Direction of Arrival Estimation for Wildlife Protection

Gustav Zetterqvist

Licentiate Seminar

2024-10-18

Supervisor: Fredrik Gustafsson
Co-supervisor: Gustav Hendeby

Opponent: Andreas Jakobsson (LU)
Examiner: Svante Gunnarsson

LINKÖPING UNIVERSITY

WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM

Background

- We are entering the sixth mass extinction.

Background

- We are entering the sixth mass extinction.
- Caused by human population growth and increased consumption

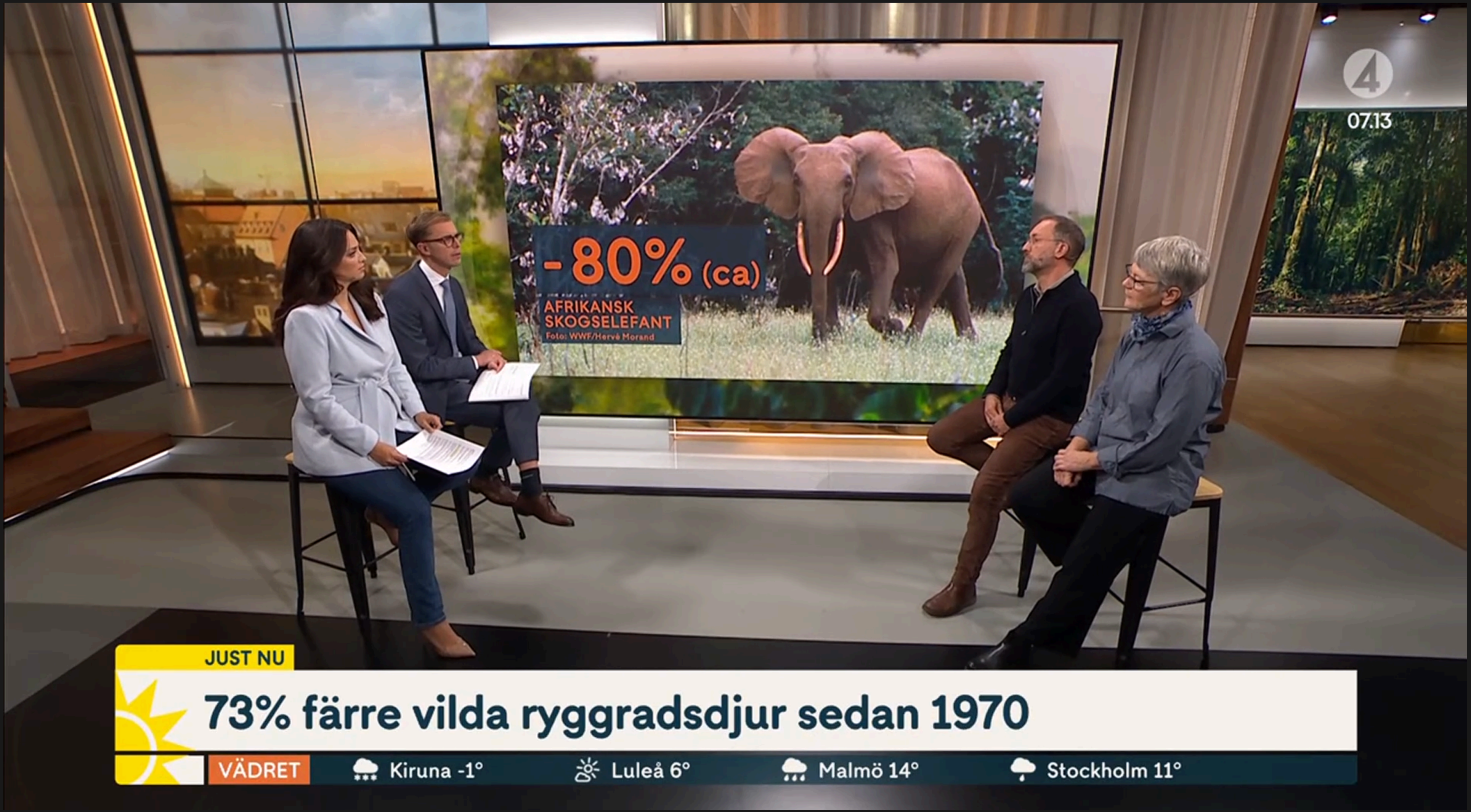
Background

- We are entering the sixth mass extinction.
- Caused by human population growth and increased consumption
- Elephant population has decreased by 98% since the 1500s.

Background

- We are entering the sixth mass extinction.
- Caused by human population growth and increased consumption
- Elephant population has decreased by 98% since the 1500s.
- Decreasing rate between 2010-2014 was 8%.

Background



Nyhetsmorgon 10th of October

Example 1 - Elephant tracking

- Human-elephant conflict endangers people, property, and crops globally.

Example 1 - Elephant tracking

- Human-elephant conflict endangers people, property, and crops globally.
- A geophone-based system detects seismic waves from elephant footsteps to estimate direction.

Example 1 - Elephant tracking

- Human-elephant conflict endangers people, property, and crops globally.
- A geophone-based system detects seismic waves from elephant footsteps to estimate direction.
- Real-time tracking helps warn residents and guide elephants back to their habitats.

Example 1 - Elephant tracking

Example 2 - Acoustic Surveillance for Savanna Protection

- Wildlife protection requires localizing poachers and analyzing sound events on the savanna.

Example 2 - Acoustic Surveillance for Savanna Protection

- Wildlife protection requires localizing poachers and analyzing sound events on the savanna.
- Microphone arrays estimate the direction of rifle shots to locate poachers.

Example 2 - Acoustic Surveillance for Savanna Protection

- Wildlife protection requires localizing poachers and analyzing sound events on the savanna.
- Microphone arrays estimate the direction of rifle shots to locate poachers.
- Key sounds monitored: elephant trumpets, a woman's scream, and police sirens.

Example 2 - Acoustic Surveillance for Savanna Protection

- Wildlife protection requires localizing poachers and analyzing sound events on the savanna.
- Microphone arrays estimate the direction of rifle shots to locate poachers.
- Key sounds monitored: elephant trumpets, a woman's scream, and police sirens.
- The system offers a complete overview, enhancing response and conservation efforts.

Outline

1. Problem Formulation

Outline

1. Problem Formulation
2. Classical Direction of Arrival

Outline

1. Problem Formulation
2. Classical Direction of Arrival
3. Elephant DOA

Outline

1. Problem Formulation
2. Classical Direction of Arrival
3. Elephant DOA
4. Directional Sensitivity DOA

Outline

1. Problem Formulation
2. Classical Direction of Arrival
3. Elephant DOA
4. Directional Sensitivity DOA
5. Conclusions & Future Work

Problem Statement

- How can a conventional DOA estimator be used to estimate the **DOA of elephant footsteps?** And how can we **differentiate** between an elephant footstep and other seismic signals?

Problem Statement

- How can a conventional DOA estimator be used to estimate the **DOA of elephant footsteps?** And how can we **differentiate** between an elephant footstep and other seismic signals?
- How can the **directional sensitivity** of the sensors be utilized to estimate the DOA, and how should the directional sensitivity be **modeled**?

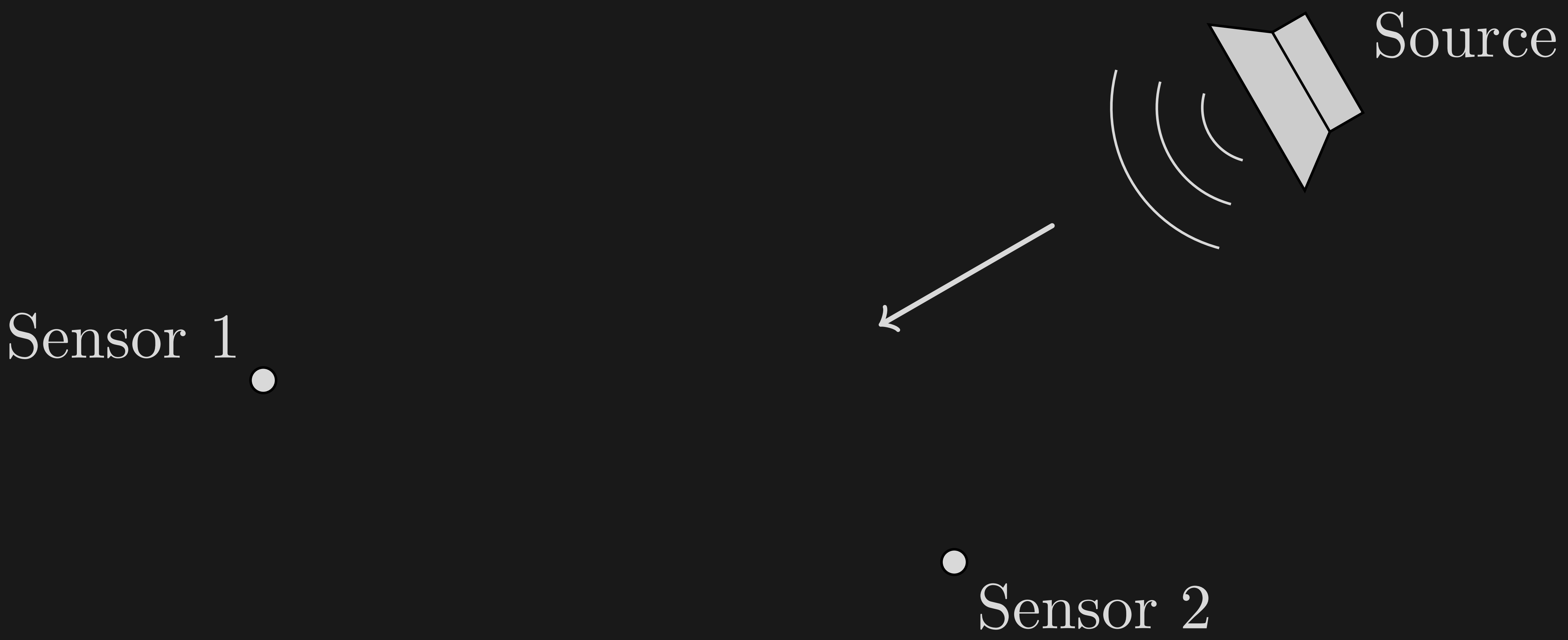
Problem Statement

- What are the **theoretical limits** of the directional sensitivity based DOA estimator, and how does the **frequency content** of the signal affect the directional sensitivity?

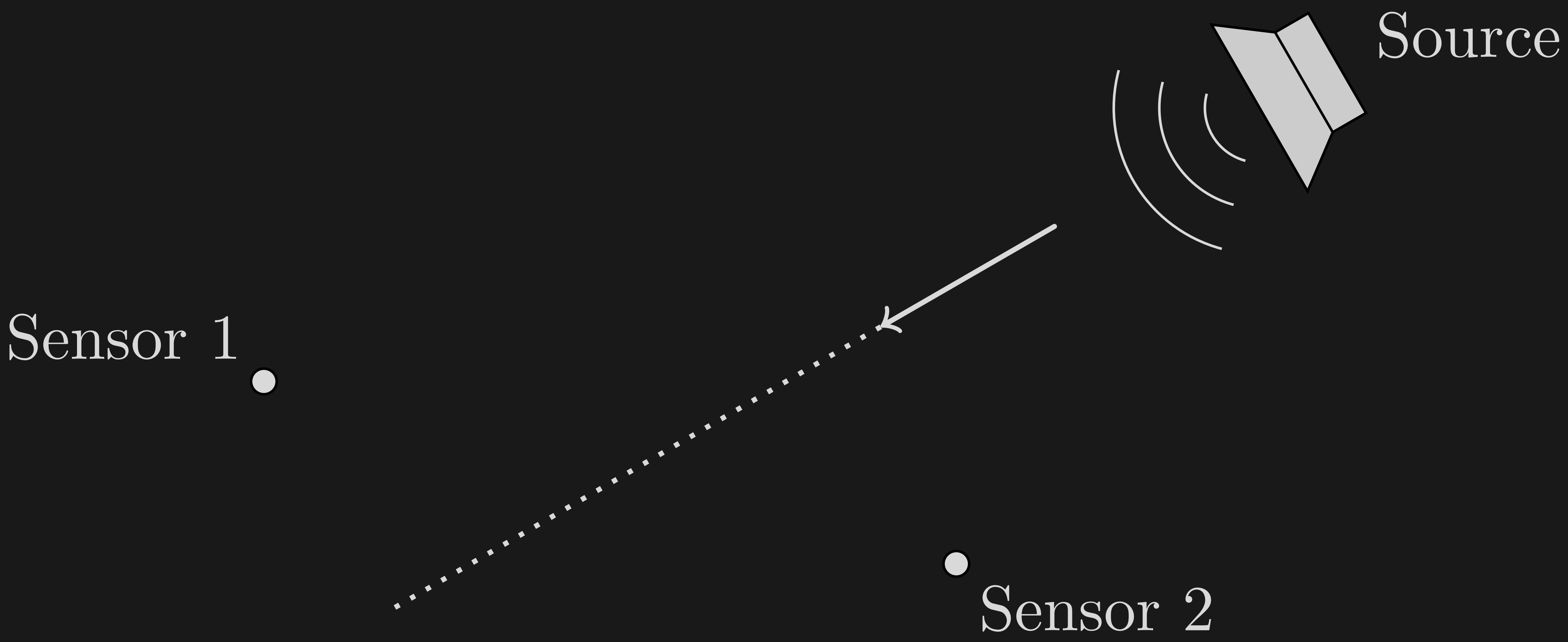
Problem Statement

- What are the **theoretical limits** of the directional sensitivity based DOA estimator, and how does the **frequency content** of the signal affect the directional sensitivity?
- How does the directional sensitivity based DOA estimate **compare to state-of-the-art** methods using narrowband and realistic signals?

