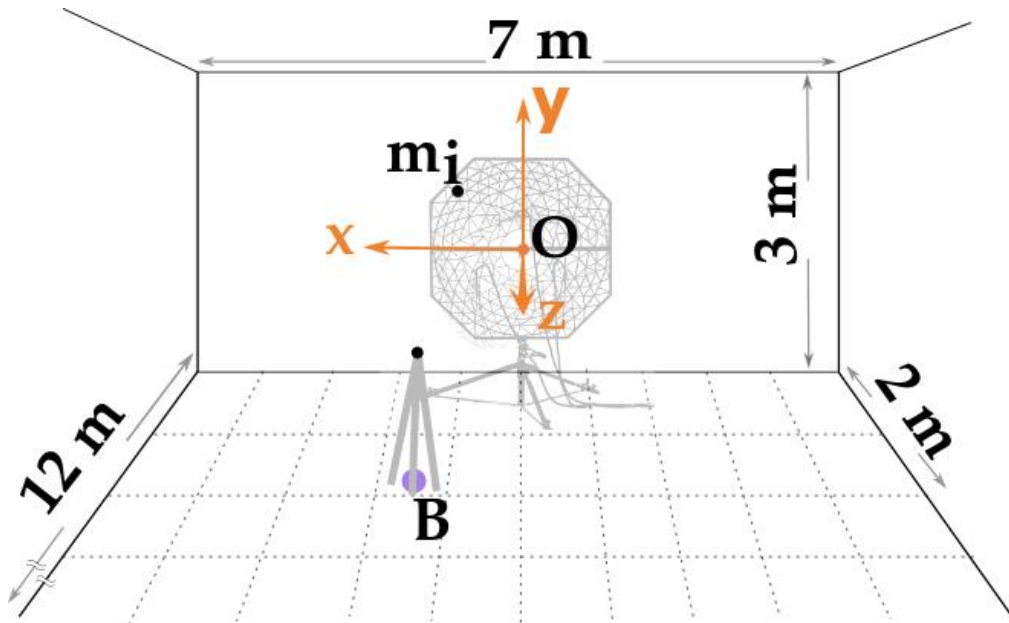
SRP



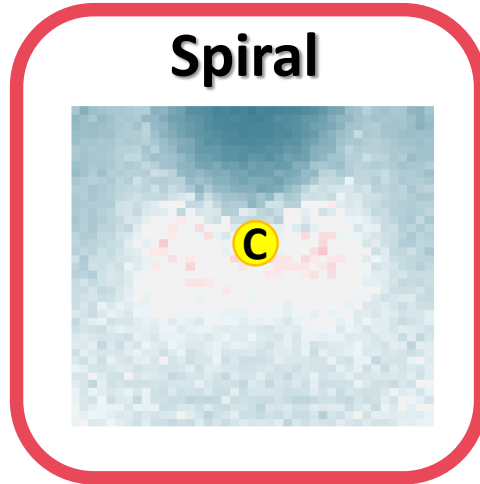
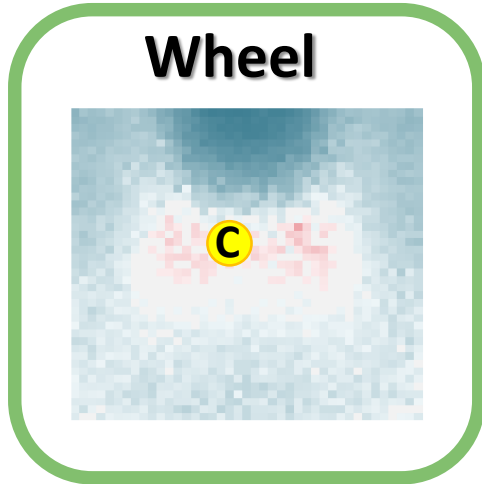
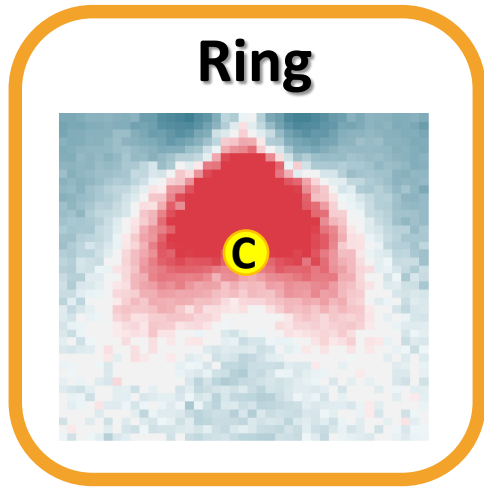
# Mid Range Localization