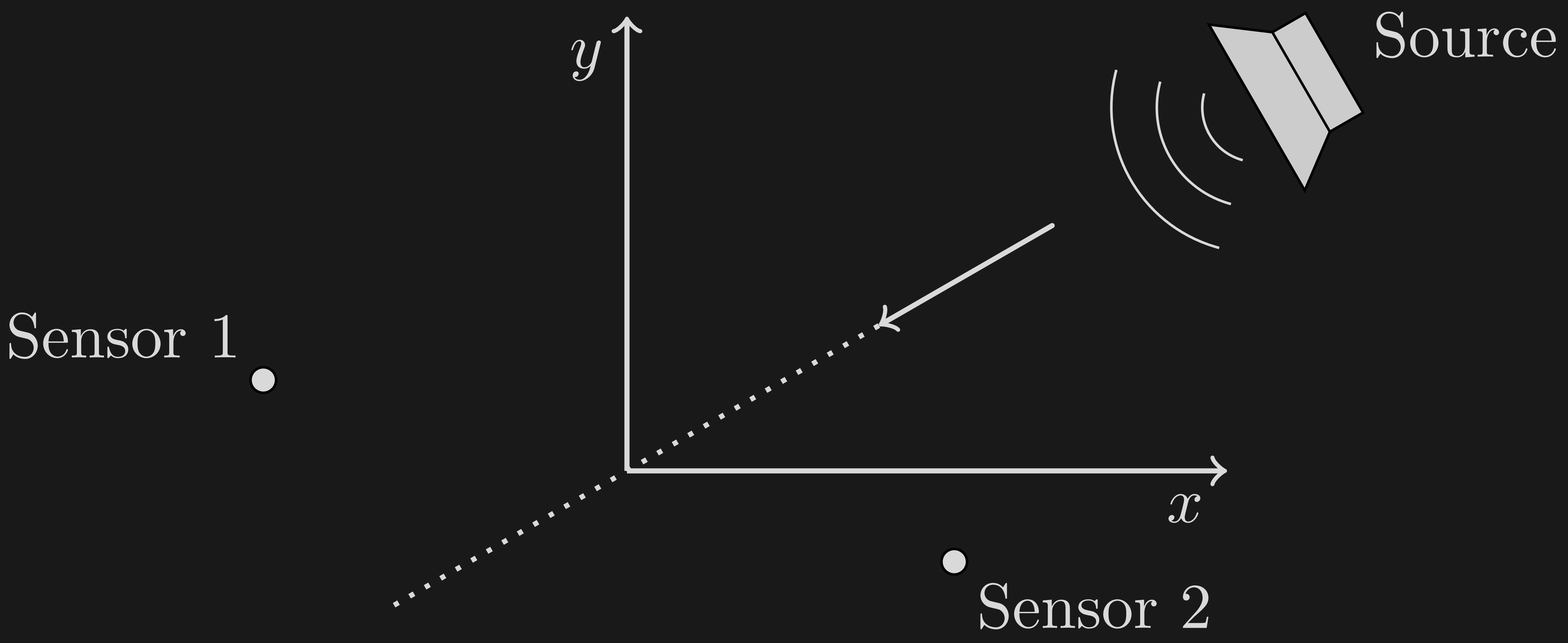
Classical DOA Estimation



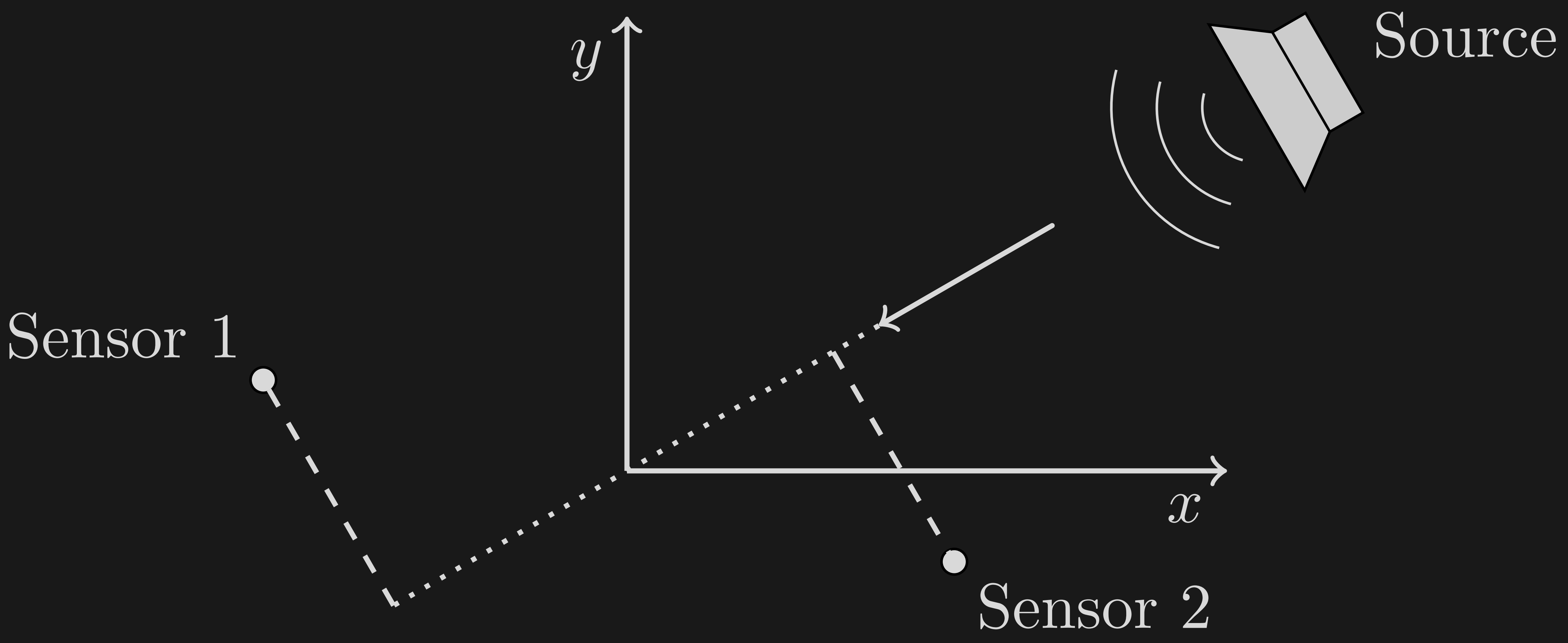
Classical DOA Estimation



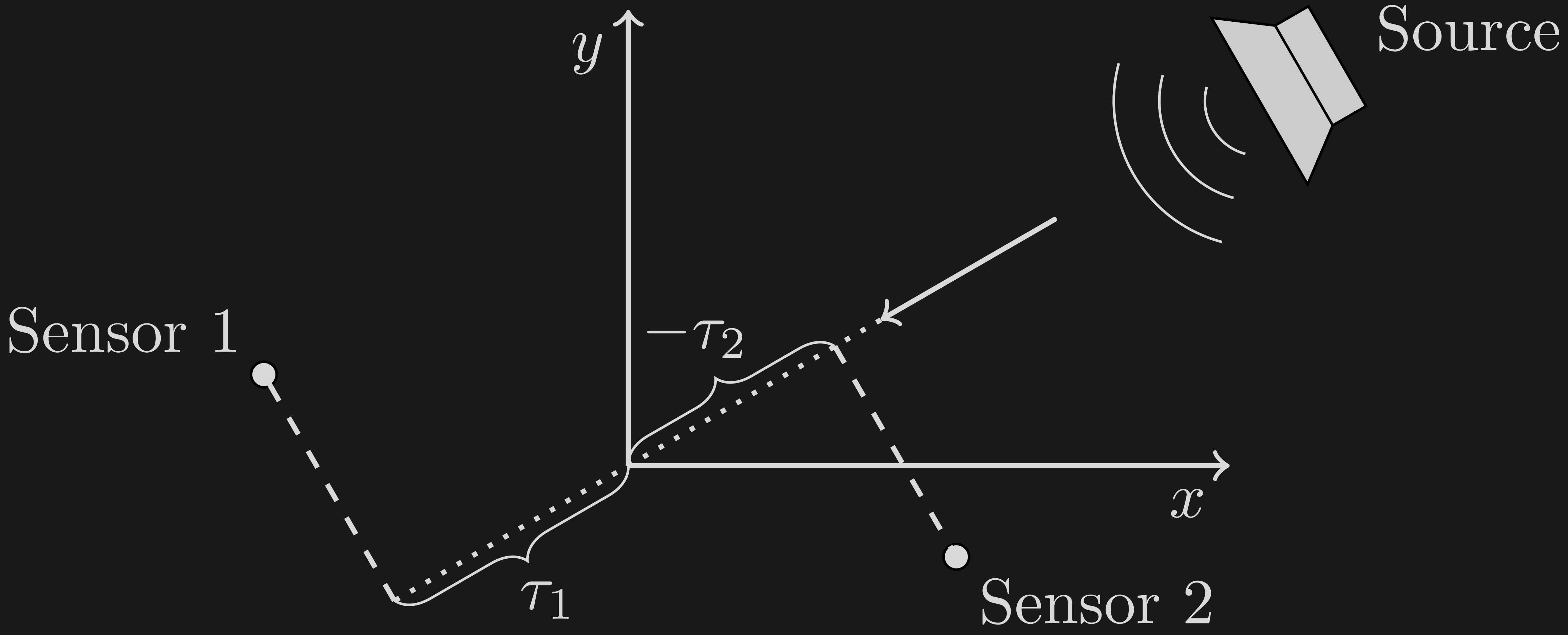
Classical DOA Estimation



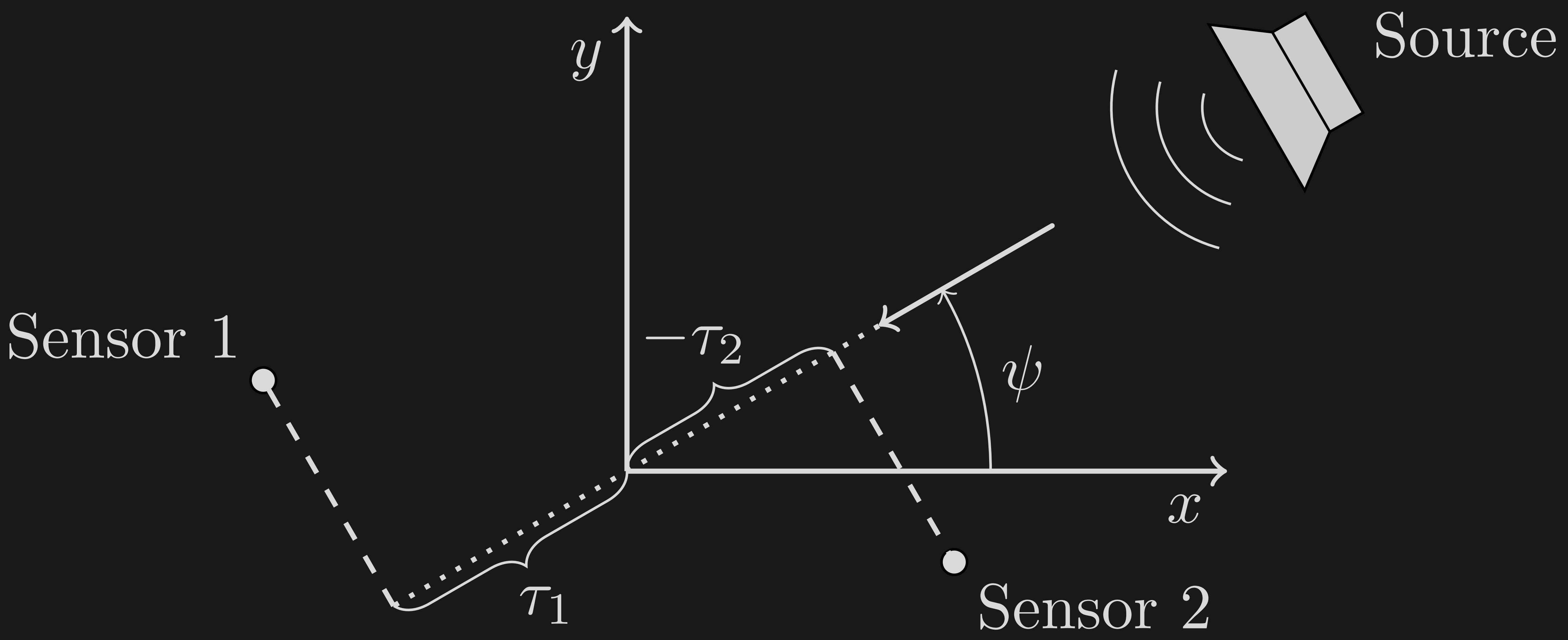
Classical DOA Estimation



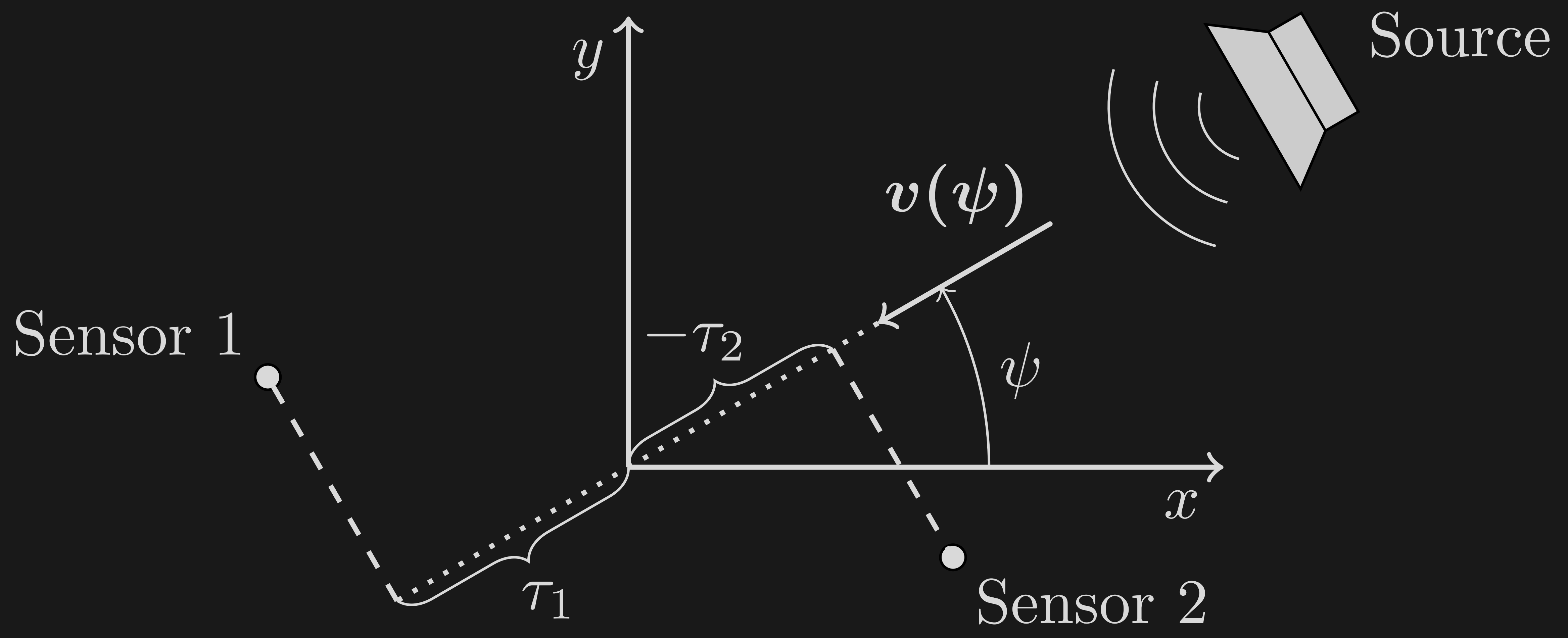
Classical DOA Estimation



Classical DOA Estimation



Classical DOA Estimation



Classical DOA Estimation

$$y_m(t) = s(t - \tau_m(\psi)) + e_m(t)$$

Classical DOA Estimation

$$y_m(t) = s(t - \tau_m(\psi)) + e_m(t)$$

$$\tau_m = \frac{\boldsymbol{p}_m^T \boldsymbol{v}(\psi)}{c}$$

$\boldsymbol{p}_m$: position of sensors

$\boldsymbol{v}(\psi)$: direction vector

c: speed of sound

ψ: DOA

Classical DOA Estimation

$$y_m(t) = s(t - \tau_m(\psi)) + e_m(t)$$

$$\xrightarrow{\mathcal{F}} Y_m(f) = S(f)e^{-j2\pi f \tau_m(\psi)} + E_m$$

$$\tau_m = \frac{\boldsymbol{p}_m^T \boldsymbol{v}(\psi)}{c}$$

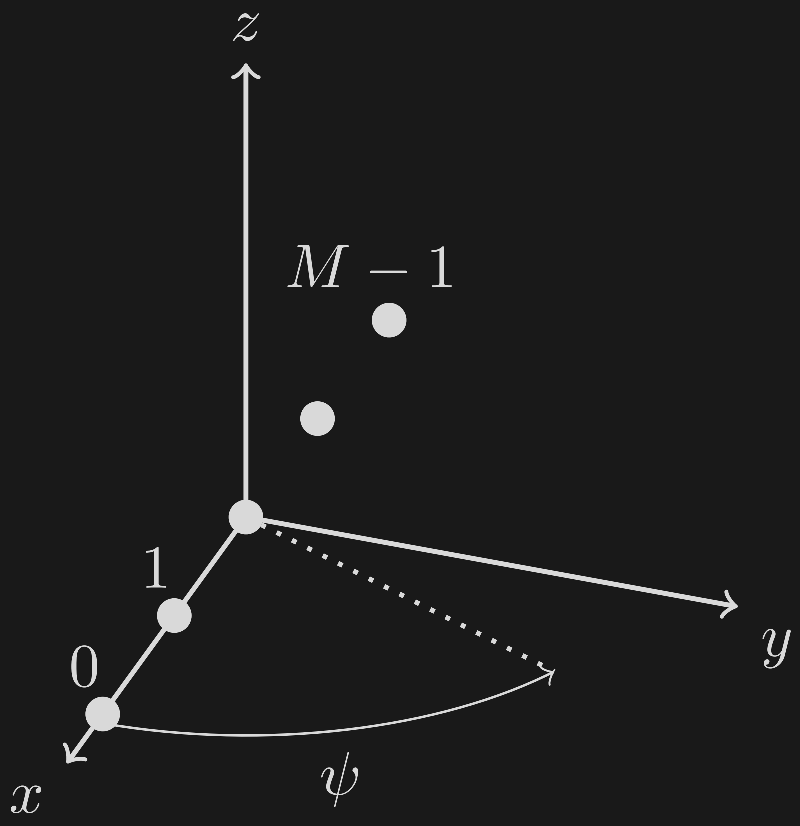
$\boldsymbol{p}_m$: position of sensors

$\boldsymbol{v}(\psi)$: direction vector

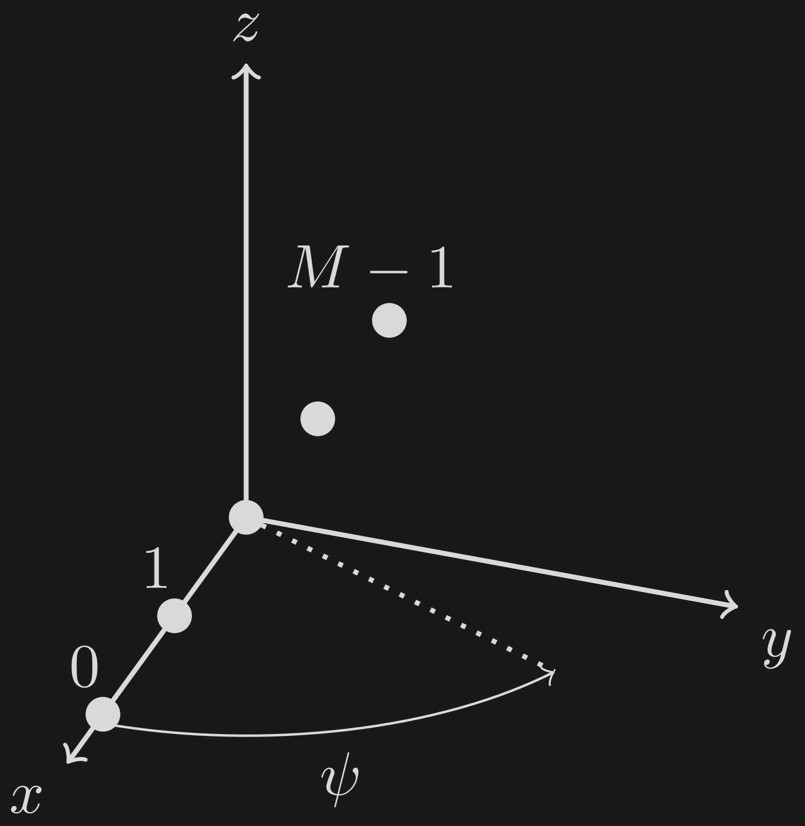
c : speed of sound

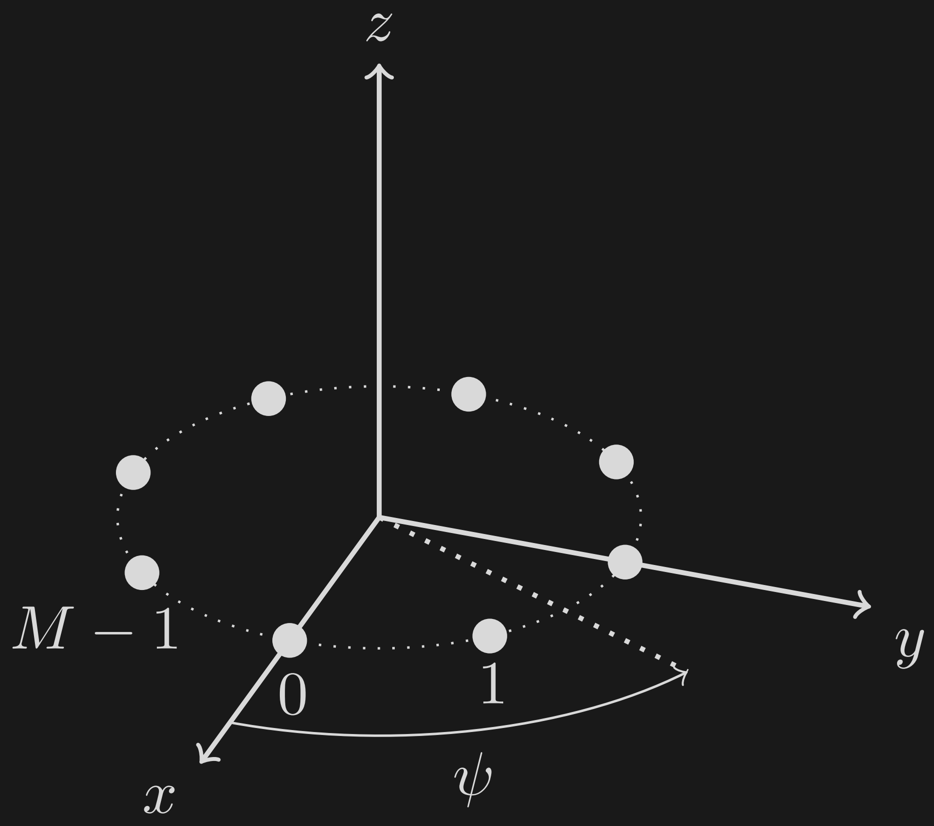
ψ : DOA

Array Structures



Array Structures





Estimation Metods

- Delay-and-Sum Beamformer
- MUSIC
- Bartlett
- MVDR (Capon)
- MCCC

Estimation Metods

- Delay-and-Sum Beamformer
  - MUSIC
  - Bartlett
  - MVDR (Capon)
  - MCCC

Delay-and-Sum Beamformer

$$\text{Signal model} : y_m[k] = s\left[k - \delta_m\right.$$

Delay-and-Sum Beamformer

$$\text{Signal model} : y_m[k] = s\left[k - \delta_m\right.$$

Delay-and-Sum Beamformer:

$$\hat{\psi} = \arg\max_{\psi} \sum_{k=1}^{N} \left| \frac{1}{M} \sum_{i=1}^{M} y_m[k + \right.$$

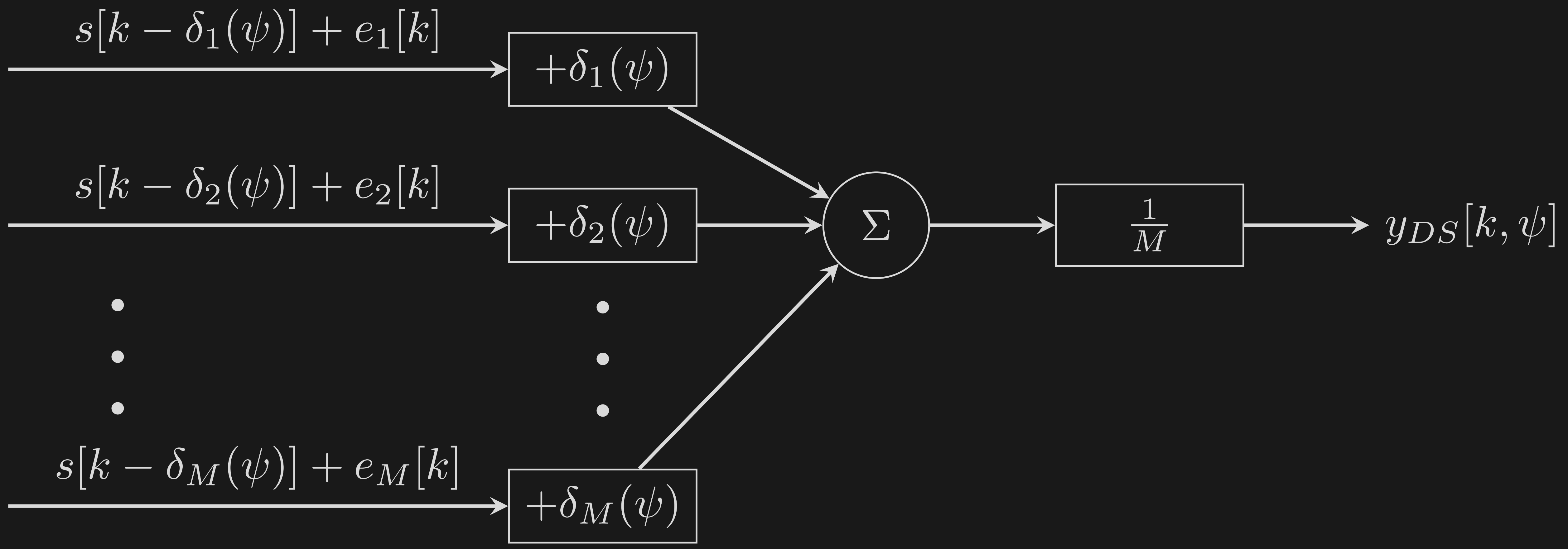
Delay-and-Sum Beamformer

$$\text{Signal model : } y_m[k] = s\,[k - \delta_m$$

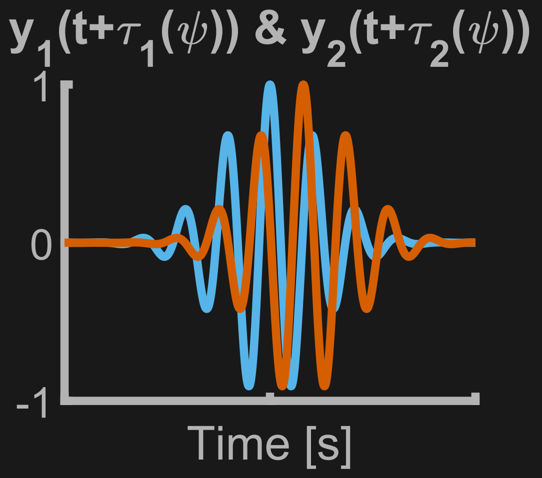
Delay-and-Sum Beamformer:

$$\hat{\psi} = \arg\max_{\psi} \sum_{k=1}^{N} \Bigg| \underbrace{\frac{1}{M} \sum_{i=1}^{M} y_m[k +}_{y_{DS}[k,\psi]}$$

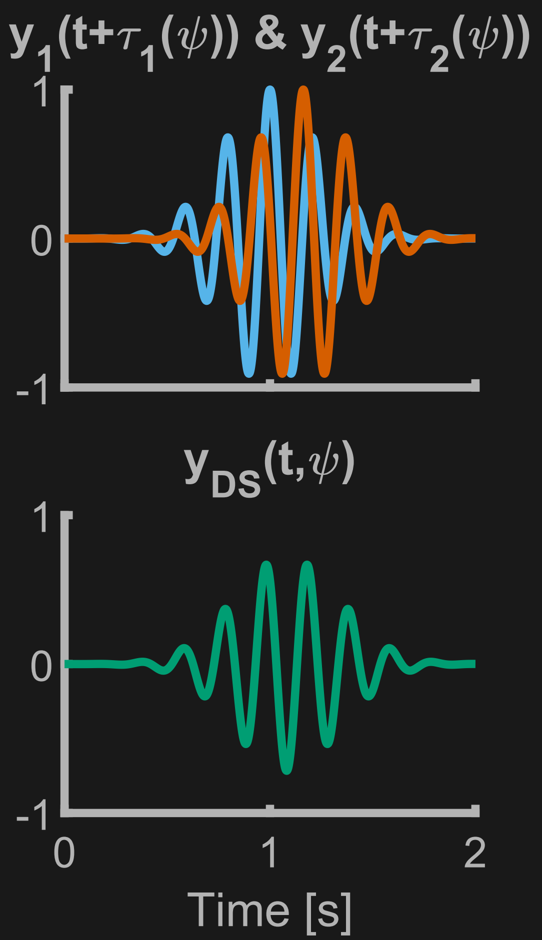
Delay-and-Sum Beamformer



Direction of Arrival Estimation

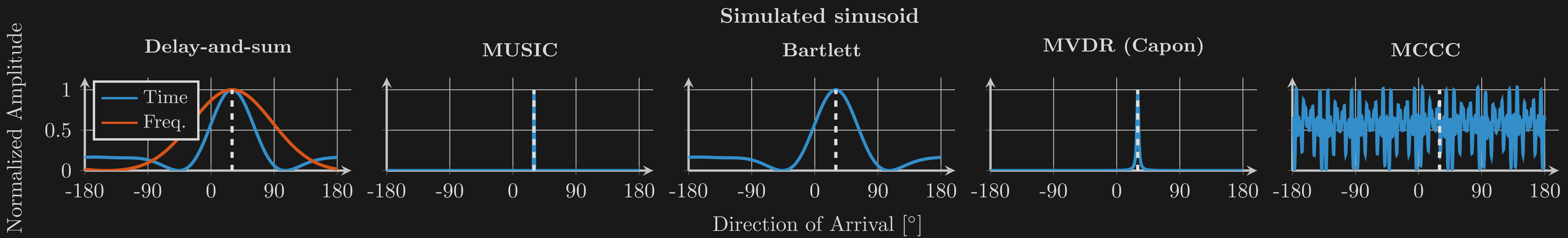


Direction of Arrival Estimation

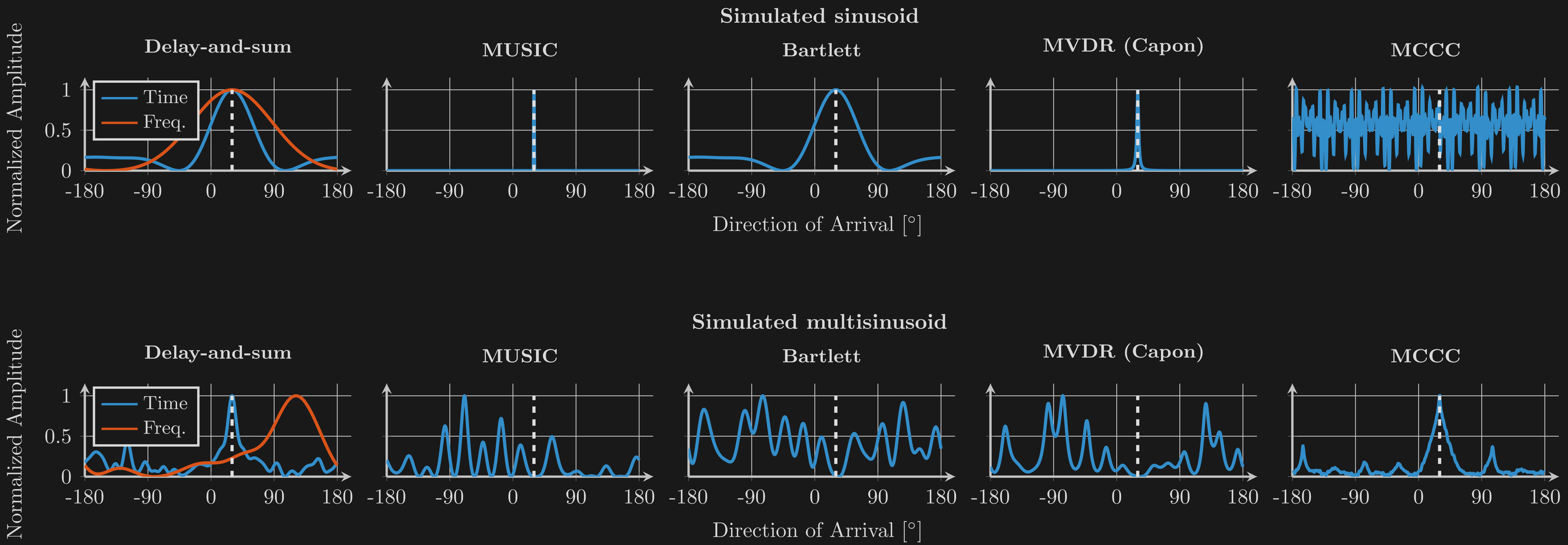


Direction of Arrival Estimation

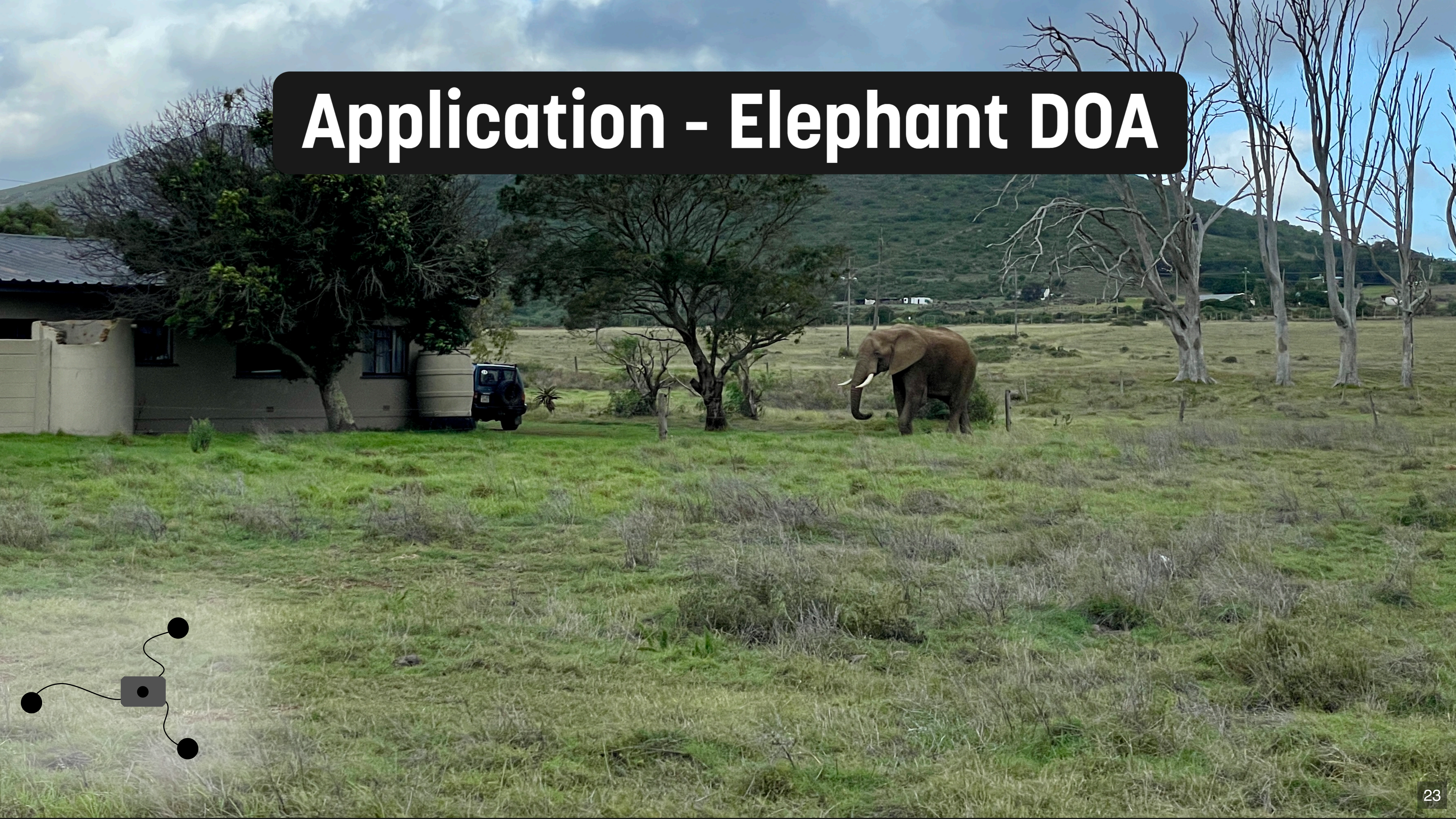
Method Comparision



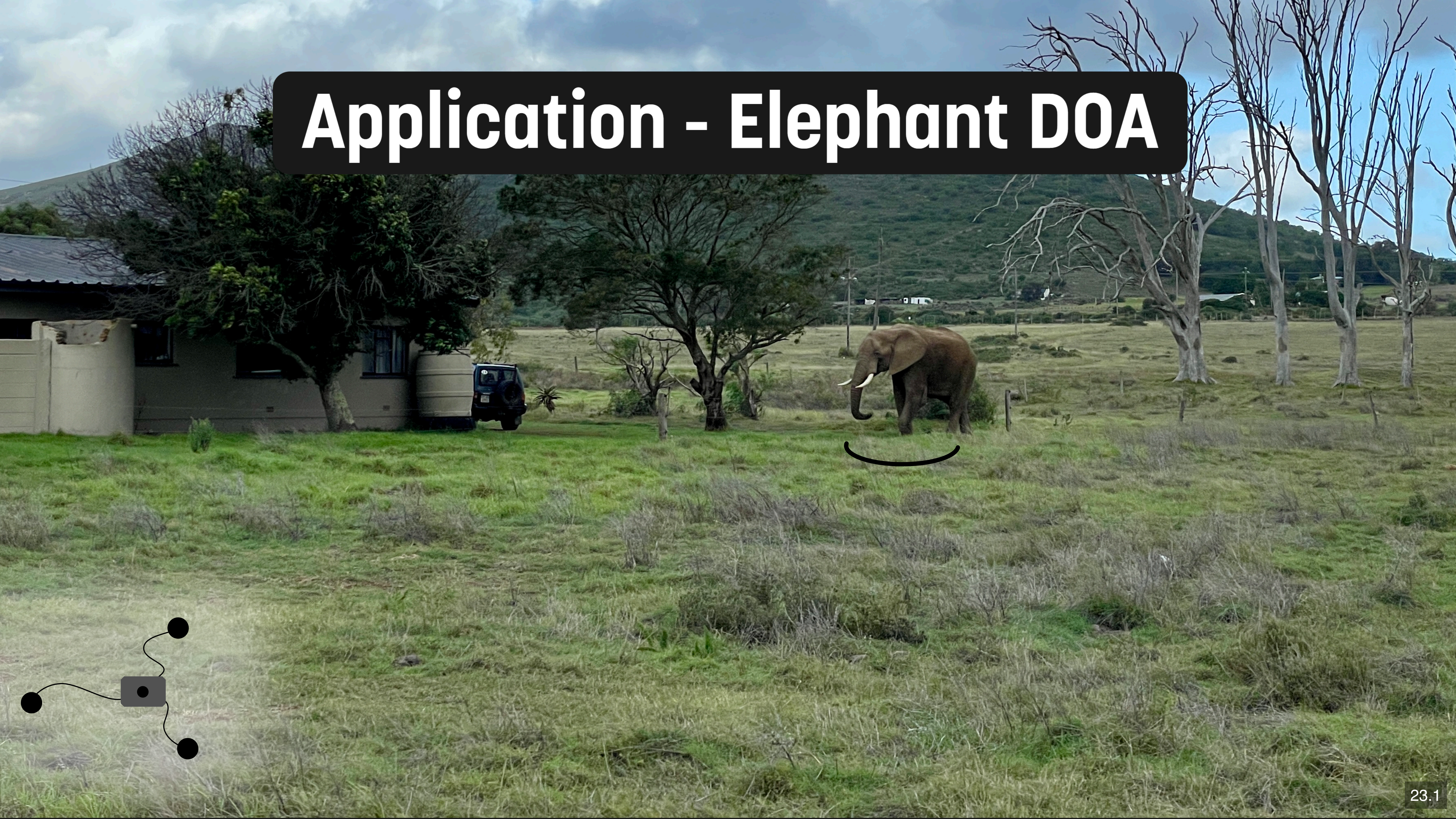
Method Comparision

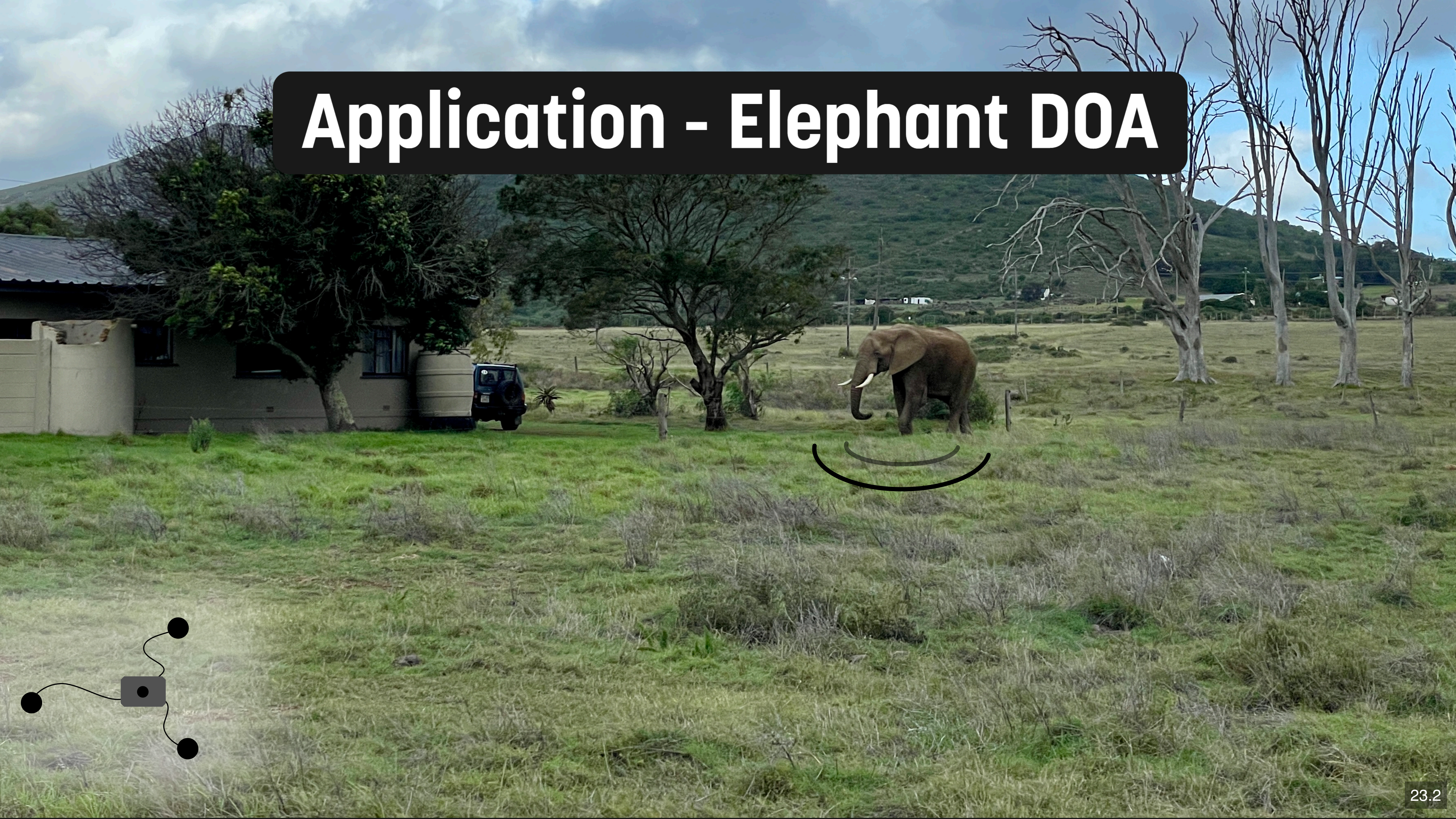


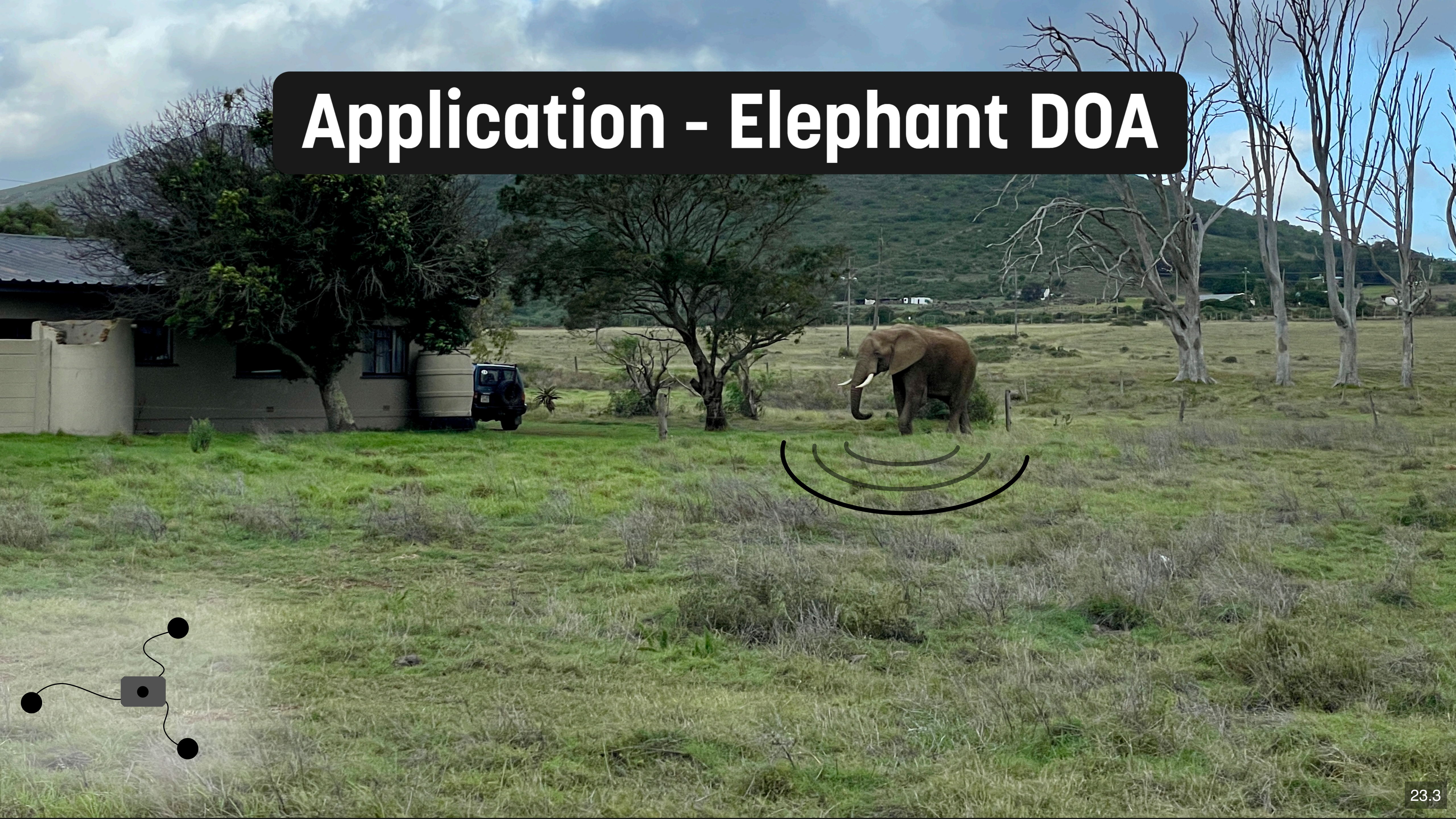
Application - Elephant DOA



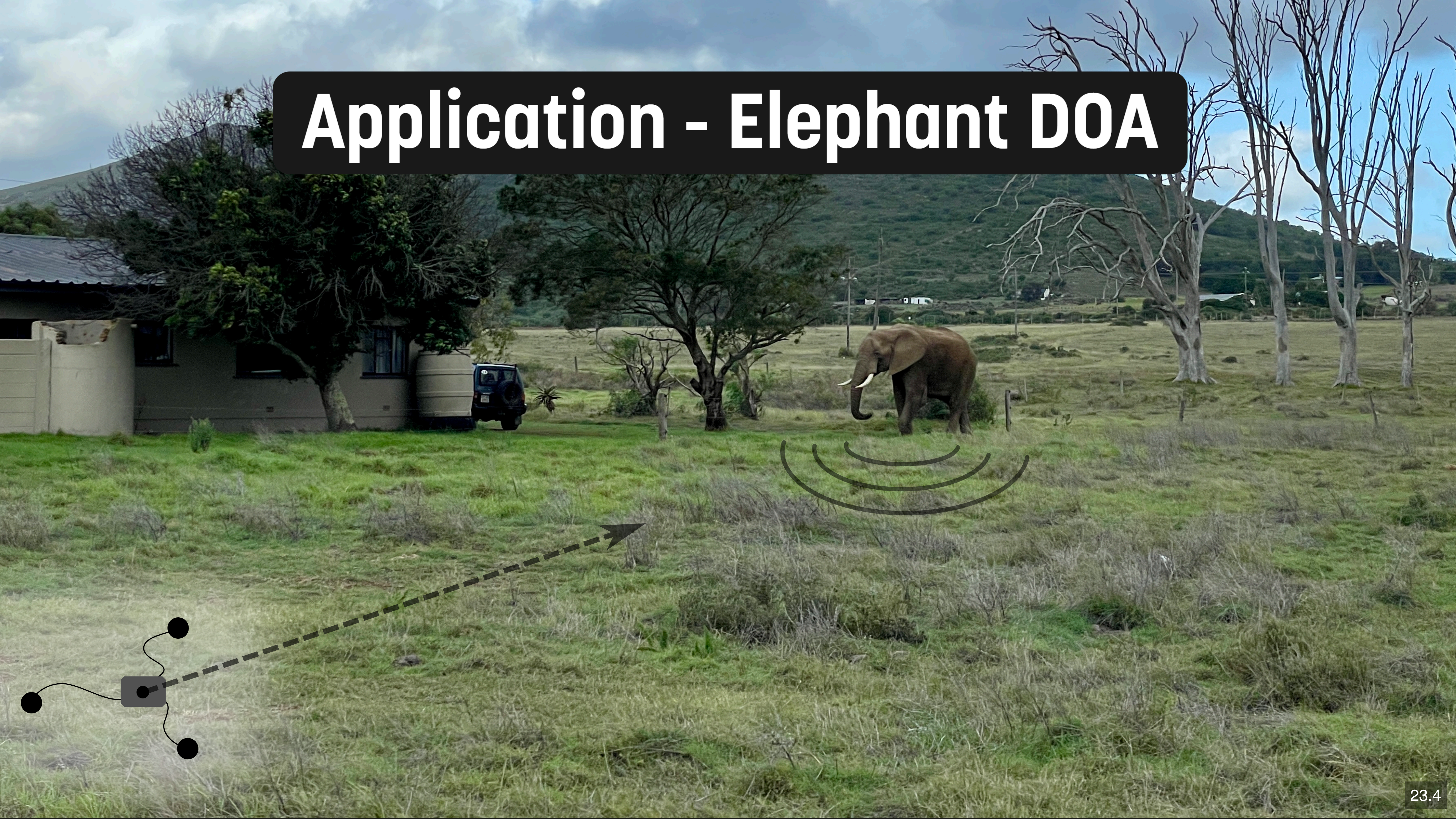
Application - Elephant DOA




Application - Elephant DOA
23.2

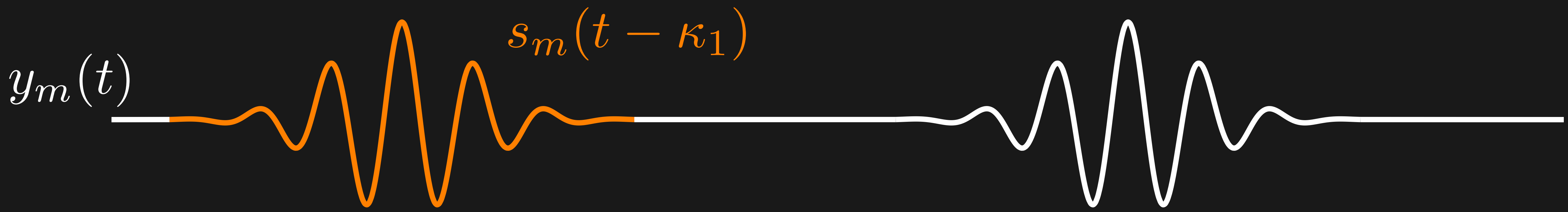

Application - Elephant DOA
23.3

Application - Elephant DOA

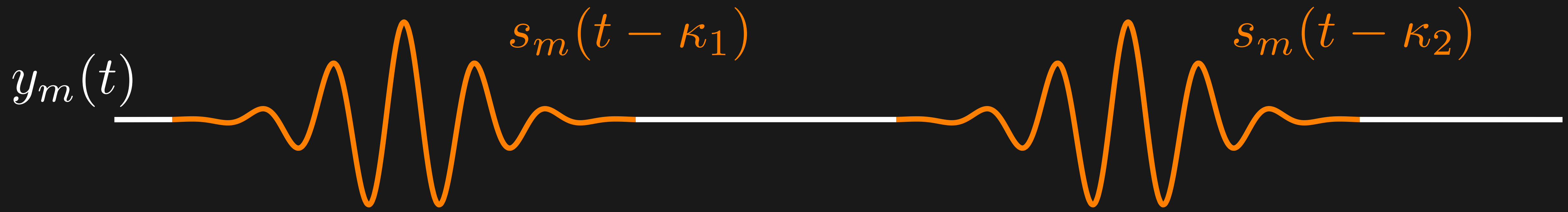


Signal Model

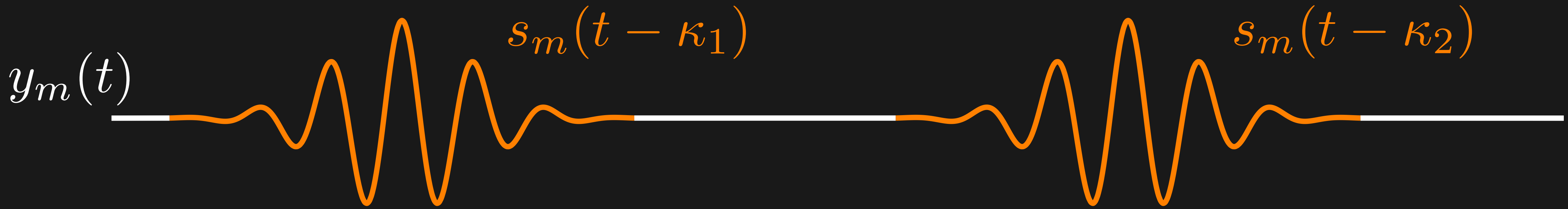
Signal Model



Signal Model

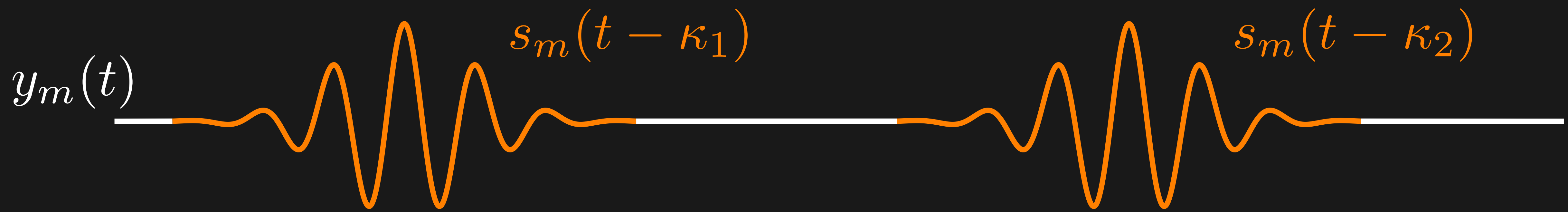


Signal Model



$$y_m(t) = \sum_n s_m\left(t-\kappa_n\right) + e_m(t),$$

Signal Model



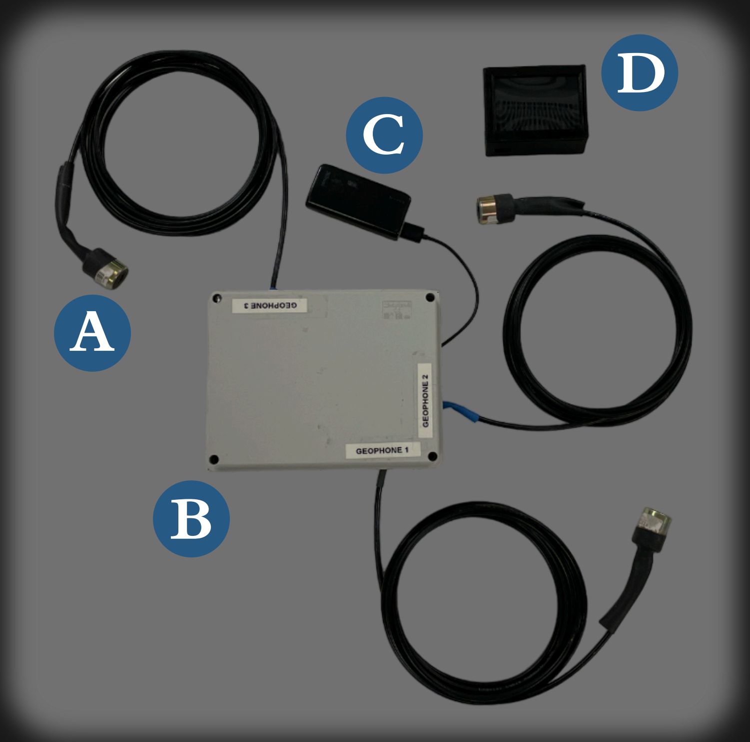
$$y_m(t) = \sum_n s_m\left(t - \kappa_n\right) + e_m(t),$$

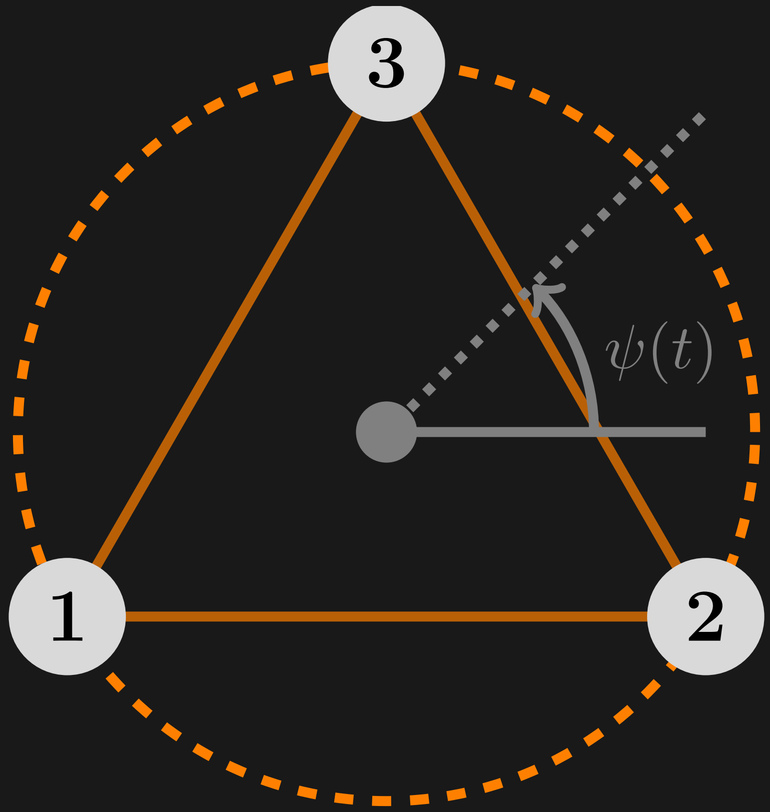
$s_m(t-\kappa_n)$: nth footstep for geophone m