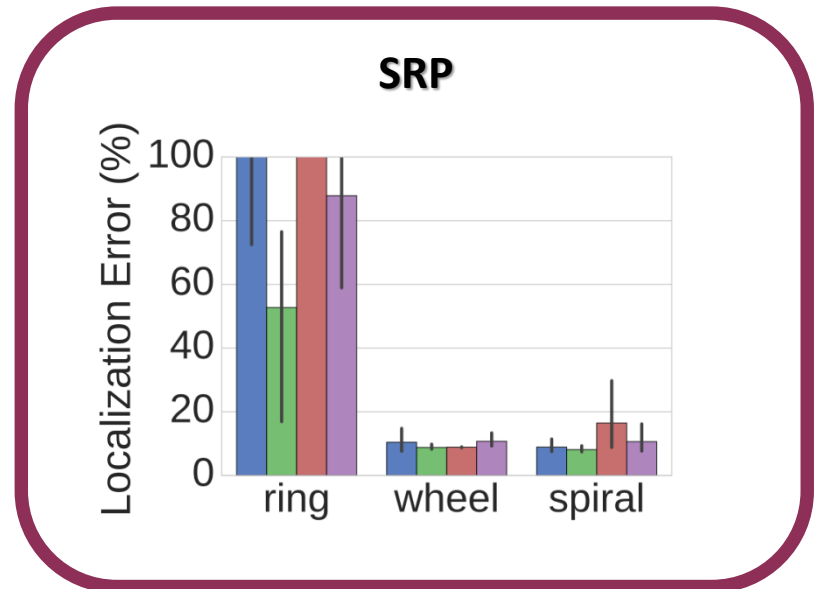
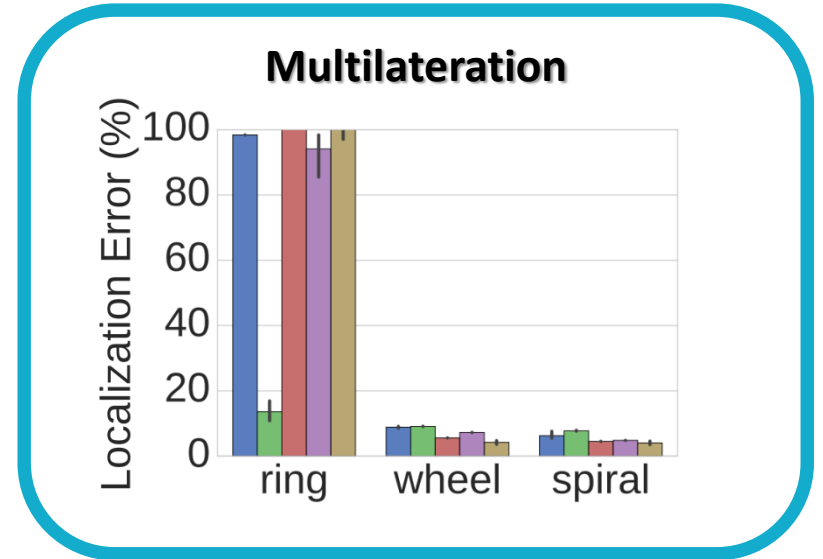
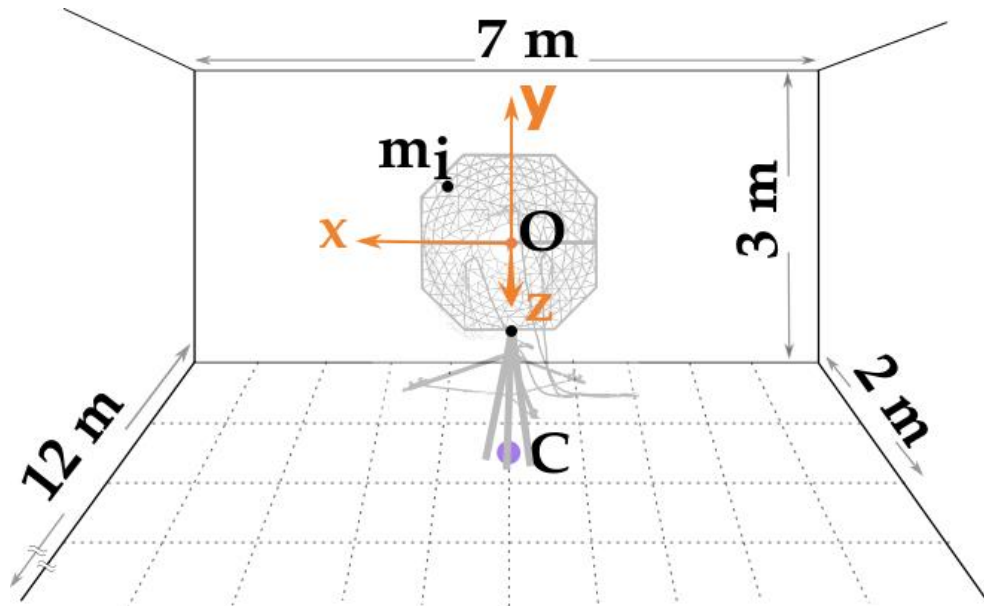
**B: (1.5,-0.32,2.0)**



# Facing The Microphone Array

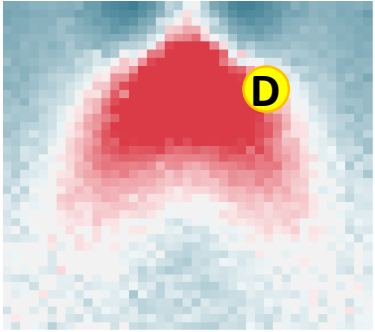


**C: (0.0, -0.32, 1.5)**

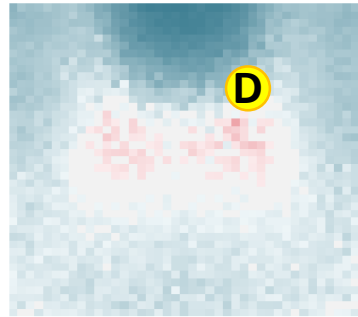


# Mid Range Localization

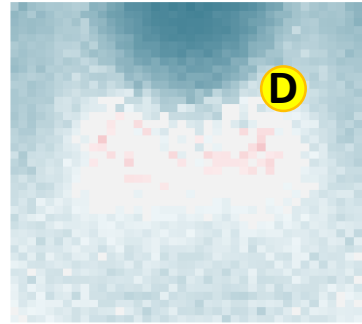
Ring



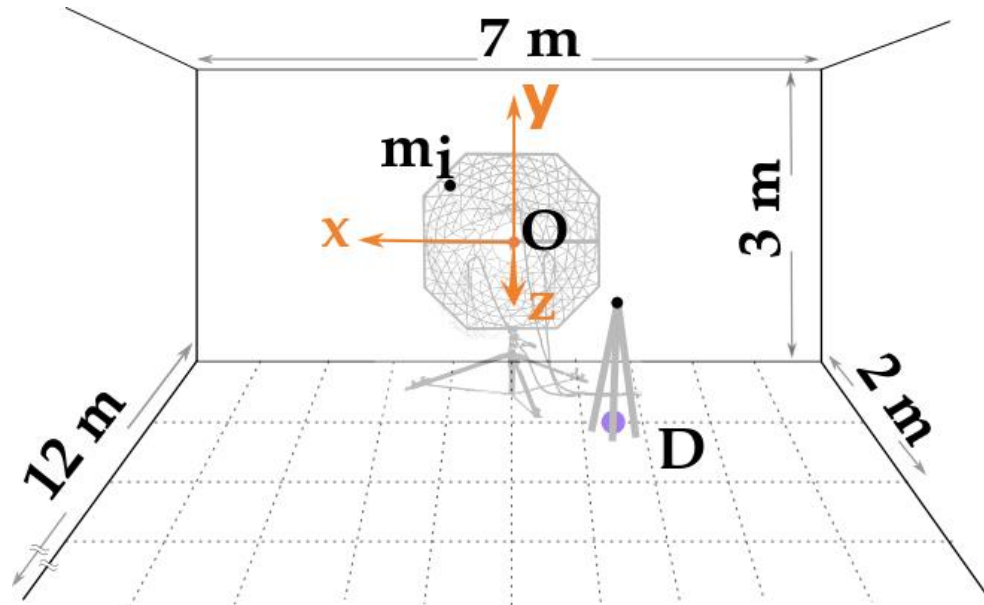
Wheel



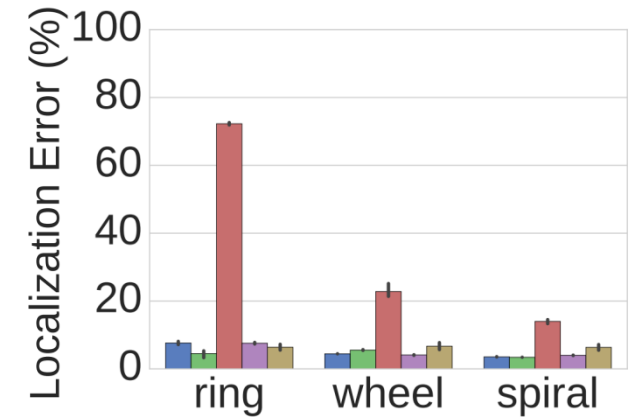