κ_n: time for footstep n

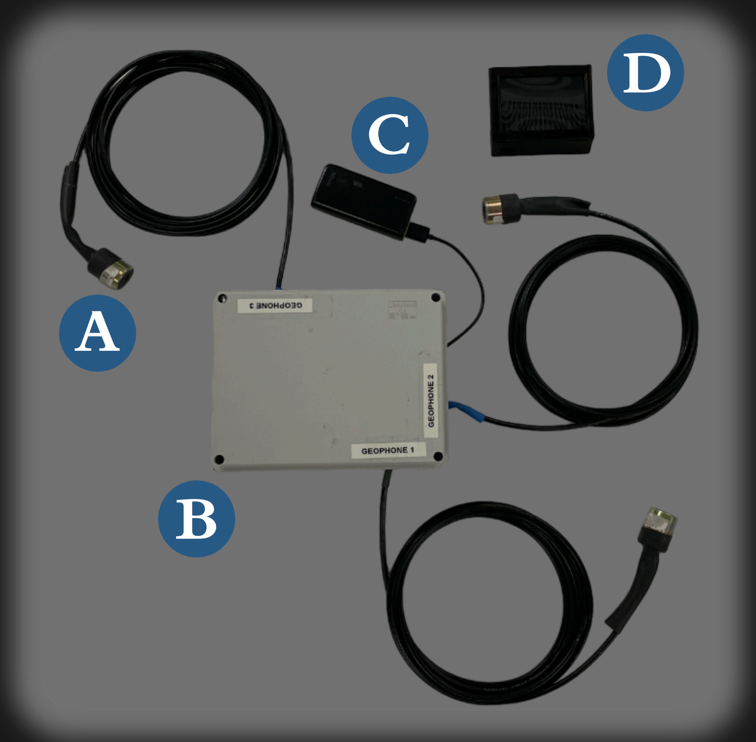
$e_m(t)$: ambient noise

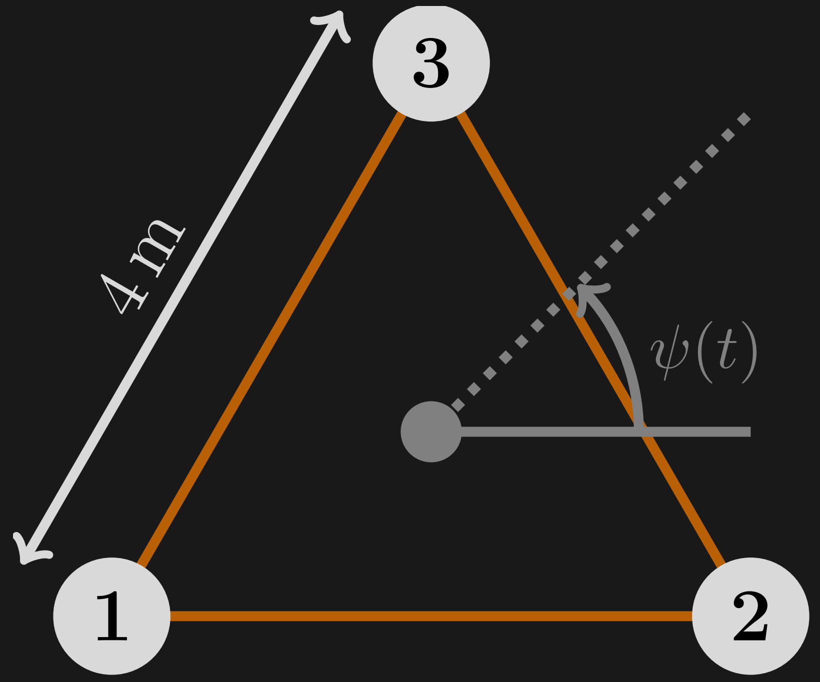
Experimental setup





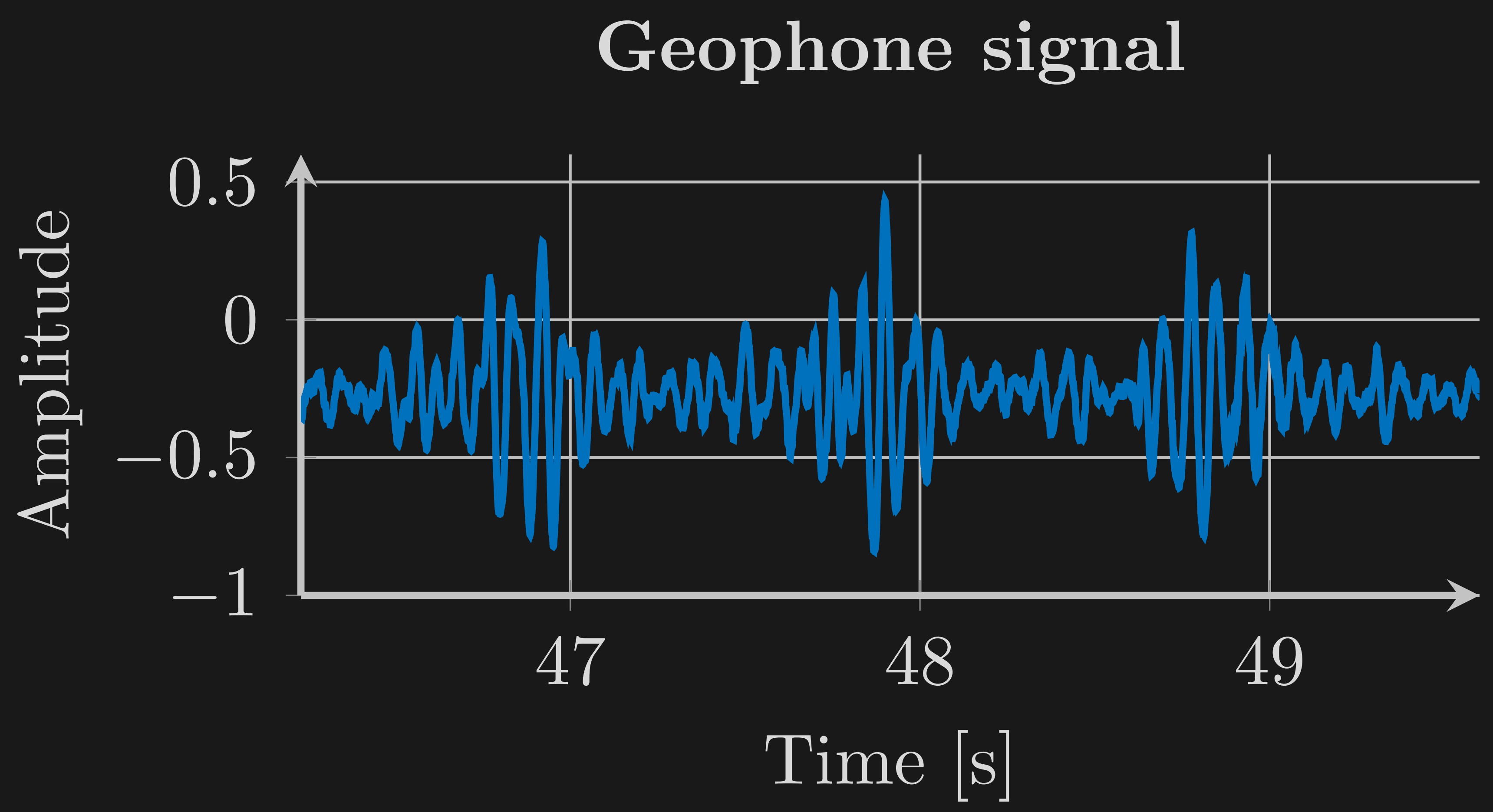
Experimental setup

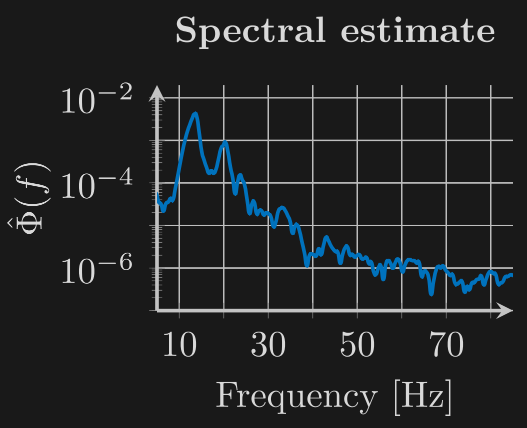




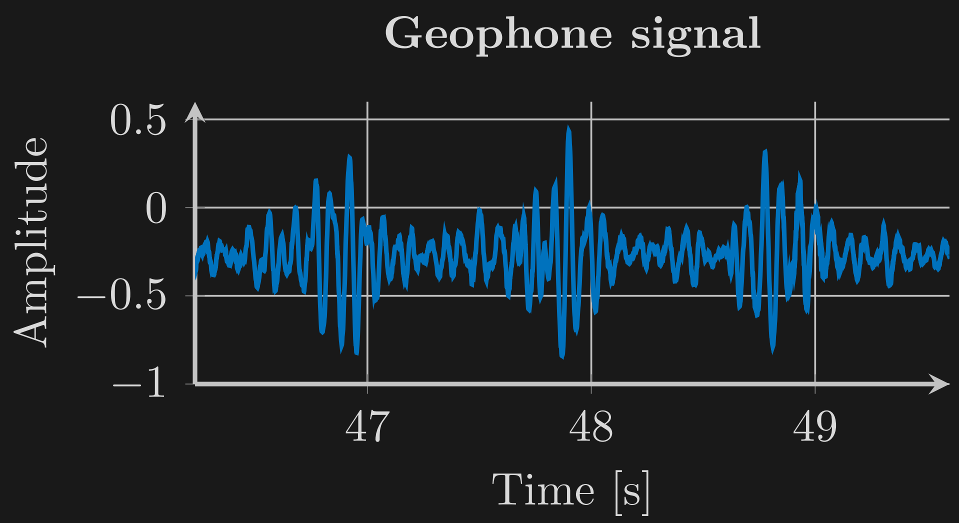
Experimental setup

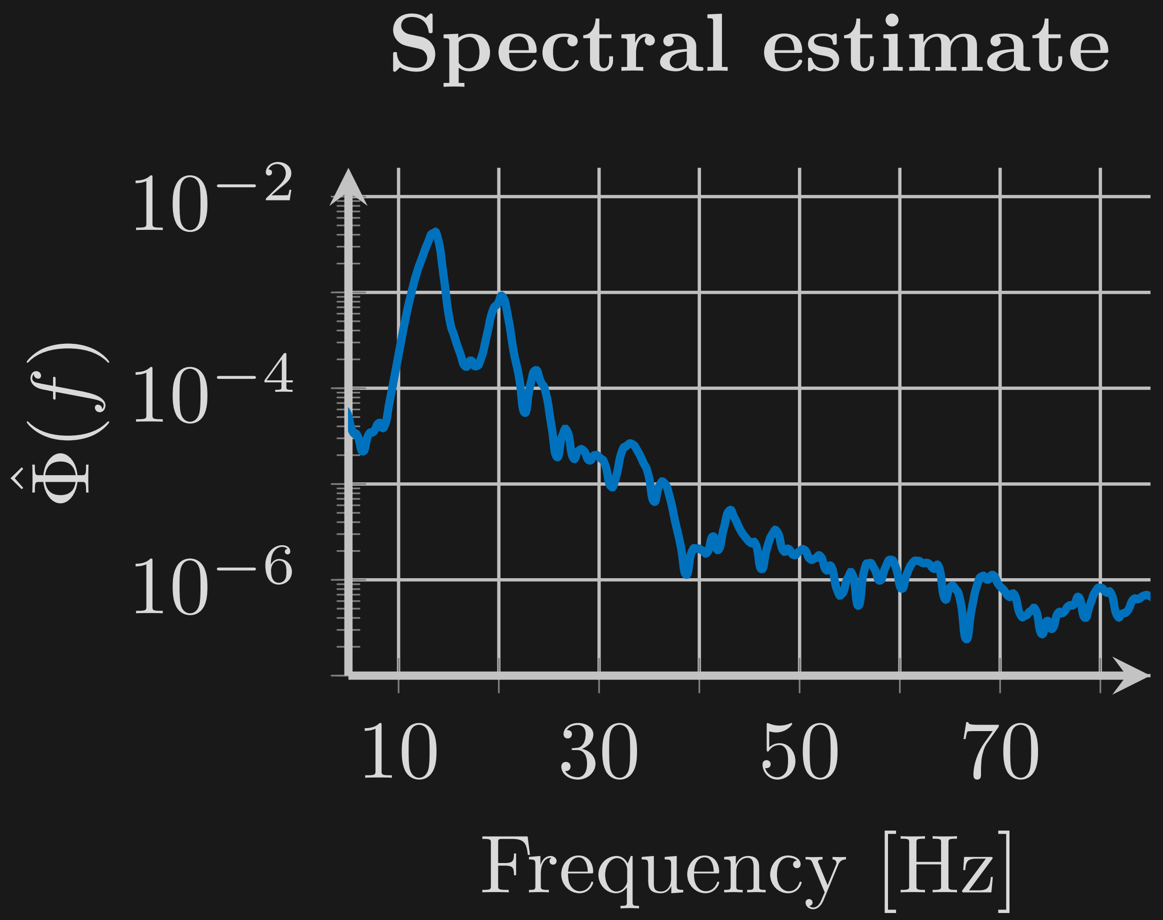
Signal Characteristics



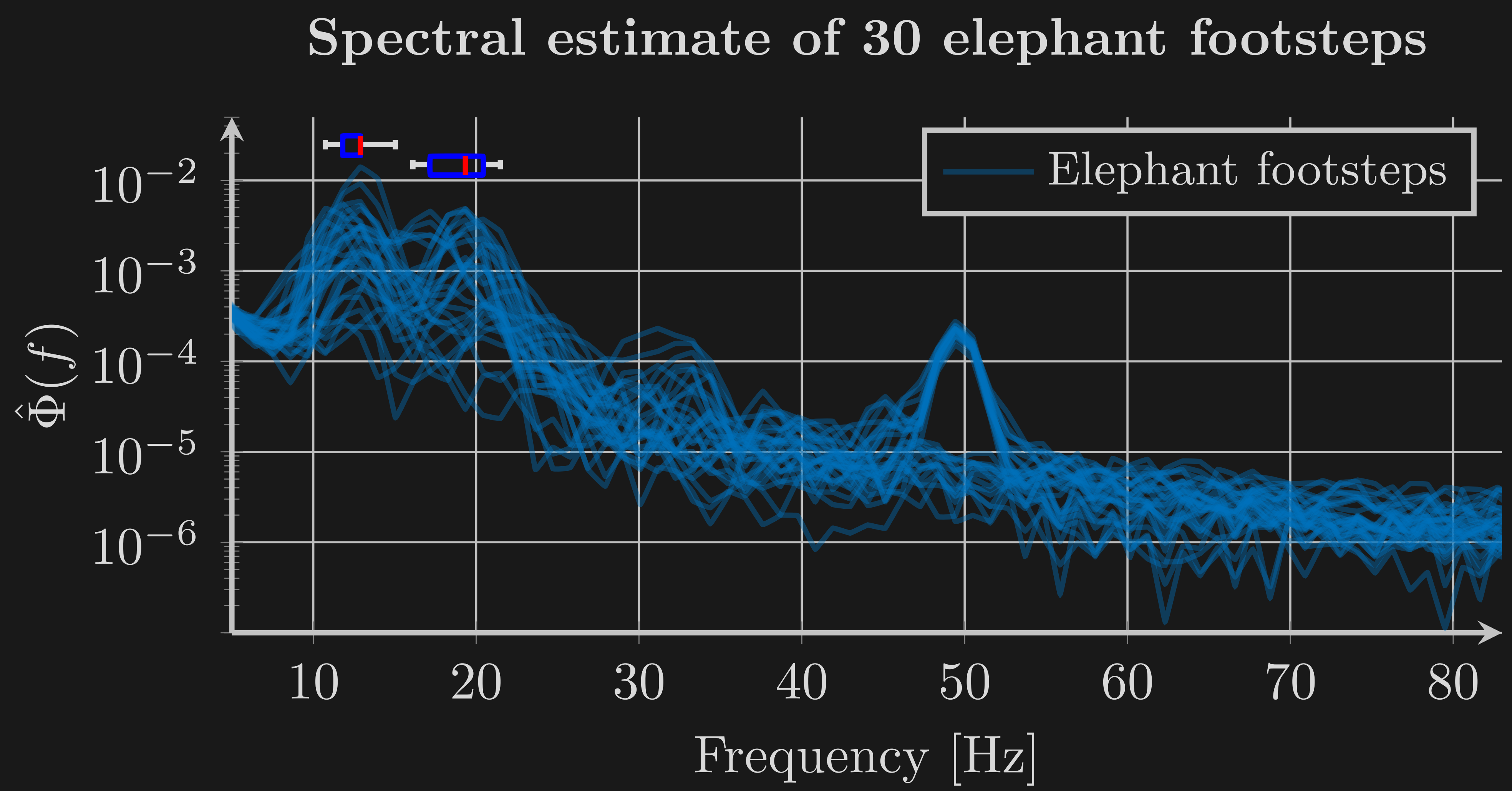


Signal Characteristics

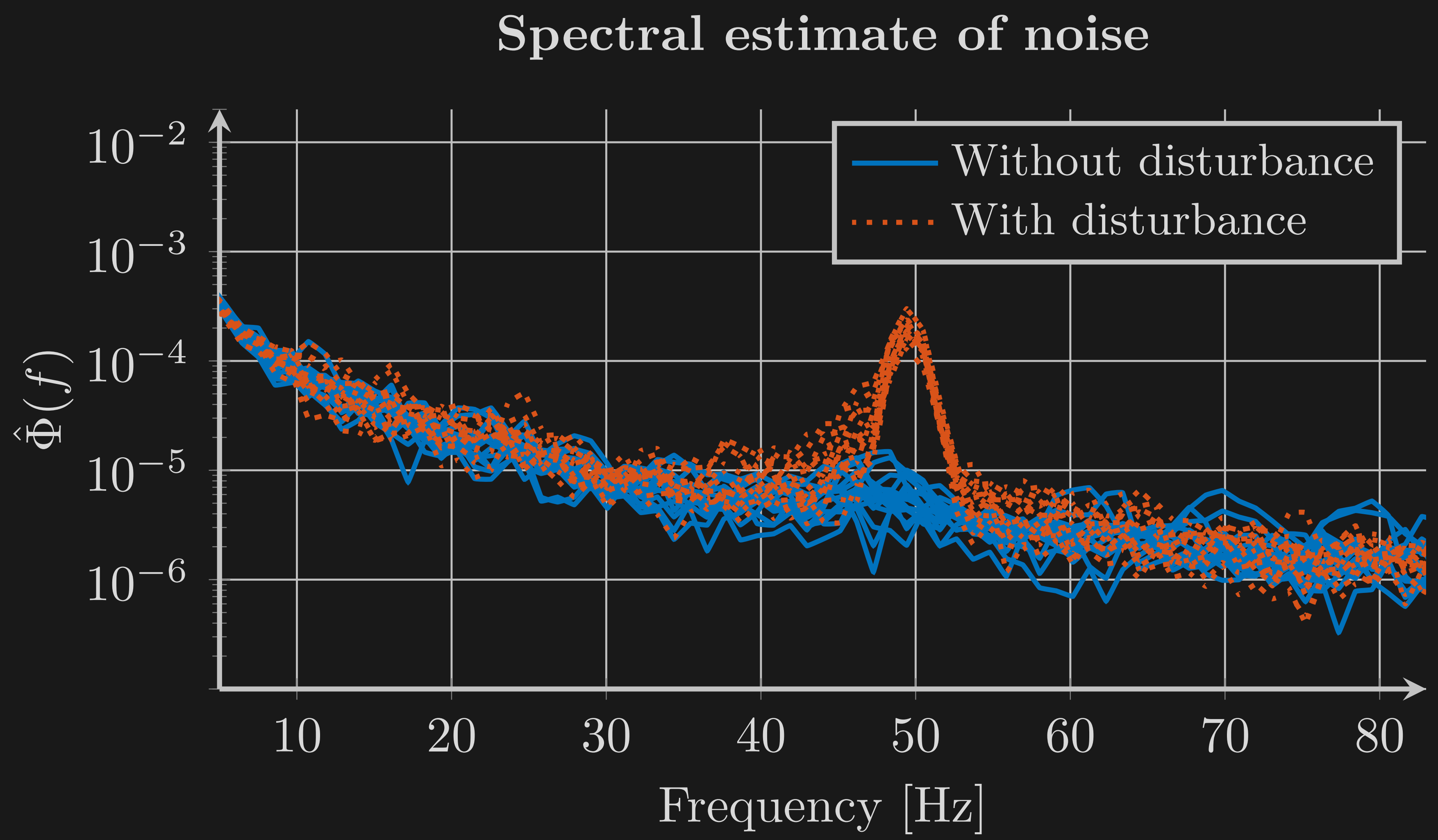




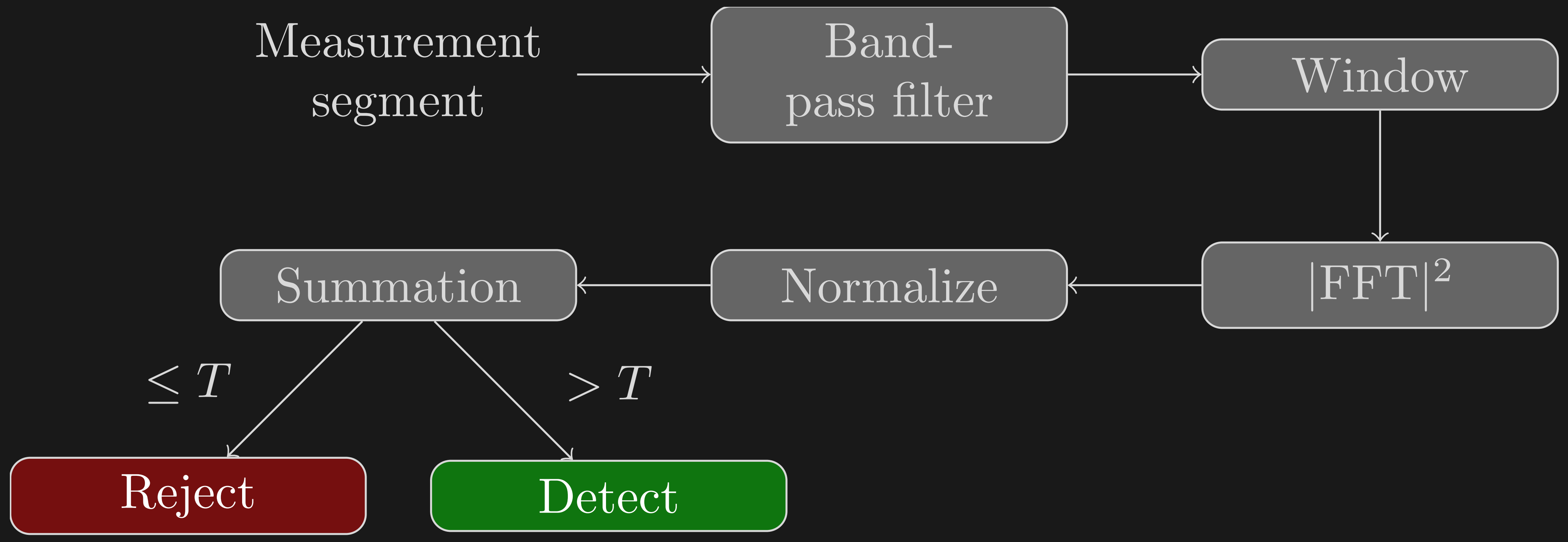
Signal Characteristics



Noise Characteristics



Elephant Footstep Detection



Pre-processing

- Bandpass filter between 4 and 30 Hz
- Hanning Window

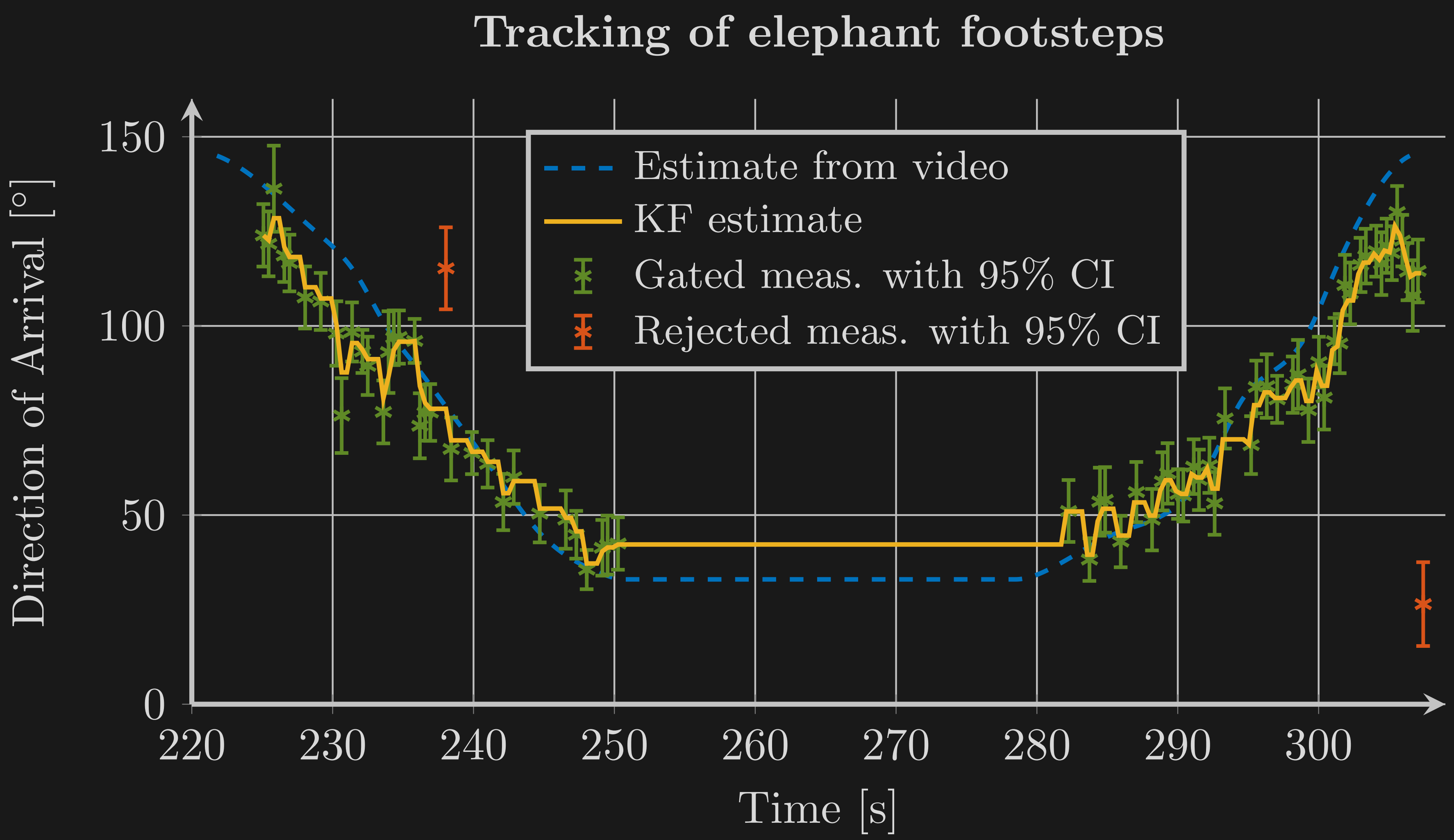
Signal Feature

- Fast Fourier transform.
- Normalized amplitude.
- Frequency content between 8 and 23 Hz.

$$\frac{\text{Energy in 8-23 Hz}}{\text{Energy in 4-30 Hz}}$$

Detection Results

Tracking Results



Tracking Results

Tracking Results

Conclusions

- Detection and DOA estimation of elephants

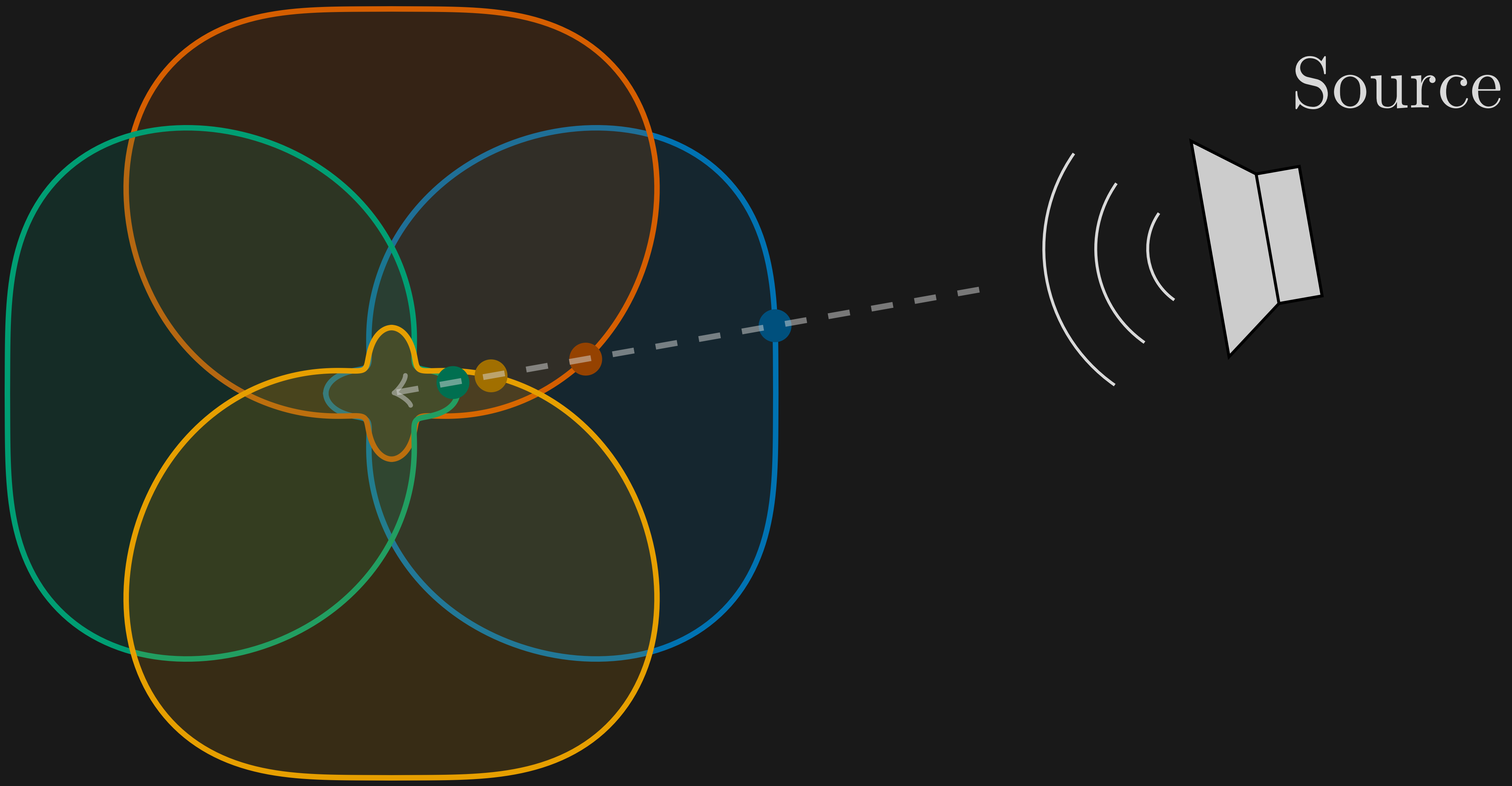
Conclusions

- Detection and DOA estimation of elephants
- Works well within 15-40 meters

Conclusions

- Detection and DOA estimation of elephants
- Works well within 15-40 meters
- Direction of arrival estimate is accurate

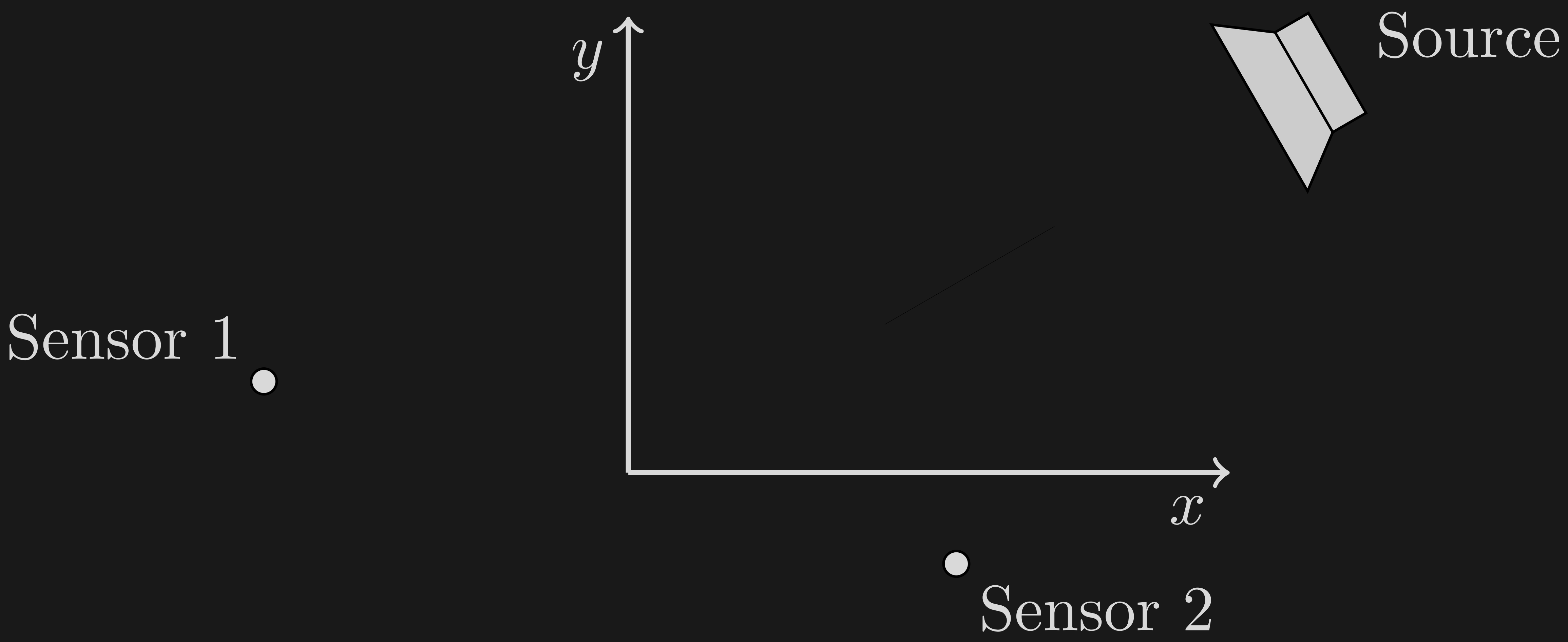
Directional Sensitivity DOA



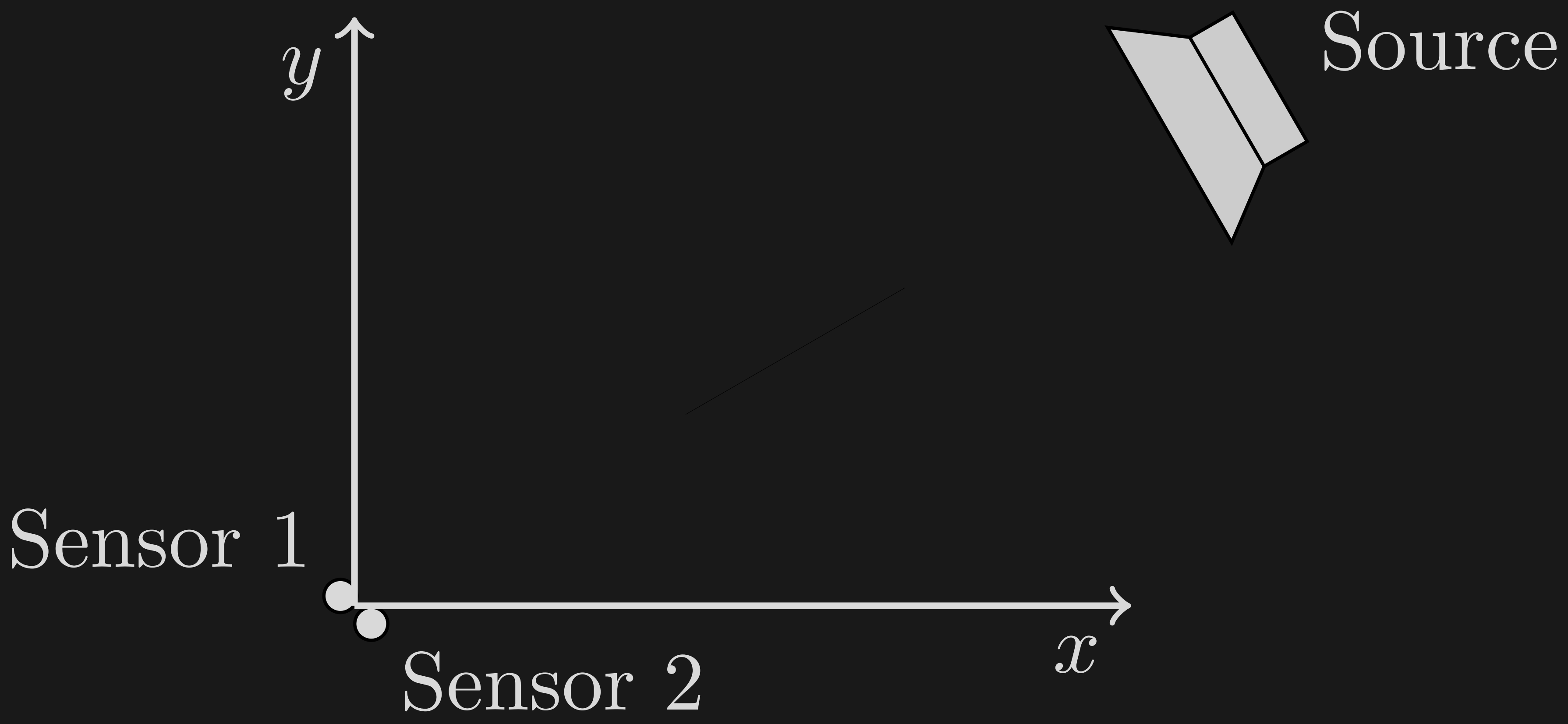
Motivation

- Classical methods rely on time-differences.
- Requries separation of the sensors in space.
- Placing them to far appart leads to ambiguity in DOA.
- Placing them to close decrease DOA resolution.

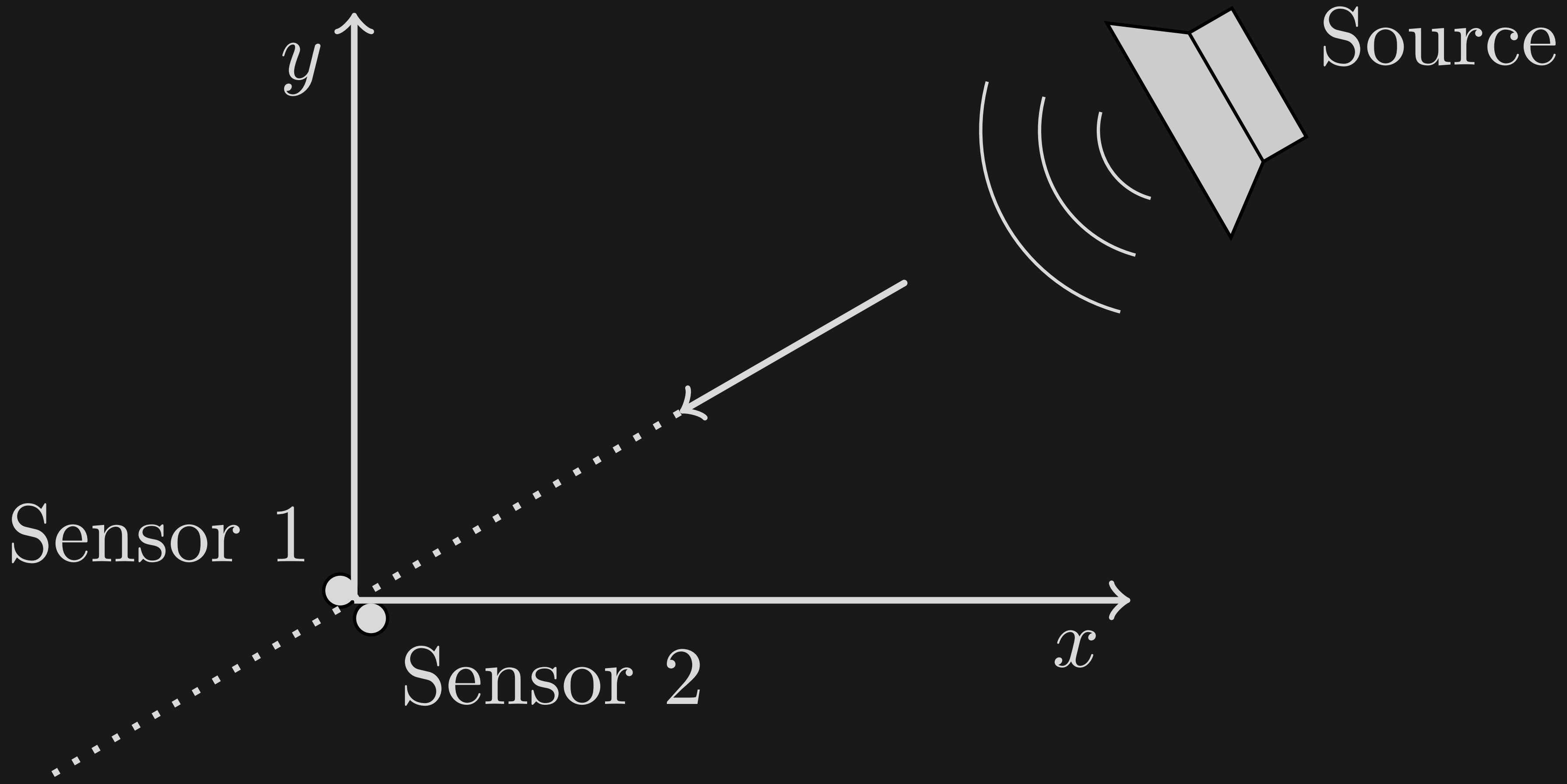
Motivation



Motivation



Motivation

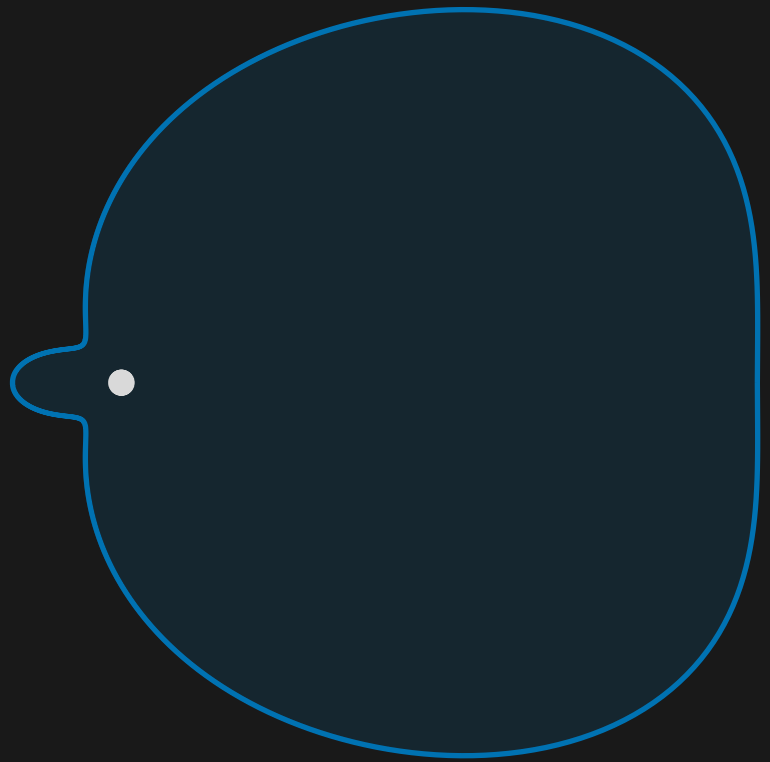


General Idea

General Idea

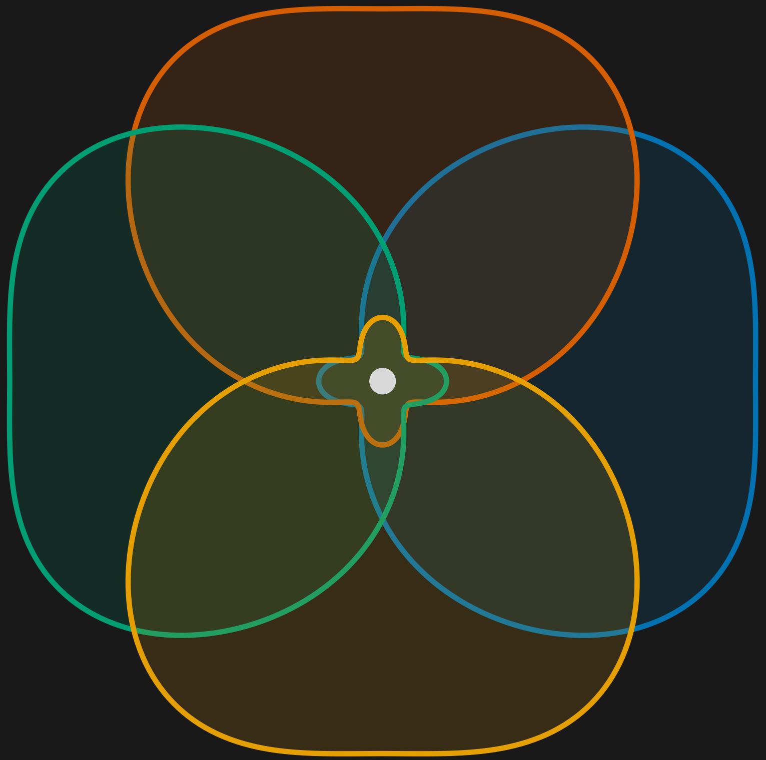
General Idea

General Idea

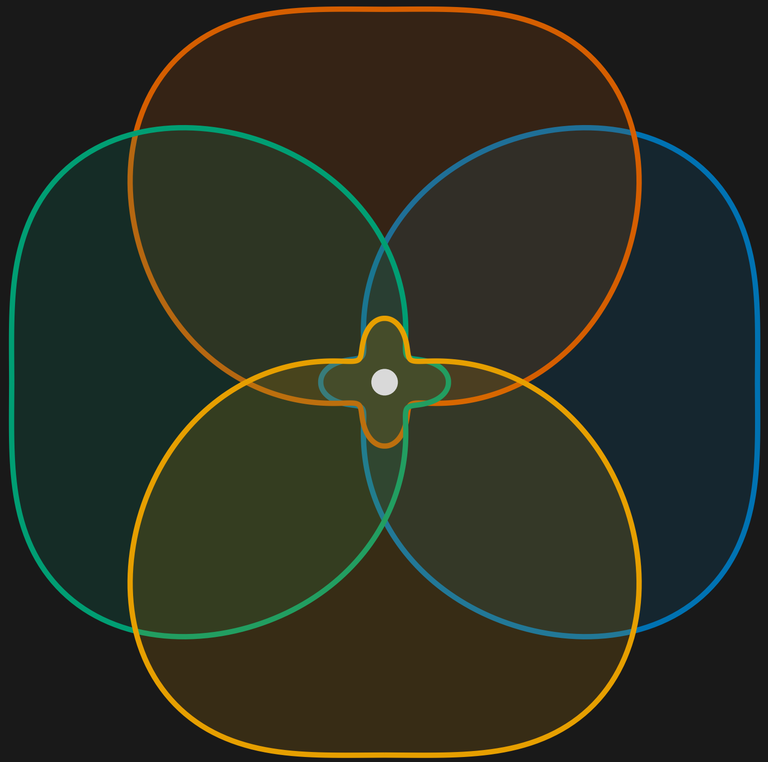


General Idea

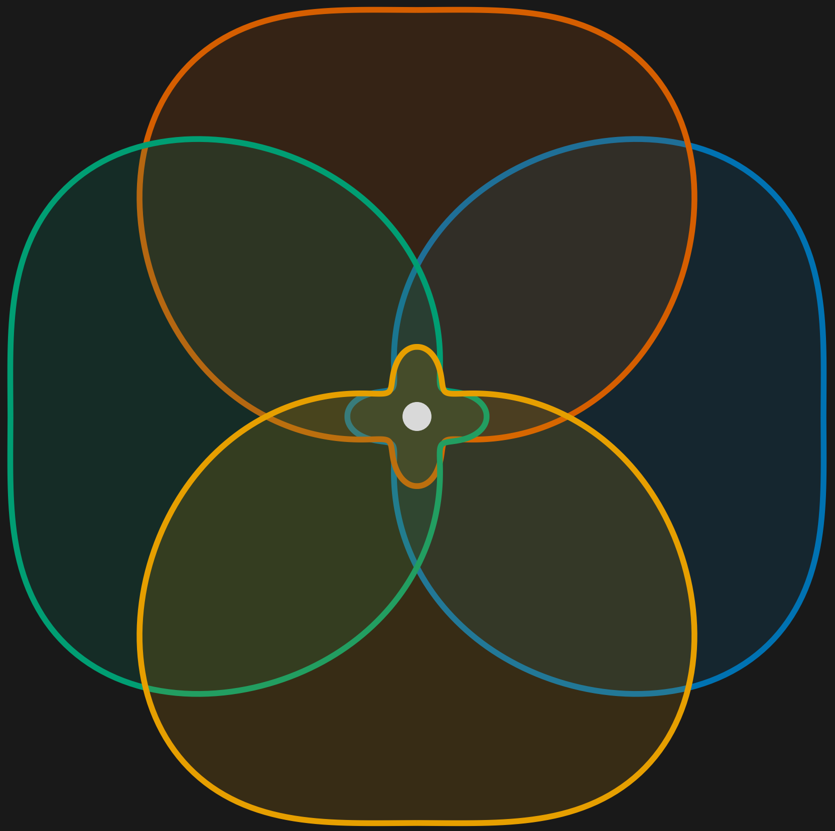
General Idea



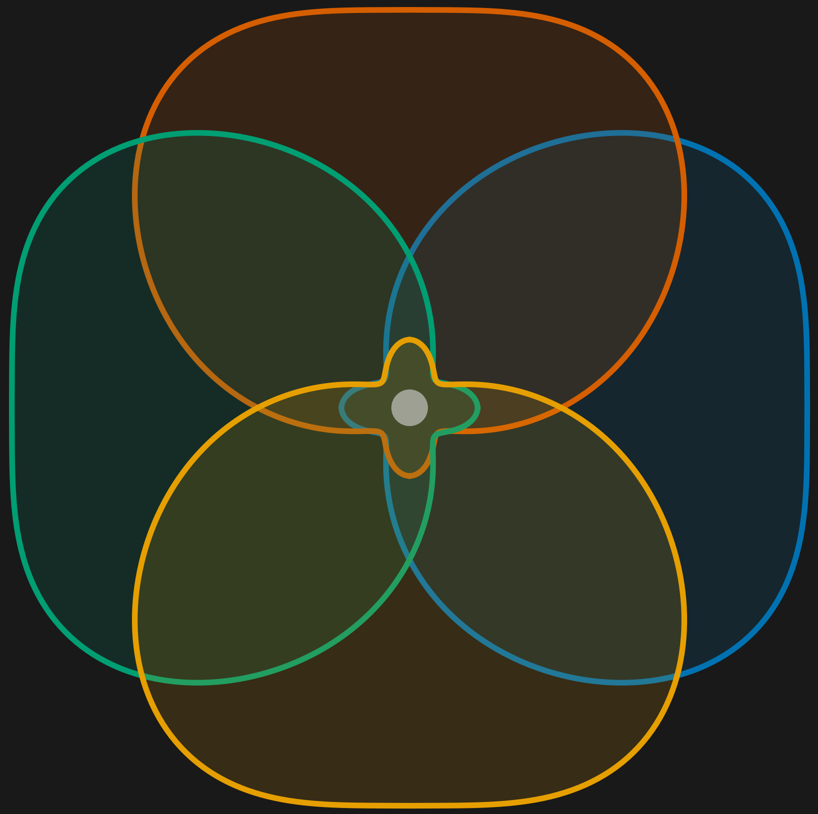
General Idea

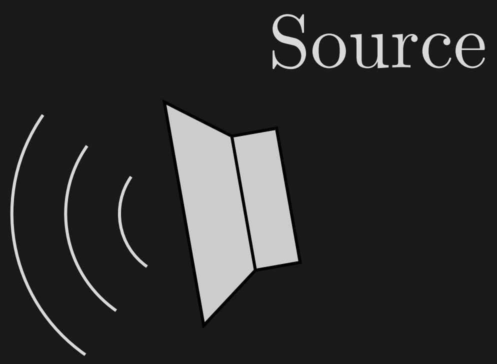


General Idea

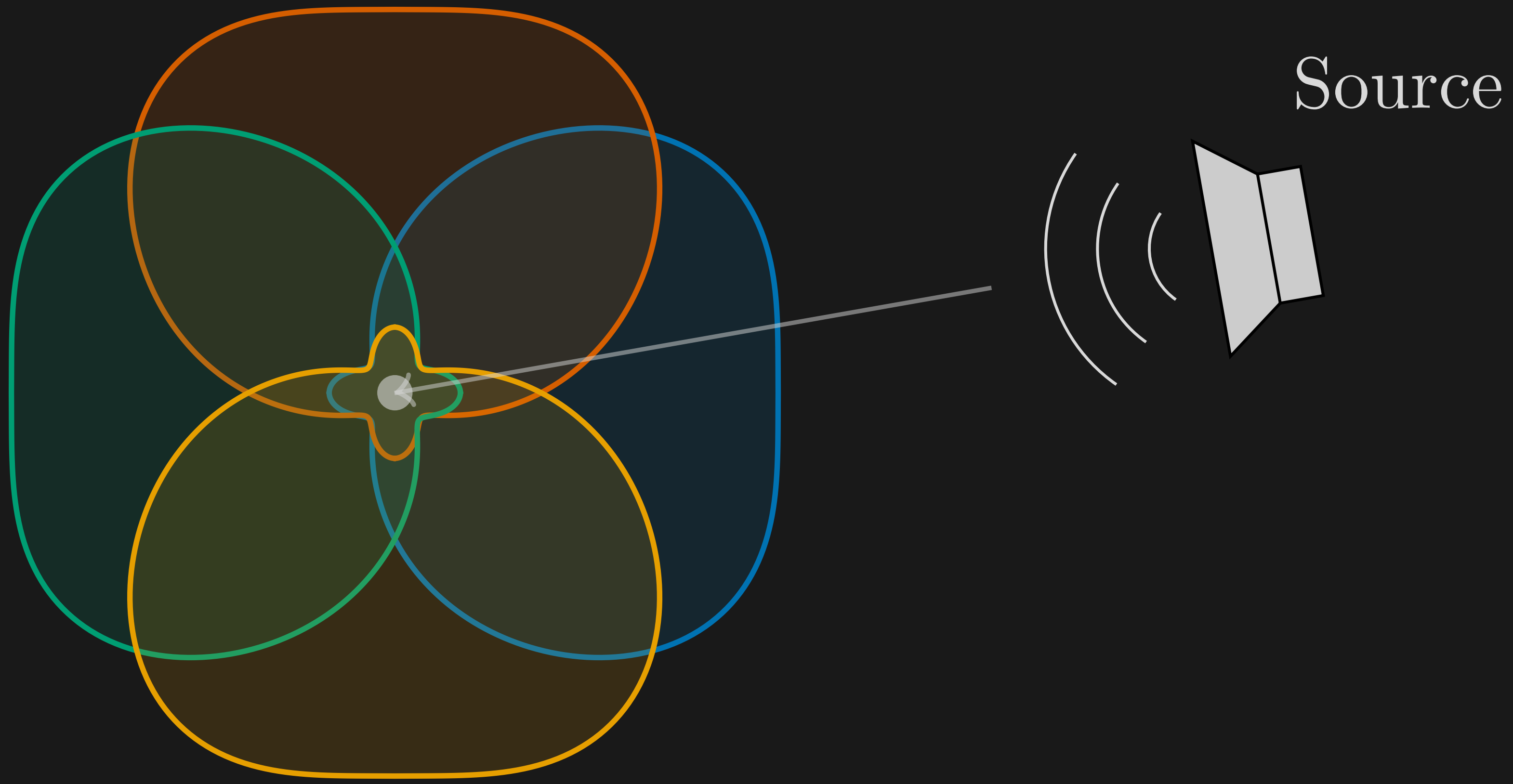


General Idea

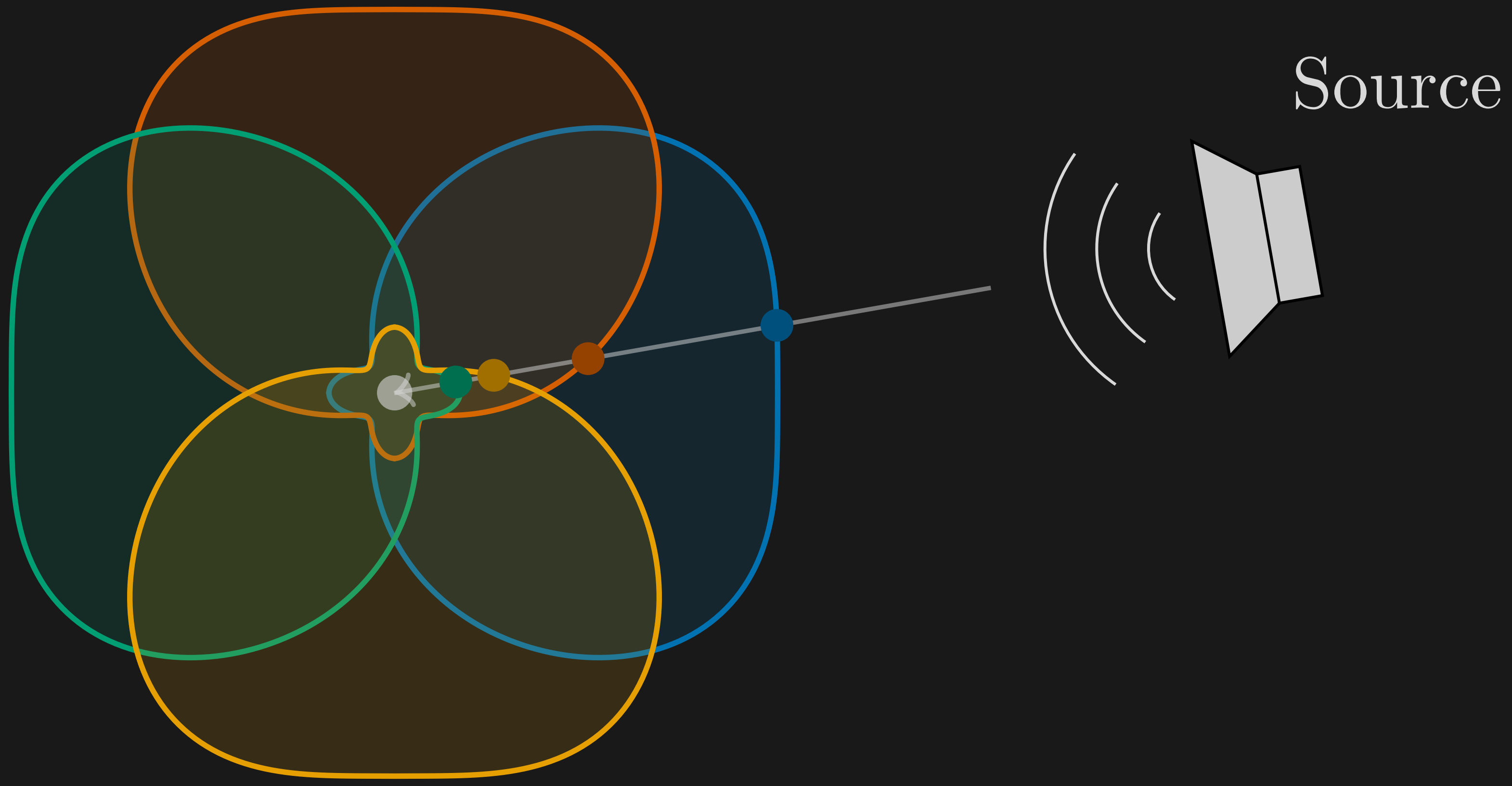




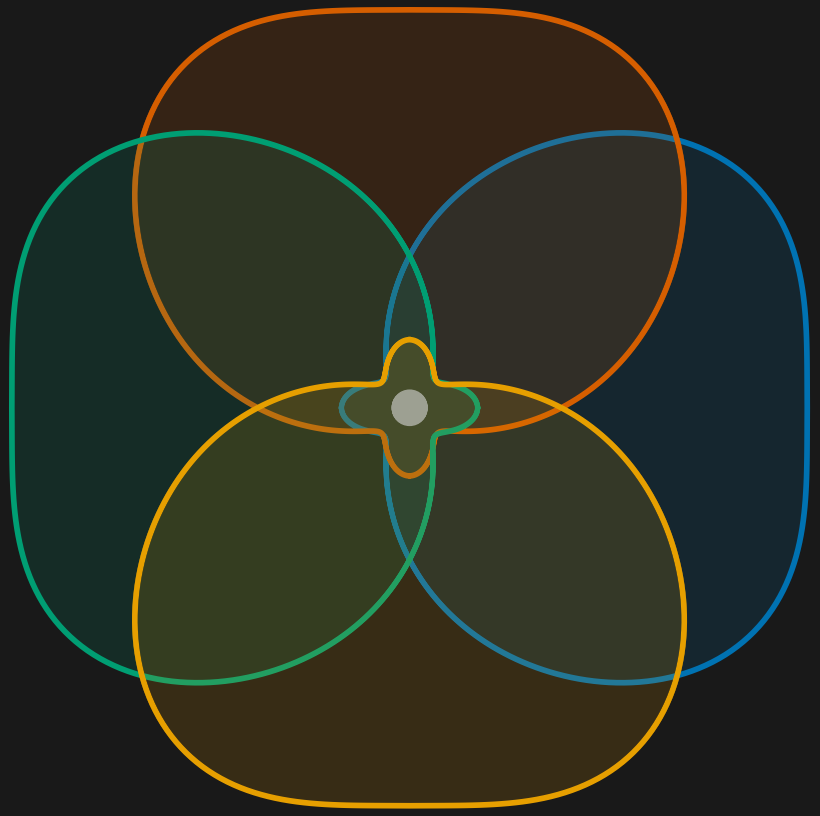
General Idea



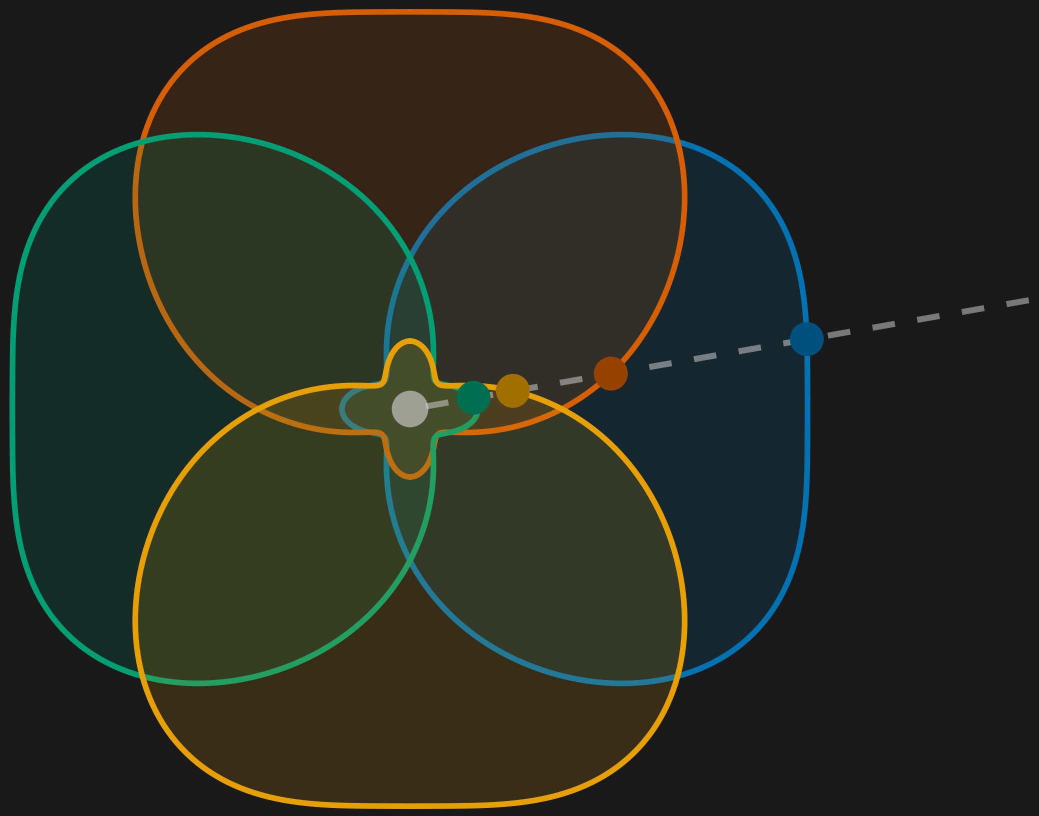
General Idea



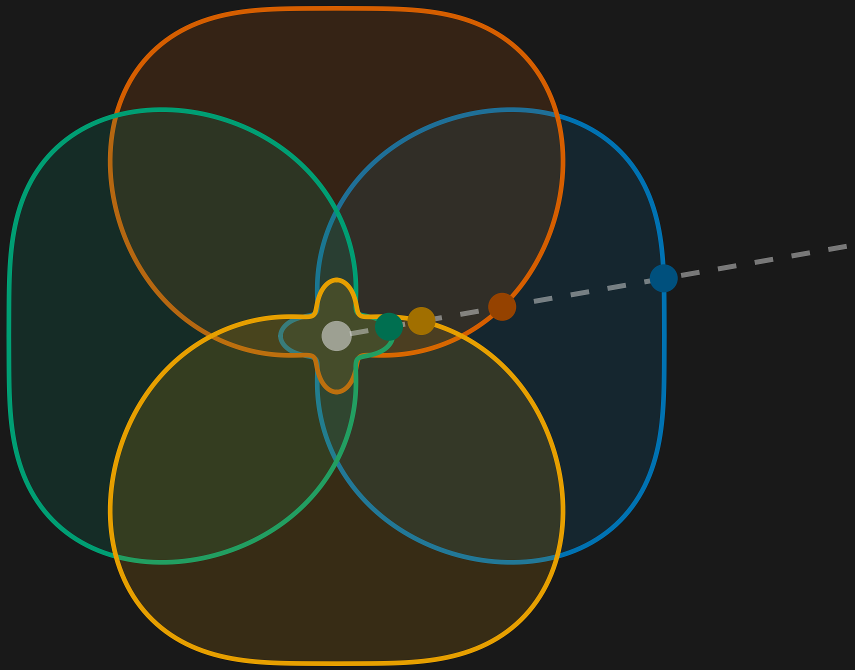
General Idea

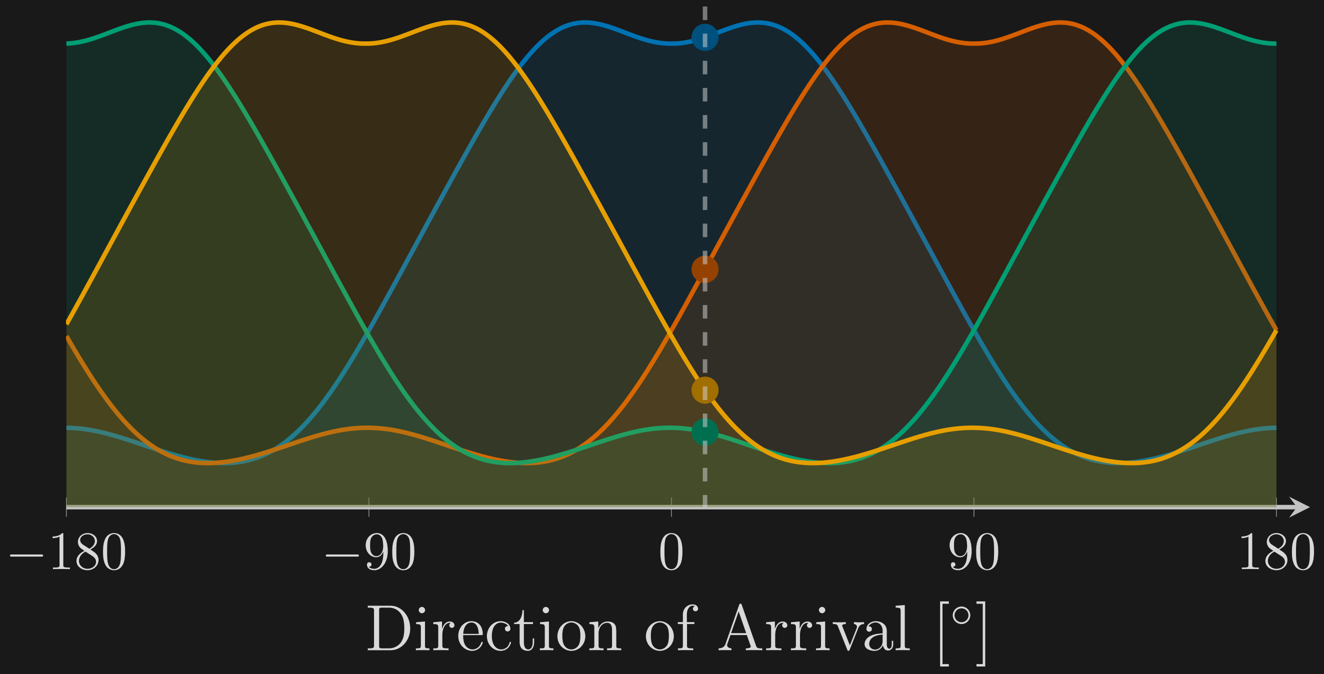


General Idea



General Idea





Signal model

$$y_m[l] = s[l - \delta_m] + e[l]$$

Signal model

$$y_m[l] = s[l - \delta_m] + e[l]$$

$$P_m = \sum_{l=1}^{L} y_m^2[l]$$

Signal model

$$y_m[l] = s_m[l] + e[l]$$

$$P_m = \sum_{l=1}^{L} y_m^2[l]$$

Signal model

$$y_m[l] = s_m[l] + e[l]$$

$$P_m = \underbrace{\frac{1}{L}\sum_{l=1}^{L} s_m^2[l]}_{P_m^s} + \underbrace{\frac{1}{L}\sum_{l=1}^{L} 2 s_m[l] e_m[l]}_{P_m^{sw}} + \underbrace{\frac{1}{L}\sum_{l=1}^{L} e_m^2[l]}_{v_m},$$

Signal model

$$y_m[l] = s_m[l] + e[l], \quad e[l] \sim \mathcal{N}(0, \sigma_m^2)$$

$$P_m = \underbrace{\frac{1}{L}\sum_{l=1}^{L} s_m^2[l]}_{P_m^s} + \underbrace{\frac{1}{L}\sum_{l=1}^{L} 2 s_m[l] e_m[l]}_{P_m^{sw}} + \underbrace{\frac{1}{L}\sum_{l=1}^{L} e_m^2[l]}_{v_m},$$

Signal model

$$y_m[l] = s_m[l] + e[l], \quad e[l] \sim \mathcal{N}(0, \sigma_m^2)$$

$$P_m = \underbrace{\frac{1}{L}\sum_{l=1}^{L} s_m^2[l]}_{P_m^s} + \underbrace{\frac{1}{L}\sum_{l=1}^{L} 2s_m[l]e_m[l]}_{P_m^{sw} \to 0} + \underbrace{\frac{1}{L}\sum_{l=1}^{L} e_m^2[l]}_{v_m},$$

Signal model

$$y_m[l] = s_m[l] + e[l], \quad e[l] \sim \mathcal{N}(0, \sigma_m^2)$$

$$P_m = \underbrace{\frac{1}{L}\sum_{l=1}^{L} s_m^2[l]}_{P_m^s} + \underbrace{\frac{1}{L}\sum_{l=1}^{L} 2 s_m[l] e_m[l]}_{P_m^{sw} \to 0} + \underbrace{\frac{1}{L}\sum_{l=1}^{L} e_m^2[l]}_{v_m \sim \mathcal{N}\left(\sigma_m^2, \frac{2\sigma_m^4}{L}\right)},$$

Signal model

$$P_m = P_m^s + v_m$$

Signal model

$$P_m = P_m^s + v_m$$

$$P_m(\psi) = \alpha g_m h(\psi, \boldsymbol{\theta}_m) + v_m$$

Signal model

$$P_m = P_m^s + v_m$$

$$P_m(\psi) = \alpha g_m h(\psi, \boldsymbol{\theta}_m) + v_m$$

α: absolute power level

g_m: gain of microphone m

$h(\psi, \boldsymbol{\theta}_m)$: directivity pattern

Fourier series Model

$$P_m(\psi) = \alpha g_m h(\psi, \boldsymbol{\theta}_m) + v_m$$

Fourier series Model

$$P_m(\psi) = \alpha g_m h(\psi, \boldsymbol{\theta}_m) + v_m$$

$$\begin{aligned} h\left(\psi, \boldsymbol{\theta}_m\right) &= \theta_0^m + \sum_{d=1}^{D} \theta_{d,c}^m \cos(d\psi) + \theta_{d,s}^m \sin(d\psi) \\ &= \boldsymbol{\Phi}(\psi)\boldsymbol{\theta}_m \end{aligned}$$

Training

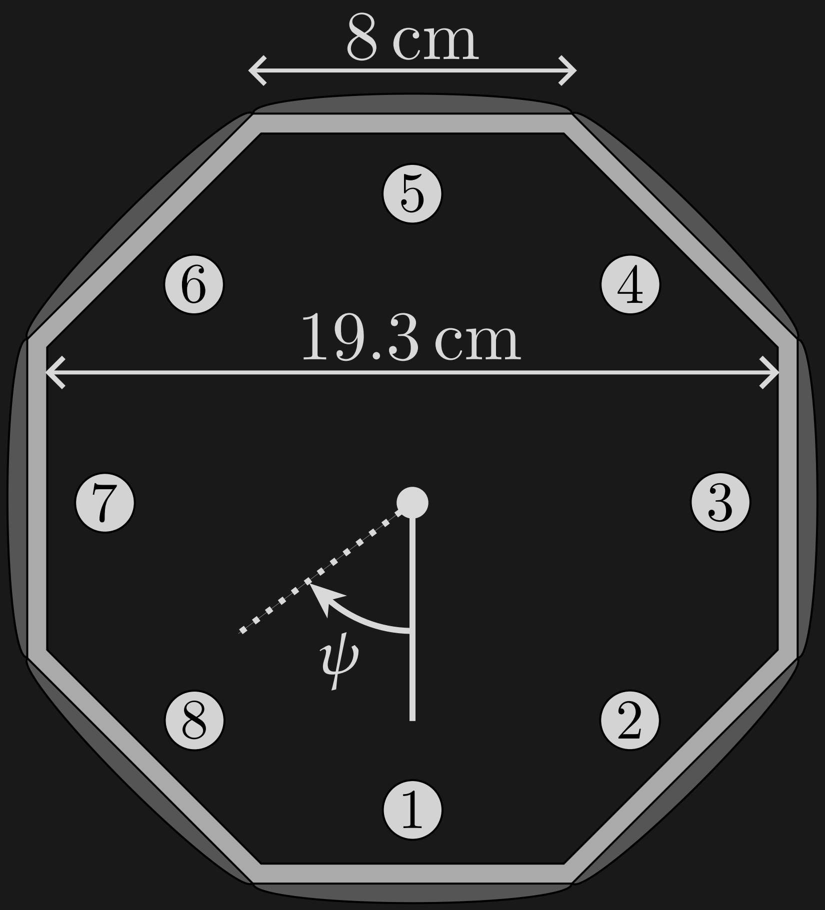
The parameters $\boldsymbol{\theta}_m, g_m$ are estimated by minimizing the error between the predicted and the measured power.