Spiral



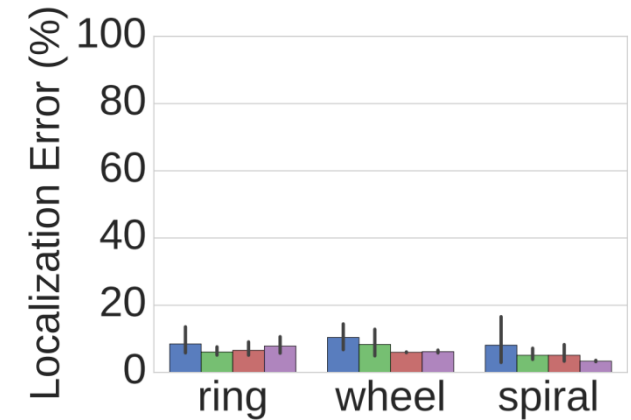
$D: (-1.5, -0.32, 1.0)$



Multilateration



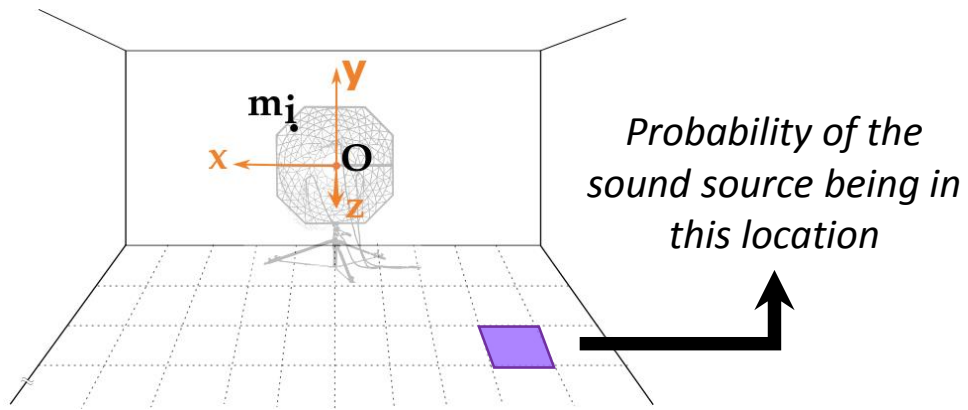
SRP



# What Happened To The Speed?

Using a microphone array, we calculate Time Difference of Arrivals (TDOA) before we can estimate the acoustic source location