Training

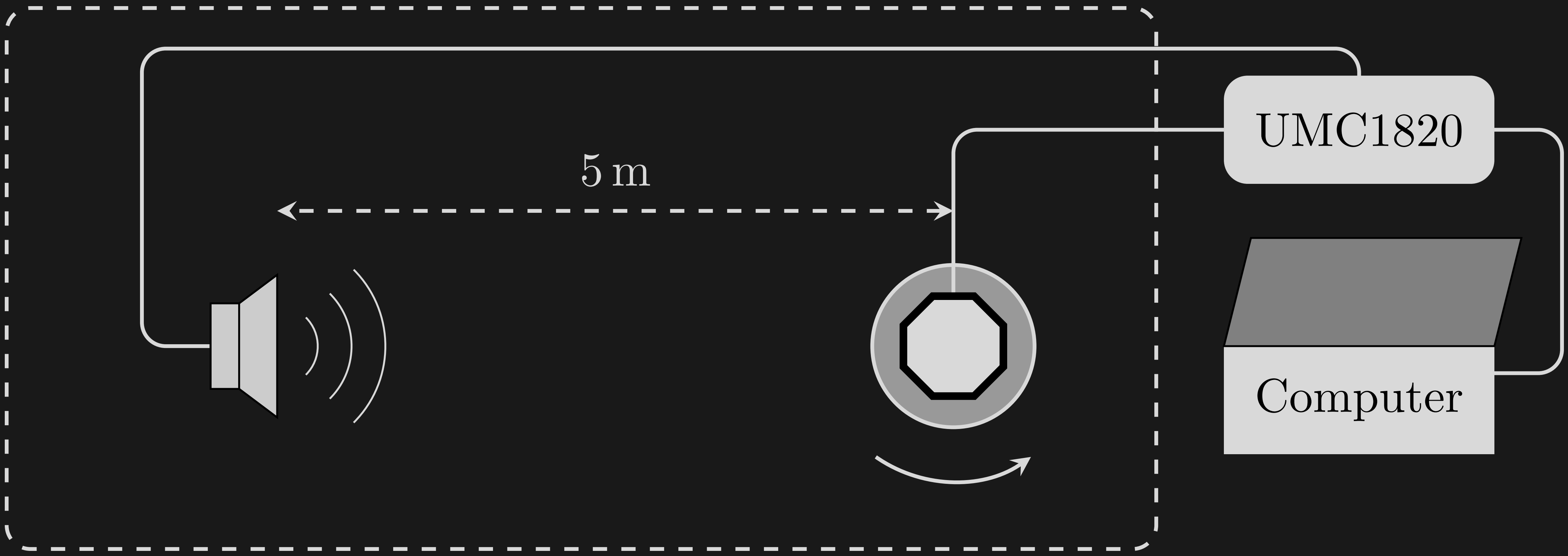
The parameters θ_m, g_m are estimated by minimizing the error between the predicted and the measured power.

$$
\begin{aligned}
\underset{\boldsymbol{\alpha},\, \boldsymbol{g}_{1:M},\, \boldsymbol{\theta}_{1:M}}{\text{minimize}} \quad & \sum_{m=1}^{M} \frac{1}{\eta_m} \sum_{k=1}^{K} \left(P_m(\psi_k) - \left(\alpha g_m h\left(\psi_k, \boldsymbol{\theta}_m\right) + \sigma_m^2\right)\right)^2 \\
\text{subject to} \quad & \alpha > 0 \\
& g_1 = 1 \\
& h\left(\psi_m, \boldsymbol{\theta}_m\right) = 1, \qquad \forall\, m = 1, 2, \ldots, M,
\end{aligned}
$$

Experimental Setup



Experimental Setup



19 different sounds

No.	Sound	Signal
1	Wideband noise	0:00 / 0:10
2-11	Sinusoids	0:00 / 0:10
12	Hovering drone	0:00 / 0:05
13	Elephant trumpet	0:00 / 0:02
14	Police Siren	0:00 / 0:10

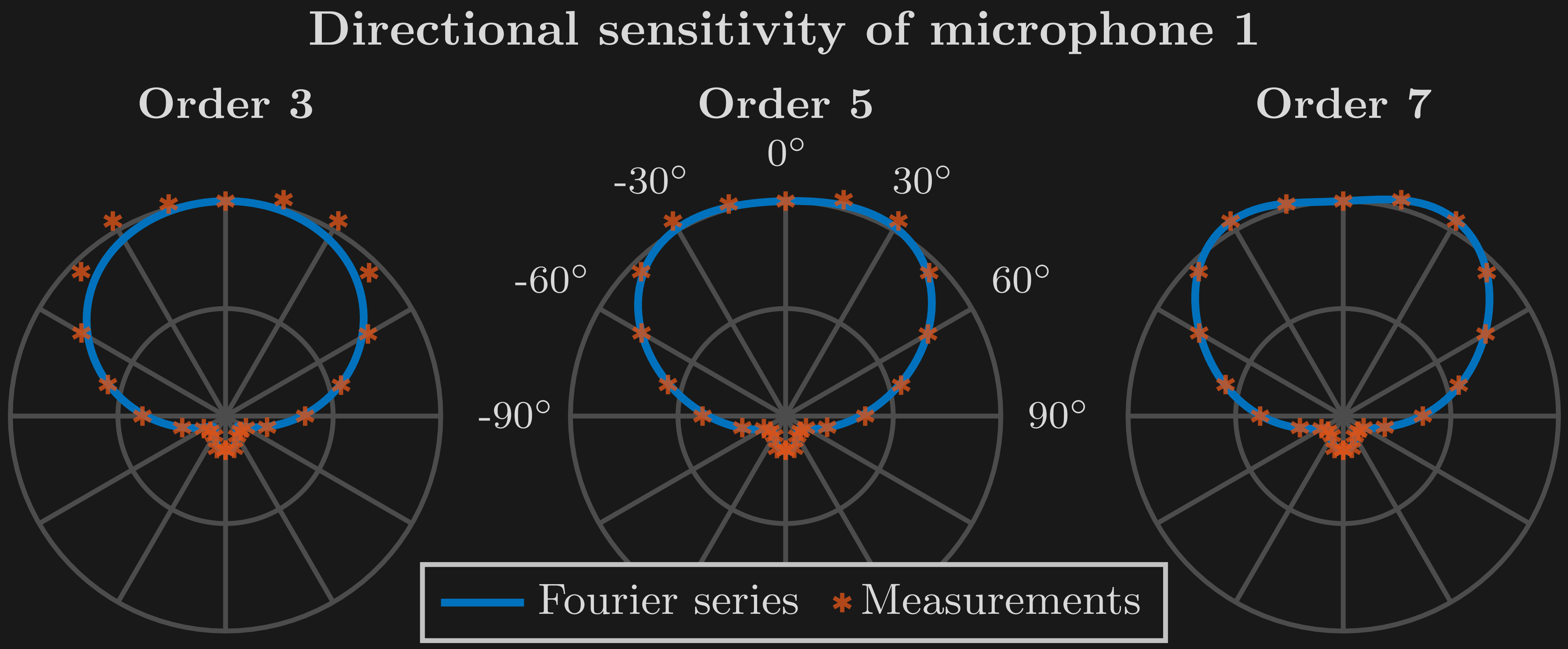
19 different sounds

No.	Sound	Signal
15	Woman Scream	0:00 / 0:01
16	Wideband noise low	0:00 / 0:10
17	Wideband noise high	0:00 / 0:10
18	Gunshot	0:00 / 0:00
19	Background noise	0:00 / 0:26

Data summary

	Training		Estimation	
	Signals	**# of angles**	**Signals**	**# of angles**
Anechoic	Signal 1	24	Signal 1-18	24
Outdoor	-	-	Signal 1-17	10

Fourier Series Order



DOA Estimation

$$\hat{P}_m(\psi) = \alpha \underbrace{\hat{g}_m \mathbf{\Phi}(\psi) \hat{\boldsymbol{\theta}}_m}_{\hat{h}_m(\psi)} + \sigma_m^2$$

DOA Estimation

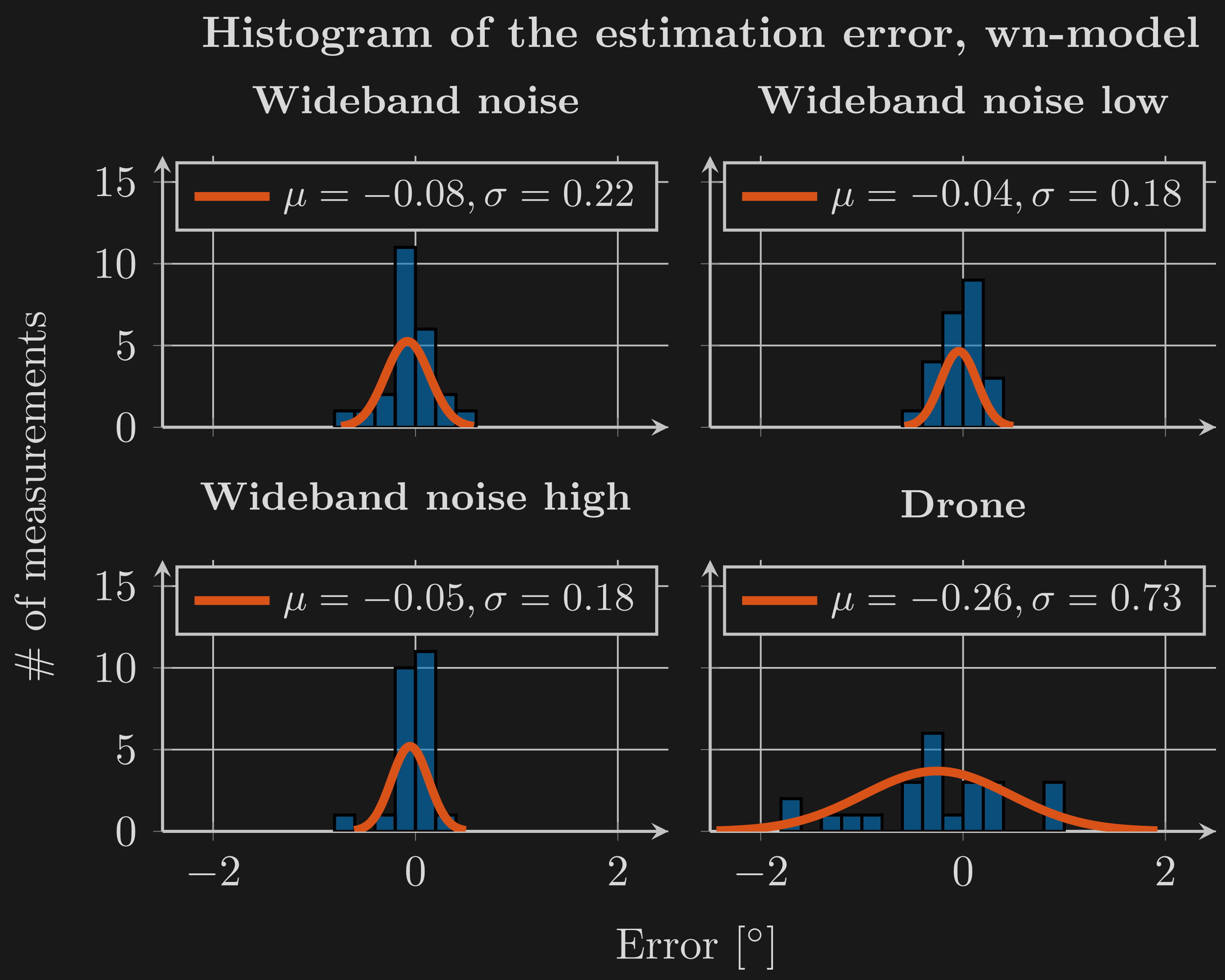
$$\hat{P}_m(\psi) = \alpha \hat{g}_m \boldsymbol{\Phi}(\psi) \hat{\boldsymbol{\theta}}_m + \sigma_m^2$$

$$\hat{h}_m(\psi)$$

$$\hat{\psi} = \arg\max_{\psi} \frac{\left(\sum_{m=1}^{M} \hat{h}_m(\psi)(P_m - \sigma_m^2)\right)^2}{\sum_{m=1}^{M} \hat{h}_m^2(\psi)}$$

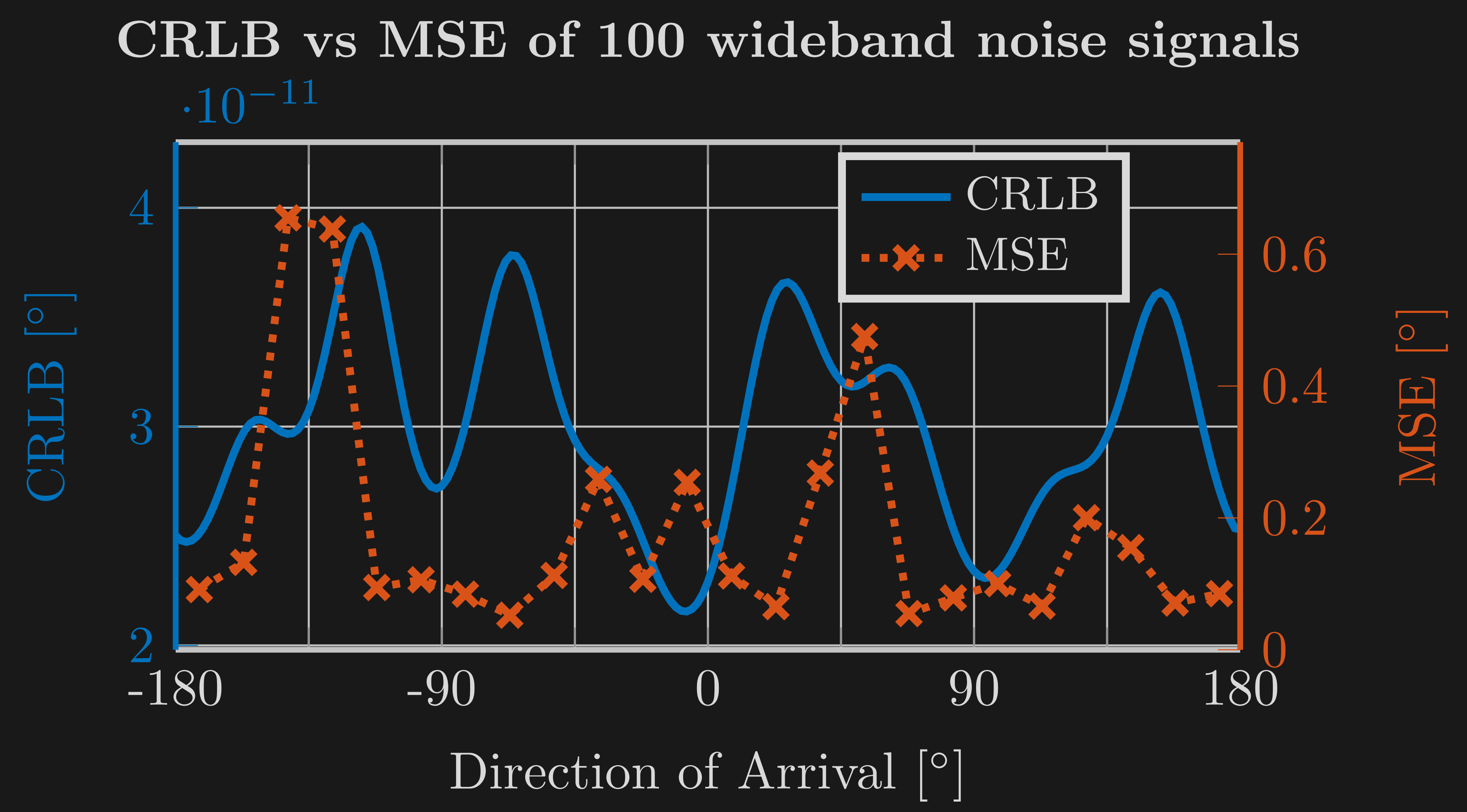
$$\hat{\alpha} = \frac{\sum_{m=1}^{M} \hat{h}_m(\hat{\psi})(P_m - \sigma_m^2)}{\sum_{m=1}^{M} \hat{h}_m^2(\hat{\psi})}$$

Result



Histogram of the estimation error, wn-model
Wideband noise
Wideband noise low
Wideband noise high
Drone
$\mu = -0.08, \sigma = 0.22$
$\mu = -0.04, \sigma = 0.18$
$\mu = -0.05, \sigma = 0.18$
$\mu = -0.26, \sigma = 0.73$
# of measurements
Error [°]
0
5
10
15
−2
0
2