## Steered Response Power (SRP)

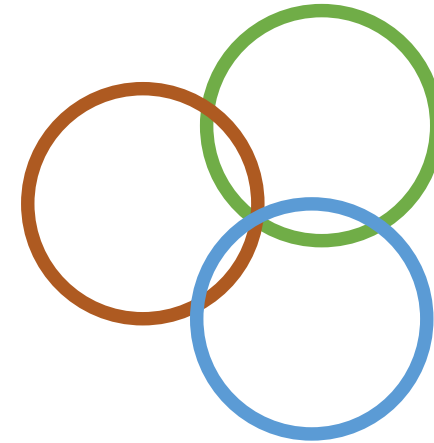


Most likely position amongst a grid of candidate locations

✓ Accurate

✗ Slow

## Multilateration



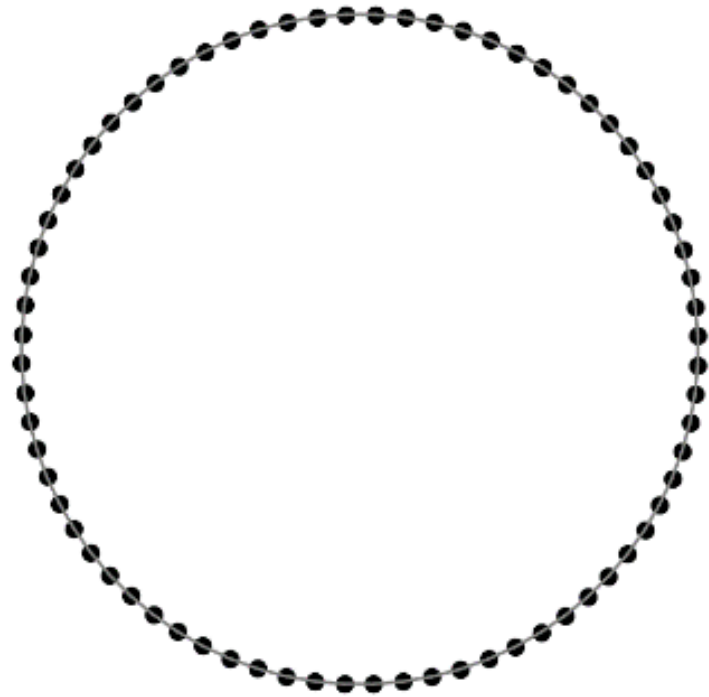
Infer the source position via least squares optimization

✓ Fast

✗ Non-convex function, local minima



# How Many Microphone Pairs To Use?



**72 microphones**

**Array 16  
Microphones**

**120  
Pairs**

**Array 32  
Microphones**

**496  
Pairs**

**Array 72  
Microphones**

**2556  
Pairs**

# Using 2556 Microphone Pairs

## Accuracy (%)

Signal	SRP	Multilateration
Chirp	14.7 (25.9)	12.1 (23.2)
Gunshot	11.0 (13.3)	6.4 (3.5)
Dogbark	16.0 (28.5)	48.5 (44.6)
Speech	13.2 (21.1)	12.9 (22.5)

## Time (minutes)

Signal	SRP	Multilateration
Chirp	3 (0.2)	4.5 (0.03)
Gunshot	2.58 (0.2)	2.4 (0.02)
Dogbark	2.49 (0.1)	2.4 (0.02)
Speech	2.63 (0.1)	2.5 (0.02)

# Using 100 Microphone Pairs

## Accuracy (%)

Signal	SRP	Multilateration
Chirp	14.7 (25.9)	14.2 (25.9)
Gunshot	11.0 (13.3)	9.6 (12.8)
Dogbark	16.0 (28.5)	58.9 (38.8)
Speech	13.2 (21.1)	15.2 (23.5)

## Time (minutes)

Signal	SRP	Multilateration
Chirp	3 (0.2)	0.5(0.01)
Gunshot	2.58 (0.2)	0.4 (0.02)
Dogbark	2.49 (0.1)	0.4 (0.02)
Speech	2.63 (0.1)	0.4 (0.02)

# Conclusions

## **QUESTION #1:**

**Can localization be accurate and fast at the same time?**

Yes it can! Direct optimization yields errors similar to the Steered Response Power (SRP) method with 6 times less computation

## **QUESTION #2:**

**How does the microphone configuration affects the localization accuracy?**

Circular arrays are the least desirable configuration