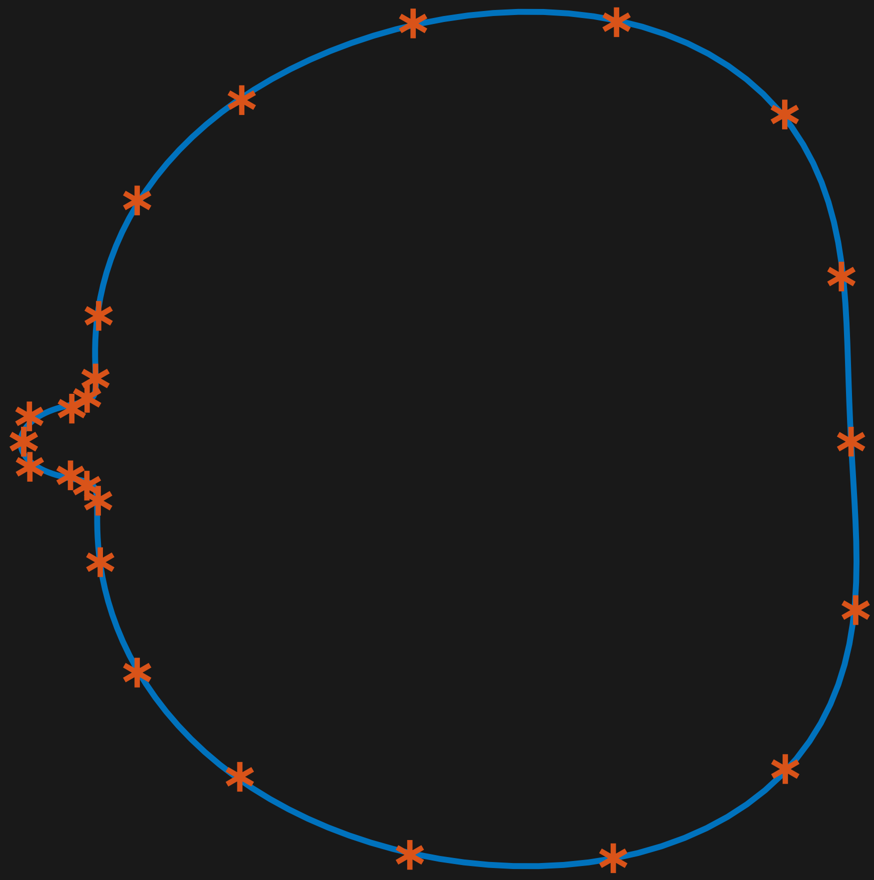
Cramér-Rao Lower Bound



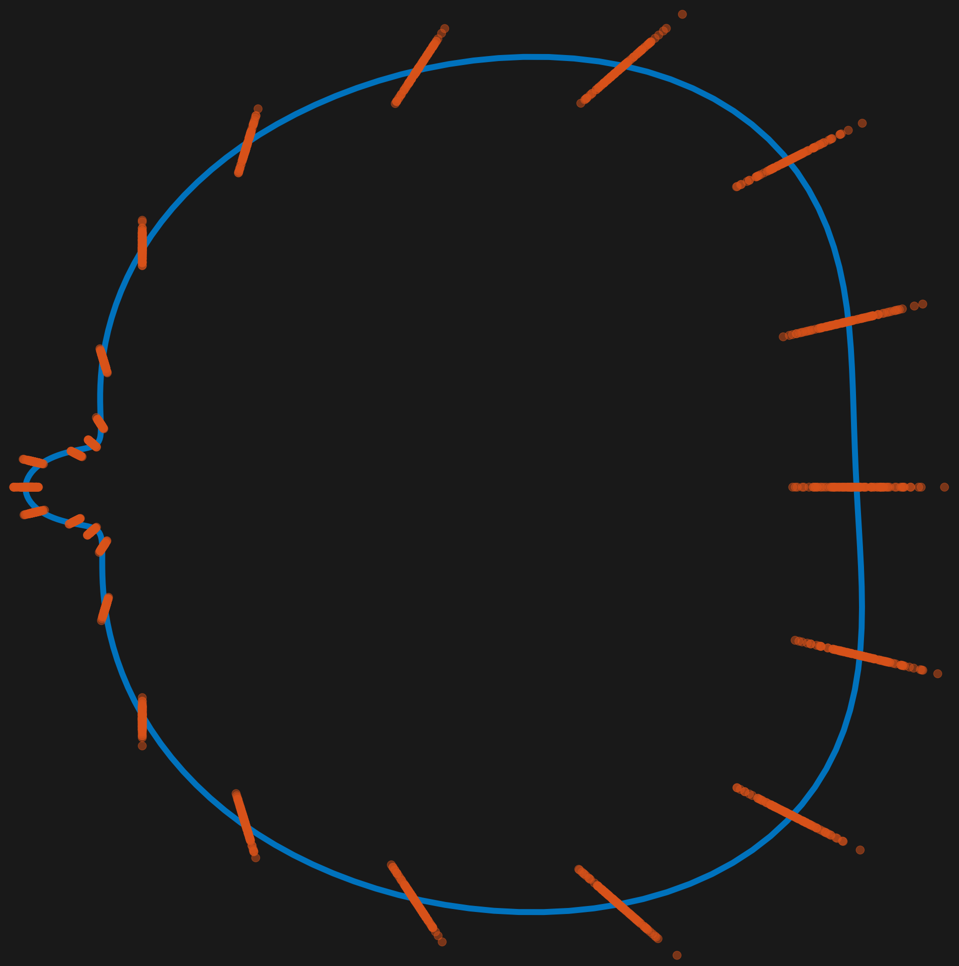
Noise distribution

$$P_m = \alpha g_m h(\psi, \boldsymbol{\theta}_m) + v_m, \quad v_m \sim$$

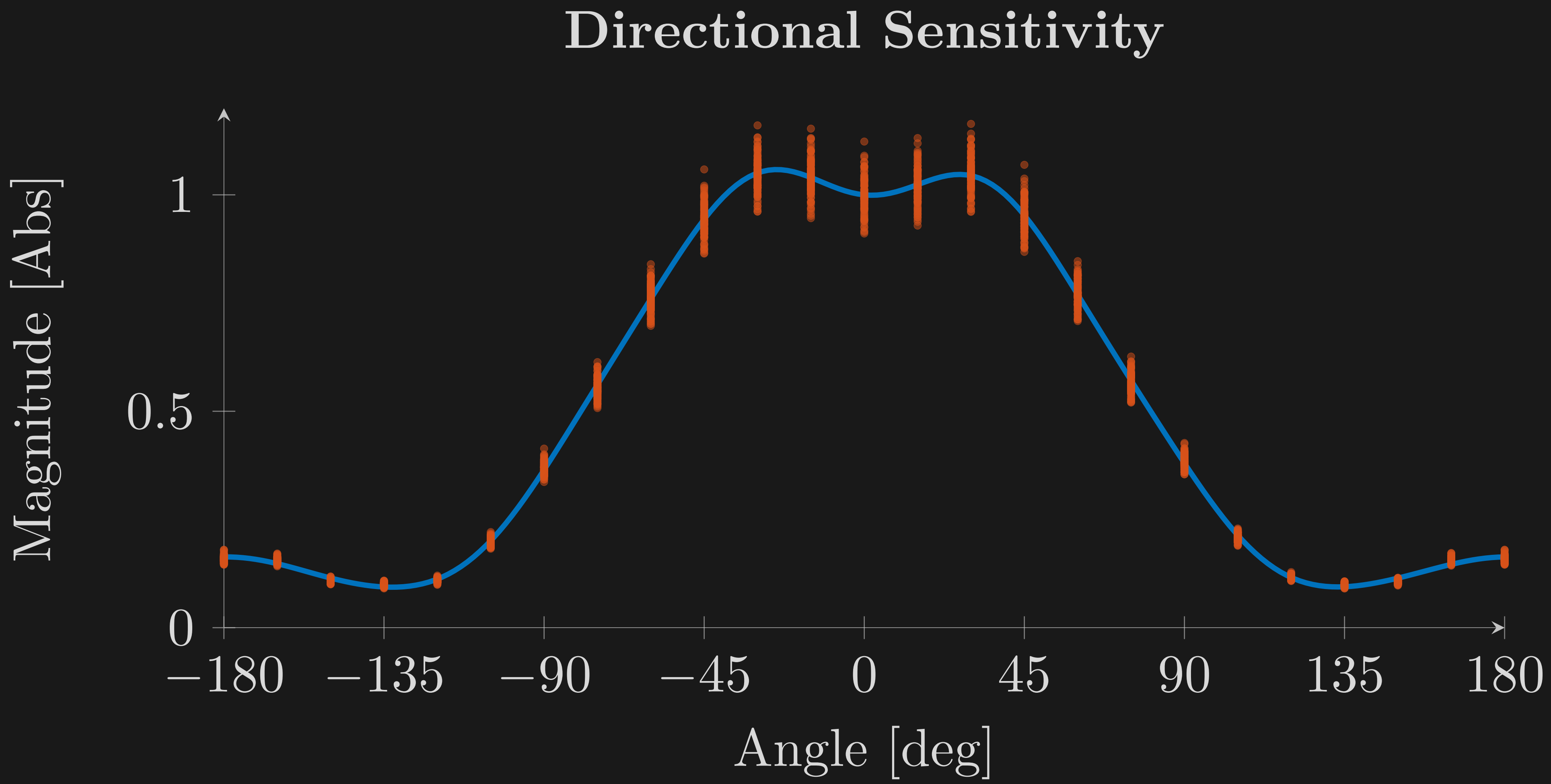
Noise distribution



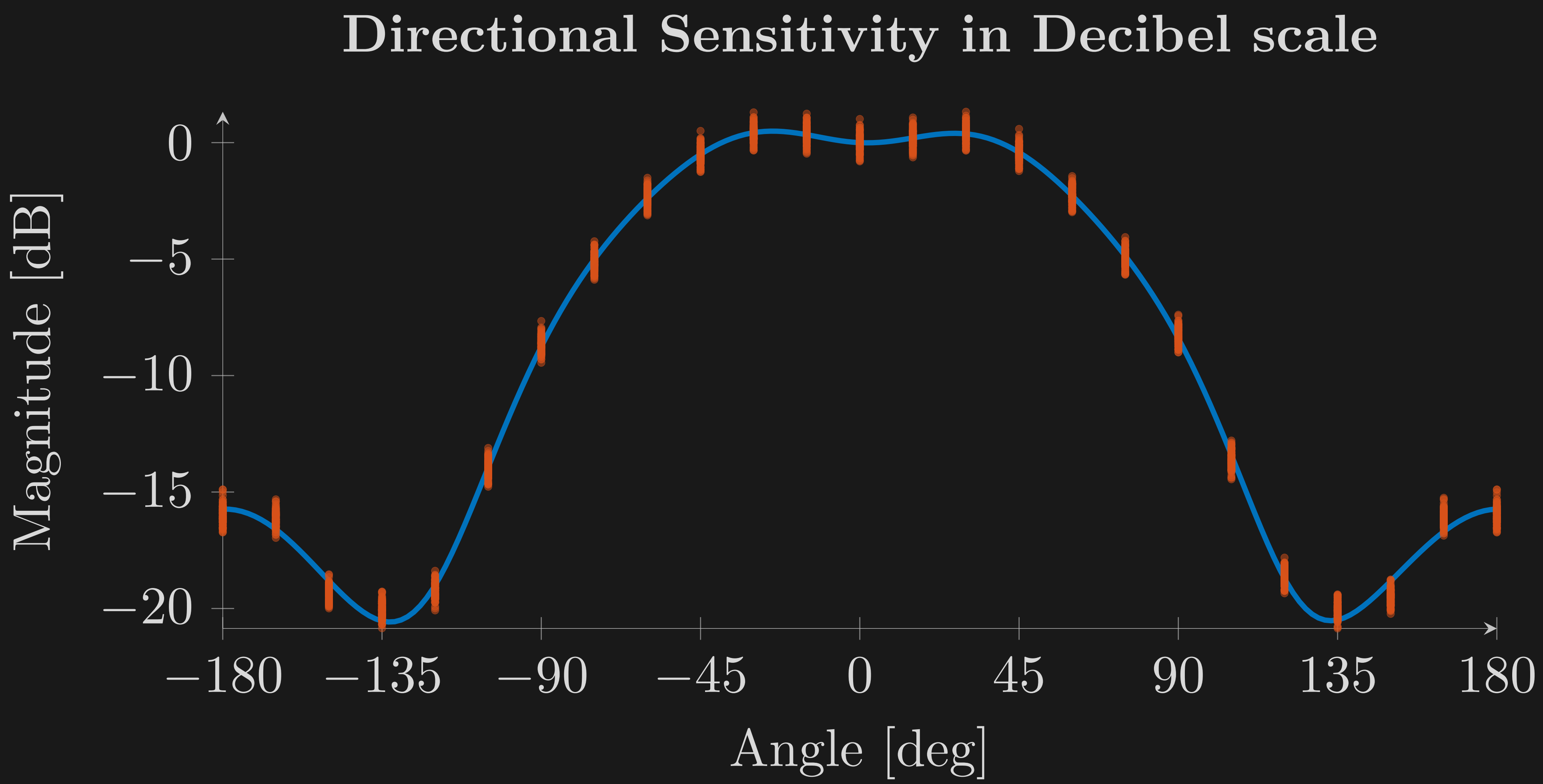
Noise distribution



Noise distribution



Noise distribution

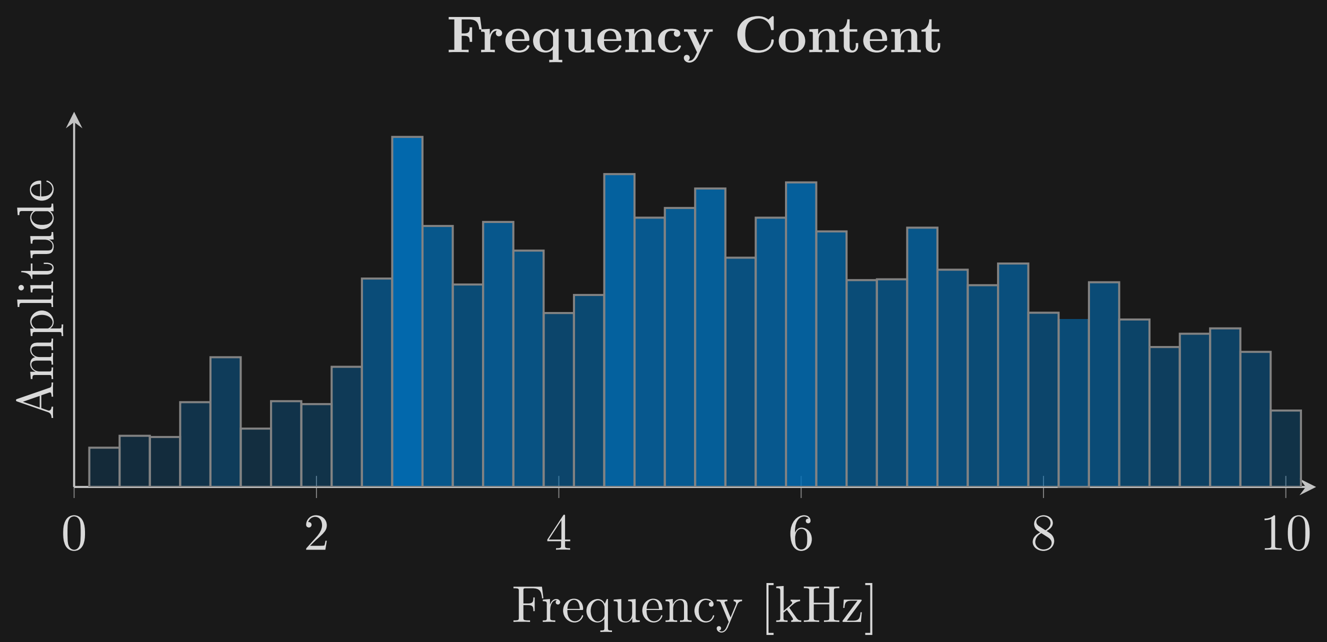


Noise distribution

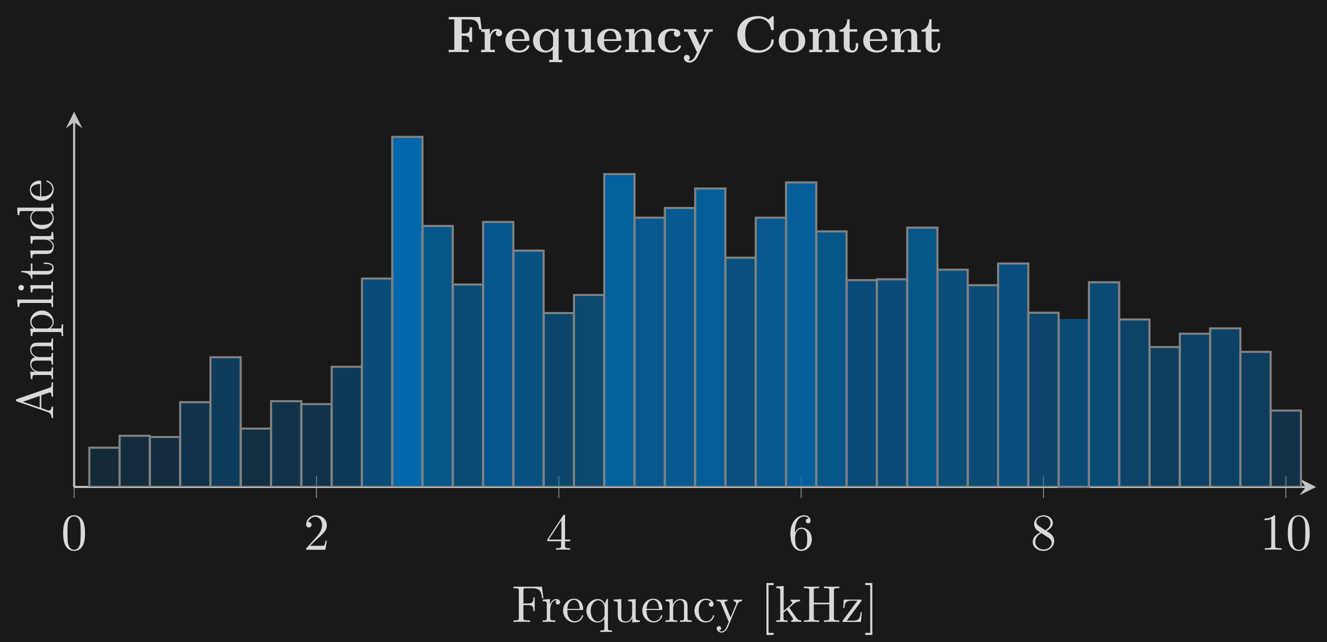
Introduce model in dB scale as

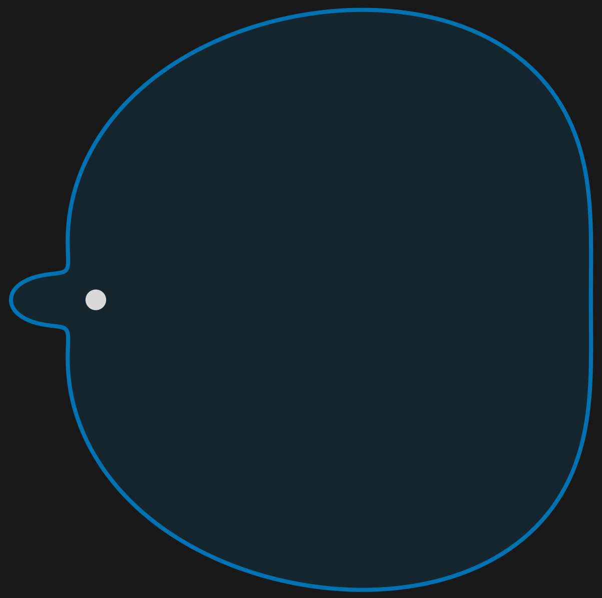
$$\bar{P}_m = \bar{\alpha} + \bar{g}_m + h(\psi, \bar{\boldsymbol{\theta}}_m) + \bar{v}_m,$$

Frequency Dependency



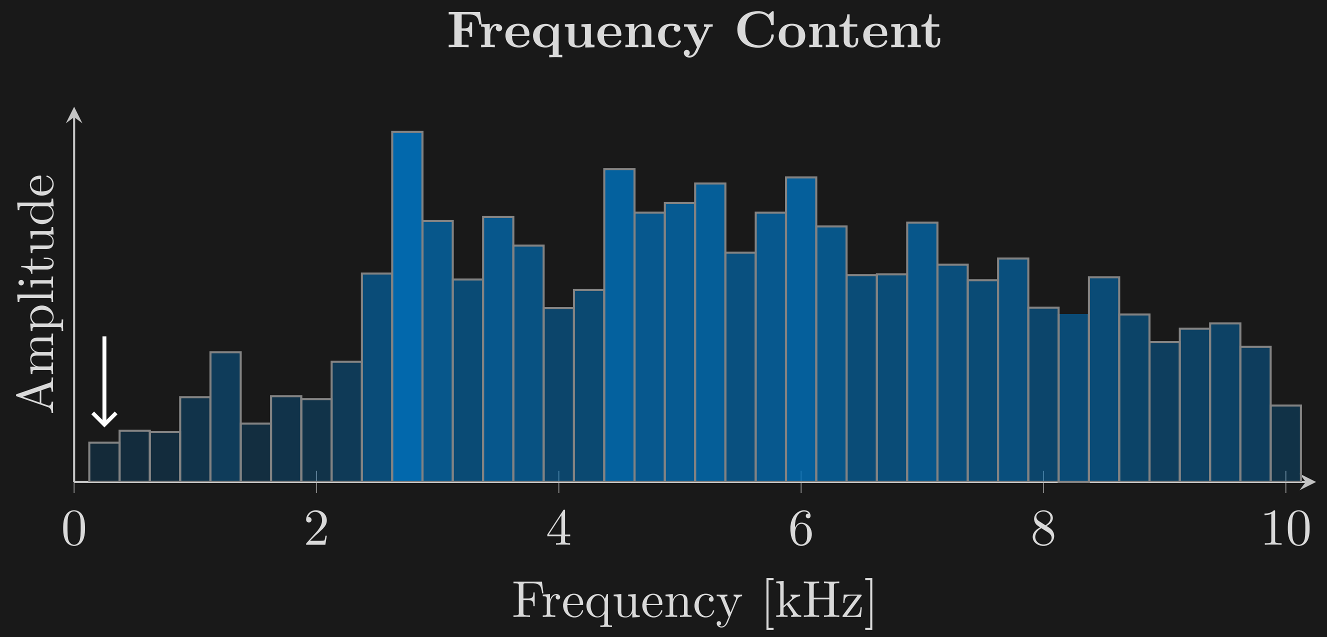
Frequency Dependency





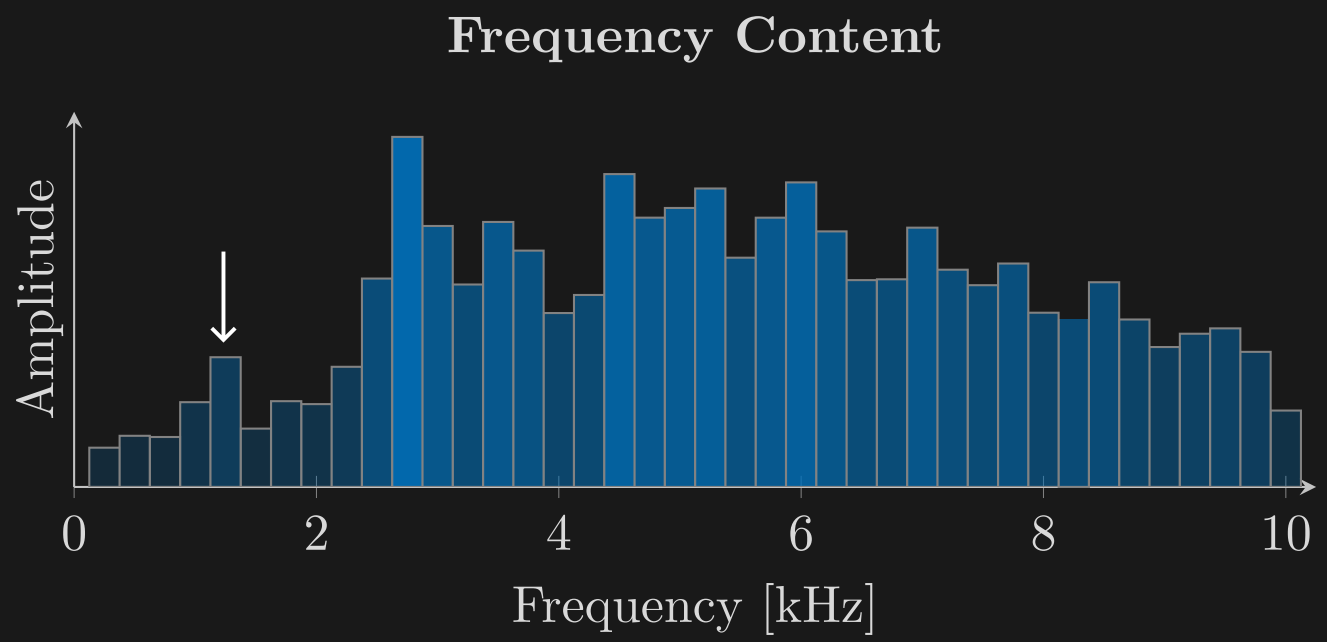
Wideband noise

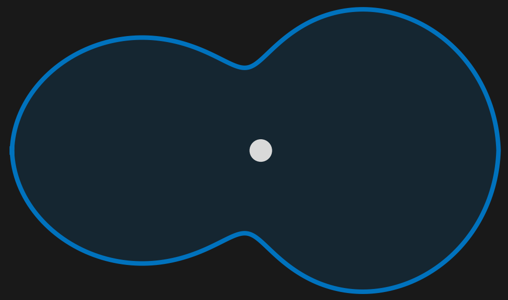
Frequency Dependency



250 Hz

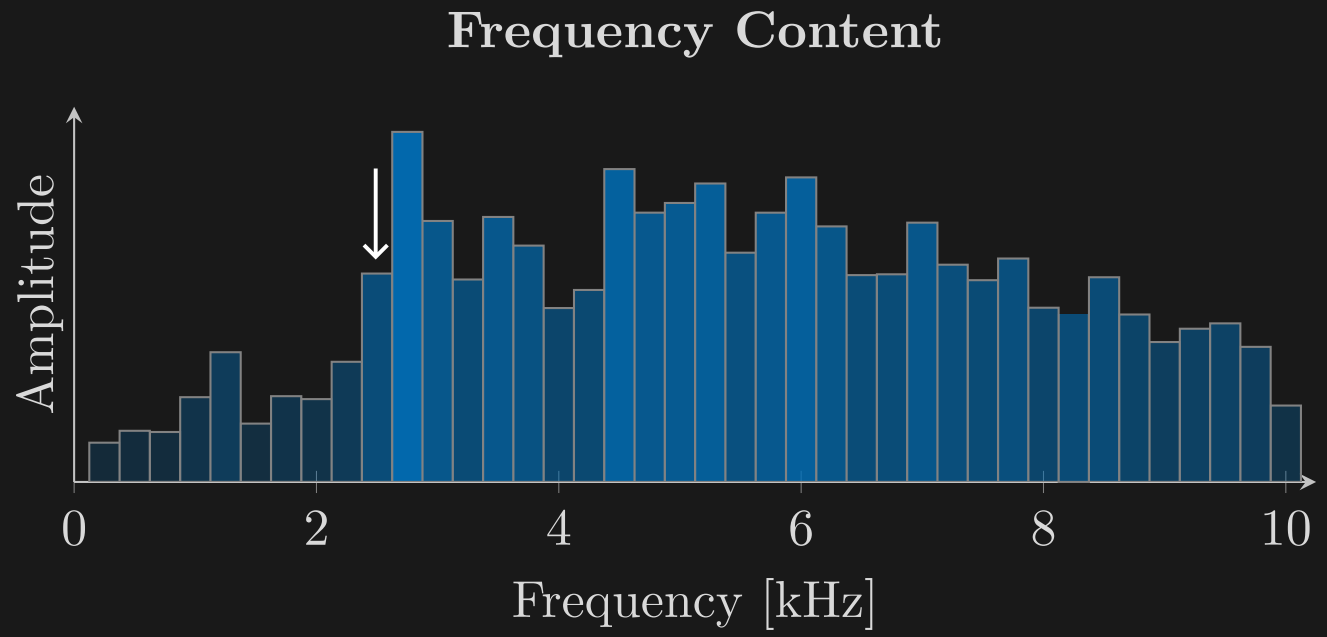
Frequency Dependency

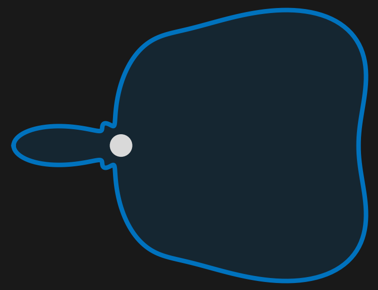




1250 Hz

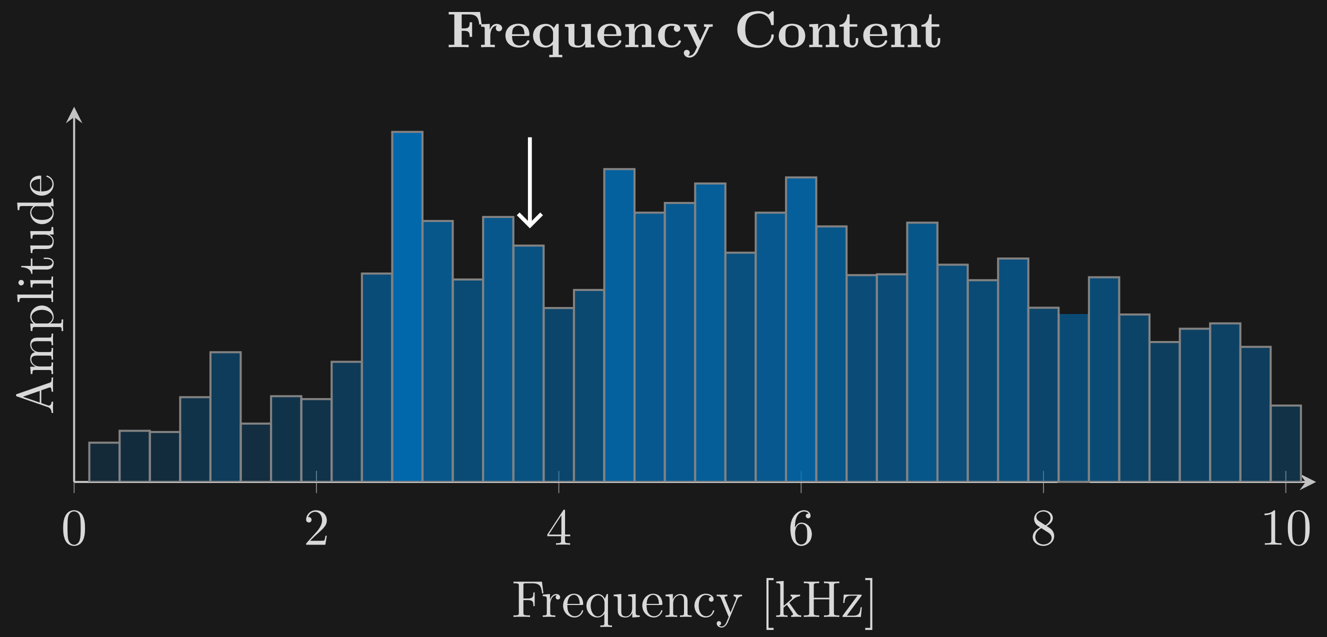
Frequency Dependency

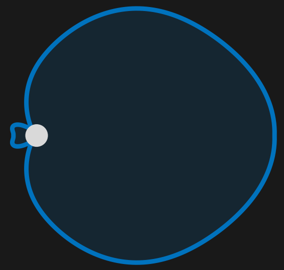




2500 Hz

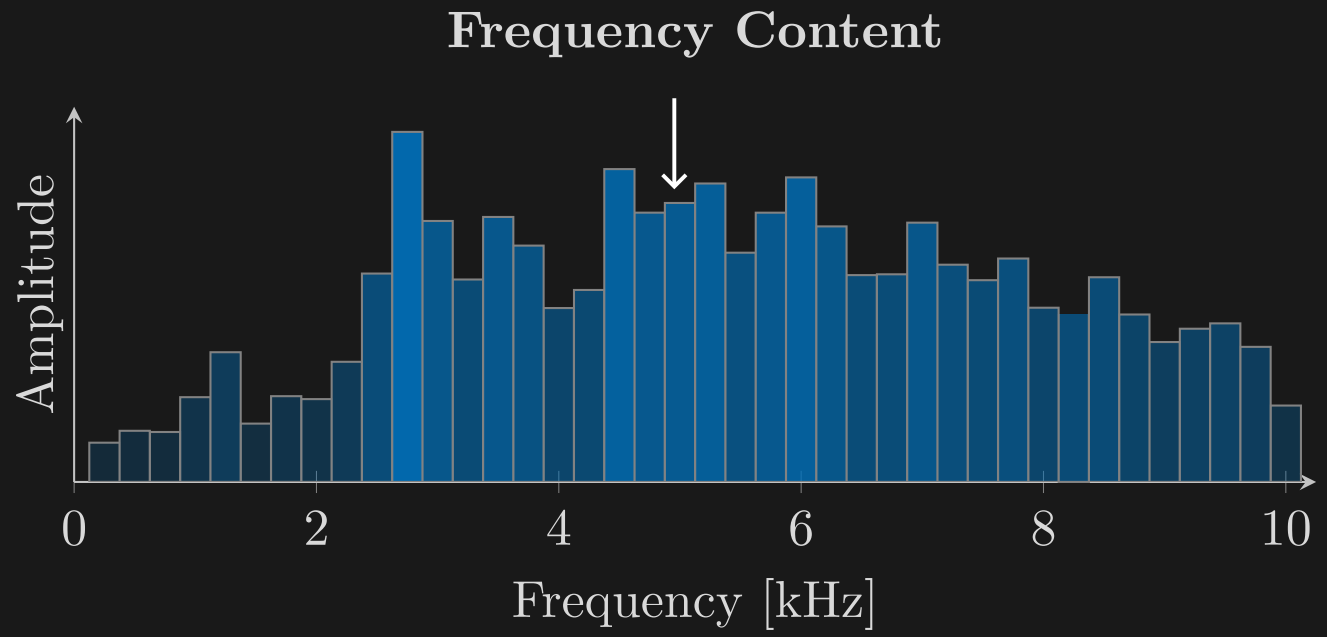
Frequency Dependency

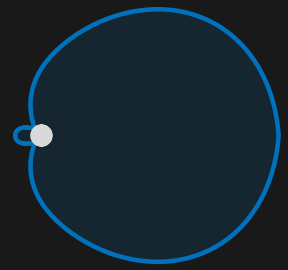




3750 Hz

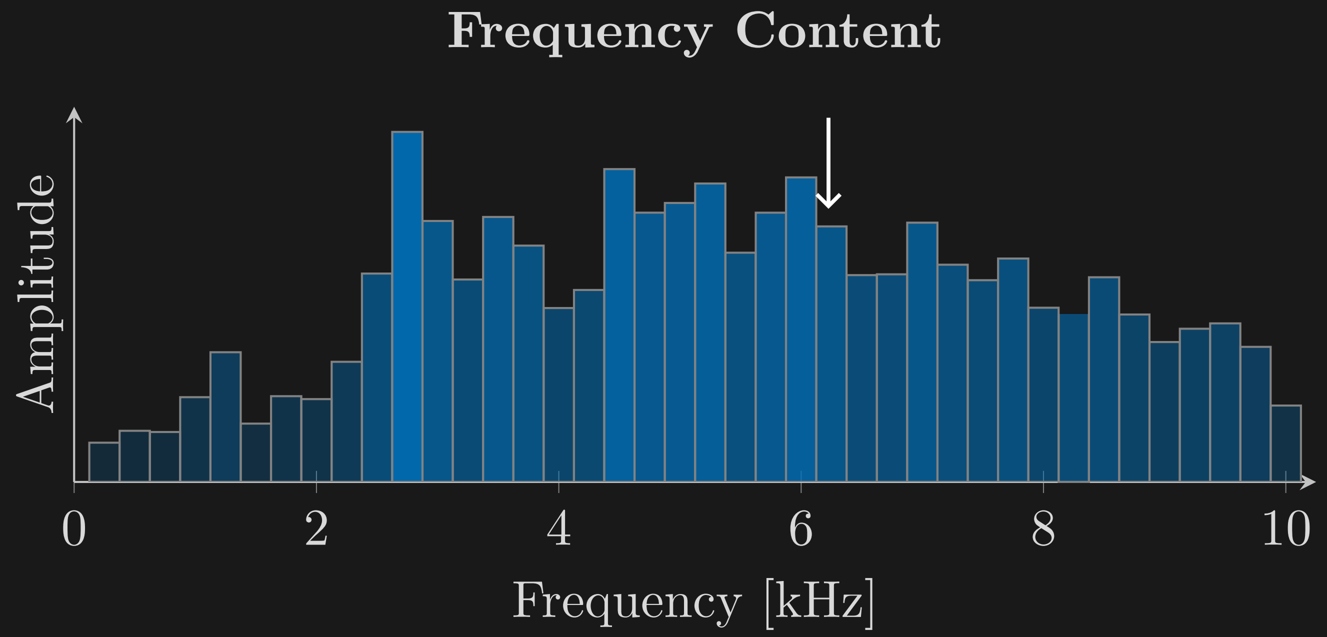
Frequency Dependency

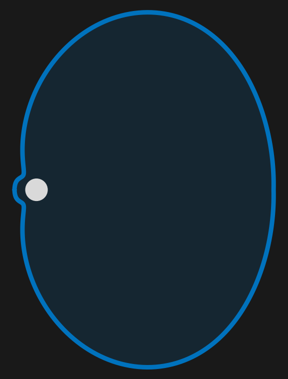




5000 Hz

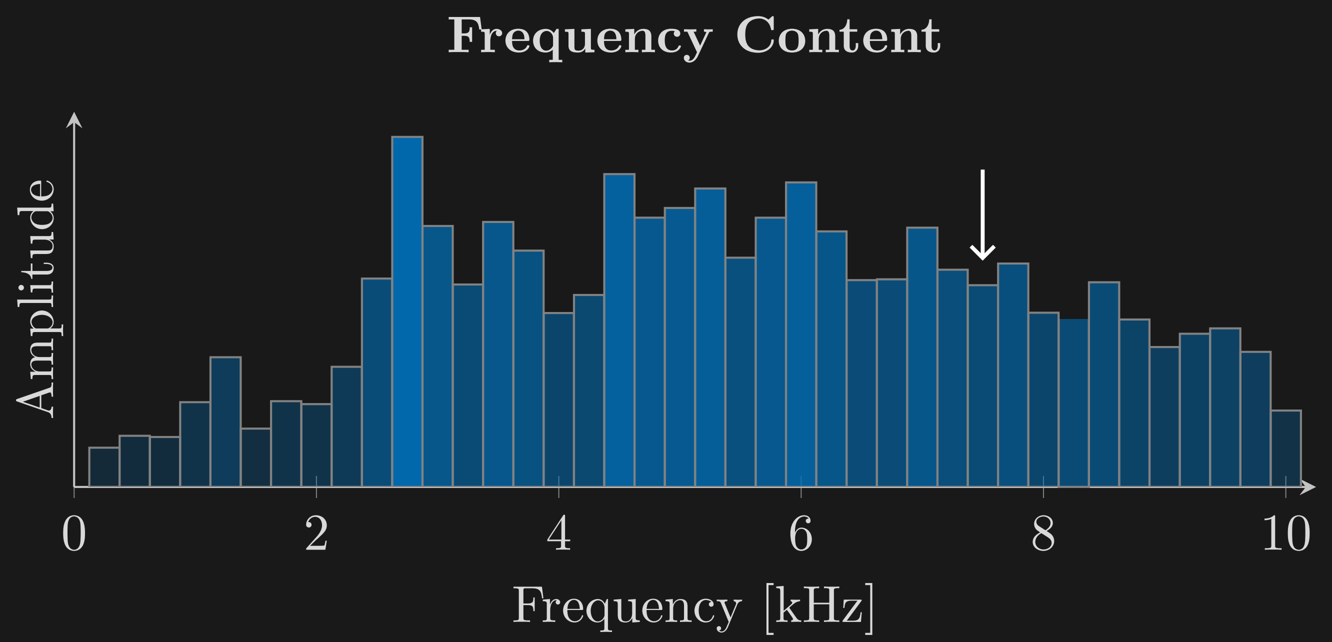
Frequency Dependency

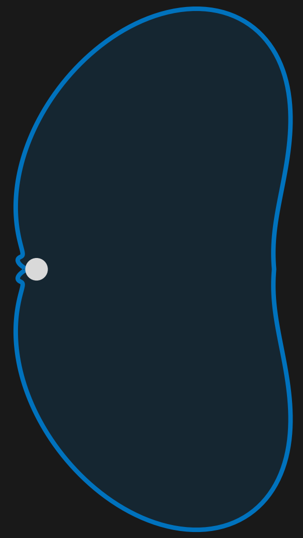




6250 Hz

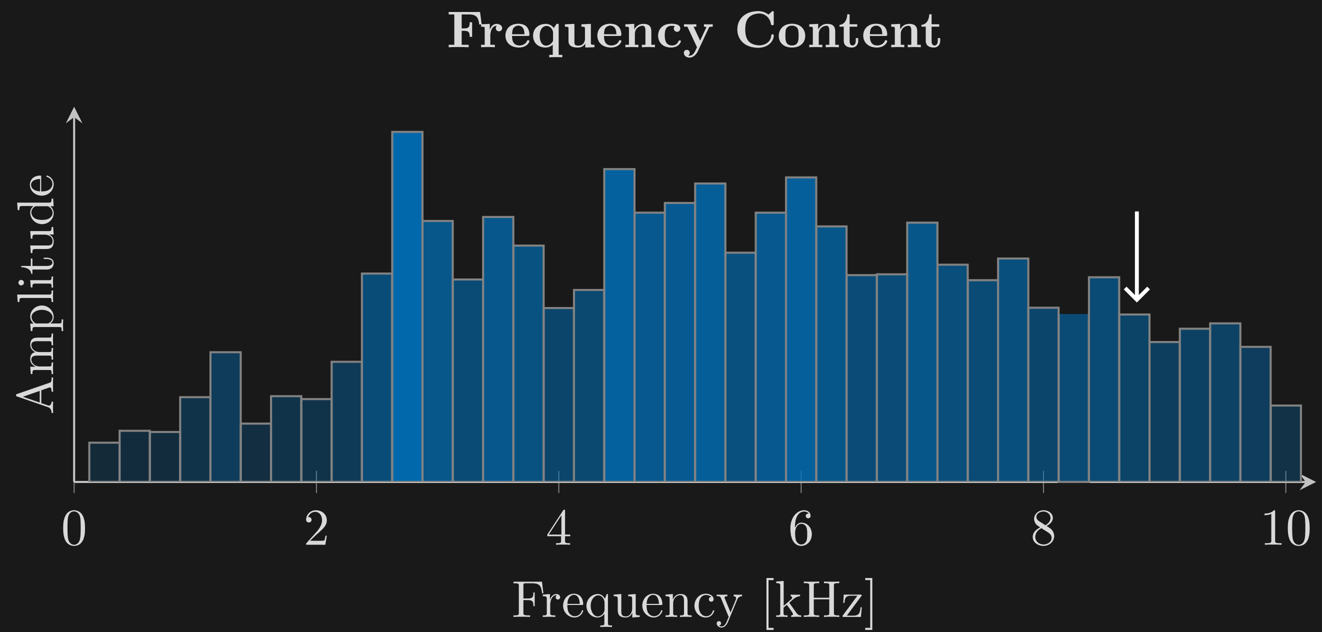
Frequency Dependency

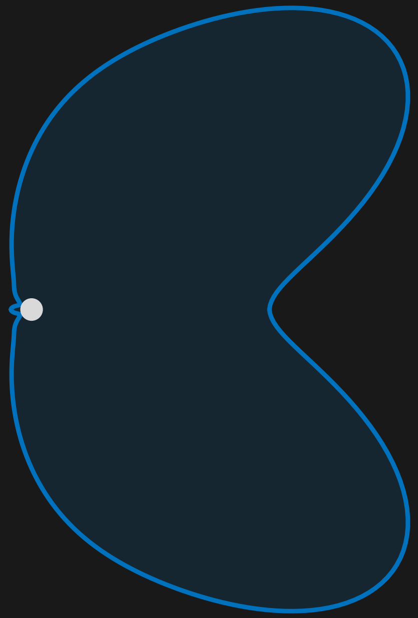




7500 Hz

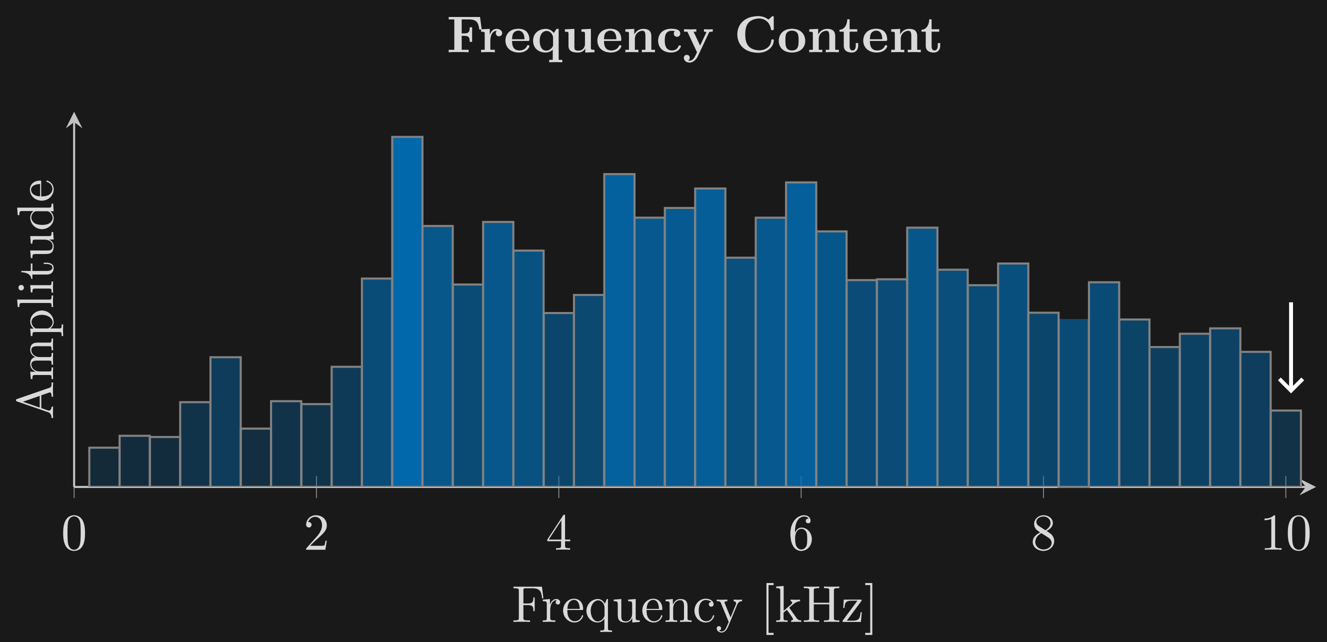
Frequency Dependency

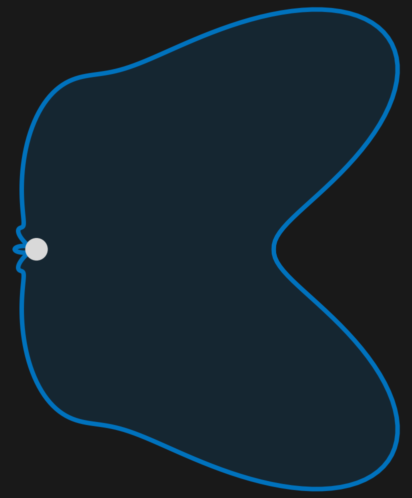




8750 Hz

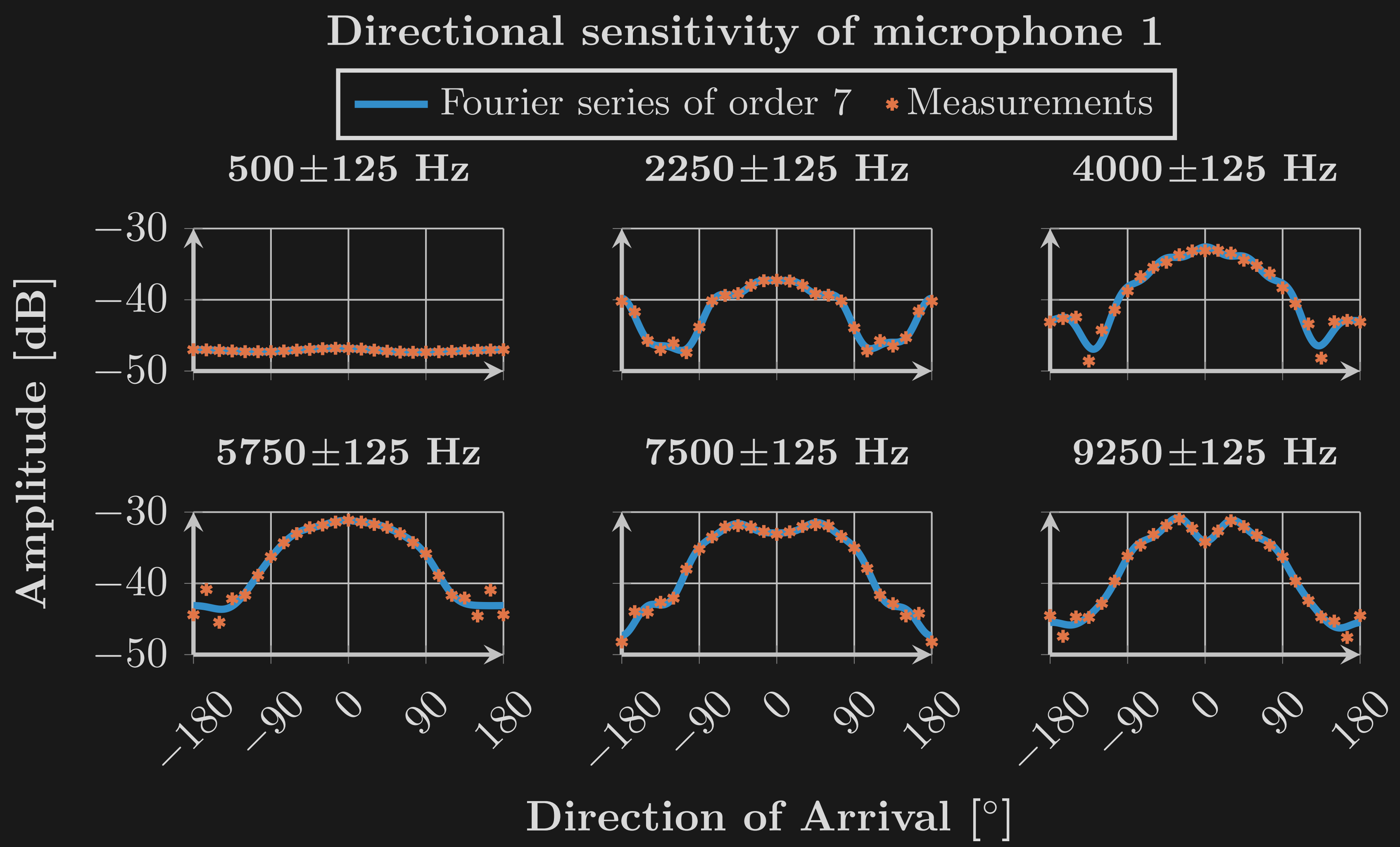
Frequency Dependency



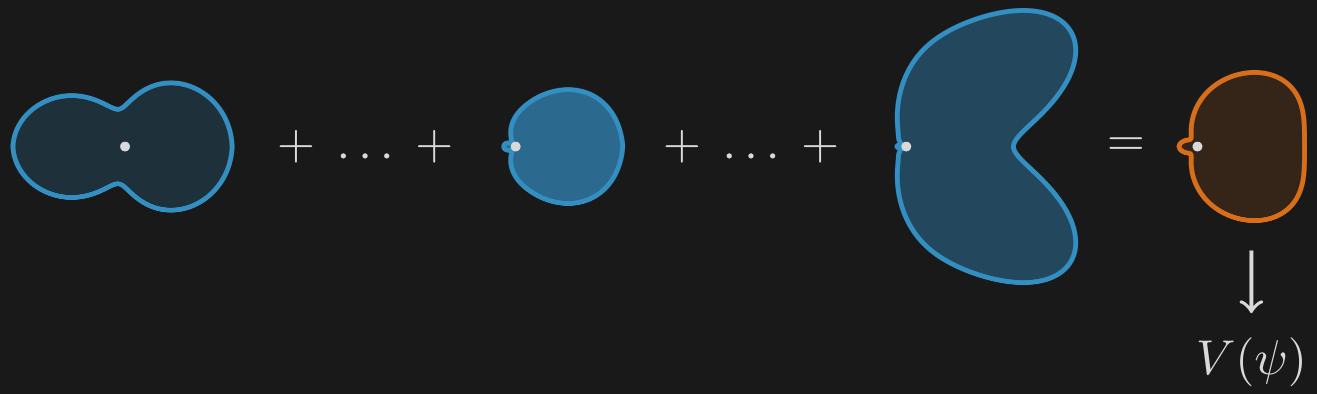


10 000 Hz

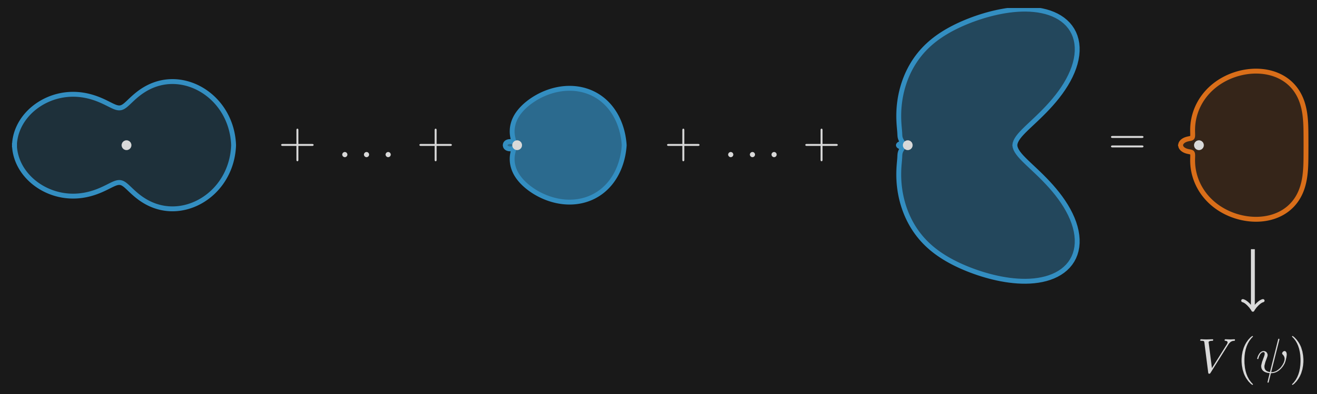
Frequency Dependency

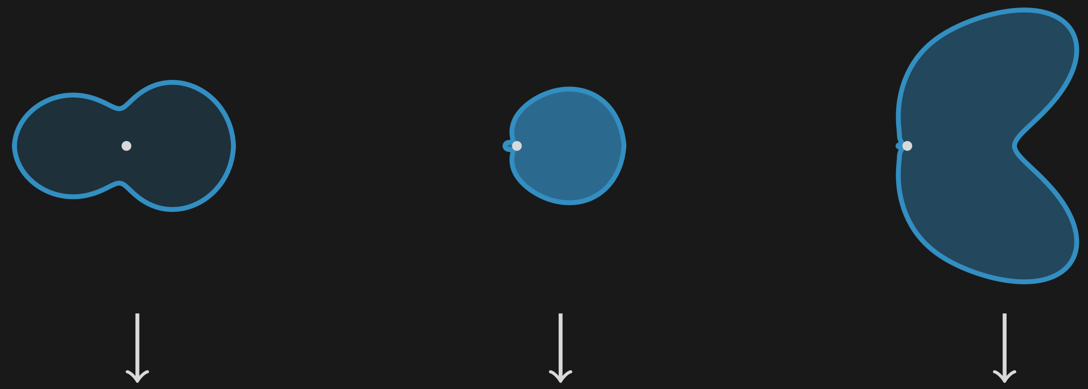


Frequency Dependency



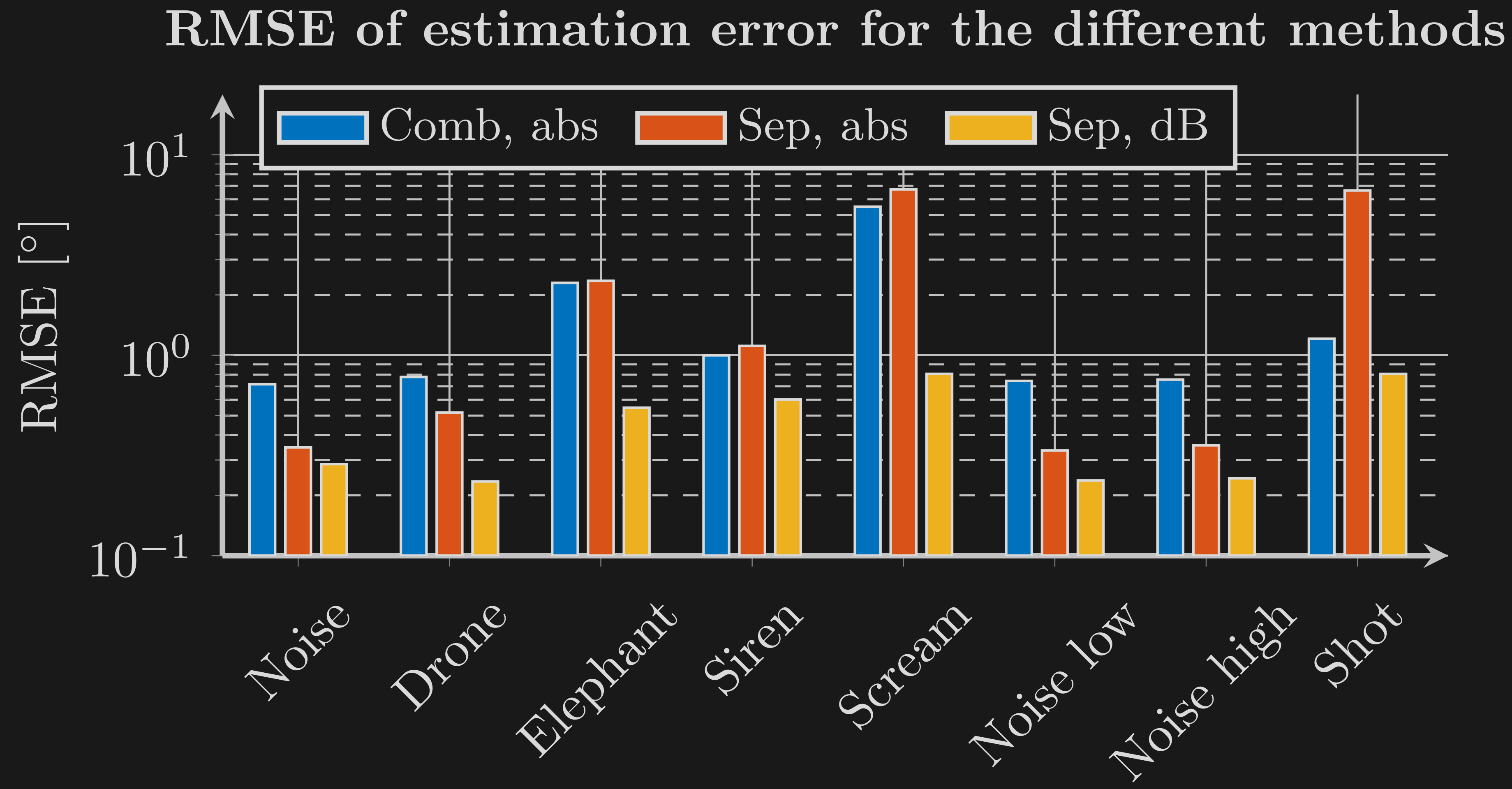
Frequency Dependency





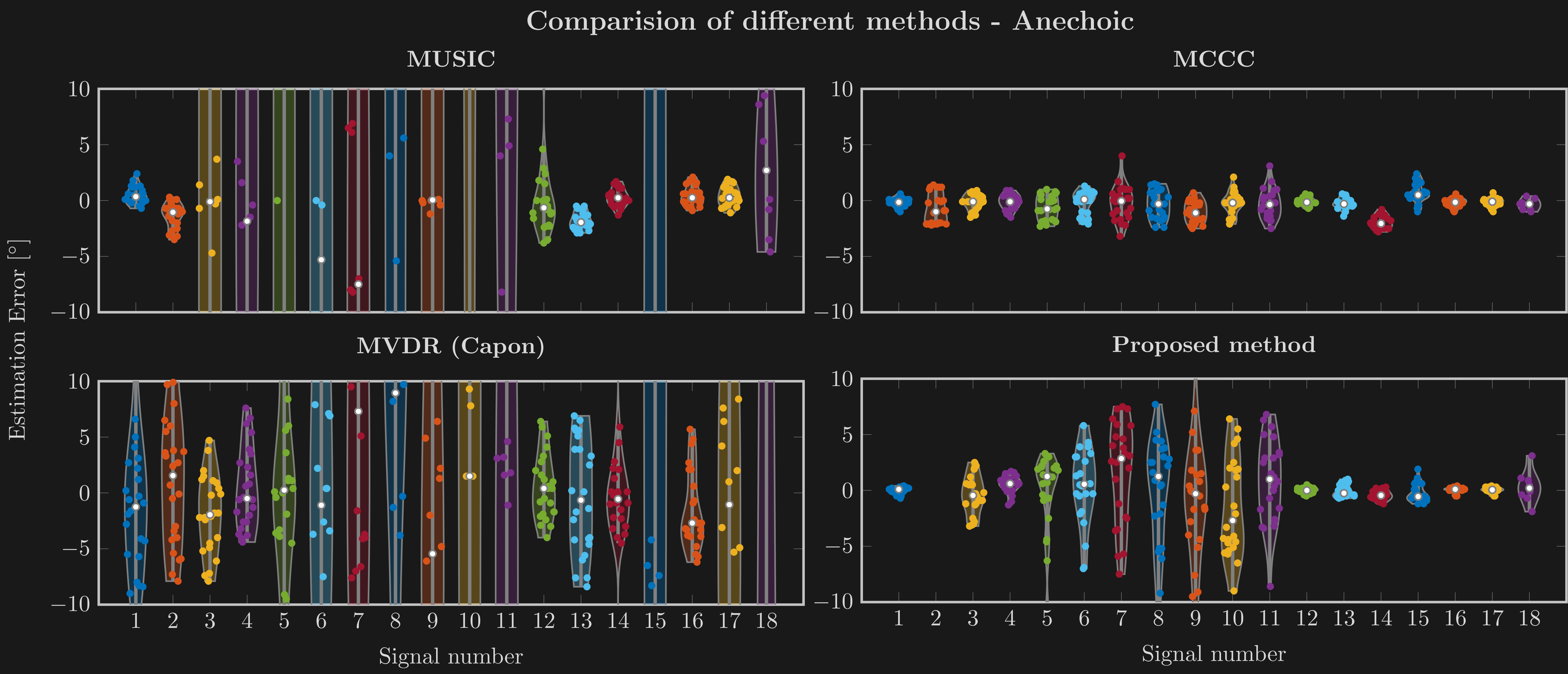
$$w_1 V_1(\psi) + \ldots + w_n V_n(\psi) + \ldots + w_F V_F(\psi) = V(\psi)$$

Results



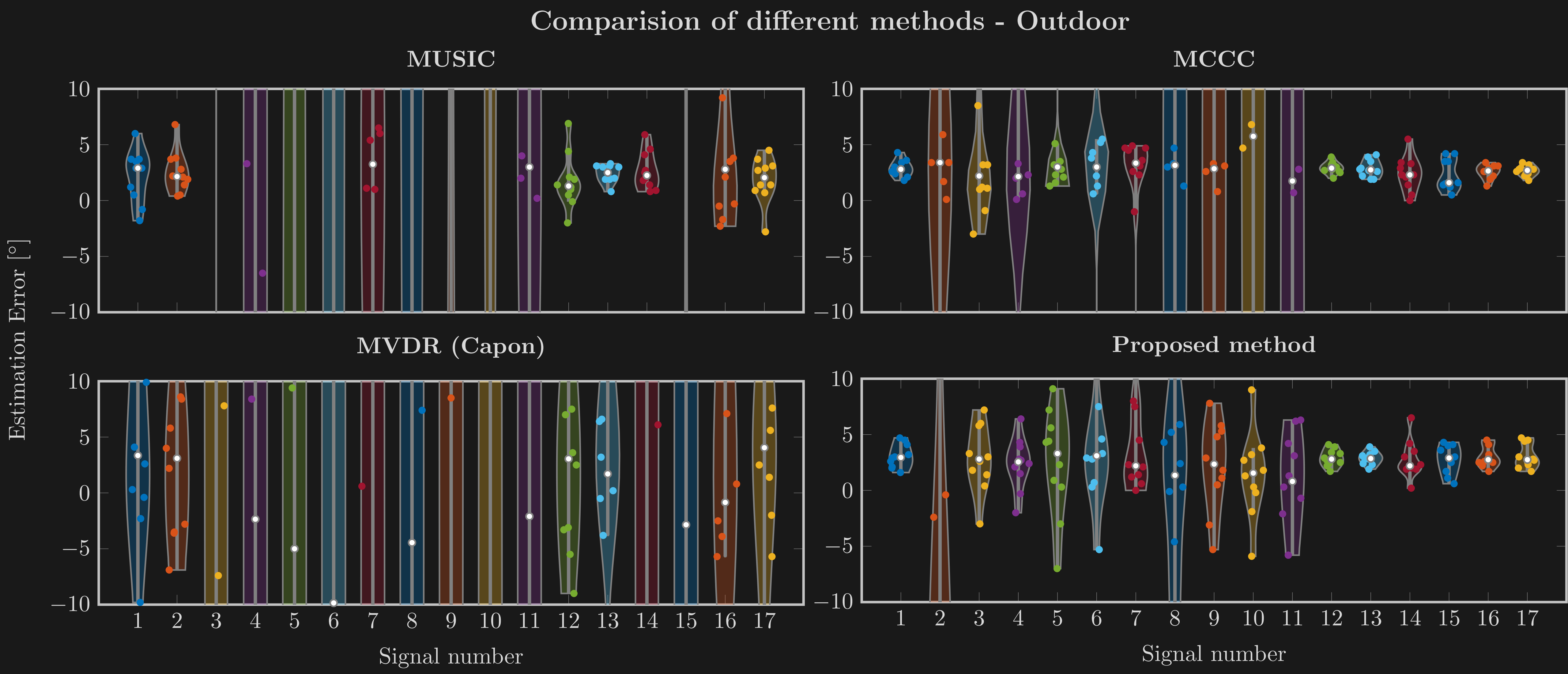
RMSE of estimation error for the different methods
Comb, abs
Sep, abs
Sep, dB
RMSE [°]
10^1
10^0
10^{-1}
Noise
Drone
Elephant
Siren
Scream
Noise low
Noise high
Shot

Results



Comparision of different methods - Anechoic
MUSIC
MCCC
MVDR (Capon)
Proposed method
Estimation Error [°]
10
5
0
−5
−10
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Signal number

Results

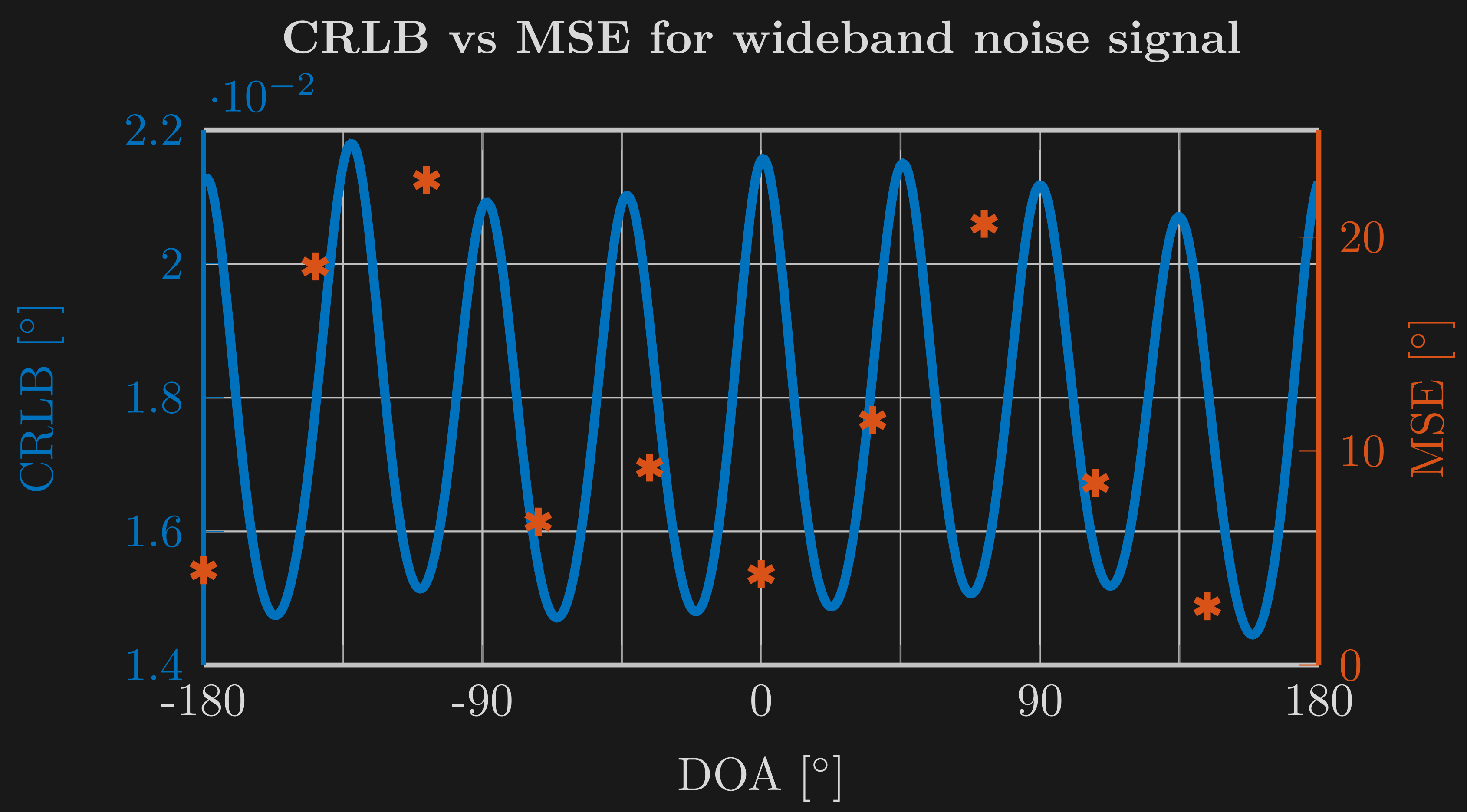


Comparision of different methods - Outdoor
MUSIC
MCCC
MVDR (Capon)
Proposed method
Estimation Error [°]
Signal number
Signal number
10
5
0
−5
−10
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

Results

	Signal No.	Proposed	MUSIC	MVDR	MCCC
Anechoic	1-17	11.62	60.04	27.07	**0.99**
	1, 12-17	**0.45**	15.33	22.51	0.79
	18	1.42	9.68	92.43	**0.58**
Outdoor	1-17	**5.14**	74.25	76.08	42.03
	1, 12-17	3.10	29.15	59.23	**2.78**

Cramér-Rao Lower Bound



Conclusions

- **DOA Estimation Method**: Utilizes directional sensitivity; requires calibration.

Conclusions

- **DOA Estimation Method**: Utilizes directional sensitivity; requires calibration.
- **Successful Experiments**: Eight-microphone array estimated DOA for 18 signal types, effective for frequencies above 1000 Hz.

Conclusions

- **DOA Estimation Method**: Utilizes directional sensitivity; requires calibration.
- **Successful Experiments**: Eight-microphone array estimated DOA for 18 signal types, effective for frequencies above 1000 Hz.
- **Robust Performance**: Maintains accuracy over time and temperature; enables high-resolution DOA estimation with smaller arrays.

Future Work

- Investigate the effect of **reverberations** to potentially improve robustness and accuracy in complex acoustic environments.

Future Work

- Investigate the effect of **reverberations** to potentially improve robustness and accuracy in complex acoustic environments.
- Address **model errors** in microphone directional sensitivity to refine performance in diverse conditions.

Future Work

- Investigate the effect of **reverberations** to potentially improve robustness and accuracy in complex acoustic environments.
- Address **model errors** in microphone directional sensitivity to refine performance in diverse conditions.
- Test with more **complex signals** (e.g., speech) to evaluate real-world effectiveness.

Future Work

- Explore the method's applicability to detecting elephant footsteps, considering **variations in individual elephants** and multi-target tracking of herds.

Future Work

- Explore the method's applicability to detecting elephant footsteps, considering **variations in individual elephants** and multi-target tracking of herds.
- Study the influence of ground vehicles and **other large mammals** (e.g., humans, giraffes, rhinos) on detection accuracy.

Future Work

- Explore the method's applicability to detecting elephant footsteps, considering **variations in individual elephants** and multi-target tracking of herds.
- Study the influence of ground vehicles and **other large mammals** (e.g., humans, giraffes, rhinos) on detection accuracy.
- Assess the **method's range limitations** to improve its design for larger or open environments.

Thank you!